Università degli Studi di Padova

Dipartimento di Biologia

SCUOLA DI DOTTORATO DI RICERCA IN BIOSCIENZE E BIOTECNOLOGIE

INDIRIZZO: BIOTECNOLOGIE

CICLO: XXVI

INTEGRATED ANALYSIS OF mRNA AND miRNA IN HUMAN DIFFERENTIATING MUSCLE CELLS

Direttore della Scuola : Ch.mo Prof. Giuseppe Zanotti

Coordinatore d'indirizzo: Ch.mo Prof. Fiorella Lo Schiavo

Supervisore :Ch.mo Prof. Giorgio Valle

Dottorando : Rusha Guha

Table of Contents

Acknowledgements

With sincerest of gratitude I would like to thank my supervisor Prof Giorgio Valle for his valuable guidance through my PhD. I also extend my heartfelt thanks to him for being such an enormous support both at professional and personal level. I feel truly grateful and very lucky for having the opportunity to work with him.

I whole heartedly thank Dr. Georgine Faulkner who is a visiting scientist of our lab and who has provided precious inputs to our project with her great experience of working in the field of muscle biology. I have had a chance to refine my practical skills as researcher with her help. I feel so blessed for having worked with Prof Giorgio Valle and Dr. Georgine Faulkner for their endless help and support through the entire thick and thins I have experienced during my stay in Padova and also for extending their care to make my life comfortable while being away from my country.

I would like to thank my colleague and a fellow co PhD student Lisa Marchioretto who has also worked on the muscle project but towards a different goal. It would have been very difficult to settle in a new country which I did not know the language of without her help. I thank her for her kindness and would like to wish her the best for her PhD work.

This PhD project would not have been complete without the excellent technical support, so I do thank the entire technical team of our lab. I also thank Nicola Vitulo a bioinformatician of our lab for his contribution towards analyzing the RNA seq data my research project. Working with the team of Prof Giorgio Valle was a great and enriching experience.

Finally I thank my family, my parents my sister for being there like a backbone support, for all their blessings and good wishes which has given me the courage to sail through the journey of PhD.

Last but not the least I would like to thank the program of Erasmus Mundus because of which I got the opportunity to work in the University of Padova.

Abstract

Muscles are responsible for the movement of body and take up roughly half of the person's body weight. Skeletal muscles are the only voluntary muscle tissue in human body, being controlled consciously.

Skeletal muscle cells form when many smaller progenitor cells lump themselves together to constitute long, straight, multinucleated fibers. The proliferating muscle cells are called myoblasts. Myoblasts fuse together to form multinucleated non-proliferating cells called myotubes. The transition from proliferative to differentiated state involves a complete shift of the cell's transcriptome based on a network of regulation at the molecular level. To uncover the intricacies of molecular activities involved in the process of muscle differentiation it is essential to have deeper and through understanding of the transcriptome in its entirety. RNA seq, which is a high through put sequencing technology, gives us the opportunity to reach into the deepest level of the transcriptome.

By means of the RNA seq technology I obtained the in-depth view of the transcriptome of human muscle cells from the proliferative to the differentiated stage. The analysis was extended also to small RNAs, to have a full picture of the transcriptome. I performed the analysis with the specific objectives of identifying the expressed genes and finding out differential gene expression, of both mRNA and miRNA. I also investigated the crosstalk between mRNA and miRNA, employing two separate methods : miRNA over expression and Ago2 immunoprecipitation followed by RNA seq.

The over expression experiments were carried out with 5 different miRNAs and in all cases I found more genes turned up than turned down. These results suggest that miRNA might either have a role in mRNA stabilization or it could play a part in a double negative mechanism by inhibiting some negative factor. By comparing the genes down regulated after miRNA over expression with Ago2-enriched genes we have found several candidate genes which are most likely under the down regulatory control of miRNA.

Skeletal muscle has enormous plasticity and can endure a lot of stress. To study this aspect of muscle we performed mechanical stretch of differentiating muscle cells. With RNA seq we got the in-depth view of the transcriptome as a response to stretch. We performed the analysis at two time points after stretch and found that stretch triggered

immune response genes soon after, but enhanced muscle structural-protein genes expression over a prolonged course of time when the immune response takes a back seat.

I believe that our thorough transcriptome analysis, including miRNA and mRNA interaction studies during myogenesis, contributes towards the better understanding of the process regulating muscle development.

Riassunto

I muscoli sono responsabili dei movimenti del corpo e costituiscono circa metà del peso di una persona. I muscoli scheletrici sono i soli tessuti muscolari volontari, essendo controllati coscientemente.

Le cellule del muscolo scheletrico si formano quando diverse piccole cellule progenitrici si conglobano tra loro per formare lunghe affusolate fibre multinucleate. Le cellule muscolari proliferanti sono chiamate mioblasti. I mioblasti si fondono tra loro per formare cellule multinucleate e non proliferanti, chiamate miotubi. La transizione dallo stato proliferante a quello differenziato implica un completo cambiamento del trascrittoma cellulare, basato su una rete di regolazione a livello molecolare. Per scoprire il groviglio di attività molecolari implicate nel processo di differenziamento muscolare è essenziale avere una più profonda ed estesa comprensione del transcrittoma nella sua interezza. L'RNA seq è una tecnologia di sequenziamento massivo che ci offre l'opportunità di accedere ai livelli più approfonditi del trascrittoma.

Con la tecnologia dell'RNA seq ho ottenuto una precisa visione del trascrittoma delle cellule muscolari umane, sia allo stadio proliferativo che a quello differenziato. Per avere una visione completa del trascrittoma, l'analisi è stata estesa anche agli *small* RNA. Ho svolto queste analisi con l'obiettivo specifico di identificare i geni espressi e di evidenziare in particolare quelli differenzialmente espressi, sia per quanto riguarda gli mRNA che i miRNA. Ho anche analizzato il *crosstalk* tra mRNA e miRNA, impiegando due diversi metodi: sovraespressione dei miRNA e immunoprecipitazione di Ago2, seguita da RNA seq.

Gli esperimenti di sovraespressione sono stati condotti con 5 diversi miRNA e in tutti i casi ho trovato più geni che hanno aumentato il loro livello piuttosto di geni che l'hanno diminuito. Questi risultati suggeriscono che i miRNA potrebbero avere un ruolo nella stabilizzazione degli mRNA, oppure potrebbero avere un ruolo in un doppio meccanismo negativo, inibendo a loro volta fattori negativi. Confrontando i geni che diminuiscono di livello dopo la sovraespressione di miRNA con i geni arricchiti dall'immunoprecipitazione con Ago2, abbiamo trovato diversi geni candidati per essere sotto il controllo inibitorio dei miRNA.

Il muscolo scheletrico ha una grande plasticità e può sopportare un notevole stress. Per studiare questo aspetto del muscolo abbiamo sottoposto le cellule in differenziamento a

stiramento meccanico. Con l'RNA seq abbiamo ottenuto un'approfondita visione del trascrittoma in risposta allo stiramento. Abbiamo effettuato queste analisi a due diversi tempi dopo lo stiramento ed abbiamo trovato che nel periodo immediatamente successivo allo stimolo vengono sovraespressi geni implicati nella risposta immunitaria, mentre successivamente sono attivati i geni codificanti proteine muscolari strutturali, quando allo stesso tempo la risposta immunitaria viene inibita.

Sono convinta che la nostra approfondita analisi del trascrittoma che include l'interazione di mRNA e miRNA durante la miogenesi, possa contribuire ad una maggiore comprensione di processi che regolano lo sviluppo muscolare.

Chapter 1
Introduction

1.1 Skeletal muscle

Skeletal muscle is one of the most highly organised structures in the biological world, and is primarily involved in the execution of voluntary movement (1). Skeletal muscle is composed of a number of muscle fibre types that differ with respect to their contractile, metabolic and molecular properties (2). The characteristics of a skeletal muscle is its contractile function

1.1.1 – Structure

Skeletal muscle is composed of many individual muscle fibres (3). Each fibre is covered in endomysium and beneath this, resides the cell membrane or sarcolemma. Structurally, each muscle fibre is composed of many protein bundles called myofibrils, which in turn comprise of alternate dark and light staining filaments (3). The sarcomere is composed of the thin (actin) filaments, the thick (mostly myosin) filaments, and the giant filamentous molecule titin. Titin alos known as connectin is the molecular spring that is responsible for the passive elasticity of the muscle. It connects the Z line to the M line in the sarcomere. The thin filaments are anchored in the Z-line, where they are cross-linked by α-actinin. The thick filament is located centrally in the sarcomere and constitutes the sarcomeric A-band. The myosin heads, interact with actin during activation. In the A-band titin is inextensible due to its strong interaction with the thick filament. The distance from one Z-line to the next is defined as one sarcomere, the smallest integral contractile unit (1).

1.1.2 Myogenesis

The formation of skeletal muscle involves a series of steps in which multipotential mesodermal precursor cells become committed to a muscle cell fate and then proliferate as myoblasts until they encounter an environment lacking mitogens, at which point they exit the cell cycle and differentiate (4). Skeletal, cardiac, and smooth muscle are each derived from mesodermal precursor cells in different regions of the embryo. Although these three different muscle cell types express many of the same muscle-specific genes, each type is unique with respect to the spectrum of muscle genes they express, their morphology, their ability to divide, and their contractile properties. Therefore, if any shared myogenic program exists it must be modified by different regulatory factors to generate the diversity of three muscle cell types.

Adult myogenic cells are derived mainly from muscle satellite cells that are specialized myogenic cells found during late fetal development. In mice the myogenic precursors in the dermomyotome express Pax3, Pax7 and low levels of the myogenic determination factor Myf5. The paired-domain transcription factors Pax3 and Pax7 act upstream of the primary myogenic basic helix–loop–helix (bHLH) transcription factors Myf5 and MyoD, in myogenic specification. When myogenic determination genes MyoD and Myf-5 are activated the muscle precursor cells are committed to become myoblasts and migrate into the adjacent embryonic connective tissue, or mesenchyme and express other muscle specific genes such as myogenin and MRF4, that after a period of proliferation, induce the fusion of myoblasts into multinucleated and highly specialized skeletal muscle cells called myotubes or myofibres (5). A typical myofibre is cylindrical, large (measuring 1-40 mm in length and 10-50 μm) and multinucleated (containing as many as 100 nuclei). Muscle fiber is a cylindrical multinucleate cell composed of numerous myofibrils that contracts when stimulated. Muscle fibres are the basic contractile units of skeletal muscle and are individually surrounded by a layer of connective tissue and grouped into bundles to form skeletal muscle (6, 7).

1.1.3 Differentiation

Skeletal muscle development involves an initial period of myoblast replication followed by a phase where some myoblasts continue to proliferate while other undergoes terminal differentiation. The process of differentiation involves permanent cessation of DNA synthesis, activation of muscle specific gene function and fusion of single cells into multi nucleated muscle fibers (51). Skeletal muscle differentiation is a tightly regulated process that requires the coupling of muscle-specific gene expression with the terminal withdrawal from the cell cycle. The MyoD family of basic helix-loop-helix (bHLH) skeletal muscle specific transcription factors plays a pivotal role in initiating skeletal muscle differentiation. MyoD and Myf-5 are the most important members of this protein family and are expressed in proliferating and undifferentiated cells, whereas the expression of other bHLH transcription factors such as myogenin and MRF-4 occurs only during differentiation or in adult mature skeletal muscle, respectively. These four myogenic bHLH proteins are known as myogenic regulatory factors (MRFs). The MRFs are able to bind DNA both in the form of homodimers as well as in the form of heterodimers with ubiquitously expressed transcription factors called E proteins. Their binding sites, called E boxes, share the consensus sequence CANNTG. The binding of the bHLH/E protein to

DNA is essential for the activation of the muscle specific differentiation program (8). The differentiation process starts during embryogenesis, when the appropriate environmental stimuli are encountered. In the myotome of the embryo, the MRF's present in determined myoblasts (MyoD and Myf-5) initiate a cascade of events that leads to the activation of other transcription factors, the MEF2 family, which are necessary for the transcription of myogenin and other skeletal muscle specific genes. Myogenin itself can also activate MEF2, creating a positive regulatory loop that ensures the maintenance of appropriate levels of these proteins in differentiating skeletal muscle (8, 9). Amongst the earliest muscle genes to be expressed in the myotome are desmin, Titin and α-actin. The expression of the myosin heavy chain gene occurs almost a day after the accumulation of the α-actin protein. Committed myoblasts initiate their transformation into differentiated myotubes by first expressing all the major structural proteins. Then, the other muscle genes are activated following a strict temporal regulation during the embryonic, fetal and postnatal development (10). At the same time a system intervenes to allow the cell to exit from the cell cycle thus permitting tissue specific gene expression, cell fusion and the formation of multinucleated myotubes (8). These events involve both muscle specific transcription factors and ubiquitous cell cycle regulatory proteins. In fact MyoD, that represents the major coordinator of skeletal muscle differentiation, is also able to induce the expression of p21, a potent inhibitor of Cyclin-Dependent Kinases (Cdks), thus forces the cell cycle withdrawal (11). This event inhibits the Retinoblatoma protein (pRb) phosphorylation thus promoting its activation and allowing it to sequester the E2F transcription factor, thus blocking cell cycle progression (12). Moreover MyoD and pRb can directly interact with each other (13).

Many other cell cycle regulatory proteins are involved in this complex picture. In fact the p53 tumor suppressor protein is essential in the process of skeletal muscle differentiation, since p53-impaired cells fail to differentiate (14, 15) even though cell cycle withdrawal takes place in a p53-independent manner. Indeed the p53 tumor suppressor is crucial to elevate un-phosphorylated pRb levels to a threshold sufficient to terminally maintain the cell in G0/G1 (17) and to activate together with MyoD the expression of late muscle differentiation markers (18, 19). The last step in the differentiation process is represented by myoblasts fusion. This event correlates with fibronectin secretion onto the extracellular matrix, to which differentiating cells attach using the α5β1 integrin. After that, myoblasts start aligning. This step is mediated by several cell membrane

glycoproteins, including several cadherins and CAMs. Ultimately cell fusion occurs, mainly through the action of a set of metalloproteinases called meltrins and multinucleated differentiated myofibres are formed (20).

1.1.4 Skeletal muscle contraction

Muscle contraction consists of a cyclical interaction between myosin and actin driven by the concomitant hydrolysis of adenosine triphosphate (ATP) (21). Myosin and actin, the components which respectively form the thick and thin filaments were amongst the first proteins to be purified with reference to muscle function (22). The hexametric protein myosin comprises two heavy chains (220 KDa) and four light chains (20-25 KDa), and forms the thick filaments. The terminus of myosin forms a globular head region required for the hydrolysis of ATP and binding of actin. The four light myosin chains are located between the globular head and the carboxy-terminal rod region (1). Thin muscle filaments are comprised of several proteins; however actin (43 KDa) is by far the most abundant constituent. Thick and thin regions of physical overlap form in which globular myosin heads project from the thick filaments to interact with thin actin filaments. ATP hydrolysis mediates a conformational change in the globular myosin heavy chain head region, resulting in an interaction between the globular head and actin further along the filament and inducing a shortening of the muscle fibre.

Multiple isoforms of myosin heavy chain (MHC) exist, which comprise a family of molecular motors able to modulate the speed of skeletal muscle contraction (23). The contractile speed of a particular muscle fibre may therefore be determined, in part, by the isoform of MHC protein which it expresses. The sarcomeric MHC family consists of at least eight known isoforms, each encoded by a distinct gene located in two multigenic regions on two separate chromosomes (24). Six genes are encoded by a $300 - 600Kb$ segment on human and mouse chromosomes 17 and 11 respectively, in a cluster arrangement in the order MyH3/MyH2, MyH1/MyH 4, MyH 8/ MyH13 (5' $-$ 3'). The MyH2, MyH1 and MyH4 genes encode the protein isoforms commonly termed MHC IIA, IIX and IIB. Of the eight sarcomeric isoform genes of MHC, four are known to be expressed in adult skeletal muscle: one "slow-twitch" (Type I) muscle associated MHC isoform is encoded by the MyH7β gene and three "fast-twitch" (Types IIA, IIX and IIB) muscle associated isoforms, associated with increasing contractile speed. A combination

of the latter "fast-twitch" isoforms account for over 90% of MHC in adult skeletal muscle (25).

1.1.5 Skeletal muscle metabolism

Muscle contraction is driven by the hydrolysis of adenosine triphosphate (ATP) (21) which may be derived from the metabolism of fatty acids (β-oxidation) or carbohydrates (glycolysis). The β-oxidation of fatty acids takes place in the mitochondria of muscle fibres. Endothelial lipoprotein lipase (LPL) is involved in the transport of fatty acids from the circulatory system into the myocellular compartment (26). Hormone sensitive lipase (HSL) liberates free fatty acids from the intramyocellular lipid (IMCL) pool which are transported into the mitochondrion for β-oxidation by carnitine palmitoyltransferase 1 (CPT-1). The production of ATP via the β-oxidation of fatty acids is an oxygen-dependent process.

Carbohydrates reach the myofibre from the circulatory system and may be stored as glycogen, converted to triglycerides, or metabolised via glycolysis. In contrast to β-oxidation, the metabolism of carbohydrates via glycolysis is an oxygen independent process; however metabolism under these conditions leads to the production of lactate (26).

1.2 miRNA and its role in general

A microRNA is a small non-coding RNA molecule (22 nucleotides) found in plants, animals, and some viruses, which functions in transcriptional ans post-transcriptional regulation of gene expression. Encoded by nuclear DNA in plants and animals and by viral DNA in certain viruses whose genome is based on DNA, miRNAs function via base pairing with complementary sequences within mRNA molecules, usually resulting in gene silencing via translational repression or target degradation. The human genome may encode over 1000 miRNAs, which may target about 60% of mammalian genes and are abundant in many human cell types.

1.2.2   miRNA biogenesis

miRNAs are transcribed into long transcripts mainly by RNA polymerase II, although there are also evidences implicating RNA polymerase III in the transcription of some miRNAs (27). Most of them are polyadenylated in its 3'end and capped at its 5'extremity,

like mRNAs. In the canonical pathway (Fig 10) these primary-miRNAs (pri-miRNAs) folds into a stem loop structure that will be further processed by two RNAse III endonuclease, Drosha and Dicer. In the nucleus, the hairpin structure is cleaved from the flanking regions originating the precursor-miRNA (pre-miRNA) that is ≈70 nucleotide long. This first processing step is catalyzed by Drosha that is helped by a cofactor, DGCR8 (DiGeorge syndrome critical region gene 8) (Pasha in Drosophila). This complex is called the Microprocessor. DGCR8 contains two dsRNA-binding domains that directly interact with the stem-loop and with the flanking region, serving as a molecular anchor to Drosha that carries out the cleavage reaction. The cleavage produces highly exact extremities and is highly regulated. The production of pre-miRNAs not always requires the participation of the microprocessor complex. In fact, a rare alternative pathway has been identified initially in fly and nematodes, but also present in mammals (28, 29, 30). This pathway uses the splicing machinery to liberate introns that mimic the features of pre-miRNAs. These structures are called mirtrons. After being spliced they enter the normal miRNA processing pathway.

Pre-miRNAs are then exported to the cytoplasm by Exportin-5 in a Ran-GTP dependent way, where they will be further processed. In the cytoplasm the terminal loop of the pre-miRNA is cleaved originating a mature dsRNA of approximately 22 nucleotides of length. This step is carried out by Dicer. The PAZ domain of Dicer binds to the 3'overhangs of the pre-miRNAs and this binding determines the cleavage site since that Dicer's catalytic sites are located precisely two helical turns away from the PAZ domain (bound to the pre-miRNA). In this step, Dicer is assisted by the Tar RNA Binding Protein – TRBP (know as Loquacious in Drosophila), another dsRNA binding protein. At the end of this last processing step Argonaute 2 is recruited to the complex Dicer/TRBP leading to the unwinding of the duplex. At this stage one of the strands, the mature miRNA, is preferentially incorporated into the complex that will repress target gene expression – the RNA-Induced Silencing Complex (RISC). The choice of the mature strand is based on the thermodynamic stability of the two ends of the duplex. The complementary strand (miRNA*) in most cases is degraded. However there are growing evidences that both strands can be incorporated into the RISC complex in a functional way (31).

The key proteins of the RISC complex are the Argonautes (AGO). These proteins contain three highly conserved domains, PAZ, MID and PIWI domains, that interact with the miRNAs. In mammals there are four AGO that function in miRNA repression (Ago1 to

Ago4). Different AGO proteins seems to have different specificity to the miRNA or siRNA pathway. Another crucial factor for miRNA-repression is the GW182 protein. This protein interact directly with AGO proteins and are thought to be the effectors of AGO. There are other proteins interacting with RISC to modulate miRNA function. This is the case of Fragile X Mental Retardation Protein (FMRP), which binds RNA molecules and might modulate translation. Also RNA Helicase RCK/p54, which is a p-body component, is thought to be essential to induce repression. Finally TRIM32 was recently seen to bind the RISC components enhancing in this way its activity. However further studies are required for a better understanding of the proteins that modulate this complex process of miRNA-mediated repression (32, 33, 34).

1.2.3 mRNA degradation

There are two mechanism for mRNA degradation, the first being when mRNAs undergo endolytic cleavage by Ago2 (37) and the second is when mRNAs undergo poly(A) removal by deadenylases (38).

1.2.4 Translational activation

Some studies indicate that miRNAs can stimulate translation under specific conditions. Two proteins are important for this ability of miRNAs to activate growth arrested mammalian cells. One is Ago2, the other one is FXR1, an RNA binding protein homologus to fragile X mental retardation protein FMR1/ FMRP. Surprisingly it was found that human Ago2 activates translation of target mRNAs on cell cycle arrest caused by serum starvation or contact inhibition, while it normally repressed translation of same target mRNAs in proliferating cells. FXR1 associates with Ago2 and helps to mediate the positive influence of miRNAs on translation (35, 36). Like translational repression, such activation requires base pairing between mRNA and seed region of miRNA.

1.3 miRNA target prediction

One of the most critical part in the study of miRNAs is the identification of the target genes they regulate. The study of the molecular mechanism implicated in target recognition, together with computational approaches were soon translated into basic principles used in the development of bioinformatic tools that could predict miRNAs targets.

The majority of animal miRNAs displays only modest base-pairing to their targets in contrast to what happens in plants, where the base pairing is perfect. Historically, miRNAs are known to regulate the 3'UTR of the target genes. This was experimentally demonstrated with the first miRNAs identified and it was also assumed as an in silico convenience that was further confirmed experimentally. But the use of these predictive algorithms has left underestimated the possibility that miRNAs might regulate other regions, such as the 5'UTR or even the coding region of the mRNAs. In fact, experiments involving artificial and natural mRNAs have shown that their 5'UTR can be targeted by miRNAs (39). Recent reports have also started to address the possibility that miRNAs can target the Open Reading Frame (ORF) of certain genes.

The interaction between miRNA and mRNA are through base pairing - most of the times imperfect base pairing. The most important region of the miRNA is the so called "seed" region. According to the seed "rule", the interaction between miRNA and mRNA requires a contiguous and perfect (or nearly perfect) Watson-Crick base pairing of the 5' nucleotides 2-8 of the miRNA. Another point to take into consideration when discussing the miRNA/mRNA interaction is the presence of multiple sites in the same 3'UTR. In fact this seems to lead to a more efficient mRNA repression. Some algorithms also take into account the conservation between related species of the miRNA-binding site. Finally one must consider that mRNAs have a secondary structure that might block miRNAs binding (40, 41, 42).

Giving different weight to each of these parameters, several algorithms were developed to predict miRNAs targets. The most known and robust ones are: TargetScan, PicTar, Miranda and PITA.

Although the principles used for target recognition are quite well established and accepted, and considering that different algorithms have different sensibilities, validation of the predicted targets is always required.

1.4 miRNA in skeletal muscle tissue

One of the first evidences that miRNAs might play a crucial role in adult skeletal muscle came from a study in sheep (43). The aim of the study was identifying the gene responsible for the hypertrophic phenotype of the Texel sheep. The authors found a point mutation in the 3'UTR of the myostatin gene that creates a new binding site for the miRNA-1 and miRNA-206. In these breed of animals, myostatin, a negative regulator of

muscle growth, is down-regulated by these two muscle-specific miRNAs inducing an exacerbated muscle growth.

Important evidence highlighting the essential role that miRNAs play in muscle came from loss-of-function experiments. In this case, O'Rourke et al (44), generated a muscle specific conditional knock-out of Dicer, the RNAse III enzyme required for miRNA maturation. In these mice, the expression of Cre recombinase was under the control of MyoD regulatory elements, and therefore started to be expressed from embryonic day 9.5. All Dicer skeletal muscle mutants died just after birth. They showed severe defects in skeletal muscle embryonic development that was mainly translated in muscle hypoplasia with hypertrophy of the few remaining fibers. This hypoplasia was attributed to an increase apoptosis rather than a defect in myofibers formation. These results are similar to the ones obtained by (45, 46) in which the knock-down of the muscle specific miRNA-1 in Drosophila caused arrest in embryogenesis with disorganized muscle development and aberrant expression of muscle-specific genes. Altogether these results anticipate a fundamental role of miRNAs in different aspects of skeletal muscle biology.

1.5 RNA seq and Transcriptome analysis

The transcriptome is the set of all RNA molecules, including mRNA, rRNA, tRNA, and other noncoding RNA produced in one or a population of cells. In multicellular organisms, nearly every cell contains the same genome and thus the same genes. However, not every gene is transcriptionally active in every cell. Different cells show different patterns of gene expression. These variations underlie the wide range of physical, biochemical, and developmental differences seen among various cells and tissues and may play a role in the difference between health and disease. Thus, by collecting and comparing transcriptomes of different types of cells or tissues, researchers can gain a deeper understanding of what constitutes a specific cell type and how changes in transcriptional activity may reflect or contribute to disease. RNA-seq is a new method in RNA sequencing to study mRNA expression. By considering the transcriptome, it is possible to generate a comprehensive picture of what genes are active at various stages of development. Not until a decade back the enormous complexity of human genome has been realized (47). Transcritome analysis has traditionally focused on cytoplasmic poly(A)+ RNA, which excluded non coding part of the genome (eg, tRNA, linc RNA, miRNA , siRNA) and hence a very large segment of important information remained

inadequately considered. But now that the non coding part of the genome is emerging as the most important regulatory population of the cell, advanced technique oriented towards delving deeper inside the transcriptome with better sensitivity to trace minor changes in gene expression and for discovering new transcripts.

Next generation sequencing provides the way for the much sought after need of understanding the complexity of transcriptome. With a dynamic range to detect subtle changes in expression level in a hypothesis-neutral environment, next generation sequencing helps provide an understanding of biological response to stimuli or environmental changes. The potential of RNA seq technology for studying transcriptome has been described by Wang et al (48). The application of RNA seq technology is vast from finding differential gene and transcript expression to revealing unannotated transcripts (49, 50).

1.6 Outline of the thesis

In the present thesis, full transcriptome analysis of human skeletal muscle during development, through proliferation to differentiation has been presented in four chapters. Chapter 2 deals with the entire transcriptome sequence of human myoblast and myotube cells along with adult skeletal muscle tissue. The transcriptome analysis includes miRnome sequencing also. Entire list of differentially expressed genes has been presented. miRNA candidates have also been found. Genes of specific protein families with substantial expression has been shown. To understand the biological meaning of the genes, GO and pathway analysis results have been presented. Argonaute 2 immunoprecipitation has been performed to capture the genes bound to Ago2 and hence supposedly bound to the RISC complex,that gave us the probable candidate genes under miRNA repressing action.

In Chapter 3 differentially upregulated miRNA in myotubes presented in chapter 1 have been analyzed for their functional role. Genes enriched with Ago2 IP (chapter 1) were used for the analysis of target recognition of the miRNA.

Chapter 4 shows the transcriptome behaviour of myotubes differentiated under 3D culture conditions and their similarity and difference in comparison with muscle tissue and 2D culture has also been presented.

In Chapter 5 The response to mechanical stretch of differentiating muscle cells have been presented. Prompt response and delayed response of the transcriptome to mechanical stretch has been studied and presented in this chapter.

1.7References:

1. McNally, E. M., K. A. Lapidos, et al. (2006). Skeletal Muscle Structure and Function. Principles of Molecular Medicine**: 674-681.

2 Pette, D. and R. S. Staron (2001). "Transitions of muscle fiber phenotypic profiles." Histochemistry and Cell Biology **115**(5): 359-372.

3 Dubowitz, V. (2007). Muscle biopsy: a practical approach, Elservier.

4 McKinsey TA, Zhang CL, Olson EN (2001) Control of muscle development by dueling HATs and HDACs. Curr Opin Genet Dev 11: 497-504

5 Chen JCJ, Goldhamer DJ (2003) Skeletal muscle stem cells. Reprod Biol Endocrinol 1: 101-107.

6 Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2002) Molecular Biology of the Cell. Fourth Edition, Garland Science.

7 Lodish H, Baltimore D, Berk A, Zipursky SL, Matsudaria P and Darnell J (1995) Molecular Cell Biology – Third edition, Scientific American Books.

8 Lassar AB, Skapek SX, Novitch B (1994) Regulatory mechanisms that coordinate skeletal muscle differentiation and cell cycle withdrawal. Curr Opin Cell Biol 6: 788-794.

9 Weintraub H, Davis R, Tapscott S, Thayer M, Krause M, Benezra R, Blackwell K, Turner D, Rupp R, Hollenberg S, Zhuang Y and Lassar A (1991) The MyoD gene family: nodal point during specification of the muscle cell lineage. Science 251: 761-766.

10 Buckingham M (1992) Making muscle in mammals. Trends Genet 8: 144-149 Carson JA, Nettleton D, Reecy JM (2002) Differential gene expression in the rat soleus muscle during early work overload-induced hypertrophy. FASEB Journal 16: 207-219.

11 Halevy O, Novitch BG, Spicer DB, Skapek SX, Rhee J, Hannon GJ, Beach D, Lassar AB (1995) Correlation of terminal cell cycle arrest of skeletal muscle with induction of p21 by MyoD. Science 267: 1018-1021.

12 Wiman KG (1993) The retinoblastoma gene: role in cell cycle control and cell differentiation. FASEB J. 7: 841-845.

13 Gu W, Schneider JW, Condorelli G, Kaushal S, Mahdavi V and Nadal-Ginard B (1993) Interaction of myogenic factors and the retinoblastoma protein mediates muscle cell commitment and differentiation. Cell 72: 309-324.

14 Soddu S, Blandino G, Scardigli R, Coen S, Marchetti A, Rizzo MG, Bossi G, Cimino L, Crescenzi M and Sacchi A (1996) Interference with p53 protein inhibits hematopoietic and muscle differentiation. J Cell Biol 134: 193-204.

15 Mazzaro G, Bossi G, Coen S, Sacchi A, Soddu S (1999) The role of wild-type p53 in the differentiation of primary hemopoietic and muscle cells. Oncogene 18: 5831-5835.

16 Porrello A, Cerone MA, Coen S, Gurtner A, Fontemaggi G, Cimino L, Piaggio G,

17 Sacchi A and Soddu S (2000) p53 regulates myogenesis by triggering the differentiation activity of pRb. J Cell Biol 151: 1295-1303.

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

19 Magenta A, Cenciarelli C, De Santa F, Fuschi P, Martelli F, Caruso M and Felsani A (2003) MyoD stimulates RB promoter activity via the CREB/p300 nuclear transduction pathway Mol Cell Biol 23: 2893-2906.

20 Gilbert SF (1997) Developmental Biology – Fifth edition, Sinauer Associates, Inc. Publishers, Sunderland, Massachuttes.

21 Rayment, I., H. Holden, et al. (1993). "Structure of the actin-myosin complex and its implications for muscle contraction." Science 261(5117): 58-65.

22 Block, S. M. (1996). "Fifty Ways to Love Your Lever: Myosin Motors." Cell 87(2): 151-157.

23 Rinaldi, C., F. Haddad, et al. (2008). "Intergenic bidirectional promoter and cooperative regulation of the IIx and IIb MHC genes in fast skeletal muscle." American Journal of Physiology-Regulatory integrative and comparative physiology 295(1): R208-R218.

24 Weiss, A. and L. A. Leinwand (1996). "The mammalian myosin heavy chain gene family." Annual Review of Cell and Developmental Biology 12: 417-439.

25 Allen, D. L., C. A. Sartorius, et al. (2001). "Different pathways regulate expression of the skeletal myosin heavy chain genes." Journal of Biological Chemistry 276(47): 43524-43533.

26 Fluck, M. (2006). "Functional, structural and molecular plasticity of mammalian skeletal muscle in response to exercise stimuli." J Exp Biol 209(12): 2239-2248.

27 Borchert,G.M., Lanier,W., and Davidson,B.L. (2006). RNA polymerase III transcribes human microRNAs. Nat. Struct. Mol. Biol. 13, 1097-1101.

28 Berezikov,E., Chung,W.J., Willis,J., Cuppen,E., and Lai,E.C. (2007). Mammalian mirtron genes. Mol. Cell 28, 328-336.

29 Okamura,K., Hagen,J.W., Duan,H., Tyler,D.M., and Lai,E.C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. Cell 130, 89-100.

30 Ruby,J.G., Jan,C.H., and Bartel,D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. Nature 448, 83-86.

31 Guo,L. and Lu,Z. (2010). The fate of miRNA* strand through evolutionary analysis: implication for degradation as merely carrier strand or potential regulatory molecule? PLoS. One. 5, e11387.

32 Carthew,R.W. and Sontheimer,E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell 136, 642-655.

33 Fabian,M.R., Sonenberg,N., and Filipowicz,W. (2010). Regulation of mRNA translation and stability by microRNAs. Annu. Rev. Biochem. 79, 351-379.

34 Krol,J., Loedige,I., and Filipowicz,W. (2010). The widespread regulation of microRNA biogenesis, function and decay. Nat. Rev. Genet. 11, 597-610.

35 Vasudevan S, Tong Y, Steitz JA. Switching from repression to activation: miRNAs can upregulate translation. Science 2007: 318: 1931-40.

36 Vasudevan S, Steitz JA: AU rich element upregulation of translation by FXR1 and argonaute 2. Cell 2007: 128:1105-18.

37 Yekta S, Shih IH, Bartel DP. microRNA mediated cleavage of HOXB8 mRNA. Science 304: 594 – 596 (2004).

38 Wu L, Fan J, Belasco JG. MicrRNA directs rapid deadenylation of mRNA. Proc Natl Acad USA. USA 103: 4034 – 4039 (2006).

39 Lytle,J.R., Yario,T.A., and Steitz,J.A. (2007). Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. Proc. Natl. Acad. Sci. U. S. A 104, 9667-9672.

40 Brodersen,P. and Voinnet,O. (2009). Revisiting the principles of microRNA target recognition and mode of action. Nat. Rev. Mol. Cell Biol. 10, 141-148.

41 Du,T. and Zamore,P.D. (2005). microPrimer: the biogenesis and function of microRNA. Development 132, 4645-4652.

42 Fabian,M.R., Sonenberg,N., and Filipowicz,W. (2010). Regulation of mRNA translation and stability by microRNAs. Annu. Rev. Biochem. 79, 351-379.

43 Clop,A., Marcq,F., Takeda,H., Pirottin,D., Tordoir,X., Bibe,B., Bouix,J., Caiment,F., Elsen,J.M., Eychenne,F., Larzul,C., Laville,E., Meish,F., Milenkovic,D., Tobin,J., Charlier,C., and Georges,M. (2006). A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. Nat. Genet. 38, 813-818.

44 O'Rourke,J.R., Georges,S.A., Seay,H.R., Tapscott,S.J., McManus,M.T., Goldhamer,D.J., Swanson,M.S., and Harfe,B.D. (2007). Essential role for Dicer during skeletal muscle development. Dev. Biol. 311, 359-368.

45 Kwon,C., Han,Z., Olson,E.N., and Srivastava,D. (2005). MicroRNA1 influences cardiac differentiation in Drosophila and regulates Notch signaling. Proc. Natl. Acad. Sci. U. S. A 102, 18986-18991.

46 Sokol,N.S. and Ambros,V. (2005). Mesodermally expressed Drosophila microRNA-1 is regulated by Twist and is required in muscles during larval growth. Genes Dev. 19, 2343-2354.

47 Frith MC, Pheasant M and Mattick JS. The mazing complexity of the human transcriptome. Eur J Hum Genetics (2005) 13: 894 – 897.

48 Wang Z, Gerstein M, Snyder M (2009). RNA-seq: A revolutionary tool for the transcriptomics. Nature Rev. Genetics 10(1): 57 – 63.

49 Trapnell C, Williams BA, Pertea G, Mortazavi A, Gordon B, Marijke J, Salzberg, Steven L, Barbara J, Pachter L (2010). Transcript assembly and quantification by RNA seq reveals unannotaed transcripts and isoform switching during cell differentiation. Nat Biotechnol 28(5): 511-515.

50 Trapnell C, Roberts A, Goff L et al (2012). Differential gene and transcript expression analysis of RNA seq experiments with Tophat and Cufflinks. Nat Protoc 7(3): 562 - 78.

51 Clegg CH, Linkhart TA, Olwin BB, Hauschka SD (1987). Growth factor control of skeletal muscle differentiation: commitment to terminal differentiation occurs in G1 phase and is repressed by fibroblast growth factor. J Cell Biol. 105 : 949 – 956.

Chapter - 2

Myoblast, myotube and skeletal muscle transcriptome, miRnome dynamics and functional profiling

2.1 Introduction:

Understanding the dynamics of muscle transcriptome during muscle growth and development is necessary to uncover the complex mechanisms underlying muscle development. The formation of skeletal muscle during vertebrate embryogenesis requires commitment of mesodermal precursor cells to skeletal muscle lineage, withdrawal of myoblasts from cell cycle and transcriptional activation of dozens of muscle structural genes. MyoD, myogenin, Myf5 and MRF4 act to establish myoblast identity and to control terminal differentiation. Myogenic bHLH factors interact with components of the cell cycle machinery to control withdrawal from the cell cycle and act combinatorial with other transcription factors to induce skeletal muscle transcription. Elucidation of these aspects of the myogenic program is leading to a detailed understanding of the regulatory circuits controlling muscle development. Skeletal muscle differentiation entails the coupling of muscle-specific gene expression to terminal withdrawal from the cell cycle. Several models have recently been proposed which attempt to explain how regulated expression and function of myogenic transcription factors ensures that proliferation and differentiation of skeletal muscle cells are mutually exclusive processes. Skeletal muscle is the dominant organ system in locomotion and energy metabolism. Postnatal muscle grows and adapts largely by remodelling pre-existing fibres, whereas embryonic muscle grows by the proliferation of myogenic cells.

The role of miRNA in muscle development is widely acknowledged now. Many groups of researchers have shown the important regulatory roles miRNA play during muscle development. miR-145 and 143 are known to regulate smooth muscle plasticity (7). Sun et al have shown the involvement of miR-24 in skeletal muscle differentiation (8). miR-181 has been shown to target Hox-A11 during myoblast differentiation (9). miR-1, miR-133 and miR-206 are skeletal muscle specific miRNA and have been extensively studied by many researchers (10, 11, 12).

To understand the complicated molecular networks working behind muscle development we need to obtain a thorough knowledge of the transcriptome first before the interactive networks are unravelled. We have attempted to do so by carrying out transcriptome sequencing at a massive scale. We performed the experiments in replica to have more reliable results.

The advent of next generation sequencing technology gives us the opportunity of capturing the full transcriptomic panorama of a cell at a given point of time. We used

NGS technology to our advantage and performed RNA seq of proliferating and differentiating primary human muscle cells along with adult human skeletal tissue. The functional significance of the genes differentially expressed from one developmental stage to next was studied by using Gene Ontology and pathway analyses. We also evaluated miR transcriptome profiles during the proliferative and differentiating stages of primary human muscle development. Since muscle cells ultimately grow into skeletal tissue we considered that it was worth to study the miRNA profile of adult human skeletal muscle tissue also, to understand how the miRNA population modulates itself from the cellular level to the tissue which is the ultimate functional form in the body. We were also interested in understanding the regulatory roles played by miRNAs. But we wanted to do it independent of any prediction algorithms as there are biases in prediction and also a huge number of false positives. So, we performed Ago2-RIP (Argonaute 2 RNA immunoprecipitation) experiments to capture the mRNA population that is associated with Ago-2 protein. Since Ago2 is a component of RISC complex which carries the miRNA and eventually chops off the mRNA or inhibits its translation, the mRNA associated with Ago2 most likely presents the population which is controlled by miRNA. We did the RNA seq of the immunoprecipitates and performed differential expression analysis in comparison with non Ago2 pull down RNA population. The Ago2 RIP was performed for myoblasts and myotubes. The mRNA population enriched with Ago2 was compared with mRNA differentially down regulated in myoblasts and myotubes. We did this to find out the population of genes repressed. The genes found in common were subjected to GO analyses and we saw interesting results from the analyses. The schematic representation of the experimental design is shown below:
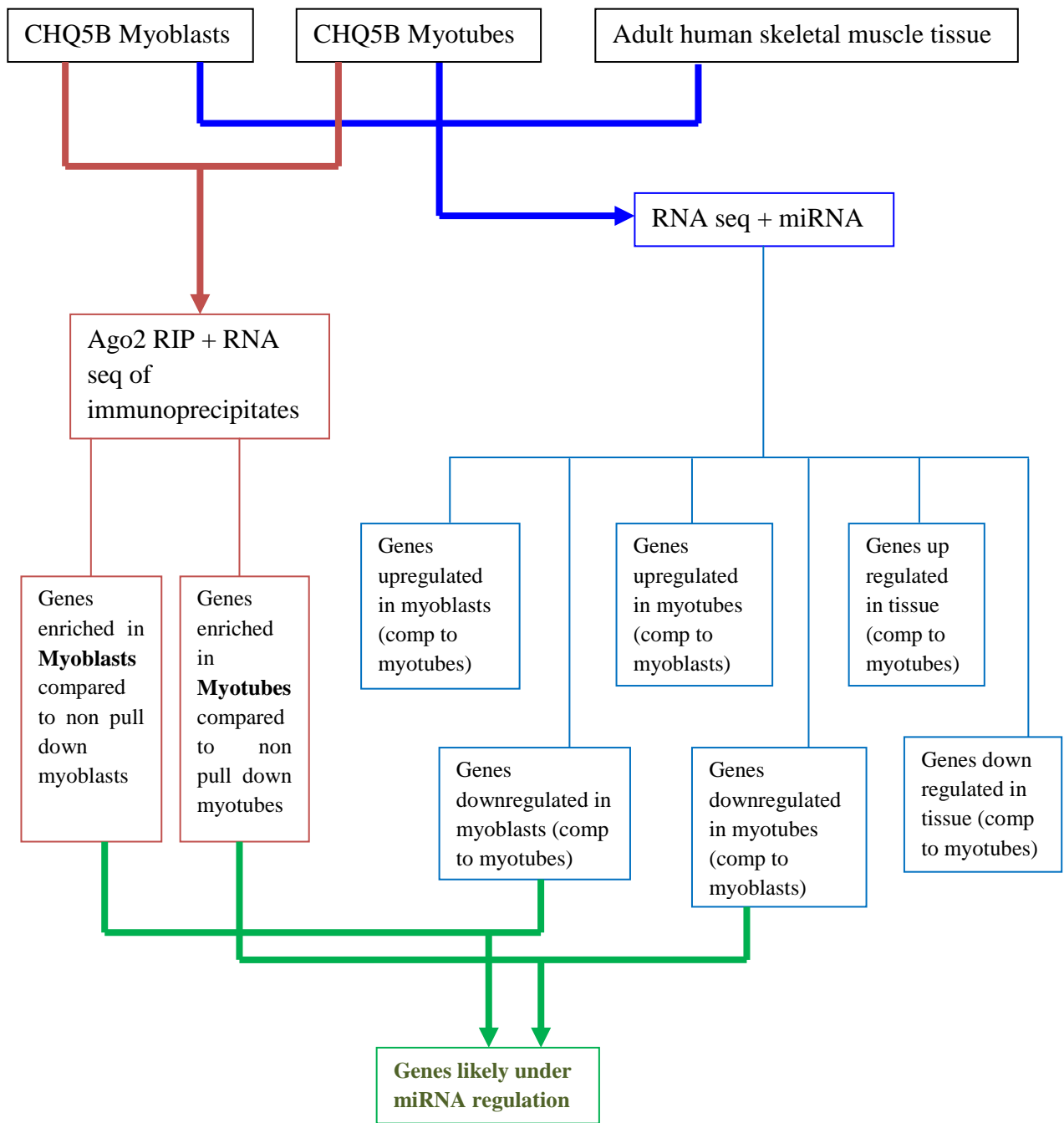
Schematic representation of experimental plan. All the RNA seq, miRNA seq and Ago2 pull downs were performed in replica. Arrows in green show the comparison of genes enriched in Ago2 pull down versus genes down regulated. The genes found common from the comparison are most likely repressed by miRNA.

2.2 Materials and methods:

2.2.1 Cell Culture (Human primary skeletal muscle cells (CHQ5B)):

CHQ5B primary human myoblasts were kindly provided by Dr. V. Mouly (URA, CNRS, Paris, France) (24). CHQ5B human myoblasts were isolated from the quadriceps of a newborn (5 days post-natal) without any sign of neuromuscular disorders and the protocols used for this work were in full agreement with the current legislation on ethical rules. This strain of cells can achieve 55-60 divisions before reaching proliferative senescence.

Growth conditions: DMEM (Gibco - Invitrogen) supplemented with 20% Fetal Bovine Serum (Gibco, Life Technologies) and 50µg/ml gentamycin.

Differentiation conditions: DMEM supplemented with 2% horse serum (GibcoBRL) and 50µg/ml gentamycin.

Differentiation of myoblasts into myotubes has to be induced by serum withdrawal without letting the myoblast culture reach confluency, as that reduces the myoblast population in the culture. Myotube formation can be observed after two days since serum withdrawal.

2.2.2 RNA extraction from cells:

2.2.2.1. a Poly (A) extraction from cells:

Polyadenylated RNA was extracted directly from cells using QuickPrep micro mRNA purification kit (Amersham Biosciences). Cells were scrape collected after PBS washing, snap frozen in dry ice and stored at -80°C until use. 400µl extraction buffer (buffered aqueous solution containing guanidium thiocyanate and N-lauroyl sarcosine) was added to the pelleted cells and vortexed until homogenous suspension was achieved. This suspension was diluted with 800 µl of elution buffer (10mM Tris HCl, pH 7.5, 1mM EDTA) and mixed using the vortex. The mixture was centrifuged at 12,000g for 1 minute and the clear cellular homogenate was added to the Oligo(dT) beads (25mg/ml oligo dT cellulose suspended in buffer) pellet. Cellular homogenate and oligo (dT) beads were incubated together for 5 minutes at 70 -75 °C. This causes the denaturation of RNA and enhances the binding of poly (A) tail of RNA with the oligo dT beads. The sample was incubated at room temperature for 30 minutes with gentle agitation. Supernatant was

removed by centrifugation at 12,000g for 30 seconds. The oligo (dT) cellulose pellet was washed 4 times using high salt buffer (10mM TrisHCl, pH 7.5, 1mM EDTA, 0.5M NaCl) followed by washing with low salt buffer (10mM TrisHCl, pH 7.5, 1mM EDTA, 0.1M NaCl) for another 4 times. High salt conditions allow the annealing the poly(A) tail to the oligo d(T). The low salf buffer removes the poly(A)- RNAs (eg. tRNA and rRNA). These washings remove contaminating DNA, RNA proteins. The oligo dT beads were transferred to the microspin column (polypropylene minicolumns) and suspended in pre heated 100 -200 µl elution buffer (10mM Tris HCl, pH 7.5, 1mM EDTA) or pre heated RNase free water (Sigma) which releases the poly(A)$^{+}$ RNA and the tube was centrifuged at 12,000g for 30 seconds. The eluate contained poly (A) RNA which was stored at -80°C until downstream processing.

2.2.2.1. b Poly (A) enrichment from total RNA:

Poly(A) was enriched from total RNA preparations using Dynabeads mRNA direct kit which relies on the pairing between the poly(A)tail of mRNa and the oligo dT sequence linked to the surface of the beads. Dynabeads oligo (dT)$_{25}$ are uniform, 2.8 µm diameter, superparamagnetic, polystyrene beads with 25 nucleotide long chains of oligodeoxythymidine covalently attached to the bead surface via a 5' linker group. RNA extraction was performed accoding to manufacturer's instructions. The resultant mRNA was quantified using nanodrop and Qubit fluorometer. Quality assessment was performed with agilent bioanalyzer.

2.2.2.2 Extraction of total ribonucleic acids (Total RNA):

Prior to extraction, all glassware was autoclaved to ensure sterility and inactivation of contaminating proteins such as nucleases. When possible, sterile plastic was used instead of glassware. To minimize the loss of nucleic acis, DNA Lobind molecular biology grade 1.5 ml tubes (Eppendorf) were extensively used. Total RNA was extracted from snap-frozen cells. 1ml TRIzol reagent (Invitrogen) was used to lyse cells grown per 10cm^2 dish. The lysates were incubated with TRIzol reagent at room temperature for 5 minutes to allow complete dissociation of nucleoprotein complexes. TRIzol is a monophasic solution of phenol and guanidine isothiocyanate which maintains the integrity of RNA while disrupting cells and dissolving cell components. 0.2 ml chloroform was used per 1ml TRIzol reagent and tubes were shaken by hand for 15 seconds. This mixture was left for incubation for 2 – 3 minutes at room temperature which was followed by

centrifugation at 12,000g for 15 minutes at 4°C. Addition of chloroform followed by centrifugation separates the solution into aqueous phase and organic phase. RNA remains exclusively in the aqueous phase. The colorless upper aqueous phase containing the RNA is then transferred to a new 1.5 ml tube. An equal volume of 70% ethanol is added to the aqueous phase to obtain a final ethanol concentration of 35% and this mixture is mixed well by vortexing. This was followed by binding, washing and elution with the column of PureLink RNA mini kit (Life Technologies). Colum binding, washing and elution was performed according to manufacture's instructions. RNA was extracted either in DNase, RNase free water (Sigma) or Tris-EDTA buffer (10mM TrisHCl, pH = 7.5, 1mM EDTA) and stored at -80°C until analysis.

2.2.2.3 Small RNA enrichment from total RNA:

miRNA enrichment from total RNA sample was performed using PureLink miRNA isolation kit. 20 -30 µg of total RNA was used for small RNA enrichment. Total RNA was suspended in 90µl Nuclease free water (Sigma) to which 300µl binding buffer (L3) and 210µl 100% ethanol were added. This mixture was mixed by vortexing and then loaded onto a spin cartridge. The spin cartridge was kept in a collection tube and spun at 12,000g for 1 minute. The flow through was collected in a fresh 1.5ml tube. 700µl 100% ethanol was added to the flow through, mixed by vortexing and loaded onto another spin cartridge in a collection tube and spun at 12,000g for 1 minute. The flow through was discarded and the small RNA bound to the spin cartridge was washed twice using 500 µl wash buffer (W5) and by spinning at 12,000g for 1 minute.

Small RNA was recovered by placing the cartridge in a clean recovery tube to which 50 µl sterile, RNase free water (Sigma) was added and incubated for 1 minute. The spin cartridge was then spun at 16,000g for 1 minute and the eluted small RNA was stored at -80°C. The quality assessment of the small enriched samples was performed using Agilent small RNA chip.

2.2.3 RNA quantification and quality assessment

2.2.3a. RNA quantification:

The concentration of RNA was determined using both Qubit fluorometer (Invitrogen) and NanoDrop system (Thermo Scientific). Qubit fluorometer quantification method is based on fluorescence dyes that bind specifically to DNA, RNA or protein where as NanoDrop

quantification is based on UV absorbance measurements. Since all nucleic acids RNA, ssDNA, dsDNA absorb at 260 nm, they will contribute to the total absorbance of the sample. The degree of contamination in RNA sample preparation using NanoDrop is estimated by A_{260}/A_{280} and A_{260}/A_{230} ratio. A_{260}/A_{280} ratio of ~2.0 is generally accepted as "pure" for RNA. If the ratio is appreciably lower, it may indicate the presence of protein, phenol or other contaminants that absorb strongly at or near 280 nm. A_{260}/A_{230} ratio is a secondary measure of nucleic acid purity. The 260/230 values for "pure" nucleic acid are often higher than the respective 260/280 values. Expected 260/230 values are commonly in the range of 2.0-2.2. If the ratio is appreciably lower than expected, it may indicate the presence of contaminants which absorb at 230 nm.

However, NanoDrop suffers a drawback. UV absorbance readings indiscriminately measure anything that absorbs at 260 nm, including DNA, RNA, protein, degraded nucleic acids, and free nucleotides. Whereas, the Qubit® Quantitation Platform, in contrast, utilizes specifically designed fluorometric technology using Molecular Probes® dyes to measure the concentration of the specific molecules of interest. These fluorescent dyes emit signals only when bound to specific target molecules, even in the presence of free nucleotides or degraded nucleic acids. Samples were prepared for RNA assay following manufacturer's instructions. Fluorometric assay yielded quantification of nucleic acids could be compared with the data obtained using the spectrophototmeter.

2.2.3b. RNA Quality assessment:

RNA integrity which is a critical first step in obtaining meaningful gene expression data was determined using Agilent bioanalyzer and RNA Nano and Pico lab chip kits (Agilent Technologies). The Agilent Bioanalyzer is a microfluidics-based platform for sizing, quantification and quality control of DNA, RNA, proteins and cells. Profiles generated on the Agilent bioanalyzer yield information on concentration, allow a visual inspection of RNA integrity, and generate ribosomal ratios. Using electrophoretic separation on microfabricated chips, RNA samples are separated and subsequently detected via laser induced fluorescence detection. The bioanalyzer software generates an electropherogram and gel-like image and displays results such as sample concentration and the ribosomal ratio. The Agilent 2100 bioanalyzer provides a better assessment of RNA intactness by showing a detailed picture of the size distribution of RNA fragments. The RIN (RNA Integrity Number) software algorithm allows for the classification of eukaryotic total RNA, based on a numbering system from 1 to 10, with 1 being the most degraded profile

and 10 being the most intact. The labels in-between are used to indicate progressing degradation states of the RNA sample.

RNA samples were run either on RNA 6000 Nano or RNA 6000 Pico chips and Small RNA chip. Agilent small RNA chip detects RNA < 150 nucleotides long. All RNA samples were diluted according to the concentration range of respective chips. Following table gives the analytical specifications of the RNA chips.

2.2.4 Cell lysate preparation:

Cell lysate is a solution of cellular proteins resulting when cells are lysed (broken apart) under conditions which preserve the protein's structure and function. Cell lysae preparation protocol is as follows:

The culture medium was removed and cells were washed twice with cold 1X PBS, Thorough washing is essential as insufficient washing will contaminate lysate with media components (especially serum/BSA) which may erroneously elevate protein concentration. All PBS was removed completely to prevent unwanted dilution of the final product. Cells were detached into 10 ml cold PBS with rubber policeman and the cell suspension was transferred into 15ml tube. Cells were pelleted by centrifugation at 1350 rpm for 10 minutes. Supernatant was discarded. The pellet was snap frozen in dry ice. Frozen cell pellets were used for cell lysate preparation. The frozen pellet of cells was resuspended in 300 μl lysis buffer with protease and RNase inhibitors. Resuspended pellet was homogenized by pipetting up and down, to break up the pellet thoroughly and was allowed to stand on ice for 30 minutes with intermittent vortexing every 10 minutes. The resulting mixture was centrifuged at 13,000g for 15 minutes at 4°C. This separates the debris (pellet) from total protein (supernatant). The supernatant was collected in a new 1.5 ml tube.

Protein concentration was measure using A-280 absorbance of Nanodrop (Thermoscientific).

2.2.5 Protein Analysis

2.2.5a SDS-PAGE

Protein samples were analyzed by SDS-PAGE. Resolving gels were made at 12% of a 37.5:1 mix of acrylamide/bis-acrylamide (Protogel) in 375 mM Tris HCl, pH 8.8, 0.1% v/v SDS and and 4.9 ml water. Polymerization was induced by addition of 0.1% w/v APS

and 15 µl TEMED. The separating gel was poured into a gel former. The stacking gel comprised of 4% of 37.5:1 mix of acrylamide/bis-acrylamide (Protogel) in 125mM Tris HCl pH 6.8, 0.2% v/v SDS and water. 0.1% w/v APS and 0.1% v/v TEMED were added for crosslinking. Stacking gel was poured on the top of separating gel. A gel- comb of corresponding thickness and desired number of wells was inserted in the stack gel and the assembly was left until the gel was set. Protein samples were diluted in 5X gel loading buffer and denatured at 95 °C for 5 minutes. Electrophoresis was carried out at 140V, 25mA in running buffer.

2.2.5b Western Blotting

Following electrophoresis, acrylamide gels were soaked in transfer buffer. Immobilon-P Transfer membrane (0.45 µm), two pieces of filter paper and two pieces of sponge cut to the size of the gel were also soaked in transfer buffer. Immobilon-P Transfer membrane was soaked in methanol and water briefly before being left in transfer buffer. The PDVF membrane was placed on top of the pre-soaked gel and housed between two pieces of filter paper and sponge in the western blotting cassette of Mini Trans-Blot Electrophoretic Transfer Cell (*Bio-Rad*). The assembly was secured in a transfer tank (Mini-Protean III, BioRad) filled with transfer buffer, ensuring the gel-side of the assembly was orientated towards the cathode (and thus ensuring that proteins were transferred from the gel to the PDVF membrane as the current passed through the apparatus). Proteins were transferred to PDVF membrane at a current of 100mA for O/N. A magnetic stir-bar was placed into the tank and the whole apparatus placed on a magnetic stirrer at moderate speed to prevent the precipitation of glycine from the transfer buffer. Precipitated glycine is known to adhere to the nitrocellulose membrane, increasing background signal.

Electroblotted PDVF membrane was blocked for 3 hours at RT in Bloking buffer (Millipore) with gentle shaking. Membrane was incubated with Anti AGO2 antibody (primary antibody) diluted 1:2500 in Bloking buffer (Millipore) for 3 hours with gentle mixing. The membrane was first washed with TBS twice for 5 minutes each wash and then with TBS-T for a total of 30 minutes, changing solution every 5 minutes. Following washing, membrane was incubated with horseradish peroxidise conjugated anti Rat antibody diluted 1:1000 in bloking buffer (Millipore) for 1 hour with gentle mixing. Following 30 minutes of washing as mentioned before, immobilized protein-antibody complexes were visualized by chemiluminescence using ECL plus kit (GE Healthcare) following manufacturer's instructions. The membrane was exposed to photographic film

(Hyperfilm ECL, GE Healthcare) in a dark room, developed in developer (Kodak), rinsed in water and fixed in fixer (Kodak). Molecular weights of the protein bands were estimated by comparison with protein standards.

2.2.6 Antibodies:

Primary Antibodies: Anti Argonaute2 Rat monoclonal antibody (Sigma).

Secondary Antibodies: Anti - rat HRP (DAKO P0450, produced in rat): Anti rat immunoglobulins conjugated with the horseradish peroxidase used in chemiluminescence detection (ECL GE Healthcare Biosciences).

2.2.7 RNA Immunoprecipitation (RIP)

RNA immunoprecipitation (RIP) is a very powerful procedure for the study of RNA binding proteins (RBPs) and their RNA targets in ribonucleoprotein (RNP) complexes. Antibodies raised against specific RBPs are used to coprecipitate RNPs, i.e., the RBP along with its RNA partner. The RNA can then be identified by next generation sequencing, or if testing for a specific RNA, by RT-PCR.

RIP of AGO2 from CHQ5B myoblasts and myotubes were performed using the following protocol:

Cell Lysis: Complete lysis buffer (containing protease inhibitor cocktail, RNase inhibitor and DTT) was added to cell pellet and incubated on ice for 30 minutes. RIP lysis reaction was centrifuged for 15 minutes at 13,000g at 4°C.

Antibody binding to protein G sepharose: 10 μg of AGO2 antibody was prebound to protein G sepharose beads by overnight incubation at 4°C with constant and gentle rotation in a 1.5 ml tube. In order to remove excess antibody not bound to the protein G beads, the beads were washed using washing buffer thrice.

Antigen binding to antibody: 200 ul of cell lysate was added to the precoated sepharose beads and incubated for 1 hour at 4°C with constant rotation. This was followed by washing with washing buffer.

RNA extraction: Total RNA from immunoprecipitated complexes was extracted using Trizol and Purelink RNA Mini kit (Invitrogen). RNA was eluted in 50 μl RNase free water and stored at -80°C. RNA was quantified by Nanodrop (Thermo sceintific) and

Qubit fluorometer (Invitrogen). The quality of RNA was assessed using Agilent bioanalyzer.

2.2.8 Transcriptome and miRnome sequencing using Applie Biosystem's SOLiD and Ion proton technology:

First step of sequencing is library preparation. Whole transcriptome and small RNA library preparation protocol are same for both the sequencing techniques. Following is the description of library preparation procedure.

2.2.9 Commercial RNA:

RNA	COMPANY
Total RNA isolated from adult human skeletal muscle tissues	Amsbio/ Zyagen
Total RNA isolated from adult human skeletal muscle tissues	Stratagene

2.3. Results:

To have the thorough picture of muscle transcriptomics we constructed mRNA and miRNA libraries for both cultured primary skeletal muscle cells and skeletal muscle tissue. We then compared the transcriptomic and miRnome profiles of myoblasts, myotubes and skeletal muscle tissue. The findings are detailed in the following sections.

2.3.1 Human myoblast, myotube and skeletal muscle tissue differential gene expression profile

Upon differentiation of myoblasts into myotubes a total of 1010 genes were significantly (FDR=0.05) over expressed ($\geq$1.5 $\log_2$ fold change) and 1593 genes were down regulated ($\leq$ -1.0 $\log_2$ fold change). Skeletal muscle tissue showed considerable over expression of 2481 genes ($\geq$ 1.6 $\log_2$ fold change) and under expression of 2243 gens ($\leq$ -1.5 $\log_2$ fold change) with respect to myotubes (figure 2-1).
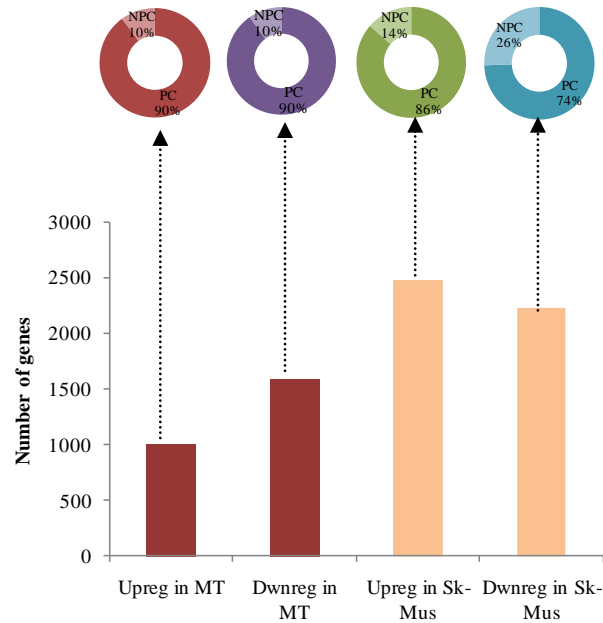
Figure 2-1: mRNA differential expression profile of myotube and skeletal muscle tissue. Number of genes found differentially up and down regulated is shown by the bar plots. The doughnut plot at the top of each bar plot shows the percentage of protein coding and non-protein coding genes. (NPC- Non protein coding, PC- Protein coding, Upreg in MT- Upregulated in myotubes, Dwnreg in MT- Down regulated in myotubes, Upreg in Sk Mus- Upregulated in skeletal muscle tissue, Dwnreg in Sk Mus- Down regulated in skeletal muscle tissue).

We observed a correlation (R^2 value) of 0.775 between the transcriptome of myoblasts and myotubes, whereas the transcriptome of skeletal muscle tissue showed a correlation (R^2 value) of 0.562 with myotubes. Figure 2-2 (A&B) shows the plots of comparison between myoblasts, myotubes and skeletal muscle tissue. As is apparent from the plots, the degree of variance between the transcriptomes of skeletal muscle tissue and myotubes is higher than the variance between myoblasts and myotube transcriptomes. The smear plots in figure 2-3 (A&B) show that the differential expression of genes is much higher in skeletal muscle tissue.

15% and 22.5% of protein coding genes showed more than 5X fold increase in myotubes and skeletal muscle tissue respectively. The genes have been listed in supplementary table S1-5 and S1-6. In myotubes MYBPC1, MYH8, DPYSL5, CASQ2, ATP1B4, CACNG1 and MYH3 showed ≥ 8X fold increase. Whereas HBB, MB, PDK4, THBS4, SMTNL2, LBP, CD74, FABP4, C8orf22, CD93, LPL, RNASE1, RORC, CA3, PPP1R1A, DARC and IDI2 were the most prominent over expressed genes showing ≥ 9X fold increase in

skeletal muscle tissues when compared to myotubes. However, 9% and 19% genes showed > -4X fold and -6X fold decrease in myotubes and skeletal muscle tissue respectively.

A   B
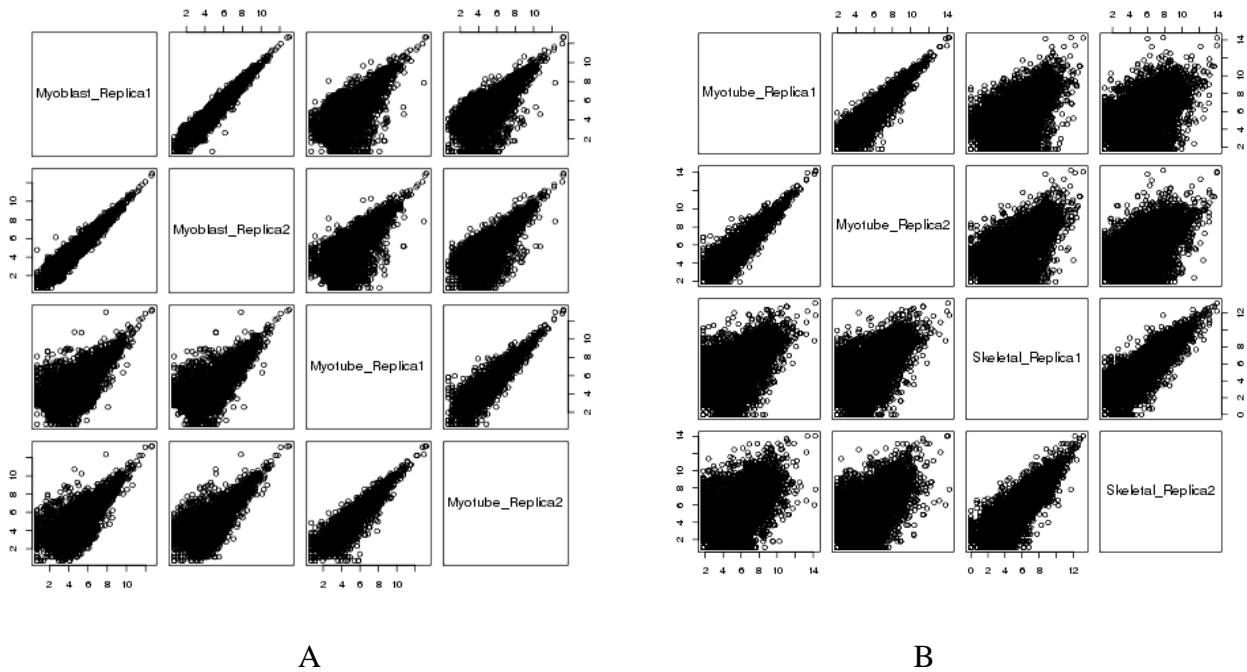
Figure 2-2: (A) Pairs-plot between the transcriptome of myoblast and myotube. These plots compare the transcriptome of myoblasts and myotubes. We performed the RNAseq in replica for both myoblasts and myotubes, the comparison of the transcriptome of the replicas are also shown here. (B) Pairs-plot between the transcriptome of myotube and skeletal muscle tissue. These plots compare the transcriptome of myotubes and skeletal muscle tissue. We performed the RNAseq in replica for both myotubes and skeletal muscle tissue, the comparison of the transcriptome of the replicas are also shown here.

**A**   **B**
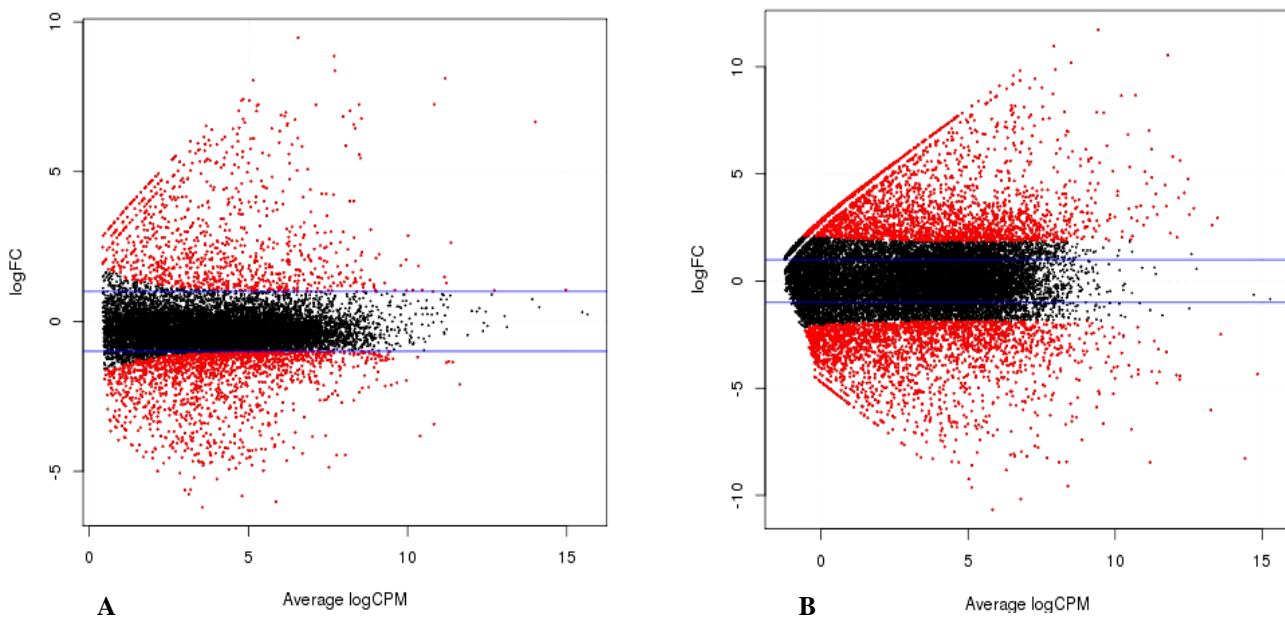
Figure 2-3: (A) Smear plot showing differential expression of genes between myoblasts and myotubes. (B) Smear plot showing differential expression of genes between myotubes and skeletal muscle tissue. Log fold-change of gene expression is shown by the vertical axis or Y-axis.Genes below the zero level (on negative scale) show genes differentially down regulated and vice-versa. Red dots represent genes showing differential expression. Black dots represent genes which are not differentially expressed between the two samples. (Log FC- log_2 Fold change, Average log CPM –log of count per million)

2.3.1.1 Protein family genes upregulated in myotubes

Members of many protein families were over expressed together, upon myotube formation. All those family specific genes have been listed in supplementary table S2-7. We observed some ABC transporter genes of *ATP binding cassette family (ABC superfamily)* were differentially over expressed. ABC transporter genes of ABCA1, 3, 6, 8 and 9 from ABCA subfamily and ABCG1 showed >1.5x fold increase. mRNA for *apolipoproteins* D, E and apolipoprotein L1, 3, 4, 6 were also upregulated. ABS 2, 4, 14 and 16 genes of *ankyrin repeat and SOCS box-containing (ASB) family* were over expressed. Some sub units of *voltage dependent calcium channels (VDCC)* were found upregulated. Alpha 1S subunit of L-type calcium channel (CACNA1S), alpha 1H subunit of T-type calcium channel (CACNA1H), beta subunit isoform 1, 4 (CACNB1, CACNB4) and gamma subunit genes CACNG 1, 4, 6 were >2x upregulated. Alpha and beta subunit genes of *voltage gated sodium channels* showed >2x increased expression. Genes for *coiled coil domain containing* proteins, CCDC141, 3, 39 and 69 also increased upon differentiation. *Cyclin dependent kinase inhibitors* CDKN1A, 1C and 3B expression was also found increased. Genes of collagen type I, III, IV, VIII, X, XI, XV, XVIII, XIX and XXI were found upregulated. Genes for members of *insulin like growth factor binding protein family*, IGFBP 2, 5, 7 were upregulated. A wide range of interleukins were also turned up. Subunits of *Voltage gated potassium channel* genes and some other genes of potassium channel subfamily were over expressed. Almost all the members of *myosin class II,* members of other myosin classes and many myosin light chain genes were strongly upregulated. Some protocadherin genes of *protocadherin subfamily of cadherins*, which have their role in homophilic adhesion and signalling, were also turned up. Many members of *Solute carrier family (SLC group),* which are membrane transport proteins had increase in their expression. Many zinc finger protein genes showed slight

upregulation. Genes for *FOX (Forkhead box) proteins* which are a family of transcription factors were upregulated.

2.3.1.2 Protein family genes upregulated in skeletal muscle tissue

Similar to myotubes many members of *ATB binding cassette (ABC superfamily)* (ABCG1, ABCC 2, 6, 9, ABCB4, ABCA 5, 10) were > 2.5X fold increased in skeletal muscle tissue. *Acyl CoA synthetase long, medium and short chain fatty acid family members* (ACSL1, 5, ACSM5, ACSS1 and 2) were found to be >2 fold upregulated. *Adenylate cyclase isoforms* (ADCY 2, 4, 5 and 9) and *Aldehyde dehydrogenase 1, 2, 5, 8 family* members (ALDH1A2, A3, L1 ALDH2, ALDH5A1 and ALDH8A1) were also >2 fold over expressed. Members of *ankyrin repeat and SOCS box-containing (ASB) family* members (ASB2, 8, 10, 11, 12, 15, 16) were ≥ 3X fold increased. Isoforms of *α-CA (carbonic anhydrase) family* (CA1, 2, CA3 muscle specific, CA4, 8, 11) showed ≥ 3X fold increase, with CA3 showing 9X fold over expression. *Chemokine (C-C motif) ligand* 8, 18, 21 were > 2.5X fold increased. 26 *CD (cluster of differentiation)* molecules showed > 2X fold up regulation. Cluster of differentiation is defined as a subset of cellular surface receptors (epitopes) that identify cell type and stage of differentiation, which are recognised by antibodies. *C-type lectin domain superfamily members* (CLEC10A, 12A, 14A, 1A, 5A, 7A, 2B, 4E and 4G) which play diverse functions such as cell adhesion, cell – cell signalling, glycoprotein turnover, roles in inflammation and immune response also were upregulated. *Cytochrome C oxidase (COX)* is the terminal enzyme of the mitochondrial respiratory chain. COX subunits (COX4I2, COX5A, COX6A2, COX6C, COX7A1, COX7B and COX7C) were ≥ 2X fold upregulated. *CXC chemokine receptors* are integral membrane proteins that specifically bind and respond to cytokines of CXC chemokine family. CXC chemokine is one of the subfamily of chemokines. Chemokines are a family of cytokines or signalling proteins secreted by cells. *CXC ligands* (CXCL1, 2, 3, 14) and *CXC receptors* (CXCR1, 2, 4) were ≥ 2X fold upregulated. *DEAD box proteins* are involved in assortment of metabolic processes that typically involve RNAs. *DEAD box polypeptides* (DDX11, 12, 51 and 59) also showed ≥ 2X fold increase. *Dual specificity phosphatase* is a form of phosphatase that can act upon tyrosine or serine/threonine residues. Many DUSP enzymes (DUSP1, 13, 16, 2326, 27, 3 and 8) were ≥ 2X fold upregulated. Many *Eukaryotic translation initiation factors (eIF)* which are involved in the initiation phase of eukaryotic translation were upregulated in skeletal muscle tissue. *Hemoglobin subunit members* (HBA1, 2, HBB, HBD, HBG1, 2) were

found to be ≥ 2.5X fold increased. Some genes of *Troponin complex protein* which is integral to muscle contraction were also ≥ 2X fold increased. Genes of some *interleukins (ILs),* and *interleukin receptors* were upregulated. The *major histocompatibility complex (MHC)* is a set of cell surface molecules that mediate interactions of leukocytes, also called white blood cells (WBCs), which are immune cells, with other leukocytes or body cells. MHC determines compatibility of donors for organ transplant as well as one's susceptibility to an autoimmune disease via crossreacting immunization. Many MHC genes were also found upregulated. Members genes of *Leukocyte Ig-like receptors (LIRs)* which are a family of immunoreceptors expressed predominantly on monocytes and B cells and at lower levels on dendritic cells and natural killer (NK) cells also showed increase in expression. *NADH dehydrogenase (ubiquinone)* is an enzyme of the respiratory chains. It catalyzes the transfer of electrons from NADH to coenzyme Q10 (CoQ10) and, in eukaryotes, it is located in the innermitochondrial membrane. It is one of the entry enzymes of oxidative phosphorylation in the mitochondria. Many NADH enzyme subunit genes were increased in expression. Many *Phosphoprotein phosphatase 1 (PP1)* genes which belongs to a certain class of phosphatases known as protein serine/ threonine phosphatases also turned up their expression. PP1 has been found to be important in the control of glycogen metabolism, muscle contraction, cell progression, neuronal activities, splicing of RNA, mitosis, cell division, apoptosis, protein synthesis, and regulation of membrane receptors and channels. A huge number of genes belonging to *solute carrier family (SLC)* were over expressed. SLC group members transport diverse type of solutes which include charged and uncharged organic molecules as well as inorganic ions. Many *zinc finger proteins* were also upturned significantly. (Details to be found in supplementary table S2-8).

2.3.1.3 Functional annotation and pathway analysis of differentially expressed genes

Using the online tool DAVID and KEGG pathways we found the GO terms and pathways over represented in myotubes relative to the myoblast and in skeletal muscle compared to myotubes.

Genes over expressed in myotubes were those involved in biological processes related to skeletal muscle development (muscle system process, muscle organ development, muscle contraction and other striated muscle specific activities). Whereas genes significantly

37

down regulated were mainly associated with cell cycle regulation (nuclear division, mitosis). For the upregulated genes in skeletal muscle tissue we identified, immune response, defence response, inflammatory response were significantly over represented GO biological process terms. Subsets of functional groups are shown in supplementary tables (S2-1 to S2-4).

Differentially over expressed genes were, analysed using KEGG pathway. Pathways are shown in table 2-1 and 2-2. Highest number of genes belonged to metabolic pathways, which is described as a set of complex metabolic networks (1).

KEGG pathways	Number of genes	Gene names
Metabolic pathways	48	ACSL1, ADC, ADH1B, AGL, AKR1C3, ALOX15B, AMPD1, AOC3, AOX1, B3GALT2, CKB, CKM, CKMT2, COX6A2, CYP19A1, CYP2J2, DHCR7, ENO3, GALNT16, GATM, GGT5, HMGCS1, HSD11B1, IDI1, INPP4B, INPP5K, ISYNA1, MAN1C1, MSMO1, ND3, ND5, NDUFA4L2, NNMT, PGAM2, PHOSPHO1, PIK3C2B, PIPOX, PLA2G16, PLA2G4C, PLCE1, PTGDS, PTGS1, PYGM, SQLE, ST6GAL1, TM7SF2, TYMP, WBSCR17
Hypertrophic cardiomyopathy (HCM)	30	ACTC1, CACNA1S, CACNB1, CACNB4, CACNG1, CACNG4, CACNG6, DAG1, DMD, IGF1, ITGA1, ITGA4, ITGA6, ITGA7, ITGB8, LAMA2, MYH6, MYH7, MYL2, MYL3, PRKAA2, PRKAB2, PRKAG3, RYR2, SGCD, SGCG, TGFB3, TNNC1, TNNT2, TTN
Pathways in cancer	30	ARNT2, BCR, CDKN1A, CDKN2B, COL4A1, COL4A4, DAPK1, DAPK2, EGLN3, EPAS1, FGF13, FGF9, FOXO1, FZD4, FZD9, IGF1, ITGA6, JUP, LAMA2, LAMA5, PAX8, PDGFA, PDGFB, RUNX1T1, RXRG, STAT5A, TGFB3, WNT10A, WNT11, ZBTB16
Dilated cardiomyopathy	29	ACTC1, ADCY1, ADCY2, CACNA1S, CACNB1, CACNB4, CACNG1, CACNG4, CACNG6, DAG1, DMD, IGF1, ITGA1, ITGA4, ITGA6, ITGA7, ITGB8, LAMA2, MYH6, MYH7, MYL2, MYL3, RYR2, SGCD, SGCG, TGFB3, TNNC1, TNNT2, TTN
Focal Adhesion	26	ACTN2, CAV3,CCND2, COL11A1, COL1A1, COL3A1, COL4A1, COL4A4, COMP, IGF1, ITGA1, ITGA4, ITGA6, ITGA7, ITGB8, LAMA2, LAMA5, MYL2, MYLK4, MYLPF, PDGFA, PDGFB, PDGFD, PPP1R12B, RAPGEF1, SHC2
PI3K-Akt signaling pathway	25	BCL2L11; CCND2; CDKN1A; COL11A1; COL1A1; COL3A1; COL4A1; COL4A4; COMP; DDIT4; FGF13; FGF9; IGF1; ITGA1; ITGA4; ITGA6; ITGA7, ITGB8LAMA2, LAMA5, PDGFA, PDGFB, PDGFD, PPP2R3B, PRKAA2
Regulation of actin cytoskeleton	21	ACTN2; ARHGEF6; CHRM4; CYFIP2; FGF13; FGF9; ITGA1,ITGA4; ITGA6; ITGA7; ITGB8; LIMK2; MRAS; MYH14; MYL2; MYLK4,MYLPF; PDGFA; PDGFD; PPP1R12B
Arrhythmogenic right ventricular cardiomyopathy	20	ACTN2; CACNA1S; CACNB1; CACNB4; CACNG1,CACNG4; CACNG6; DAG1,DMD; ITGA1; ITGA4; ITGA6,ITGA7; ITGB8; JUP; LAMA2; PKP2;

KEGG pathways		Gene Name
(ARVC)		RYR2; SGCD; SGCG
HTLV-I infection	20	ADCY1; ADCY2; CCND2; CDKN1A; CDKN2B; EGR2; FZD4; FZD9; ICAM1; IL15RA; IL1R1; MRAS; PDGFA; PDGFB; STAT5A; TGFB3; TP53INP1;VCAM1; WNT10A; WNT11
MAPK signalling pathway	19	CACNA1H; CACNA1S; CACNB1; CACNB4; CACNG1; CACNG4; CACNG6; DUSP4; FGF13; FGF9; GADD45G; IL1B; IL1R1; MEF2C; MRAS; PDGFA; PDGFB; PLA2G4C; TGFB3

***Table 2-1: KEGG pathways enriched for genes differentially up-regulated in Myotubes.** The table shows the name of the enriched pathways, the number of genes involved in corresponding pathways and the gene names.*

KEGG pathways	**Number of Genes**	**Gene Name**
Metabolic pathways	153	ACACB; ACADM; ACADS; ACAT1; ACO2; ACSL1; ACSL5; ACSM5;ACSS1; ACSS2; ADH1A; ADH1B; ADH4; ADH5, ADSSL1;AGPAT9;AKR1B10;AKR1C3;ALAS2;ALDH1A2;ALDH1A3; ALDH2;ALDH5A1; ALDOA;ALOX5;AMPD1; AMPD3;AMY1A; AMY2A; AMY2B; AOC3;ARG1; ARG2;ATP5E; B3GALT4; B3GNT5; B4GALNT1;BDH1;CD38;CDO1; CES1; CHDH; CKM; CKMT2; CMBL; CMPK2; COX17; COX4I2; COX5A; COX5B;COX6A2;COX6C;COX7B;COX7C;CYP1A1; CYP27A1;DCXR;DGAT1;DGAT2;DGKG;DHRS3;DHRS4;EPHX2; FBP2;FUT1;GALNT6;GAMT;GAPDH;GCNT2;GLUL;GOT2;GPT; HADH;HADHA;HADHB;HK3;HMGCS2;HSD17B8;HSD3B1;HYAL1; IDH2;IDH3A;IDI2;IDO1;IMPA2,KL;KYNU;MAN2A2;MAOA; MAOB;MGAT3; MGAT4A; MLYCD; MOCS1; NDUFA12,NDUFA2; NDUFA3; NDUFA4; NDUFA4L2; NDUFA5; NDUFA6;NDUFB1; NDUFB10;NDUFB3;NDUFB4;NDUFB8;NDUFB9;NDUFS3;NMNAT3; NNMT; NOS1; NT5C1A; NT5M; PCK1; PFKM; PGAM2; PGM1; PIP5K1B;PLA2G2A; PLA2G4C; PLA2G4F; PLA2G5; PLB1; PLCB2; PLCD4; PLCG2; PNMT;POLR2I; POLR2J; POLR2J2; POLR3GL; PRODH;PYCR1; PYGM; SDHB; SDHD; SHMT1; ST3GAL3; ST6GALNAC1; ST6GALNAC3;TBXAS1; TPO; UGP2; UPP1; UQCRB; UQCRC1; UQCRFS1
Epstein-Barr virus infection	50	BCL2; CD38; CDKN1B; FGR; HDAC5; HLA-B; HLA-C; HLA-DPA1; HLA-DPB1;HLA-DQA2; HLA-DQB1;HLA-DRA; HLA-DRB4; HLA-E; HLA-F; HLA-G; HSPA1A;HSPA2; HSPA6;IL10RA; MAP2K6; MAP3K14; MAPK12; NEDD4; NFKBIA; PIK3CB; PIK3R5; PLCG2; POLR2I; POLR2J; POLR2J2;POLR3E; POLR3GL;PSMC1; PSMC4; PSMC5; PSMC6; PSMD12;PSMD13; PSMD14; PSMD3; PSMD4; PSMD7; PSMD8; SHFM1; SPI1; SYK; TNFAIP3; YWHAE
Cytokine-cytokine receptor interaction	47	AMHR2; CCL18; CCL21; CCL8; CCR1; CD40; CNTFR; CSF1R; CSF2RA; CSF2RB; CSF3; CSF3R; CX3CL1; CXCL1; CXCL14;

		CXCL2; CXCL3; CXCR4; EGF; EPOR; FIGF; FLT1; FLT4; GHR; IL10RA; IL12RB2; IL15RA; IL17RB; IL18; IL18RAP; IL1R2; IL2RG; IL3RA; IL6R; IL8; INHBB; KDR; OSM; PDGFB; RELT; TNF; TNFRSF14; TNFRSF1B; TNFRSF4; TNFSF10; TPO
Alzheimer's disease	43	ATP2A1; ATP2A2; ATP5D; ATP5E; ATP5H; ATP5J; COX4I2; COX5A; COX5B; COX6A2; COX6C; COX7A1;COX7B;COX7C; CYCS; GAPDH; LPL;MAPT; NDUFA12; NDUFA2; NDUFA3; NDUFA4;NDUFA4L2; NDUFA5; NDUFA6; NDUFB1; NDUFB10; NDUFB3; NDUFB4; NDUFB8; NDUFB9;NDUFS3;NOS1; PLCB2; PPP3CC; RYR3; SDHD; SNCA; TNF; UQCRB; UQCRC1; UQCRFS1
MAPK signaling pathway	43	ARRB2; ATF4; CACNA2D3; CACNB2; CD14; DUSP1; DUSP16; DUSP3; DUSP8; ECSIT; EGF; FGF10; FOS; GADD45A; GADD45B; GADD45G; HSPA1A; HSPA2; HSPA6; IL1R2; MAP2K6; MAP3K14; MAPK12; MAPKAPK3; MAPT; MKNK2; NFATC1; NR4A1; NTRK1; NTRK2; PDGFB; PLA2G4C; PLA2G4F; PPM1B; PPP3CC; PRKCB; PTPN5; RASGRF2; RASGRP2; RASGRP3; RPS6KA1; TNF; ZAK
Huntington's disease	43	ATP5D; ATP5E; ATP5H; ATP5J; COX4I2; COX5A; COX5B; COX6A2; COX6C; COX7A1; COX7B; COX7C; CYCS; NDUFA12; NDUFA2; NDUFA3, NDUFA4; NDUFA4L2; NDUFA5; NDUFA6; NDUFB1; NDUFB10; NDUFB3; NDUFB4; NDUFB8; NDUFB9; NDUFS3; PLCB2; POLR2I; POLR2J; POLR2J2; PPARGC1A; PPIF; SDHB; SDHD; SLC25A4; SOD1;SOD2; TAF4B; UQCRB; UQCRC1; UQCRFS1; VDAC1
PI3K-Akt signaling pathway	42	ANGPT2; ATF4; BCL2; CDKN1B; CHAD; COL4A3; CSF1R; CSF3; CSF3R; EFNA1; EGF; EIF4B; EIF4EBP1; EPOR; FGF10; FIGF; FLT1; FLT4; GHR; GNG5; IGF1;IL2RG; IL3RA; IL6R; INSR; ITGA9; ITGB4; KDR; NR4A1; OSM; PCK1; PDGFB; PIK3CB; PIK3R5; PPP2R5A; PRKAA2; RHEB; SYK; THBS4; TLR2;YWHAE

Table 2-2: KEGG pathways enriched for genes differentially up-regulated in Skeletal muscle tissue. The table shows the name of the enriched pathways, the number of genes involved in corresponding pathways and the gene names.

2.3.2 miRNA expression profiling of myoblasts, myotubes and skeletal muscle tissue

miRNA are well conserved and evolutionarily ancient component of genetic regulation in eukaryotic organisms [5,6]. So, it was of significant interest for us to carry out an in-depth study of miRNA expression profiles at each developmental stage of skeletal muscle (i.e. from mono nucleated myoblasts to multinucleated myotubes to skeletal muscle tissue). Differential expression of miRNA was found using edgeR software with an FDR (false discovery rate) of 0.05. We performed the miRNA-seq in replica for each sample to have more robust and reliable set of results.

2.3.2.1 miRNA profile of myotubes

Using NGS we found that more than 400 miRNA were expressed at a substantial level in myoblasts and myotubes. With edgeR software we found the differential expression of miRNA between myoblasts and myotubes. The correlation coefficient (R^2 value) of miRNA expression between myoblasts and myotubes was found to be 0.895, which indicates high correlation between the two samples. We found 24 miRNA were over expressed and 16 miRNA were under expressed in myotubes. Table 2-3 shows the list of differentially expressed miRNA in myotubes.

hsa-miR-1, 133 and 206 (2,3,4) have their role very well documented in muscle growth and differentiation. We found that along with miR-1, miR-133a and miR-206, miR-139-5p and miR-4488 were > 5 $\log_2$ fold upregulated upon differentiation into myotubes.

Over Expressed			*Under Expressed*		
miRNA	**Log$_2$ FC**	**P value**	**miRNA**	**Log$_2$ FC**	**P value**
hsa-miR-1	7.91	2.94E-18	hsa-miR-222-5p	-3.59	1.74E-06
hsa-miR-133a	7.06	1.68E-15	hsa-miR-155-3p	-3.56	1.33E-05
hsa-miR-139-5p	6.12	2.830E-11	hsa-miR-155-5p	-3.51	2.18E-06
hsa-miR-4488	5.75	1.23E-11	hsa-miR-15b-5p	-3.21	1.38E-05
hsa-miR-206	5.55	6.03E-12	hsa-miR-221-3p	-3.09	1.74E-05
hsa-miR-133b	4.60	2.89E-09	hsa-miR-7p	-2.67	0.00021
hsa-miR-95	3.94	0.00012	hsa-miR-15a-3p	-2.57	0.0004
hsa-miR-486-5p	3.67	7.14E-05	hsa-miR-3065-5p	-2.55	0.00208
hsa-miR-26b-5p	2.92	8.57E-05	hsa-miR-221-5p	-2.48	0.00051
hsa-miR-668	2.89	0.00163	hsa-miR-154-3p	-2.37	0.00086
hsa-miR-181a-5p	2.62	0.00021	hsa-miR-18a-3p	-2.26	0.00199
hsa-miR-1264	2.52	0.00127	hsa-miR-1285-3p	-2.16	0.00296
hsa-miR-378d	2.44	0.0036	hsa-miR-409-3p	-2.14	0.00207
hsa-miR-501-5p	2.37	0.00092	hsa-miR-16-2-3p	-2.09	0.00276
hsa-miR-660-5p	2.36	0.00075	hsa-miR-376b-3p	-2.06	0.00339
hsa-miR-188-5p	2.34	0.00096	hsa-miR-29b-3p	-2.01	0.00356
hsa-miR-128	2.33	0.00086			
hsa-miR-362-5p	2.23	0.00152			
hsa-miR-532-3p	2.22	0.00146			
hsa-miR-26a-5p	2.05	0.00297			
hsa-miR-664a-3p	2.05	0.00334			
hsa-miR-188-3p	2.02	0.00439			
hsa-miR-532-5p	1.94	0.00488			

A   B
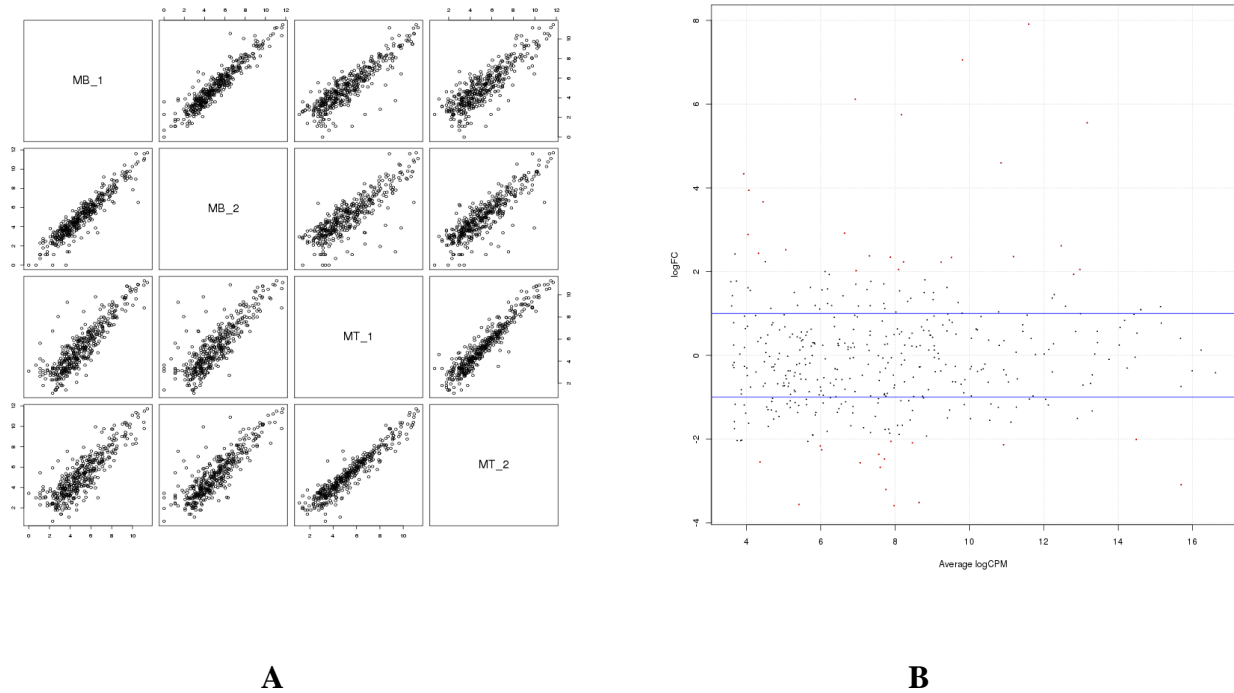
*Figure 2-4: (A) **Pair-plot of miRnome between myoblasts and myotubes.** Pairs plot compares the miRnomes of myoblasts and myotubes. The plots show the extent of similarity and differences between the samples compared. The pairs plot between replica of samples are also shown (MB_1 – Myoblast replica 1, MB_2 – Myoblast replica 2, MT_1 – Myotube replica 1, MT_2 – Myotube replica 2). (B) **Smear-plot between miRnome of myoblast and myotube.** Smear plot shows the differential expression of miRNA between myoblast and myotube. Dots above zero show differentially upregulated miRNA and vice versa. (logFC- log$_2$ Fold change, Average log CPM- ....)*
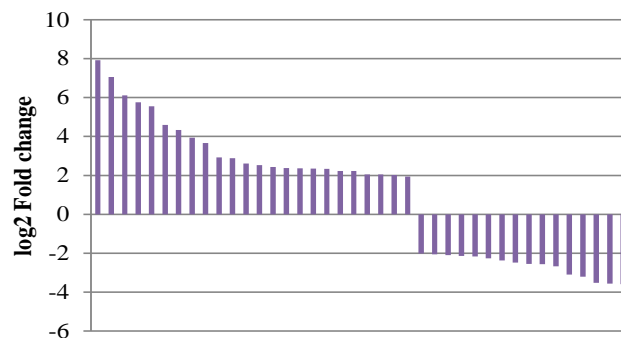
Figure 2-5: miRNA expression profile of myotubes. Shown are the differentially expressed miRNA in myotubes as compared to myoblasts.

2.3.2.2 miRNA expression profile in skeletal muscle tissue

The miRnome expression profile of skeletal muscle tissue was quite different and distinct from that of myotubes. The correlation coefficient (R^2 value) was 0.432, which indicates a very poor correlation between the miRNA expression profiles of myotubes and skeletal muscle tissue. Differential expression profile of miRNA in skeletal muscle tissue is shown in figure 2-7 and by Pairs and scatter plots also (figure 2-6 A & B).

90 miRNA were significantly over expressed and 191 miRNA were considerably down regulated in skeletal muscle tissue as compared to myotubes. The list of all these miRNA is shown in supplementary tables S2-5 and S2-6.

Bar plot of figure 2-8 shows all those miRNA which showed > 5 $\log_2$ FC in skeletal muscle tissue as compared to myotubes.

A   **B**
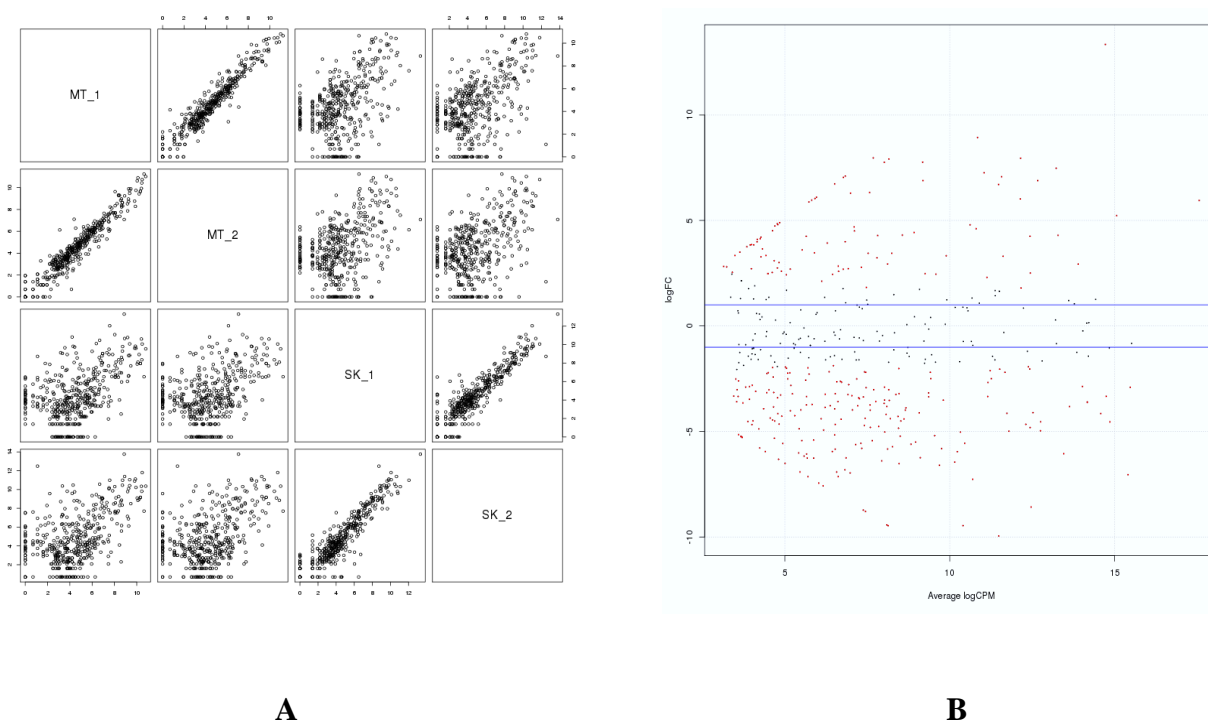
Figure 2-6: (A) Pairs-plot between miRnome of myotubes and skeletal muscle tissue. Pairs plot compares the miRnomes of myoblasts and myotubes. The plots show the extent of similarity and differences between the samples compared. The pairs plot between replica of samples are also

Figure2-7: miRNA expression profile in skeletal muscle tissue. Shown are all the differentially expressed miRNA in skeletal muscle tissue as compared to myotubes.

Both -3p and -5p arms of miR-101, miR-126, miR-142, miR-144, miR-30b, miR-338, miR-340, miR-378a, miR-486 and miR-499a were differentially upregulated in skeletal muscle tissue. Many IsomiRs were differentially over and under expressed in skeletal muscle tissue. Table 2-4 shows the list of those isomiRs.

IsomiRs Upregulated	*IsomiRs Downregulated*
hsa-miR-133a, b	hsa-miR-125a, b
hsa-miR-146a, b	hsa-miR-130a, b
hsa-miR-208a, b	hsa-miR-181a, b, c
hsa-miR-30a, b, c, d, e	hsa-miR-193a, b
hsa-miR-378a, c, d, f, i	hsa-miR-199a, b
hsa-miR-517a, b	hsa-miR-27a, b
	hsa-miR-301a, b
	hsa-miR-323a, b
	hsa-miR-34a, b, c
	hsa-miR-376a, b, c
	hsa-miR-487a, b

	hsa-miR-92a, b
	hsa-miR-99a, b

Table 2-4: IsomiRs differentially up and down regulated in skeletal muscle tissue. *Enlisted are the isomiRs found over and under expressed in skeletal muscle tissue.*
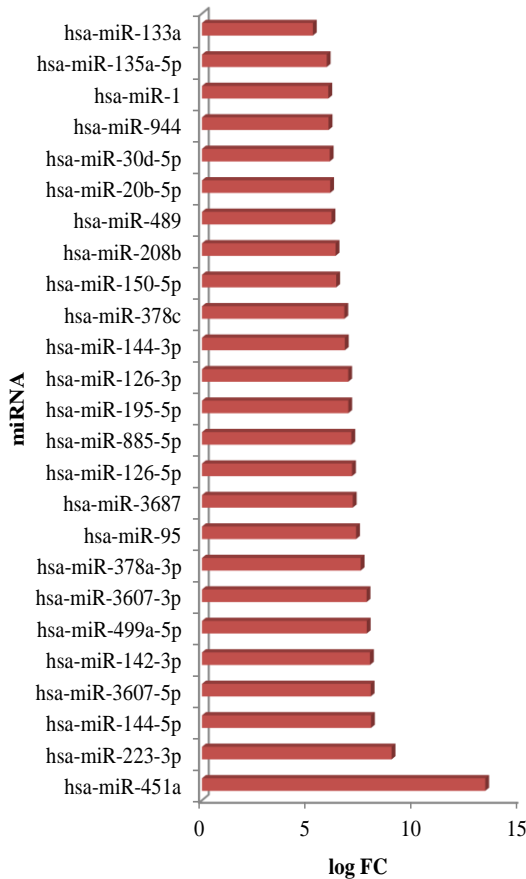
Figure 2-8: miRNA showing more than 5 log $_2$ fold increase in skeletal muscle tissue. miRNA showing more than 5 log_2 FC in skeletal muscle tissue as compared to myotubes are represented in this bar graph. Y-axis shows the miRNA and X-axis shows log_2 fold change.

There are miRNA found commonly expressed differentially in both myotubes and skeletal muscle tissue. Some of these miRNA were upregulated in myotubes whereas down regulated in skeletal muscle tissue and vice versa. Those miRNA have been listed in table 2-5.

miRNA	Up reg in Myotube *(log$_2$FC)*	Up reg in Skeletal muscle *(log$_2$FC)*	Down reg in Myotube *(log$_2$FC)*	Down reg in Skeletal muscle *(log$_2$FC)*
hsa-miR-1	7.912101	5.948486	--	--
hsa-miR-133a	7.058043	5.222945	--	--
hsa-miR-139-5p	6.117224	2.454135	--	--

hsa-miR-4488	5.748539	--	--	-9.42245
hsa-miR-133b	4.597984	2.927984	--	--
hsa-miR-95	3.941105	7.256696	--	--
hsa-miR-486-5p	3.669232	4.304105	--	--
hsa-miR-26b-5p	2.918038	3.321551	--	--
hsa-miR-181a-5p	2.615645	--	--	-4.66044
hsa-miR-378d	2.435125	3.44239	--	--
hsa-miR-501-5p	2.374571	--	--	-3.97934
hsa-miR-660-5p	2.356814	--	--	-2.11302
hsa-miR-188-5p	2.344865	--	--	-4.80874
hsa-miR-362-5p	2.230706	--	--	-3.51081
hsa-miR-532-3p	2.225932	--	--	-5.49949
hsa-miR-188-3p	2.022257	--	--	-5.83725
hsa-miR-532-5p	1.937058	--	--	-8.57967
hsa-miR-222-5p	--	--	-3.58772	-4.49821
hsa-miR-221-3p	--	--	-3.08959	-4.10406
hsa-miR-15a-3p	--	--	-2.56554	-4.67183
hsa-miR-221-5p	--	--	-2.47864	-4.65151
hsa-miR-154-3p	--	--	-2.36679	-2.23824
hsa-miR-409-3p	--	--	-2.13655	-4.50168
hsa-miR-376b-3p	--	--	-2.05503	-2.57024

Table2-5: miRNAs found common between myotube and skeletal muscle tissue. *Listed are the miRNA differentially expressed both in myotube and skeletal muscle tissue. miRNA highlighted in blue are over expressed in myotube but under expressed in skeletal muscle tissue.*

2.3.3 Ago2 associated mRNA

We wanted to investigate the mRNA repertoire which is associated with argonaute 2 (Ago2) protein, which is a member of RISC complex in order to have a better idea of mRNA pool which is likely under the regulation of miRNA. We performed Ago2-RNA immunoprecipitation for both myoblast and myotube cells, in replica. The RIP sample was enriched for poly(A) RNA using oligo dT column and sequenced using SOLiD sequencing. We then compared these results with the non Ago2 pull down RNA population from myoblasts and myotubes. The mRNAs which were enriched with Ago2 pull down as compared to non-pull down most likely are the sequences which are directly under the control of miRNA regulation.

We found 1493 and 1979 genes were enriched with Ago2 pulldown in myoblasts and myotubes respectively. The comparison between pull down and non pull down samples

are shown by pairs plot of figure 2-9. Immunoblotting result of Ago2-RIP is shown in figure 2-10.

A **B**
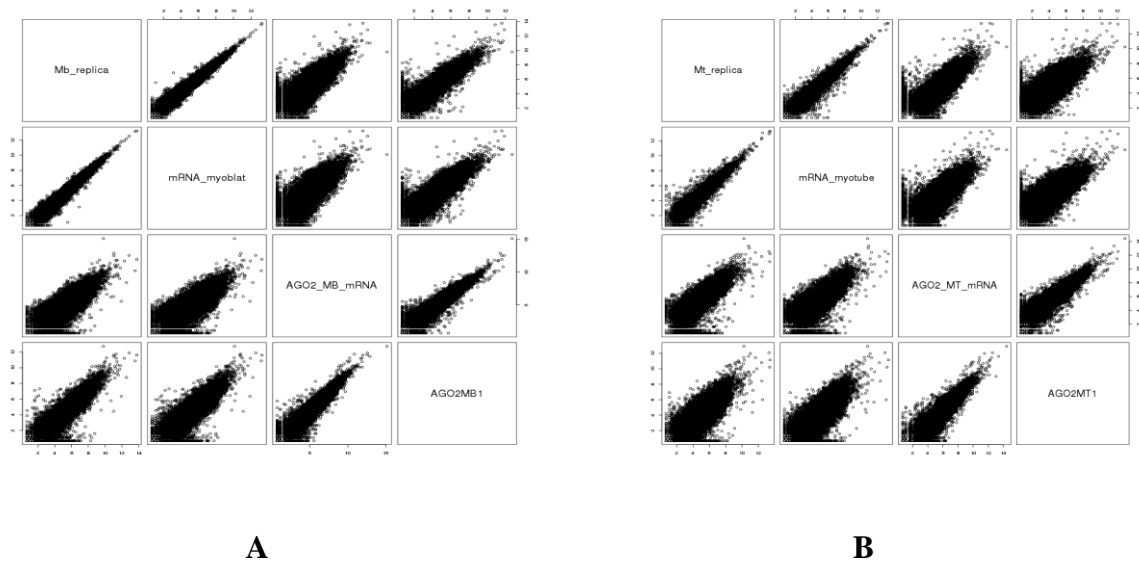
Figure 2-9: Pairs plot between Ago2 pull down and non-pull down RNA seq results for (A) Myoblasts (B) Myotubes (Mb_replica+ mRNA_myoblast = name of non pull down myoblast RNA seq, AGO2_MB_mRNA + AGO2MB1 = name of Ago2 pull dowm RNA seq for myoblasts, Mt_replica+mRNA_myotube = name of non pull down myotube RNA seq, AGO2_MT_mRNA+AGO2MT1 = name of Ago2 pull dowm RNA seq for myotubes).

85 kD →

Ago2 IP samples   Cell lysates
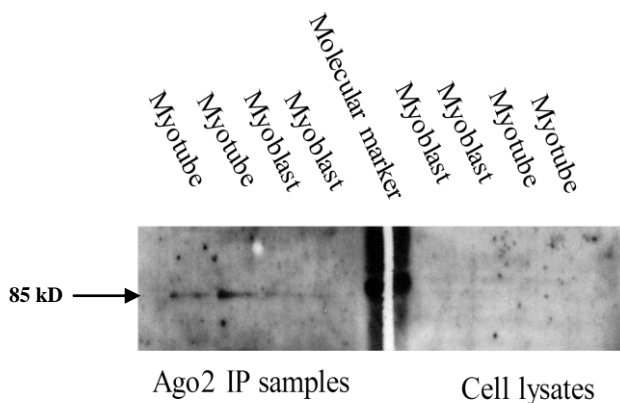
Figure 2-10: ECL image of Ago2 pull down in CHQ5B myoblast and myotube cells. We observed that the Ago2 concentration seemed higher in myotubes than in myoblasts.

2.3.3.1 Ago2 enriched mRNA in myoblasts

To understand the list of genes better, we found out those genes which were enriched with the Ago2 pull down in myoblast and genes which were under expressed in myoblasts. Our rationale behind this is, genes down regulated by miRNA should be associated with

Ago2 protein and hence should be enriched when pulled down with argonaute, when their actual expression in the cell remains low if is down regulated by miRNA. We found 173 such genes in myoblasts. List can be found from supplementary table S2-11. We also carried out biological process enrichment using online software DAVID for these genes and found that they were skeletal muscle process associated biological terms like muscle system process, muscle contration, muscle tissue development etc (supplementary table S2-13).

2.3.3.2 Ago2 enriched mRNA in myotubes

Following the rationale mentioned in the above section, we did the similar analysis for myotubes and found 134 genes down regulated in myotubes but enriched with Ago2 pull down in myotubes (supplementary table S2-12). GO analysis of biological process terms showed enrichment of DNA metabolic process, protein polymerization and carbohydrate catabolic process terms (supplementary table S2-14).

Although there are a huge number of genes found enriched with Ago2 pull down, but we do not have a convincing explanation for them, so we have only highlighted those genes which our experimental evidence suggest to be the direct candidates under miRNA regulation.

2.4 Discussion:

Aiming at understanding the molecular basis of skeletal muscle development, we report the high-throughput sequencing analysis of mRNA and small RNA transcriptome of primary human muscles. Transcriptome analysis is useful in understanding the relation between genome and functioning of the cells (13). In this study we have addressed the mRNA and miRNA constitution of primary human muscle cells at both the myoblast and myotube stage of development. We did the profiling of mRNA and miRNA for adult skeletal muscle as well. The expression level of both mRNA and miRNA displayed a broad range. Our study provides the large scale characterization of the transcriptome and its modulation during the stages of muscle cell proliferation and differentiation and of skeletal muscle tissue.

We found far more number of genes differentially expressed between myotubes and skeletal muscle tissue. The overlap between of transcriptomes of myotubes and skeletal muscle tissue was less (R^2 value = 0.56) as compared to the extent of overlap between

myoblasts and myotubes (R^2 value = 0.77). This was expected as the transition from a cellular stage to next is not as complicated as from becoming tissue from cell. On differentiation we found many genes for specific protein families were upregulated. ABC superfamily members, which are transmembrane proteins and utilise the energy of ATP to carry out biological process of substrate translocation, RNA translation and DNA repair (14, 15) were upregulated in myotubes as well as skeletal muscles. Genes from ASB family were also upregulated. Ankyrin and SOCS box containing proteins have their role in fixing contractile apparatus to muscle membrane (16). VDCC subunit genes have their role in skeletal muscle differentiation were also upregulated (17). Genes for Voltage gated sodium and potassium channels were also turned up at this level of muscle development. Some Fork head (FOX) protein genes showed upregulation. Role of FoxO in preventing muscle wasting is well documented (18, 19). In skeletal tissue we found members of troponin complex protein genes upregulated. Along with many protein families whose role in muscle is known we found many chemokine family, interleukin, interleukin receptors and major histocompatibility complex family member genes were significantly upregulated in skeletal muscle tissue, but they were not turned up in myotubes.

With the pathway analysis we found, metabolic pathways were the most pronounced pathways with highest number of altered genes, after differentiation and at the tissue level. This shows that cells undergo major metabolic changes at each stage of development from proliferative stage to becoming tissue. Many altered genes in myotubes and skeletal muscle belonged to PI3k-Akt and MAPK signalling pathways, which are very well documented pathways in muscle (20, 21, 22). We performed GO analysis separately for genes differentially up and down regulated. The Grouping into GO functional classes provided insights into the function of differentially expressed genes. GO analysis revealed several interesting classes. Genes involved in muscle system process, striated muscle contraction, muscle organ development, contractile fiber, I band, sarcomere, biological and cellular compartment terms were remarkable upregulated on myotube formation. Significant cell cycle function terms like cell cycle processes like M-phase, nuclear division, mitosis, cell cycle process, chromosome, condensed chromosome contained genes down regulated upon myotube formation. So, the GO analysis reflected the transition of entire transcriptome from proliferative phase of myoblast to muscle specific non dividing state of myotubes. The interesting thing we found with gene

ontology for skeletal muscle tissue was the marked enrichment of biological process related to immune response, defence response, response to wounding, inflammatory response along with other biological process of muscle contraction, cellular respiration, electron transport chain. Muscle cells after differentiation although did not show any immunogenic capacity. GO terms of extracelluar matrix organization, extracellular parts were enriched for genes down regulated in skeletal muscle tissue. May be the immunogenic capacity is only acquired after becoming tissue and not in the cells by themselves.

miRNA analysis showed more miRNA were down regulated in skeletal muscle tissue in comparison with myotubes than were in myotubes compared to myoblasts. We also looked for miRNA common between myotubes and skeletal muscle tissue. We found many miRNA which are differentially upregulated after myotube formation but are not studied adequately yet. So we focused on their functional analysis, which is discussed in the following chapter.

Since RNA seq gives a huge amount of data and a lot of aspects look very interesting as well worth investigating further. But distilling down that vast amount of information into one single interesting thing is extremely difficult. Here we had the entire knowledge of muscle transcriptome from cells to tissue and we also had the knowledge of miRNA expression which are important modulators of post transcriptional regulation. In order to understand this huge amount of information in a biologically meaningful way, we wanted to find out how the two set of RNAs (i.e mRNA and miRNA) interact with each other. The straight forward way to do so was by performing Ago2 RIP and then carry out RNA seq analysis. This provides a simple and direct experimental method to identify mRNA targets. We performed Ago2-RIP in myoblasts and myotubes. We then extracted the poly(A) from the immunoprecipitated complexes and performed RNA seq analysis. To find out the genes enriched with Ago2 we compared sequences with the total gene pool of non Ago2 pull down. We found several genes enriched with Ago2. We then compared the enriched genes, with the genes turned down in myoblasts and myotubes. We found 171 such genes in myoblasts and 132 genes in myotubes. We found many muscle structural genes were enriched with Ago2 in myoblasts including MYOG which is a differentiation marker. Since myoblast is still a proliferative state where in muscle specific genes are not turned up (as we have shown with our RNA seq results), Ago2 enrichment of muscle structural genes indicate them to be under miRNA regulatory control. With GO analysis

of these genes we found them enriched for biological processes of muscle contraction, muscle development, regulatory processes. Genes enriched with Ago2 in myotubes consisted of transcription factors like PAX7, MYF5, MYF6, HOXA11and HOXA9. Pax7 is known to be down regulated for muscle differentiation (23). HoxA11 is known to be under miR-181 regulation in muscle (9). GO of genes enriched with Ago2 showed significant enrichment for DNA metabolic processes. Since the cells upon differentiation cease to multiply, genes related to DNA replication, ligation or strand elongation are probably turned down. Our experimental evidences and literature curation shows the genes enriched with Ago2 pull down in myoblasts and myotubes are very likely to be under miRNA control. Although, it will certainly need further more experimental evidence to verify and validate which miRNA act on which mRNA. A transcriptome wide gene comparison between cellular development stages and adult tissue has not been described before to our knowledge.

In summary, we have demonstrated the total pool of differentially expressed genes in human myoblasts, myotubes and adult skeletal muscle. We have also found differential expression of miRNA and with Argonaute-2 pull down we have been able to find those genes which are most likely under the control of miRNA regulation.

2.5 <u>Reference:</u>

1. Schuster S, Fell DA, Dandekar T (2000): A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nat Biotechnol, 18(3)**:**326–332.
2. Chen J F et al, (2005): The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genetics 38: 228-233.
3. McCarthy J J, (2008): MicroRNA 206: The skeletal muscle specific myomiR. Biochimica et Biophysica Acta (BBA) Gene regulatory mechanisms 1779 (11): 682-691.
4. Koutsoulidou A, Mastroyiannopoulos N P, Furling D, Uney J B and Phylactou L A, (2011): Expression of miR-1, miR-133a, miR-133b and miR-206 increases during development of human skeletal muscle. BMC Dev Biol 11(34).
5. Molnár A, Schwach F, Studholme D J, Thuenemann E C, Baulcombe D C, (2007): miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. Nature 447 (7148): 1126–9.

6. Lee C T, Risom T, Strauss W M. (2007): Evolutionary conservation of microRNA regulatory circuits: an examination of microRNA gene complexity and conserved microRNA-target interactions through metazoan phylogeny. DNA Cell Biol. 26 (4): 209–18.

7. Cordes KR. et al. (2009): miR-145 and miR-143 regulate smooth muscel cell fate and plasticity. Nature 460: 705 – 710.

8. Sun Q, Zhang Y, Yang G, Chen X, Zhang Y, Cao G, Wang J, Sun Y, Zhang P, Fan M, Shao N, Yang X. (2008): Transforming growth factor-β- regulated miR-24 promotes skeletal muscle differentiation. Nucl Acids Res. 36(8): 2690 – 2699.

9. Naguibneva I, Zazoua MA, Polesskaya A, Ait-Si-Ali S, Groisman R, Souidi M, Cuvellier S and Bellan AH.(2005): The micro-RMA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. Nature Cel. Biol. 8(3):278 – 284.

10. Chen JF, Tao Y, Li J, Deng Z, Yan Z, Xiao X and Wang DZ. (2010): microRNA-1 and microRNA-206 regulate skeletal muscle satellite cell proliferation and differentiation by repressing Pax7. J Cell Biol. 190(5): 867 – 879.

11. Chen JF, Mandel E, Thomson JM, Wu Q, Callis TE, Hammond SM, Conlon FL and Wang DZ. (2006): The role of microRNA-1 and microRNA -133 in skeletal muscle proliferation and differentiation. Nature Genetics 38 (2): 228 -233.

12. Kim HK, Lee YS, Sivaprasad U, Malhotra A and Dutta A. (2006): Muscle specific microRNA-206 promotes muscle differentiation. J Cell Biol. 174(5): 677-687.

13. Adams JU. (2008): Transcriptome: connecting the genome to gene function. Nature education 1(1): 195.

14. Davidson AL, Dassa E, Orelle C, Chen J. (2008): Structure, function and evolution of bacterial ATP-binding cassette systems. Microbiol. Mol. Biol. Rev.72 (2): 317–64.

15. Goffeau, A.; Hertogh B; and Baret PV. (2004): ABC Transporters. In: Encyclopedia of Biological Chemistry. 1: 1–5.

16. Tee JM and Peppelenbosch MP.(2010): Anchoring skeletal muscle development and disease: the role of ankyrin repeat domain containing proteins in muscle physiology. Crit Rev Biochem Mol Biol 45(4):318-330.

17. Bidaud I, Monteil A, Nargeot J, Lory P. (2006): Properties and role of voltage-dependent calcium channels during mouse skeletal muscle differentiation. J Mus Res Cell Motility 27(1):75-81.

18. Demontis F and Perrimon N. (2010): FOXO/4E-BP signalling in drosophila muscles regulate organism-wide proteostasis during aging. Cell 143(5): 813-825.

19. Reed SA, Sabdesara PB, Senf SM and Judge AR. (2010): Inhibition of FoxO transcriptional activity prevents muscle fiber atrophy during cachexia and induces hypertrophy. 26(3): 987 – 1000.

20. Kramer H and Goodyear LJ. (2007): Exercise, MAPK and NF-κB signalling in skeletal muscle. J App. Physiol. 103(1):388-395.

21. Keren A, Tamir Y, Bengal E. (2006): The p38 MAPK signalling pathway: A major regulator of skeletal muscle development. Mol Cell Endocri. 252(1):224-230.

22. Sakamoto K, Aschenbach G, Hirshman MF, Goodyear LJ (2003): Akt signalling in skeletal muscle : regulation by exercise and passive stretch. Am J Physiol-Endocri and Metabol 285:E1081-1088.

23. Zammit PS, Relaix F, Nagata Y, Ruiz AP, Collins CA, Partridge TA and Beauchamp JR. (2006): Pax7 and myogenic progression in skeletal muscle satellite cells. J Cell Sci 119: 1824 – 1832.

24. Guigal N, Rodriguez M, Cooper RN, Dromaint S, Santo JP, Mouly V, Boutin JA and Galizzi JP (2002) Uncoupling protein 3 (UCP3) mRNA expression in reconstituted in human muscle after myoblast transplantation in $RAG2^{-/-}/\gamma^{-}_{c}/C5^{-}$ immunodeficient mice. J Biol. Chem. 277 (49): 47407 – 47411.

Chapter-3

Functional analysis of hsa-miR-139-5p, hsamiR-532-5p, hsa-miR-660-5p and hsa-miR-92a-3p in human skeletal myoblasts

3.1 Introduction:

Based on the miRnome studies performed on skeletal muscle cells, we found out many differentially expressed miRNA between myoblasts and myotubes. In this chapter I focus on the functional analysis of hsa-miR-206, hsa-miR-139-5p, hsa-miR-532-5p, hsa-miR-660-5p and hsa-miR-92a-3p. The reason behing selection hsa-miR-139-5p, hsa-miR-532-5p and hsa-miR-660-5p is their differential upregulation in myotubes and they are less studied miRNAs with regards to muscle biology. hsa-miR-206 however is widely reported as an important and specific miRNA during muscle cell differentiation and proliferation (1,2,3), so I chose to use hsa-miR-206 as a positive control for the experiment. hsa-miR-92a-3p is upregulated in myotubes however not differentially, but this miRNA is one of the highly expressed miRNAs thereby making itself a molecule of interest for deeper functional analysis. The functional significance of miRNAs was studied by creating transient transfectants with miRNA mimics. To investigate the influence of miRNAs, we transfected miRNAs into human skeletal muscle cells (CHQ5B cells) and performed the genome wide expression analysis using NGS. We decided to perform the whole transcriptome analysis at two time points of 12 hours and 24 hours after miRNA mimic transfection. At 12 h time point we wanted to see the immediate impact and after 24 hours the prolonged effect of miRNA over expression. We had a negative control with which we compared each miRNA to find out the differential expression of genes caused by miRNA over expression. Our intention was not to find out the likely targets for these miRNAs, but to observe the overall impact these miRNA cast on the myoblast cells and if their over expression can induce any differentiation (since they are upregulated upon differentiation), even when the cells are left under growing conditions in 20% FBS medium.

Both hsa-miR-660 and hsa-miR-532 belong to mir-188 family and their host gene is located in X chromosome (table3-1). Upregulation of hsa-miR-660-5p, upon differentiation has been reported in porcine skeletal muscle [4] and also in human primary myoblasts (5). hsa-miR-660 also shows upregulation Facioscapulohumeral muscular dystrophy (FSHD) as compared to normal human myoblast cells (6). However no function of hsa-miR-660-5p in muscle development has been described yet. So we wanted to investigate the role of this miRNA. hsa-miR-532-5p is also a poorly studied miRNA with regards to muscle biology. It is however known to be expressed in mesenchymal stem cells which promote tissue repair/regeneration through release of

growth factors, cytokines and extracellular matrix molecules (7). Repair and regeneration are very important aspects of muscle development.

miRNA	Genomic position / Coordinates	Gene Family
hsa-miR-139	Chr 11: 72326107 – 72326174 [-]	mir - 139
hsa-miR- 92a	Chr 13: 92003568 – 92003645 [+]	mir - 25
hsa-miR-206	Chr 6: 52009147 – 52009232 [+]	mir - 1
hsa-miR- 532	chrX: 49767754 – 49767844 [+]	mir – 188
hsa-miR- 660	chrX: 49777849 – 49777945 [+]	mir – 188

Table 3-1: Genome context of miRNA

miR-92a which is a member of mir 17-25 cluster, has recently been shown to be involved in post transcriptional regulation of Hand 2 gene in heart muscle (8). miR-92a has also been found involved some primary muscular disorders (9). But in-depth knowledge of role played by miR-92a in muscle development is lacking. Role of hsa-miR-139-5p during skeletal muscle development has also not been investigated yet.

In this chapter I have tried to shed light on the functional impact of these miRNAs, which remain poorly studied and investigated in human skeletal muscle development so far.

Experimental plan is depicted below:
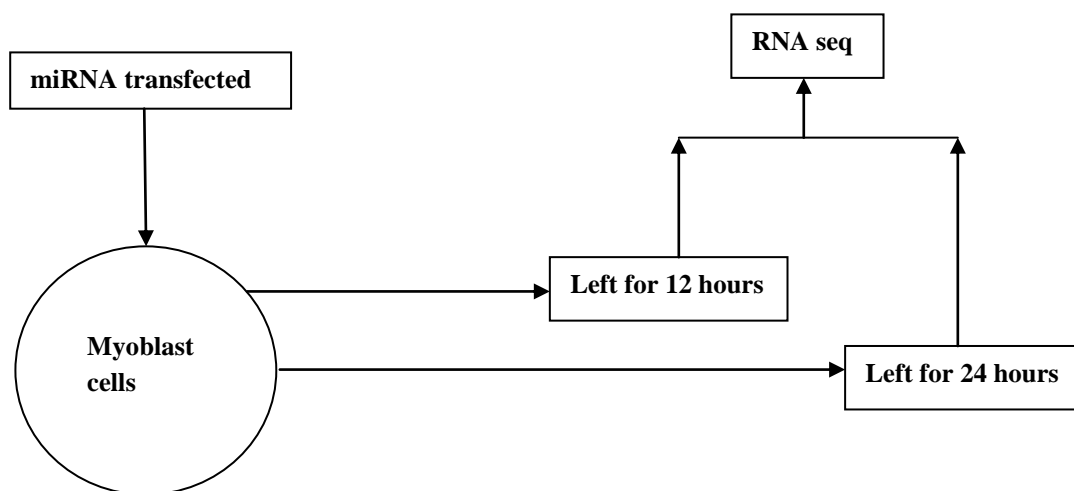
3.2 Materials and methods:

3.2.1 Cell culture

CHQ5B myoblast cells were cultured in 6cm petri dishes. $5.9*10^5$ cells were cultured in one dish. Cells were maintained in 20% FBS medium throughout.

3.2.2 Transient transfection with oligonucleotides

CHQ5B myoblasts were transfected with miR-206, miR-139-5p, miR-532-5p, miR-660-5p and miR-92a-3p mirVana mimics at a 40nM concentration, or with a non-targeting negative control miRNA (miR-NC) oligos (Ambion) using Lipofectamine RNAiMax (Invitrogen) according to the manufacturer's instructions. Transfected cells were harvested for *in vitro* studies after culturing for 12h and 24h.

3.2.3 RNA extraction and Whole transcriptome sequencing

Poly(A)$^+$ RNA was extracted from the cells after 12h and 24h of transfection using Amersham kit (details has been mentioned in chapter I). Quality of RNA was tested using agilent technologies and quantified using Qbit. The mRNA was then processed for sequencing using manufacturer's instructions and sequencing was performed using Applied biosystem's SOLiD technology.

Differential gene expression analysis was performed using edgeR software with a P value of 0.1 considered statistically significant.

3.3 Results:

3.3.1 miRNA expression upon muscle differentiation

Upon differentiation of CHQ5B myoblasts into myotubes we found many miRNA differentially expressed. In details this has been discussed in chapter I. We found hsa-miR-139-5p, hsa-miR-532-5p, hsa-miR-660-5p and hsa-miR-206 were differentially upregulated upon myotube formation with an FDR 0f 0.05. hsa-miR-92a-3p was upregulated although not differentially, but its level of expression was very high and prominent therefore we chose to study it. Table 3-2 shows $\log_2$ FC for each miRNA. The bar plot in figure 3-1 show the expression level of these miRNAs in both myoblast and myotubes.

miRNA	Log$_2$ FC	P value
hsa-miR-139-5p	6.12	2.83E-11
hsa-miR-532-5p	1.93	0.004881
hsa-miR-660-5p	2.35	0.000749
hsa-miR-92a-3p	0.96	0.152295
hsa-miR-206	5.55	6.03E-12

Table 3-2: miRNA differentially upregulated in myotubes with their log$_2$ fold change values and P-values.
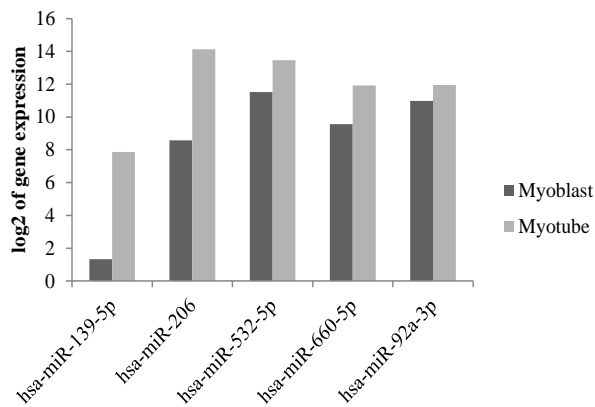
Figure 3-1: Bar plot showing miRNA expression in myoblasts and myotubes. From this plot the differential upregulation of these miRNAs can be seen.

3.3.2 Transient miRNA over expression caused higher number of genes upregulated and less genes down regulated

Differentially expressed genes for each miRNA transfection were found as compared to negative control with an FDR of 0.1. Whole transcriptome analysis showed that, each miRNA over expression caused both up regulation and down regulation of several genes. We found that each over expression caused, higher number of genes differentially upregulated as compared to the number of genes down regulated for both time points of 12 hours and 24 hours (figure 3-2).

	139-12H	206-12H	532-12H	660-12H	92a-12H	139-24H	206-24H	532-24H	660-24H	92a-24H
up	48	153	102	0	11	347	419	312	536	427
down	15	86	26	2	3	325	144	75	162	271
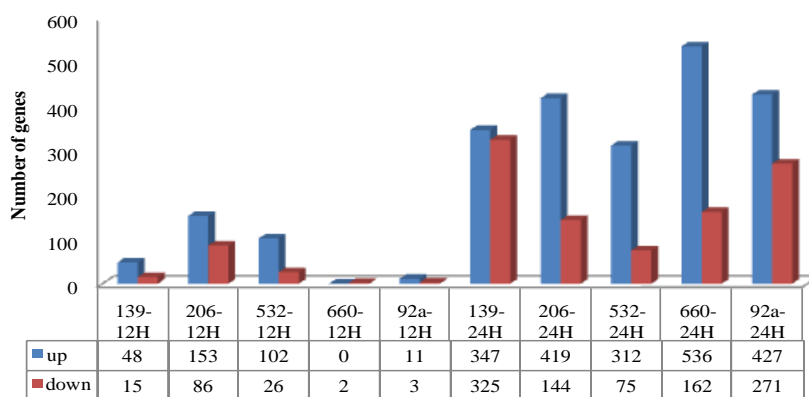
Figure 3-2: Bar plot showing gene expression after 12 and 24 hours of transfection. *139-12H/24H = hsa-miR-139-5p transfection effect after 12h/24h, 206-12H/24H = hsa-miR-206 transfection effect after 12h/24h, 532-12H/24H = hsa-miR-532-5p transfection effect after 12h/24h, 660-12H/24H = hsa-miR-660-5p transfection effect after 12h/24h, 92a-12H/24H = hsa-miR-92a-3p transfection effect after 12h/24h. The number of genes showing differential expression is also shown in the boxes below the bars.*

24 hours post transfection with miR-92a and miR-660, highest number of genes was found affected, and miR-660 over expression showed highest number of genes upregulated (table 3-3).

miRNA	Up regulated	Down regulated	Total genes affected
hsa-miR-139	347	325	672
hsa-miR- 92a	427	271	698
hsa-miR-206	419	144	563
hsa-miR- 532	312	75	387
hsa-miR- 660	536	162	698

Table 3-3: 24h post miRNA over expression effect. *Shown are the numbers of genes up and down regulated for each miRNA. The last column shows the total number of genes altered after 24 h.*

As a general expectation, far fewer genes were found affected after 12 hours of miRNA mimic transfection. 12 hours post transfection with miRNA-206 showed highest number of genes with altered expression (table 3-4).

miRNA	Up regulated	Down regulated	Total genes affected
hsa-miR-139	48	15	63
hsa-miR- 92a	11	3	14
hsa-miR-206	153	86	239
hsa-miR- 532	102	26	128
hsa-miR- 660	0	2	2

Table 3-4: 12h post miRNA over expression effect. *Shown are the numbers of genes up and down regulated for each miRNA. The last column shows the total number of genes altered after 24 h.*

3.3.3 Differential gene expression, Pathway and GO analysis for genes affected by each miRNA

We used edgeR software for differential gene expression analysis. KEGG pathway database (http://www.genome.jp/kegg/pathway.html) using the Homo sapiens reference

pathway and DAVID online software (http://david.abcc.ncifcrf.gov) were used to determine the cellular pathways and biological functions that could be influenced by over expression of miRNAs in CHQ5B myoblast cells. The pathways and processes associated with altered genes for each miRNA have been discussed in details in the following sections.

3.3.3a) hsa-miR-139-5p analysis

We observed that after miR-139-5p transfection, a total of 63 and 672 genes were affected post 12 and 24 hours of transfection respectively. After 12 hours of transfection, 15 snoRNA genes were found down regulated and only 48 genes were upregulated. Genes found most repressed and considerably over expressed after transient transfection have been listed in supplementary table S3-1. Targetscan database, predicts NRK, MYCBP and HDX as targets of miR-139-5p, which we also found were -5.55, -5.77 and -6.06 $\log_2$ folds down regulated respectively after 24 hours. This observation is suggestive that these three genes might be directly repressed by miR-139-5p.

For genes differentially regulated after 24 hours of miR-139-5p over expression, KEGG pathways with ≥ 7 differentially expressed genes have been listed in table 3-5, but in fact many more pathways were detected by KEGG. The 7 pathways with at least 7 differentially expressed genes are listed along with their log FC values.

KEGG Pathways	Differentially expressed genes (with their log FC)
hsa01100 Metabolic pathways (28)	ALDH4A1 (-6.38), ALG9 (5.42), AOC3(-6.76), ARG2(6.62), ATP6V1E2(5.77), B3GALNT1(-4.42), B4GALT6(5.39), BDH1(-5.91), CBR3(6.64), CDA(5.48), COQ2(4.89), COQ3(-6.30), COX6A1(-4.39), DCXR(5.61), GCLC(5.24), HSD11B1(2.99), MGAM(-5.37), NDUFA9(6.62), NDUFS6(4.41), PAFAH2(6.55), PIGH(-5.49), PIGV(5.56), POLR3K(5.24), PPT2(5.61), PTGDS(5.18), SHMT1(-5.37), UPRT(5.77), ZNRD1(5.56)
hsa04151 PI3K-Akt signaling pathway (10)	CCNE1(6.19), COL11A2 (5.70), CSF3 (5.73), FGF16 (-5.49), IGF1(5.39), ITGA8(-5.61), LPAR6(6.19), NGFR(-3.93), PKN3(-5.77), TLR4(5.73)
hsa05203 Viral carcinogenesis (9)	C3 (8.26), CCNE1(6.19), CDKN2A (5.94), HIST1H2BF (5.39), HIST1F2BJ (5.33), HIST1H2BN (5.84), HIST1H4E (5.50), KAT2B (5.42), LYN (-5.33)
hsa04062 Chemokine signaling pathway (7)	ADCY4 (4.78), CXCL3(6.45), CXCL5(6.73), CXCL6(4.18), GRK5(4.75), ITK(-5.88), LYN(-5.33)
hsa05016 Huntington's disease	BBC3(5.98), COX6A1(-4.39), NDUFA9(6.62), NDUFS6 (4.41), PPARG(-

(7)	6.33), SLC25A4(6.00), TFAM (5.96)
hsa05322 Systemic lupus erythematosus (7)	C3(8.26), C4A (-5.84), C4B(-5.40), HIST1H2BF(5.39), HIST1H2BJ(5.33), HIST1H2BN(5.84), HIST1H4E(5.50)
hsa05133 Pertussis (7)	C3(8.26), C4A (-5.84), C4B(-5.40), CXCL5 (6.73), CXCL6(4.18), NOD1 (5.68), TLR4 (5.73)

Table3-5: KEGG pathways for genes found altered after 24 hours of hsa-miR-139-5p over expression. *The table shows the pathway names along with the genes involved. The LHS column gives the pathway names with the number of genes involved mentioned with in the brackets. In the RHS column, in bracket the log_2 Fold change values of corresponding genes are shown. Values in negative show down regulation and vice versa.*

Pathway and gene ontology analysis shows that hsa-miR-139-5p over expression after 24 hours, might influence genes involved in metabolism, PIK3-Akt signalling (which regulates fundamental cellular functions such as transcription, translation, proliferation, growth and survival), viral carcinogenesis, chemokine signalling, transcription regulation and chemokine activity. However, after 12 hours hsa-miR-139-5p seems to affect biological processes of regulation of coagulation, regulation of response to stimulus and cellular response to stress. (supplementary table S3-6).

3.3.3b) hsa-miR-532-5p analysis

128 and 327 genes were found altered after 12 and 24 hours of miR-532 over expression. Most prominently affected genes are listed in supplementary table S3-2. After 12 hours only some snoRNAs were found down regulated but no protein coding genes were turned down.

For genes differentially regulated after 24 hours of miR-139-5p over expression, KEGG pathways with ≥ 6 differentially expressed genes have been listed in table 3-6. The 7 pathways with at least 6 differentially expressed genes are listed along with their log FC values.

24 HOURS	
KEGG Pathways	Differentially expressed genes (with their log FC)
hsa01100 Metabolic pathways (18)	ALG10(5.32), ALG9(5.65), AOC3(-6.48), BPGM(5.86), CBR3(4.52), COX7B(5.50), CYP2R1(5.21), GALNT12(4.27), GCLC(6.01), IDI1(6.47), NDUFA9(5.65), NDUFS6(4.24), PAFAH2(4.74), PIGV(4.77), POLR3F(5.21), QPRT (-4.07),

	SDHD(4.07), TH(-4.83)
hsa05203 Viral carcinogenesis (7)	C3(6.56), CCNE1(5.62), CDKN2A(5.43), HIST1H2BD(4.93), HIST1H4E(6.28), KAT2B(5.21), REL(5.62)
hsa04060 Cytokine-cytokine receptor interaction (7)	ACVR2A(5.62), CXCL1(5.04), CXCL3(5.04), CXCL5(5.05), CXCL6(4.82), EGF(5.36), IL15 (5.17)
hsa05016 Huntington's disease (7)	BBC3(3.49), COX7B(5.50), NDUFA9(5.65), NDUFS6(4.24), SDHD(4.07), SLC25A4(5.04), TFAM(5.65)
hsa05012 Parkinson's disease (6)	COX7B(5.50), NDUFA9(5.65), NDUFS6(4.24), SDHD(4.07), SLC25A4(5.04), TH(-4.83)
hsa04062 Chemokine signaling pathway (6)	CXCL1(5.04), CXCL3(5.04), CXCL5(5.05), CXCL6(4.82), GRK5(5.56), ITK(-5.60)
12HOURS	
KEGG Pathways	**Differentially expressed genes (with their log FC)**
hsa01100 Metabolic pathways (10)	ALG9(5.86), CBR3(5.70), DCXR(5.88), FUT4(5.91), IDI1(4.56), NDUFA9(5.53), PAFAH2(6.07), POLR3F(5.76), PTGS2(4.24), UPRT(5.93)

Table3-6: KEGG pathways for genes found altered after 12 and 24 hours of hsa-miR-532-5p over expression. The table shows the pathway names along with the genes involved. The LHS column gives the pathway names with the number of genes involved mentioned with in the brackets. In the RHS column, in bracket the log_2 Fold change values of corresponding genes are shown. Values in negative show down regulation and vice versa.

Genes for almost 22 zinc finger proteins, myosin heavy chain 15 (MYH15), some histone cluster protein and chemokine ligand were upregulated after 24 hours of miR-532-5p over expression.

Pathway and gene ontology analysis shows that hsa-miR-532-5p over expression after 24 hours, might influence (both by up regulating and down regulating) genes involved in metabolism, viral carcinogenesis, cytokine-cytokine receptor interaction, regulation of transcription, chemokine activity, response to hypoxia. However, after 12 hours hsa-miR-532-5p seems to affect along with metabolic pathways, biological processes of regulation of coagulation, regulation of response to stimulus and cellular response to stress. (supplementary table S3-7).

3.3.3c) hsa-miR-660-5p analysis

We observed that after 12 hours of miR-660-5p transfection, no alteration to gene expression was caused. Only 2 genes were down regulated after 12 hours. However 24 hours post-transfection, 698 genes showed differential expression. Thereby, indicating that miR-660-5p probably takes a gradual course of action into the cells, as compared to other miRNAs which we have studied. Supplementary Table S3-3 lists the genes found most significantly influenced by miR-660-5p over expression.

For genes differentially regulated after 24 hours of miR-660-5p over expression, KEGG pathways with ≥ 6 differentially expressed genes have been listed in table 3-7. The pathways with at least 6 differentially expressed genes are listed along with their log FC values.

24 HOURS	
KEGG Pathways	**Differentially expressed genes (with their log FC)**
hsa01100 Metabolic pathways (27)	AGMAT(-4.54), ALG14(4.80), ALG9(5.67), AOC3(-3.43), CBR3(6.23), COX10(4.00), COX7B(4.34), CYP2R1(5.99), DCXR(6.09), FUT4(5.73), GCLC(6.36), MCAT(4.69), NDUFA9(4.55), NDUFS6(4.71), PAFAH2(6.56), PIGP(5.63), PIGV(4.48), POLR3K(5.57), PPT2(5.27), SDHD(3.84), SLC27A5(-4.31), SMPD2(4.30), SORD(-4.60), SPR(-3.41), TPK1(-5.47), UPRT(5.42), ZNRD1(4.85)
hsa05203 Viral carcinogenesis (9)	C3(5.49), CCNE1(5.70), CDKN2A(5.85), HIST1H2BD(5.30), HIST1H2BN(5.07), HIST1H4A(4.97), IRF9(4.91), KAT2B(5.85), REL(5.46)
hsa05016 Huntington's disease (7)	BBC3, COX7B, NDUFA9(4.55), NDUFS6(4.71), SDHD(3.84), SLC25A4, TFAM(5.99)
hsa04151 PI3K-Akt signaling (7)	CCNE1(5.70), FGF16(-5.03), NGF(4.74), PPP2R2B(5.67), PPP2R5A(3.37), TLR4(5.99), VWF(-4.66)
hsa05200 Pathways in cancer (7)	AXIN2(-3.58), CCNE1(5.70), CDKN2A(5.85), FGF16(-5.03), FLT3LG(-4.50), LEF1(5.38), MITF(5.02)
hsa05012 Parkinson's disease (6)	COX7B(4.34), NDUFA9(4.55), NDUFS6(4.71), SDHD(3.84), SLC25A44.74), SNCAIP(5.34)
hsa03010 Ribosome (6)	FAU(3.84), MRPS14(5.41), RPL18A(3.79), RPL41(4.55), RPS26(5.02), RPS7(3.54)
hsa05010 Alzheimer's disease (6)	APH1B(3.47), CASP7(-5.93), COX7B(4.34), NDUFA9(4.55), NDUFS6(4.71), SDHD(3.84)

Pathway and gene ontology analysis shows that hsa-miR-660-5p over expression after 24 hours, might influence (both by up regulating and down regulating) genes involved in metabolism, viral carcinogenesis, Huntington disease, PI3K-Akt pathway, regulation of transcription, regulation of cell death and regulation of apoptosis (supplementary table S3-8). hsa-miR-660-5p caused many zinc finger proteins upregulated after 24 hours, and GO analysis indicate the positive effect miR-660-5p has on transcriptional regulation of CHQ5B myoblast cells.

3.3.3d) hsa-miR-92a-3p analysis

Like hsa-miR-660-5p, has-miR-92a-3p also showed very little near negligible effect on transcriptome after 12 hours as only 14 were found affected but after 24 hours of miRNA over expression 698 genes showed differential expression. Genes found most affected after 12 and 24 hours of miRNA over expression have been listed in table S3-4.

For genes differentially regulated after 24 hours of miR-92a-3p over expression, KEGG pathways with ≥ 8 differentially expressed genes have been listed in table 3-8. The 7 pathways with at least 6 differentially expressed genes are listed along with their log FC values.

24 HOURS	
KEGG Pathways	Differentially expressed genes (with their log FC)
hsa01100 Metabolic pathways (27)	AKR1B10(-4.81), ALG10(5.09), ALG3(-4.43), ALG9(6.78), ATP6V1E2(5.39), ATP6V1G1(-4.25), BPGM(4.20), COX6A1(-4.70), CTH(6.00), CYP2R1(5.66), FUK(-5.99), GALNT12(5.87), GCLC(5.92), ISYNA1(-5.52), MCEE(-5.47), NDUFAB1(-5.52), NDUFS6(4.90), PAFAH2(5.87), PIGL(-5.79), PIGV(4.34), POLR3F(4.78), POLR3K(5.35), ST3GAL6(5.46), TGDS(-4.81), TPK1(-5.54), UPRT(5.75), ZNRD1(6.27)

hsa05203 Viral carcinogenesis (17)	BAK1(-5.65), C3(8.55), CCNE1(5.35), CDKN2A(5.75), GTF2B(3.84), HIST1H2BC(6.04), HIST1H2BJ(5.09), HIST1H4E(4.72), HLA-F(4.45), HLA-G(5.87), IRF9(6.07), KAT2B(5.46), LYN(-5.40), POLB(-5.56), RBL1(-5.69), REL(6.20)
hsa04060 Cytokine-cytokine receptor interaction (9)	CSF3(7.14), CXCL1(3.71), CXCL3(5.23), CXCL5(5.44), CXCL6(4.54), IL10RB(-5.03), IL15(5.84), TNFSF12(5.57), TNFSF4(-5.38)
hsa04510 Focal adhesion (9)	BCL2(6.20), BIRC3(5.18), CAV3(-5.52), COL11A2(5.09), IGF1(5.72), ITGA8(-6.13), LAMB3(3.74), MYLK2(-6.03), PAK3(5.35)
hsa05166 HTLV-I infection (8)	ADCY4(5.10), CDKN2A(5.75), HLA-F(4.45), HLA-G(5.87), IL15(5.84), KAT2B(5.46), POLB(-5.56), SLC25A4(6.00)
hsa04062 Chemokine signaling pathway (8)	ADCY4(5.10), CXCL1(3.71), CXCL3(5.23), CXCL5(5.44), CXCL6(4.54), GRK5(4.31), ITK(-5.50), LYN(-5.40)

Table3-8: KEGG pathways for genes found altered after 24 hours of hsa-miR-92a-3p over expression. *The table shows the pathway names along with the genes involved. The LHS column gives the pathway names with the number of genes involved mentioned with in the brackets. In the RHS column, in bracket the log_2 Fold change values of corresponding genes are shown. Values in negative show down regulation and vice versa.*

Pathway and gene ontology analysis shows that hsa-miR-92a-3p over expression after 24 hours, might influence (both by up regulating and down regulating) genes involved in metabolic pathways, viral carcinogenesis, cytokine-cytokine receptor interaction, focal adhesion, regulation of transcription, mitochondrial transport, protein targeting to mitochondrion and mitochondrion organization (supplementary table S3-9.).

3.3.3e) hsa-miR-206 analysis (Positive control)

Compared to rest of the miRNAs, hsa-miR-206 showed highest genes having changed their expression after 12 hours, with 239 genes found influenced. After 24 hours of miRNA over expression 563 genes showed altered expression. Genes most prominently influenced are shown in tableS3-5.

For genes differentially regulated after 24 hours of miR-206 over expression, KEGG pathways with ≥ 8 differentially expressed genes have been listed in table 3-9. The 7

pathways with at least 6 differentially expressed genes are listed along with their log FC values.

24 HOURS	
KEGG Pathways	**Differentially expressed genes (with their log FC)**
hsa01100 Metabolic pathways (21)	ALG9(5.04), AOC3(-5.35), CDA(4.99), CKM(5.48), COX10(4.90),CYP2R1(2.50), DCXR (5.96), GALNT3 (5.55), GCLC(5.82), GK(5.19), IDI1(3.69), NDUFA9(5.61), PAFAH2(5.76), PIGH(-4.08), POLR3F(5.51), POLR3K(6.04), PPT2(4.58), PYGM(4.99), SMPD2(5.11), UPRT(6.28), ZNRD1(6.38)
hsa05203 Viral carcinogenesis (8)	C3(7.09), CASP8(-3.52), CCNE1(5.19), CCNE2(6.28), CDKN2A(6.09), KAT2B(5.19), LYN(-4.38),REL(5.04)
hsa05200 Pathways in cancer (8)	CASP8(-3.52), CCNE1(5.19), CCNE2(6.28), CDKN2A(6.09), FGF16,PPARG,TGFB2(3.72), WNT3(5.28)
hsa05166 HTLV-I infection (7)	ADCY4(4.02), CDKN2A(6.09), CHEK2(-4.80), KAT2B(5.19), SLC25A4(5.04), TGFB2(3.72), WNT3(5.28)
hsa04110 Cell cycle (7)	CCNE1(5.19), CCNE2(6.28), CDC14A(6.57), CDK7(5.02), CDKN2A(6.09), CHEK2(-4.80), TGFB2(3.72)
hsa04151 PI3K-Akt signaling pathway (7)	CCNE1(5.19), CCNE2(6.28), CSF3(5.04), FGF16(-4.08), PPP2R2B(6.04), SGK3(4.75), TLR4(4.99)
hsa04020 Calcium signaling pathway (7)	ADCY4(4.02), CACNA1H(5.04), CHRM3(5.48), PTGER3(5.19), SLC25A4(5.04), TBXA2R(5.48), TNNC2(5.28)
12 HOURS	
KEGG Pathways	**Differentially expressed genes (with their log FC)**
hsa01100 Metabolic pathways (13)	ALG9(5.45), CBR3(6.00), COX7B(4.65), GALNT3(5.58), GCLC(6.21), IDI1(4.40), NDUFS6(5.04), PAFAH2(6.25), POLR3F(5.07), PTGS2(3.98), QPRT(-5.35), TPK1(-5.65), ZNRD1(5.42)
hsa04151 PI3K-Akt signaling pathway (5)	CCNE1(5.45), CCNE2(5.72), COL11A2(5.61), JAK3(-5.22), TLR4(6.07)
hsa05203 Viral carcinogenesis (5)	C3(6.27), CCNE1(5.45), CCNE2(5.72), HIST1H2BC(5.42), JAK3(-5.22)

Table3-9: KEGG pathways for genes found altered after 12 and 24 hours of hsa-miR-206 over expression. *The table shows the pathway names along with the genes involved. The LHS column gives the pathway names with the number of genes involved mentioned with in the brackets. In the RHS column, in bracket the log$_2$ Fold change values of corresponding genes are shown. Values in negative show down regulation and vice versa.*

GO analysis for genes turned up after 12 hours showed that biological process terms of positive regulation of smooth muscle contraction, muscle contraction and BP terms of transcription, transcription regulation, muscle contraction were enriched for genes turned up after 24 hours (supplementary table S 3-10).

hsa-miR-206 has a very well documented role in muscle developmental biology. With our transient over expression studies, we found out the repertoire of genes which got effected by the over expression of this miRNA for 12 and 24 hours. Genes which were over expressed at 12 hours, showed enrichment for biological processes specific for muscle development like, positive regulation of muscle contraction, positive regulation of smooth muscle contraction, electron transport chain. This result stands in parallel to the well-established involvement of miR-206 in muscle development, as the over expression of miRNA enhanced muscle developmental processes.

3.3.4 Common effects exerted by all miRNAs

We observed that number of genes with altered expression after 12 hours of transfection were considerably low than those affected after 24 hours for each miRNA, as shown in the bar graph of figure 3-3.

Genes down regulated after 12 hours of each miRNA over expression were mostly snoRNAs, as shown in the pie chart (figure 3-4 A). A greater percentage of over expressed genes were protein coding, others were either non protein coding genes or genes coding for uncharacterized proteins. The bar plot (figure 3-4 D) shows the %age of protein coding genes among all over expressed genes.

We were interested to find if there are any genes commonly perturbed by two or more miRNA of all the five miRNAs we have investigated. We found that almost ¼ th of the total genes upregulated after 24 hours of each miRNA over expression, were commonly affected by either all five miRNA or by any two miRNAs (figure3-4 B). The genes which are affected by only any one miRNA largely comprised of non-coding genes. The observation remained almost similar for genes down regulated after 24 hours (figure3-4 C). Since very limited number of genes was differentially affected after 12 hours, we focused on results of 24 hour transfection to find out common affect exerted by all the miRNAs together.

We also looked for genes which remained either supressed or over expressed through 12h and 24h. We found limited number of genes remained common between both time points for each miRNA.
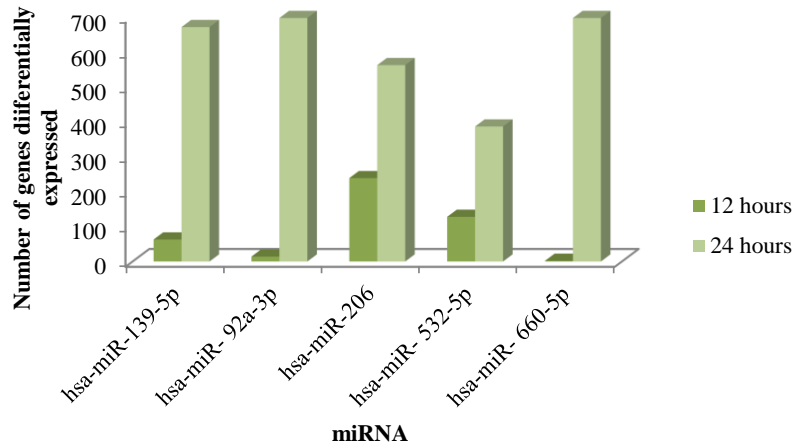

Figure3-3: Bar plot comparing the effect on gene expression after 12h and 24h of miRNA over expression. *X-axis gives the number of genes altered and y-axis gives the name of miRNA.*
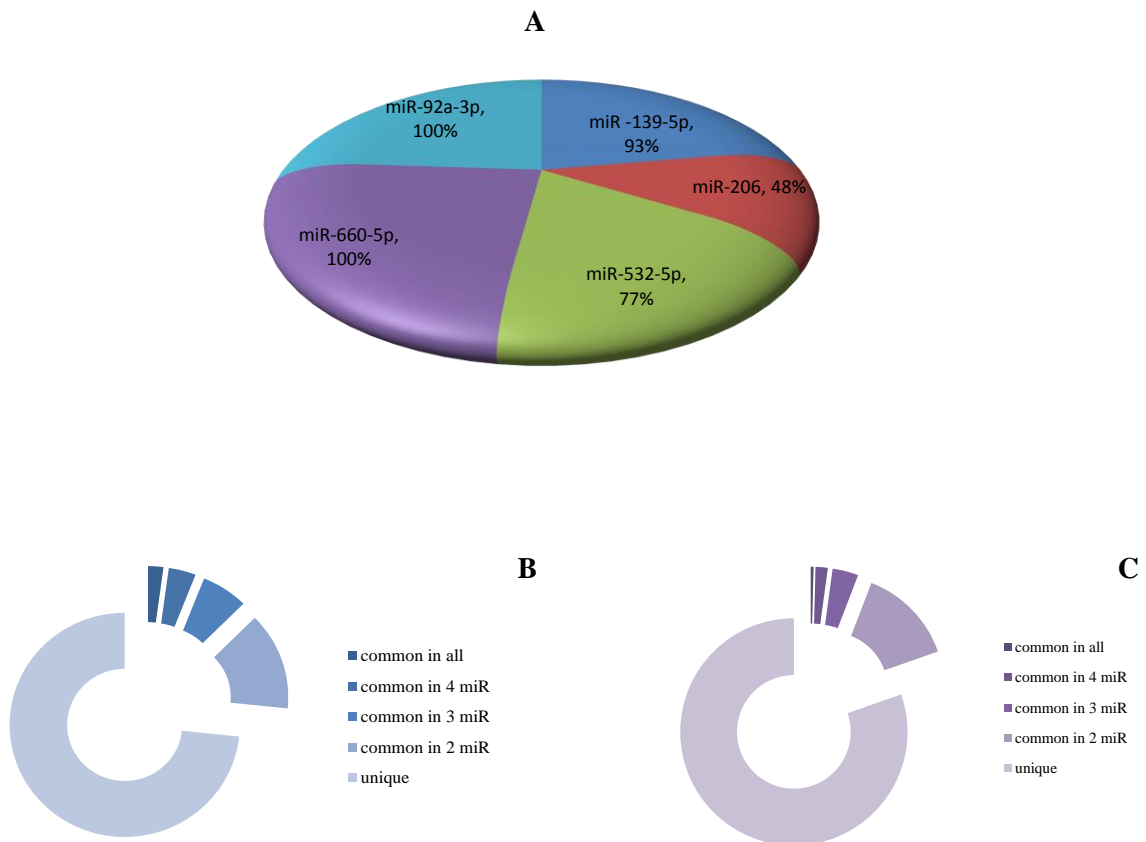
D
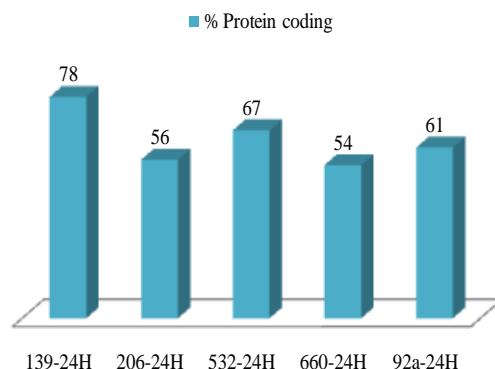
Figure3-4: (A) Pie chart showing percentage of snoRNA amongst total genes down regulated after 12 hours of miRNA over expression. (B) Doughnut plot representing proportion of upregulated genes, commonly perturbed by more than one miRNA. (C) Doughnut plot representing proportion of downregulated genes, commonly perturbed by more than one miRNA.(D) Percentage of protein coding genes amongst up regulated after 24h of miRNA over expression. 139-24H = hsa-miR-139-5p transfection effect after 24h, 206-24H = hsa-miR-206 transfection effect after 24h, 532-24H = hsa-miR-532-5p transfection effect after 24h, 660-24H = hsa-miR-660-5p transfection effect after 12h/24h, 92a-24H = hsa-miR-92a-3p transfection effect after 24h.

3.3.5 Comparison of miR down regulated genes with Ago2 enriched genes in myoblasts

In the previous chapter, I have described the genes enriched with Ago2 pull down. Here I have compared the genes down regulated by each miRNA after 24 hours with the genes enriched with Ago2 in myoblasts. We found some genes common and they are listed in table 3-10.

miRNA	Genes down regulated	Protein encoded and function
hsa-miR-139-5p	GYG2	Encodes the muscle glycogenin. Glycogenin is a self-glucosylating protein involved in the initiation reactions of glycogen biosynthesis.
	MYH7	Myosins are a large family of motor proteins that share the common features of ATP hydrolysis, actin binding and potential for kinetic energy transduction.
	COBL	Plays an important role in the reorganization of the actin cytoskeleton.
	CDK18	May play a role in signal transduction cascades in terminally differentiated cells
	HRC	May play a role in the regulation of calcium sequestration or release in the SR of skeletal and cardiac muscle
	MYBPH	Binds to myosin; probably involved in interaction with thick myofilaments in the A-band

	SHF	Adapter protein which may play a role in the regulation of apoptosis in response to PDGF
	HES6	This gene encodes a member of a subfamily of basic helix-loop-helix transcription repressors. Members of this gene family regulate cell differentiation in numerous cell types.
	AQP3	Acts as a glycerol transporter in skin
	VGLL2	May act as a specific coactivator for the mammalian TEFs. May play a role in the development of skeletal muscles
	MSC	Transcription repressor capable of inhibiting the transactivation capability of TCF3/E47. May play a role in regulating antigen-dependent B-cell differentiation
	MYLPF	myosin light chain, phosphorylatable, fast skeletal muscle
	CAV3	regulate voltage-gated potassium channels. Plays a role in the sarcolemma repair mechanism of both skeletal muscle and cardiomyocytes that permits rapid resealing of membranes disrupted by mechanical stress
hsa-miR-532-5p	COBL	Plays an important role in the reorganization of the actin cytoskeleton.
	GMPR	Catalyzes the irreversible NADPH-dependent deamination of GMP to IMP. It functions in the conversion of nucleobase, nucleoside and nucleotide derivatives of G to A nucleotides, and in maintaining the intracellular balance of A and G nucleotides
	LMOD3	GO annotations related to this gene include tropomyosin binding. An important paralog of this gene is TMOD1.
hsa-miR-660-5p	LMOD3	GO annotations related to this gene include tropomyosin binding. An important paralog of this gene is TMOD1.
	ARPP21	Isoform 2 may act as a competitive inhibitor of calmodulin-dependent enzymes such as calcineurin in neurons
hsa-miR-92a-3p	COBL	Plays an important role in the reorganization of the actin cytoskeleton.
	CDK18	May play a role in signal transduction cascades in terminally differentiated cells
	SHF	Adapter protein which may play a role in the regulation of apoptosis in response to PDGF
	VGLL2	May act as a specific coactivator for the mammalian TEFs. May play a role in the development of skeletal muscles.
	UCP2	UCP are mitochondrial transporter proteins that create proton leaks across the inner mitochondrial membrane, thus uncoupling oxidative phosphorylation from ATP synthesis. As a result, energy is dissipated in the form of heat
	CAV3	regulate voltage-gated potassium channels. Plays a role in the sarcolemma repair mechanism of both skeletal muscle and cardiomyocytes that permits rapid resealing of membranes disrupted by mechanical stress
hsa-miR-206	HRC	May play a role in the regulation of calcium sequestration or release in the SR of skeletal and cardiac muscle
	SHF	Adapter protein which may play a role in the regulation of apoptosis in response to PDGF
	AQP3	Acts as a glycerol transporter in skin

Table 3-10: Genes down regulated by miRNA and enriched in Ago2 pull down. *The gene names and the function of proteins encoded by them (taken from Gene cards) have been shown in corresponding columns.*

With this analysis we found that hsa-miR-139-5p turns down genes like MYH7, MYBPH, MYLPF, CAV3 and CDK 18 which are upregulated on myotube formation as they have muscle specific roles. hsa-miR-532-5p caused down regulation of LMOD3 gene which binds to tropomyosin. hsa-miR-660-5p also inhibits LMOD3 gene. hsa-miR-92a-3p is found to bring under expression of COBL, CDK18, SHF, VGLL2, UCP2 and CAV3 genes all of which have muscle related functions. hsa-miR-206 inhibited or down regulated HRC genes.

These results indicate that although they are turned up in myotubes, but in myoblasts they have a role to keep down the expression of genes which have role in muscle differentiation.

3.4 Discussion:

In this chapter I have tried to elucidate, the overall impact of hsa-miR-139-5p, hsa-miR-532-5p, hsa-miR-660-5p, hsa-miR-92a-3p and hsa-miR-206 over expression on the transcriptome of human skeletal myoblast cells. We chose to study these miRNA as we found them significantly upregulated upon differentiation into myotubes. Also these remain very little studied in human skeletal muscle. Our goal was not to study the target genes specifically, but to find out the overall affect these miRNA have on skeletal muscle development. The direct effect of one miRNA on a cell might be limited, but that little impact actually is far wider at the molecular level, as the message gets conveyed from one molecule to another. So, we considered studying the total impact at molecular level which these miRNA have on the muscle cells. We chose to do the analysis of gene expression at two time points of 12h and 24h post miRNA over expression to verify if the impact of perturbation on transcriptome remains steady or changes with time.

The rationale of these experiments was based on the observation that some miRNAs are found at a much higher level after differentiation of myoblasts into myotubes. Therefore, the impact of each miRNA on gene regulation could be "dissected" by overexpressing it in myoblasts and looking of the effect on the transcriptome. One of our

presumptions was that the over expression of some miRNAs could induce differentiation of transfected myoblast cells (as these miRNA are upregulated upon muscle

differentiation). However with the transcriptomic analysis that is not what we found. On comparing the genes down regulated by each miRNA with the genes enriched with Ago2 pull down in myoblasts we found some common genes. These genes have muscle specific function. So we found that although these miRNAs are turned up upon differentiation, when they are over expressed in myoblast they do not induce differentiation, rather they keep down genes which have a role in muscle development and differentiation. We can infer that since from the surrounding (which has 20%FBS) the cells do not get the message to undergo differentiation and mere upregulation of these miRNA was not a strong trigger for the cells to differentiate. This also shows that even though they are expressed at a much lower level in myoblasts, they still have a role in maintaining the proliferative state of the cells, by keeping down genes involved in muscle differentiation.

Also, with each miRNA transfection we observed greater number of genes upregulated than genes down regulated at both the time points of 12 and 24 hours. This finding of ours goes very much in accordance with the review by Shobha Vasudevan, which talks about the post transcriptional upregulation by miRNA (10). Also, recent research is showing the role of miRNA in upregulation of genes (11,12,13). These findings make it more concrete that miRNA have role in over expression of mRNA and not only in turning down gene expression. It is difficult for us to ascertain if the genes found up regulated were directly acted upon by the miRNA or whether what we observed was an implied effect, but it is possible that there may be a direct effect of the miRNAs. Our whole transcriptome analysis showed that, with each miRNA transfection and at both the time points of 12h and 24h, higher numbers of genes were differentially turned up as compared to the total number of genes differentially down regulated. This stands contradictory to the more widely accepted concept for any miRNA, which is to down regulate genes. We however do not know if prolonged time periods of observation would have changed the picture by turning down more genes and less genes over expressed. We observed that total number of genes altered after miRNA over expression escalated at 24 hours as compared to 12 hours, also very limited genes remained either over or under expressed throughout 12 – 24 hours, indicating that miRNA takes a gradual course of action inside the cells and the effect changes with time. After 12h the overall change in protein coding gene expression was very little to discern any biological meaning from them. So, our findings mainly come from 24h time point.

With pathway analysis for genes with altered expressions, we found that metabolic pathways were majorly and consistently influenced by each miRNA over expression. Also, genes over expressed after 24 hours of miRNA transfection were mainly involved in biological process of transcription and transcriptional regulation. Genes over expressed after hsa-miR-92a showed enrichment of "response to wounding" along with transcriptional regulation. hsa-miR-206 is very well characterised for its role in muscle development (1,2,3), we used it as a positive control for our experiment. We found that genes over expressed after 12h and 24h of miR-206 over expression were enriched for biological process terms of positive regulation of muscle contraction, positive regulation of smooth muscle contraction, heart contraction, muscle contraction etc. This is a good evidence to show reliability of our experiment and its results. But the interesting part is that genes "upregulated" were enriched for muscle specific function terms after miR-206 over expression. This finding of our's indicates that miRNA causes upregulation of genes. The set of genes perturbed due to miRNA over expression was different for each miRNA, but their over-all effect remained quite similar to an appreciable extent.

Our main finding from this experiment are: (1) Each of the 5 miRNAs caused more genes turned up as compared to turned down after 12h and 24h of over expression, we however cannot comment how many of these genes were directly affected and how many came out as a secondary effect. (2) After 24hours, genes turned up by each miRNA were involved on biological process of transcriptional regulation and transcription. Except for miR-206, which showed enrichment of muscle specific processes. (3) Metabolic pathways mainly and consistently were perturbed due to each miRNA over expression. (4) With the comprehensive analysis for all the 5 miRNAs, we find that they all participate towards transcriptional regulation. But we did not see the transcriptome of myoblasts shifting towards myotubes after miRNA perturbation, except for hsa-miR-206 which is a well-documented miRNA in muscle development and we used it as a positive control for our study. (5) Comparison analysis of genes down regulated by these miRNA after 24h with Ago2 enriched genes in myoblasts showed some common genes. We found these genes were involved in muscle differentiation and development. Hence it shows that although these miRNA are less expressed in myoblasts, they still have a role to play in maintaining the cell proliferation by preventing upregulation of genes for muscle development. Mere upregulation of miRNA did not cause the cells to differentiate.

3.5 Reference

1. Anderson C., Catoe H. & Werner R. (2006): MIR-206 regulates connexin43 expression during skeletal muscle development. Nucleic Acids Research 34, 5863–71.

2. Kim H.K., Lee Y.S., Sivaprasad U. *et al.* (2006): Muscle-specific microRNA miR-206 promotes muscle differentiation. Journal of Cell Biology 174, 677–87.

3. Rosenberg M.I., Georges S.A., Asawachaicharn A. *et al.* (2006): MyoD inhibits Fstl1 and Utrn expression by inducing transcription of miR-206. The Journal of Cell Biology 175, 77–85.

4. Nielsen M, Hansen J.H, Hedegaard J, Nielsen R.O, Panitz F, Bendixen C, Thomsen B. (2010): MicroRNA identity and abundance in porcine skeletal muscles determined by deep sequencing. Animal Genetics 41(2): 159-168.

5. Dmitriev P et al. (2013): Simultaneous miRNA and mRNA transcriptome profiling of human myoblasts reveals a novel set of myogenic differentiation associated miRNAs and their targets. BMC Genomic 14:265.

6. Dmitriev P et al. (2013): Defective regulation of microRNA target genes in myoblasts from Fascioscapulohumeral dystrophy patients. J Biol Chem 288: 34989 – 35002.

7. Guo L, Zhao R.C.H and Wu Y. (2011): The role of miRNAs in self renewal and differentiation of mesenchymal stem cells. Exp Hematology 39:608 – 616.

8. Dirkx E et. al. (2013): Nfat and miR-25 cooperate to reactivate the transcription factor Hand2 in heart failure. Nat Cel Biol 15: 1282 – 1293.

9. Eisenberg Iris et.al.(2007): Distinctive pattern of miRNA expression in primary muscular disorders. PNAS 104(43): 17016 : 17021.

10. Vasudevan S. (2011): Post transcriptional upregulation by microRNAs. WIREs RNA 2011. doi: 10.1002/wrna.121.

11. Ma F, Liu X, Li D, Wang P, Li N, Lu L and Cao X. (2010): MicroRNA-466l upregulates IL-10 expression in TLR-triggered macrophages by antagonizing RNA-binding protein tristetraprolin-mediated IL-10 mRNA degradation. J Immunol. 1;184 (11):6053-9.

12. Huang V, Place RF, Partnoy V, Wang Ji, Qi Z, Jia Z, Yu A, Shuman M, Yu J And Li LC. (2011): Upregulation of Cyclin B1 by miRNA and its implications in cancer. Nucl Acids Res. doi:10.1093/nar/gkr934.

13. Vasudevan S, Tong Y, Steitz JA. (2007): Switching from Repression to Activation: MicroRNAs Can Up-Regulate Translation. Science 318(5858):1931-1934.

Chapter - 4

Whole transcriptome analysis of differentiation of human skeletal muscle cells on three-dimensional scaffold

This work was performed in collaboration with Professor Nicola Elvassore of "Dipartimento Di Ingegneria Industriale, Universita Degli Studi di Padova". The work was carried out with Susi Zatti, a then PhD student who worked with Prof. Elvassore. This work was undertaken mainly as an interest of prof. Elvassore to test the influence of 3D hydrogel matrix on human skeletal muscle cells. Our interest was to elucidate the influence 3D culture exerts on the transcriptome of skeletal muscle cells.

With the next generation sequencing facility in our lab we performed the RNA seq of muscle cells grown in 3D culture. The RNA seq result analysis was also performed by us. The cell culture and immunofluorescence of cells in 3D culture were performed in the Dipartimento Di Ingegneria Industriale. CHQ5B cells were used for this experiment.

4.1 Introduction:

Myogenesis is the formation of muscular tissue, particularly during embryonic cevelopment. Muscle fibers form from the fusion of myoblasts into multi-nucleated fibers called myotubes. In the early development of an embryo these myoblasts will proliferate if enough fibroblast growth factor (FGF) is present. Whne the FGF runs out, the myoblasts cease division and secrete fibronectin onto extracellular matrix. The second stage involves the alignment of the myoblasts into the myotubes. The third stage is the actual cell fusion itself. In this stage, calcium ions are critical for development. Fusion is mediated by a set of metalloproteinases called meltrins. Myocyte enhancer factors (MEFs) promote myogenesis.

It is known that cells as classical 2D cultures loose many of the hallmarks of their in vivo counterparts. In 2D cell culture, cells are grown on flat dishes made of polystyrene plastic that is very stiff and unnatural. 3D cell-culture models have recently garnered great attention because they often promote levels of cell differentiation and tissue organization not possible in conventional 2D culture systems. In 3D cell culture, cells attach to one another and form natural cell-to-cell attachments. The cells and the extracellular matrix that they synthesize and secrete in three dimensions is the natural material to which cells are attached. It is flexible and pliable like natural tissues. It is made of complex proteins in their native configuration and so provides important biological instructions to the cells. In this 3D cell culture environment, cells can exert forces on one another and can move and migrate as they do in vivo. The close proximity of cells in 3D also enables surface adhesion molecules and surface receptors on one cell to bind to surface adhesion molecules and surface receptors on an adjacent cell. This coupling in 3D also maximizes cell-to-cell communication and signaling that is critical for cell function. Not too surprising, the phenotype or function of cells grown in 3D is more complex and closer to the functions of native tissues than cells grown in 2D (1, 27).

Since cells grown in vitro are dramatically perturbed by their micro environment, genome wide gene expression analysis is of paramount importance, as it provides the insight of genes and pathways affected in cells as compared to their tissue of origin. We have performed the transcriptome analyses of human skeletal muscle cells grown in 2D and 3D culture models along with adult skeletal muscle tissue and identified differentially

expressed genes, to develop a better understanding of how the cells adapt to in vitro microenvironments.

4.2. Materials and Methods

4.2.1 Cell Culture

Human primary skeletal muscle cells (CHQ5B):

CHQ5B primary human myoblasts were kindly provided by Dr. V. Mouly (URA, CNRS, Paris, France). CHQ5B human myoblasts were isolated from the quadriceps of a newborn (5 days post-natal) without any sign of neuromuscular disorders and the protocols used for this work were in full agreement with the current legislation on ethical rules. These cells can achieve 55-60 divisions before reaching proliferative senescence.

Growth conditions: DMEM (Gibco - Invitrogen) supplemented with 20% Fetal Bovine Serum (Gibco, Life Technologies) and 50μg/ml gentamycin.

Differentiation conditions: DMEM supplemented with 2% horse serum (GibcoBRL) and 50μg/ml gentamycin.

Differentiation of myoblasts into myotubes has to be induced by serum withdrawal when the myoblasts are just below confluency.. Myotube formation can be observed after two days after serum withdrawal.

2D culture was carried out in 10cm petri dishes. For 3D culture of cells, at day 0 cells were initially suspended in matrigel 50% (v/v in DMEM) and injected inside the μ-channels (about $2x10^4$ cells / channel). 24 hours after injection, the differentiation medium was added and myoblasts began to fuse forming bundles of myotubes. 3D culture was performed at the DIPIC (Engineering Department, University of Padova).

4.2.2 RNA extraction

4.2.2a Poly (A)[+] RNA from CHQ5B cells in 2D culture:

Polyadenylated RNA was extracted directly from cells using QuickPrep micro mRNA purification kit (Amersham Biosciences). Cells were scraped and collected after PBS washing, snap frozen in dry ice and stored at -80°C until use. 400μl extraction buffer (buffered aqueous solution containing guanidium thiocyanate and N-lauroyl sarcosine) was added to the pelleted cells which were vortexed until homogenous suspension was achieved. This suspension was diluted with 800 μl of elution buffer (10mM Tris HCl, pH

7.5, 1mM EDTA) and mixed using the vortex. The mixture was centrifuged at 12,000g for 1 minute and the clear cellular homogenate was added to the Oligo(dT) beads (25mg/ml oligo dT cellulose suspended in buffer) pellet. Cellular homogenate and oligo (dT) beads were incubated together for 5 minutes at 70 -75 °C. This causes the denaturation of RNA and enhances the binding of the poly (A) tail of RNA with the oligo dT beads. The sample was incubated at room temperature for 30 minutes with gentle agitation. The supernatant was removed by centrifugation at 12,000g for 30 seconds. The oligo (dT) cellulose pellet was washed 4 times using high salt buffer (10mM TrisHCl, pH 7.5, 1mM EDTA, 0.5M NaCl) followed by washing with low salt buffer (10mM TrisHCl, pH 7.5, 1mM EDTA, 0.1M NaCl) for another 4 times. High salt conditions allow the annealing of the poly(A) tail to the oligo d(T). The low salt buffer removes the poly(A)-RNAs (eg. tRNA and rRNA). These washings also remove contaminating DNA, RNA proteins. The oligo dT beads were transferred to the microspin column (polypropylene mini-columns) and suspended in pre heated 100 -200 µl elution buffer (10mM Tris HCl, pH 7.5, 1mM EDTA) or preheated RNase free water (Sigma) which releases the poly(A)$^+$ RNA and the tube was centrifuged at 12,000g for 30 seconds. The eluate contained poly (A) RNA which was stored at -80°C until downstream processing.

4.2.2b Poly (A)$^+$ RNA from CHQ5B cells in 3D culture:

Cells were collected from 3D hydrogel matrix in extraction buffer of QuickPrep micro mRNA purification kit (Amersham Biosciences) after cutting open the matrix channels under sterile conditions. Poly(A) RNA was extracted was carried out using the same procedure as mentioned in section 3.2.2a.

4.2.3 RNA quantification, quality assessment and RNA seq

Same as mentioned in materials and methods section of Chapter 2. Sequencing produced millions of short reads which were aligned on the human genome using PASS program (9).

4.3 Result:

Anchorage of muscle cells to the extracellular matrix is crucial for a range of fundamental biological processes including migration, survival and differentiation. Three-dimensional

(3D) culture has been proposed to provide a more physiological *in vitro* model of muscle growth and differentiation than routine 2D cultures. We have characterized cell-matrix interactions in 3D muscle culture and analyzed their consequences on cell differentiation. Our goal was to develop a comprehensive understanding of how 3D culture modulates the muscle differentiation using NGS technology.

4.3.1 Cell Observation in 3D culture

CHQ5B cells were grown on tissue culture plastic and matrigel surface. In 3D culture, after 24 hours of differentiation medium addition myoblasts began to fuse forming bundles of myotubes, which progressively aligned generating elastic, elongated structure. In 3D culture, after the switch to the differentiation medium, cells fused to form long, multinucleated myotubes that were well aligned along the longitudinal axis of the gel (figure 4-2). This pattern contrasted with the large branched myotubes and random cell orientation usually observed in routine 2D cultures (figure 4-1).
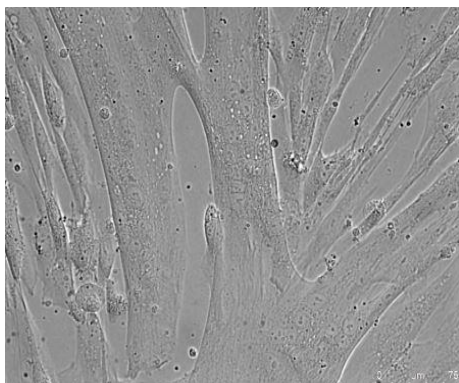

Figure 4-1: Phase contrast image of myotubes cultured and differentiated on 2D plastic surface. Cells were 10 days differentiated.
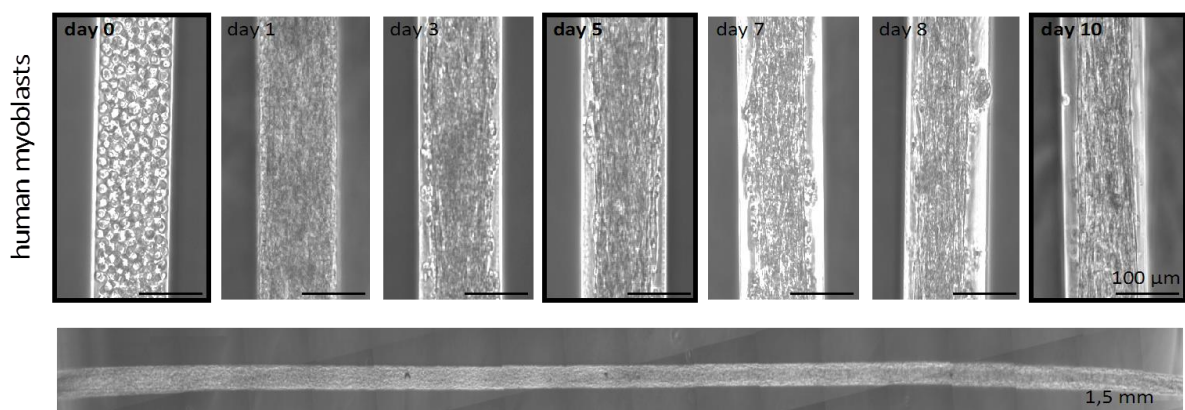
Myotube formation was assessed by immunostaining for the presence of myosin heavy chain, α-actinin and dystrophin on 3D myotubes bundles obtained 5 and 10 days after myoblasts injection in the micrometric channels and extracted from the hydrogel scaffolds (figure 4-3).
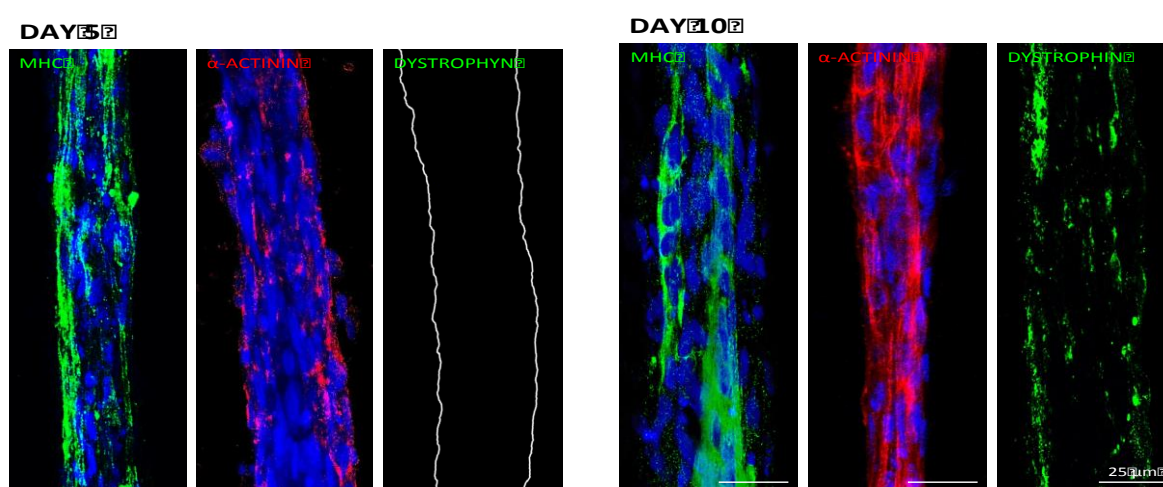

Figure 4-3: Confocal microscopic images of myosin heavy chain, α-actinin and dystrophin staining of CHQ5B cells during the process of differentiation in 3D culture. Images show the cross section through the middle of channels of matrigel for 5 and 10 days of differentiation.

4.3.2 Transcriptome analysis of 2-D, 3-D cultivated myotubes and human adult skeletal muscle tissue

We performed an extensive comparison of transcriptomes of 2D, 3D cultured myotubes and human adult skeletal muscle to address two things, firstly to find out the transcriptional variability that happen due to the change in culture conditions and secondly to address the correlation, 2D and 3D cultured myotubes share with the skeletal muscle tissue. Both cell culture conditions were compared with each other as well as with the adult skeletal muscle tissue. By comparing the transcriptomes we found out differentially expressed genes. Differentially regulated genes were identified using edgeR

(2) combined with false discovery rate (FDR) method (3). A statistical significance of p < 0.05 was used for the comparisons.

We found 1053 genes (731 up, 322 down regulated) differentially expressed in 3D myotubes when compared to 2D. 1570 genes (480 up, 1090 down regulated) and 4724 genes (2481 up, 2243 down regulated) were differentially expressed in skeletal muscle tissue when compared to 3D and 2D cultured myotubes respectively (figure 4-4). Our results showed that transcriptome of myotubes cultured under 3D conditions show greater similarity to adult skeletal muscle tissue transcriptome, when compared to 2D cultured cells, as can be seen from the smear plots shown in figure 4-5. The extent of overlap between transcriptomes of two culture conditions and the skeletal tissue is shown by the venn diagram of figure 4-6. Minimal overlap was found between 2Dvs3D and 3DvsSkeletal muscle. 248 genes were found in common between genes differentially upregulated in 3D and skeletal muscle tissue when compared to 2D cultured cells.
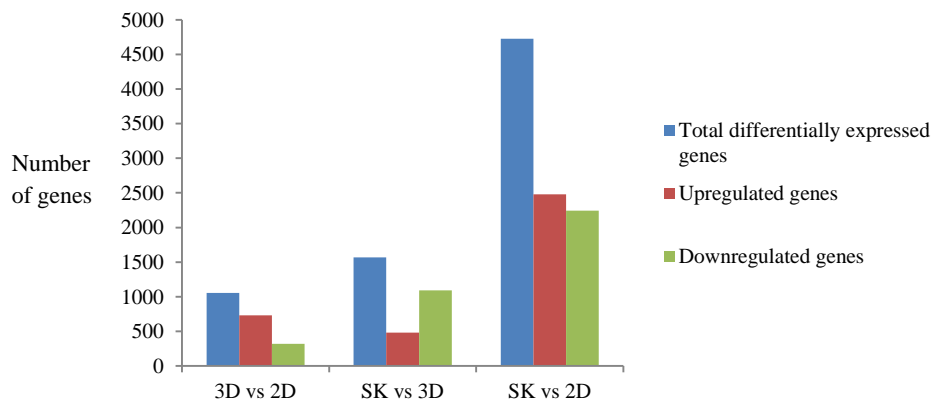

Figure 4-4: Bar plot showing differentially up and down regulated genes in 2D, 3D culture and skeletal muscle tissue. *The bar plot shows genes differentially up and down regulated. The differential expression of genes was found by comparing two samples. (3D vs 2D – 3D culture vs 2D culture, SK vs 3D – Skeletal muscle vs 3D culture, SK vs 2D – Skeletal muscle vs 2D).*

A   **B**   **C**
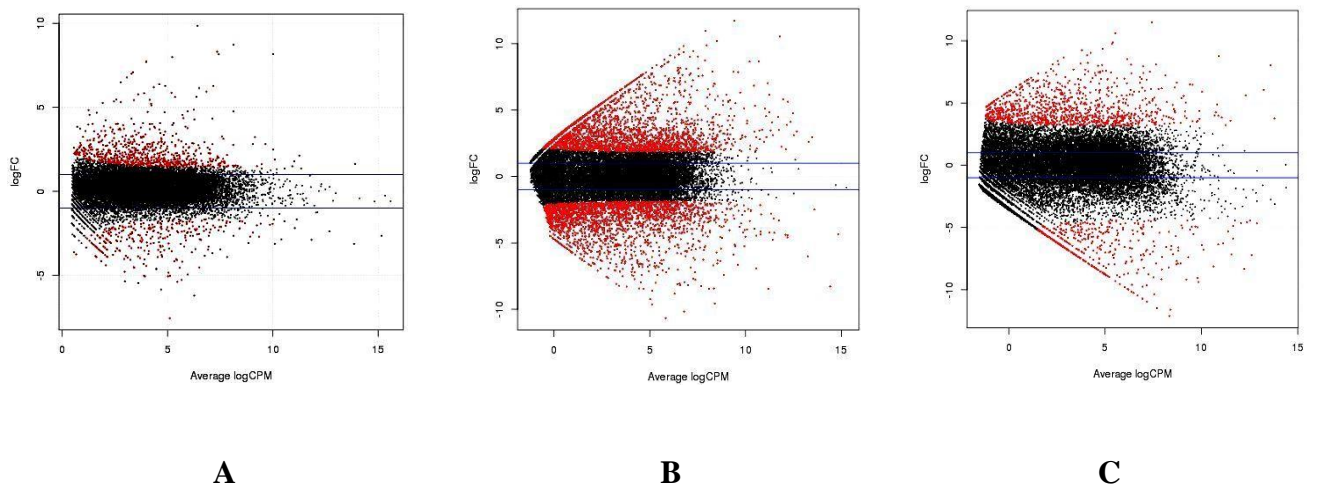
***Figure 4-5: Scatter plot showing differential gene expression in 2D, 3D cultures and skeletal muscle tissue.** (A) 2D vs 3D culture (B) 2D vs Skeletal muscle tissue (C) 3D vs Skeletal muscle tissue. The plots show that highest numbers of differentially expressed genes are found when 2D culture is compared with Skeletal muscle tissue.*
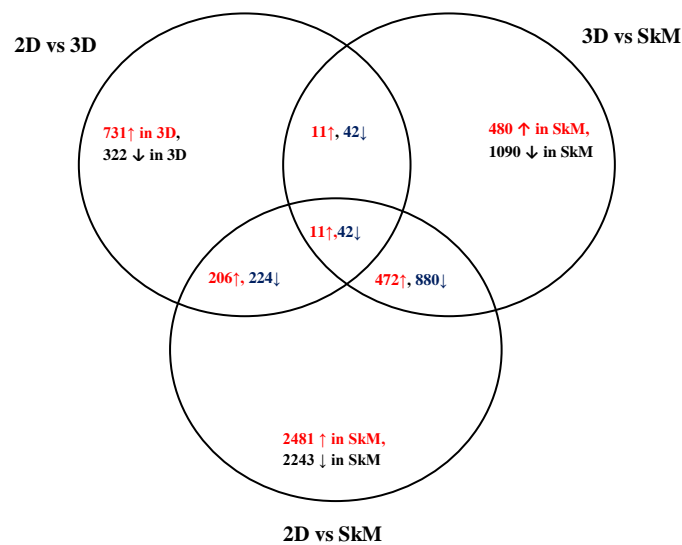
***Figure 4-6: Venn diagram shows overlap between three sample pairs (2D vs 3D, 3D vs SkM, 2D vs SkM).** Shown in red are the number of genes upregulated and shown in black are the genes down regulated. Maximum overlap is found between 3D vs SkM and 2D vs SkM. (2D – 2D cultured myotubes, 3D – 3D cultured myotubes, SkM – Skeletal muscle).*

83

4.3.3 Gene Ontology (GO) and Pathway analysis

To understand the underlying biological message, the differentially expressed genes were subjected to GO and pathway analyses. DAVID (Database for annotation, visualization and integrated discovery), a software which can cluster genes according to their functions[8] was used to analyze all the differentially expressed genes.

The most significantly enriched biological process terms found by GO analysis for genes differentially upregulated in skeletal muscle tissue compared to 3D and 2D cultured myotubes, largely remained the same. The biological process terms showing prominent enrichment (p value < 0.05) included immune, defence and inflammatory responses and response to wounding. Up regulated genes in skeletal tissue compared to 2D culture, also showed considerable enrichment of cellular respiration, respiratory ETC (electron transport chain) and oxidative reduction. All the enriched biological processes, cellular compartment and molecular function terms for genes differentially over expressed in skeletal tissue have been listed in supplementary tables (S4-1A, S4-3A). For genes down regulated in skeletal muscle tissue compared to 2D and 3D culture conditions, biological process terms of cell adhesion, biological adhesion, extracellular structure organization showed substantial enrichment (supplementary table S4-1B, S4-3B).

For genes upregulated in 3D culture compared to 2D culture, immune related response and apoptosis regulation were the most prominent biological process terms enriched. Genes down regulated in 3D culture were enriched for biological processes of cell adhesion, biological adhesion and extracellular structure organization (supplementary table S4-1A and B). Genes upregulated in 2D cultured myotubes compared to myoblasts shows enrichment for biological processes of muscle process development, striated muscle contraction and muscle cell differentiation (supplementary table S4-4A and B). Upregulated genes found common between 3D and skeletal muscle as compared to 2D culture showed enrichment for biological processes of immune, defense, inflammatory responses and response to wounding.

Pathway analysis using KEGG database showed a common enrichment for metabolic processes for genes upregulated in skeletal muscle tissue as compared to both 2D and 3D cultured cells. A lot of genes were found participating in immune related pathways (table 4-1 A and B).

Pathways	Number of genes
Metabolic pathways	28
Cytokine-cytokine receptor interaction	19
Phagosome	19
Staphylococcus aureus infection	17
Tuberculosis	16
Cell adhesion molecules (CAMs)	15
Osteoclast differentiation	15
Systemic lupus erythematosus	15
PI3K-Akt signaling pathway	15

Pathways	Number of genes
Metabolic pathways	169
Epstein-Barr virus infection	57
Cytokine-cytokine receptor interaction	51
HTLV-I infection	46
MAPK signaling pathway	46
Huntington's disease	46
PI3K-Akt signaling pathway	45
Alzheimer's disease	45
Tuberculosis	44
Phagosome	40
Osteoclast differentiation	40
Parkinson's disease	40
Pathways in cancer	38
Oxidative phosphorylation	36

A   **B**

Table 4-1: (A) KEGG Pathways for gene upregulated in skeletal muscle tissue as compared to 3D cultured myotubes. The number of genes involved in each pathway is shown in the corresponding column. (B) KEGG Pathways for gene upregulated in skeletal muscle tissue as compared to 2D cultured myotubes. . The number of genes involved in each pathway is shown in the corresponding column.

4.3.4 Transcriptome of 3-D compared with 2-D cultured myotubes

Some important differences were noted on comparison of the transcriptome of both culture conditions. With a FDR of 0.05, we found 731 genes were up regulated and 322 genes were down regulated in CHQ5B myotubes differentiated in 3D culture system. 83% of the upregulated genes and 63 % of down regulated genes were protein coding (Figure 4-7A). Genes found most significantly up regulated in 3D culture conditions were known myokines, a variety of cytokines and other immunologically relevant molecules. 5% of upregulated genes showed a fold change of > 5x.

The most prominent group of genes showing upregulation were contraction induced myokines, interleukins, cytokines and other immune related molecules under 3D culture conditions. Genes showing strongest upregulation were CFS3, IL8, CHI3L1, CXCL1 and CXCL6 (table 4-3). 8% of significantly over expressed genes were of those producing immune response molecules.

Genes of IL8 and IL24 showed highest fold increase of 8x. Other interleukin genes like IL1A, IL1B, IL24, IL32, IL33, IL6, IL8 and IL11 were also remarkably upregulated. Chemokines like (C-X-C motif) CXCL1, CXCL12, CXCL2, CXCL3, CXCL5, CXCL6

and (C-C motif) CCL2, CCL20 were also noticeably upregulated. Many other mRNAs for immunologically relevant molecules were also differentially upregulated.

Terminal differentiation of muscle cells are governed by a network of four MRFs: myogenic factor 5 (MYF5), muscle specific regulatory factor 4 (MRF4 also known as MYF6), myogenic differentiation protein (MYOD1) and myogenin (MYOG). MRFs are transcriptional factors that activate many downstream genes to initiate muscle cell differentiation. MYF5 and MYOD1 act as determination genes, whereas myogenin is essential for the terminal differentiation of committed myoblasts (10). Three of the four MEF2 (myogenic enhancer factor) proteins (MEF2A, MEF2C and MEF2D) are expressed in skeletal muscle[12]. MEF2 directly interacts with MYOD (11). SMAD7 by interacting with MYOD also, regulates differentiation of skeletal muscle cells (13). Under 3D culture condition, MYF6, MYOD1, MYOG, MEF2A, MEF2D and SMAD7 were found upregulated (Figure 4-7B).

A   B
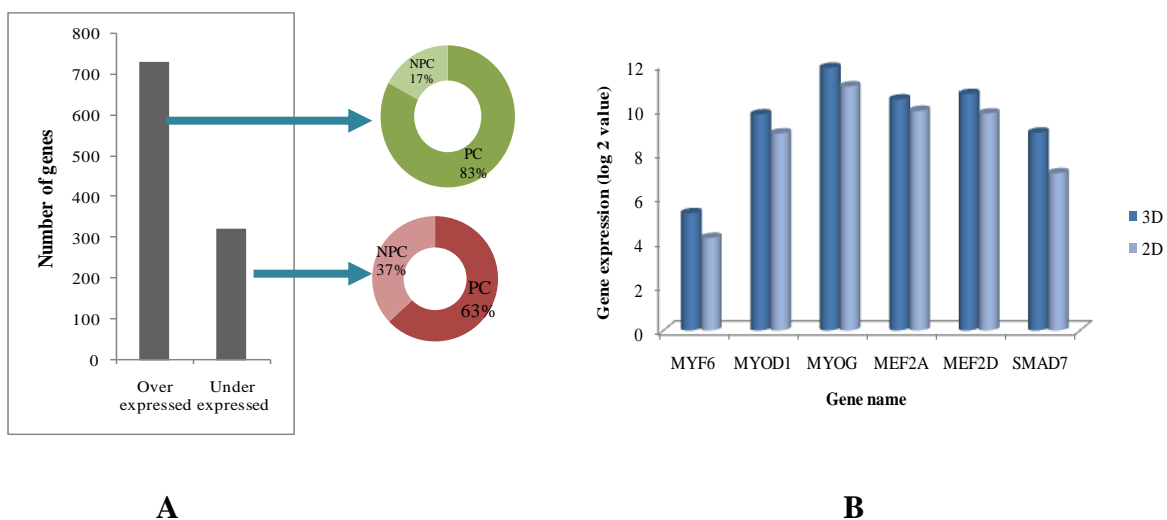
*Figure 4-7: (A) Differential gene expression profile of 3D cultured cells when compared to 2D culture. Doughnut curves show the distribution of protein coding (PC) and non-protein coding (NPC) gene for differentially up and down regulated genes. (B) Muscle specific gene expression 2D and 3D culture conditions. From the bar curve it is seen that these genes are over expressed in 3D culture as compared to 2D culture.*

Adhesion proteins like Integrin, ADAM12, CD9, CD81, M-cadherin and VCAM-1 play important role in myoblast fusion and myotube formation[14]. Integrins are transmembrane receptors which have the dual function of attachment of cells to the ECM and signal transduction from ECM to cells. Alpha and beta subunit genes of integrin were up regulated. ITGA8, ITGB3, ADAM12 and CD81 showed upregulation whereas other genes did not show any change. Collagen and laminin which play important role in formation of muscle specific extracellular matrix[15] were also found to be upregulated in 3D culture. Myoglobin showed 3x fold increase. Matrix metallopeptidases (MMP) have their essential role documented in skeletal muscle healing and extracellular matrix reorganization (16) were also found upregulated in 3D culture. Members of fibulin gene family, which are secreted glycoproteins that become incorporated into a fibrillar extracellular matrix when expressed by cultured cells, were turned up (24). Table 4-2 enlists all these genes.

Gene Symbol	Description	logFC
COL13A1	collagen, type XIII, alpha 1	2
COL22A1	collagen, type XXII, alpha 1	4
COL24A1	collagen, type XXIV, alpha 1	3
COL3A1	collagen, type III, alpha 1	2
COL4A3	collagen, type IV, alpha 3	3
COL7A1	collagen, type VII, alpha 1	3
ITGA8	integrin, alpha 8	2
ITGB3	integrin, beta 3	2
ADAM12	ADAM metallopeptidase domain 12	1
CD81	CD81 molecule	1
FMOD	fibromodulin	2

Gene Symbol	Description	logFC
FBLN1	fibulin 1	2
FBLN7	fibulin 7	3
GJB2	gap junction protein, beta 2, 26kDa	6
LAMA4	laminin, alpha 4	3
LAMB3	laminin, beta 3	2
MB	myoglobin	3
MMP1	matrix metallopeptidase 1	4
MMP11	matrix metallopeptidase 11	1
MMP2	matrix metallopeptidase 2	2
MMP3	matrix metallopeptidase 3	5
MMP9	matrix metallopeptidase 9	8

Table 4-2: Genes upregulated in 3D cultured myotubes as compared to 2D cultured myotubes. Corresponding columns give the description of each gene and their log$_2$ FC.

Gene symbol	Description	logFC	PValue
CSF3	Colony stimulating factor 3 (granulocyte)	10	1.86E-44
IL8	Interleukin 8	9	1.90E-44
CA9	Carbonic anhydrase IX	9	1.17E-28
IL24	Interleukin 24	9	1.60E-28
CXCL1	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	8	4.32E-41
PCSK1	Proprotein convertase subtilisin/kexin type 1	8	8.28E-34
CXCL6	Chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2)	8	2.84E-40
CHI3L1	Chitinase 3-like 1 (cartilage glycoprotein-39)	8	3.05E-42
IL1RN	Interleukin 1 receptor antagonist	8	2.09E-24
MMP9	Matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	8	6.04E-24
TREM1	Triggering receptor expressed on myeloid cells 1	8	1.67E-22
SLC16A6	Solute carrier family 16, member 6 (monocarboxylic acid transporter 7)	7	1.42E-20
CXCL3	Chemokine (C-X-C motif) ligand 3	7	7.25E-26
COLEC12	Collectin sub-family member 12	6	1.43E-24
PTPRN	Protein tyrosine phosphatase, receptor type, N	6	6.96E-26
TFPI2	Tissue factor pathway inhibitor 2	6	1.05E-28
STC1	Stanniocalcin 1	6	2.52E-27
PTGES	Prostaglandin E synthase	6	1.59E-21
TMEM132B	Transmembrane protein 132B	6	1.97E-24
BDKRB2	Bradykinin receptor B2	6	2.05E-26
NR4A2	Nuclear receptor subfamily 4, group A, member 2	6	1.42E-20
ANO1	Anoctamin 1, calcium activated chloride channel	6	2.35E-14
TNFRSF1B	Tumor necrosis factor receptor superfamily, member 1B	6	7.87E-17
GJB2	Gap junction protein, beta 2, 26kDa	6	3.99E-16
FAM65C	Family with sequence similarity 65, member C	6	3.14E-19
CXCL5	Chemokine (C-X-C motif) ligand 5	6	2.13E-23
PTGS2	Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	6	2.97E-24
FAIM3	Fas apoptotic inhibitory molecule 3	6	3.06E-11
CCL20	Chemokine (C-C motif) ligand 20	6	7.87E-17
BDKRB1	Bradykinin receptor B1	6	1.52E-20
PDE4B	Phosphodiesterase 4B, cAMP-specific (phosphodiesterase E4 dunce homolog, Drosophila)	5	9.33E-20
RAB27B	RAB27B, member RAS oncogene family	5	1.58E-19
FAM57B	Family with sequence similarity 57, member B	5	7.99E-12
IL1B	Interleukin 1, beta	5	1.33E-19
PENK	Proenkephalin	5	2.65E-17
MMP3	Matrix metallopeptidase 3 (stromelysin 1, progelatinase)	5	5.61E-19
LIF	Leukemia inhibitory factor (cholinergic differentiation factor)	5	1.82E-19
IL11	Interleukin 11	5	8.20E-18
OLFM2	Olfactomedin 2	5	1.37E-17
ANXA10	Annexin A10	4	2.02E-11
KCNJ15	Potassium inwardly-rectifying channel, subfamily J, member 15	4	4.84E-14
HSPA6	Heat shock 70kDa protein 6 (HSP70B\')	4	4.06E-06
PLA2G2A	Phospholipase A2, group IIA (platelets, synovial fluid)	4	1.64E-07
C1QTNF1	C1q and tumor necrosis factor related protein 1	4	1.68E-16
RPS6KA1	Ribosomal protein S6 kinase, 90kDa, polypeptide 1	4	6.30E-09
MMP1	Matrix metallopeptidase 1 (interstitial collagenase)	4	1.19E-16
AJAP1	Adherens junctions associated protein 1	4	1.85E-09
IL6	Interleukin 6 (interferon, beta 2)	4	3.58E-16
ADCY4	Adenylate cyclase 4	4	7.00E-12
RFX8	RFX family member 8, lacking RFX DNA binding domain	4	1.06E-11
WNT7B	Wingless-type MMTV integration site family, member 7B	4	1.45E-07
FER1L4	Fer-1-like 4 (C. elegans)	4	1.10E-13
FAM20A	Family with sequence similarity 20, member A	4	3.90E-14
CDCP1	CUB domain containing protein 1	4	1.16E-14
IRAK2	Interleukin-1 receptor-associated kinase 2	4	1.28E-13
EREG	Epiregulin	4	9.73E-14
ATP2A3	ATPase, Ca++ transporting, ubiquitous	4	8.16E-12
NTN1	Netrin 1	4	4.97E-10

Table 4-3: Genes showing ≥ 4 log $_2$ fold increase under 3D culture conditions as compared to 2D culture condition. *Table gives the gene name with their description, log$_2$FC and corresponding P-Value.*

4.4 Discussion

With this experiment of my PhD work, I have studied the modulation of the transcriptome of human skeletal muscle cells under 2D and 3D culture conditions. The transcriptome analysis shows that the transcriptome of 3D cultured cells shares a greater degree of similarity with that of adult skeletal muscle tissue. By comparing the transcriptomes of the two different culture conditions and adult skeletal muscle tissue we have identified all of the genes that were differentially up- and downregulated. From this comparison we found many immune related genes were prominently over-expressed under 3D culture conditions. There is a growing set of evidence supporting the role of immune specific molecules in muscle growth and regeneration (26). The elevated immunogenic response of 3D cultured myotubes was of significant interest to us.

It is known that the process of myofiber formation requires precise regulation of myoblast fusion (18, 19). Myoblast fusion is a highly complicated event during which myoblasts fuse with each other to form myotubes which in turn enlarge the myofiber. Our study demonstrates the impact of a 3D culture environment on the differentiation of human skeletal myoblast cells (CHQ5B cells). We addressed the influence of the 3D culture system by comparing the entire transcriptome of myoblast cells differentiated under 2D and 3D culture conditions. The whole transcriptomes were sequenced using NGS technology. Muscle cell differentiation is a concerted process which involves a number of events like myoblast proliferation, migration, elongation and fusion. Cells in tissue are in a 3D micro-environment having characteristic biophysical and biochemical signals. The normal function of most cells and tissues depend on interaction with neighbouring cells and the extracellular matrix. It is difficult for 2D monolayer cultures to mimic the physiological situation *in vivo*. Hence, recapitulating the function of ECM and 3D cell interactions is an important aspect of investigating molecular mechanisms of differentiation.

On comparing 2D and 3D cultures we found an upregulation of muscle differentiation markers like MYOD1, MYOG, MRF6 (Figure 4-7B) under 3D culture condition, which

is consistent with a previous report (17). Baeza-Raja and Munoz-Canoves demonstrated the positive role of NFκB during skeletal muscle differentiation and its activation which is required for IL6 production which is a promyogenic molecule (20). In C2C12 myotubes the inflammatory cytokine IL-1 isoforms, IL-1α and IL-1β have been shown to stimulate nuclear factor κB (NF κB) signalling (22). We found two NFκB members NFκB2, RELB and IL6, IL-1α, IL-1β all considerably upregulated under 3D culture conditions. Armand et al. have shown the involvement of the Calcineurin/NFAT signalling pathway in myogenin expression during skeletal muscle development. They showed the role of two NFAT family members NFATc2 and NFATc3 and their synergistic cooperation with MyoD during skeletal muscle development (21). We found NFATc2 and NFATc4 were over expressed in 3D cultured myotubes along with MRF members. In primary mouse muscle cells a large number of chemokines were found to be expressed during cell-cell fusion and their role in cell migration and positioning myocytes in correct spatial pattern necessary for cell fusion has been speculated by Griffin et. al (23). Although the in depth knowledge of the role pf chemokines and their mechanism of action during muscle development is lacking, high level of expression of chemokine ligands during muscle development is a strong indication of that chemokines play some important role in myogenesis. In accordance with the findings of Griffin et al, we also found several chemokine ligands were overexpressed in 3D culture conditions (supplementary information).

The most interesting finding of our work was the exclusive and prominent ove - expression of genes for myokines, interleukins and several immunogenic response molecules in 3D culture. Within the last decade skeletal muscle has been identified as a secretory organ (6) and the cytokines secreted by muscle cells are collectively termed as myokines (4). The role of interleukins and myokines are well documented in exercising muscles. Myokines also have their roles in regulation of cellular expression, differentiation, tissue regeneration and repair, as well as immunomodulation and embryogenesis. Positive effects of myokines are associated with their transient production and short term action (5). Pedersen and co-workers have recently shown the role played by contractile activity in regulating the expression of cytokines by skeletal muscle (7). In order to understand our findings better we compared the transcriptomes of both 2D and 3D cultures with the adult human skeletal muscle tissue transcriptome. We found that myotubes cultured under 3D conditions shared greater similarity to the transcriptome of

adult skeletal muscle tissue, as compared to 2D cultured cells (Figure 4-5). GO analysis of genes differentially over-expressed in skeletal muscle tissue when compared to both 2D and 3D culture showed a common enrichment for the biological processes of immune response, inflammatory response, defence response and response to wounding. Pathway analysis showed, metabolic pathways dominated along with many immune related pathways (Table 4-1). Our results of full transcriptome analyses of muscle cells cultured under 2D and 3D culture systems and skeletal muscle tissue, show that development of immunogenic capacity is a prominent process and an important aspect of muscle development, as has been mentioned by Pillon et al that skeletal muscle undergoes continuous repair as a result of contractile activity which requires the participation of myokines and anti-inflammatory input (25).

GO analysis showed that genes responsible for muscle oxidative metabolism like cellular respiration, electron transport chain, generation of precursor metabolites and energy, oxidative reduction were under-expressed under 2D culture conditions. Whereas, immune specific genes were over-expressed and immune related processes were enriched for genes differentially over-expressed in 3D culture and skeletal muscle.

In summary, we have thoroughly studied the transcriptomic behaviour of CHQ5B cells' differentiation under 2D and 3D culture conditions. We compared the transcriptomes of both culture conditions with adult human skeletal muscle tissue and found that cells differentiated under 3D conditions share a higher degree of similarity with the skeletal muscle tissue transcriptome. Cells differentiated in 2D culture showed enrichment for muscle specific processes like muscle development and contraction. Whereas, cells differentiated under 3D culture conditions showed elevated immunogenic response, as many cytokine, myokine, and interleukin genes were significantly over-expressed, which indicates that the immunogenic capacity that is innate to skeletal muscle for both development and regeneration is more pronounced in 3D culture conditions that 2D conditions. Skeletal muscle tissue transcriptome also showed immunogenic response as the most prominent biological process. Our study shows how 3D cultured myotubes mimic actual skeletal tissue better than 3D culture systems. Our results demonstrate that skeletal muscle cell differentiation is enhanced in a 3D culture system compared to 2D culture system.

4.5 References:

1. Yamada K.M & Cukierman E. (2007): Modeling tissue morphogenesis and cancer in 3D. Cell 130: 601 – 610.

2. Robinson M.D., McCarthy D.J. & Smyth G.K. edgeR (2010): a Bioconductor package for differential analysis of digital gene expression data. Bioinformatics 26(1): 139-140.

3. Benjamini Y. & Hochberg Y. (1995): Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc B Met 57: 289–300.

4. Pedersen, B. K. et al. (2003): Searching for the exercise factor: is IL-6 a candidate? J. Muscle Res. Cell Motil. 2: 113–119 .

5. Munoz-Canoves P et al. (2013): IL-6 myokine signaling in skeletal muscle: a double-edged sword? The FEBS Journal 17: 4131–4148.

6. Pedersen, B.K & Febbraio, M.A. (2012): Muscles, exercise and obesity: Skeletal muscle as a secretory organ. Nat. Rev. Endo. 8: 457 – 465.

7. Pedersen B.K. (2012): Muscular Interleukin-6 and Its Role as an Energy Sensor. Med Sci Sports Exerc. 44(3): 392-6.

8. Huang W, Sherman B.T. & Lempicki R.A. (2009): Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 4(1): 44- 57.

9. Campagna D et al. (2009): PASS: A program to align short sequences. Bioinformatics 25 (7): 967-968.

10. Kassar-Duchossoy L. et al. (2004): Mrf4 determines skeletal muscle identity in Myf5:MyoDdouble-mutant mice. Nature 431: 466–471.

11. Penn, B. H., Bergstrom, D. A., Dilworth, F. J., Bengal, E. & Tapscott, S. J. A (2004): MyoD-generated feed-forward circuit temporally patterns gene expression during skeletal muscle differentiation. Genes Dev. 18: 2348–2353.

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

13. Kollias HD, Perry RLS, Miyake T, Azaz A, John C and McDermott JC (2006): Smad7 promotes and enhances skeletal muscle differentiation. Mol. Cell Biol. 16: 6248 – 62.

14. Przewoz´niak M, Czaplicka I, Czerwin´ ska AM, Markowska-Zagrajek A, Moraczewski J, et al. (2013): Adhesion Proteins -An Impact on Skeletal Myoblast Differentiation. PLoS ONE 8(5): e61760.

15. Nusgens B. et al., (1986): Metabolic changes in extracellular matrix during differentiation of myoblasts of the L6 line and of a myo- non-fusing mutant. Exp Cell Res. 162: 51-62.

16. Bellayr I, Holden K, Mu X, Pan H, Li Y (2013): Matrix metalloproteinase inhibition negatively affects muscle stem cell behaviour. Int J Clin Exp Pathol. 6(2): 124 – 41.

17. Chiron S, Tomczak C, Duperray A, Lainé J, Bonne G, et. al. (2012): Complex interactions between human myoblasts and the surrounding 3D fibrin based matrix. PLoS ONE 7(4): e36173.

18. Pavlath G.K., (2010): Spatial and functional restriction of regulatory molecules during mammalian myoblast fusion. Exp Cell Res 316: 3067–3072.

19. Abmayr S.M, Pavlath G.K, (2012): Myoblast fusion: lessons from flies and mice. Development 139: 641–656.

20. Baeza-Raja B, Munoz-Canoves P. (2004): p38 MAPK-induced nuclear factor-kappaB activity is required for skeletal muscle differentiation: role of interleukin-6. Mol Biol Cell 15: 2013–2026.

21. Armand A.S, Bourajjaj M, Martinez-Martinez S, et al. (2008): Cooperative synergy between NFAT and MyoD regulates myogenin expression and myogenesis. J. Biol Chem 283: 29004-10.

22. Li W., Moylan J. S, Chambers M. A, Smith J, Reid M.B, (2009): Interlekin-1 stimulates catabolism in C2C12 myotubes. Am J Physiology 297: C706-C714.

23. Griffin C.A., Apponi L.H, Long K.K., Pavlath G.K, (2010): Chemokine expression and control of muscle cell migration during myogenesis. J Cel Science 123: 3052-3060.

24. Timpl R, Sasaki T, Kostka G, Chu ML, (2003): Fibulins: a versatile family of extracellular matrix proteins. Nat. Rev. Mol. Cell Biol. 4 (6): 479–489.

25. Pillon N.J, Bilan P.J, Fink L.N, Klip A, (2013): Cross-talk between skeletal muscle and immune cells: muscle-derived mediators and metabolic implications. Am J Physiol Endocrinol Metab. 304: E453–E465.

26. 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(5): R1173–R1187.

27. Pampaloni F, Reynaud E.G & Stelzer E.H.K. (2007): The third dimension bridges the gap between cell culture and live tissue. Nat. Rev. Mol. Cell Biol. 8, 839 – 845.

Chapter 5

Stretching stress response of differentiating human myotubes

This work was also performed in collaboration with Professor Nicola Elvassore of "Dipartimento

Di Ingegneria Industriale, Universita Degli Studi di Padova". The work was carried out with

Lia Prevedello, a PhD student working with Prof. Elvassore. We undertook this collaboration mainly as our interest to elucidate skeletal muscle response to mechanical stretch. We cultured the cells on PDMS membrane devices for stretching and with the next generation sequencing facility in our lab we performed the RNA seq of muscle cells after stretching. The RNA seq result analysis was also performed by us. PDMS membrane device preparation and mechanical stretching were performed in the Dipartimento Di Ingegneria Industriale. CHQ5B cells were used for this experiment.

5.1 Introduction:

Skeletal muscle is highly adaptable and responds to changes in loading through exercise or resistance training through a number of mechanisms resulting in increased muscle mass and changes in contractile phenotype. The remarkable ability to adjust to changes in physiological requirements involves a concerted control of transcription regulation. Stretch stimulates skeletal muscle growth (5). Our study was aimed to investigate the molecular events that happen as a part of adaptation in skeletal muscle cells to mechanical stretch. We tried to decipher the process of adaptation by studying the changes in the transcriptome using NGS technology. We tried to capture both the immediate response and delayed response to mechanical stretch by investigating the transcriptome at two time points: 3h post stretch and 40h post stretch. We compared the transcriptome of stretched cells with unstretched controls. It is known that mRNA decay rates in human vary from being as short as 2h to more than 50h (6). For example dystrophin transcript has a half-life of 15 ± 2.5h in human muscle (7). Our reason behind doing RNA seq analysis at 3h and 40h post stretch was to capture short lived mRNAs which might be turned up as a response to stretch and also we to assess the long lasting effect of stretch on skeletal muscle cells. To have more confidence in our results we performed two replicas for each time point.

Experimental plan: We cultured human primary myoblast cells (CHQ5B) on PDMS membrane devices (described in materials and methods). Differentiation was induced by serum withdrawal. 24 hours after initiation of differentiation, mechanical stretching was performed for 30 minutes continuously (stretching conditions described in materials and methods). Post stretch, cells were left under differentiating conditions for 3h and 40h. We did mRNA extractions at these time points. So, our transcriptomic analyses were carried out at 28h and 68h of myotube formation. We had controls i.e. non stretched cells differentiated for 28h and 68h. RNA seq was performed for stretched and control cells. Our experimental design is depicted in the following figure:

```
                    ┌──────────────────────┐        ┌──────────────────────────────────┐
                    │ Cells stretched      │        │ mRNA extraction after 3h (replica 1) = │
                ┌──▶│ continuous for 30    │───────▶│ 28 hours off differentiation     │
                │   │ minutes              │        └──────────────────────────────────┘
                │   └──────────────────────┘
                │   ┌──────────────────────┐        ┌──────────────────────────────────┐
                │   │ Cells stretched      │        │ mRNA extraction after 3h (replica 2) = │
                ├──▶│ continuous for 30    │───────▶│ 28 hours off differentiation     │
                │   │ minutes              │        └──────────────────────────────────┘
┌────────────┐  │   └──────────────────────┘
│ Cells 24h in│ │   ┌──────────────────────┐        ┌──────────────────────────────────┐
│ differentiation│─┤   │ Cells stretched      │        │ mRNA extraction after 40h (replica 1) = │
└────────────┘  ├──▶│ continuous for 30    │───────▶│ 68 hours of differentiation      │
                │   │ minutes              │        └──────────────────────────────────┘
                │   └──────────────────────┘
                │   ┌──────────────────────┐        ┌──────────────────────────────────┐
                │   │ Cells stretched      │        │ mRNA extraction after 40h (replica 2) = │
                └──▶│ continuous for 30    │───────▶│ 68 hours of differentiation      │
                    │ minutes              │        └──────────────────────────────────┘
                    └──────────────────────┘
```

Cells 24h in differentiation

Cells stretched continuous for 30 minutes

mRNA extraction after 3h (replica 1) = 28 hours off differentiation

mRNA extraction after 3h (replica 2) = 28 hours off differentiation

mRNA extraction after 40h (replica 1) = 68 hours of differentiation

mRNA extraction after 40h (replica 2) = 68 hours of differentiation

Cells in differentiation for a total of 28hours (control for 3h post stretch)

Cells in differentiation for a total of 68hours (control for 40h post stretch)

RNA seq

Schematic representation of the experiment

5.2 Materials and Methods:

5.2.1 Cell culture and stretch

CHQ5B cells were cultured in Dulbecco's modified Eagle's medium supplemented with 20% fetal bovine serum and 100µg/ml gentamycin in a humidified incubator containing 6% CO_2 at 37ºC. For stretching cells were then cultured for an additional 24 hours in 2% horse serum.

PDMS membrane for stretching:

In order to carry out stretch, a special elastic surface made out of PDMS (Polydimethylsiloxane), Sylgard 184 (Dow corning) in a 1:10 ratio of curing agent (cross linking agent) and base (prepolymer) was fabricated for culturing the cells. The membrane was very soft with high elasticity. The surface area of the membrane for cell culture was $3*3cm^2$. Myotubes when cultured on plastic cell culture dishes acquire a branched and spread out structure, but is not ideal to carry out a uniaxial stretch since the cells are not aligned unidirectional. To obviate this, channels of 100µm diameter were imprinted on the surface of the PDMS membrane. For cell culture, the PDMS membranes were treated with laminin. Laminin treatment procedure was as follows:

For one cell culture, the membrane was treated with 4.5µg ($0.5µg/cm^2$) of laminin solution made in PBS. Each membrane was treated with 3ml of laminin solution and then allowed to dry under laminar hood for overnight. The following day, after the PBS was evaporated the membrane was washed once with PBS to remove any leftover salt crystals. Cells were grown on laminin treated PDMS membranes for the purpose of stretching.

5×10^5 CHQ5B myoblast cells were cultured on one PDMS membrane, followed by their differentiation (figure 4-1B).

Stretching conditions:

Cultured cells were subjected to uniaxial continuous mechanical stretch for 30 minutes in strain providing instrument (ElectroForce Planar Biaxial 4 Motor TestBench, BOSE). Cells were 10% stretched at 0.2 Hz. Stretching was performed under incubation at 37ºC with 5% CO_2.

(The stretching device with CO_2 incubator mounted is shown in figure 5-1A).
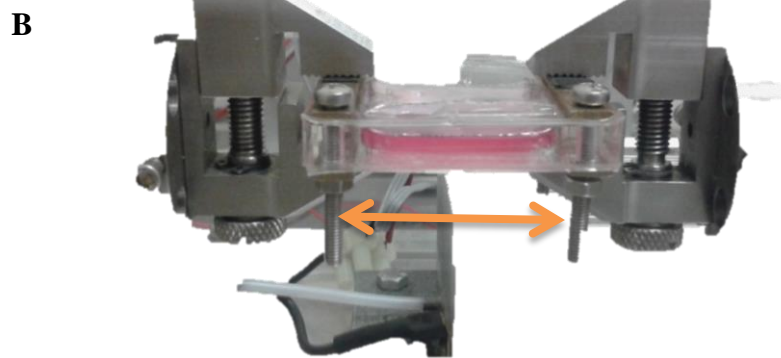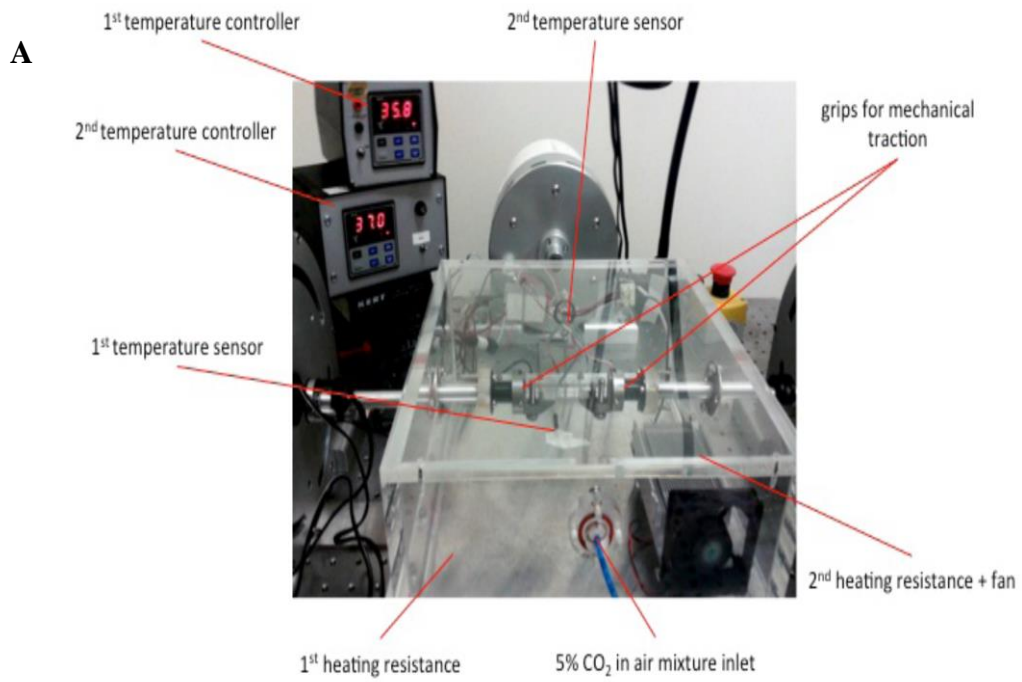
Figure 5-1: (A) ElectroForce Planar Biaxial 4 Motor TestBench instrument for stretching with incubator mounted on it. Various parts are labeled. (B) PDMS membrane with cells fitted between the clamps of the stretching instrument. Arrow shows the direction in which stretching was carried out.
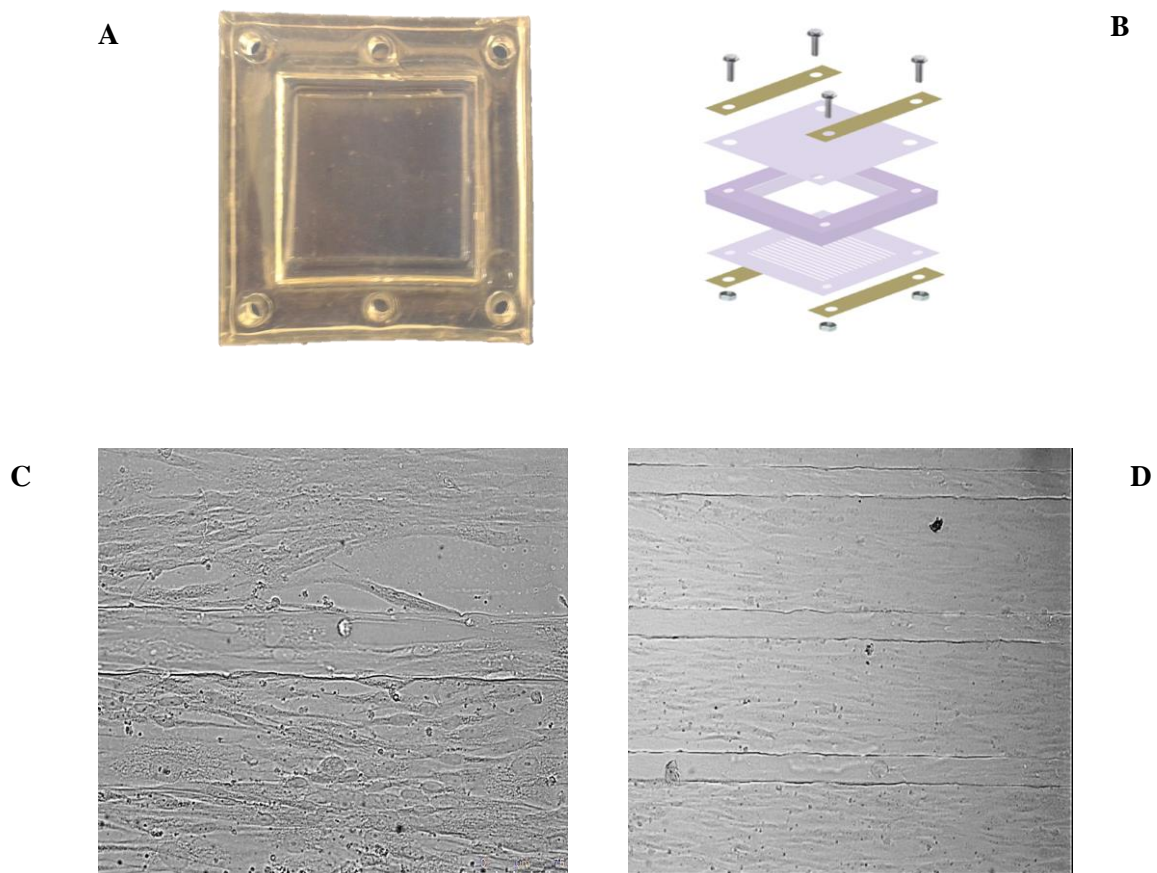
Figure 5-2: (A) PDMS membrane device used for cell culture and stretching. (B) Cartoon of assembly of the device. Microscopic image of CHQ5B cells growing on PDMS membrane with channels (C) Image taken at 20X magnification. Multinucleated myotubes can be seen. (D) Image taken at 10X magnification. Channels and unidirectionally aligned cells can be seen.

5.2.3 RNA extraction and transcriptome sequencing

We wanted to study the immediate and late response of the CHQ5B cells to mechanical stretch by transcriptomic analysis. So, mRNA was extracted using Amersham kit. mRNA were extracted at 3h and 40h post stretch. mRNA was also extracted from corresponding controls. After quantification and quality assessment, RNA samples were proceded with RNA seq protocol (RNA seq procedure mentioned in details in materials and methods section of chapter I).

5.2.4 Immunofluorescence

Primary human muscle cells (CHQ5B) were grown on PDMS membrane and differentiated for 3 days. Cells were washed 3 times with PBS then fixed with ice cold methanol for 15 minutes. Then the membrane were treated with 0.1M glycine for 5 minutes and then 0.05% Tween-20 and 5% BSA (bovine serum albumin) for 1 hour to permeabilize the cells and to block any nonspecific signals. All the washings and antibody dilutions were performed in PBS with 0.1% BSA and 0.05% Tween-20. The samples were incubated with following primary antibody for overnight: Myosin monoclonal antibody (MF20)- (1:20 dilution). The secondary antibody was anti-mouse fluorescein isothiocyanate (FITC) conjugated (Sigma, F 4018). The nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI) (0.33 µg/ml). Finally after several washes, the coverslip was mounted on the membrane using glycerol-PBS (50:50) as the mounting medium and examined under a fluorescence microscope (Leica) at both 10X and 20X magnifications.

5.3 Results

5.3.1 Differential gene expression analysis

Analysis of differentially expressed genes after 3h of stretch in comparison with the control identified 142 genes differentially expressed (88 up (> 1.4X), 54 down (< -1.4X)). We observed 10 times more genes were affected after 40 hours of stretching when compared to control. Total 1447 genes showed differential expression, with more genes being down regulated (969) and 478 genes were upregulated (figure 5-3D). When 40h post stretch was compared with 3h post stretch, highest numbers of genes (1980 genes) were found altered (1093 up, 887 down). The scatter plots of figure 5-3A,B,C show the effect of stretching on transcriptome at 3h and 40h post stretching.

Number of genes with differential expression between 27h and 65h of differentiation of cells that were not stretched was far less than cells that were stretched. (figure 5-3E).
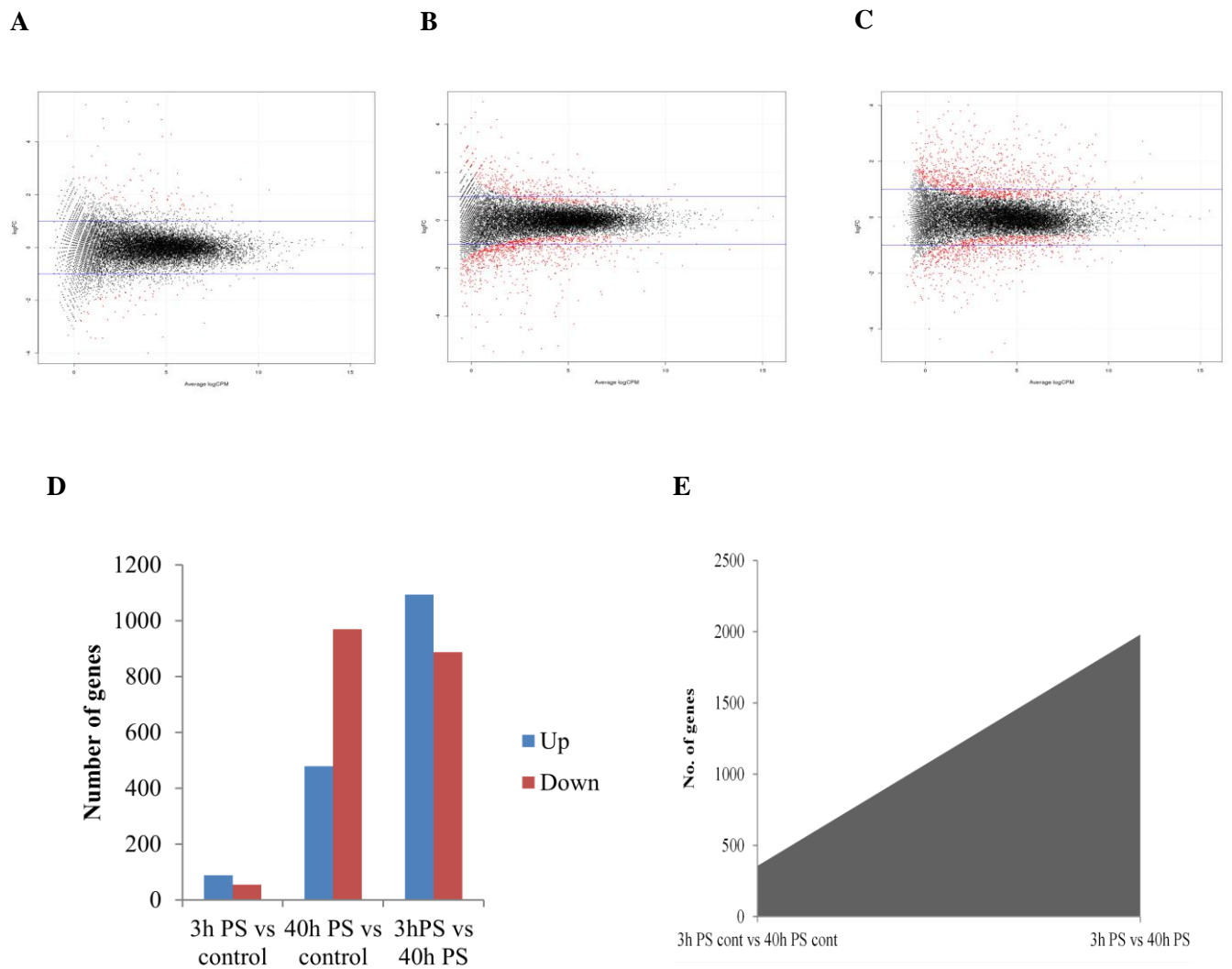
Figure 5-3: Differential gene expression analysis shown through scatter and bar plots. (A) Differential gene expression between 3h post stretch and control. (B) Differential gene expression between 40h post stretch and control. (C) Differential gene expression between 3ha nd 40h post stretch. Red dots represent differentially expressed gene. (D) Bar plot shows differentially up and down regulated genes between two samples (3h PS vs control – 3h post stretch vs control, 40h PS vs control – 40h post stretch vs control, 3hPS vs 40h PS – 3h post stretch vs 40h post stretch). (E) Plot shows genes differentially expressed between 3h and 40h post stretch control samples and 3h and 40h post stretch samples. (3h PS cont vs 40h PS cont – 3h post stretch control vs 40h post stretch control, 3hPS vs 40hPS – 3h post stretch vs 40h post stretch)

5.3.2 GO and pathway analysis

To evaluate the biological functions of the observed transcriptomic changes, we searched for enriched GO terms and also performed pathway analysis using KEGG database. For genes upregulated after 3h of stretch in comparison with control, GO terms related to

immune response, response to wounding, cell proliferation, taxis and chemotaxis were enriched. Pathway analysis was done for all the genes differentially expressed and it showed enrichment of cytokine-cytokine receptor interaction, TNF signaling pathway and chemokine signaling pathway (figure 5-4A and supplementary table – S5-2 A & B).

Genes up regulated after 40 hours of stretching vs control showed enrichment for biological process terms of L-serine metabolic process, response to nutrient levels, cellular response to insulin stimulus. However genes down regulated after 40 hours were enriched for biological processes of response to wounding, inflammatory response, locomotory behavior, defense response (supplementary table – S5-3 A & B). Pathway analysis of genes differentially affected after 40 hours showed a clear enrichment of metabolic pathways (figure 5-4B).

Differentially expressed genes 40h after stretch in comparison with 3h after stretch, showed enrichment of muscle specific biological process terms of muscle contraction, muscle system process, muscle organ development (supplementary table – S5-4 A & B). Pathway analyses showed most of the effected genes were involved in metabolic pathways, followed by pathways in cancer, PI3k-Akt pathway (fig 5-4 C).

GO and pathway analysis shows that initially after stretching immunogenic response if triggered which goes down with the passage of time. The long term effect of stretching is mostly reflected through the change in metabolic behavior of the cells.
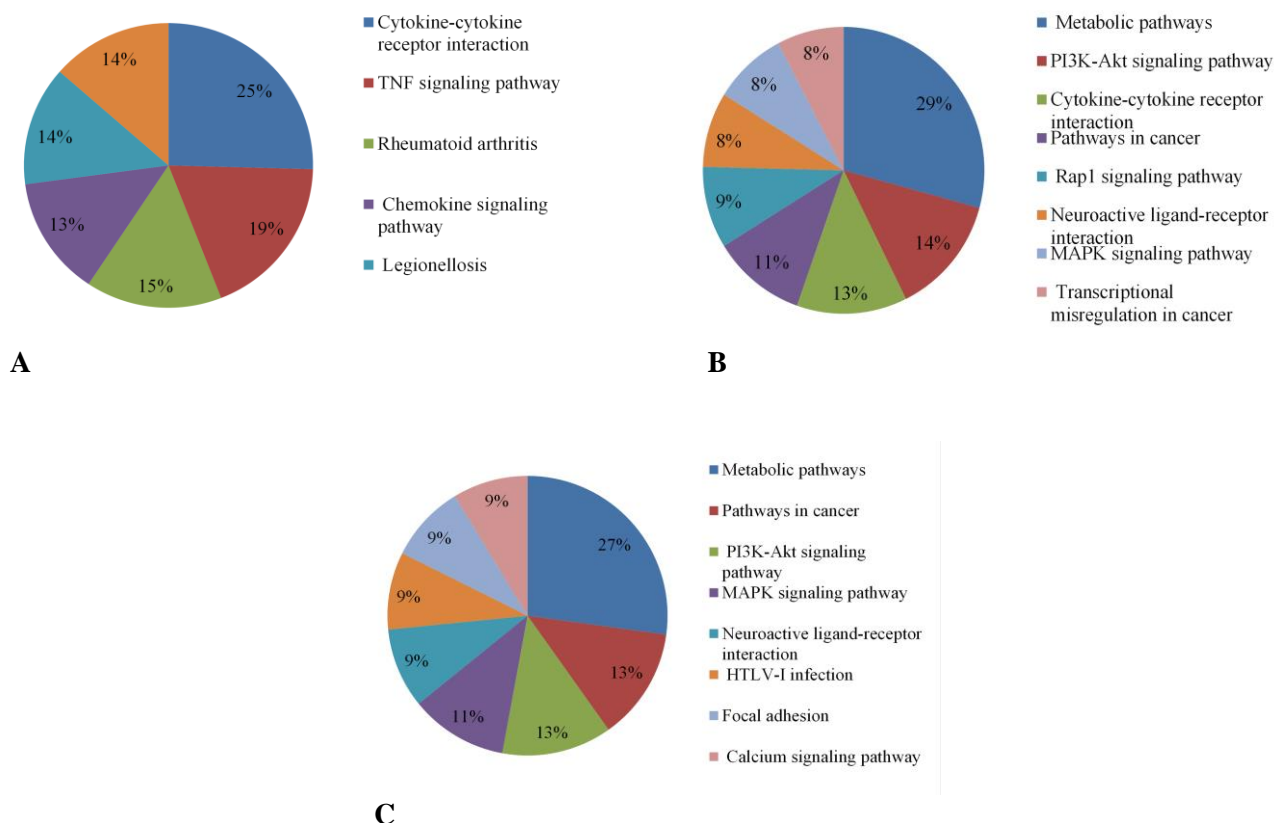
Figure 5-4: KEGG pathway analysis (A) Genes differentially affected after 3h of stretch compared to control. (B) Genes differentially affected after 40h of stretch compared to control. (C) Genes differentially affected after 40h of stretch compared to 3h post stretch. Percentage of genes belonging to each pathway is shown.

5.3.3 Myokine genes over expression as an immediate response to stretch

RNA seq analysis performed after 3 hours of stretch to capture the immediate response to stretching at transcriptome level showed many cytokine ligands and myokine genes strongly upregulated: CXCL3(5x fold up), CXCL1, CXCL6, IL1B and IL8 (>4x fold up), IL11, IL24, IL6, CXCL5 and CCL2 genes were >2x fold increased. MYH15 and MYH13 genes were >2x fold enriched. But these genes were found down regulated after 40h of stretch (table 5-1). However after 40h of stretch other set of myokine and interleukin genes were turned up which were not found effected after 3h of stretch. These genes with their log$_2$ fold change have been listed in table 5-1.

	3H		**40H**	
	log FC	P value	log FC	P value
CXCL1	4.848795943	2.89E-23	-2.9852132	1.51E-26
CXCL2	5.512684907	3.19E-19	-2.687482	1.30E-16
CXCL3	4.884179874	4.52E-12	-2.736439	1.18E-11
CXCL5	2.354070196	1.15E-06	-2.2668929	1.09E-11
CXCL6	4.19420982	1.56E-19	-2.454953	4.80E-20
IL11	2.792468051	7.03E-11	-2.7482501	3.75E-23
IL1B	4.769195706	1.97E-17	-2.0189763	3.54E-11
IL1RN	2.977031682	1.72E-10	-2.2203845	4.57E-14
IL24	3.465339337	3.09E-07	-2.9086675	6.28E-09
IL6	2.910441617	3.10E-11	-1.8053958	1.45E-11
IL8	4.28280788	9.07E-21	-4.5190076	2.09E-49
IRAK2	1.948365024	1.56E-06	--	--
IGSF22	--	--	3.10924783	4.36E-16
IGSF3	--	--	0.7216452	0.0037057
IGSF9B	--	--	2.48386769	1.93E-12
IL16	--	--	1.14243003	0.0003661
IL17B	--	--	2.65482873	3.03E-13
IL20RA	--	--	1.06653578	0.0015112
IL2RG	--	--	1.60235459	0.0002765
IL32	--	--	1.48386962	1.05E-09
IL34	--	--	1.03281938	2.91E-05
IL1R1	--	--	-0.7472284	0.0016459

Table 5-1: Myokine, interleukin and chemokine genes affected due to stretch. Shown are genes that were over expressed after 3h of stretch but decreased significantly after 40h of stretch. Also shown are those genes which were upregulated after 40h of stretch but not differentially affected after 3h. (colums left blank show no differential expression)

5.3.4 3h post stretch response analysis

Transcriptome analysis showed CXL3, CSF3 and EDNRB genes most prominently upregulated. They were > 5x up regulated. 88 genes were differentially upregulated after 3h of stretch. Genes upregulated >2x fold have been shown in supplementary table S5-5. To figure out, if this effect of gene upregulation is long lasting or just transient, we compared the genes upregulated after 3h of stretch with genes under expressed after 40 hours of stretch. The noticeable observation we made was, 72 genes out of 88 genes significantly up regulated after 3h of stretch were considerably down regulated after 40h of stretching. The list of these genes is shown in table 5-2.

Another interesting observation was that the myokine genes which were immediately turned up following stretch were also found upregulated after almost 3 days of differentiation of the cells which were not stretched. Meaning that, stretching caused upregulation of myokine genes after 3h of stretch (i.e. 27 hours of differentiation) which are turned up after around 65 hours of normal process of differentiation without any

stretching. We observed this for CSF and EDNRB3 genes also which were prominently over expressed 3h after stretch. This observation is illustrated in figure 5-5.
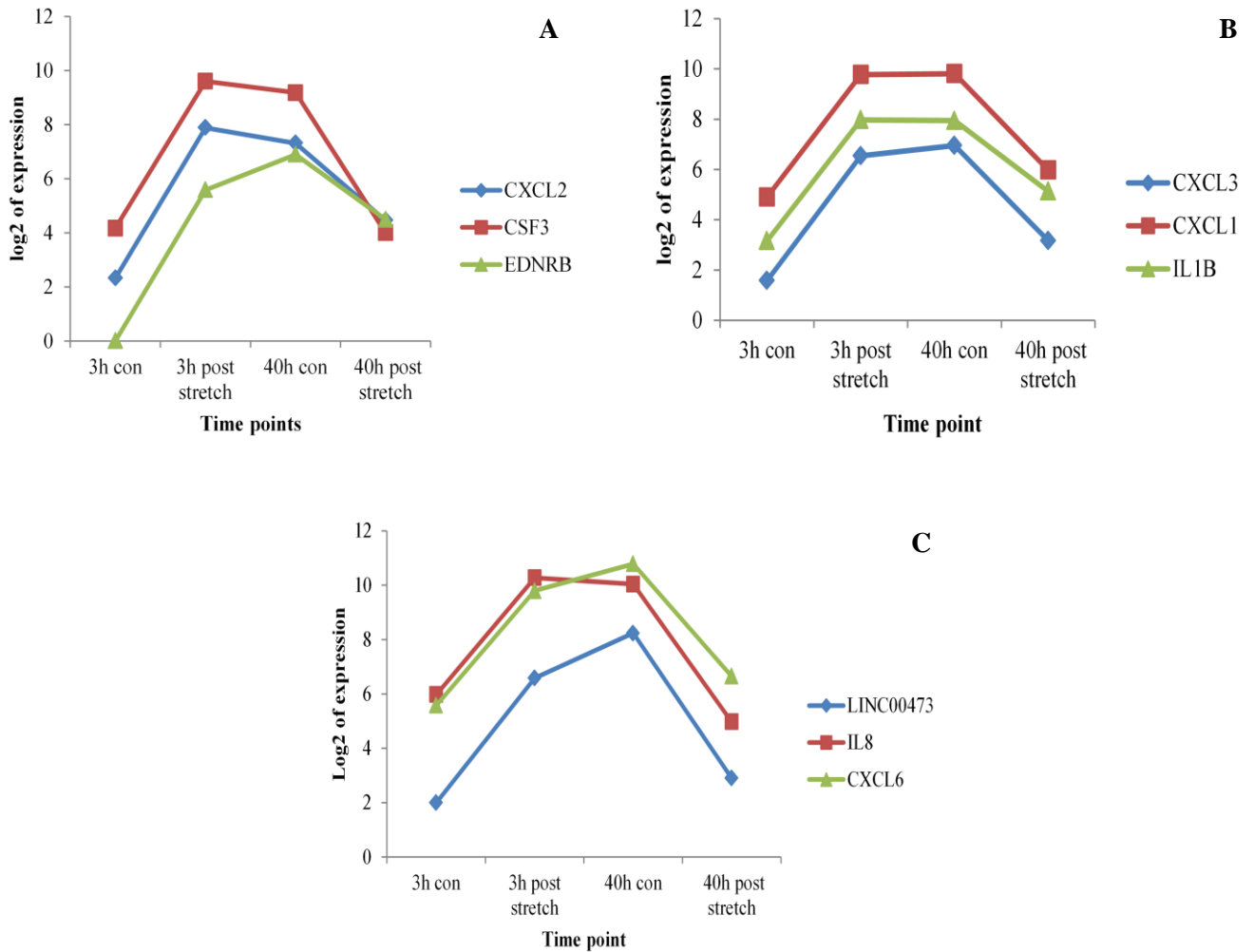

Figure 5-5: Plot showing pattern of gene expression between control for 3h post stretch (i.e 27 hours of differentiation), 3h post stretch, control for 40 h stretch (i.e. 67 hours of differentiation) and 40 h post stretch samples. *It is clear that these genes are over expressed 3h after the stretch but goes down after 40h of stretch (A) CXCL2, CSF3, EDNRB genes are shown. CXCL2 and CSF3 are highest expressed 3h post stretch (B) CXCL3, CXCL1 and IL1B are optimally expressed after 40h of stretch (C) LINC00473 (linc RNA), IL8 and CXCL6 also show similar pattern (3h con – 3h post stretch control, 40h con – 40h post stretch control.*

Genes	3h Post stretch	40h Post stretch	Genes	3h Post stretch	40h Post stret
CXCL2	5.512684907	-3.026448191	CD70	2.030848184	-1.405430!
CSF3	5.410659841	-5.350238613	MYH13	1.979979239	-1.446628
EDNRB	5.406211792	-2.57390689	G0S2	1.964905042	-1.210100!
CXCL3	4.884179874	-3.941003866	IER3	1.948697481	-1.56332(
CXCL1	4.848795943	-4.01302678	IRAK2	1.948365024	-1.019333(
IL1B	4.769195706	-2.98825512	C3	1.937700428	-2.969863!
LINC00473	4.529320167	-5.48352517	PCSK1	1.923027999	-2.693928(
IL8	4.28280788	-5.234705151	SLC39A8	1.905798494	-1.743416.
CXCL6	4.19420982	-4.314193355	TMEM217	1.901758556	-1.309088(
SLCO4A1	3.84117393	-2.579578042	CTSS	1.890578066	-1.301174.
FAM65C	3.531013342	-2.441232228	NR4A3	1.875092886	-1.900732
IL24	3.465339337	-3.942528602	MMP3	1.873753184	-4.55649
C2CD4A	3.112784614	-5.470447947	UCN2	1.821017939	-1.899682
IL1RN	2.977031682	-2.694581066	PDE7B	1.820773804	-2.032223!
IL6	2.910441617	-3.456590455	DRAM1	1.812948779	-1.417751.
EREG	2.793276793	-1.508141925	PTHLH	1.799682582	-2.923697!
IL11	2.792468051	-3.841027878	NAMPT	1.795155491	-1.26747
BDKRB1	2.712777687	-1.853560522	ASPHD2	1.793178718	-1.362912
SNTG2	2.627048854	-2.036966686	HSD11B1	1.790478209	-2.211390.
SLC5A3	2.581017714	-1.177632932	PPARGC1A	1.730702471	-1.001360
NR4A2	2.567804952	-3.875003835	MRPS6	1.644311168	-1.218310.
TNFAIP6	2.538932341	-1.57285873	AKR1C1	1.639186131	-1.928805
BMP6	2.538381129	-4.539614282	TFPI2	1.614974588	-2.477358!
CXCL5	2.354070196	-3.910932972	PDGFRL	1.583593455	-2.117027(
RNF144B	2.352708744	-1.26415099	RCAN1	1.574135016	-3.089265.
SLC7A2	2.328450642	-1.856775785	PDE4B	1.535515291	-1.925875.
PRSS35	2.259027876	-3.348193926	LIF	1.53041881	1.53041!
PTGS2	2.22176717	-3.580656456	FGF7	1.501875197	-2.679617
SOD2	2.172065892	-1.390377426	RASD1	1.48353965	-1.084192
MYH15	2.170123588	-2.03077112	GFPT2	1.472341711	-1.105996
CCL2	2.129136483	-1.861376277	WTAP	1.470211441	-0.671087!
NTN1	2.101380092	-1.21245295	SMOX	1.449895728	-1.895261
BDKRB2	2.076448304	-1.712124129	PGM2L1	1.441838463	-1.258429!
GPR183	2.057097704	-2.660423737	PADI2	1.439834302	-1.039229
STC1	2.049768217	-3.216828029	NFKBIA	1.434852932	-1.153151.
NAMPTL	2.045305703	-1.902067569	TMEM158	1.413559274	-1.298442

Table 5-2: Genes upregulated after 3h of stretch but significantly down regulated after 40h of stretch.

We wanted to find, if stretching enhances the process of muscle differentiation. In order to find that we compared genes upregulated after 3 hours of stretching with those genes which were over expressed after almost 3days of differentiation. We found 50 common

genes. The list can be found from supplementary table S5-1. The $\log_2$ fold change for all these genes were very similar, except for few genes like FGF7, MYH15, MMP3, BMP6, RCAN1, CXCL5, CXCL6, HSD11B1 and PADI2 which were more upregulated after 3 days of differentiation as compared to 3h post stretch.

5.3.5 40h post stretch response analysis

We compared the transcriptome of cells after 40 hours of stretching with the control transcriptome. 478 genes were upregulated. Genes more than 2x fold increased are shown in supplementary table S4-6. GRIN2A, ANO1 and HRK were the most prominently upregulated genes after 40h of stretch which showed >4x fold increase. Muscle specific genes showed very little upregulation. Muscle specific genes with their descriptions and fold change in bracket are mentioned below:

MYBPC1 – Myosin binding protein C, slow type (1.29), MYL1 – Myosin, light chain 1, alkali, skeletal, fast (1.41), MYLPF – Myosin light chain, phosphorylatable, fast skeletal muscle (0.71), MYO1D - Myosin ID (0.80), MYO1G – Myosin IG (1.61), MYOM2 – Myomesin (M protein) 2 (0.81), MYOZ2 – myozenin (0.90). Aldehyde dehydrogenase genes (ALDH1L2, ALDH3A1, ALDH5A1) were up regulated. Integrin genes (ITGA6, 7, ITGB6, ITBG1BP2, 3) were over expressed compared to 3h post stretched cells, however these genes were not upregulated after 65 hours of normal differentiation. Genes for collagen (COL15A1, COL4A1, COL6A6) and Insulin like growth factor (IGF2) were turned up compared to control. Voltage gated sodium channel gene (SCN7A) was 2.2x fold over expressed after 40h of stretch. Apolipoprotein genes (APOL5, 6) were >1.5x and APOM gene was 3.3x upregulated. Many myokine, cytokine genes were prominently down regulated as has been mentioned already.

C2CD4A, CSF3, SDIM1 and IL8 were most prominently down regulated genes. Many other interleukin and myokine genes were also found down regulated as has been previously mentioned.

5.3.6 Stretching propels the process of muscle cell differentiation at transcriptomic level

Aside from understanding the immediate and long term response of differentiating muscle cells to mechanical stretch, we also tried to find if stretch enhances the process of differentiation. To answer that, we compared the transcriptome of muscle cells 3h after

stretch with that of 40h after stretch. The myokine gene expression alteration has already been mentioned in previous section. As an obvious expectation we found after 40 hours of stretch the differentiating cells which further advances into differentiation, turns up many muscle specific genes. We found 24 muscle specific genes upregulated after 40h of stretch, whereas only 14 muscle specific genes were turned up in the control cells for 40h in comparison with 3h control cells. (data shown in table 5-3). Stretching caused upregulation of muscle differentiation marker myogenin (MYOG), which was not found upregulated in non-stretched cells.

Immunofluorescence showed upregulation of myosin light chain expression after 40h of stretch (figure 5-6).

Up in 40h PS vs 3h PS			Up in 40h con vs 3h con		
	log FC	P value		log FC	P value
MURC	0.950714	8.01E-05	MTUS1	2.487763	7.00E-05
MUSK	1.096964	7.86E-06	MYBPC2	3.376779	2.15E-09
MYBPC1	2.117868	4.09E-11	MYBPH	2.539496	2.80E-06
MYBPC2	3.314625	6.67E-38	MYH13	2.644173	2.73E-05
MYBPH	2.215364	1.58E-19	MYH15	3.053119	7.74E-07
MYH2	1.55618	1.63E-08	MYH2	2.220217	6.94E-05
MYH3	2.719724	1.19E-27	MYH3	3.108129	2.10E-08
MYH6	0.969583	0.000597	MYH7B	3.095218	7.11E-08
MYH7	1.265555	1.40E-07	MYH8	3.349778	3.06E-09
MYH7B	3.416475	1.26E-35	MYL2	2.188593	5.97E-05
MYH8	2.592981	1.20E-24	MYL6B	1.743602	0.00096
MYL1	2.939824	1.40E-29	MYLPF	2.330874	1.47E-05
MYL2	1.545134	1.08E-09	MYOM1	2.491142	6.65E-06
MYL3	1.874788	3.62E-06	MYOT	4.867177	2.51E-05
MYL4	1.787602	1.63E-13	MYOZ2	2.663385	3.54E-05
MYL6B	1.728081	9.00E-13			
MYLK4	1.320341	4.13E-05			
MYLPF	2.960675	9.31E-32			
MYO5B	1.192062	2.22E-05			
MYOG	0.707438	0.002681			
MYOM1	2.814417	1.02E-26			
MYOM2	2.119037	5.55E-16			
MYOM3	1.92738	3.21E-15			
MYOT	1.573054	0.002911			
MYOZ2	3.261387	1.11E-24			

Table 5-3: Muscle specific genes found upregulated after 40h of stretch vs 3h stretched cells and in control for 40h stretch vs control for 3h stretched samples.
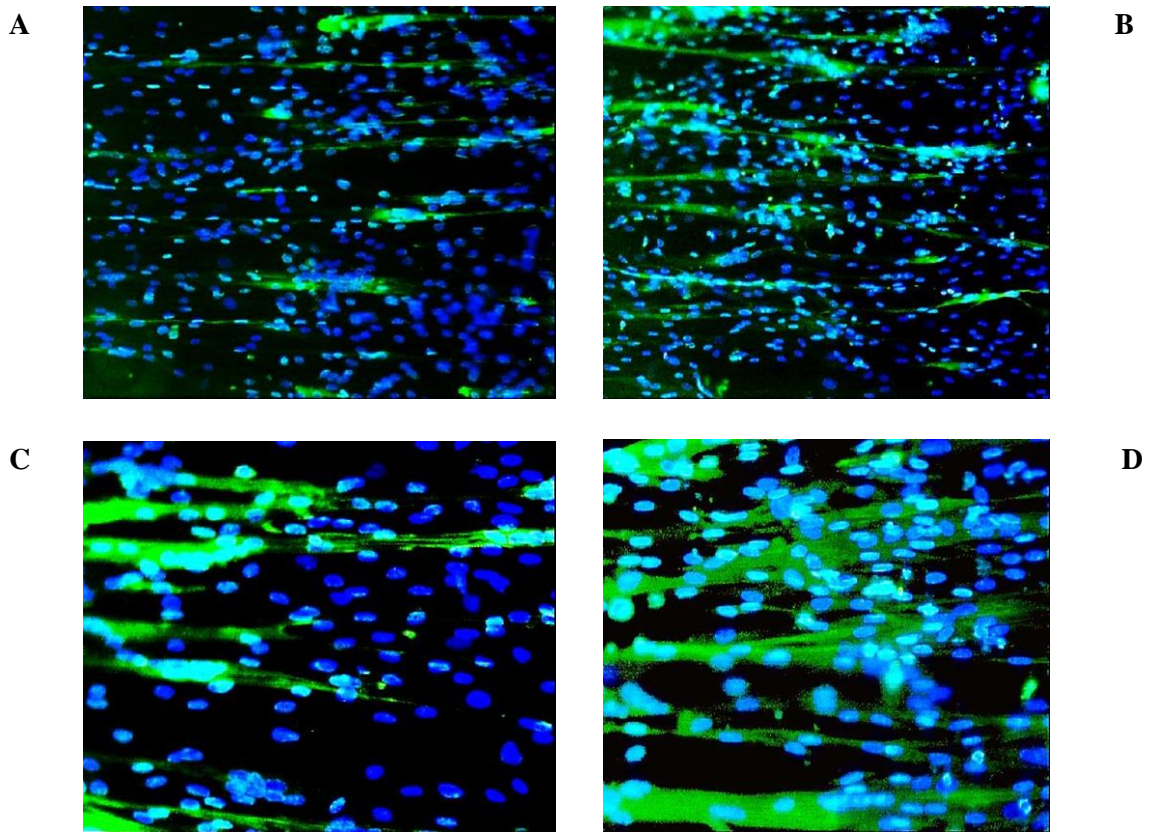
Figure 5-6: Immunofluorescence analysis of myosin light chain on (A) control cells for 40h post stretch-10X image (B) 40h post stretch- 10X image (C) Control cells for 40h post stretch – 20X (D) 40h post stretch – 20X image.

5.4 Discussion

The major aim of our study was to investigate the adaptation at transcriptomics level of growing human myotubes when stretched. We performed the study at two time points of 3h and 40h post stretch. We found far more genes affected after 40h of stretching than after 3hours. After 3hours mainly genes for myokines, chemokine liagnds and interleukin were upregulated. 79% of genes upregulated after 3h of stretch were considerably down regulated after 40h of stretch, indicating that immediate response to stretch was transient. After 40 h of stretch however we found IL16, 17B, 20RA, 2RG, 32, 34 and 1R1 genes upregulated which were not upregulated as an immediate response to stretch. IL1B and IL6 genes which were upregulated after 3h of stretch declined after 40h. IL1B is known to stimulate IL6 production in skeletal muscle cells (8) Despite their role in muscle development, their down regulation after 40h of stretch might indicate that the usual process of muscle development gets perturbed due to stretch. CSF3 gene was strongly

upregulated shortly after muscle stretch. CSF3 gene encodes for cytokine. Colony stimulating factor is known to enhance muscle proliferation and strength following skeletal muscle injury (11). CXCL1, 2, 3 and 6 genes were also upregulated as a quick response to stretch.

We observed that many more myosin family genes were upregulated after 40 hours of stretch compared to its unstretched control (table 5-3). Brunello E et al, has shown that skeletal muscle resists stretch by binding of myosin motor domain to actin(1). They have shown that a stretch, strains attached myosin motor domain which promotes the attachment of second motor domain, which in turn allows skeletal muscle to resist external stretch without increasing the force per motor. This stands as a possible explanation for our observation of more myosin family genes expressed after stretching as that assists the cell to withstand the stress caused due to stretch. Although commenting about it to be surely happening without the knowledge of what happens at the proteomic level is difficult, but with Schwanhäusser et al. study performed on mammalian tissue culture cell, showing that a good correlation between mRNA and protein levels exists, and that mRNA levels are the most important factor when predicting the protein levels (10) does provide more confidence to our interpretations. Genes for membrane spanning protein integrins and extracellular matrix collagen proteins were upregulated after 40h of stretch compared to the control. Integrins are present at Z-line and have the capacity to bind extracellular and intracellular molecules, which provides a link between basement membrane and cytoskeleton (2). They play a role in force transmission to the contractile apparatus of muscle (3). We found an upregulation of apolipoprotein genes after 40h of stretch. Apolipoproteins' have been shown to be derived from C2C12 myocytes (4), however we do not know the correlation between stretching and upregulation of apolipoprotein genes, but their over expression does indicate of some role being played by them as a response to stretch.

With pathway analysis we found that metabolic pathways were most prominently affected pathways as a long term response to stress. However after 3h of stretch cytokine-cytokine receptor pathway was the most prominent pathway activated. We also observed PI3k-Akt and MAPK signaling pathway enrichment after 40h of stretch. PI3K-Akt pathway activation is known to stimulate muscle differentiation and has a role in muscle hypertrophy (9). MAPK signaling pathway is stimulated upon exercise in human skeletal muscles (12).

In this chapter I have elucidated the changes that happen at transcriptomic level, of differentiating myotubes shortly and 40h after stretching. Shortly after stretching limited number of genes were found perturbed and also that effect was transient as after 40h of stretch the gene expression land scape changed. Immunogenic response was an immediate response to stretch as was found by gene ontology, pathway analysis and also by the upregulation of chemokine ligand, myokine genes. We have also shown that some genes which were turned up as a quick response to stretch (3h post stretch) were upregulated after 68h of normal differentiation. This indicates that stretch stress hastens some molecular aspects of muscle differentiation. The evidences we have presented here suggests, that stretch is immediately perceived as an injury by differentiating muscle cells and so its innate immune response is triggered. But at the same time, muscle growth process is enhanced. Over a period of time the immune response goes down but muscle structural protein genes are upregulated. Stretching affects cells' metabolism over the course of time. We have been able to pin point those genes which come out as an instantaneous or long standing response to stretch.

5.5 Reference

1. Brunello E, Reconditi M, Elangovan R, Linari M, Sun Y, Narayanan T, Panine P, Piazzesi G, Irving M and Lombardi V. (2007): Skeletal muscle resists stretch by rapid binding of second motor domain of myosin to actin. PNAS 104 (50): 20114 – 20119.

2. McDonald KA, Lakonishok M, Horwitz AF. (1995): Alpha v and alpha 3 integrin subunits are associated with myofibrils during myofibrillogenesis. J Cell Sci. **108**:975–983.

3. Deyne PG (2001): Application of passive stretch and its implications for muscle fibers. Phys Therapy 81(2):819 -827.

4. Zhang H. et al. (2011): Proteome of skeletal muscle lipid droplet reveals association with mitochondrion and Apolipoprotein A-I. J. Proteome Res. 10(10): 4757 – 4786.

5. Holly RG, Barnett JG, Ashmore CR, Taylor RG, Mole PA. (1980): Stretch induces growth in chicken wing muscles: a new model of stretch hypertrophy. Am J Physio 238: C62-71.

6. Yang E, Nimwegen E, Zavolan M, Rajewsky N, Schroeder M, magnasco M and Darnell JE, (2003): Decay rates of human mRNAs: Corelation with functional characteristics and sequence attributes. Genome Res. 13:1863- 1872.

7. Tennyson CN, Shi Q, Worton RG. (1996): Stability of human dystrophin transcript in muscle. Nucleic Acid Res. 24(15): 3059 – 3064.

8. Luo G, DD Hershko, BW Robb, Wary CJ, Hasselgren PO. (2003): IL-1beta stimulates production in cultured skeletal muscle cells through activation of MAP kinase signaling pathway and NF-kappa B. Am J Physiol Regul Integr Comp Physiol 285(5):R1249 – 54.

9. Glass DJ. (2010): PI3 kinase regulation of skeletal muscle hypertrophy and atrophy. Curr top Microbiol Immunol. 346: 267-78.

10. Schwanhüusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W. & Selbach, M. (2011) Global quantification of mammalian gene expression control, Nature.473: 337-342.

11. Stratos I, Rotter R, Eipel C, Mittlmeier and Vollmar B. (2007): Granulocyte-colony stimulating factor enhances muscle proliferation and strength following skeletal muscle injury in rats. J App Physiol 103(5):1857 – 1863.

12. D. Aronson, Volan MA, Dufresne SD, Zangen D, Fielding RA and Goodyear LJ (1997) : Exercise stimulates the mitogen activated protein kinase pathway in human skeletal muscle. J Clin Invest 99 (6): 1251 – 1257.

Summary and Conclusions

My PhD project was thoroughly focused on human skeletal muscle transcriptome. During the three years of my PhD I have focused on understanding various aspects of muscle development molecular level. First step towards that was having a thorough and exhaustive idea of muscle transcriptome. Transcriptomic analysis included miRNA profiling also. I have studied muscle development from the stage of myoblast where cells are in their proliferative stage, and then they differentiate into myotubes when the proliferation stops and cells fuse to become multinucleated specific muscle cells. Although there is a lot of research already conducted on the various transcritomic modulations that occur during the transition from myoblasts to myotubes, but not that everything has been unraveled yet. Next generation sequencing is a more mature and powerful technique for studying transcriptomics. NGS have been extensively for my PhD project. Using NGS we have obtained the complete picture of the transcriptome as it occurs in myoblasts and myotubes. Differential gene expression analysis was performed using edgeR software. That gave the whole bunch of genes which change from one developmental stage to another. By performing GO analysis, the functional aspects of genes differentially expressed were studied. But the skeletal muscle tissue is the functional form in human body. So to understand the difference that occurs from the cell to the tissue had to be understood. Hence adult skeletal muscle tissue transcriptome were sequenced and compared with myotubes followed by GO analysis. At this level the major finding was, upon myotube formation, muscle specific genes and muscle structural genes get strongly upregulated whereas cell cycle specific genes take a back seat, as after myotubes do not have the proliferative capacity. The transcriptome of skeletal muscle how ever showed upregulation of many interleukins, cytokines and histocompatibility complex genes which were not up regulated at the myotube stage. Skeletal muscle is known to have its own immunogenic capacity. However at this stage of my research work it was difficult to know how and when exactly the muscle tissue acquires the immunogenic capacity, because at the differentiated stage no immune specific genes were noticeable expressed.

More than thousands of genes were differentially expressed between myoblasts and myotubes. It is very unlikely that all of them are under miRNA regulation. It was of substantial importance to find out the genes which are directly under the control of miRNA regulation. Our approach towards finding out these genes was by carrying out Ago2-RNA immunoprecipitation. Argonaute 2 is a component of RISC complex, using which miRNA exerts its regulatory control on mRNA. By Ago-RIP, RNA bound to Ago2 in myoblasts and myotubes were immunoprecipiated and the RNA were then sequences using NGS. Immunoprecipitation ameliorates genes bound to Ago2. I compared this pool of immunoprecipitated genes with total genes expressed in myoblasts and myotubes. This comparison showed many genes enriched with Ago2. Since argonaute 2 has also been recently shown to have its role in stabilizing mRNA, it was not possible to accept that, enriched pool of RNA were under down regulatory action of miRNA. To fish out the most likely gene candidates associated with Ago2 and under miRNA down regulatory control, the enriched genes were compared with genes differentially down regulated in myoblasts and myotubes. This gave very interesting results. First of all out of the 1000 genes enriched, only 90 genes corroborated with genes which are turned down at either stage of muscle development, which is a very small percentage suggesting that ago2 associated miRNA down regulates a limited population of genes. Secondly I found, Genes enriched in Ago2 pull down in myoblasts were many skeletal muscle structural genes. That is a reliable finding as myoblasts do not express muscle structural genes prominently. In myotubes Ago2 enriched, transcription factor genes like Pax7, HoxA11, HoxA9 which are known to be down regulated upon muscle differentiation. There are many other genes associated with Ago2 which have not been elucidated in muscle biology. There is a need of further experiments to find the accurate interaction of miRNA with these mRNAs.

With the miRnome sequencing aside from obtaining the entire population of miRNA and their differential expression in myoblasts, myotube and skeletal muscle tissue, I found couple of miRNA which were significantly turned up in myotubes as compared to myoblasts. In literature very little information exists about these miRNA's role in muscle development. So, analyzing their functional role was of importance. I found miR-139, miR-532, miR-660 and miR-92a very interesting due to the limited knowledge available about them and the fact that they were differentially upregulated in myotubes. I used miR-206 as a positive control because of its well documented role in muscle biology. A

scramble sequence was used as negative control. A large scale experiment was designed to perform their functional analysis. Myoblast cells were transfected with each of these miRNA mimics. The transfection was carried out for 12h and 24h. The time points were to observe if the effect of miRNA intensifies with the passage of time or it changes. RNA seq of transfected cells at 12h and 24h was carried out. This simple yet exhaustive approach, gave bunch of interesting results again. The first observation was, very few genes altered after 12h as compared to 24h after each miRNA transfection, indicating the effect of miRNA intensifies with time. Another consistent observation was, more genes were upregulated and less genes were down regulated at both the time points and with each miRNA. With positive control miR-206, muscle structural genes were upregulated after 12h and 24h. This is a very strong evidence to suggest that miRNA has a role in stabilizing mRNA. We do not say all the genes found affected are being directly acted upon by miRNA, but at large these miRNA were found to cause gene upregulation than down regulation. Pathway analysis showed all of these miRNA perturbed metabolic pathways. Genes upregulated after each miRNA over expression showed involvement in transcription and transcriptional regulation. Distinct pool of genes was affected by each miRNA. Comparison of genes down regulated after 24h of miRNA over expression with Ago2 enriched mRNA in myoblasts, showed that genes involved in muscle development like transcription factors and structural genes were down regulated by these 5 miRNA in myoblasts. I did not find any did not find any enhancement of differentiation after the over expression of these 5 miRNA. Rather the results show that these miRNA although expressed at a lower level in myoblasts have their distinct function in myoblast as they keep genes involved in muscle differentiation and development, turned down. It could also be hypothesized that the pool of genes down regulated by these miRNA in myotube would be different from the genes kept down in myoblasts. Meaning that same miRNA at different level of cellular development plays different roles.

Muscle cell development under 3D culture conditions was also studied. 3D culture mimics the tissue environment and hence is known to be a comfortable niche for cells to grow. Muscle cells were allowed to differentiate for 10 days under 3D culture condition. After which RNA seq was performed. In order to understand the influence of culture conditions, we compared RNA seq results of 2D and 3D cultured myotubes. In order to see if 3D can mimic skeletal muscle tissue, we compared the RNA seq results of 3D cultured myotubes with adult human skeletal muscle tissue. This experiment showed that

when cells are cultured under 3D conditions they show immune response by upregulation of myokines and Interleukin genes. 3D cultured myotubes showed greater similarity to adult skeletal muscle compared to 2D cultured myotubes. As mentioned earlier, adult muscle tissue showed upregulation of chemokines, interleukins and major histocompatibility genes. This is an indication that 3D culture provides cells with more comfortable environment as compared to 2D culture and their transcriptome shifts more towards the skeletal muscle tissue state.

Muscle is known for its remarkable plasticity. Lastly I studied the effect of mechanical stretch on molecular events during differentiation process. Uniaxial stretch of cells early in differentiation was performed continuously for 30 minutes. The cells were either used for RNA seq analysis immediately after stretching or were left for 40h into differentiation to study the long term effect of stretch. Immmediately after stretch interleukin, chemokine and myokine genes were upregulated but that effect was transient as after 40h of stretch these genes were prominently down regulated. Stretching caused a higher population of myosin gene expression compared to non-stretched cells. So the finding from the experiment was immunogenic response enhancement as a quick response to stretch and muscle specific genes upregulation over a course of time post stretch. Immunofluorescence showed upregulation of myosin light chain after stretching. Muscle cells seemed to have developed a power of resistance and hence a quick differentiation.

During my PhD, I have tried to develop a comprehensive understanding of muscle transcriptome during various developmental stages, miRNA regulation inferred from miRNA over expression and Ago2-RIP experiments, effect of culture conditions and finally response to mechanical stretch.

Chapter – 2

Supplementary Information

Term	Gene Count	PValue	Benjamini
GO:0003012~muscle system process	56	2.05E-31	5.89E-28
GO:0006936~muscle contraction	53	1.17E-30	1.68E-27
GO:0007517~muscle organ development	54	4.09E-24	3.91E-21
GO:0006941~striated muscle contraction	23	1.89E-17	1.36E-14
GO:0007155~cell adhesion	80	8.31E-13	4.77E-10
GO:0022610~biological adhesion	80	8.65E-13	4.14E-10
GO:0060537~muscle tissue development	29	5.81E-12	2.38E-09
GO:0044057~regulation of system process	47	6.83E-12	2.45E-09
GO:0014706~striated muscle tissue development	28	1.01E-11	3.23E-09
GO:0008016~regulation of heart contraction	22	5.03E-11	1.45E-08
GO:0055001~muscle cell development	18	5.59E-10	1.46E-07
GO:0030239~myofibril assembly	12	6.80E-10	1.63E-07
GO:0055002~striated muscle cell development	17	1.44E-09	3.19E-07
GO:0031032~actomyosin structure organization	13	2.12E-09	4.36E-07
GO:0042692~muscle cell differentiation	24	1.34E-08	2.56E-06
GO:0008015~blood circulation	30	2.00E-08	3.58E-06
GO:0003013~circulatory system process	30	2.00E-08	3.58E-06
GO:0051146~striated muscle cell differentiation	20	2.86E-08	4.83E-06
GO:0006937~regulation of muscle contraction	18	3.75E-08	5.99E-06
GO:0030029~actin filament-based process	34	5.53E-08	8.35E-06
GO:0010927~cellular component assembly involved in morphogenesis	13	6.12E-08	8.79E-06
GO:0060538~skeletal muscle organ development	17	6.34E-08	8.67E-06
GO:0007519~skeletal muscle tissue development	17	6.34E-08	8.67E-06

Table S2-1A: GO terms of biological processes for gene differentially upregulated in myotubes. *The number of genes belonging to each term (gene count), enrichment (p-value) and false discovery rate (benjamini value) are shown in the corresponding columns.*

Cellular compartment Term	Gene Count	PValue	Benjamini
GO:0043292~contractile fiber	56	2.07E-39	7.99E-37
GO:0030016~myofibril	53	3.85E-38	7.42E-36
GO:0044449~contractile fiber part	51	3.41E-35	4.39E-33
GO:0030017~sarcomere	47	6.82E-34	6.58E-32
GO:0031674~I band	27	9.06E-20	6.99E-18
GO:0015629~actin cytoskeleton	56	1.34E-19	8.62E-18
GO:0031012~extracellular matrix	61	1.07E-17	5.90E-16
GO:0005578~proteinaceous extracellular matrix	58	2.36E-17	1.14E-15
GO:0030018~Z disc	21	1.33E-14	5.67E-13
GO:0044421~extracellular region part	104	6.79E-14	2.62E-12
GO:0016459~myosin complex	23	3.02E-13	1.06E-11
GO:0032982~myosin filament	13	8.65E-12	2.78E-10
GO:0016529~sarcoplasmic reticulum	17	1.79E-11	5.31E-10
GO:0016528~sarcoplasm	17	4.45E-11	1.23E-09
GO:0042383~sarcolemma	21	5.03E-11	1.29E-09
GO:0005859~muscle myosin complex	12	5.32E-10	1.28E-08
GO:0016460~myosin II complex	12	2.04E-09	4.63E-08
GO:0005865~striated muscle thin filament	10	7.44E-09	1.60E-07
GO:0031672~A band	10	1.63E-08	3.30E-07
GO:0044420~extracellular matrix part	23	7.17E-08	1.38E-06

Table S2-1B: GO terms of cellular compartments for gene differentially upregulated in myotubes. The number of genes belonging to each term (gene count), enrichment (p-value) and false discovery rate (benjamini value) are shown in the corresponding columns.

Term	Gene Count	PValue	Benjamini
GO:0000279~M phase	123	1.63E-53	5.72E-50
GO:0022403~cell cycle phase	138	3.09E-53	5.44E-50
GO:0007049~cell cycle	193	1.09E-52	1.27E-49
GO:0022402~cell cycle process	150	4.12E-44	3.63E-41
GO:0000087~M phase of mitotic cell cycle	92	8.38E-44	5.90E-41
GO:0000280~nuclear division	91	1.26E-43	7.38E-41
GO:0007067~mitosis	91	1.26E-43	7.38E-41
GO:0048285~organelle fission	91	6.45E-42	3.24E-39
GO:0000278~mitotic cell cycle	114	5.41E-40	2.38E-37
GO:0051301~cell division	97	1.40E-36	5.49E-34
GO:0006259~DNA metabolic process	128	4.00E-35	1.41E-32
GO:0006260~DNA replication	69	5.83E-29	1.87E-26
GO:0007059~chromosome segregation	41	9.35E-24	2.74E-21
GO:0006281~DNA repair	76	1.91E-22	5.16E-20
GO:0006974~response to DNA damage stimulus	85	3.77E-20	9.49E-18
GO:0051726~regulation of cell cycle	76	2.86E-18	6.71E-16
GO:0006323~DNA packaging	42	1.39E-17	3.07E-15
GO:0051276~chromosome organization	94	2.89E-17	5.98E-15

Table S2-2A: GO terms of biological processes for gene differentially down-regulated in myotubes. The number of genes belonging to each term (gene count), enrichment (p-value) and false discovery rate (benjamini value) are shown in the corresponding columns.

Term	Gene Count	PValue	Benjamini
GO:0005694~chromosome	140	1.15E-49	6.02E-47
GO:0044427~chromosomal part	126	3.09E-48	8.07E-46
GO:0000793~condensed chromosome	60	1.57E-32	2.73E-30
GO:0000775~chromosome, centromeric region	57	1.42E-30	1.86E-28
GO:0043232~intracellular non-membrane-bounded organelle	330	3.22E-26	3.37E-24
GO:0043228~non-membrane-bounded organelle	330	3.22E-26	3.37E-24
GO:0000779~condensed chromosome, centromeric region	37	1.08E-23	9.40E-22
GO:0000777~condensed chromosome kinetochore	34	1.29E-22	9.61E-21
GO:0000776~kinetochore	38	8.64E-22	5.65E-20
GO:0005819~spindle	50	5.40E-20	3.14E-18
GO:0000785~chromatin	52	3.51E-15	1.86E-13
GO:0000228~nuclear chromosome	46	5.05E-15	2.38E-13
GO:0015630~microtubule cytoskeleton	94	4.50E-14	1.96E-12
GO:0032993~protein-DNA complex	31	2.67E-13	1.08E-11
GO:0005856~cytoskeleton	175	1.18E-12	4.41E-11
GO:0005654~nucleoplasm	124	5.83E-12	2.03E-10
GO:0005657~replication fork	17	1.76E-10	5.75E-09
GO:0044454~nuclear chromosome part	33	2.32E-10	7.13E-09
GO:0031981~nuclear lumen	173	2.59E-10	7.51E-09
GO:0000786~nucleosome	23	4.00E-10	1.10E-08
GO:0044430~cytoskeletal part	125	4.58E-10	1.20E-08

Table S2-2B: GO terms of cellular compartments for gene differentially down-regulated in myotubes. The number of genes belonging to each term (gene count), enrichment (p-value) and false discovery rate (benjamini value) are shown in the corresponding columns.

Term	Gene Count	PValue	Benjamini
GO:0006955~immune response	145	2.22E-16	8.64E-13
GO:0006952~defense response	123	1.91E-12	3.72E-09
GO:0009611~response to wounding	109	6.97E-12	9.04E-09
GO:0015980~energy derivation by oxidation of organic compounds	46	8.16E-12	7.94E-09
GO:0006954~inflammatory response	77	1.20E-11	9.35E-09
GO:0003012~muscle system process	50	1.63E-11	1.05E-08
GO:0045333~cellular respiration	36	1.85E-11	1.03E-08
GO:0022904~respiratory electron transport chain	28	5.76E-11	2.80E-08
GO:0006091~generation of precursor metabolites and energy	73	9.43E-11	4.08E-08
GO:0055114~oxidation reduction	120	2.54E-10	9.88E-08
GO:0006936~muscle contraction	45	2.91E-10	1.03E-07
GO:0022900~electron transport chain	36	2.87E-09	9.31E-07
GO:0010033~response to organic substance	125	1.62E-08	4.85E-06
GO:0042775~mitochondrial ATP synthesis coupled electron transport	23	1.69E-08	4.71E-06
GO:0042773~ATP synthesis coupled electron transport	23	1.69E-08	4.71E-06
GO:0009725~response to hormone stimulus	72	2.61E-07	6.76E-05
GO:0007242~intracellular signaling cascade	187	4.74E-07	1.15E-04
GO:0002684~positive regulation of immune system process	52	5.53E-07	1.27E-04
GO:0070482~response to oxygen levels	36	9.78E-07	2.11E-04
GO:0034097~response to cytokine stimulus	25	1.07E-06	2.20E-04
GO:0048878~chemical homeostasis	90	1.18E-06	2.30E-04
GO:0009719~response to endogenous stimulus	75	1.45E-06	2.69E-04

Table S2-3A: GO terms of biological processes for gene differentially up-regulated in skeletal muscle tissue. The number of genes belonging to each term (gene count), enrichment (p-value) and false discovery rate (benjamini value) are shown in the corresponding columns.

Term	Gene Count	PValue	Benjamini
GO:0005743~mitochondrial inner membrane	77	3.06E-13	1.58E-10
GO:0043292~contractile fiber	41	1.28E-11	3.28E-09
GO:0005740~mitochondrial envelope	91	1.43E-11	2.45E-09
GO:0019866~organelle inner membrane	77	1.49E-11	1.91E-09
GO:0005739~mitochondrion	182	4.05E-11	4.16E-09
GO:0031966~mitochondrial membrane	86	4.17E-11	3.57E-09
GO:0030016~myofibril	38	5.69E-11	4.18E-09
GO:0070469~respiratory chain	30	1.18E-10	7.58E-09
GO:0044429~mitochondrial part	111	1.15E-09	6.54E-08
GO:0031226~intrinsic to plasma membrane	192	1.59E-09	8.18E-08
GO:0005746~mitochondrial respiratory chain	26	1.75E-09	8.19E-08
GO:0044449~contractile fiber part	36	1.78E-09	7.61E-08
GO:0030017~sarcomere	33	2.02E-09	8.00E-08
GO:0044459~plasma membrane part	308	8.94E-09	3.28E-07
GO:0044455~mitochondrial membrane part	37	9.24E-09	3.17E-07
GO:0005887~integral to plasma membrane	185	1.06E-08	3.40E-07
GO:0031967~organelle envelope	107	1.63E-07	4.93E-06
GO:0031975~envelope	107	1.93E-07	5.50E-06
GO:0005829~cytosol	194	6.07E-07	1.64E-05

Table S2-3B: GO terms of cellular compartments for gene differentially up-regulated in skeletal muscle tissue. *The number of genes belonging to each term (gene count), enrichment (p-value) and false discovery rate (benjamini value) are shown in the corresponding columns.*

Term	Gene Count	PValue	Benjamini
GO:0022610~biological adhesion	157	3.33E-29	1.17E-25
GO:0007155~cell adhesion	156	9.20E-29	1.62E-25
GO:0043062~extracellular structure organization	59	1.17E-21	1.38E-18
GO:0030199~collagen fibril organization	22	1.18E-16	9.76E-14
GO:0030198~extracellular matrix organization	39	5.03E-15	3.51E-12
GO:0000902~cell morphogenesis	76	4.77E-13	2.80E-10
GO:0016337~cell-cell adhesion	64	9.19E-13	4.61E-10
GO:0032989~cellular component morphogenesis	81	9.28E-13	4.08E-10
GO:0001501~skeletal system development	68	1.01E-11	3.93E-09
GO:0048858~cell projection morphogenesis	56	5.20E-11	1.83E-08
GO:0000904~cell morphogenesis involved in differentiation	55	1.38E-10	4.41E-08
GO:0032990~cell part morphogenesis	56	3.04E-10	8.92E-08
GO:0051960~regulation of nervous system development	46	6.89E-10	1.86E-07
GO:0006928~cell motion	84	6.93E-10	1.74E-07
GO:0030030~cell projection organization	70	9.70E-10	2.27E-07
GO:0007156~homophilic cell adhesion	36	1.37E-09	3.02E-07
GO:0048812~neuron projection morphogenesis	48	2.41E-09	4.98E-07
GO:0007409~axonogenesis	45	2.68E-09	5.24E-07
GO:0031589~cell-substrate adhesion	30	2.74E-09	5.07E-07
GO:0048667~cell morphogenesis involved in neuron differentiation	46	1.17E-08	2.07E-06
GO:0031175~neuron projection development	52	1.99E-08	3.33E-06
GO:0050767~regulation of neurogenesis	39	2.87E-08	4.59E-06
GO:0060284~regulation of cell development	44	5.52E-08	8.44E-06
GO:0007411~axon guidance	29	9.75E-08	1.43E-05

Table S2-4A: GO terms of biological processes for gene differentially down-regulated in skeletal muscle tissue. The number of genes belonging to each term (gene count), enrichment (p-value) and false discovery rate (benjamini value) are shown in the corresponding columns.

Term	Gene Count	PValue	Benjamini
GO:0031012~extracellular matrix	117	5.20E-37	2.60E-34
GO:0005578~proteinaceous extracellular matrix	111	4.03E-36	1.01E-33
GO:0044420~extracellular matrix part	53	1.41E-23	2.35E-21
GO:0044421~extracellular region part	181	2.68E-21	3.34E-19
GO:0005581~collagen	23	8.48E-15	8.42E-13
GO:0005604~basement membrane	31	3.86E-12	3.21E-10
GO:0005576~extracellular region	275	4.42E-12	3.15E-10
GO:0044459~plasma membrane part	281	5.77E-09	3.60E-07
GO:0042995~cell projection	108	1.27E-07	7.07E-06
GO:0030054~cell junction	84	6.02E-07	3.00E-05
GO:0045202~synapse	63	9.06E-07	4.11E-05
GO:0005886~plasma membrane	427	1.06E-06	4.39E-05
GO:0005583~fibrillar collagen	9	2.00E-06	7.69E-05
GO:0031224~intrinsic to membrane	583	1.28E-05	4.57E-04
GO:0005856~cytoskeleton	174	2.09E-05	6.94E-04
GO:0043005~neuron projection	57	2.21E-05	6.88E-04
GO:0031226~intrinsic to plasma membrane	156	2.31E-05	6.79E-04
GO:0015629~actin cytoskeleton	47	3.95E-05	0.001095806
GO:0005783~endoplasmic reticulum	127	4.08E-05	0.001071897
GO:0005887~integral to plasma membrane	151	5.44E-05	0.001356608
GO:0005911~cell-cell junction	36	7.04E-05	0.001672679
GO:0016021~integral to membrane	558	8.50E-05	0.001926237
GO:0031594~neuromuscular junction	10	8.55E-05	0.001852799

Table S2-4B: GO terms of cellular compartments for gene differentially down-regulated in skeletal muscle tissue. *The number of genes belonging to each term (gene count), enrichment (p-value) and false discovery rate (benjamini value) are shown in the corresponding columns.*

Gene symbol	Gene Name	logFC	logCPM	PValue
MYBPC1	Myosin binding protein C, slow type	10.52561963	6.55427022	5.35E-54
MYH8	Myosin, heavy chain 8, skeletal muscle, perinatal	9.105050991	7.6905291	2.90E-56
DPYSL5	Dihydropyrimidinase-like 5	9.095176909	5.12990888	1.66E-41
CASQ2	Calsequestrin 2 (cardiac muscle)	8.447736933	7.71247102	5.36E-53
ATP1B4	ATPase, (Na+)/K+ transporting, beta 4 polypeptide	8.243092299	4.81407202	9.50E-38
CACNG1	Calcium channel, voltage-dependent, gamma subunit 1	8.19672089	4.76672992	2.48E-37
MYH3	Myosin, heavy chain 3, skeletal muscle, embryonic	8.125375304	11.1839675	5.90E-56
BCAS1	Breast carcinoma amplified sequence 1	7.782751149	5.29884052	2.05E-39
SLN	Sarcolipin	7.758755566	5.01867456	9.20E-38
C10orf71	Chromosome 10 open reading frame 71	7.568519904	3.62288972	9.98E-29
TNNC2	Troponin C type 2 (fast)	7.524757295	5.25165123	2.71E-38
AIF1L	Allograft inflammatory factor 1-like	7.509015939	5.24032444	3.42E-38
TMOD1	Tropomodulin 1	7.41793347	4.68367627	7.24E-35
MYL1	Myosin, light chain 1, alkali; skeletal, fast	7.331087183	7.11114121	1.07E-44
NEB	Nebulin	7.254654979	10.8380388	1.94E-48
PKHD1	Polycystic kidney and hepatic disease 1 (autosomal recessive)	7.227730871	3.81264343	2.88E-29
HS6ST2	Heparan sulfate 6-O-sulfotransferase 2	7.126533649	4.86267808	6.17E-35
SRL	Sarcalumenin	7.118254804	6.17133235	1.45E-40
LMOD2	Leiomodin 2 (cardiac)	7.084187973	4.81449193	1.60E-34
SPARCL1	SPARC-like 1 (hevin)	7.055105125	3.11993498	1.17E-24
ACTN2	Actinin, alpha 2	7.051418832	8.19387195	7.57E-45
UNC45B	Unc-45 homolog B (C. elegans)	7.024828159	6.01108818	1.40E-39
CKM	Creatine kinase, muscle	7.001967809	6.13203426	8.07E-40
WIPF3	WAS/WASL interacting protein family, member 3	6.976161084	3.56487996	3.00E-27
NPNT	Nephronectin	6.872240864	6.46252798	3.27E-40
MYH7	Myosin, heavy chain 7, cardiac muscle, beta	6.871043964	7.96534527	3.01E-43
CLSTN2	Calsyntenin 2	6.802453504	8.54362361	1.89E-43
SCN7A	Sodium channel, voltage-gated, type VII, alpha	6.765444442	4.68636724	8.60E-33
TNNI2	Troponin I type 2 (skeletal, fast)	6.763664689	3.36331213	1.40E-25
COMP	Cartilage oligomeric matrix protein	6.751676528	4.49307584	8.94E-32
APOBEC2	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 2	6.733654252	3.71034726	1.59E-27
PACSIN1	Protein kinase C and casein kinase substrate in neurons 1	6.695491572	4.22845672	3.17E-30
TNNT3	Troponin T type 3 (skeletal, fast)	6.691717312	6.22101377	2.01E-38
EGLN3	Egl nine homolog 3 (C. elegans)	6.673093423	6.21052328	2.47E-38
TTN	Titin	6.66281633	14.0134449	1.15E-43
PDE2A	Phosphodiesterase 2A, cGMP-stimulated	6.610112061	3.59660777	1.42E-26
MYLPF	Myosin light chain, phosphorylatable, fast skeletal muscle	6.590210234	8.30065327	1.69E-41
HRC	Histidine rich calcium binding protein	6.577802837	5.7210592	3.11E-36
TRDN	Triadin	6.561824442	2.64837753	6.00E-21
SDK2	sidekick cell adhesion molecule 2	6.540743622	4.6351138	1.05E-31
ACHE	Acetylcholinesterase (Yt blood group)	6.513051704	2.5983779	1.35E-20
TRIM54	Tripartite motif-containing 54	6.49922448	3.10655192	1.64E-23
SMYD1	SET and MYND domain containing 1	6.492638749	5.97123271	1.42E-36
RYR1	Ryanodine receptor 1 (skeletal)	6.462874633	8.35480386	1.40E-40
ZBTB16	Zinc finger and BTB domain containing 16	6.43994824	2.5271181	4.43E-20
ZBTB7C	Zinc finger and BTB domain containing 7C	6.438572518	2.53559843	4.43E-20
MYBPC2	Myosin binding protein C, fast type	6.43487569	3.96965664	4.37E-28
EEF1A2	Eukaryotic translation elongation factor 1 alpha 2	6.423727913	3.4141413	4.62E-25
SMPX	Small muscle protein, X-linked	6.397576138	4.14038468	8.22E-29
KLHL40	Kelch repeat and BTB (POZ) domain containing 5	6.392006129	3.3809343	8.21E-25
GPRC5C	G protein-coupled receptor, family C, group 5, member C	6.306410363	3.59737211	9.26E-26
MYH6	Myosin, heavy chain 6, cardiac muscle, alpha	6.29333424	3.28371154	4.92E-24
XIRP2	Xin actin-binding repeat containing 2	6.250066766	5.15247886	1.46E-32
MYOZ2	Myozenin 2	6.249661936	5.15142613	1.46E-32
TMEM8C	Transmembrane protein 8-like	6.233700197	5.04736966	4.18E-32
BEST3	Bestrophin 3	6.228117804	3.76676942	2.07E-26
LDB3	LIM domain binding 3	6.198956955	6.54695102	3.32E-36
C1orf105	Chromosome 1 open reading frame 105	6.161907578	2.78512925	5.88E-21
SEP4	Septin 4	6.147776331	4.64997682	3.54E-30
TSPAN7	Tetraspanin 7	6.095432115	4.47241738	3.30E-29
CACNA1S	Calcium channel, voltage-dependent, L type, alpha 1S subunit	6.067787979	5.96176776	4.75E-34
C20orf166-AS1	Chromosome 20 open reading frame 200	6.046403221	3.59187855	5.86E-25

Gene Symbol	Gene Name	logFC	logCPM	PValue
NRAP	Nebulin-related anchoring protein	5.96172242	2.08276922	8.68E-17
ARPP21	Cyclic AMP-regulated phosphoprotein, 21 kD	5.94516413	6.35813268	3.96E-34
COX6A2	Cytochrome c oxidase subunit VIa polypeptide 2	5.90783981	2.91542904	4.00E-21
MYH2	Myosin, heavy chain 2, skeletal muscle, adult	5.89734875	5.04702777	2.03E-30
MYBPH	Myosin binding protein H	5.89118148	8.05419447	7.76E-36
TNNC1	Troponin C type 1 (slow)	5.88516375	6.67948196	2.75E-34
CAPN6	Calpain 6	5.88104873	2.00664143	3.13E-16
IGFBP2	Insulin-like growth factor binding protein 2, 36kDa	5.87155565	3.63656735	1.21E-24
MYOT	Myotilin	5.85897236	1.97937269	4.35E-16
CDK18	PCTAIRE protein kinase 3	5.80818522	4.80810424	3.37E-29
SPTB	Spectrin, beta, erythrocytic	5.80504174	4.81080108	3.19E-29
CASZ1	Castor zinc finger 1	5.78150037	4.29498216	3.64E-27
C1orf95	Chromosome 1 open reading frame 95	5.73712146	1.8693068	2.85E-15
PDGFB	Platelet-derived growth factor beta polypeptide (simian sarcoma viral (v-sis) oncogene homolog)	5.73399044	1.87210806	2.85E-15
KLHL31	Kelch-like 31 (Drosophila)	5.71722695	5.42121369	1.54E-30
COBL	Cordon-bleu homolog (mouse)	5.71175686	4.63244391	3.86E-28
SERPINA3	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	5.71113908	4.21350244	1.59E-26
KREMEN2	Kringle containing transmembrane protein 2	5.66427737	3.44424439	4.61E-23
ASB4	Ankyrin repeat and SOCS box-containing 4	5.65811696	1.79432902	8.38E-15
C3	Complement component 3	5.657463	6.26096421	4.30E-32
MYOM1	Myomesin 1, 185kDa	5.64626123	5.62964264	1.21E-30
CCDC3	Coiled-coil domain containing 3	5.64431778	3.75721769	2.20E-24
HFE2	Hemochromatosis type 2 (juvenile)	5.6418403	3.5946819	1.20E-23
ADH1B	Alcohol dehydrogenase 1B (class I), beta polypeptide	5.62893065	3.1809202	9.79E-22
COL15A1	Collagen, type XV, alpha 1	5.62682101	4.78905466	3.23E-28
FREM2	FRAS1 related extracellular matrix protein 2	5.60554187	1.74580343	1.76E-14
F13A1	Coagulation factor XIII, A1 polypeptide	5.58598874	8.47913014	1.24E-33
TSPAN33	Tetraspanin 33	5.57542753	3.97315368	5.06E-25
CARNS1	ATP-grasp domain containing 1	5.54951087	3.94445614	8.73E-25
DCX	Doublecortin	5.52784945	2.17857144	1.73E-16
SCN3B	Sodium channel, voltage-gated, type III, beta	5.52690515	2.55227061	2.19E-18
NEURL	Neuralized homolog (Drosophila)	5.4983339	2.15285239	2.71E-16
TRIM63	Tripartite motif-containing 63	5.48547638	1.64074328	9.98E-14
RAI2	Retinoic acid induced 2	5.46937593	1.62968728	1.22E-13
TNNI1	Troponin I type 1 (skeletal, slow)	5.46086278	8.52238311	1.30E-32
RBFOX1	RNA binding protein, fox-1 homolog (C. elegans) 1	5.42283629	4.98626307	9.25E-28
CAMK2B	Calcium/calmodulin-dependent protein kinase II beta	5.41388698	2.0788736	9.26E-16
CACNG6	Calcium channel, voltage-dependent, gamma subunit 6	5.4026796	2.72956108	6.39E-19
NDRG2	NDRG family member 2	5.37838247	4.1145739	1.04E-24
CRMP1	Collapsin response mediator protein 1	5.37051382	4.5454588	3.57E-26
KCNQ4	Potassium voltage-gated channel, KQT-like subfamily, member 4	5.36400912	3.14588132	1.07E-20
SCN4A	Sodium channel, voltage-gated, type IV, alpha subunit	5.36155467	4.20105575	5.97E-25
APOL4	Apolipoprotein L, 4	5.34337695	2.38623354	4.39E-17
C20orf166	Chromosome 20 open reading frame 166	5.32979922	1.50031333	7.79E-13
TAS1R1	Taste receptor, type 1, member 1	5.32524364	2.3718704	5.72E-17
MYOM3	Myomesin family, member 3	5.30979562	6.29147339	1.24E-29
NAT8L	N-acetyltransferase 8-like (GCN5-related, putative)	5.29593031	2.87553607	2.64E-19
LMOD3	Leiomodin 3 (fetal)	5.28706194	4.66880044	4.23E-26
C7	Complement component 7	5.28358148	1.94723367	6.45E-15
RRAGD	Ras-related GTP binding D	5.26181681	2.595551	6.60E-18
PRKAG3	Protein kinase, AMP-activated, gamma 3 non-catalytic subunit	5.25339332	3.04221108	6.45E-20
A2M	Alpha-2-macroglobulin	5.20571684	5.27371068	4.29E-27
PPM1E	Protein phosphatase 1E (PP2C domain containing)	5.14715617	1.82721004	5.13E-14
MYL4	Myosin, light chain 4, alkali; atrial, embryonic	5.09854942	6.85517061	9.32E-29
CXCR4	Chemokine (C-X-C motif) receptor 4	5.09337917	1.29605446	1.79E-11
KCNE1L	KCNE1-like	5.09331464	2.15711582	1.93E-15
ACTA1	Actin, alpha 1, skeletal muscle	5.07544361	6.51765936	2.86E-28
CHRND	Cholinergic receptor, nicotinic, delta	5.06266164	4.9441719	2.39E-25
ONECUT2	One cut homeobox 2	5.05638209	3.47202822	5.59E-21
GPR123	G protein-coupled receptor 123	5.04667902	3.03379001	3.30E-19
XPNPEP2	X-prolyl aminopeptidase (aminopeptidase P) 2, membrane-bound	5.03904277	1.23292295	3.67E-11
TXLNB	Taxilin beta	5.01997749	4.46652157	4.55E-24

Table S2-5: Genes up regulated ≥ 5log₂ FC in myotubes as compared to myoblasts. *The table shows the gene names, their description, corresponding log₂ fold change and P value. Fold change = value describing the average level of increase or decrease in gene expression.*

Gene Symbol	Gene Name	logFC	logCPM	P Value
HBB	Hemoglobin, beta	12.47123082	8.858841055	2.10E-30
MB	Myoglobin	12.06569951	9.427084423	2.96E-24
PDK4	Pyruvate dehydrogenase kinase, isozyme 4	10.57340817	11.79275026	9.16E-22
THBS4	Thrombospondin 4	10.43519147	8.508393596	8.43E-21
SMTNL2	Smoothelin-like 2	10.38062652	6.76924324	2.21E-19
LBP	Lipopolysaccharide binding protein	10.16190804	6.553929019	7.29E-19
CD74	CD74 molecule, major histocompatibility complex, class II invariant chain	9.814947785	6.766449059	1.15E-18
FABP4	Fatty acid binding protein 4, adipocyte	9.724925075	7.281185917	8.20E-19
C8orf22	Chromosome 8 open reading frame 22	9.677300468	6.072002497	1.03E-17
CD93	CD93 molecule	9.558594127	6.532219477	6.71E-18
LPL	Lipoprotein lipase	9.552570827	6.525922948	6.93E-18
RNASE1	Ribonuclease, RNase A family, 1 (pancreatic)	9.369074301	5.768535715	5.53E-17
RORC	RAR-related orphan receptor C	9.33919356	6.313184686	2.22E-17
CA3	Carbonic anhydrase III, muscle specific	9.272824151	7.105442875	7.88E-18
PPP1R1A	Protein phosphatase 1, regulatory (inhibitor) subunit 1A	9.230591275	7.515850874	5.35E-18
DARC	Duffy blood group, chemokine receptor	9.170661329	5.572047025	2.66E-20
IDI2	Isopentenyl-diphosphate delta isomerase 2	9.000731982	5.976259759	1.41E-16
HLA-DRA	Major histocompatibility complex, class II, DR alpha	8.967118433	5.37075946	4.94E-16
NRAP	Nebulin-related anchoring protein	8.966507925	10.95605903	8.43E-23
RPL3L	Ribosomal protein L3-like	8.954709543	5.931169296	1.81E-16
VWF	Von Willebrand factor	8.84626447	7.887939801	2.46E-17
SLPI	Secretory leukocyte peptidase inhibitor	8.804188224	5.783788299	4.11E-16
ARHGEF15	Rho guanine nucleotide exchange factor (GEF) 15	8.755606132	5.159143092	4.88E-19
PLVAP	Plasmalemma vesicle associated protein	8.684552363	5.666837419	7.87E-16
TRIM63	Tripartite motif-containing 63	8.677230365	10.20928442	2.86E-17
PLA2G2A	Phospholipase A2, group IIA (platelets, synovial fluid)	8.637471489	5.619632932	1.02E-15
UBE2QL1	Probable ubiquitin-conjugating enzyme E2 FLJ25076	8.420121719	6.475905977	5.88E-16
CXorf36	Chromosome X open reading frame 36	8.402936471	4.817300806	1.06E-14
GIMAP5	GTPase, IMAP family member 5	8.355982153	4.768119904	7.97E-18
CASQ1	Calsequestrin 1 (fast-twitch, skeletal muscle)	8.340145347	6.606601119	6.18E-16
S100A9	S100 calcium binding protein A9	8.263388568	4.683952454	2.25E-14
HBA2	hemoglobin, alpha 2	8.249494194	4.670537529	2.43E-14
PLIN4	KIAA1881	8.241009099	6.912255649	8.04E-16
ALDH1L1	Aldehyde dehydrogenase 1 family, member L1	8.239615404	4.65460834	2.56E-14
PPP1R27	Dysferlin interacting protein 1	8.237884791	6.507123727	1.08E-15
ASB2	Ankyrin repeat and SOCS box-containing 2	8.231732035	7.53142713	6.06E-16
STAB1	Stabilin 1	8.192430568	5.165661282	7.93E-15
VSIG4	V-set and immunoglobulin domain containing 4	8.178776663	4.602967131	3.56E-14
TMEM140	Transmembrane protein 140	8.168695426	6.456552967	1.81E-15
S100A8	S100 calcium binding protein A8	8.157484513	4.581227059	3.99E-14
C1QB	Complement component 1, q subcomponent, B chain	8.150948215	4.57507432	4.14E-14
RBP7	Retinol binding protein 7, cellular	8.146153141	4.571682185	4.25E-14
HLA-DRB1	major histocompatibility complex, class II, DR beta 1	8.136590014	5.103369487	5.97E-18
MAOB	Monoamine oxidase B	8.128065466	5.120925684	1.62E-14
PLIN5	Lipid storage droplet protein 5	8.118111124	4.535711133	4.94E-14
CLIC5	Chloride intracellular channel 5	8.115175437	5.084609479	1.21E-14
TSPAN8	Tetraspanin 8	8.047680317	4.471413729	7.24E-14
C8orf4	Chromosome 8 open reading frame 4	8.02277163	6.557926103	3.14E-15
SLCO4A1	Solute carrier organic anion transporter family, member 4A1	7.980227589	4.955583524	2.50E-14
FAM134B	Family with sequence similarity 134, member B	7.949206214	8.281469739	1.90E-15
NOTCH4	Notch homolog 4 (Drosophila)	7.948844258	5.328726951	6.80E-18
APOD	Apolipoprotein D	7.904334554	9.376509176	1.96E-15
S100A1	S100 calcium binding protein A1	7.888986354	6.493093942	5.45E-15
CNTFR	Ciliary neurotrophic factor receptor	7.882565672	4.303228177	1.76E-13
PYGM	Phosphorylase, glycogen, muscle	7.87810194	9.613524494	2.18E-15
IL6R	Interleukin 6 receptor	7.857862095	7.147592851	4.96E-15
GJA4	Gap junction protein, alpha 4, 37kDa	7.857128776	4.284347776	2.55E-16
TIE1	Tyrosine kinase with immunoglobulin-like and EGF-like domains 1	7.852526793	4.283070675	2.07E-13
SMTNL1	Smoothelin-like 1	7.836414908	5.221175358	3.19E-14
ESAM	Endothelial cell adhesion molecule	7.79205772	5.009751611	4.66E-14
GIMAP4	GTPase, IMAP family member 4	7.787708881	4.219682772	2.94E-13
BTNL9	Butyrophilin-like 9	7.766338812	4.768040763	1.15E-13
ANKRD2	Ankyrin repeat domain 2 (stretch responsive muscle)	7.762131262	6.71406599	9.71E-15
CDH5	Cadherin 5, type 2 (vascular endothelium)	7.750879999	5.810271185	2.32E-14
MYH14	Myosin, heavy chain 14	7.746239362	7.075156732	8.42E-15
HLA-DPA1	Major histocompatibility complex, class II, DP alpha 1	7.735630377	4.165825665	5.89E-16
MPZL2	Myelin protein zero-like 2	7.717293343	4.155121689	4.30E-13
SLC16A6	Solute carrier family 16, member 6 (monocarboxylic acid transporter 7)	7.692615927	4.133025758	4.91E-13
HBA1	Hemoglobin, alpha 1	7.687427078	4.125397072	5.06E-13

Gene Symbol	Gene Name	logFC	logCPM	P Value
MYF6	Myogenic factor 6 (herculin)	7.663543898	7.756905545	9.26E-15
C5orf27	Chromosome 5 open reading frame 27	7.643110636	5.504817284	2.13E-17
KRT6A	Keratin 6A	7.633743229	4.056213689	1.19E-15
UCP3	Uncoupling protein 3 (mitochondrial, proton carrier)	7.625563712	7.552315592	1.25E-14
CYP4B1	Cytochrome P450, family 4, subfamily B, polypeptide 1	7.597172156	4.028331552	8.22E-13
CPVL	Carboxypeptidase, vitellogenic-like	7.577806355	4.022965147	9.07E-13
MYLK2	Myosin light chain kinase 2	7.563164343	6.289859143	3.54E-14
ABRA	Actin-binding Rho activating protein	7.549433199	6.013811625	4.02E-14
SEPP1	selenoprotein P, plasma, 1	7.523062007	7.17924044	4.75E-18
MYH11	Myosin, heavy chain 11, smooth muscle	7.51426217	6.606042422	3.23E-14
WNK2	WNK lysine deficient protein kinase 2	7.494484099	3.927450505	1.43E-12
SMOC2	SPARC related modular calcium binding 2	7.473825218	4.485730386	5.54E-13
FCGR3A	Fc fragment of IgG, low affinity IIIa, receptor (CD16a)	7.419011628	3.861184049	5.23E-15
KANK3	KN motif and ankyrin repeat domains 3	7.415057898	3.839156695	5.42E-15
PLN	Phospholamban	7.39994606	6.005273463	7.95E-14
ITM2A	Integral membrane protein 2A	7.393068631	3.839477225	2.46E-12
FHL5	Four and a half LIM domains 5	7.365682106	3.80875615	2.84E-12
LYVE1	Lymphatic vessel endothelial hyaluronan receptor 1	7.35234553	5.088806037	2.84E-13
GPR116	G protein-coupled receptor 116	7.350820527	5.966896021	1.19E-13
SHANK3	SH3 and multiple ankyrin repeat domains 3	7.335427525	6.765479029	2.01E-17
ADCY2	Adenylate cyclase 2 (brain)	7.335053367	6.254138563	1.02E-13
EMCN	Endomucin	7.326200227	4.53128377	6.68E-13
GIMAP8	GTPase, IMAP family member 8	7.305654295	3.75342028	3.93E-12
CP	Ceruloplasmin (ferroxidase)	7.305582556	3.759774902	3.93E-12
MAPT	Microtubule-associated protein tau	7.294799141	6.575754228	1.00E-13
RGS5	Regulator of G-protein signaling 5	7.293433026	7.252859649	7.61E-14
MAOA	Monoamine oxidase A	7.263321441	5.244883016	2.04E-16
SLC2A4	Solute carrier family 2 (facilitated glucose transporter), mer	7.258021836	6.33238803	1.44E-13
FBXO40	F-box protein 40	7.232176694	4.63065885	8.34E-13
SLC22A3	Solute carrier family 22 (extraneuronal monoamine transpo	7.212206782	5.312664681	3.57E-13
GSTT1	Glutathione S-transferase theta 1	7.205384542	4.202449564	3.81E-15
C1orf170	Chromosome 1 open reading frame 170	7.177085999	4.173604939	4.13E-15
PTPN3	protein tyrosine phosphatase, non-receptor type 3	7.167030051	6.52848554	1.91E-13
C1QA	Complement component 1, q subcomponent, A chain	7.161653746	3.624002668	8.43E-12
KLHL38	Kelch-like 38 (Drosophila)	7.123685945	5.985597535	3.19E-13
SLC11A1	solute carrier family 11 (proton-coupled divalent metal ion t	7.120634514	3.587500612	1.05E-11
GIMAP6	GTPase, IMAP family member 6	7.086559199	3.543842278	1.26E-11
C1QC	Complement component 1, q subcomponent, C chain	7.07911828	3.545615209	1.31E-11
ICAM2	Intercellular adhesion molecule 2	7.074025574	4.485003703	1.95E-12
MYH1	Myosin, heavy chain 1, skeletal muscle, adult	7.041318515	8.363589383	2.14E-13
LRRC38	Leucine rich repeat containing 38	7.04060791	3.48943618	1.61E-11
SHE	Src homology 2 domain containing E	7.039645661	3.499848466	1.62E-11
CMYA5	cardiomyopathy associated 5	7.031852559	11.15716203	1.86E-13
ECSCR	endothelial cell surface expressed chemotaxis and apoptosi	7.029702298	4.042592869	1.14E-14
RNF157	Ring finger protein 157	7.022872419	6.578906285	3.67E-13
RGL3	Ral guanine nucleotide dissociation stimulator-like 3	7.01828485	4.214928557	5.79E-15
STEAP4	STEAP family member 4	7.009831908	5.595104446	7.58E-13
TMEM52	Transmembrane protein 52	7.004314001	3.999787185	4.81E-12

Table S2-6: Genes up regulated ≥ 7log₂ FC in Skeletal muscle as compared to myotubes. *The table shows the gene names, their description, corresponding log₂ fold change and P value. Fold change = value describing the average level of increase or decrease in gene expression.*

Gene	Description	log FC	P Value
ABC superfamily			
ABCA1	ATP-binding cassette, sub-family A (ABC1), member 1	2.00652285	3.99E-07
ABCA3	ATP-binding cassette, sub-family A (ABC1), member 3	3.01033754	1.29E-10
ABCA6	ATP-binding cassette, sub-family A (ABC1), member 6	3.212978429	1.36E-12
ABCA8	ATP-binding cassette, sub-family A (ABC1), member 8	1.122990965	0.00412888
ABCA9	ATP-binding cassette, sub-family A (ABC1), member 9	1.68084996	2.83E-05
ABCG1	ATP-binding cassette, sub-family G (WHITE), member 1	2.937221445	0.00289602
Apolipoproteins			
APOD	Apolipoprotein D	3.951442744	1.43E-10
APOE	Apolipoprotein E	4.005260913	8.13E-11
APOL1	Apolipoprotein L, 1	1.354669951	0.00072394
APOL3	Apolipoprotein L, 3	2.108419831	4.66E-06
APOL4	Apolipoprotein L, 4	5.343376951	4.39E-17
APOL6	Apolipoprotein L, 6	1.427253416	0.00049957
ASB family			
ASB14	Ankyrin repeat and SOCS box-containing 14	3.361238075	0.00024988
ASB16	Ankyrin repeat and SOCS box-containing 16	3.158012684	0.00083551
ASB2	Ankyrin repeat and SOCS box-containing 2	2.690931483	0.00184679
ASB4	Ankyrin repeat and SOCS box-containing 4	5.658116963	8.38E-15
ATPases family			
ATP1A2	ATPase, Na+/K+ transporting, alpha 2 (+) polypeptide	4.531382055	4.52E-23
ATP1B4	ATPase, (Na+)/K+ transporting, beta 4 polypeptide	8.243092299	9.50E-38
ATP2A1	ATPase, Ca++ transporting, cardiac muscle, fast twitch 1	3.85587447	4.01E-18
ATP8A1	ATPase, aminophospholipid transporter (APLT), class I, type 8A, member 1	1.999379667	4.54E-05
Voltage dependent calcium channels			
CACNA1H	Calcium channel, voltage-dependent, T type, alpha 1H subunit	3.128384668	2.76E-13
CACNA1S	Calcium channel, voltage-dependent, L type, alpha 1S subunit	6.067787979	4.75E-34
CACNB1	Calcium channel, voltage-dependent, beta 1 subunit	2.717803259	4.97E-11
CACNB4	Calcium channel, voltage-dependent, beta 4 subunit	1.794493258	1.12E-05
CACNG1	Calcium channel, voltage-dependent, gamma subunit 1	8.19672089	2.48E-37
CACNG4	Calcium channel, voltage-dependent, gamma subunit 4	3.845002792	3.33E-11
CACNG6	Calcium channel, voltage-dependent, gamma subunit 6	5.402679597	6.39E-19
Coiled coil			
CCDC141	Coiled-coil domain containing 141	3.796533274	3.06E-18
CCDC3	Coiled-coil domain containing 3	5.644317781	2.20E-24
CCDC39	Coiled-coil domain containing 39	4.106459981	8.48E-13
CCDC69	Coiled-coil domain containing 69	3.289532815	7.78E-10
CDK inhibitors			
CDKN1A	Cyclin-dependent kinase inhibitor 1A (p21, Cip1)	1.042024475	0.00617358
CDKN1C	Cyclin-dependent kinase inhibitor 1C (p57, Kip2)	3.148407867	2.51E-13
CDKN2B	Cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)	2.645930427	9.64E-11
Collagen			
COL10A1	Collagen, type X, alpha 1	4.162185829	1.18E-13
COL11A1	Collagen, type XI, alpha 1	1.377232327	3.29E-04
COL15A1	Collagen, type XV, alpha 1	5.626821006	3.23E-28
COL18A1	Collagen, type XVIII, alpha 1	1.345534735	5.04E-04
COL19A1	Collagen, type XIX, alpha 1	4.660905862	2.58E-20
COL1A1	Collagen, type I, alpha 1	1.051390897	5.65E-03
COL21A1	Collagen, type XXI, alpha 1	2.463083359	7.58E-09
COL3A1	Collagen, type III, alpha 1	1.030148803	6.68E-03
COL4A1	Collagen, type IV, alpha 1	1.043655096	6.04E-03
COL4A4	Collagen, type IV, alpha 4	3.001945637	4.26E-08
COL8A2	Collagen, type VIII, alpha 2	2.025984696	4.63E-07
Cytochrome oxidases			
COX6A2	Cytochrome c oxidase subunit VIa polypeptide 2	5.90783981	4.00E-21
COX7A1	Cytochrome c oxidase subunit VIIa polypeptide 1 (muscle)	1.966619983	1.30E-05
FOX family			
FOXO1	Forkhead box O1	3.641978493	2.53E-16
FOXO4	Forkhead box O4	2.06747158	8.81E-07
FOXO6	forkhead box O6	3.270501254	5.72E-10
FOXS1	Forkhead box S1	4.506860662	8.37E-12

Gene	Description	log FC	P Value
IGFBP family			
IGFBP2	Insulin-like growth factor binding protein 2, 36kDa	5.87155565	1.21E-24
IGFBP5	Insulin-like growth factor binding protein 5	1.13151665	0.0030116
IGFBP7	Insulin-like growth factor binding protein 7	1.35478777	0.00051494
Interleukins			
IL16	Interleukin 16 (lymphocyte chemoattractant factor)	2.81201577	1.05E-05
IL17B	Interleukin 17B	4.28616711	3.49E-09
IL18R1	Interleukin 18 receptor 1	4.78285044	2.88E-17
IL1B	Interleukin 1, beta	1.80204674	0.00010703
IL1R1	Interleukin 1 receptor, type I	1.53200093	7.92E-05
IL1RL1	Interleukin 1 receptor-like 1	2.75750413	5.85E-09
IL20RA	Interleukin 20 receptor, alpha	2.41081326	4.72E-05
IL32	Interleukin 32	4.01357442	2.48E-19
IL34	Interleukin 34	2.36307467	7.78E-06
Integrin family			
ITGA4	Integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)	1.57974364	4.58E-05
ITGA6	Integrin, alpha 6	2.01021194	2.75E-07
ITGA7	Integrin, alpha 7	1.19872387	0.00172607
ITGB1BP2	Integrin beta 1 binding protein (melusin) 2	3.9410166	8.21E-16
ITGB8	Integrin, beta 8	1.23157078	0.0024167
Myosin Family			
MYH1	Myosin, heavy chain 1, skeletal muscle, adult	4.35781903	6.41E-11
MYH14	Myosin, heavy chain 14	2.70469526	0.00184679
MYH15	Myosin, heavy chain 15	1.90419568	1.34E-05
MYH2	Myosin, heavy chain 2, skeletal muscle, adult	5.89734875	2.03E-30
MYH3	Myosin, heavy chain 3, skeletal muscle, embryonic	8.1253753	5.90E-56
MYH6	Myosin, heavy chain 6, cardiac muscle, alpha	6.29333424	4.92E-24
MYH7	Myosin, heavy chain 7, cardiac muscle, beta	6.87104396	3.01E-43
MYH7B	Myosin, heavy chain 7B, cardiac muscle, beta	4.39299426	6.85E-14
MYH8	Myosin, heavy chain 8, skeletal muscle, perinatal	9.10505099	2.90E-56
MYO18A	Myosin XVIIIA	1.15349376	0.00283372
MYO18B	Myosin XVIIIB	1.72915026	9.62E-06
MYOG	Myogenin (myogenic factor 4)	3.65975412	5.39E-18
MYOM1	Myomesin 1, 185kDa	5.64626123	1.21E-30
MYOM2	Myomesin (M-protein) 2, 165kDa	3.14670901	1.13E-10
MYOM3	Myomesin family, member 3	5.30979562	1.24E-29
MYOT	Myotilin	5.85897236	4.35E-16
MYOZ1	Myozenin 1	4.20231231	3.73E-11
MYOZ2	Myozenin 2	6.24966194	1.46E-32
MYL1	Myosin, light chain 1, alkali; skeletal, fast	7.33108718	1.07E-44
MYL2	Myosin, light chain 2, regulatory, cardiac, slow	4.90131287	5.76E-21
MYL3	Myosin, light chain 3, alkali; ventricular, skeletal, slow	3.22475787	9.32E-06
MYL4	Myosin, light chain 4, alkali; atrial, embryonic	5.09854942	9.32E-29
MYL6B	Myosin, light chain 6B, alkali, smooth muscle and non-muscle	2.58926962	1.93E-10
MYLK4	Myosin light chain kinase family, member 4	1.6450026	0.00356855
MYLPF	Myosin light chain, phosphorylatable, fast skeletal muscle	6.59021023	1.69E-41
Voltage gated potassium channels			
KCNAB1	Potassium voltage-gated channel, shaker-related subfamily, beta member 1	3.43739997	1.57E-06
KCNB1	Potassium voltage-gated channel, Shab-related subfamily, member 1	3.07806228	1.32E-12
KCNH2	Potassium voltage-gated channel, subfamily H (eag-related), member 2	3.64872678	1.06E-11
KCNH3	Potassium voltage-gated channel, subfamily H (eag-related), member 3	3.15520523	8.36E-04
KCNJ11	Potassium inwardly-rectifying channel, subfamily J, member 11	3.79872264	1.20E-05
KCNJ12	Potassium Inwardly-Rectifying Channel, Subfamily J, Member 12	3.88551255	1.05E-14
KCNJ6	Potassium Inwardly-Rectifying Channel, Subfamily J, Member 6	3.55194048	7.02E-14
KCNK3	Potassium channel, subfamily K, member 3	4.27839168	1.46E-07
KCNQ1	Potassium voltage-gated channel, KQT-like subfamily, member 1	3.54483705	7.76E-05
KCNQ4	Potassium voltage-gated channel, KQT-like subfamily, member 4	5.36400912	1.07E-20
KCNT1	Potassium channel, subfamily T, member 1	3.91598402	4.10E-06
KCNT2	Potassium channel, subfamily T, member 2	1.11692583	7.31E-03
KCTD12	Potassium channel tetramerisation domain containing 12	1.11785533	3.64E-03

Gene	Description	log FC	P Value
Photocadherins			
PCDH1	Protocadherin 1	2.75606522	5.23E-11
PCDH10	Protocadherin 10	1.06018493	0.00591171
PCDH19	Protocadherin 19	2.78054673	2.64E-09
PCDHB10	Protocadherin beta 10	1.42863769	0.00169714
PCDHB11	Protocadherin beta 11	1.79713389	0.00066847
PCDHB9	Protocadherin beta 9	1.81056918	0.00017762
Sodium channel family			
SCN3B	Sodium channel, voltage-gated, type III, beta	5.52690515	2.19E-18
SCN4A	Sodium channel, voltage-gated, type IV, alpha subunit	5.36155467	5.97E-25
SCN5A	Sodium channel, voltage-gated, type V, alpha subunit	2.27143739	1.17E-08
SCN7A	Sodium channel, voltage-gated, type VII, alpha	6.76544444	8.60E-33
Solute carrier family			
SLC12A7	Solute carrier family 12 (potassium/chloride transporters), member 7	1.0704181	0.00554857
SLC12A8	Solute carrier family 12 (potassium/chloride transporters), member 8	1.45806763	3.63E-04
SLC15A3	Solute carrier family 15, member 3	1.9590115	5.24E-05
SLC16A10	Solute carrier family 16, member 10 (aromatic amino acid transporter)	2.75370367	6.75E-03
SLC22A15	Solute carrier family 22, member 15	2.32193842	1.17E-07
SLC25A34	Solute carrier family 25, member 34	1.56250591	7.58E-03
SLC2A4	Solute carrier family 2 (facilitated glucose transporter), member 4	2.9282485	2.90E-03
SLC2A5	Solute carrier family 2 (facilitated glucose/fructose transporter), member 5	4.80732433	1.13E-16
SLC44A3	Solute carrier family 44, member 3	2.12159942	8.01E-05
SLC46A3	Solute carrier family 46, member 3	2.29290053	5.66E-08
SLC4A11	Solute carrier family 4, sodium borate transporter, member 11	1.88492284	2.48E-03
SLC6A17	Solute carrier family 6, member 17	4.13871697	1.58E-17
SLC7A7	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 7	2.64344258	8.83E-09
SLC8A3	Solute carrier family 8 (sodium/calcium exchanger), member 3	2.11093408	5.53E-07
Anion transporter family			
SLCO3A1	Solute carrier organic anion transporter family, member 3A1	1.16129944	4.11E-03
SLCO5A1	Solute carrier organic anion transporter family, member 5A1	3.54156986	6.08E-10
Troponin			
TNNC1	Troponin C type 1 (slow)	5.88516375	2.75E-34
TNNC2	Troponin C type 2 (fast)	7.5247573	2.71E-38
TNNI1	Troponin I type 1 (skeletal, slow)	5.46086278	1.30E-32
TNNI2	Troponin I type 2 (skeletal, fast)	6.76366469	1.40E-25
TNNT1	Troponin T type 1 (skeletal, slow)	2.89446139	1.41E-12
TNNT2	Troponin T type 2 (cardiac)	4.26848428	9.74E-23
TNNT3	Troponin T type 3 (skeletal, fast)	6.69171731	2.01E-38
Zinc finger protein family			
ZBTB16	Zinc finger and BTB domain containing 16	6.43994824	4.43E-20
ZBTB20	Zinc finger and BTB domain containing 20	1.21956619	0.00168255
ZBTB46	Zinc finger and BTB domain containing 46	1.22763624	0.00705204
ZBTB47	Zinc finger and BTB domain containing 47	1.14530515	0.00299628
ZBTB7C	Zinc finger and BTB domain containing 7C	6.43857252	4.43E-20
ZEB2	Zinc finger E-box binding homeobox 2	1.50755684	9.33E-05
ZFHX2	Zinc finger homeobox 2	1.88481832	8.00E-06
ZFP106	Zinc finger protein 106 homolog (mouse)	2.86451808	1.24E-12
ZNF138	Zinc finger protein 138	1.18044266	0.0055102
ZNF219	Zinc finger protein 219	2.27876911	3.41E-07
ZNF385B	Zinc finger protein 385B	4.1439872	5.31E-07
ZNF474	Zinc finger protein 474	1.74503596	0.00793409
ZNF556	Zinc finger protein 556	2.68892637	9.25E-09
ZNF704	Zinc finger protein 704	1.1976498	0.00219241

Table S2-7: Protein families and their member genes found upregulated in myotubes. *Gene names, their description, corresponding log$_2$ fold change and P Value are shown. Fold change = value describing the average level of increase or decrease in gene expression.*

Gene	Description	logFC	P Value
ABC family			
ABCA10	ATP-binding cassette, sub-family A (ABC1), member 10	2.9458942	3.14E-04
ABCA5	ATP-binding cassette, sub-family A (ABC1), member 5	2.59479069	9.38E-04
ABCB4	ATP-binding cassette, sub-family B (MDR/TAP), member 4	2.80748421	2.64E-03
ABCC2	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	2.48981292	2.27E-03
ABCC6	ATP-binding cassette, sub-family C (CFTR/MRP), member 6	3.02883646	1.20E-03
ABCC9	ATP-binding cassette, sub-family C (CFTR/MRP), member 9	3.22173142	6.21E-05
ABCG1	ATP-binding cassette, sub-family G (WHITE), member 1	5.5777641	1.05E-09
Acyl CoA synthetase family			
ACSL1	Acyl-CoA synthetase long-chain family member 1	2.82502005	3.40E-04
ACSL5	Acyl-CoA synthetase long-chain family member 5	3.95303812	1.75E-07
ACSM5	Acyl-CoA synthetase medium-chain family member 5	3.95577118	2.42E-05
ACSS1	Acyl-CoA synthetase short-chain family member 1	2.8732394	3.23E-04
ACSS2	Acyl-CoA synthetase short-chain family member 2	2.45325815	1.62E-03
Adenylate cyclase family			
ADCY2	Adenylate cyclase 2 (brain)	7.33505337	1.02E-13
ADCY4	Adenylate cyclase 4	4.92850141	3.30E-08
ADCY5	Adenylate cyclase 5	3.85185914	1.36E-05
ADCY9	Adenylate cyclase 9	2.1236005	5.70E-03
Adenylate dehydrogenase family			
ALDH1A2	Aldehyde dehydrogenase 1 family, member A2	4.38718616	6.37E-07
ALDH1A3	Aldehyde dehydrogenase 1 family, member A3	4.34840649	6.43E-09
ALDH1L1	Aldehyde dehydrogenase 1 family, member L1	8.2396154	2.56E-14
ALDH2	Aldehyde dehydrogenase 2 family (mitochondrial)	2.39134661	2.10E-03
ALDH5A1	Aldehyde dehydrogenase 5 family, member A1	2.3615759	2.75E-03
ALDH8A1	Aldehyde dehydrogenase 8 family, member A1	2.52132673	7.34E-03
Ankyrin repeat and SOCS box			
ASB10	Ankyrin repeat and SOCS box-containing 10	5.70416324	1.73E-08
ASB11	Ankyrin repeat and SOCS box-containing 11	6.89531559	1.24E-11
ASB12	Ankyrin repeat and SOCS box-containing 12	3.682685	0.00023406
ASB15	Ankyrin repeat and SOCS box-containing 15	5.69148411	1.87E-08
ASB16	Ankyrin repeat and SOCS box-containing 16	3.19378634	1.26E-04
ASB2	Ankyrin repeat and SOCS box-containing 2	8.23173203	6.06E-16
ASB8	Ankyrin repeat and SOCS box-containing 8	3.26022634	5.04E-05
Carbonic anhydrase			
CA1	Carbonic anhydrase I	5.4109504	7.67E-08
CA11	Carbonic anhydrase XI	3.0805614	3.17E-04
CA2	Carbonic anhydrase II	6.54816359	5.49E-11
CA3	Carbonic anhydrase III, muscle specific	9.27282415	7.88E-18
CA4	Carbonic anhydrase IV	5.63979758	2.41E-08
CA8	Carbonic anhydrase VIII	6.01705375	3.48E-09
Cluster of differentiation			
CD14	CD14 molecule	3.3670475	6.24E-05
CD163	CD163 molecule	6.25179297	4.19E-11
CD177	CD177 molecule	2.73805414	8.38E-03
CD1D	CD1d molecule	3.25644776	1.35E-03
CD209	CD209 molecule	4.3730824	1.15E-05
CD247	CD247 molecule	2.2646182	0.00929282
CD300A	CD300a molecule	4.12453877	3.50E-05
CD300E	CD300e molecule	3.34367414	0.00020939
CD300LG	CD300 molecule-like family member g	5.84980578	8.34E-09
CD33	CD33 molecule	3.61007528	3.23E-04
CD34	CD34 molecule	3.75031192	6.23E-06
CD36	CD36 molecule (thrombospondin receptor)	2.82485659	3.44E-04
CD37	CD37 molecule	4.1412272	3.33E-05
CD38	CD38 molecule	3.11015215	2.38E-03
CD4	CD4 molecule	4.30503652	4.85E-06
CD40	CD40 molecule, TNF receptor superfamily member 5	5.17925012	3.19E-08
CD48	CD48 molecule	3.2059943	0.00158057
CD52	CD52 molecule	3.07233296	0.00258542
CD53	CD53 molecule	5.52192082	4.40E-08
CD55	CD55 molecule, decay accelerating factor for complement (Cromer blood group)	4.75169334	3.46E-10
CD68	CD68 molecule	4.61549752	1.09E-06
CD72	CD72 molecule	2.35152105	9.84E-03
CD74	CD74 molecule, major histocompatibility complex, class II invariant chain	9.81494778	1.15E-18
CD79B	CD79b molecule, immunoglobulin-associated beta	3.70589459	7.40E-05
CD84	CD84 molecule	3.53032159	4.51E-04
CD86	CD86 molecule	4.2836447	1.71E-05
CD9	CD9 molecule	2.00932849	8.99E-03
CD93	CD93 molecule	9.55859413	6.71E-18

Gene	Description	logFC	P Value
Chemokine			
CCL18	Chemokine (C-C motif) ligand 18 (pulmonary and activation-regulated)	2.86718189	0.00522974
CCL21	Chemokine (C-C motif) ligand 21	5.33146316	1.15E-07
CCL8	Chemokine (C-C motif) ligand 8	2.99896228	3.34E-08
CX3CL1	Chemokine (C-X3-C motif) ligand 1	6.11182873	1.29E-10
CXCL1	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	2.31960616	4.07E-03
CXCL14	Chemokine (C-X-C motif) ligand 14	6.78958378	6.10E-11
CXCL2	Chemokine (C-X-C motif) ligand 2	2.72980697	6.16E-04
CXCL3	Chemokine (C-X-C motif) ligand 3	2.53510313	0.00248838
CXCR1	chemokine (C-X-C motif) receptor 1	3.58551361	3.69E-04
CXCR2	chemokine (C-X-C motif) receptor 2	3.33910363	9.93E-04
CXCR4	Chemokine (C-X-C motif) receptor 4	3.13016199	0.00010446
C-type lectin domain superfamily			
CLEC10A	C-type lectin domain family 10, member A	5.00014548	5.91E-07
CLEC12A	C-type lectin domain family 12, member B	3.04420242	2.81E-03
CLEC14A	C-type lectin domain family 14, member A	5.39662379	5.14E-09
CLEC1A	C-type lectin domain family 1, member A	6.16968176	1.58E-09
CLEC2B	C-type lectin domain family 2, member B	6.56528689	1.08E-11
CLEC4E	C-type lectin domain family 4, member E	3.46059763	0.00059599
CLEC4G	C-type lectin domain family 4, member G	3.1939359	1.71E-03
Cytochrome C oxidases			
COX14	Cytochrome C Oxidase Assembly Homolog 14 (S. Cerevisiae)	2.1255235	0.0059006
COX17	COX17 cytochrome c oxidase assembly homolog (S. cerevisiae)	2.83606514	0.00039046
COX4I2	Cytochrome c oxidase subunit IV isoform 2 (lung)	4.25801031	1.96E-05
COX5A	Cytochrome c oxidase subunit Va	3.14715034	8.44E-05
COX5B	Cytochrome c oxidase subunit Vb	2.85479705	0.00030538
COX6A2	Cytochrome c oxidase subunit VIa polypeptide 2	2.9649563	0.0001942
COX6C	Cytochrome c oxidase subunit VIc	2.01928513	8.39E-03
COX7A1	Cytochrome c oxidase subunit VIIa polypeptide 1 (muscle)	4.26636426	4.22E-07
COX7B	Cytochrome c oxidase subunit VIIb	2.42362431	1.82E-03
COX7C	Cytochrome c oxidase subunit VIIc	2.85942034	2.97E-04
DEAD box polypeptides			
DDX11	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11	2.98065259	2.95E-04
DDX12P	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 12	2.71414394	0.00081472
DDX51	DEAD (Asp-Glu-Ala-Asp) box polypeptide 51	2.08732987	2.16E-03
DDX59	DEAD (Asp-Glu-Ala-Asp) box polypeptide 59	2.13479023	5.77E-03
Dual specificity phosphatases			
DUSP1	Dual specificity phosphatase 1	2.43103068	0.0017531
DUSP13	Dual specificity phosphatase 13	3.46011274	2.10E-05
DUSP16	Dual specificity phosphatase 16	1.99087977	0.00976306
DUSP23	Dual specificity phosphatase 23	2.30930355	0.00325827
DUSP26	Dual specificity phosphatase 26 (putative)	6.377995	9.28E-12
DUSP27	Dual specificity phosphatase 27 (putative)	1.806805	0.00680276
DUSP3	Dual specificity phosphatase 3	2.01902904	8.33E-03
DUSP8	Dual specificity phosphatase 8	3.52743617	1.30E-06
Translation initiation factors			
EIF2B5	Eukaryotic translation initiation factor 2B, subunit 5 epsilon, 82kDa	1.92643186	1.17E-02
EIF3C	Eukaryotic translation initiation factor 3, subunit C-like	3.57334047	1.19E-05
EIF3CL	Eukaryotic translation initiation factor 3, subunit C-like	3.6106506	1.00E-05
EIF3I	Eukaryotic translation initiation factor 3, subunit I	1.95292	1.05E-02
EIF3J	Eukaryotic translation initiation factor 3, subunit J	2.0069088	8.78E-03
EIF3K	Eukaryotic translation initiation factor 3, subunit K	2.4377475	1.72E-03
EIF4B	Eukaryotic translation initiation factor 4B	1.99480034	9.09E-03
EIF4E3	Eukaryotic translation initiation factor 4E family member 3	3.55252127	1.42E-05
EIF4EBP1	Eukaryotic translation initiation factor 4E binding protein 1	1.68729379	0.01131091
EIF4H	Eukaryotic translation initiation factor 4H	2.05500628	7.34E-03
Hemoglobin subunit family			
HBA1	Hemoglobin, alpha 1	7.68742708	5.06E-13
HBA2	Hemoglobin, alpha 2	8.24949419	2.43E-14
HBB	Hemoglobin, beta	12.4712308	2.10E-30
HBD	Hemoglobin, delta	4.87187541	9.68E-08
HBG1	Hemoglobin, gamma A	2.52014407	0.00774347
HBG2	Hemoglobin, gamma G	3.3275386	0.00023292
Troponins			
TNNC1	Troponin C type 1 (slow)	2.58311056	9.38E-04
TNNC2	Troponin C type 2 (fast)	3.12822565	8.99E-05
TNNI2	Troponin I type 2 (skeletal, fast)	2.90419108	2.48E-04
TNNI3	Troponin I type 3 (cardiac)	2.55649888	6.80E-03
TNNT1	Troponin T type 1 (skeletal, slow)	2.06872013	6.93E-03

Gene	Description	logFC	P Value
Human Leukocyte Antigens			
HLA-B	Major histocompatibility complex, class I, B	2.39862913	0.00042481
HLA-C	Major histocompatibility complex, class I, C	1.72053258	0.00971183
HLA-DMA	major histocompatibility complex, class II, DM alpha	5.18046634	1.23E-10
HLA-DMB	major histocompatibility complex, class II, DM beta	4.48987743	2.10E-07
HLA-DOA	Major histocompatibility complex, class II, DO alpha	5.07769502	2.71E-08
HLA-DPA1	Major histocompatibility complex, class II, DP alpha 1	7.73563038	5.89E-16
HLA-DPB1	Major histocompatibility complex, class II, DP beta 1	5.96668282	1.79E-12
HLA-DQA1	Major histocompatibility complex, class II, DQ alpha 1	4.91923047	8.79E-07
HLA-DQA2	Major histocompatibility complex, class II, DQ alpha 2	2.47505837	0.00897448
HLA-DQB1	Major histocompatibility complex, class II, DQ beta 1	5.12974277	3.14E-07
HLA-DQB2	Major histocompatibility complex, class II, DQ beta 2	5.68550507	5.82E-10
HLA-DRA	Major histocompatibility complex, class II, DR alpha	8.96711843	4.94E-16
HLA-DRB1	Major histocompatibility complex, class II, DR beta 1	8.13659001	5.97E-18
HLA-DRB5	Major histocompatibility complex, class II, DR beta 5	5.8702923	1.80E-10
HLA-DRB6	Major histocompatibility complex, class II, DR beta 4	5.81286494	2.57E-10
HLA-E	Major histocompatibility complex, class I, E	2.75504509	6.60E-05
HLA-F	Major histocompatibility complex, class I, F	3.58304928	9.18E-07
HLA-G	Major histocompatibility complex, class I, G	3.2867546	9.47E-06
HLA-L	Major histocompatibility complex, class I, C	2.28899192	0.0012913
Interleukins and receptors			
IL10RA	Interleukin 10 receptor, alpha	3.38160666	1.64E-04
IL12RB2	Interleukin 12 receptor, beta 2	6.12619943	1.99E-09
IL15RA	Interleukin 15 receptor, alpha	3.10050937	1.56E-04
IL17RB	Interleukin 17 receptor B	2.54681272	5.33E-03
IL18	Interleukin 18 (interferon-gamma-inducing factor)	5.86755258	1.95E-09
IL18RAP	Interleukin 18 receptor accessory protein	2.56031877	5.37E-03
IL1R2	Interleukin 1 receptor, type II	2.97712736	0.00056251
IL1RN	Interleukin 1 receptor antagonist	4.19829851	2.49E-05
IL2RG	Interleukin 2 receptor, gamma (severe combined immunodeficiency)	2.75906965	1.60E-03
IL33	Interleukin 33	4.5470801	2.03E-07
IL3RA	Interleukin 3 receptor, alpha (low affinity)	5.83202976	2.26E-10
IL6R	Interleukin 6 receptor	7.85786209	4.96E-15
IL8	Interleukin 8	3.67620268	1.08E-05
Leukocyte Ig-like receptor family			
LILRA2	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 2	2.69290824	0.00380202
LILRA5	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 5	4.42477924	1.37E-06
LILRA6	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 3	4.2608665	3.53E-06
LILRB1	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 1	2.74928397	8.38E-03
LILRB2	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 2	4.5113071	6.15E-06
LILRB3	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 3	3.38558934	0.00018842
LILRB4	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 4	4.31447795	2.55E-06
LILRB5	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 5	5.69412072	1.83E-08
Myosin family			
MYF6	Myogenic factor 6 (herculin)	7.6635439	9.26E-15
MYH1	Myosin, heavy chain 1, skeletal muscle, adult	7.04131852	2.14E-13
MYH11	Myosin, heavy chain 11, smooth muscle	7.51426217	3.23E-14
MYH14	Myosin, heavy chain 14	7.74623936	8.42E-15
MYH4	Myosin, heavy chain 4, skeletal muscle	1.76813628	0.0125641
MYH7	Myosin, heavy chain 7, cardiac muscle, beta	2.96055373	0.00018809
MYL12A	Myosin, light chain 12A, regulatory, non-sarcomeric	2.83211066	0.00032911
MYL2	Myosin, light chain 2, regulatory, cardiac, slow	6.83819325	5.32E-13
MYL3	Myosin, light chain 3, alkali; ventricular, skeletal, slow	5.19476784	4.93E-09
MYLK2	Myosin light chain kinase 2	7.56316434	3.54E-14
MYLK3	Myosin light chain kinase 3	4.95694399	1.26E-07
MYLK4	Myosin light chain kinase family, member 4	3.55573555	1.70E-05
MYO18A	Myosin XVIIIA	1.66587466	0.01214992
MYO1F	Myosin IF	4.23454019	3.65E-06
MYO1G	Myosin IG	2.2434837	8.48E-03
MYO5C	Myosin VC	3.13202188	1.60E-04
MYO7A	Myosin VIIA	3.14248982	2.70E-04
MYOC	Myocilin, trabecular meshwork inducible glucocorticoid response	4.56788326	4.67E-06
MYOM1	Myomesin 1, 185kDa	2.32363182	0.0026581
MYOM2	Myomesin (M-protein) 2, 165kDa	4.03571908	1.32E-06
MYOT	Myotilin	6.20318803	1.83E-11
MYOZ1	Myozenin 1	6.83759375	6.46E-13
MYOZ2	Myozenin 2	2.2757978	3.21E-03
MYOZ3	Myozenin 3	5.87218868	8.58E-10
MYPN	Myopalladin	2.09758356	0.00626124

Gene	Description	logFC	P Value
NADH dehysrogenase enzyme subunit genes			
NDUFA12	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 12	2.22047163	4.04E-03
NDUFA2	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2, 8kDa	2.12068275	5.91E-03
NDUFA3	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3, 9kDa	2.27921037	0.00336536
NDUFA4	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4, 9kDa	2.34467426	2.48E-03
NDUFA4L2	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4-like 2	1.66011784	1.32E-02
NDUFA5	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5, 13kDa	2.42638674	1.85E-03
NDUFA6	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 6, 14kDa	2.0700131	0.00703008
NDUFB1	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 1, 7kDa	2.19293298	4.58E-03
NDUFB10	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 10, 22kDa	2.05716028	7.35E-03
NDUFB3	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 3, 12kDa	2.10600554	0.00646738
NDUFB4	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 4, 15kDa	2.48851985	1.43E-03
NDUFB8	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 8, 19kDa	1.96306269	0.01122296
NDUFB9	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 9, 22kDa	3.04612433	1.31E-04
NDUFS3	NADH dehydrogenase (ubiquinone) Fe-S protein 3, 30kDa (NADH-coenzyme Q reductase)	1.95271668	0.00370649
Phosphoprotein phosphatases			
PPM1B	Protein phosphatase 1B (formerly 2C), magnesium-dependent, beta isoform	2.39963164	1.98E-03
PPM1J	Protein phosphatase 1J (PP2C domain containing)	3.37732176	2.09E-04
PPM1L	Protein phosphatase 1 (formerly 2C)-like	3.31572684	4.06E-05
PPP1R13B	Protein phosphatase 1, regulatory (inhibitor) subunit 13B	3.04678479	1.63E-04
PPP1R14A	Protein phosphatase 1, regulatory (inhibitor) subunit 14A	3.65379696	1.47E-05
PPP1R14C	Protein phosphatase 1, regulatory (inhibitor) subunit 14C	6.39977353	2.64E-14
PPP1R15A	Protein phosphatase 1, regulatory (inhibitor) subunit 15A	1.99053464	9.27E-03
PPP1R16B	Protein phosphatase 1, regulatory (inhibitor) subunit 16B	5.41114356	7.67E-08
PPP1R1A	Protein phosphatase 1, regulatory (inhibitor) subunit 1A	9.23059128	5.35E-18
PPP1R27	Dysferlin interacting protein 1	8.23788470	1.08E-15
PPP1R3A	Protein phosphatase 1, regulatory (inhibitor) subunit 3A	4.26313967	8.35E-07
PPP1R3B	Protein phosphatase 1, regulatory (inhibitor) subunit 3B	2.23168556	0.00381567
PPP1R3E	Protein phosphatase 1, regulatory (inhibitor) subunit 3E	3.46540528	1.36E-06
PPP2R5A	Protein phosphatase 2, regulatory subunit B\', alpha isoform	1.98693179	9.62E-03
PPP3CC	Protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform	2.10618229	6.16E-03
Zinc finger proteins			
ZNF331	Zinc finger protein 331	2.97575993	0.00019328
ZNF358	Zinc finger protein 358	1.80603215	0.00718559
ZNF366	Zinc finger protein 366	3.9257582	1.62E-05
ZNF367	Zinc finger protein 367	2.04682034	0.01143508
ZNF385A	Zinc finger protein 385A	3.60676209	1.26E-05
ZNF385C	Zinc finger protein 385C	3.36088271	1.88E-04
ZNF395	Zinc finger protein 395	3.42011827	1.75E-06
ZNF438	Zinc finger protein 438	2.15631895	0.00542706
ZNF511	Zinc finger protein 511	2.18256101	0.00135746
ZNF579	Zinc finger protein 579	2.28203487	0.00088703
ZNF593	Zinc finger protein 593	2.26192222	3.83E-03
ZNF672	Zinc finger protein 672	2.20348043	4.45E-03
ZNF676	Zinc finger protein 676	2.47603644	0.00897448
ZNF703	Zinc finger protein 703	2.33614072	2.58E-03
ZNF750	Zinc finger protein 750	4.3598752	3.66E-06
ZNF853	Zinc finger protein 853	2.02053324	0.00350575

Table S2-8: Protein families and their member genes found upregulated in Skeletal muscle tissue. *Gene names, their description, corresponding $\log_2$ fold change and P Values are shown. Fold change = value describing the average level of increase or decrease in gene expression.*

Overexpressed in SK			
miRNA	logFC	logCPM	PValue
hsa-miR-142-3p	13.09	14.0424	1.16E-24
hsa-miR-1323	12.76	13.71834	6.24E-24
hsa-miR-200c-5p	12.33	13.29039	5.74E-23
hsa-miR-3064-3p	12.26	13.21225	8.60E-23
hsa-miR-2964a-5p	12.25	13.20718	8.82E-23
hsa-miR-450a-5p	12.11	14.71468	1.69E-23
hsa-miR-2277-5p	11.14	12.09485	2.83E-20
hsa-miR-1262	10.60	11.56102	4.52E-19
hsa-miR-3074-3p	10.21	11.48107	2.41E-18
hsa-miR-3617-5p	10.14	11.10631	4.78E-18
hsa-miR-944	10.07	11.0318	7.03E-18
hsa-miR-1538	9.90	10.86742	1.66E-17
hsa-miR-376a-5p	9.86	13.21811	1.08E-18
hsa-miR-125b-2-3p	9.81	12.32	3.06E-18
hsa-miR-130b-5p	9.53	15.02387	3.00E-18
hsa-let-7i-5p	9.46	17.53917	3.28E-18
hsa-miR-212-5p	9.35	11.28087	7.86E-17
hsa-let-7f-2-3p	9.31	12.91795	1.75E-17
hsa-miR-1277-3p	9.07	10.35036	8.66E-16
hsa-miR-2355-3p	8.82	10.75718	1.19E-15
hsa-miR-197-5p	8.82	9.793997	4.44E-15
hsa-miR-2467-5p	8.79	9.765882	5.20E-15
hsa-miR-302d-3p	8.63	10.74662	2.73E-15
hsa-miR-15a-3p	8.58	13.08942	4.85E-16
hsa-miR-3620-3p	8.58	11.1088	1.63E-15
hsa-miR-27b-5p	8.52	13.2299	6.54E-16
hsa-miR-101-5p	8.33	10.268	1.52E-14
hsa-miR-1914-5p	8.24	9.223942	8.67E-14
hsa-let-7g-3p	8.21	11.76496	4.93E-15
hsa-miR-190a	8.15	10.0943	3.74E-14
hsa-miR-2682-5p	8.10	10.93546	1.64E-14
hsa-miR-296-5p	7.99	12.36449	1.05E-14
hsa-miR-376b-5p	7.51	11.4707	1.53E-13
hsa-miR-4762-5p	7.50	8.509241	3.65E-12
hsa-miR-3173-5p	7.43	8.736212	3.91E-12
hsa-miR-2116-3p	7.32	9.740801	1.46E-12
hsa-miR-182-5p	6.80	8.952029	2.17E-11
hsa-miR-196b-3p	6.78	7.821687	1.39E-10
hsa-miR-100-3p	6.74	10.5801	7.82E-12
hsa-miR-144-5p	6.73	7.776616	1.81E-10
hsa-miR-4750-3p	6.72	7.765837	1.87E-10
hsa-miR-148b-5p	6.67	7.706938	2.46E-10
hsa-miR-141-3p	6.64	7.703408	2.81E-10
hsa-miR-3605-5p	6.55	7.903095	3.17E-10
hsa-miR-520d-3p	6.52	7.581055	5.10E-10
hsa-miR-222-5p	6.32	11.00305	5.05E-11
hsa-miR-30a-5p	6.20	12.13525	7.29E-11
hsa-miR-195-3p	6.16	7.230134	3.11E-09
hsa-miR-1247-5p	6.12	7.182334	3.62E-09
hsa-let-7a-3p	6.12	13.24149	1.00E-10
hsa-miR-605	6.06	7.698992	2.76E-09
hsa-miR-125a-3p	6.04	10.38741	2.35E-10
hsa-miR-32-3p	6.03	8.208194	1.08E-09
hsa-miR-4803	6.01	7.093563	6.38E-09
hsa-miR-106a-5p	5.97	8.665307	9.77E-10
hsa-miR-147b	5.97	7.593464	4.64E-09
hsa-miR-610	5.93	7.012017	9.31E-09
hsa-miR-1305	5.82	6.91801	1.64E-08
hsa-miR-221-5p	5.82	10.86312	5.98E-10

miRNA	logFC	logCPM	PValue
hsa-miR-196b-5p	5.81	10.4843877	6.77E-10
hsa-miR-365a-5p	5.78	6.86110995	1.92E-08
hsa-miR-3614-3p	5.77	6.87190079	2.03E-08
hsa-miR-30b-3p	5.71	6.82142155	2.74E-08
hsa-miR-449b-5p	5.70	6.80133358	2.86E-08
hsa-miR-3691-5p	5.68	6.79889715	3.17E-08
hsa-miR-576-5p	5.64	7.48894893	1.23E-08
hsa-miR-302c-5p	5.53	6.64868037	6.66E-08
hsa-miR-3193	5.52	6.65537525	6.97E-08
hsa-miR-412	5.40	6.83073228	9.02E-08
hsa-miR-194-3p	5.38	6.51513875	1.37E-07
hsa-miR-105-3p	5.36	6.51513495	1.52E-07
hsa-miR-141-5p	5.29	6.45238805	2.08E-07
hsa-miR-219-5p	5.19	6.35387743	3.32E-07
hsa-miR-5003-3p	5.14	6.26991679	4.06E-07
hsa-miR-1914-3p	5.10	6.27502359	5.00E-07
hsa-miR-216a-3p	5.06	6.21230719	6.19E-07
hsa-let-7d-3p	4.98	10.5298418	3.40E-08
hsa-miR-4536-3p	4.85	6.07057345	1.61E-06
hsa-miR-4769-3p	4.79	6.01516482	2.07E-06
hsa-miR-1304-5p	4.78	5.96741286	2.14E-06
hsa-miR-15b-3p	4.61	9.25315184	2.41E-07
hsa-miR-876-5p	4.58	6.83839426	1.20E-06
hsa-miR-483-5p	4.58	9.22963738	2.87E-07
hsa-miR-1193	4.48	5.68178561	8.32E-06
hsa-miR-411-5p	4.48	8.02546065	7.14E-07
hsa-miR-224-3p	4.45	9.33261596	4.82E-07
hsa-miR-20a-3p	4.42	5.98272046	8.57E-06
hsa-miR-1273c	4.40	6.32899182	4.45E-06
hsa-miR-942	4.34	5.88348215	1.21E-05
hsa-miR-26a-2-3p	4.21	7.10612087	3.91E-06
hsa-miR-338-5p	4.07	5.34264943	5.14E-05
hsa-miR-4746-5p	4.03	5.30845928	5.75E-05
hsa-miR-323a-5p	4.00	5.56829695	5.38E-05
hsa-miR-1302	3.90	5.18403758	0.00010345
hsa-miR-636	3.82	5.68141933	9.22E-05
hsa-miR-676-3p	3.79	5.12111523	0.00016062
hsa-miR-3192	3.63	4.977693	0.00031586
hsa-miR-299-3p	3.58	8.97151852	2.50E-05
hsa-miR-148b-3p	3.58	10.3116991	2.19E-05
hsa-miR-3188	3.58	4.96146078	0.00036433
hsa-miR-203a	3.58	4.92123844	0.00036433
hsa-let-7b-3p	3.54	11.7190926	2.45E-05
hsa-let-7e-3p	3.52	9.5455549	3.03E-05
hsa-miR-4502	3.51	4.94129445	0.00048874
hsa-miR-520f	3.49	4.85368733	0.00052693
hsa-miR-493-3p	3.44	10.0041262	4.16E-05
hsa-miR-4798-5p	3.42	4.8596608	0.0007174
hsa-miR-105-5p	3.40	4.83349859	0.00077646
hsa-miR-376b-3p	3.40	9.15375734	5.38E-05
hsa-miR-523-5p	3.34	4.75664757	0.00091184
hsa-miR-548w	3.34	4.75664757	0.00091184
hsa-miR-4761-5p	3.33	4.72991857	0.00098939
hsa-miR-146a-5p	3.30	6.1249587	0.00023464
hsa-miR-3177-3p	3.27	4.94791704	0.00097859
hsa-miR-373-3p	3.25	4.69590584	0.00127066
hsa-miR-490-3p	3.18	4.61892687	0.00164558
hsa-miR-25-3p	3.10	12.5013113	0.00015943
hsa-miR-3939	3.08	4.54866434	0.00235552

miRNA	logFC	logCPM	PValue
hsa-miR-4763-5p	3.02	4.51355739	0.00311738
hsa-miR-4798-3p	2.99	4.52364546	0.00311738
hsa-miR-218-5p	2.98	8.55490863	0.00032927
hsa-miR-520a-3p	2.97	4.4538943	0.00343037
hsa-miR-345-3p	2.96	4.45860896	0.00377922
hsa-miR-2116-5p	2.96	5.31316159	0.00172101
hsa-miR-663b	2.96	4.38321568	0.00343037
hsa-miR-154-5p	2.95	9.54978524	0.00032332
hsa-miR-1245a	2.94	4.4340613	0.00377922
hsa-miR-193a-3p	2.94	9.35067821	0.00034592
hsa-miR-4747-5p	2.93	4.4061253	0.00416855
hsa-miR-411-3p	2.93	8.50995357	0.00039916
hsa-miR-20b-5p	2.92	4.41492646	0.00416855
hsa-miR-139-5p	2.92	11.9744184	0.0003453
hsa-miR-7-2-3p	2.92	5.44126086	0.00139124
hsa-miR-1273d	2.91	4.37987168	0.00416855
hsa-miR-139-3p	2.86	4.39546551	0.00509054
hsa-miR-153	2.86	4.41028034	0.00509054
hsa-miR-330-5p	2.86	5.04542207	0.00256959
hsa-miR-200b-5p	2.79	4.34432316	0.00624849
hsa-miR-1295a	2.78	4.28965581	0.00693664
hsa-miR-96-5p	2.71	4.29087975	0.00858426
hsa-miR-23b-5p	2.69	5.35855039	0.00340857
hsa-miR-758-5p	2.68	5.79875871	0.00227382
hsa-miR-454-3p	2.67	7.00824946	0.00138498
hsa-miR-1913	2.66	4.24754557	0.00956994
hsa-miR-184	2.66	6.9224215	0.00150927
hsa-miR-4473	2.64	4.43716616	0.00905379
hsa-miR-652-5p	2.63	7.20621098	0.00155152

miRNA	logFC	logCPM	PValue
hsa-miR-302d-5p	2.60	4.16630536	0.01194682
hsa-miR-19a-5p	2.57	4.11167819	0.01337883
hsa-miR-204-5p	2.56	7.18688582	0.00206467
hsa-miR-4789-3p	2.56	4.41551407	0.01197636
hsa-miR-486-3p	2.54	4.59164863	0.00779769
hsa-miR-519e-5p	2.53	4.13879809	0.01500604
hsa-miR-659-3p	2.52	4.11334227	0.01500604
hsa-miR-4421	2.39	4.03283823	0.02138221
hsa-miR-708-3p	2.39	4.81472546	0.00962026
hsa-miR-146b-3p	2.37	4.83249917	0.01273945
hsa-miR-421	2.22	11.2510407	0.00526125
hsa-miR-627	2.20	5.11654294	0.01398088
hsa-miR-579	2.12	5.12919694	0.01789154
hsa-miR-589-5p	2.10	4.07892531	0.04539914
hsa-miR-106b-3p	2.09	8.49595139	0.00894057
hsa-miR-381-5p	2.06	4.42427274	0.03458077
hsa-miR-423-5p	2.01	7.33578453	0.01319258
hsa-miR-196a-5p	1.94	10.7862026	0.01347665
hsa-miR-137	1.81	7.56360736	0.02279356
hsa-miR-192-3p	1.75	5.31596561	0.04365033
hsa-miR-33a-3p	1.75	7.99642901	0.02747601
hsa-miR-337-5p	1.67	8.15303818	0.0340222
hsa-miR-181b-3p	1.58	5.77642596	0.06090176
hsa-miR-379-3p	1.52	5.89138686	0.06715555
hsa-miR-1343	1.52	5.59861045	0.07189847
hsa-miR-629-5p	1.51	5.36714564	0.07691305
hsa-miR-23b-3p	1.44	12.8353726	0.06233563
hsa-let-7i-3p	1.40	9.47544566	0.07208301

Table S2-9: **List of miRNA differentially upregulated in human skeletal muscle**. List has been arranged in a descending order of log Fold Change (log $_2$ FC). Corresponding logCPM and P-values are also shown. Fold change = value describing the average level of increase or decrease in gene expression.

Under expressed in SK			
miRNA	logFC	logCPM	PValue
hsa-miR-103a-3p	-14.6378	14.89372	4.23E-28
hsa-miR-424-5p	-13.6021	14.05091	9.53E-26
hsa-miR-199b-3p	-12.9183	14.27051	6.85E-25
hsa-miR-214-3p	-12.8475	13.10246	4.50E-24
hsa-miR-125a-5p	-12.5615	14.81076	1.32E-24
hsa-miR-574-3p	-12.553	13.70111	6.75E-24
hsa-miR-221-3p	-12.1973	12.68465	1.12E-22
hsa-miR-199a-3p	-11.9607	14.27068	3.16E-23
hsa-miR-29b-3p	-11.8027	12.28981	8.63E-22
hsa-miR-181a-5p	-11.7413	12.39434	1.25E-21
hsa-miR-21-5p	-11.5562	11.80955	3.60E-21
hsa-miR-484	-11.0298	11.28145	5.48E-20
hsa-let-7f-5p	-10.8438	13.32943	8.41E-21
hsa-miR-100-5p	-10.8334	14.98165	4.12E-21
hsa-miR-532-5p	-10.688	12.81028	2.75E-20
hsa-miR-29a-3p	-10.6611	14.49543	1.22E-20
hsa-let-7d-5p	-10.2599	10.90983	2.66E-18
hsa-miR-199b-5p	-10.2004	11.43606	8.19E-19
hsa-miR-149-5p	-10.0927	10.53449	7.33E-18
hsa-miR-22-5p	-9.81518	10.06154	2.91E-17
hsa-miR-133b	-9.66748	10.59592	2.99E-17
hsa-let-7g-5p	-9.64957	12.83568	2.76E-18
hsa-miR-31-5p	-9.45395	9.892976	1.98E-16
hsa-miR-365a-3p	-9.40439	9.647139	2.42E-16
hsa-miR-210	-9.31973	9.96478	3.41E-16
hsa-miR-365b-3p	-9.21124	9.647425	6.91E-16
hsa-miR-214-5p	-9.07389	10.51858	1.78E-16
hsa-miR-93-5p	-8.97678	10.00085	6.34E-16
hsa-miR-193a-5p	-8.88846	9.361529	2.97E-15
hsa-miR-455-3p	-8.88171	9.118215	3.57E-15
hsa-miR-19a-3p	-8.83296	9.070651	4.59E-15
hsa-miR-185-5p	-8.83259	9.068597	4.61E-15
hsa-miR-532-3p	-8.79932	9.231149	5.75E-15
hsa-miR-22-3p	-8.76633	9.19978	6.83E-15
hsa-miR-23a-3p	-8.56918	12.93206	4.95E-16
hsa-miR-376a-3p	-8.29795	8.722593	7.55E-14
hsa-miR-503-5p	-8.25974	9.278344	2.53E-14
hsa-miR-145-5p	-8.23165	15.18332	1.98E-15
hsa-miR-301a-3p	-8.22082	8.446619	1.06E-13
hsa-miR-30c-5p	-8.19711	8.423203	1.20E-13
hsa-miR-362-5p	-7.97849	8.199896	3.66E-13
hsa-miR-542-3p	-7.94824	8.863566	2.08E-13
hsa-miR-10b-5p	-7.79731	11.60116	2.83E-14
hsa-miR-29c-3p	-7.73273	8.190452	1.12E-12
hsa-miR-299-5p	-7.54233	8.55231	1.00E-12
hsa-miR-543	-7.53372	7.942263	3.75E-12
hsa-miR-675-5p	-7.51507	7.926794	4.13E-12
hsa-miR-19b-3p	-7.43047	9.810735	4.40E-13
hsa-miR-181a-3p	-7.3854	8.72693	1.63E-12
hsa-miR-542-5p	-7.2943	7.497272	1.20E-11
hsa-miR-132-5p	-7.24741	7.648921	1.60E-11
hsa-miR-181a-2-3p	-7.24348	7.973042	7.98E-12
hsa-miR-342-3p	-7.24226	7.855416	5.80E-12
hsa-miR-127-3p	-7.23197	7.434752	1.63E-11
hsa-miR-362-3p	-7.19029	7.391078	2.04E-11
hsa-let-7a-5p	-7.17063	13.25172	4.84E-13
hsa-miR-191-5p	-7.16077	12.37941	5.54E-13
hsa-miR-21-3p	-7.1478	8.268272	8.47E-12
hsa-miR-222-3p	-7.05941	7.45586	4.17E-11

miRNA	logFC	logCPM	PValue
hsa-miR-130b-3p	-7.055971	9.85533969	2.26E-12
hsa-miR-27a-3p	-7.0508714	10.9060883	1.34E-12
hsa-miR-382-5p	-7.0421408	7.90579294	1.13E-11
hsa-miR-30e-3p	-6.8460587	7.03513571	1.15E-10
hsa-miR-505-3p	-6.8268334	7.82152517	3.78E-11
hsa-miR-181b-5p	-6.8110374	8.23529453	1.93E-11
hsa-miR-485-3p	-6.8077318	6.99073803	1.41E-10
hsa-miR-130a-3p	-6.6998161	12.015036	5.65E-12
hsa-miR-152	-6.6524753	9.2207852	1.72E-11
hsa-miR-744-5p	-6.624483	6.80221158	3.48E-10
hsa-miR-323a-3p	-6.5883822	6.9678963	4.47E-10
hsa-miR-99b-3p	-6.5403172	8.27190298	5.05E-11
hsa-miR-425-5p	-6.4985476	9.4327327	3.26E-11
hsa-miR-708-5p	-6.4960262	6.91657823	5.73E-10
hsa-miR-574-5p	-6.4626255	8.39342254	7.47E-11
hsa-miR-625-3p	-6.4535454	6.62402485	8.29E-10
hsa-miR-339-3p	-6.4270535	6.59472392	9.38E-10
hsa-miR-590-5p	-6.4194738	8.57693977	7.18E-11
hsa-miR-7-1-3p	-6.3907748	6.80488681	9.85E-10
hsa-miR-409-3p	-6.3894432	8.46619824	1.25E-10
hsa-miR-30a-3p	-6.3488422	7.23176042	3.13E-10
hsa-miR-654-3p	-6.2808504	7.75635186	2.10E-10
hsa-miR-494	-6.2617088	6.66986183	1.84E-09
hsa-miR-339-5p	-6.2509893	8.34826911	1.47E-10
hsa-miR-151a-3p	-6.2309262	6.39119778	2.48E-09
hsa-let-7e-5p	-6.1556282	10.9548128	9.75E-11
hsa-miR-671-5p	-6.0754567	6.43364015	5.67E-09
hsa-miR-423-3p	-6.0554389	8.3395886	5.10E-10
hsa-miR-15b-5p	-6.0334547	11.0433022	1.81E-10
hsa-miR-345-5p	-6.0221161	6.59841745	2.70E-09
hsa-miR-500a-5p	-5.9236027	6.89611443	3.52E-09
hsa-miR-134	-5.8520218	8.88424154	8.41E-10
hsa-miR-331-3p	-5.8475922	8.26468832	1.49E-09
hsa-miR-487b	-5.8216223	9.68088168	5.95E-10
hsa-miR-381-3p	-5.7613832	6.11014749	2.67E-08
hsa-miR-20a-5p	-5.7611216	9.25911108	1.01E-09
hsa-miR-1271-5p	-5.751557	6.71717493	8.23E-09
hsa-miR-128	-5.6989648	7.99920206	2.77E-09
hsa-miR-500a-3p	-5.6750099	5.81455152	3.84E-08
hsa-miR-192-5p	-5.6510271	7.04292655	6.71E-09
hsa-miR-539-5p	-5.650085	5.78098778	4.26E-08
hsa-miR-502-5p	-5.6390135	5.77370499	4.56E-08
hsa-miR-329	-5.6206063	8.21670885	2.63E-09
hsa-miR-598	-5.5954944	7.43327159	8.51E-09
hsa-miR-335-3p	-5.56325	6.64877745	2.33E-08
hsa-miR-29b-1-5p	-5.5367974	6.09949833	2.90E-08
hsa-miR-495-3p	-5.5163621	9.78873902	2.85E-09
hsa-miR-615-3p	-5.4657113	5.84539873	9.06E-08
hsa-miR-30d-5p	-5.4247997	5.97601562	4.94E-08
hsa-miR-652-3p	-5.4144901	5.54205898	1.35E-07
hsa-miR-379-5p	-5.375695	6.66817342	4.02E-08
hsa-miR-151a-5p	-5.3723332	8.17841438	1.01E-08
hsa-miR-431-5p	-5.2475439	5.36913961	2.90E-07
hsa-miR-17-5p	-5.2399584	10.4158833	9.25E-09
hsa-miR-1307-3p	-5.2155449	5.54493107	3.61E-07
hsa-miR-424-3p	-5.1695924	5.2969699	4.19E-07
hsa-miR-452-5p	-5.163788	5.27929662	4.39E-07
hsa-miR-671-3p	-5.1168698	5.23055308	5.58E-07
hsa-miR-26a-5p	-5.1030879	12.9778491	1.49E-08

miRNA	logFC	logCPM	PValue
hsa-miR-361-3p	-5.0824249	5.40581306	6.74E-07
hsa-miR-106b-5p	-5.0711918	6.14209207	2.51E-07
hsa-miR-487a	-5.0645608	5.60324902	2.85E-07
hsa-miR-224-5p	-5.053452	5.16401764	7.52E-07
hsa-miR-548t-3p	-4.9575494	5.07009415	1.14E-06
hsa-miR-2355-5p	-4.8991562	5.86108705	4.50E-07
hsa-miR-425-3p	-4.8555008	5.81625046	5.68E-07
hsa-miR-26b-5p	-4.8312268	6.29346975	4.22E-07
hsa-miR-136-3p	-4.7532615	6.38478943	4.79E-07
hsa-miR-432-5p	-4.72668	5.0419412	3.49E-06
hsa-miR-27b-3p	-4.6962091	8.84780086	1.42E-07
hsa-miR-378a-3p	-4.6763954	4.78855633	4.12E-06
hsa-miR-433	-4.6546469	4.76676912	4.69E-06
hsa-miR-382-3p	-4.6481578	4.96430551	5.14E-06
hsa-miR-17-3p	-4.6273339	4.75418577	5.34E-06
hsa-miR-1226-3p	-4.5740976	4.6810644	6.53E-06
hsa-miR-361-5p	-4.5409869	8.7815838	2.79E-07
hsa-miR-145-3p	-4.458367	4.77718028	1.19E-05
hsa-miR-1249	-4.4394337	4.80177589	1.05E-05
hsa-miR-545-3p	-4.4311671	4.79793793	1.12E-05
hsa-miR-323b-3p	-4.3828559	6.65519226	1.77E-06
hsa-miR-92b-3p	-4.3592284	8.30974117	8.23E-07
hsa-miR-1185-1-3p	-4.3508702	4.66773627	1.89E-05
hsa-miR-376c-3p	-4.3355093	11.7484071	6.27E-07
hsa-miR-6511a-3p	-4.3326548	4.64680402	2.17E-05
hsa-miR-127-5p	-4.2965974	4.66505611	2.02E-05
hsa-miR-455-5p	-4.2912925	4.42011529	2.28E-05
hsa-miR-30e-5p	-4.2628203	4.39299913	2.68E-05
hsa-miR-16-5p	-4.2459127	11.9113873	9.54E-07
hsa-miR-188-5p	-4.2142584	6.81249148	2.37E-06
hsa-miR-93-3p	-4.2021083	5.13843527	1.24E-05
hsa-miR-31-3p	-4.1420359	5.40521672	7.96E-06
hsa-miR-18a-5p	-4.1379928	7.68993601	2.72E-06
hsa-miR-103a-2-5p	-4.1329335	4.27635657	4.44E-05
hsa-miR-1248	-4.1171014	4.25266054	4.85E-05
hsa-miR-126-5p	-4.0703964	4.62626495	2.67E-05
hsa-miR-7-5p	-4.0338728	4.37900617	7.75E-05
hsa-miR-380-3p	-4.0164108	4.98528697	2.61E-05
hsa-miR-770-5p	-3.9867942	4.13066174	8.40E-05
hsa-miR-28-3p	-3.937177	8.31677045	5.59E-06
hsa-miR-590-3p	-3.925781	4.08996044	0.00011239
hsa-miR-126-3p	-3.9139226	5.52180678	2.58E-05
hsa-miR-887	-3.8946874	4.24175159	0.00014032
hsa-miR-324-5p	-3.8868115	7.65742652	7.06E-06
hsa-let-7f-1-3p	-3.8802277	6.60838703	1.44E-05
hsa-miR-370	-3.8789082	5.58865001	1.96E-05
hsa-miR-16-2-3p	-3.8655822	6.01450126	1.76E-05
hsa-miR-940	-3.8357517	5.11585972	5.01E-05
hsa-miR-625-5p	-3.8282121	4.39910557	7.75E-05
hsa-miR-3605-3p	-3.7858177	4.14731811	0.00020146
hsa-miR-99b-5p	-3.7346565	10.0214088	1.05E-05
hsa-miR-1287	-3.7048063	4.56388674	8.34E-05
hsa-miR-193b-5p	-3.703499	4.66834646	0.0001075
hsa-miR-194-5p	-3.6919681	4.26494843	0.00012853
hsa-miR-143-3p	-3.6919001	10.120633	1.28E-05
hsa-miR-502-3p	-3.5841259	7.7333962	3.07E-05
hsa-miR-206	-3.541066	13.5423221	2.35E-05
hsa-miR-320a	-3.4863274	6.86131347	5.62E-05
hsa-miR-548aa	-3.4579727	5.18820042	0.00012011

miRNA	logFC	logCPM	PValue
hsa-miR-1	-3.4265228	11.9033279	3.93E-05
hsa-miR-193b-3p	-3.3921572	11.8054865	4.58E-05
hsa-miR-3607-5p	-3.289415	4.20414114	0.000868
hsa-miR-501-5p	-3.2787436	7.44669119	0.0001156
hsa-miR-133a	-3.2542741	8.90616494	9.27E-05
hsa-miR-409-5p	-3.2491319	5.1992559	0.00028542
hsa-miR-143-5p	-3.2367312	4.75032271	0.00050462
hsa-miR-660-3p	-3.2253454	7.03522196	0.00015064
hsa-miR-337-3p	-3.1942916	10.2970316	0.00011249
hsa-miR-493-5p	-3.1438949	6.1515852	0.00026298
hsa-miR-140-5p	-3.106736	9.62487339	0.00016609
hsa-miR-186-5p	-3.0907646	6.01221161	0.00036704
hsa-miR-483-3p	-3.0147345	7.65036882	0.00029586
hsa-miR-101-3p	-2.9899959	4.05764733	0.00183124
hsa-miR-19b-1-5p	-2.959788	4.57893988	0.00112516
hsa-miR-330-3p	-2.9500996	5.67380189	0.00059032
hsa-miR-328	-2.9377425	7.52053437	0.00041354
hsa-miR-107	-2.8886188	11.5970033	0.00039004
hsa-miR-656	-2.8857062	4.27496898	0.00225983
hsa-miR-301b	-2.8569973	4.05203064	0.00343037
hsa-miR-431-3p	-2.8508505	3.96613811	0.00273745
hsa-miR-92a-3p	-2.7668653	11.241357	0.00063993
hsa-miR-181c-5p	-2.7257206	5.07700749	0.00169959
hsa-miR-132-3p	-2.639869	6.03614423	0.00182458
hsa-miR-1306-5p	-2.6296418	5.93162587	0.00197068
hsa-miR-26b-3p	-2.6205074	5.17212599	0.00282973
hsa-miR-4521	-2.6083779	7.3635329	0.00149038
hsa-miR-138-5p	-2.5986238	10.9129822	0.00124906
hsa-miR-27a-5p	-2.5755956	5.10251335	0.00293696
hsa-miR-369-3p	-2.5406885	4.59876178	0.00434414
hsa-miR-15a-5p	-2.523362	8.32693437	0.00184138
hsa-miR-154-3p	-2.4671908	5.42012385	0.00404492
hsa-miR-135b-5p	-2.4644771	3.95611485	0.00976931
hsa-miR-874	-2.4624864	3.95558918	0.00976931
hsa-miR-454-5p	-2.2948149	5.05160827	0.00828909
hsa-miR-378a-5p	-2.2900303	7.02547541	0.00480253
hsa-miR-29a-5p	-2.2659155	3.90609919	0.01327707
hsa-miR-28-5p	-2.2585286	6.47137896	0.00589355
hsa-miR-766-3p	-2.1727815	6.52136603	0.00788287
hsa-miR-99a-5p	-2.1713615	12.5748641	0.00610864
hsa-miR-660-5p	-2.113868	11.2272316	0.00751843
hsa-miR-377-3p	-2.0197397	4.15870841	0.02910462
hsa-miR-3940-3p	-1.992829	3.99752043	0.02750017
hsa-miR-758-3p	-1.9156997	3.89403246	0.03697784
hsa-miR-34b-5p	-1.803158	6.53199762	0.02576084
hsa-miR-33b-3p	-1.7872421	4.03850257	0.04931927
hsa-miR-374b-3p	-1.7866574	4.22862524	0.04856718
hsa-miR-24-3p	-1.6240624	15.9493567	0.03679896
hsa-let-7b-5p	-1.5813615	11.3879223	0.04196731
hsa-miR-655	-1.5669858	4.82168213	0.06317489
hsa-miR-98-5p	-1.4983527	9.20212116	0.05423931
hsa-miR-1301	-1.4961872	5.39874268	0.07030605
hsa-miR-140-3p	-1.4466846	7.02179492	0.06757599
hsa-miR-664a-3p	-1.4343867	8.51871513	0.06613664
hsa-miR-335-5p	-1.4130192	7.48167157	0.07267049

Table S2-10: List of miRNA differentially down regulated in human skeletal muscle. List has been arranged in a ascending order of log Fold Change (log $_2$ FC). Corresponding logCPM and P-values are also shown. Fold change = value describing the average level of increase or decrease in gene expression.

Term	Count	PValue	Benjamini
GO:0006259~DNA metabolic process	14	9.67E-05	0.09834693
GO:0051258~protein polymerization	5	4.79E-04	0.22619436
GO:0016052~carbohydrate catabolic process	6	0.0013282	0.37751655
GO:0005996~monosaccharide metabolic process	8	0.00135934	0.30501831
GO:0034621~cellular macromolecular complex subunit organization	10	0.00138651	0.25689515
GO:0019320~hexose catabolic process	5	0.00174476	0.26759442
GO:0046365~monosaccharide catabolic process	5	0.00193882	0.2566963
GO:0006260~DNA replication	7	0.00298152	0.3292604
GO:0046164~alcohol catabolic process	5	0.0031367	0.3116826
GO:0019318~hexose metabolic process	7	0.00313992	0.2857341
GO:0044275~cellular carbohydrate catabolic process	5	0.00373184	0.30488981
GO:0048589~developmental growth	5	0.00405629	0.30401117

Table S2-11: Genes found enriched with Ago2 pull down in myoblasts but which are under expressed in myoblasts.

ACTB	CCDC85B	EIF6	GOT2	LDHAP3	PGF	RPS6KA1	TDP1
ACTG1	CCNE1	ELMOD1	H3F3C	LGALS1	PLCXD1	S100A2	TK1
AGPAT9	CDC20	ENO2	HMGA1	MANBAL	POLA2	SALL2	TMEM106C
ANAPC15	CDCA8	ETV4	HMGB1P1	MAPKAPK3	POTEF	SECTM1	TMEM246
ARID3B	CDKN2C	EYA2	HOXA11	MPP2	POTEKP	SEMA3F	TMSB4X
ATOH8	CFL1	FANCE	HOXA9	MSX2	PPIA	SEMA7A	TMSB4XP8
AURKA	CLDN4	FBL	HPCAL1	MT1E	PRELP	SH3BGRL3	TPI1
AURKB	CLIC1	FIBCD1	HSPB7	MUTYH	PRPS1	SLC14A1	TSPAN15
BCL2L12	COL9A2	FLT4	IER2	MYBL2	PSG4	SLC37A4	TUBA1B
BDH1	COLEC12	FOSL1	IER5	MYF5	PSG5	SLC38A5	TUBA1C
BEX4	CPNE2	FUCA1	IGSF1	MYF6	PTMA	SMAP2	TUBA4A
C17orf53	DCTPP1	GALE	KIFC1	NGF	PTTG1	SNRPA	TUBB
C19orf10	DOLPP1	GALNT15	KRT34	NUAK2	RAB34	SNRPG	TUBB4B
C1orf51	DRAP1	GAPDH	KRTAP1-1	NUDT1	RANGAP1	STMN1	
C1orf85	DUSP2	GINS1	KRTAP2-3	PAFAH1B3	RASD2	STRA13	
CALM3	DUSP5	GINS2	LAPTM5	PAX7	REEP4	TAGLN2	
CCDC167	EDN1	GINS3	LCE3A	PGAM1	RFC2	TARBP2	

Table S2-12: Genes found enriched with Ago2 pull down in myoblasts but which are under expressed in myotubes

Term	Count	PValue	Benjamini
GO:0003012~muscle system process	28	6.34E-26	8.20E-23
GO:0006936~muscle contraction	27	1.24E-25	8.01E-23
GO:0006941~striated muscle contraction	15	4.54E-18	1.96E-15
GO:0007517~muscle organ development	23	4.59E-17	1.48E-14
GO:0014706~striated muscle tissue development	15	6.37E-12	1.65E-09
GO:0060537~muscle tissue development	15	1.26E-11	2.71E-09
GO:0044057~regulation of system process	20	1.00E-10	1.85E-08
GO:0055001~muscle cell development	11	1.07E-10	1.73E-08
GO:0055002~striated muscle cell development	10	1.28E-09	1.84E-07
GO:0042692~muscle cell differentiation	12	2.02E-08	2.62E-06
GO:0006937~regulation of muscle contraction	10	2.52E-08	2.97E-06
GO:0030239~myofibril assembly	7	3.23E-08	3.48E-06
GO:0051146~striated muscle cell differentiation	10	1.49E-07	1.48E-05
GO:0055008~cardiac muscle tissue morphogenesis	7	1.68E-07	1.55E-05

Table S2-13: GO terms of Biological processes for genes enriched with Ago2 pull down in myoblasts but which are under expressed in myoblasts. *The number of genes belonging to each term (gene count), enrichment (p-value) and false discovery rate (benjamini value) are shown in the corresponding columns.*

Term	Count	PValue	Benjamini
GO:0006259~DNA metabolic process	14	9.67E-05	0.09834693
GO:0051258~protein polymerization	5	4.79E-04	0.22619436
GO:0016052~carbohydrate catabolic process	6	0.0013282	0.37751655
GO:0005996~monosaccharide metabolic process	8	0.00135934	0.30501831
GO:0034621~cellular macromolecular complex subunit organization	10	0.00138651	0.25689515
GO:0019320~hexose catabolic process	5	0.00174476	0.26759442
GO:0046365~monosaccharide catabolic process	5	0.00193882	0.2566963
GO:0006260~DNA replication	7	0.00298152	0.3292604
GO:0046164~alcohol catabolic process	5	0.0031367	0.3116826
GO:0019318~hexose metabolic process	7	0.00313992	0.2857341
GO:0044275~cellular carbohydrate catabolic process	5	0.00373184	0.30488981
GO:0048589~developmental growth	5	0.00405629	0.30401117

Table S2-14: GO terms of Biological processes for genes enriched with Ago2 pull down in myoblasts but which are under expressed in myotubes. *The number of genes belonging to each term (gene count), enrichment (p-value) and false discovery rate (benjamini value) are shown in the corresponding columns.*

Chapter 3

Supplementary Information

12H- Over expressed		12H- Under expressed		24H-Over expressed		24H-Under expressed	
Genes	log FC	Genes	log FC	Genes	log FC	Genes	log FC
LIMCH1	7.63755928	SNORA20	-6.195360662	C3	8.26304525	SNRPGP10	-7.377975255
PAFAH2	6.696648394	SNORD114-25	-6.065659775	REP15	7.47957982	MSC	-7.345020789
IL33	6.688046426	SNORD115-23	-6.051983183	RNFT2	7.203055436	CHAF1B	-7.29413319
C3orf52	6.546412694	SNORD12C	-6.030634131	MRPL40	6.911952776	ZBTB26	-7.268000067
EID2	6.498112037	SNORD116-13	-5.701695305	ZNF283	6.911952776	PEX11A	-7.232402828
FRMPD3	6.481645658	SNORD114-3	-5.524574614	ZNF594	6.774578908	TRPV4	-7.050173427
EDN1	6.464989166	MT-TM	-5.514878124	AAMDC	6.763449406	PTPN3	-7.039913954
TLR4	6.360797672	SNORD116-27	-5.429929935	KCNG1	6.74092947	C12orf5	-6.944139784
CCNE2	6.286906591	SNORA7B	-5.360311442	CXCL5	6.732689606	VWA7	-6.92196474
OSTCP1	6.275842827	SNORD114-23	-5.343508077	LNP1	6.706476411	AOC3	-6.76887044
GDAP2	6.267826967	SNORD114-9	-5.112762496	ACTR3B	6.694806761	FAM126B	-6.756392607
MED6	6.228893631	SNORD99	-4.920576582	LIMCH1	6.694806761	ZNF696	-6.756392607
		SNORD113-9	-4.79846868	NDUFAF7	6.694806761	MRPS22	-6.718298216
		SNORD114-26	-4.677345518	LRRC23	6.689827209	CEP70	-6.695313289
				CBR3	6.647160682	LRRC66	-6.679170599
				RGAG1	6.63499917		

Table S3-1: List of genes influenced by hsa-miR-139 over expression. *Genes over and under expressed after 12 and 24 hours of miRNA over expression have been listed.*

12H- Over expressed		12H-Under expressed		24H-Over expressed		24H-Under expressed	
Genes	log FC	Genes	log FC	Genes	log FC	Genes	log FC
LIMCH1	7.618074873	SNORD116-12	-6.359999845	ESM1	8.157845403	VWA7	-6.634679311
GDAP2	6.861940003	SNORD115-23	-6.086210197	ALG9	7.268234069	AOC3	-6.481601477
EDN1	6.824176597	TTLL7-IT1	-5.769018222	ZNF548	6.68382633	OGFOD2	-6.418111626
TMEM132D	6.772236767	DICER1-AS1	-5.660710416	GABARAPL3	6.609913323	LRRC66	-6.391912128
FRMPD3	6.732016716	MIR2355	-5.567800302	C3	6.563676782	MYLK2	-6.131874172
IL33	6.483028136	C5AR1	-5.465116885	PPL	6.563676782	MIR431	-6.09986082
PAAF1	6.480133761	SNORD12C	-5.204505699	ARHGEF3	6.547929339	RGS14	-5.891253231
FOSB	6.463853491	MT-TM	-5.012109709	CCDC74A	6.547929339	MYO5C	-5.814429116
LRRC4	6.447387406	SNORD70	-4.999149219	ABCA6	6.53200811	HDHD3	-5.754001304
PTGER3	6.290069875	SNORD114-3	-4.880888359	TRMT61B	6.515909215	TTLL7-IT1	-5.73328285
TLR4	6.252652082	SNORD93	-4.875929781	TGFBR3	6.483162253	CST6	-5.690931383
TMEM128	6.233572852	SNORD117	-4.83082069	IDI1	6.476231074	LAPTM5	-5.669280114
ZNF345	6.233572852	SNORD116-29	-4.574921569	NOD1	6.432604474	SNORA33	-5.669280114
PPP1R14C	6.194640338	SNORD66	-4.550883302	ZNF345	6.415350343	HLA-F-AS1	-5.647298956
SRP9	6.178409988	SNORD114-25	-4.414739085	C9orf150	6.34419314	COBL	-5.624977701
SLC16A6	6.154627985	SNORD114-23	-4.364763237	ZNF267	6.34419314	ITK	-5.602305659
DHRS7B	6.134197819	SNORD116-1	-4.328562751	CDK7	6.307253624	STON2	-5.60068452
MED6	6.134197819	SNORD63	-4.304427648	HIST1H4E	6.288423021	BACH1-IT2	-5.53207003
SEPSECS	6.113474178	SNORD114-28	-4.234283201	CCRK	6.269343379	CA9	-5.432764875
TMEM18	6.102241751	SNORD100	-4.200843048	MYLIP	6.230410007	PEG3	-5.240613641
ZNF804A	6.092448508	SNORD114-12	-4.126167885	RFTN2	6.230410007	TCHH	-5.240613641
PAFAH2	6.071111874	SNORD114-11	-4.126092428	ATL1	6.210542094	NTRK2	-5.180660727

Table S3-2: List of genes influenced by hsa-miR-532-5p over expression. *Genes over and under expressed after 12 and 24 hours of miRNA over expression have been listed.*

24H-Over expressed		24H-Under expressed	
Genes	log FC	Genes	log FC
NOD1	6.883875938	SNORD63	-7.6509335
ZNF345	6.869677145	SNORD94	-6.9985549
PAFAH2	6.569373708	VWA7	-6.4633809
FRMPD3	6.382009516	FAM78A	-6.3103127
GCLC	6.361861984	SNORA20	-6.022578
KCNG1	6.361861984	ZNF600	-6.0073342
FAM151B	6.299674106	MYLK2	-5.9606119
ZNF554	6.299674106	CASP7	-5.9354591
UBE2D4	6.278334506	ISG20	-5.8280738
BBS12	6.256674516	BEGAIN	-5.8106165
CDC14A	6.256674516	TASP1	-5.7929453
CBR3	6.23468437	PKP3	-5.7011893
TM9SF1	6.189672241	IL1RL1	-5.6031987
ZNF700	6.189672241	SNORA71C	-5.6031987
GDAP2	6.16662834	CST6	-5.5197131
MYLIP	6.16662834	TPK1	-5.4760858
MED6	6.143210379	ADAM20	-5.4310979
DCXR	6.095202283	ITK	-5.4310979
TRMT61B	6.095202283	TMEM100	-5.3846619
PAAF1	6.070585552	STAG3L1	-5.3822117
PECR	6.070585552	SHPK	-5.2356505

Table S3-3: ***List of genes influenced by hsa-miR-660-5p over expression.*** *Genes over and under expressed after 24 hours of miRNA over expression have been listed.*

12H- Over expressed		12H-Under expressed		24H-Over expressed		24H-Under expressed	
Genes	log FC	Genes	log FC	Genes	log FC	Genes	log FC
LIMCH1	7.7339483	SNORD114-23	-5.5638624	C3	8.558029845	SNORD94	-7.0686355
EDN1	7.992112	SNORD114-11	-7.2393406	MED6	7.826986484	SNORD14E	-6.7217052
UPRT	7.2318248	SNORD116-1	-6.0513007	PHF16	7.575262204	SNORD116-13	-6.5222353
TFAM	6.3838833			FAM162A	7.426136879	SNORA49	-6.4995401
C3	7.4192955			SBSN	7.300699137	TMEM205	-6.4764822
MED6	7.2318248			CSF3	7.140958768	FAM86A	-6.3553177
TMEM18	6.5650672			DHRS7B	7.106773016	CDKL3	-6.3038513
F2RL1	6.320635			TRMT61B	7.047932594	TYSND1	-6.2906927
CHMP4BP1	7.1352375			IL16	6.974002098	ATP8B1	-6.2597098
HIF1AP1	7.1352375			MYH15	6.974002098	UCN2	-6.2504809
				COL15A1	6.896077455	ZNF624	-6.2368239
				GDAP2	6.882670766	SNORD116-12	-6.2230365
				LIMCH1	6.855477741	TPRKB	-6.1808655
				KIAA1377	6.81370232	SLC25A19	-6.1520507
				LOC387763	6.799503969	ITGA8	-6.1374246
				ZNF594	6.799503969	PFDN6	-6.1374246
				ALG9	6.785164495	SNORA20	-6.092635
				ZNF93	6.785164495	ALS2CL	-6.0773907
				ATP5S	6.726337375	AKAP3	-6.0619835
				GRK6P1	6.726337375	PQLC3	-6.0619835
				THAP1	6.711248001	STON2	-6.0550898
				CDK20	6.680587368	MYLK2	-6.0306667
				SNTG2	6.680587368	ZNF287	-6.0306667
				ARHGEF3	6.633338913	TOMM7	-6.0131132
				SQRDL	6.600957775	FUK	-5.998655

151

Table S3-4: List of genes influenced by hsa-miR-92a-3p over expression. *Genes over and under expressed after 12 and 24 hours of miRNA over expression have been listed.*

12H- Over expressed		12H-Under expressed		24H-Over expressed		24H-Under expressed	
Genes	log FC	Genes	log FC	Genes	log FC	Genes	log FC
LIMCH1	7.5190141	SNORD12C	-6.9506323	PAAF1	7.3326135	SLC25A39P1	-5.7570606
THG1L	6.9583401	SNORD114-15	-6.815252	MED6	7.3220525	VWA7	-5.5111513
EDN1	6.7959506	VWA7	-6.6467353	MYL4	7.153593	SNORD116-13	-5.4999365
MED6	6.7826679	RNF207	-6.5425109	ZNF345	7.1174051	AOC3	-5.3580845
BBS12	6.5825999	C5AR1	-6.3637533	C3	7.0927655	PRDM11	-5.3203297
IL33	6.5662468	SNORD116-8	-6.3267163	UNC45B	6.9897702	OGFOD2	-5.2945996
LOC344967	6.6278688	SNORA7B	-6.2941345	RTN4IP1	6.9353772	SNORD116-12	-5.2007487
GDAP2	6.5038443	C19orf60	-6.2797983	LOC344967	6.8643687	DEPDC1B	-5.1727732
ZNF804A	6.3502503	SNORA20	-6.2058993	SEPSECS	6.7896835	EIF4A1	-5.1585794
AQP1	6.2951901	SNORD9	-6.0956369	KCNB1	6.774271	COA6	-5.1442446
C3	6.2763597	SCN4A	-6.062521	ZNF594	6.774271	SNORD114-15	-5.1192467
CCRK	6.2763597	SNORA11	-6.0286271	TRIM45	6.7109199	TTLL1	-5.055109
PAFAH2	6.2572802	MT-TV	-5.9402359	AQP1	6.6276067	AKAP3	-5.0397026
TMEM132D	6.2572802	PKP3	-5.884479	FAM162A	6.6276067	TRIM47	-4.9601016
WARS2	6.2572802	AKR1E2	-5.8654041	WDSUB1	6.6276067	LIMD2	-4.9395268
ZNF345	6.2572802	SNORA71C	-5.7864774	LIMCH1	6.6103508	SNORD115-23	-4.9269893
TFAM	6.2238164	SNORD114-25	-5.7737961	LOC284600	6.6103508	PPARG	-4.9269893
GCLC	6.2183471	RAB11B-AS1	-5.6813303	CDC14A	6.575207	PCDHGB2	-4.8930991
INSC	6.2183471	HLA-F-AS1	-5.6593489	IKZF2	6.5208322	SNORA11	-4.8930991
TFB1M	6.2183471	TPK1	-5.6593489	EID2	6.4643275	ZNF230	-4.8930991
CALB2	6.1783341	LEAP2	-5.6370275	CCDC107	6.4253895	LINC00847	-4.8758507
NOD1	6.1579037	IQCH-AS1	-5.6143552	ZNRD1	6.3853713	MRPS36P1	-4.8758507
SLFN12	6.1579037	MIR1185-2	-5.4953202	HOXB5	6.3649381	SNORD116-24	-4.8583936

Table S3-5: *List of genes influenced by hsa-miR-206 over expression. Genes over and under expressed after 12 and 24 hours of miRNA over expression have been listed.*

12H Over expressed		
Term	Count	PValue
GO:0010558~negative regulation of macromolecule biosynthetic process	6	0.00126
GO:0030193~regulation of blood coagulation	3	0.001401
GO:0031327~negative regulation of cellular biosynthetic process	6	0.00141
GO:0009890~negative regulation of biosynthetic process	6	0.00155
GO:0050818~regulation of coagulation	3	0.001815
GO:0046328~regulation of JNK cascade	3	0.004501
GO:0010605~negative regulation of macromolecule metabolic process	6	0.004562
24H-Over expressed		
Term	Count	PValue
GO:0050878~regulation of body fluid levels	6	0.017902
GO:0046942~carboxylic acid transport	6	0.020912
GO:0034622~cellular macromolecular complex assembly	9	0.021242
GO:0015849~organic acid transport	6	0.021552
GO:0032103~positive regulation of response to external stimulus	4	0.031253
GO:0022900~electron transport chain	5	0.034176
GO:0034621~cellular macromolecular complex subunit organization	9	0.038417
GO:0043281~regulation of caspase activity	4	0.052902
GO:0001501~skeletal system development	8	0.056007
24H Under expressed		
Term	Count	PValue
GO:0006942~regulation of striated muscle contraction	4	0.001568
GO:0044057~regulation of system process	11	0.00173
GO:0009266~response to temperature stimulus	5	0.011825
GO:0045449~regulation of transcription	39	0.02138
GO:0007517~muscle organ development	7	0.025356
GO:0050873~brown fat cell differentiation	3	0.026648
GO:0051241~negative regulation of multicellular organismal process	6	0.030884
GO:0006350~transcription	32	0.033887
GO:0006937~regulation of muscle contraction	4	0.041365
GO:0014741~negative regulation of muscle hypertrophy	2	0.041908
GO:0014819~regulation of skeletal muscle contraction	2	0.041908

Table S3-6: Enriched Biological process terms for has-miR-139-5p. (12H over expresses – genes over expressed after 12 hours of miRNA transfection, 24H Over expressed - genes over expressed after 24 hours of miRNA transfection, 24H under expressed – genes under expressed after 24hours).

12H- Over expressed		
Term	**Count**	**PValue**
GO:0030193~regulation of blood coagulation	3	0.007486
GO:0050818~regulation of coagulation	3	0.009632
GO:0051240~positive regulation of multicellular organismal process	5	0.011594
GO:0032101~regulation of response to external stimulus	4	0.019783
GO:0070374~positive regulation of ERK1 and ERK2 cascade	2	0.021541
GO:0070372~regulation of ERK1 and ERK2 cascade	2	0.021541
GO:0006390~transcription from mitochondrial promoter	2	0.021541
24H- Over expressed		
Term	**Count**	**PValue**
GO:0045449~regulation of transcription	48	3.56E-04
GO:0006350~transcription	38	0.002945
GO:0032103~positive regulation of response to external stimulus	5	0.005774
GO:0006355~regulation of transcription, DNA-dependent	32	0.00759
GO:0022900~electron transport chain	6	0.009238
GO:0051252~regulation of RNA metabolic process	32	0.010373
GO:0006954~inflammatory response	10	0.011173
24H- Under expressed		
Term	**Count**	**PValue**
GO:0001666~response to hypoxia	3	0.031119
GO:0070482~response to oxygen levels	3	0.034164
GO:0042462~eye photoreceptor cell development	2	0.04062
GO:0001754~eye photoreceptor cell differentiation	2	0.042608
GO:0042461~photoreceptor cell development	2	0.04855

Table S3-7: Enriched Biological process terms for has-miR-532-5p. (12H over expresses – genes over expressed after 12 hours of miRNA transfection, 24H Over expressed - genes over expressed after 24 hours of miRNA transfection, 24H under expressed – genes under expressed after 24hours).

24H- Over expressed		
Term	**Count**	**PValue**
GO:0006350~transcription	53	6.38E-04
GO:0045449~regulation of transcription	62	7.84E-04
GO:0051252~regulation of RNA metabolic process	45	0.002691
GO:0006355~regulation of transcription, DNA-dependent	44	0.00307
GO:0010941~regulation of cell death	24	0.00564
GO:0007005~mitochondrion organization	8	0.006676
GO:0042981~regulation of apoptosis	23	0.009518
GO:0043067~regulation of programmed cell death	23	0.010604
GO:0034621~cellular macromolecular complex subunit organization	13	0.01238
24H- Under expressed		
Term	**Count**	**PValue**
GO:0019438~aromatic compound biosynthetic process	2	0.063268
GO:0007267~cell-cell signaling	6	0.064888

Table S3-8: Enriched Biological process terms for has-miR-660-5p. (24H Over expressed - genes over expressed after 24 hours of miRNA transfection, 24H under expressed – genes under expressed after 24hours).

155

24H- Over expressed		
Term	Count	PValue
GO:0009611~response to wounding	17	0.002267
GO:0045449~regulation of transcription	50	0.00754
GO:0006350~transcription	42	0.00857
GO:0060548~negative regulation of cell death	12	0.009907
GO:0031668~cellular response to extracellular stimulus	5	0.010958
GO:0010941~regulation of cell death	20	0.014399
GO:0007568~aging	6	0.016777
GO:0043066~negative regulation of apoptosis	11	0.022297
GO:0001816~cytokine production	4	0.024157
GO:0043069~negative regulation of programmed cell death	11	0.024276
GO:0042981~regulation of apoptosis	19	0.024488
GO:0043067~regulation of programmed cell death	19	0.026694
GO:0051252~regulation of RNA metabolic process	35	0.028806
24H- Under expressed		
Term	Count	PValue
GO:0006839~mitochondrial transport	6	1.85E-04
GO:0006626~protein targeting to mitochondrion	4	0.002221
GO:0070585~protein localization in mitochondrion	4	0.002221
GO:0017038~protein import	6	0.003375
GO:0007005~mitochondrion organization	6	0.004215
GO:0006605~protein targeting	7	0.006271
GO:0045449~regulation of transcription	30	0.020093
GO:0006886~intracellular protein transport	8	0.025179
GO:0006350~transcription	25	0.027261
GO:0006355~regulation of transcription, DNA-dependent	22	0.027492
GO:0014819~regulation of skeletal muscle contraction	2	0.030402
GO:0014741~negative regulation of muscle hypertrophy	2	0.030402

Table S3-9: Enriched Biological process terms for has-miR-92a-3p. (24H Over expressed - genes over expressed after 24 hours of miRNA transfection, 24H under expressed – genes under expressed after 24hours).

12H- Over expressed		
Term	Count	PValue
GO:0045987~positive regulation of smooth muscle contraction	3	0.004381
GO:0045933~positive regulation of muscle contraction	3	0.005949
GO:0051240~positive regulation of multicellular organismal process	6	0.01136
GO:0032101~regulation of response to external stimulus	5	0.011708
GO:0019229~regulation of vasoconstriction	3	0.015172
GO:0030193~regulation of blood coagulation	3	0.016917
GO:0006940~regulation of smooth muscle contraction	3	0.018744
GO:0050818~regulation of coagulation	3	0.021632
GO:0022900~electron transport chain	4	0.025383
12H- Under expressed		
Term	Count	PValue
GO:0006873~cellular ion homeostasis	3	0.048626
GO:0055082~cellular chemical homeostasis	3	0.050038
GO:0050801~ion homeostasis	3	0.057079
GO:0019725~cellular homeostasis	3	0.071879
24H- Over expressed		
Term	Count	PValue
GO:0006350~transcription	49	2.14E-04
GO:0045449~regulation of transcription	56	4.81E-04
GO:0008016~regulation of heart contraction	6	0.004232
GO:0051252~regulation of RNA metabolic process	39	0.005251
GO:0006936~muscle contraction	8	0.005375
GO:0006355~regulation of transcription, DNA-dependent	38	0.006344
GO:0032874~positive regulation of stress-activated MAPK cascade	3	0.006419
GO:0032872~regulation of stress-activated MAPK cascade	3	0.007951
GO:0007588~excretion	5	0.008391
GO:0003012~muscle system process	8	0.008826
24H- Under expressed		
Term	Count	PValue
GO:0007267~cell-cell signaling	8	0.00692
GO:0007268~synaptic transmission	5	0.025645
GO:0046903~secretion	5	0.026201
GO:0044057~regulation of system process	5	0.028791
GO:0033574~response to testosterone stimulus	2	0.033432
GO:0048545~response to steroid hormone stimulus	4	0.03564

Table S3-10: Enriched Biological process terms for has-miR-206. (12H over expressed - genes over expressed after 12 hours of miRNA transfection, 12H Under expressed - genes under

expressed after 12hours, 24H Over expressed - genes over expressed after 24 hours of miRNA transfection, 24H under expressed – genes under expressed after 24hours).

Chapter – 4

Supplementary Information

Biological process Terms	% of genes	PValue
GO:0009611~response to wounding	8	4.26E-10
GO:0006954~inflammatory response	5.92	8.61E-10
GO:0006955~immune response	9.28	1.11E-09
GO:0006952~defense response	8.32	7.94E-09
GO:0010033~response to organic substance	8.64	2.15E-07
GO:0043066~negative regulation of apoptosis	5.28	8.70E-07
GO:0006916~anti-apoptosis	3.84	8.88E-07
GO:0043069~negative regulation of programmed cell death	5.28	1.18E-06
GO:0042981~regulation of apoptosis	8.96	1.25E-06
GO:0060548~negative regulation of cell death	5.28	1.26E-06
GO:0043067~regulation of programmed cell death	8.96	1.68E-06
GO:0010941~regulation of cell death	8.96	1.90E-06
GO:0001944~vasculature development	4.16	2.50E-06
GO:0006959~humoral immune response	2.24	3.18E-06
GO:0001568~blood vessel development	4	5.30E-06
GO:0045944~positive regulation of transcription from RNA polymerase II promoter	5.12	6.63E-06

Cellular Compartment Terms	% of genes	PValue
GO:0044421~extracellular region part	15.04	9.71E-21
GO:0005615~extracellular space	11.68	5.65E-18
GO:0005576~extracellular region	21.44	5.51E-15
GO:0005578~proteinaceous extracellular matrix	5.44	1.74E-08
GO:0031012~extracellular matrix	5.6	3.25E-08
GO:0000267~cell fraction	9.92	8.16E-05
GO:0005625~soluble fraction	4.16	8.27E-05
GO:0044420~extracellular matrix part	2.08	6.94E-04

Molecular Function Terms	% of genes	PValue
GO:0005125~cytokine activity	4.48	5.33E-10
GO:0008083~growth factor activity	3.52	1.30E-07
GO:0008009~chemokine activity	1.44	1.44E-04
GO:0042379~chemokine receptor binding	1.44	2.27E-04
GO:0005509~calcium ion binding	8.16	6.75E-04
GO:0043565~sequence-specific DNA binding	5.76	0.001707

Table S4-1A: Enriched GO terms for genes upregulated in 3D cultured myotubes compared to 2D cultured myotubes. Terms for Biological processes, Cellular compartments and Molecular functions are shown with %age of genes belonging to each

category along with respective P values. The analysis was performed using DAVID online software.

Biological Process Terms	% of genes	PValue
GO:0007155~cell adhesion	11.20689655	7.42E-08
GO:0022610~biological adhesion	11.20689655	7.63E-08
GO:0016337~cell-cell adhesion	6.034482759	7.20E-06
GO:0043062~extracellular structure organization	4.74137931	8.83E-06
GO:0016126~sterol biosynthetic process	2.586206897	3.09E-05
GO:0051094~positive regulation of developmental process	5.172413793	1.78E-04
GO:0030199~collagen fibril organization	2.155172414	2.32E-04

Cellular compartment Terms	% of genes	PValue
GO:0031012~extracellular matrix	9.482758621	2.07E-10
GO:0005578~proteinaceous extracellular matrix	8.189655172	1.56E-08
GO:0005576~extracellular region	18.96551724	8.96E-06
GO:0044421~extracellular region part	11.63793103	1.79E-05
GO:0044420~extracellular matrix part	3.448275862	3.23E-04
GO:0005886~plasma membrane	25.43103448	0.002537
GO:0005581~collagen	1.724137931	0.006775

Molecular Function Terms	% of genes	PValue
GO:0005198~structural molecule activity	9.482758621	2.02E-06
GO:0005201~extracellular matrix structural constituent	3.879310345	2.82E-06
GO:0019838~growth factor binding	3.017241379	7.54E-04
GO:0005509~calcium ion binding	8.620689655	0.002793
GO:0030246~carbohydrate binding	4.74137931	0.003664
GO:0005272~sodium channel activity	1.724137931	0.004285

Table S4-1B: Enriched GO terms for genes down-regulated in 3D cultured myotubes compared to 2D cultures myotubes. *Terms for Biological processes, Cellular compartments and Molecular functions are shown with %age of genes belonging to each category along with respective P values. The analysis was performed using DAVID online software.*

Biological Process Terms	% of genes	P Value
GO:0006952~defense response	10.6888361	1.06E-11
GO:0006955~immune response	10.9263658	1.32E-10
GO:0006954~inflammatory response	6.650831354	6.00E-09
GO:0009611~response to wounding	8.313539192	4.40E-08
GO:0001568~blood vessel development	5.225653207	1.84E-07
GO:0003012~muscle system process	4.275534442	2.70E-07
GO:0001944~vasculature development	5.225653207	2.78E-07
GO:0001525~angiogenesis	3.800475059	1.35E-06
GO:0048514~blood vessel morphogenesis	4.513064133	1.52E-06
GO:0006936~muscle contraction	3.800475059	2.05E-06
GO:0002684~positive regulation of immune system process	4.750593824	2.10E-06
GO:0048584~positive regulation of response to stimulus	4.513064133	7.39E-06

Cellular Compartment Terms	% of genes	P Value
GO:0044459~plasma membrane part	23.99049881	6.27E-11
GO:0005886~plasma membrane	34.44180523	2.07E-10
GO:0031226~intrinsic to plasma membrane	15.67695962	6.98E-10
GO:0005887~integral to plasma membrane	15.20190024	1.92E-09
GO:0009986~cell surface	6.175771971	1.33E-06
GO:0031224~intrinsic to membrane	41.09263658	5.78E-06
GO:0005615~extracellular space	8.551068884	2.59E-05
GO:0016529~sarcoplasmic reticulum	1.90023753	3.72E-05
GO:0043292~contractile fiber	3.087885986	3.81E-05
GO:0016528~sarcoplasm	1.90023753	5.21E-05
GO:0016021~integral to membrane	38.71733967	8.09E-05

Molecular Function Term	% of genes	P Value
GO:0019955~cytokine binding	3.087885986	8.16E-06
GO:0008092~cytoskeletal protein binding	6.413301663	1.17E-04
GO:0004896~cytokine receptor activity	1.90023753	2.63E-04
GO:0019865~immunoglobulin binding	1.187648456	3.16E-04
GO:0008307~structural constituent of muscle	1.662707838	3.84E-04
GO:0004089~carbonate dehydratase activity	1.187648456	4.13E-04
GO:0019864~IgG binding	0.950118765	6.34E-04
GO:0019838~growth factor binding	2.375296912	7.36E-04

Table S4-2A: Enriched GO terms for genes up-regulated in Skeletal muscle compared to 3D cultured myotubes. Terms for Biological processes, Cellular compartments and Molecular functions are shown with %age of genes belonging to each category along with respective P values. The analysis was performed using DAVID online software.

Biological process Terms	% of genes	PValue
GO:0007155~cell adhesion	9.709821429	4.60E-18
GO:0022610~biological adhesion	9.709821429	5.03E-18
GO:0043062~extracellular structure organization	3.459821429	3.58E-11
GO:0030199~collagen fibril organization	1.5625	9.59E-11
GO:0001501~skeletal system development	4.910714286	1.01E-10
GO:0030198~extracellular matrix organization	2.678571429	1.57E-10
GO:0007267~cell-cell signaling	7.142857143	2.44E-10
GO:0032963~collagen metabolic process	1.450892857	9.77E-10
GO:0044259~multicellular organismal macromolecule metabolic pr	1.450892857	4.00E-09
GO:0006928~cell motion	5.691964286	1.80E-08
GO:0001503~ossification	2.455357143	3.70E-08
GO:0044236~multicellular organismal metabolic process	1.450892857	4.09E-08
GO:0031175~neuron projection development	3.794642857	4.81E-08
GO:0000902~cell morphogenesis	4.575892857	8.81E-08
GO:0048812~neuron projection morphogenesis	3.348214286	9.59E-08

Cellular compartment Terms	% of genes	PValue
GO:0031012~extracellular matrix	9.040178571	1.58E-31
GO:0005578~proteinaceous extracellular matrix	8.59375	1.04E-30
GO:0044421~extracellular region part	15.29017857	3.04E-29
GO:0005576~extracellular region	21.98660714	2.91E-21
GO:0044420~extracellular matrix part	4.017857143	4.01E-18
GO:0005581~collagen	2.008928571	1.85E-13
GO:0005615~extracellular space	8.147321429	3.75E-09
GO:0044459~plasma membrane part	18.30357143	2.60E-07
GO:0005583~fibrillar collagen	0.892857143	5.74E-07
GO:0031226~intrinsic to plasma membrane	11.16071429	1.83E-06
GO:0031224~intrinsic to membrane	37.94642857	2.19E-06
GO:0005604~basement membrane	1.785714286	8.23E-06

Molecular function Terms	% of genes	PValue
GO:0005201~extracellular matrix structural constituent	2.678571429	1.52E-12
GO:0005509~calcium ion binding	9.933035714	2.20E-12
GO:0005125~cytokine activity	3.348214286	8.31E-09
GO:0004222~metalloendopeptidase activity	2.232142857	1.19E-07
GO:0008083~growth factor activity	2.566964286	2.29E-06
GO:0008237~metallopeptidase activity	2.678571429	5.75E-06
GO:0022843~voltage-gated cation channel activity	2.34375	7.03E-06
GO:0005198~structural molecule activity	5.580357143	9.18E-05

Table S4-2B: Enriched GO terms for genes down-regulated in Skeletal muscle compared to 3D cultured myotubes. Terms for Biological processes, Cellular compartments and Molecular functions are shown with %age of genes belonging to each

163

category along with respective P values. The analysis was performed using DAVID online software.

Biological Process Terms	% of genes	PValue
GO:0006955~immune response	7.635597683	2.22E-16
GO:0006952~defense response	6.477093207	1.91E-12
GO:0009611~response to wounding	5.739863086	6.97E-12
GO:0015980~energy derivation by oxidation of organic compounds	2.422327541	8.16E-12
GO:0006954~inflammatory response	4.054765666	1.20E-11
GO:0003012~muscle system process	2.632964718	1.63E-11
GO:0045333~cellular respiration	1.895734597	1.85E-11
GO:0022904~respiratory electron transport chain	1.474460242	5.76E-11
GO:0006091~generation of precursor metabolites and energy	3.844128489	9.43E-11
GO:0055114~oxidation reduction	6.319115324	2.54E-10
GO:0006936~muscle contraction	2.369668246	2.91E-10
GO:0022900~electron transport chain	1.895734597	2.87E-09
GO:0010033~response to organic substance	6.582411796	1.62E-08
GO:0042775~mitochondrial ATP synthesis coupled electron transport	1.21116377	1.69E-08
GO:0042773~ATP synthesis coupled electron transport	1.21116377	1.69E-08

Cellular compartment Terms	% of genes	PValue
GO:0005743~mitochondrial inner membrane	4.054765666	3.06E-13
GO:0043292~contractile fiber	2.159031069	1.28E-11
GO:0005740~mitochondrial envelope	4.791995787	1.43E-11
GO:0019866~organelle inner membrane	4.054765666	1.49E-11
GO:0005739~mitochondrion	9.583991575	4.05E-11
GO:0031966~mitochondrial membrane	4.528699315	4.17E-11
GO:0030016~myofibril	2.001053186	5.69E-11
GO:0070469~respiratory chain	1.579778831	1.18E-10
GO:0044429~mitochondrial part	5.845181675	1.15E-09
GO:0031226~intrinsic to plasma membrane	10.11058452	1.59E-09
GO:0005746~mitochondrial respiratory chain	1.369141654	1.75E-09
GO:0044449~contractile fiber part	1.895734597	1.78E-09
GO:0030017~sarcomere	1.737756714	2.02E-09
GO:0044459~plasma membrane part	16.21906266	8.94E-09
GO:0044455~mitochondrial membrane part	1.948393892	9.24E-09

Molecular Function Terms	% of genes	PValue
GO:0019955~cytokine binding	1.632438125	5.21E-07
GO:0016655~oxidoreductase activity	1.000526593	1.05E-06
GO:0030246~carbohydrate binding	3.528172722	2.37E-06
GO:0003954~NADH dehydrogenase activity	0.895208004	3.43E-06
GO:0008137~NADH dehydrogenase (ubiquinone) activity	0.895208004	3.43E-06
GO:0050136~NADH dehydrogenase (quinone) activity	0.895208004	3.43E-06
GO:0016651~oxidoreductase activity, acting on NADH or NADPH	1.263823065	9.58E-06
GO:0004896~cytokine receptor activity	0.947867299	2.95E-05
GO:0046983~protein dimerization activity	4.58135861	5.74E-05
GO:0008092~cytoskeletal protein binding	4.318062138	5.94E-05

Table S4-3A: Enriched GO terms for genes up-regulated in Skeletal muscle compared to 2D cultured myotubes. *Terms for Biological processes, Cellular compartments and Molecular functions are shown with %age of genes belonging to each category along with respective P values. The analysis was performed using DAVID online software.*

Biological Function Terms	% of genes	PValue
GO:0022610~biological adhesion	8.855047941	3.33E-29
GO:0007155~cell adhesion	8.798646362	9.20E-29
GO:0043062~extracellular structure organization	3.327693175	1.17E-21
GO:0030199~collagen fibril organization	1.240834743	1.18E-16
GO:0030198~extracellular matrix organization	2.199661591	5.03E-15
GO:0000902~cell morphogenesis	4.286520023	4.77E-13
GO:0016337~cell-cell adhesion	3.609701072	9.19E-13
GO:0032989~cellular component morphogenesis	4.568527919	9.28E-13
GO:0001501~skeletal system development	3.835307389	1.01E-11
GO:0048858~cell projection morphogenesis	3.158488438	5.20E-11
GO:0000904~cell morphogenesis involved in differentiation	3.102086858	1.38E-10
GO:0032990~cell part morphogenesis	3.158488438	3.04E-10
GO:0051960~regulation of nervous system development	2.594472645	6.89E-10

Cellular Compartment Terms	% of genes	PValue
GO:0031012~extracellular matrix	6.598984772	5.20E-37
GO:0005578~proteinaceous extracellular matrix	6.260575296	4.03E-36
GO:0044420~extracellular matrix part	2.9892837	1.41E-23
GO:0044421~extracellular region part	10.20868584	2.68E-21
GO:0005581~collagen	1.297236323	8.48E-15
GO:0005604~basement membrane	1.748448957	3.86E-12
GO:0005576~extracellular region	15.51043429	4.42E-12
GO:0044459~plasma membrane part	15.84884377	5.77E-09
GO:0042995~cell projection	6.091370558	1.27E-07
GO:0030054~cell junction	4.737732657	6.02E-07
GO:0045202~synapse	3.553299492	9.06E-07

Molecular Function Terms	% of genes	PValue
GO:0005509~calcium ion binding	8.403835307	1.03E-14
GO:0005201~extracellular matrix structural constituent	1.917653694	3.06E-14
GO:0019838~growth factor binding	1.46644106	2.00E-06
GO:0043167~ion binding	24.42188381	4.06E-06
GO:0005178~integrin binding	1.015228426	5.66E-06
GO:0004222~metalloendopeptidase activity	1.410039481	5.71E-06
GO:0043169~cation binding	24.02707276	6.30E-06
GO:0046872~metal ion binding	23.80146644	7.16E-06
GO:0003774~motor activity	1.692047377	8.51E-06
GO:0003779~actin binding	2.932882121	1.84E-05
GO:0032403~protein complex binding	2.030456853	2.35E-05
GO:0005518~collagen binding	0.73322053	2.69E-05

Table S4-3B: Enriched GO terms for genes down-regulated in Skeletal muscle compared to 2D cultured myotubes. *Terms for Biological processes, Cellular compartments and Molecular functions are shown with %age of genes belonging to each*

category along with respective P values. The analysis was performed using DAVID online software.

Biological function terms	% of genes	PValue
GO:0003012~muscle system process	6.263982103	2.05E-31
GO:0006936~muscle contraction	5.928411633	1.17E-30
GO:0007517~muscle organ development	6.040268456	4.09E-24
GO:0006941~striated muscle contraction	2.572706935	1.89E-17
GO:0007155~cell adhesion	8.948545861	8.31E-13
GO:0022610~biological adhesion	8.948545861	8.65E-13
GO:0060537~muscle tissue development	3.243847875	5.81E-12
GO:0044057~regulation of system process	5.257270694	6.83E-12
GO:0014706~striated muscle tissue development	3.131991051	1.01E-11
GO:0008016~regulation of heart contraction	2.460850112	5.03E-11
GO:0055001~muscle cell development	2.013422819	5.59E-10
GO:0030239~myofibril assembly	1.342281879	6.80E-10
GO:0055002~striated muscle cell development	1.901565996	1.44E-09
GO:0031032~actomyosin structure organization	1.454138702	2.12E-09
GO:0042692~muscle cell differentiation	2.684563758	1.34E-08

Cellular compartment terms	% of genes	PValue
GO:0043292~contractile fiber	6.263982103	2.07E-39
GO:0030016~myofibril	5.928411633	3.85E-38
GO:0044449~contractile fiber part	5.704697987	3.41E-35
GO:0030017~sarcomere	5.257270694	6.82E-34
GO:0031674~I band	3.020134228	9.06E-20
GO:0015629~actin cytoskeleton	6.263982103	1.34E-19
GO:0031012~extracellular matrix	6.823266219	1.07E-17
GO:0005578~proteinaceous extracellular matrix	6.487695749	2.36E-17
GO:0030018~Z disc	2.348993289	1.33E-14
GO:0044421~extracellular region part	11.63310962	6.79E-14
GO:0016459~myosin complex	2.572706935	3.02E-13
GO:0032982~myosin filament	1.454138702	8.65E-12
GO:0016529~sarcoplasmic reticulum	1.901565996	1.79E-11
GO:0016528~sarcoplasm	1.901565996	4.45E-11
GO:0042383~sarcolemma	2.348993289	5.03E-11
GO:0005859~muscle myosin complex	1.342281879	5.32E-10
GO:0016460~myosin II complex	1.342281879	2.04E-09

Molecular function terms	% of genes	PValue
GO:0008307~structural constituent of muscle	2.796420582	1.61E-21
GO:0008092~cytoskeletal protein binding	8.165548098	1.30E-17
GO:0003779~actin binding	5.928411633	7.59E-15
GO:0005509~calcium ion binding	10.29082774	7.70E-12
GO:0019838~growth factor binding	2.572706935	3.01E-09
GO:0005198~structural molecule activity	6.935123043	8.74E-08
GO:0005261~cation channel activity	3.914988814	2.77E-07
GO:0015267~channel activity	4.921700224	9.07E-07
GO:0022838~substrate specific channel activity	4.8098434	9.18E-07
GO:0022803~passive transmembrane transporter activity	4.921700224	9.67E-07
GO:0005216~ion channel activity	4.697986577	1.04E-06
GO:0022843~voltage-gated cation channel activity	2.572706935	1.55E-06
GO:0005244~voltage-gated ion channel activity	3.020134228	1.72E-06
GO:0022832~voltage-gated channel activity	3.020134228	1.72E-06

Table S4-4A: Enriched GO terms for genes up-regulated in 2D cultured myotubes compared to myoblasts. *Terms for Biological processes, Cellular compartments and Molecular functions are shown with %age of genes belonging to each category along with respective P values. The analysis was performed using DAVID online software.*

Biological process terms	% of genes	PValue
GO:0000279~M phase	9.030837004	1.63E-53
GO:0022403~cell cycle phase	10.13215859	3.09E-53
GO:0007049~cell cycle	14.17033774	1.09E-52
GO:0022402~cell cycle process	11.01321586	4.12E-44
GO:0000087~M phase of mitotic cell cycle	6.754772394	8.38E-44
GO:0000280~nuclear division	6.681350954	1.26E-43
GO:0007067~mitosis	6.681350954	1.26E-43
GO:0048285~organelle fission	6.681350954	6.45E-42
GO:0000278~mitotic cell cycle	8.370044053	5.41E-40
GO:0051301~cell division	7.121879589	1.40E-36

Cellular compartment terms	% of genes	PValue
GO:0005694~chromosome	10.27900147	1.15E-49
GO:0044427~chromosomal part	9.251101322	3.09E-48
GO:0000793~condensed chromosome	4.405286344	1.57E-32
GO:0000775~chromosome, centromeric region	4.185022026	1.42E-30
GO:0043232~intracellular non-membrane-bounded organelle	24.22907489	3.22E-26
GO:0043228~non-membrane-bounded organelle	24.22907489	3.22E-26
GO:0000779~condensed chromosome, centromeric region	2.716593245	1.08E-23
GO:0000777~condensed chromosome kinetochore	2.496328928	1.29E-22
GO:0000776~kinetochore	2.790014684	8.64E-22

Molecular function terms	% of genes	PValue
GO:0032559~adenyl ribonucleotide binding	12.48164464	9.57E-09
GO:0008094~DNA-dependent ATPase activity	1.468428781	1.35E-08
GO:0005524~ATP binding	12.26138032	1.89E-08
GO:0001882~nucleoside binding	12.99559471	5.34E-08
GO:0030554~adenyl nucleotide binding	12.70190896	8.20E-08
GO:0001883~purine nucleoside binding	12.84875184	8.89E-08
GO:0003677~DNA binding	17.1071953	6.64E-07
GO:0032553~ribonucleotide binding	13.95007342	1.09E-06
GO:0032555~purine ribonucleotide binding	13.95007342	1.09E-06
GO:0000166~nucleotide binding	16.22613803	4.52E-06

Table S4-4B: Enriched GO terms for genes down-regulated in 2D cultured myotubes compared to myoblasts. *Terms for Biological processes, Cellular compartments and Molecular functions are shown with %age of genes belonging to each category along with respective P values. The analysis was performed using DAVID online software.*

Chapter-5

Supplementary information

Gene	Description	3h PS up	65h differentiation
SLC7A2	solute carrier family 7 (cationic amino acid transporter, y+ system), member 2	2.328450642	2.137481112
MYH13	myosin, heavy chain 13, skeletal muscle	1.979979239	2.644173185
FAM65C	family with sequence similarity 65, member C	3.531013342	3.051521466
PTGS2	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxyg	2.22176717	2.951015151
CXCL2	Chemokine (C-X-C motif) ligand 2	5.512684907	5.067445613
PTHLH	Parathyroid hormone-like hormone	1.799682582	2.572029144
IL11	Interleukin 11	2.792468051	3.133896828
BDKRB1	Bradykinin receptor B1	2.712777687	2.195771912
SLCO4A1	Solute carrier organic anion transporter family, member 4A1	3.84117393	3.761933979
PDGFRL	Platelet-derived growth factor receptor-like	1.583593455	1.851711302
CSF3	Colony stimulating factor 3 (granulocyte)	5.410659841	5.031639087
CCL2	Chemokine (C-C motif) ligand 2	2.129136483	2.671566168
PADI2	Peptidyl arginine deiminase, type II	1.439834302	3.521512229
HSD11B1	Hydroxysteroid (11-beta) dehydrogenase 1	1.790478209	4.286628487
TNFAIP6	Tumor necrosis factor, alpha-induced protein 6	2.538932341	2.944623678
G0S2	G0/G1switch 2	1.964905042	2.047857717
CXCL6	Chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2)	4.19420982	5.218094397
EREG	Epiregulin	2.793276793	1.974534911
IL1B	Interleukin 1, beta	4.769195706	4.859993511
C3	Complement component 3	1.937700428	2.736318121
DRAM1	Damage-regulated autophagy modulator	1.812948779	2.024651147
EDNRB	Endothelin receptor type B	5.406211792	6.823623593
IL6	Interleukin 6 (interferon, beta 2)	2.910441617	3.778830002
IL1RN	Interleukin 1 receptor antagonist	2.977031682	2.810853914
FGF7	Fibroblast growth factor 7 (keratinocyte growth factor)	1.501875197	2.571632478
MYH15	Myosin, heavy chain 15	2.170123588	3.053119007
PRSS35	Protease, serine, 35	2.259027876	2.716687564
MMP3	Matrix metallopeptidase 3 (stromelysin 1, progelatinase)	1.873753184	3.535860852
BMP6	Bone morphogenetic protein 6	2.538381129	3.26537318
NR4A2	Nuclear receptor subfamily 4, group A, member 2	2.567804952	2.776175814
STC1	Stanniocalcin 1	2.049768217	2.080516736
RCAN1	Regulator of calcineurin 1	1.574135016	3.045071858
IL24	Interleukin 24	3.465339337	3.55215553
CXCL3	Chemokine (C-X-C motif) ligand 3	4.884179874	5.422684782
CXCL5	Chemokine (C-X-C motif) ligand 5	2.354070196	3.444303802
CXCL1	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	4.848795943	4.915784824
BDKRB2	Bradykinin receptor B2	2.076448304	1.782412063
IL8	Interleukin 8	4.28280788	4.079484963
PDE7B	Phosphodiesterase 7B	1.820773804	2.055418149
SNTG2	Syntrophin, gamma 2	2.627048854	2.990883083
TMEM217	Transmembrane protein 217	1.901758556	2.430629034
PCSK1	Proprotein convertase subtilisin/kexin type 1	1.923027999	2.562661379
PDE4B	phosphodiesterase 4B, cAMP-specific	1.535515291	1.872138117
AKR1C1	aldo-keto reductase family 1, member C1	1.639186131	1.771265386
C2CD4A	C2 calcium-dependent domain containing 4A	3.112784614	3.19213768
LINC00473	long intergenic non-protein coding RNA 473	4.529320167	6.298072296
NAMPTL	nicotinamide phosphoribosyltransferase-like	2.045305703	1.914285881

Table S5-1: List of genes differentially upregulated after 3h of stretch and also after 65 hours of differentiation without any stretching. *The fold change is very similar for both the conditions. Genes which had higher fold change after 65 hours of differentiation compared to 3h post stretch have been highlighted in yellow.*

Biological process Terms	Count	PValue
GO:0009611~response to wounding	22	4.65E-14
GO:0006954~inflammatory response	16	4.16E-11
GO:0042127~regulation of cell proliferation	19	3.02E-08
GO:0042330~taxis	10	8.60E-08
GO:0006935~chemotaxis	10	8.60E-08
GO:0007626~locomotory behavior	12	9.06E-08
GO:0006955~immune response	17	1.73E-07
GO:0006952~defense response	16	2.37E-07
GO:0007610~behavior	14	4.06E-07
GO:0008284~positive regulation of cell proliferation	13	7.52E-07
GO:0030595~leukocyte chemotaxis	6	9.87E-07
GO:0060326~cell chemotaxis	6	1.29E-06
GO:0002237~response to molecule of bacterial origin	7	4.05E-06
GO:0050900~leukocyte migration	6	8.79E-06
GO:0019221~cytokine-mediated signaling pathway	6	2.42E-05
GO:0007267~cell-cell signaling	13	3.32E-05
GO:0032496~response to lipopolysaccharide	6	3.84E-05
GO:0033135~regulation of peptidyl-serine phosphorylation	4	4.06E-05
GO:0051384~response to glucocorticoid stimulus	6	4.09E-05
GO:0031960~response to corticosteroid stimulus	6	6.19E-05
GO:0030593~neutrophil chemotaxis	4	8.98E-05

Table S5-2A: Biological process terms for gene differentially <u>up regulated</u> after 3h of stretch compared to control.

Biological process Term	Count	PValue
GO:0001568~blood vessel development	5	0.0017461
GO:0001944~vasculature development	5	0.00190767
GO:0001525~angiogenesis	4	0.00380625
GO:0044092~negative regulation of molecular function	5	0.00532261
GO:0045765~regulation of angiogenesis	3	0.00799216
GO:0048514~blood vessel morphogenesis	4	0.01013978
GO:0007178~transmembrane receptor protein serine/threonine kinase signaling pathway	3	0.02039152
GO:0051789~response to protein stimulus	3	0.02189858
GO:0043405~regulation of MAP kinase activity	3	0.03644919
GO:0045859~regulation of protein kinase activity	4	0.03682286
GO:0043549~regulation of kinase activity	4	0.04012315
GO:0051338~regulation of transferase activity	4	0.0444557
GO:0000165~MAPKKK cascade	3	0.05877736

Table S5-2B: Biological process terms for gene differentially <u>down regulated</u> after 3h of stretch compared to control.

Term	Count	PValue
GO:0006563~L-serine metabolic process	4	3.50E-04
GO:0015804~neutral amino acid transport	5	5.82E-04
GO:0031667~response to nutrient levels	12	0.00133158
GO:0032869~cellular response to insulin stimulus	7	0.0017778
GO:0043434~response to peptide hormone stimulus	10	0.00264518
GO:0032868~response to insulin stimulus	8	0.00287545
GO:0009991~response to extracellular stimulus	12	0.00316594
GO:0042060~wound healing	11	0.00347644
GO:0009719~response to endogenous stimulus	17	0.00449708
GO:0007584~response to nutrient	9	0.00515468
GO:0060284~regulation of cell development	11	0.00573151
GO:0043067~regulation of programmed cell death	27	0.00603367
GO:0010941~regulation of cell death	27	0.00630437
GO:0006869~lipid transport	9	0.00634503

Table S5-3A: Biological process terms for gene differentially <u>up regulated</u> after 40h of stretch compared to control.

Term	Count	PValue
GO:0009611~response to wounding	54	3.92E-12
GO:0006954~inflammatory response	35	1.12E-08
GO:0007610~behavior	42	6.09E-08
GO:0007626~locomotory behavior	30	1.07E-07
GO:0006952~defense response	48	3.67E-07
GO:0001501~skeletal system development	31	8.42E-07
GO:0006935~chemotaxis	19	1.23E-05
GO:0042330~taxis	19	1.23E-05
GO:0008284~positive regulation of cell proliferation	33	2.30E-05
GO:0007204~elevation of cytosolic calcium ion concentration	15	2.86E-05
GO:0009719~response to endogenous stimulus	32	3.73E-05
GO:0016477~cell migration	25	3.87E-05
GO:0048545~response to steroid hormone stimulus	20	4.38E-05
GO:0060326~cell chemotaxis	9	4.98E-05
GO:0043062~extracellular structure organization	18	5.68E-05
GO:0042127~regulation of cell proliferation	50	5.85E-05
GO:0051480~cytosolic calcium ion homeostasis	15	6.30E-05
GO:0030198~extracellular matrix organization	14	6.65E-05
GO:0007267~cell-cell signaling	41	6.82E-05

Table S5-3B: Biological process terms for gene differentially <u>down regulated</u> after 40h of stretch compared to control.

Term	Count	PValue
GO:0006936~muscle contraction	43	1.40E-23
GO:0003012~muscle system process	44	8.54E-23
GO:0007517~muscle organ development	39	9.40E-15
GO:0006941~striated muscle contraction	19	8.24E-14
GO:0014706~striated muscle tissue development	21	7.40E-08
GO:0060537~muscle tissue development	21	1.71E-07
GO:0007519~skeletal muscle tissue development	15	3.59E-07
GO:0060538~skeletal muscle organ development	15	3.59E-07
GO:0044057~regulation of system process	33	1.47E-06
GO:0030049~muscle filament sliding	6	1.27E-05
GO:0033275~actin-myosin filament sliding	6	1.27E-05
GO:0070252~actin-mediated cell contraction	6	1.27E-05
GO:0008016~regulation of heart contraction	14	1.36E-05
GO:0030048~actin filament-based movement	8	2.63E-05
GO:0043462~regulation of ATPase activity	7	2.64E-05
GO:0006937~regulation of muscle contraction	13	3.32E-05
GO:0042692~muscle cell differentiation	17	3.50E-05

Table S5-4A: Biological process terms for gene differentially <u>up regulated</u> after 40h of stretch compared to 3h post stretch.

Term	Count	PValue
GO:0000279~M phase	48	4.14E-14
GO:0000280~nuclear division	38	1.49E-13
GO:0007067~mitosis	38	1.49E-13
GO:0000087~M phase of mitotic cell cycle	38	2.64E-13
GO:0048285~organelle fission	38	5.32E-13
GO:0022403~cell cycle phase	51	4.35E-12
GO:0007049~cell cycle	75	4.38E-12
GO:0000278~mitotic cell cycle	44	5.07E-10
GO:0022402~cell cycle process	56	1.52E-09
GO:0051301~cell division	37	3.69E-09
GO:0007059~chromosome segregation	17	1.33E-07
GO:0009611~response to wounding	48	4.15E-07
GO:0000070~mitotic sister chromatid segregation	11	1.13E-06
GO:0000819~sister chromatid segregation	11	1.49E-06
GO:0051726~regulation of cell cycle	34	1.91E-06
GO:0008284~positive regulation of cell proliferation	38	6.17E-06
GO:0000226~microtubule cytoskeleton organization	20	8.13E-06
GO:0006954~inflammatory response	32	9.73E-06

Table S5-4B: Biological process terms for gene differentially <u>down regulated</u> after 40h of stretch compared to 3h post stretch.

Gene name	Description	logFC	PValue
CXCL2	Chemokine (C-X-C motif) ligand 2	5.512684907	3.19E-19
CSF3	Colony stimulating factor 3 (granulocyte)	5.410659841	1.07E-25
EDNRB	Endothelin receptor type B	5.406211792	6.43E-09
CXCL3	Chemokine (C-X-C motif) ligand 3	4.884179874	4.52E-12
CXCL1	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	4.848795943	2.89E-23
IL1B	Interleukin 1, beta	4.769195706	1.97E-17
IL8	Interleukin 8	4.28280788	9.07E-21
RBPMS2	RNA binding protein with multiple splicing 2	4.218398729	0.000124
CXCL6	Chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2)	4.19420982	1.56E-19
SLCO4A1	Solute carrier organic anion transporter family, member 4A1	3.84117393	5.45E-09
FAM65C	Family with sequence similarity 65, member C	3.531013342	1.02E-13
IL24	Interleukin 24	3.465339337	3.09E-07
C2CD4A	Family with sequence similarity 148, member A	3.112784614	3.18E-08
IL1RN	Interleukin 1 receptor antagonist	2.977031682	1.72E-10
IL6	Interleukin 6 (interferon, beta 2)	2.910441617	3.10E-11
EREG	Epiregulin	2.793276793	2.23E-11
IL11	Interleukin 11	2.792468051	7.03E-11
BDKRB1	Bradykinin receptor B1	2.712777687	1.19E-09
SLC16A6	Solute carrier family 16, member 6 (monocarboxylic acid transporter 7)	2.639617028	8.65E-07
SNTG2	Syntrophin, gamma 2	2.627048854	9.13E-06
SLC5A3	Mitochondrial ribosomal protein S6	2.581017714	1.57E-10
NR4A2	Nuclear receptor subfamily 4, group A, member 2	2.567804952	2.46E-08
TNFAIP6	Tumor necrosis factor, alpha-induced protein 6	2.538932341	7.19E-09
BMP6	Bone morphogenetic protein 6	2.538381129	2.95E-08
MT1F	Metallothionein 1F	2.522836927	7.37E-05
LRRN3	Leucine rich repeat neuronal 3	2.435681285	9.63E-06
CXCL5	Chemokine (C-X-C motif) ligand 5	2.354070196	1.15E-06
RNF144B	Ring finger protein 144B	2.352708744	1.14E-05
SLC7A2	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 2	2.328450642	5.92E-07
PRSS35	Protease, serine, 35	2.259027876	2.79E-08
CH25H	Cholesterol 25-hydroxylase	2.239103642	1.61E-05
PTGS2	Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	2.22176717	1.34E-07
SOD2	Superoxide dismutase 2, mitochondrial	2.172065892	4.38E-08
MYH15	Myosin, heavy chain 15	2.170123588	1.10E-05
CCL2	Chemokine (C-C motif) ligand 2	2.129136483	4.13E-07
NTN1	Netrin 1	2.101380092	7.85E-06
BDKRB2	Bradykinin receptor B2	2.076448304	1.95E-07
PDPN	Podoplanin	2.069010605	6.49E-07
GPR183	G protein-coupled receptor 183	2.057097704	4.37E-05
STC1	Stanniocalcin 1	2.049768217	4.47E-07
CD70	CD70 molecule	2.030848184	7.61E-05
MYH13	Myosin, heavy chain 13, skeletal muscle	1.979979239	0.00014

Table S5-5: Genes differentially <u>up regulated (>2x fold)</u> after 3h of stretch compared to control.

Gene name	Description	logFC	PValue
GRIN2A	Glutamate receptor, ionotropic, N-methyl D-aspartate 2A	4.943056	6.39E-09
ANO1	Anoctamin 1, calcium activated chloride channel	4.503801	1.58E-07
HRK	Harakiri, BCL2 interacting protein (contains only BH3 domain)	4.26102	2.15E-06
RAMP2	Receptor (G protein-coupled) activity modifying protein 2	3.886529	4.38E-05
FOXS1	Forkhead box S1	3.870176	8.77E-10
CRYM	Crystallin, mu	3.680122	2.35E-06
FAM46C	Family with sequence similarity 46, member C	3.629949	0.000282
SLC14A1	Solute carrier family 14 (urea transporter), member 1 (Kidd blood group)	3.456989	7.70E-12
KIF17	Kinesin family member 17	3.107831	0.000107
APOM	Apolipoprotein M	3.072581	0.000107
HSD17B2	Hydroxysteroid (17-beta) dehydrogenase 2	3.039269	0.000142
FCRLB	Fc receptor-like B	2.999371	0.000189
PLA2G3	Phospholipase A2, group III	2.972284	4.49E-05
CLDN14	Claudin 14	2.814028	2.87E-13
MMRN2	Multimerin 2	2.806173	0.000201
RAB39B	RAB39B, member RAS oncogene family	2.77584	0.000257
TRIB3	Tribbles homolog 3 (Drosophila)	2.763747	3.89E-23
PCDH1	Protocadherin 1	2.74497	1.02E-19
CHRM4	Cholinergic receptor, muscarinic 4	2.660895	0.001896
IGSF22	Immunoglobulin superfamily, member 22	2.631733	1.64E-09
AKR1B10	Aldo-keto reductase family 1, member B10 (aldose reductase)	2.583304	2.11E-14
SUSD4	Sushi domain containing 4	2.575234	5.42E-10
ST6GAL2	ST6 beta-galactosamide alpha-2,6-sialyltranferase 2	2.56925	4.22E-06
LCE3A	Late cornified envelope 3A	2.514626	0.001482
RIMS4	Regulating synaptic membrane exocytosis 4	2.478178	0.001902
ZDHHC22	Zinc finger, DHHC-type containing 22	2.476296	0.00244
OAS3	2\'-5\'-oligoadenylate synthetase 3, 100kDa	2.468994	2.30E-07
CCDC148	Coiled-coil domain containing 148	2.441424	0.00244
SH2D7	SH2 domain containing 7	2.439529	0.00244
GJD4	Gap junction protein, delta 4, 40.1kDa	2.415563	4.81E-15
IL13RA2	Interleukin 13 receptor, alpha 2	2.405915	0.003129
ADM2	Adrenomedullin 2	2.393414	1.83E-15
C3orf14	Chromosome 3 open reading frame 14	2.391017	2.50E-06
ELOVL2	Elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 2	2.313104	0.001252
FAM211A	Chromosome 17 open reading frame 76	2.311901	0.001252
SIGLEC15	Sialic acid binding Ig-like lectin 15	2.303024	8.30E-05
SLC47A2	Solute carrier family 47, member 2	2.26658	0.000103
SCN7A	Sodium channel, voltage-gated, type VII, alpha	2.262267	0.000103
LPAR4	Lysophosphatidic acid receptor 4	2.248113	0.002003
SLN	Sarcolipin	2.243661	7.61E-10
OPCML	Opioid binding protein/cell adhesion molecule-like	2.216203	0.002531
CPNE7	Copine VII	2.178742	9.57E-06
SLCO4C1	Solute carrier organic anion transporter family, member 4C1	2.170537	0.000242
FAM129A	Family with sequence similarity 129, member A	2.146806	3.75E-16
KCNN4	Potassium intermediate/small conductance calcium-activated channel, subfamily N, member 4	2.145676	1.67E-07
SPP1	Secreted phosphoprotein 1	2.13923	3.92E-14
ICA1L	Islet cell autoantigen 1,69kDa-like	2.125392	0.000141
IL17B	Interleukin 17B	2.120046	1.30E-07

Table S5-6: *Genes differentially <u>up regulated (>2x fold)</u> after 40h of stretch compared to control.*