FIBRE TYPE SPECIFIC DNA METHYLATION IN HUMAN SKELETAL MUSCLE

presented by
ANDREW PALMER

M.Sc. ETH Zurich
B.Ex.Phys. UNSW

Thesis submitted for the fulfillment of the
requirements for the degree of
DOCTOR OF PHILOSOPHY

Victoria University, Australia
Institute for Health and Sport (iHeS)

April 2024

ABSTRACT

Skeletal muscle is a heterogeneous tissue comprised of diverse cell types and is highly adaptable and sensitive to stimuli, such as exercise. Epigenetics is the study of mechanisms that impact chromosomal regions to alter gene activity without an underlying change in the sequence of Deoxyribonucleic acid (DNA). One of the most widely studied epigenetic mechanisms is DNA methylation (DNAm), which plays an important role in perpetuating cell identity. To date, studies have shown that exercise causes widespread DNAm changes in skeletal muscle. However, these studies have remained at the whole tissue-level. This approach limits interpretation of results due to, 1) the influence of cell identities on DNAm, 2) the variable cellular composition of tissue and 3) the specific responses of each cell type to exercise. This thesis investigates the genome-wide DNAm profiles of Type I, Type IIa and Type IIx multi-nucleated muscle fibres from human muscle at baseline and after twelve weeks of High Intensity Interval Training (HIIT). It describes the development of a novel method for the simultaneous extraction of DNA and protein from pooled human skeletal muscle fibre fragments (Chapter 4). The thesis also compares the methylome and proteome assessed in skeletal muscle fibre types using the same biological samples (Chapter 5 & 6), and in response to twelve weeks of HIIT (Chapter 7). This research will contribute to a deeper understanding of fibre-type specific muscle biology and molecular adaptations to exercise, with potential to improve treatment strategies for muscle related diseases and maintenance of muscle health.

Student Declaration

"I, Andrew Palmer, declare that the PhD thesis entitled Fibre Type Specific DNA Methylation in Human Skeletal Muscle is no more than 80,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work".

"I have conducted my research in alignment with the Australian Code for the Responsible Conduct of Research and Victoria University's Higher Degree by Research Policy and Procedures.

Signature  Date: 29/03/2024

Ethics Declaration "All research procedures reported in the thesis were approved by the human ethics and research committee at Victoria University (HRE13-223, HRE21-122)."

Signature  Date: 29/03/2024

Acknowledgements

The work presented here is the culmination of a number (quite a large number) of years of study. Throughout the entire process, I am deeply indebted to a number of people who have assisted me and provided support and guidance.

I would firstly like to thank my principal supervisor Prof. Nir Eynon for all of his guidance and support and for the freedom he afforded me to develop and shape my research questions. I am certain that under different guidance I would not have had the support to pursue the work presented here. I would also like to thank my co-supervisors Dr. Danielle Hiam and Dr. Sarah Voisin, both of whom provided me with guidance and words that assured me the ideas I developed were both possible and that I would be able to achieve them.

In this work I had the opportunity to collaborate with the Monash Proteomics and Metabolomics Platform. I would like to thank the team headed by Assoc. Prof. Ralf Schittenhelm, including: Dr. Joel Steele, Dr. Han Lee and Iresha Hanchapola for their expertise in proteomics. In particular, I am grateful for Joel's constant answering of emails and for seeding the idea to utilise the dot blot apparatus in the method.

I would also like to thank all the members of the Eynon Lab throughout my PhD; Dr Javier Alvarez-Romero, Dr Shanie Landen, Dr Macsue Jacques, Dr Patrice Jones, Dr Megan Taylor, Dr. Sarah Cornish, Ethan Mooreland and Dr Kirsten Seale. A big thankyou goes to Javier and Shanie who aided me during multiple participant testing and biopsy sessions. Although later in on the action both Patrice and Ethan stepped up to help me towards the later stages of training, testing and biopsy days. A big big thanks to Meg, who spent countless hours assisting in the biochemistry lab and pouring over the written manuscript. I would also like to extend many thanks to Dr Esther García-Domínguez and Dr Javier Botella-Ruiz. Javi thanks for the support, discussions, ideas and fibre isolation expertise. Esther your skills, knowledge, enthusiasm and love of science spurred me on to finishing the PhD, thanks for all your assistance, especially the fibre isolation sessions. A special thanks goes to Dr Andrew Garnham, your biopsy expertise and guidance for all things participant related, and commitment to many projects. I would also

like to acknowledge all the others who made the time for me during my PhD, Melpomeni Malamouli, Dr. Alba Moreno-Asso, Carlie Bauer and Dr Nicholas Saner.

To family and friends, I am extremely grateful for your support over countless years. To Geoff and Jeanne (Dad and Mum), you have never stopped supporting me through my endless endeavours and without your support I would have had no chance to pursue a PhD. To Laura you have been my biggest supporter throughout this journey, and I am grateful that even when my confidence and will wavered yours never ceased. To Julia and Rianna, being stuck inside for +200 days was made all the easier by having you both willing to listen to the ongoing PhD saga. Jonty and Javi, you guys were forced to be a sporting outlet within the 5km radius of my PhD, thanks for always being available for a quick hit of the outdoors and sport.

Finally, I am most grateful to my enthusiastic and committed participants. Without you, this project, thesis, and science would not progress. I thank you all for your time, energies and commitment.

COVID-19 IMPACT STATEMENT

I started my PhD in March 2020, and shortly thereafter COVID-19 restrictions came into effect. My study required collection of fresh muscle samples and I implemented a 16 week training and testing intervention. According to university policy, and in line with state and national government restrictions, I was unable to commence any training or data collection in the first twelve months of my PhD due to a pause on all human based research. Further, I was deprioritised to enter the biochemistry laboratories to pilot work as I had not passed confirmation of candidature. These restrictions put me fourteen months behind on data collection. Despite this, I recruited participants and implemented all COVID-19 safe requirements to ensure the safety of myself and participants. In 2021 during the training and recruitment phase of this PhD, I was also required to train participants one at a time with extra time before and after sessions to ensure a COVID-19 safe environment, in normal circumstances it is possible to train multiple individuals in the research space. These policies added additional time to intervention studies conducted during 2021 and early 2022. This resulted in a slower than normal recruitment and training phase. Additionally, I have also experienced significant delays in ordering essential reagents and materials due to the impact of COVID-19 on supply chains throughout 2020-2021. I do not believe COVID-19 has impacted the quality of my research but it has led to challenges such as sample collection, motivation and delays in my PhD timeline. It also led to minimal time to complete all desired analyses and further experiments due to the four year timeframe being compacted into two and a half years (with a realistic loss of 18 months due to the COVID-19 disruptions).

Contents

List of Figures

List of Tables

Abbreviations

5mC	5-methylcytosine
ATP	adenosine triphosphate
BH	Benjamani-Hochberg
BMI	body mass index
BMIQ	beta-mixture quantile normalisation
BP	biological process
CC	cellular component
CpG	CpG dinucleotide pairs
CV	coefficient of variation
DEP	differentially expressed protein
detp	Detection P-value
DMA	differential methylation analysis
DMP	differentially methylated position
DMR	differentially methylated region
DNA	deoxyribonucleic acid
DNAm	DNA methylation
ECM	extracellular matrix
EPIC	Infinium MethylationEPIC Beadchip
EWAS	Epigenome wide association studies
FCS	functional class scoring
FDR	false discovery rate
Fe-S	iron sulfide
gDNA	genomic DNA
GO	gene ontology
GO-ORA	Gene ontology overrepresentation analysis
GSEA	gene set enrichment analysis
GXT	graded exercise test
HAD	3-hydroxyacyl-CoA-dehydrogenase
HIIT	High Intentisy Interval Training

IgG	Immunoglobulin G
IgM	Immunoglobulin M
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC-MS/MS	liquid chromatography with tandem mass spectrometry
LFQ	label-free quantification
LT	lactate threshold
MDS	multidimensional scaling
MF	molecular function
mitoDMP	mitochondrial DMP
MLC	myosin light chain
MNAR	missing not at random
MS	mass spectrometry
MSC	Mesenchymal stem cell
MYH	myosin heavy chain
MyoD	myoblast determination protein 1
NHMRC	National Health and Medical Research Council
ORA	overrepresentation analysis
PC	principal component
PCA	principal component analysis
PCr	phosphocreatine
PCR	principal component regression
QC	quality control
RNA	ribonucleic acid
rpm	revolutions per minute
RRBS	reduced representation bisulfite sequencing
RT	Room temperature
SC	Satellite cell
SDH	succinate dehydrogenase
SDS-PAGE	SDS gel electrophoresis
SNP	single nucleotide polymorphism

T2D	Type 2 diabetes
TI	Type I
TII	Type II
t-SNE	t-distributed stochastic neighbor embedding
VO2	maximal oxygen uptake
WM	whole muscle
Wpeak	peak aerobic power

Chemical Abbreviations

5-Aza	5-Azacytidine
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
PBS	Phosphate-Buffered Saline
PCI	Phenol Choloroform Isoamyl Alcohol
PFA	Paraformaldhyde
PVDF	Polyvinylidene Fluoride
SDS	Sodium Dodecyl Sulfate
TBS	Tris-Buffered Saline
TBS-T	Tris-Buffered Saline with Tween
TRIS	Trishydroxymethylaminomethane
Tris-HCl	Trishydroxymethylaminomethane Hydrochloric Acid

List of Symbols

gram	g
nanogram	ng
milligram	mg
millilitre	mL
watts	W
molar	M

1 Introduction

"...there is nothing more interesting than the theory of cells.... In reality, this genesis of cells is something that closely concerns everybody, for all of us, like plants and animals, consist of and have our origin in cells." Ernst Haeckel

[1, 2]

Skeletal muscle is one of the largest organs in the human body, comprising approximately 40-60% of body mass, and is vital for movement, posture and whole-body metabolism. Muscle is known to adapt and change throughout life and regular exercise can cause profound changes in muscle structure and function. At the cellular level, the coordination of specialised cell types enable skeletal muscle function. These cells and their interactions underpin skeletal muscle health and adaptation and are known to respond differently to exercise.

Muscle fibres are multi-nucleated cells that contract to enable movement. Skeletal muscle fibres have long been characterised as slow (Type I) and fast (Type II) fibres. Each fibre type has unique contractile properties and their proportions differ between muscles, biopsy location and individuals, particularly the sexes. In addition, muscle tissue contains a range of mono-nucleated cells such as vascular, immune and stem cells. Despite each cell type having the same stable deoxyribonucleic acid (DNA) sequence, they have distinct functions and morphologies. These differences are partly explained by epigenetic mechanisms.

Epigenetics is the study of mechanisms which alter DNA structure, access and transcription without altering the underlying nucleotide sequence. DNA methylation (DNAm), the chemical addition of a methyl group to cytosine nucleotides, is an important epigenetic mechanism that influences cell-type specificity. To date, the study of DNAm in skeletal muscle has predominantly focused on alterations at the whole tissue level. This may lead to misleading conclusions and may mask important DNAm changes occurring in a subset of skeletal muscle cells.

The ability to study DNAm in purified-fibre-types of skeletal muscle helps to address common limitations of global tissue studies. Limitations include the following: 1) fibre-type proportions of tissues differ between individuals and between muscle samples of the same individual; 2) skeletal muscle fibres are categorised by their functional properties into fast and slow fibres and their proportions can vary due to age, sex, training status or biopsy location; 3) DNAm differences between cell types are larger than DNAm differences of the same cell-type between individuals 4) skeletal muscle cells are known to respond to exercise in a fibre-type-specific manner. Only one study to date has investigated the methylome of slow and fast skeletal muscle fibres in humans. This study showed that muscle fibre types have a distinct methylome but was limited to a sample size of one sample per fibre type [3].

This thesis contains the following chapters. In chapter 2, I provide a literature review on the connection between cell type identity and epigenetics, and discuss the importance of studying fibre-type-specific DNAm in human skeletal muscle at rest and in response to exercise. Chapter 3 provides an overview of the ethics, participants, intervention and methods encompassed in this thesis. In chapter 4, I have developed a novel method to study both the methylome and proteome of human Type I and Type II skeletal muscle fibres. In chapter 5, I have utilised this method to investigate fibre-type-specific DNAm differences in Type I and Type II muscle fibres using both male and female samples. In chapter 6, I have studied the fibre-type-specific proteome of Type I and Type II fibres from the same samples utilised in chapter 5 and compared the methylome and proteome in a fibre-type-specific manner. In chapter 7, I assessed how exercise can shift the methylome and proteome of Type I and Type II muscle fibres. Finally, chapter 8 provides a discussion of the results presented in this thesis and offers suggestions for future research.

2 Background

The background literature in this thesis is organised into five sections with a sixth section outlining the aims and hypotheses:

Section 1 - An overview of the importance of DNA and epigenetics in cell type specification with a focus on the role of DNAm in cell type determination and identity.

Section 2 - A discussion on the importance of skeletal muscle in health and disease and the beneficial role of exercise. It also provides an overview of skeletal muscle cell type composition and the role of studying a complex tissue in terms of cell types.

Section 3 - A discussion of the role of DNAm in skeletal muscle as a mediator between environment and genetics.

Section 4 - An analysis of the current knowledge of fibre-type specific DNAm and the fibre-type specific responses to exercise and a discussion of fibre-type specific proteome differences.

Section 5 - An outline of the open questions in consideration of the literature discussed.

Section 6 - The overall aims and hypotheses of the thesis.

2.1 Section 1: From DNA to Multicellular Organisms

2.1.1 DNA

It is a profound moment when two reproductive cells fuse providing all the information to produce an organism. This initial cell (known as a zygote) is set on a course of development that will culminate in a new multicellular organism. By the late 19th century it was hypothesised that the cell and it's nucleus contain this information for life [2]. Ernst Haeckel first proposed that the cell nucleus was the site that handles heritable characteristics [4], suggesting that the cell nucleus is pivotal in passing information from one generation to the next. At the end of the 19th century Walther Flemming was studying a substance he called chromatin and showed that during cellular division (mitosis) the chromatin in the nucleus divides ending in two daughter nuclei [5].

Around the same time Johann Friedrich Miescher had isolated a crude mixture of DNA and protein that he called nuclein [6]. The nuclein was likely chromatin as experiments confirmed that it contained both histone proteins [7] and nucleic acids [8]. The chemical nature of DNA was determined to be made of the four nucleic acids; adenine (A), cytosine (C), guanine (G) and thymine (T) [9]. The DNA was tightly associated with proteins in the nucleus in structures called chromosomes, and the chromosome theory was proposed stating that genes (a region of DNA that codes for ribonucleic acid (RNA) or protein) are found on the chromosomes and are Mendel's units of inheritance [10, 11]. Avery and others [12] showed that DNA was able to transform one bacteria to another type, finally providing experimental evidence that DNA was the hereditary material containing the information of life.

The DNA structure was revealed to be the four nucleic acids arranged in a helical structure (Figure 2.1) that immediately provided a mechanism for DNA to divide and replicate on cell division [13, 14, 15]. DNA is like a musical score that holds all the notes required for life to develop, play and become immortalised and in 2001 after 142 years of decoding the music of life, the first score of *Homo Sapiens* was published in the form of the human genome project [16, 17]. Since this major scientific milestone there has been an exponential increase in the understanding of the human genome.

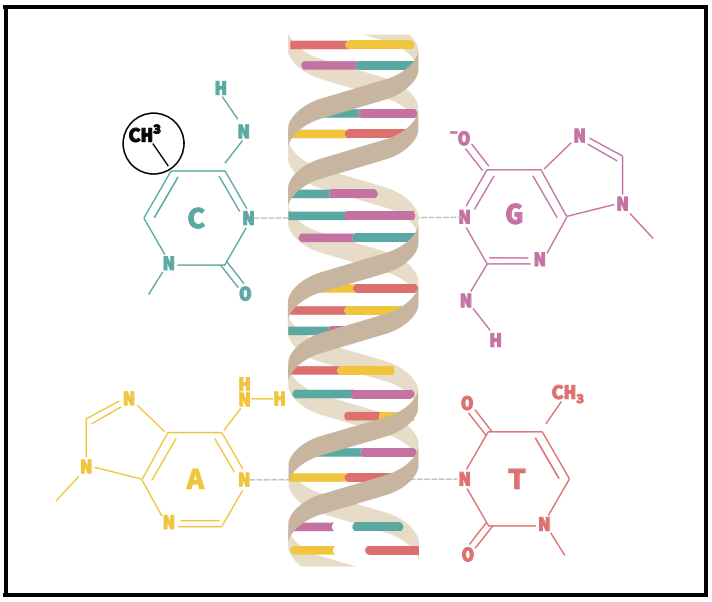


Figure 2.1: Helical Structure of DNA

A representation of DNA with a methyl group in black on the fifth carbon of cytosine.

2.1.2 Epigenetics

The entire score of a human is in the DNA of the very first cell. The human body builds itself through multiple rounds of mitosis as the cells take on specific roles coded into the DNA. Despite containing the same DNA, cells start to express different genes differentiating into the different cell types that make up a multicellular organism. This process is significantly influenced by the chemical modifications on DNA that can alter the expression of gene sets without altering the underlying DNA sequence. This has been termed "Epigenetics". If DNA is the musical score, epigenetics can be considered the orchestra that generates the complexity contained in DNA.

Epigenetics had its foundation in two separate ideas [18]. Waddington [19] initially described epigenetics as the mechanisms by which phenotypes arise from a specific genotype and Nanney [20] added that epigenetic control systems may perpetuate expression patterns in cells and that these can persist during cellular division. This persistence has led others to describe epigenetics as a property that can enable a cell to retain information of a previous event. Lappalainen and Greally *"define an epigenetic property as that of a cell mediated by genomic regulators, conferring on the cell the ability to remember a past event"* [21]. Furthermore, attempts have been made to capture the diverse meanings used in the field such as the definition suggested by Bird that epigenetics is *"the structural adaptations of*

chromosomal regions so as to register, signal or perpetuate altered activity states"[22]. Despite differences in definitions, two central questions in epigenetics research are: 1) how does a multicellular organism form from a zygote that contains the entire genetic code [23, 24]; 2) how can cell identity be maintained throughout the organisms' lifespan [25]? The alteration of chromatin structure via histone modifications, the emerging role of non-coding RNAs in gene silencing, and the widely studied DNAm are all epigenetic mechanisms that orchestrate the development of a multicellular organism.

Histone modifications

Each cell contains approximately 2 metres of DNA yet is packaged in to a nucleus with a diameter of 2-20 µm. The efficiency with which the DNA is packaged into the nucleus is due to the wrapping of DNA around histone proteins [7, 26, 27]. Histone proteins are organised in octamers consisting of two each of H2A, H2B, H3 and H4 histones [28]. Histones have small sections of amino acid tails that can be modified via the attachment of molecules [29]. Histone modifications have been shown to influence RNA synthesis [30] and interact to either compact DNA in chromosomes (heterochromatin) or loosen them (euchromatin) influencing gene expression [31].

Non Coding RNAs

Only 1-2% of the DNA sequence codes for proteins, leaving the vast majority as non coding DNA. These regions of the genome contain non-coding RNA (ncRNA) genes that can regulate translation [32]. MicroRNAs, PIWI-interacting RNAs, small interfering RNAs, enhancer RNAs and long non-coding RNAs are all types of ncRNAs that influence gene expression [33]. ncRNAs can modulate epigenetic regulation but are less well understood than histone modifications and DNAm [34].

DNA Methylation

DNAm is a chemical modification where a methyl group is bonded to the fifth carbon of cytosine nucleotides of DNA [35] (Figure 2.1). The accumulation of 5-methylcytosine (5mC) in humans occurs most frequently in CpG dinucleotide pairs (CpG) [36, 37]. Although CpGs are relatively sparse throughout the mammalian genome, accounting for 1% of nucleotide pairs [38], CpG DNAm is important in gene expression and regula-

tion [39]. The discovery of methylated CpGs and DNA methyltransferases [40] provide evidence that epigenetic patterns can be maintained during cell division [41]. This mechanism enables a cell to pass on it's identity to a daughter cell and illustrates an important role of DNAm in cell differentiation and cell specification. DNAm is the focus of this thesis and will be discussed in further detail.

2.1.3 DNAm

Developmental biologists have studied how cells divide and differentiate to form multicellular organisms, and have identified the importance in the coordinated and controlled expression of gene sets to enable the development of distinct cell types with similar morphologies and functions. One mechanism by which DNAm influences cellular identity is through gene regulation. The genomic location and frequency of methylated and unmethylated cytosines play an integral role in this process. Regions dense in CpGs are called CpG islands [42] and are often found in promoter and enhancer regions of genes and are usually unmethylated [43]. Methylation of CpG islands normally represses gene expression [44]. In contrast, DNAm in gene bodies seems to be non-repressive [45], highlighting the complex role of DNAm. Further, emphasising this complexity is that DNAm displays the contradictory nature of being both stable (that is, perpetuates cell identity) and flexible, with the ability to play a dynamic role during an organism's lifespan. [46, 47]. Epigenetic regulation of the genome enables cells to develop functional diversity [48] and DNAm is a stable epigenetic mark that enables cellular identity to be maintained [49].

DNAm can be likened to the conductor in the orchestra, that has potential to direct higher levels of epigenetic modifications [50] that dictate cellular identity. The picture of an orchestra and conductor is a helpful analogy of the interplay between histone modifications and DNAm. Both play an important role in the production of the music. Likewise histone modifications can help direct methylation of DNA and DNAm offers a scaffold for which histone modifications can be reestablished after cellular division [51].

2.1.4 Cell Type

Robert Hooke coined the term cell in the 15th century when studying cork under magnification [52], and shortly thereafter Antony von Leuwenheuk described cells as living

entities. Further observations led to the development of the cell theory; which postulates that all living organisms are made up of one or more cells and find their origin in a cell [53]. In the analogy of DNA as a musical score, cells can be considered the musicians that create the music. The musicians can be clustered into different sections with different instruments. This is just like defining the different cell types in the human body some similar to others and some quite different, but all of them reading off a specific part of the music to create the whole musical piece.

There is large diversity in cell types in humans and despite containing the same genetic sequence and starting from the same single cell, these cell types express different sets of genes and proteins. The set of genes and proteins distinguishes the shape, size and function of cells allowing the generation of complex tissues and organisms [54]. There is mounting evidence that a small number of distinct transcription factors can even force a cell to switch fate [55]. The definition of cell type has been proposed to contain three parts: a cell type is defined based on phenotype; a cell type has a particular historical lineage; and cell types have particular cell states which can respond to stimuli [56]. Making things more difficult, even within a cell-type there is heterogeneity and cells can be further divided into sub-populations of cells [57, 58]. Skeletal muscle fibre types are a key example of such a subdivision with three distinct fibres types [59] (to be further reviewed in later sections).

The fate of a cell's identity is greatly influenced by epigenetic mechanisms [60]. As earlier highlighted, epigenetic mechanisms are key to maintaining altered activity states of genes required for cell type determination [22]. The Roadmap Epigenome Project aims to characterise epigenetic patterns of the multitude of cell types in the human body indicating the importance of epigenetic regulation of cell type [61]. A basic function of DNAm is in altering gene transcription state. During germ line development there is large de novo methylation that helps cell specification [62] and experiments have shown that DNAm is important during cell lineage determination [63]. A key aspect of cell identity is that a particular cell state can remain stable over time [56]. DNAm in somatic cells tend to be more stable than during development [64] and DNAm patterns are a form of memory that is passed on to daughter cells [65, 66]. It is no surprise then that studies of DNAm are often strongly confounded by cell type and sub cell type populations of samples. Greater differences in DNAm have been reported between cell types of a given

individual than between individuals for a given cell type [67]. In recent years there has been a surge in the number of studies reporting the generation of cell type atlases of DNA methylation [49, 68].

2.1.5 Key Point 1

Complex tissues are composed of numerous cell types each with a specific function. Even within a set cell-type there exists sub-cell types and specific cell states. The coordinated expression of gene sets helps dictate cell identity, and DNA methylation is a crucial epigenetic modification that assists in maintaining a stable cell identity through coordination of gene transcription.

2.2 Section 2: Skeletal Muscle

2.2.1 Skeletal Muscle in Health and Disease

Skeletal muscle is vital for movement, posture and metabolism and comprises approximately 40-60% of body mass. One of the first major milestones in human development is the ability to walk and the capacity to maintain this is fundamental to autonomy and health. To initiate movement skeletal muscle uses specialised cells, called muscle fibres (also known as a muscle cell), that contract and relax pulling on the skeleton. The skeletal muscle fibres comprise the largest part of muscle tissue and are the focus of this thesis, however a brief discussion of skeletal muscle composition is provided. The deterioration of muscle mass and muscle fibres with age or disease and the concomitant rise in musculoskeletal and metabolic diseases highlights the importance of studying skeletal muscle health [69, 70, 71, 72].

A distinguishing feature of skeletal muscle is its high proclivity for adaptation and regeneration [73]. This response is largely driven by the stem cell of muscle, termed the satellite cell [74], which maintains the ability to differentiate toward a mature myogenic cell or divide, forming a daughter cell [75, 76, 77]. Satellite cells play an important role in muscle maintenance and growth [78], and contribute to fibre regeneration and repair [73, 79]. Upon insult, quiescent satellite cells are activated and fuse with muscle fibres contributing new myonuclei to the fibres [75]. Another feature of muscle is the sub-

stantial requirement for blood delivery at rest and during activity [80, 81]. As a result, skeletal muscle is highly vascularised to enable efficient delivery of oxygen and nutrients, and removal of waste [82, 83]. There is growing evidence that the endothelium plays a role in skeletal muscle regeneration and maintenance [73, 84]. Interactions between the mono-nucleated cells and multi-nucleated muscle fibres have been reported during muscle regeneration [73], however, their interplay and contribution to muscle maintenance and adaptation remains relatively unknown.

Muscle is an important metabolic organ with a basal metabolic rate approximated to use 20% of energy at rest [85]. Muscle is a major store of glycogen and protein both important for vital bodily functions [86, 87]. A decline in skeletal muscle mass is linked to a decline in metabolic health [88]. There has been a rise in the prevalence and burden of metabolic syndrome [89]: defined as a disease associated with a clustering of the following conditions; increases in insulin resistance, excessive adipose deposits, hypertension and an imbalance of lipids [90]. Skeletal muscle is one of the most important tissues for insulin stimulated glucose uptake and dysregulation of insulin action in muscle is a main contributor to metabolic syndrome [91]. Further studies of skeletal muscle metabolism may provide insight in how to counteract the burden of metabolic disease.

The decline in physical performance with age is associated with a decline in muscle mass [69]. However, the decline in muscle mass is not the sole contributor to this decline, with other aspects of muscle health such as strength, muscle integrity and altered muscle metabolism also contributing. Evidence suggests that mitochondrial dysfunction with age is also a factor in sarcopenia [92]. There is a need to further understand the molecular mechanisms underpinning skeletal muscle health throughout the lifespan.

2.2.2 Adaptations of Skeletal Muscle to Exercise

Outside of the changes occurring during development, growth and ageing, skeletal muscle shows a dynamic plasticity in response to environmental stimuli, in particular muscle shows a potent response to exercise. As a result exercise is an ideal model to study skeletal muscle adaptations and muscle biology.

The adaptation of muscle to exercise has been extensively studied in both mouse and human models. Exercise induces angiogenesis, changes the composition of muscle-

fibres, alters substrate utilisation in muscle, and stimulates muscle and endocrine cross-talk [93]. Skeletal muscle undergoes both morphological and metabolic changes due to exercise, with an increase in oxidative capacity, muscle mass and muscle strength due to both resistance and endurance training [94, 95, 96]. Exercise can lead to increases in muscle mass [97], improvements in endurance capacity [98], and evidence is accumulating that exercise can be used to treat a range of chronic diseases [99]. At a molecular level, recent studies have highlighted the transcriptomic and proteomic changes in skeletal muscle after exercise [100, 101]. These studies are revealing the beneficial effects of exercise in skeletal muscle health. The importance of exercise is highlighted by findings that physical inactivity is a major risk factor for poor health and in the establishment of a global action plan on physical activity [102]. There is a continued need to understand how exercise and physical activity leads to improved muscle health.

2.2.3 Cell-Type Composition of Skeletal Muscle

Skeletal muscle is a heterogeneous organ comprised of diverse cell types, such as multi-nucleated muscle fibres, and a population of mononucleated cells [103, 104, 105, 106, 107]. Muscle forms from the fusion of myogenic progenitor cells as a result of the coordinated expression of myogenic factors [108]. These progenitor cells differentiate and fuse to form muscle fibres or retain stem-like properties to initiate future repair and growth [76, 77]. This enables muscle to maintain plasticity outside of development with a large ability to regenerate and adapt to environmental cues. Although the majority of the structure and protein of muscle tissue comes from muscle fibres, myonuclei may only make up 50% of the nuclei in muscle tissue [109].

Human skeletal muscle consists of three distinct muscle fibre types categorised by their expression of distinct myosin heavy chain (MYH); Type I (slow-twitch, oxidative), Type IIa (fast-twitch, oxidative-glycolytic-intermediate), Type IIx (fast-twitch, glycolytic) fibres, and a further collection of hybrid fibres [59]. Muscle contraction and function is primarily driven by muscle fibres [110], but is also dependent on the additional functions of the mono-nucleated cells resident within the muscle. Three important structures supporting muscle fibres are the extracellular matrix (ECM) [111], the muscle vasculature and the nerve supply. These structures and their cell types play a coordinated role in the

regeneration of injured skeletal muscle as outlined recently [73]. The following section briefly summarises the resident cells in skeletal muscle (Figure 2.2).

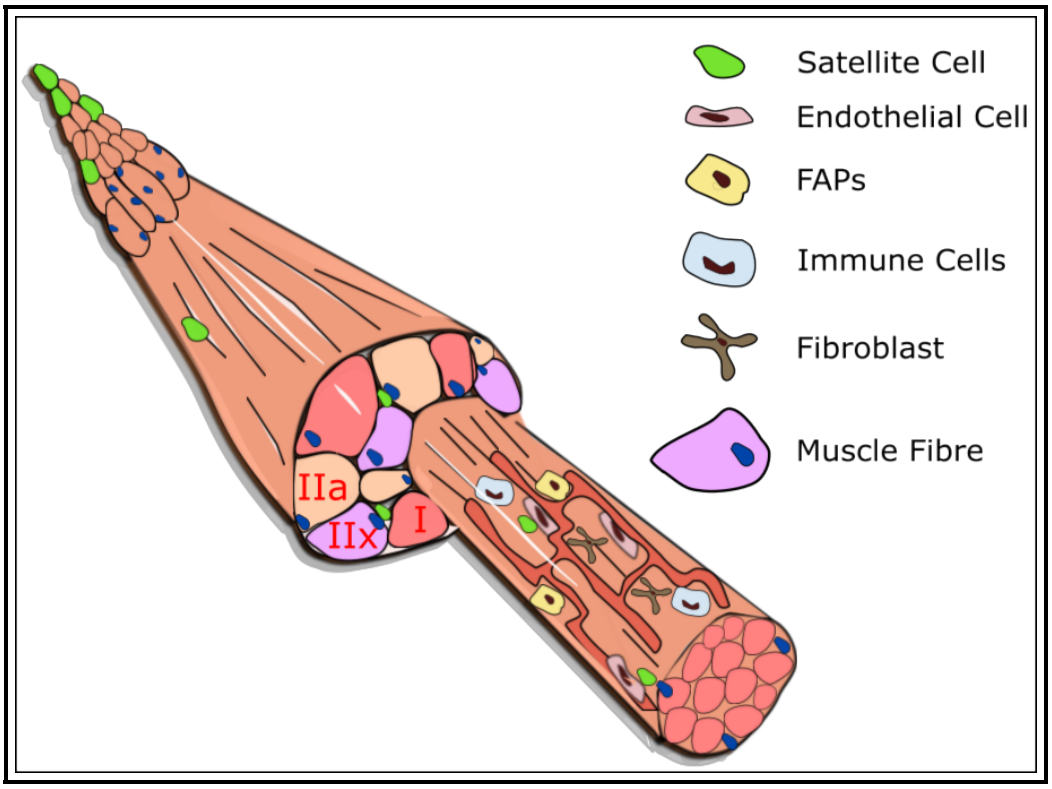


Figure 2.2: Cell-Types of Skeletal Muscle

A representation of skeletal muscle and key resident cell types.

Muscle Fibres

Muscle fibres are multi-nucleated cells that form when myoblasts fuse [112]. The resultant nuclei of the muscle fibre are called myonuclei and are located at the periphery of muscle fibres. Human skeletal muscle consists of three distinct muscle fibre types as categorised by their expression of MYH; Type I, Type IIa, Type IIx, and a further collection of hybrid fibres [59]. Muscle fibres contain the highly specialised contractile proteins actin and myosin that form cross-bridges leading to muscle contraction. There are morphological and functional differences in the action of myosin and actin in the different fibre types. Muscle fibres are the main focus of this work and are outlined in further detail in Section 4.

Fibroblasts

Although muscle fibres spatially constitute the majority of muscle tissue, the ECM plays an important role [113]. Skeletal muscle fibres are surrounded by and intertwined with a structure of support comprising approximately 10% of muscle mass [111]. The ECM contains multiple cell types: stem-like cells, endothelial cells, resident immune cells, fibroblasts and nerve-related cells. Fibroblasts secrete collagens that maintain the ECM enabling it to support the micro-environment of muscle and enable movement [113, 114].

Satellite Cells

A distinguishing feature of skeletal muscle is its high propensity for adaptation and regeneration [73]. Muscle contains multiple progenitor cell types that can form various tissue structures assisting in the process of adaptation. These cells are referred to as muscle stem cells and include satellite cells, mesenchymal stem cells and fibroadipogenic progenitor cells. Satellite cells (SCs) lie between the sarcolemma and the outer membrane of muscle fibres [74]. The satellite cell is called the stem cell of muscle, with the ability to differentiate toward a mature myogenic cell or to divide into a daughter cell [75, 76, 77, 115]. During skeletal muscle development, SCs provide the source of mature muscle cells that can fuse to form muscle fibres. SCs constitute 2-5% of adult skeletal muscle and can be identified by the expression of paired box protein 7 (Pax7) [116]. In postnatal skeletal muscle, SCs normally remain quiescent, but after muscle insult become activated and contribute to fibre regeneration and repair [79], and are indispensable for muscle regeneration [117].

Mesenchymal Stem Cells

Mesenchymal stem cells (MSCs) were first isolated and identified from bone marrow [118, 119] and have since been shown to have myogenic, angiogenic and neurogenic potential [120]. MSCs have been isolated from skeletal muscle [121, 122] and can differentiate into multiple cell lineages. These cells have the potential to contribute to skeletal muscle regeneration and have application for skeletal muscle tissue engineering [123].

Fibro-adipogenic Progenitor and Twist Progenitor Cells

Fibro-adipogenic progenitor cells (FAPs) are a mesenchymal cell population resident in skeletal muscle [124, 125]. FAPs have the potential to form both adipocytes and fibroblasts and contribute to muscle regeneration [124]. Twist 2 positive cells found in the interstitium of skeletal muscle may be involved in muscle regeneration [126, 127].

Endothelial Cells and Pericytes

Skeletal muscle is highly vascularised [128] enabling oxygen and nutrient delivery and removal of waste. Two cell types associated with blood vessels in muscle are endothelial cells and pericytes. Endothelial cells line the walls of blood vessels and may comprise a quarter of muscle nuclei [109]. Angiogenesis and neo-vascularisation in skeletal muscle in coordination with endothelial cells may contribute to muscle regeneration [129, 130].

Pericytes surround endothelial cells of small blood vessels [131] and maintain a close association in rodent and human skeletal muscle [132]. In skeletal muscle pericytes and satellite cells lie in proximity and pericytes have the potential to fuse with muscle fibres during regeneration [133, 134]. Evidence suggests that pericytes contribute to postnatal myofiber growth and satellite cell quiescence [135].

Immune Cells

Immune cells are produced in specialised immune organs but reside in other tissues [136]. Macrophages reside in all tissues [136] and have been identified in both healthy and regenerating mouse and human skeletal muscle [137, 138, 139]. Macrophages may lead to the activation of satellite cells and promote muscle growth and repair [140, 141]. Immune cells identified in muscle are neutrophils, macrophages, regulatory T-cells, eosinophils, B cells and natural killer cells [73, 103, 106]. The finding that macrophages support the co-ordination of myogenesis and angiogenesis during muscle regeneration highlights the need to study muscle cell type interactions in skeletal muscle [142].

2.2.4 Key Point 2

Skeletal muscle is an important tissue for maintaining a high quality of life. Muscle is a complex tissue composed of numerous cell types many of which provide muscle it's abil-

ity to increase in size. Skeletal muscle fibres make up a large portion of muscle and can be further divided in to distinct fibre types. The improvement in sensitivity to measure molecules from low input biological material has led to numerous studies investigating differences between groups of cells. There is growing interest in assessing cell type and fibre type differences and interactions in human skeletal muscle.

2.3 Section 3: DNA Methylation & Skeletal Muscle

2.3.1 DNAm Regulation of Skeletal Muscle

DNAm is an important epigenetic modification that coordinates the development of cell types [66]. Skeletal muscle is an important model for studying the role of DNAm in cell specification and tissue (muscle) development. The forced reprogramming of mouse embryonic cells to striated muscle like cells after treatment with 5-Azacytidine (5-Aza) provided some of the earliest evidence of the importance of DNAm to cell identity [143]. The action of 5-Aza was the inhibition of DNAm. Experiments showed that with a higher concentration of 5-Aza there was an increased inhibition of DNAm and a subsequent increase in the number of reprogrammed muscle cells [144]. The application of 5-Aza with the forced expression of myoblast determination protein 1 (MyoD) was able to convert somatic fibroblast cells to myoblasts [145]. More recent work has reported an improved efficiency to directly reprogram mouse fibroblasts to muscle progenitor cells by using key epigenetic modulators [146, 147]. Indicating that DNAm is regulating cell identity during differentiation in skeletal muscle [148].

These data provide strong experimental evidence explaining the role of DNAm in cell type identity in skeletal muscle. DNAm in skeletal muscle appears to function by changing both the transcription of RNA and translation of proteins. Seminal work has shown that cells with a lower level of methylation in enhancer regions of MyoD in mouse have have high levels of MyoD expression [149]. Further experiments have linked a hypomethylation in muscle cells in myogenic genes to increased protein expression of myogenic proteins [150, 151]. For example, the demethylation of the Myf5 super enhancer in muscle cells increases gene expression of Myf5 and initiates the myogenic program [148]. Numerous cell culture studies have established a clear link between DNAm and altered gene activity in muscle [152]. A recent atlas of skeletal muscle development confirms that

DNAm affects gene expression during muscle development [153]. Outside of development a number of studies have investigated DNAm and transcription in skeletal muscle in a number of different contexts.

2.3.2 Factors associated with DNAm changes in Skeletal Muscle

The study of DNAm in skeletal muscle has identified a number of factors associated with changes in DNAm, such as age, diet, disease and exercise.

Ageing

Studies have shown there is a difference in methylomes between old and young skeletal muscle with increased hypermethylation in older muscle [154, 155], and this has also been found in differentiating human-derived cell cultures [156]. DNAm increases globally during ageing in both mouse and human myogenic cell cultures and increased DNAm in muscle stem cell quiescence pathways reduces the capacity for stem cell self-renewal in older muscle [157]. Epigenetic discordance between mono-zygotic twins increases during ageing in skeletal muscle as a result of environmental stressors and stimuli experienced differently by the twins throughout their lives [158]. Our laboratory and others have recently reported widespread DNAm changes associated with ageing in key muscle genes [159, 160].

Sex

There are significant sex differences in skeletal muscle; such as size [161], metabolic [162] and functional differences [163]. At the molecular level there are global differences in the transcriptome between males and females [164, 165]. Experiments assessing human derived myoblasts have shown differences between male and females in DNA methylation and gene expression in genes involved with energy metabolism [166]. Our group has reported extensive DNAm differences between males and females in skeletal muscle, with a greater overall hypomethylation in male muscle and DNAm sex differences in genes involved in substrate metabolism [167].

Health and Disease

The role of DNAm in health and disease is a burgeoning field of study. Numerous papers report on associations between DNAm and a number of diseases, such as cardiovascular disease and Type 2 diabetes (T2D) [168]. There is evidence showing that DNAm plays a role in type 2 diabetes [169]. DNAm differences have been reported in skeletal muscle of individuals with and without a family history of T2D [170], and twins discordant for T2D have differences in DNAm [171]. Targeted analysis of DNAm differences in skeletal muscle between normal and diabetic participants revealed hypermethylation in PGC1-α in diabetic participants with a corresponding reduction in mitochondria [172]. Differences in DNAm in diabetic individuals may be linked to diet induced changes in DNAm. Differences in skeletal muscle DNAm have been reported after six weeks of a high fat diet compared to a control diet [173]. Although associations have been identified between health outcomes and DNAm within skeletal muscle, there is a large need for further investigations into the role of DNAm and skeletal muscle health.

Exercise and Physical Activity

Epigenetics is thought to mediate muscle tissue changes in response to exercise. Seminal work by Barres and others [174] investigating DNAm in human skeletal muscle after an acute bout of exercise reported a global decrease in DNAm, which was exercise dose-dependent and resulted in increased gene expression of key genes in muscle metabolism (PGC1α, PDK4 and PPARδ). Using genome-wide methods, further studies have investigated the impact of both endurance and resistance training on DNAm [170, 175, 176, 177] Some have reported no overall global changes in DNAm as a result of exercise [176], while others reported an increase in hypomethylation after training [170, 177]. Endurance exercise increased DNAm at enhancers of the myogenic regulatory factors (MRFs), which was unexpected as increased methylation should lead to transcriptional repression [175]. The study reporting no changes in DNAm after 12 weeks of HIIT, did however report gene expression changes [176]. This indicates an incomplete understanding of the interplay between DNAm and gene expression in the context of exercise, and a reason to further study DNAm in muscle. It has been reported that there is a greater hypomethylation after a second resistance training intervention compared with an initial intervention [177, 178], suggesting that the methylome encodes a cellular memory of previous exercise. However, the capacity for DNAm to mediate a memory of previous exercise remains

to be firmly established. Literature so far suggests that exercise may alter DNAm patterns in muscle.

Another consideration when studying skeletal muscle adaptations to exercise is the different response of females and males to varied stimulus. It has been reported that there are sex differences in response to both resistance exercise [179] and endurance exercise [180] between males and females. For example, it appears that females have higher expression of key metabolic enzymes after endurance training [180]. Different responses to exercise in male and female muscle may be driven by differences in DNAm and exercise may lead to sex-specific DNAm changes. Our laboratory has conducted a DNAm assessment of sex specific changes to exercise, and reported no changes in DNAm after four weeks of exercise [181]. However, due to notable sex differences in DNAm at baseline, future studies that involve exercise should consider sex-specific differences and aim to recruit both males and females.

2.3.3 Key Point 3

DNAm plays an important role in cell identity and specification in skeletal muscle. Along with this, a number of factors are associated with changes in DNAm in whole skeletal muscle samples, however, the interpretation of DNAm changes in bulk muscle is challenging due to the important role of DNAm in both fibre sub-populations and cell types. The following section describe the importance of studying skeletal muscle in a fibre-type specific manner by highlighting that skeletal muscle fibre-type differences are apparent in multiple contexts, suggesting the need to extend the study of skeletal muscle DNAm to muscle fibre types.

2.4 Section 4: Skeletal Muscle Fibres

2.4.1 Skeletal Muscle Fibres in Focus

Skeletal muscle initiates and maintains movement, and is vital for maintaining a high quality of life. The energy demands of skeletal muscle increases dramatically during physical activity with an estimated 100 fold increase with exercise [182]. The performance of muscle is largely driven by skeletal muscle fibres as previously described [183]. In

human muscle these fibres have been characterised into three distinct sub-types each with a distinctive functional and metabolic phenotype [59]. The demands during long sustained or short powerful movements rely on the capacity of either Type I slow or Type II (TIIa/TIIx) fast skeletal muscle fibre types respectively. Type I fibres are more predominant in endurance athletes and Type IIa/Type IIx fibres are more abundant in power-based athletes [184]. Skeletal muscle fibres make up the bulk of muscle tissue although a number of other cell types are present in muscle [103, 106]. It is estimated that muscle fibres contribute 50-70% of the nuclei in skeletal muscle depending on the muscle and the animal [109, 185, 186, 187]. Skeletal muscle fibres have been extensively studied due to their importance to muscle performance, and this section briefly describes the differences between Type I, Type IIa and Type IIx fibres.

Size, function & metabolic differences between fibre types

Initial studies of skeletal muscle fibres were conducted on cross-sections of muscle. The cross-sectional approach to study skeletal muscle fibres restricts studies predominantly to questions of size and distribution. Type II muscle fibres are larger than Type I fibres in male muscle with fibre size observed in the order TIIa > TIIx > TI [188, 189, 190], however in females the results are mixed [188, 191, 190]. A number of studies indicate fibre type size differences in the order TI > TIIa > TIIx in female muscle [188]. Moving beyond cross-sectional area, a number of studies have extensively studied single fibre functional and metabolic properties in human skeletal muscle. Type IIa and Type IIx muscle fibres have higher maximum shortening velocities and maximum power output compared to Type I fibres [192, 193]. The metabolic properties of skeletal muscle fibre types have been well documented. The volume of mitochondria differs in muscle fibre types with a higher mitochondria volume in the following order TI > TIIa > TIIx [194], and Type I fibres have more interconnected mitochondrial networks than Type II fibres [195, 196]. These mitochondrial differences are accompanied by differences in energy metabolism between the muscle fibre types [183]. Type I muscle fibres rely on oxidative metabolic systems and have been shown to have higher levels of succinate dehydrogenase (SDH) and 3-hydroxyacyl-CoA-dehydrogenase (HAD) than Type II fibres [197], and Type II fast fibres have higher levels of glycogenolysis enzymes than Type I fibres [198]. The differences in contractile and metabolic properties of the muscle fibre types enable muscle

to perform varied movements from long sustained to short powerful contractions, and there is evidence that the muscle fibre types utilise energy systems differently.

2.4.2 Factors influencing fibre-type in skeletal muscle

Ageing

The effects of ageing appear to differ between the fibre types. Skeletal muscle mass decreases during ageing and is attributed to a loss of motor-neuron number which results in a loss of muscle fibre number and a reduction in muscle fibre size [199, 200]. The evidence has been conflicting regarding Type I muscle fibre changes in size with ageing, a recent review concludes that Type I fibre size does not change with ageing [201]. In contrast, Type IIa/TIIx fibre size decreases with age in both men and women [201]. The functional properties of skeletal muscle fibres aligns with the reductions in size, with a Type I fibre contractile function maintained with age, however, Type IIa/TIIx fibres have a reduction in contractile function with age [201]. Interestingly, a study examining single fibres in 90 year-olds report greater contractile properties in both Type I and Type IIa skeletal muscle fibres in advanced age compared with 20 year-olds, suggesting that the quality of remaining muscle fibres is high [202]. In addition to size and contractile function, Type II fibres have reduced metabolic properties with age [203]. Recent proteomics identified a reduction in glycolytic proteins in Type II fast fibre populations compared with Type I fibres, offering further evidence for a greater loss of function in Type II fibres with age [203].

Sex

Cell composition of muscle may also vary between males and females. Females have a tendency for higher proportions of Type I fibres compared to males and a greater area of muscle is composed of Type IIa fibres in males compared with a greater area of Type I fibres in females [190, 204]. Muscle fibre cross sectional area is larger in males than females for both Type I and Type IIa fibres [205]. There appears to be a difference in satellite cell distribution between females and males, with females showing a lower satellite cell content of Type II compared with Type I fibres and females had a lower content of SCs in Type II fibres compared to males [206]. Evidence from our laboratory suggests

differences in DNAm between the sexes in skeletal muscle and indicated that fibre type proportion differences between males and females may explain DNAm differences [167].

2.4.3 Transcriptomic & Proteomic differences between Fibre Types

The study of molecular differences between skeletal muscle fibre types have increased with high-throughput transcriptomics and proteomics. Gene expression differences between skeletal muscle fibre types have been investigated in rodent muscle [207], and a number of genes were differentially regulated between the fibre types in mouse muscle [208]. The study of single fibre and pooled fibre type proteomics has progressed in the last 10 years with improvements in technology. Single fibre and pooled fibre type proteomics has revealed fibre-type specific protein patterns in Type I, Type IIa and Type IIx muscle fibres in human skeletal muscle, with clear differences in key contractile and metabolic proteins [209, 210]. A recent pre-print reported both a transcriptomic and proteomic single fibre analysis on Type I, Type IIa and Type IIx fibres [211]. Their results confirmed that Type I and Type IIa fibres have distinct contractile and metabolic proteomes and transcriptomes, however they challenged the conventional wisdom of a distinct Type IIx fibre population [211]. The study of fibre-type proteomics and transcriptomics has the ability to identify changes that can not be seen using whole muscle samples [212]. The recent data indicates the need to also assess fibre-type specific DNAm in human skeletal muscle.

2.4.4 DNAm regulation of Skeletal Muscle Fibre Types

DNAm profiles of skeletal muscle fibre types are not well characterised. The only study to assess DNAm differences between Type I and Type IIa human muscle fibres reported DNAm differences in key muscle-related genes [3]. However, this study did not assess fibre type profiles of females or Type IIx fibres, and they did not assess individual variation. A more recent study assessed DNAm differences between fast and slow muscle fibres in mouse muscle [213] showing a similar set of differentially methylated genes as Begue and others [3]. Oe and others identified a strong hypermethylation of promoters particularly in slow fibres [213]. The importance of muscle fibre-type on DNAm is further highlighted by a comprehensive study that attempted to integrate diverse omic data sets to link DNAm with gene expression. The authors assessed the correlation between

gene expression, DNAm and physiological traits in human skeletal muscle, and found that controlling for tissue composition and fibre type reduced the number of associations identified between genes and DNAm sites [214]. The authors did not isolate single fibres to assess fibre-specific DNAm differences. Current methods to adjust for fibre type and cell type heterogeneity in DNAm studies of whole muscle tissue are limited. Commenting on results from their comprehensive study, Taylor and others [214] provide the following insight into future research, *"Overall, these results emphasise the importance of tissue/cell type composition as a component of physiological traits and the need for single cell data, either for the study of samples or as a source of cell type signatures for more accurate estimates of tissue composition"*. [214]. As earlier discussed, recent work from our laboratory made a similar observation that fibre type proportion was an underlying explanation for sex differences in DNAm, indicating the importance of fibre-type identity on DNAm patterns.

2.4.5 Fibre-Type-Specific Changes of Muscle to Exercise

Size & distribution

The role of exercise on morphological changes such as fibre size, proportion and distribution have been extensively investigated using microscopy throughout the latter half of the 20th century. In 1972, Gollnick and others [215] compared Type I and Type II fibre percentages in men with a history of sustained participation in a variety of sports. Their study showed that men who took part in endurance like sports had a higher percentage of slow-twitch fibres (Type I) compared to resistance/sprint trained men or untrained men [215]. Inversely resistance/sprint trained men had a higher percentage of fast-twitch fibres (Type II) than endurance trained men. The results suggest that the mode of exercise may influence the structure of muscle tissue in a fibre specific manner.

Subsequent intervention studies have shown that skeletal muscle fibres have the ability to shift in response to endurance and resistance based exercise training. There is a tendency that the proportions of Type IIx fibres decrease with a subsequent increase in Type IIa fibres after high intensity and moderate endurance training [216, 217, 218, 219, 194, 220, 221, 222]. Both short and long term resistance training results in an increase in fibre size with a preference for an increase in Type II fibre area [223, 224, 225, 226], and a similar muscle fibre shift away from Type IIx toward Type IIa [227, 228, 229, 230, 231, 232, 233]. Hybrid and intermediate fibres also change as a result of exercise [234, 235].

Studies also confirm that resistance and exercise training cause a shift in fibres in hybrid fibre populations, with an increase in Type I and Type IIa fibre populations however no change in Type IIx [236].

Contractile function

Observational studies assessing lifelong exercise on single muscle fibre properties have shown that despite the ageing cohorts, exercise leads to higher force production in muscle fibres [237]. Gries and others [238] showed that in elderly women who had undertaken lifelong exercise fibre power, strength and speed were maintained despite reductions in Type IIa fibre size. Resistance exercise improves contractile properties of single muscle fibres with improvements in contractile velocity and power in both Type I and Type IIa fibres in men [239] and increases in absolute power in older women [240]. Resistance training has been shown to improve the specific force and shortening velocity of single fibres in females [241]. However, evidence for a change in contractile properties in the elderly after resistance exercise is conflicting with some showing improved contractile properties and others showing no improvements [242, 243]. There is also evidence that endurance exercise leads to fibre specific changes in contractile function. Marathon training decreased both Type I and Type IIa muscle fibre size whilst contractile function remained unchanged [244], however aerobic training in elderly women increased the power of Type I fibres but not Type IIa fibres [245]. In both young and elderly males 12 weeks aerobic training increased contractile power of Type I fibres, but contractile power of Type IIa fibres only increased after exercise in elderly males [246]. There is evidence that fibre contractile properties are influenced by exercise in a fibre-type-specific manner and that there are both age and sex specific adaptations.

Metabolic function

Exercise induces fibre-type specific metabolic responses of skeletal muscle, many of which are related to energy substrate utilisation. Sprint training alters adenosine triphosphate (ATP), phosphocreatine (PCr) and glycogen more in Type II fibres compared to Type I fibres [247]. Further studies showed an increased utilisation of PCr in Type IIx and Type IIa fibres with utilisation of ATP only in Type II fibres and not in Type I fibres during short maximal exercise [248]. Glycogen content of Type II fibres is dramatically decreased

compared to Type I fibres in response to high intensity exercise [249, 250]. The fibre-type differences in metabolic substrate utilisation during exercise are firmly established. Studying purely whole muscle limits the ability to determine fibre-type specific metabolic changes to exercise and future work should investigate the role of DNAm in regulation of fibre-type metabolism.

Proteome & Transcriptome

A number of studies have reported the transcriptome and proteome differences of fibre types using either single or pooled muscle fibres [210, 251]. The study of fibre type specific mRNA changes have revealed that aerobic exercise causes an increase in the number of fibres expressing MYH7 and a reduction in the number of fibres expressing MYH2 [252]. An acute session of resistance exercise increased the expression of proteolytic genes in both Type I and Type IIa muscle fibres, with Type I fibres having a higher expression both at rest and after exercise [253]. The assessment of the transcriptome after resistance exercise showed that Type IIa fibres had a significant increase in gene expression compared with Type I fibres [254]. On the level of proteins, protein synthesis rates differ between Type I and Type II fibres after resistance exercise [255]. A recent study assessed the proteome after endurance exercise and reported differential responses of Type I and Type II muscle fibres [210]. The results of transcriptomic and proteomic studies investigating the response of muscle fibre types to exercise is beginning to reveal the underlying molecular adaptations driving global skeletal muscle changes. As of yet a global assessment of the methylome after exercise has not been investigated and is warranted considering the clear fibre-type specific responses to exercise.

2.4.6 Key Point 4

Skeletal muscle enables both fast powerful and slow sustained movements and the energy demands of muscle increase with exercise. Skeletal muscle fibre types enable these different movements and energy abilities of muscle and exercise causes fibre-type-specific responses of muscle. There is growing evidence that muscle causes specific fibre type transcriptomic and proteomic changes. To date no studies have investigated the potential shifts in methylation of muscle fibres after exercise and studying changes in DNAm in whole skeletal muscle might mask interesting fibre-type DNAm changes to exercise.

2.5 Section 5: Open Questions

Appreciating that skeletal muscle is a heterogeneous tissue and considering the role DNAm plays in perpetuating cell identities, there is a need to identify fibre-type DNAm profiles of human skeletal muscle. To our knowledge, the challenge of studying DNAm in a fibre-type specific manner in human muscle has been restricted to one individual study assessing the differences of Type I and Type II muscle fibres [3]. The study was restricted to a sample size of one and did not investigate fibre-type specific DNAm in female muscle. Future research in skeletal muscle should aim to broaden studies to include both males and females and include multiple samples. There is a pressing need to extend the study of DNAm in skeletal muscle fibre types considering the evidence that skeletal muscle function, metabolic properties, the proteome and the transcriptome differ between the fibre-types.

The nature and role of the epigenetic landscape is still developing, and although integral to cell specification, evidence is growing that the environment has an effect on the epigenome [256]. Studies should also assess the impact of the environment on the epigenome of both fibre-types. As indicated earlier, fibre-type variations are major factors to consider in DNAm studies in exercise and muscle adaptation. Studies have shown that muscle responds to exercise in a fibre-type-specific manner. For example, different modes of exercise cause fibre-type changes, such as growth [257], and shifts in fibre-type proportions [258]. Specifically, HIIT has been shown to decrease the quantity of Type IIx fibres [219], and cause enzymatic changes in Type IIa fibres [259]. The findings described in the previous sections highlight the need to study fibre-specific DNAm changes in response to exercise.

2.6 Section 6 Aims and Hypotheses

The aims of this doctoral thesis presented in the following chapters were:

1. To improve existing methods for the extraction of DNA and protein from pooled Type I, Type IIa and Type IIx skeletal muscle fibre segments.

   I expect that we can isolate DNA and protein from pooled muscle fibre types without cutting the fibre using an approach that is scalable.

2. To determine the DNAm profiles of Type I, Type IIa and Type IIx muscle fibres in male and female skeletal muscle.

   I expect that Type I, Type IIa and Type IIx muscle fibres will have distinct methylomes that are not influenced by sex and that such methylomes can be utilised as a reference for future work.

3. To identify differential shifts in DNAm between Type I, Type II muscle fibres before and after 12 weeks of HIIT in male and females.

   I predict that 12 weeks HIIT will alter the methylome in Type II muscle fibres in key genes involved in mitochondrial function and that this will be reflected in mitochondrial networks.

4. To compare fibre-type specific DNAm results with fibre-type specific proteome data extracted from the same sample before and after 12 weeks of HIIT.

3 Materials & Methods

3.1 Ethics, participants and study design

3.1.1 Ethics approval

This study was approved by the human ethics and research committee at Victoria University (HRE13-223, HRE21-122). Written and informed consent was obtained from all participants prior to commencement.

3.1.2 Participants

Fifteen healthy, non-smoking, young males (n=8) and females (n=7) volunteered to participate in the study (participant characteristics are summarised in Table 5.1). A detailed medical history was taken on the first visit and participants were excluded if they had a history of the following medical conditions: coronary heart disease, significant or chronic respiratory conditions, major musculoskeletal problems, uncontrolled endocrine or metabolic disorders or diabetes requiring insulin. Presence of the following additional criteria led to the exclusion of participants from the study: females taking hormonal contraceptives, currently pregnant or breast feeding and any participant non-compliant with the training intervention.

3.1.3 Study design

This thesis was part of the ongoing Skeletal Muscle Adaptive Response to Training (Gene SMART) study and was designed accordingly [260]. Participants completed familiarisation (see section Familiarisation visit) and exercise testing prior to the commencement of training. The complete study design is outlined in Figure 3.1. All participants completed twelve weeks of High Intentisy Interval Training (HIIT) that consisted of 36 individual sessions (three times per week) on a Velotron cycle ergometer (Velotron Dynafit Pro, Racermate, Inc., Seattle, WA, USA). Each training session consisted of six to fourteen bouts of two minute cycling intervals (performed at individually determined workloads) interspersed with one minute active rest intervals (performed at 30-60 Watts). Muscle biopsies

were taken before, at four weeks, and at twelve weeks at the completion of the training program. All muscle biopsies were collected 48 hours after testing/training.

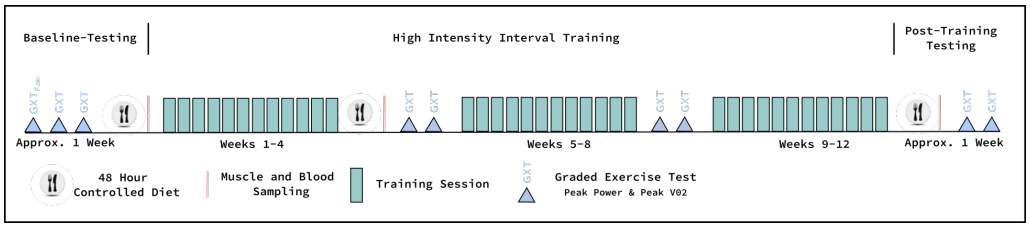


Figure 3.1: Study Overview

An overview of the study design showing testing, training and biopsy timepoints.

Standard set-up of the COSMED system

Prior to each graded exercise test (GXT) and maximal oxygen uptake (VO_{2peak}) test the COSMED gas analysers were calibrated with calibration gases and the flow turbine (COSMED) was calibrated with a 3L calibration syringe following the COSMED calibration program. The COSMED mixing chamber application was used for all $V0_2$ measurements

Familiarisation visit

On the day of the first visit, participants completed medical screening and genetic background questionnaires, and were provided the opportunity to discuss the study with research staff before giving written informed consent of participation. The participants then completed a familiarisation session of a GXT and VO_{2peak} test as outlined below. After each exercise test and biopsy visit, participants were provided a small UP and GO (Sanitarium) liquid based meal to aid in recovery.

Graded exercise test (GXT)

During the second and third visits, participants completed one GXT to exhaustion to determine baseline lactate threshold (LT) and peak aerobic power (Wpeak). The GXT was performed at the same time as the VO_{2peak} test. The test was performed on an electronically-braked cycle ergometer (Lode-Excalibur Sport, Groningen, the Netherlands) and consisted of four minute stages separated by 30 s rest periods until exhaustion. Before commencement, during each 30 s rest, and at exhaustion, capillary blood was sampled from the fingertip and analysed for whole blood lactate (La) and glucose concen-

trations on a YSI2300 STAT Plus system (Yellow Springs, Ohio, USA). For female participants the test started at either 25 or 50 depending on results from the familiarisation test. Likewise, male participants started at 30 , 60 or 120 W. Each four minute stage was increased by either 25 W or 30 W for females or males respectively. The participant wore a COSMED silicone face mask attached to the COSMED metabolic system and VO_2 was measured for two minutes during stationary rest and for the first three stages of the GXT. LT was calculated using the modified DMax method [261] using the Lactate-E macro in excel [262]. W_{peak} was calculated using the following formula:

$$W_{peak} = W_{LC} + ((W_S * (Secs_{LS}/240)). \tag{3.1}$$

W_{LC} were the Watts at the last fully completed stage.
W_S were the Watts increment of each stage (i.e 25W or 30W)}
$Secs_{LS}$ were the seconds completed of the last attempted stage.

VO_{2peak} test

After 5 minutes rest following the GXT, peak oxygen consumption (VO_{2peak}) was measured using a calibrated Quark CPET metabolic system with mixing chamber (COSMED, Rome, Italy). The face mask was returned to the participants and VO_2 was collected for two minutes during stationary rest, then for three minutes whilst cycling at the starting intensity of the GXT, and during a subsequent exercise to exhaustion stage at 105% of calculated W_{peak} obtained on the GXT. VO_{2peak} values were smoothed over a 15 second period using the COSMED software and VO_{2peak} was considered the highest value maintained over one minute during the test (every 15 second value was averaged over one minute).

High Intensity Interval Training (HIIT) phase

Our group have previously reported that 8-12 weeks of HIIT is optimal to elicit performance and biological adaptations [263, 264] therefore we chose to implement a 12 week HIIT program. All participants completed 12 weeks HIIT that consisted of three training sessions per week (36 sessions in total) on an electronically braked cycle ergometer (Velotron, Racer Mate Inc., Seattle, USA). Each session consisted of a short warm up fol-

lowed by six to 14 two minute intervals performed at individually determined workloads with one minute rest intervals. The participants cycled at individually-determined intensities calculated as LT power plus 40% (week 1) to 70% (week 4) of the difference (Δ) between their individually-determined W_{peak} (derived from the baseline GXT test results) and the power at the LT (LT + 40–70% Δ). At the completion of four and eight weeks of training two GXTs and $V0_{2peak}$ tests were completed to re-adjust individually determined intensities for the HIIT sessions. Fourteen out of fifteen participants completed the intervention and one female participant dropped out during week three of training.

Post-testing

After 12 weeks of HIIT, participants reported to the laboratory to complete two post GXT/$V0_{2peak}$ tests and the average of the two tests was calculated to determine LT, W_{peak} and $V0_{2peak}$. Values were calculated as the average of the two tests or as the highest value of one test if the difference between values were greater than 5%. Three out of the fifteen participants were unable to complete the post testing for various reasons but were able to provide a muscle sample which were included in biochemical analyses.

Controlled diet

Each participant was provided with individualised, pre-packaged meals for the 48 hours prior to their muscle biopsies. The nutrient needs of participants were calculated using the Mifflin St-Jeor equation [265] and each participant's weight, height, age and activity level. FoodWorks 10 software (Xyris, Brisbane, Australia) was then used to determine the nutritional composition of the prepared meals and to ensure participant's nutrient needs were met over the provided 48 hour diet. The content of the diets were constructed based on the current National Health and Medical Research Council (NHMRC) guidelines (50-55% carbohydrates, 15-20% protein, <30% total fat) [266].

Muscle biopsies

Muscle biopsies were taken from the participants dominant leg prior to training, at four weeks, and at completion of the training. Biopsies were collected from the *m. vastus lateralis* under local anaesthesia (5 mL, 1% Xylocaine) by an experienced Medical Doctor (Dr. Andrew Garnham) using a Bergstrom needle [267]. For female participants muscle

biopsies were taken during the first seven days of their menstrual cycle and testing and training was coordinated accordingly. For consistency all muscle biopsies were taken in the morning between 6-8am. The four week muscle biopsies were collected as part of the ongoing Gene SMART study. For fibre type experiments we used only baseline and 12 week samples as 12 weeks of training offers more reliable measures of the response to exercise than four weeks based on our groups previous work [264].

3.2 Muscle processing

3.2.1 Muscle separation and preservation

After muscle biopsies were taken, muscle was placed in a glass petri-dish on ice, the perti-dish was lined with filter paper to remove excess blood. Samples (100-200 mg) were separated into four sections: ~50-70 mg of muscle was immediately snap frozen in liquid nitrogen (stored at -80 °C), ~25-40 mg of sample was placed in ice cold PBS 1X for single fibre separations, ~10-20 mg was embedded in Tissue-Tek O.C.T Compound and frozen on thawing 2-methylbutane, and ~>60 mg was placed in Phosphate-Buffered Saline (PBS) on ice for mono-nucleated cell fraction preparation.

3.2.2 Preparation and preservation of skeletal muscle fibres

Immediately after the biopsy, muscle stored in ice cold PBS was separated into smaller bundles under a light microscope and processed in the following manner. Two thirds of the sample was moved to ice-cold RNAlater in a 1.5 mL eppendorf tube and stored at -20 °C. The remaining one third was fixed for imaging experiments as previously described (with minor alterations) [196]. Firstly, muscle fibre bundles were transferred to 4% Paraformaldhyde (PFA) in 1.5 mL eppendorfs for 30 minutes at Room temperature (RT). Muscle bundles were then transferred to fresh 4% PFA and allowed to sit for two hours at 4 °C, after fixation, muscle bundles were washed three times by slowly dipping them in PBS. Fixed bundles were then moved to a 1:1 PBS/Glycerol solution and left at 4 °C overnight. The following day fixed fibre samples were moved to -20 °C for long term storage.

3.3 Biochemistry techniques

3.3.1 Single fibre isolation, muscle digestion and lysis

Skeletal muscle fibre bundles stored in RNAlater were removed from -20 °C storage and immediately put on ice. Using a light microscope smaller bundles of muscle fibres were isolated with Dumont biological grade tweezers (Dumont, Jura, Switzerland) in a plastic 60 mm petri dish on a ceramic tray on ice. Small bundles were transferred to ice cold RNAlater and the remaining muscle was kept in RNAlater at -20 °C. Single fibres were isolated using tweezers on ice under the microscope and transferred to 8-strip PCR strip tubes and kept on ice. Single fibre fragments were approximately 4-8mm in length depending on the sample. After isolation of approximately 24 fibres, 10 mL of digestion buffer (100 mM TRIS-HCl pH=7.6, 50 mM DTT and 4% SDS) was added to each tube. PCR tubes were spun down for one minute on a table top micro-centrifuge and then vortexed for one minute on a MSA vortexer. Tubes were briefly spun again and transferred to -80 °C for storage. Before fibre-typing, muscle fibres were vortexed for 15 minutes on an MSA vortexer to aid in complete digestion of the fibres. Samples were stored at -80 °C until fibre typing had been completed as outlined below.

3.3.2 Dot-blot fibre typing

Sample preparation

A dot blot procedure of muscle fibres as previously described [268] was modified to allow high-throughput and accurate dot blotting. A nitrocellulose membrane was pre-wet in 1X Tris-Buffered Saline (TBS) for 5-10 minutes. 40-50 µL of TBS 1X was added to each well of a 96 well plate. Muscle fibre segments were thawed on a 96-tube vortexer (MSA) at 2500RPM and centrifuged. 2 µL of each single fibre lysate was added to each well of the 96-well plate and agitated on an orbital plate shaker for 10 minutes.

Dot-blot microfiltration

A 96 well Bio-Dot Microfiltration Apparatus (BioRad) was assembled with the nitrocellulose membrane whilst under vacuum. The vacuum was switched off and 50 µL of TBS was loaded into each well and filtered through by vacuum suction. 40-50 µL of pro-

tein/TBS from each fibre segment was added to the well of the dotblot apparatus and allowed to filter through for 45-60 minutes by gravity. After all the lysate had filtered through the apparatus an additional 50 µL of TBS was added to wash the remainder of the lysate onto the membrane via vacuum suction. The dot blot apparatus was disassembled and the nitrocellulose membrane was briefly washed in TBS 1X and then allowed to air dry. The membrane was washed in Tris-Buffered Saline with Tween (TBS-T) and blocked in TBS-T with 5% milk for 15 minutes at RT.

Antibody incubation and imaging

After blocking, the membranes were rinsed with TBS-T 3x5 minutes and incubated overnight with antibodies for MYH7 (A4:840 DHSB - made fresh every second day) at 1:75 in TBS-T with 1% bovine serum albumin (BSA). The next day the membrane was washed with TBST for 3x10 minutes and incubated with secondary antibody for anti-mouse Immunoglobulin M (IgM) 1:10,000 (Invitrogen, California, United States) in TBS-T with 5% milk. The membrane was again washed 3x10 minutes and exposed to a chemiiluminescent substrate for 45 seconds (BioRad, California, United States). The membranes were imaged with a BioRad ChemiDoc MP system (Biorad, California, United States) at both high and low resolution settings for between 30-120 seconds.

Stripping and re-probing

After imaging, membranes were washed briefly in TBS-T and stripped for 5-10 minutes in PIERCE Restore PLUS stripping buffer (Thermofisher(Thermofisher, Massachusetts, United States). The membranes were briefly washed in TBS and re-incubated in secondary antibody for more than 15 minutes and re-imaged. If signal remained one more round of stripping was performed. After confirmation of successful stripping the membrane was incubated overnight with antibodies for MYH2 (A4:74 Development Studies Hybridoma Bank DSHB, Iowa, United States - made fresh every few days) at 1:200 in TBS-T with 1% BSA. The next day the membrane was washed with TBS-T for 3x10 minutes and incubated with secondary antibody for anti-mouse Immunoglobulin G (IgG) 1:10,000 (Invitrogen, California, United States) in TBS-T with 5% milk and the same imaging process was completed as the day before. Following imaging for MYH2 membranes were washed in TBS and incubated in ponceau red for 10 minutes. Membranes were briefly

washed in water to remove excess ponceau stain and imaged as a ponceau protein stain using a ChemiDoc.

3.3.3 Fibre type sample pooling

After fibre typing, skeletal muscle single fibres fragments were pooled for downstream use in protein and DNA analyses. PCR strip tubes were allowed to thaw by vortexing for 5 minutes at 2,500 revolutions per minute (rpm) on an MSA vortexer. Type I and Type II samples were pooled by transferring 2 µL to a 1.5 mL eppendorf tube for protein analyses. 2 µL of ponceau positive but MYH2 and MYH7 negative fibre lysates were also pooled. 100 µL of each sample was stored at -80 °C for proteomic sample preparation. After removal of sample for protein analyses, 6 µL of stabilisation buffer (10mM TRIS-HCl pH = 8.0, 26mM EDTA) was added to each tube and briefly vortexed and spun down. DNA was extracted as outlined below.

3.3.4 Pooled fibre type DNA extractions

Day 1

75-150 single fibre samples were pooled together for each fibre (Type I and Type II) in 1.5 mL eppendorf tubes in batches of 50 fibres per eppendorf tube. 25 µg of protinase K was added to each tube and the tube was kept at 56 °C for 4 hours at 1000 rpm. The following protocol was adapted to extract DNA [269]. 200 µL of 3M sodium acetate (non adjusted pH) was added to each eppendorf tube and vortexed for one minute. 200 µL of cool Phenol Choloroform Isoamyl Alcohol (PCI) was added to each tube and then vortexed for 20 seconds and centrifuged for 10 minutes at 18000 *g* at 4°C. The upper aqueous phase was carefully removed and transferred to a new 1.5 mL eppendorf without disturbing the organic phase. 500 µL of cool isoproponol was added to the new eppendorf and mixed by inversion. genomic DNA (gDNA) was precipitated overnight in isoproponol at 4°C.

Day 2

Eppendorf tubes were centrifuged at 20000 *g* for 30 minutes at 4°C. After centrifugation a small pellet was visible and the supernatant was carefully removed without disturbing the gDNA pellet. The pellets were washed with 70% ethanol and centrifuged at 20000

g for 20 minutes at 4°C. The supernatant was again removed and the pellets allowed to air dry at 37 °C on a heating block. gDNA pellets were dissolved in TE buffer and re-suspended by gently pipetting up and down. gDNA was further dissolved by placement on a heating block at 37 °C for 30-60 minutes. gDNA concentration was assessed using a Qubit High Sensitivity 1x DNA kit on a Qubit Fluorometer 4 (Thermofisher, Massachusetts, United States) following kit instructions. DNA A260/280 values were measured on a Nanodrop One (Thermofisher, Massachusetts, United States) unless they contained less than 250 ng DNA concentration on the Qubit reading.

3.3.5 Whole muscle DNA extractions

gDNA and RNA was extracted from pre and twelve-week whole muscle samples using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen, Venlo, Netherlands). 10-15 mg of muscle was homogenised two times at 30 Hz for 30 seconds with a Qiagen TissueLyser II (Qiagen, Netherlands) in 600 µL RLT Plus Buffer (with beta-mercaptoethanol added as per manufacturers instructions). gDNA and RNA was extracted following the AllPrep Universal Kit instructions. gDNA was re-suspended in 50 µL of AE buffer provided with the kit. The extracted gDNA was passed twice over the column to increase gDNA yield.

3.3.6 DNA Agarose Gel Visualisations

gDNA was visualised by running gDNA on a 1.5% agarose gel which were prepared with 1x SYBR Safe DNA Gel Stain (Invitrogen, California, United States). 50-75 ng of gDNA was prepared in a volume of 2 µL and mixed with five parts 6X DNA loading gel (Invitrogen, California, United States). 12 µL of final volume was loaded in to each well and ran at 80V for 45 minutes. 2-3 µL of 1 Kb Plus DNA Ladder (Invitrogen, California, United States) was loaded in each gel. Agarose gel images were taken on a BioRad Chemidoc Imaging System (Biorad, California, United States).

3.3.7 DNA array sample distribution

Both balanced and randomised approaches to sample distribution on the chip can be problematic [270, 271]. Therefore, I distributed samples using a stratified randomised approach to minimise batch effects. Male/female, rest/exercise and typeI/typeII/whole-muscle samples were distributed across arrays and across array position. All samples

were bisulfite converted in a single 96 well plate. Three duplicate samples were distributed across arrays and two duplicate samples were distributed across positions in one array.

3.3.8 Bisulfite Conversion and DNA Methylation Analyses

Samples were sent to the Human Genomics Facility of the Genetic Laboratory of the Department of Internal Medicine (HuGe-F) at Erasmus MC, Erasmus, Netherlands for bisulfite conversion and DNA methylation data generation. Between 200-250 ng gDNA was bisulfite converted using the EZ-96DNA Methylation™MagPrep (ZYMO Research, California, United States) according to the manufacturer's instructions. The Illumina Infinium MethylationEPIC v2.0 Kit (Illumina, California, United States) was used for DNA methylation analysis. After elution of converted DNA the samples were hybridised to the BeadChip and imaged with a NextSeq 550 system (Illumina, California, United States).

3.3.9 Liquid Chromatography Tandem Mass Spectrometry (LC–MS/MS)

Protein Lysates (100 µL) from the pooled muscle fibre samples were prepared in a 4% SDS, 100mM Tris-HCl (pH 8.0), 50mM DTT buffer. Liquid chromatography with tandem mass spectrometry (LC-MS/MS) was performed by the Monash Proteomics and Metabolomics Platform at Monash University, Melbourne, Australia. Samples were sonicated and proteins trapped on S-trap Acclaim PepMap 100 (100 µm x 2 cm, nanoViper, C18, 5 µm, 100Å; Thermo Scientific) columns. Proteins were cleaned and then digested to peptides overnight at 36 °C with trypsin. Peptide concentrations were measured and 0.9ug of peptide was injected for LC-MS/MS. LC-MS/MS data was acquired using a Dionex Ultimate 3000 RSLCnano (Thermo Scientific) for peptide separation and measurements acquired with a Orbitrap Q-Exactive HF (Thermo Scientific) mass spectrometer.

3.4 Data Processing, Analysis and Statistics

3.4.1 Power Calculations

Epigenome wide association studies (EWAS) conducted in bulk tissue and investigating non-disease factors, such as, age, diet and exercise, typically report small, widespread effect sizes of 5-10% DNAm differences between individuals and conditions [175]. However, differences between cell types are much larger than differences between individuals for a given cell type [67]. In skeletal muscle for example 50% of all differentially methylated sites between Type I and Type IIa muscle fibres showed at least 10% methylation differences [3]. While there is no study on the differential DNAm response of the different muscle fibres after exercise to guide our power calculations, effect sizes in whole blood have been reported (from ~5% DNAm difference in whole blood to >20% DNAm difference in one specific blood cell type) [272]. We therefore expect medium to large effect sizes (>10%) both at baseline and in response to exercise training between different fibre types. Using the power calculations from Mansell and others [273], 85% of differentially methylated positions (DMPs) showing a mean difference of 10% would be detected at a power of 80%, with a sample size of n=60, that is n=10 males, n=10 females (for 3 fibre types). Due to COVID19 restrictions and a subsequent pause on human research, we were only able to recruit eight males and seven females to this study.

3.4.2 DNAm Data Preprocessing and Quality Control

All processing and analysis were performed using the R statistical computing platform [274]. Raw IDAT files were processed using the minfi package [275]. Firstly, probes with a detection P-value > 0.01, and/or bead count < 3 were removed. All non-cpg probes, probes located in the sex chromosomes and SNP-related probes were removed. Probes flagged by Illumina as inaccurate or underperforming and cross reactive EPICv1 [276] and 450K [277] probes were removed. Normalisation of type I and type II probes were performed using the Dasen normalisation method [278] as implemented in the wateRmelon R package. Principal component regression (PCR) was utilised to identify biological and technical sources of variation in DNAm and due to limited batch effects, batch correction was not implemented.

3.4.3 Differentially Methylated Probes, Regions and Gene Set Enrichment Analysis

Identifying DMPs

DMPs between fibres at rest and after 12 weeks of exercise were identified using the *limma* [279, 280] package in R by calculating empirical bayes moderated t-statistics. The model was grouped according to cell-type/tissue (TI, TII, WM) and timepoint (pre and post). Sex was included as a co-variate and duplicate correlation was added to control for the clustering of samples by participant due to the paired experiment design. As the sample size was small there was a low representation of different ethnicities in the dataset and it was not included as a co-variate in the model. The following code was used:

```
design <- model.matrix(~ 0 + group + sex, pheno_filt_subset)

cor <- duplicateCorrelation(m_val_filt_subset, design, block =
↪  pheno_filt_subset$participant)

fit <- lmFit(object = m_val_filt_subset,
             design = design,
             block = pheno_filt_subset$participant,
             correlation = cor$consensus.correlation)
```

The following contrast matrix was generated to answer the following questions:

1. What methylation differences are there between TI-pre and TII-pre fibres?
2. What methylation differences are there between WM-pre and TI-pre fibres?
3. What methylation differences are there between WM-pre and TII-pre fibres?
4. What methylation changes occur within a TI fibre after HIIT?
5. What methylation changes occur within a TII fibre after HIIT?
6. What methylation changes occur within WM after HIIT?

The contrast matrix and contrast fit code is found below.

```
contrasts <- makeContrasts(T1preVsT2pre = groupT1_Pre - groupT2_Pre,
                           T1postVsT2post = groupT1_Post - groupT2_Post,
                           T1preVsWMpre = groupT1_Pre - groupWM_Pre,
                           T2preVsWMpre = groupT2_Pre - groupWM_Pre,
                           T1VsT2 =  (groupT1_Pre + groupT1_Post) / 2 -
                               (groupT2_Pre + groupT2_Post) / 2,
                           T1preVsT1post = groupT1_Pre - groupT1_Post,
                           T2preVsT2post = groupT2_Pre - groupT2_Post,
                           WMpreVsWMpost = groupWM_Pre - groupWM_Post,
                           Interaction = (groupT1_Post - groupT1_Pre) - (groupT2_Post -
                           ↪  groupT2_Pre),
                           levels = colnames(design))

fit_2 <- contrasts.fit(fit, contrasts)

fit_3 <- eBayes(fit_2)
```

Identifying DMRs

The *DMRcate* package in R [281, 282] was applied to determine differentially methylated regions (DMRs), that is neighbouring CpGs with consistent DNAm differences. DMRs were defined as regions with ≥ 4 CpGs. Any DMPs/DMRs identified were annotated, taking into consideration the skeletal muscle chromatin state obtained from the Roadmap Epigenomics Project.

Gene Set Enrichment Analysis of DMPs and DMRs

The application of multiple gene set enrichment analysis (GSEA) approaches increases robustness of results [283]. ClusterProfiler [284, 285] and missMethyl [286] packages in R were both used for overrepresentation analysis (ORA). The *missMethyl* package [287] was modified for compatibility with the Infinium MethylationEPIC Beadchip (EPIC)v2 to conduct bias adjusted over representation of GO, Reactome and KEGG genesets. To correct for multiple testing in the identification of DMPs, DMRs and over-represented gene sets, the Benjamini and Hochberg false discovery rate (FDR) method was used [288] and multiple thresholds reported. All genesets with < 5 or > 500 genes were removed prior to ORA.

3.4.4 Proteomic Data Preprocessing and Quality Control

Raw LC-MS/MS data files were analysed using Fragpipe (V 19.1) and the MSfragger (V 3.5) search engine [289] to obtain protein identifications. The picky library was used for protein search and identifications [290]. Quality control and statistical analyses were performed using the R software. The analysis required the MSFragger combined protein file containing proteins that passed a FDR cutoff of 0.01 and a experiment design text file with sample details. First contaminant proteins were removed for all non human species. All proteins identified in at least four condition replicates were kept so that proteins with inconsistent identification were removed. The MaxLFQ values were log2 transformed, normalised using a median centring approach, grouped by condition and missing values were imputed assuming they were missing not at random (MNAR), where random values were selected from a left-shifted Gaussian distribution of 1.8 standard deviation with a width of 0.3.

3.4.5 Proteomic Data Analysis

Identifying DEPs

The limma [279] package was used to generate protein-wise linear models using empirical bayes statistics, that identified deferentially expressed proteins for each pair-wise comparison. We reported multiple cutoffs of the adjusted-p-value (Benjamani-Hochberg method [288]) along with a log2 fold change of 1 to determine regulated proteins between each fibre type. The model was grouped according to fibre-type (TI and TII) and timepoint (pre and post). Sex was included as a co-variate and duplicate correlation was added to control for the clustering of samples by participant due to the paired experiment design. We removed sample s197-pre-T2 from the analysis due to the reduced number of proteins identified.

The following code was the model design used with Limma:

```
design <- model.matrix(~ 0 + group + sex, experiment_design)

colnames(design) <- c("TI_post", "TI_pre", "TII_post", "TII_pre", "sexm")
```

```
cor <- duplicateCorrelation(data_normalised_imputed, design, block =
↪  experiment_design$sample_id)


fit <- lmFit(object = data_normalised_imputed,
             design = design,
             block = experiment_design$sample_id,
             correlation = cor$consensus.correlation)
```

```
contrasts <- makeContrasts(TIpreVsTIIpre = TI_pre - TII_pre,
                           TIpostVsTIIpost = TI_post - TII_post,
                           TIpostVsTIpre = TI_post - TI_pre,
                           TIIpostVsTIIpre = TII_post - TII_pre,
                           interaction = (TII_post - TII_pre) -  (TI_post - TI_pre),
                           ↪  levels = colnames(design))

fit2 <- contrasts.fit(fit, contrasts)
fit2 <- eBayes(fit2)
```

Gene Set Enrichment Analysis of DEPs

ClusterProfiler [284, 285] was used to conduct overrepresentation analysis of differentially expressed proteins (DEPs) in GO, KEGG and Reactome gene sets. All genesets with < 5 or > 500 genes were removed prior to ORA and functional class scoring (FCS). Multiple FDR cutoff levels for each comparison are reported. ORA results were compared to FCS methods using the fGSEA [291] package in R. For FCS genes were ranked using the following formula.

$$-log10(P.Value) * sign(logFC) \tag{3.2}$$

4 Results 1: A method for the simultaneous extraction of DNA and proteins from single and pooled skeletal muscle fragments

4.1 Introduction

To date, researchers interested in analysing protein, RNA and DNA in human skeletal muscle fibres have been studying either single muscle fibres or pooled muscle fibres based on myosin isoform content. Single muscle fibre experiments are becoming more common due to improvements in technology. However, experiments with single fibres are expensive and single muscle fibre samples contain little biological material leading to experimental challenges and poor yield. To overcome these limitations, researchers have developed methods to separate and type muscle fibre fragments [292], before pooling muscle fibre types to increase biological material [3]. The isolated fibre-type-specific DNA and protein can be used to study protein or DNAm differences between skeletal muscle fibre types or protein and DNAm changes as a result of environmental stimuli such as exercise.

The common approach to determine fibre types is to digest fibres, type them using SDS gel electrophoresis (SDS-PAGE) followed by myosin isoform content identification using antibodies targeted against MYH7 and MYH2. In order to extract fibre-type-specific DNA, fibres are first isolated and then protein extracted for fibre-type identification before pooling them for DNA extraction. This approach was utilised by Begue and others [3] to study DNAm in human skeletal muscle fibre types, where they cut single muscle fibres and used one fragment to type fibres and the remainder to extract DNA. This cutting of fibres adds additional time to the method and causes a higher chance of error when handling the biological sample. I hypothesised that the process of digesting,

typing and preparing single fibre fragments for DNA extraction could be improved and achieved in one tube without the need for cutting fibre fragments.

Begue and others [3] conducted DNAm analysis on one pooled Type I fibre sample and one pooled Type II fibre sample (pooled from n=8 male samples) using Reduced representation bisulfite sequencing (RRBS). The aim of this thesis was to study DNAm in pooled muscle fibre types using the most common assay to measure DNAm, the Infinium Methylation Arrays [293]. This method requires a minimum of 250 ng of DNA and is the most frequently used DNAm assay in skeletal muscle tissue. To achieve the required quantity of DNA, I optimised the fibre typing process to enable the processing of more fibres per sample.

Since the publication by Begue and others [3], Christiansen and others have developed a quick sample sparing method to type single fibre fragments using a dot blot method [268]. In the dot blot method, a small sample of protein lysate is blotted onto a membrane and subsequently probed with antibodies targeting MYH7, MYH2 and MYH1 [268]. This method removes the need for gel electrophoresis and can be completed in one day. As fibre typing is based on the absence or presence of myosin isoforms this method could enable a higher throughput of fibres leading to enough DNA material for DNAm measurements using the EPIC array.

The overarching aim of this chapter was to develop an optimised simple method for the isolation of DNA and proteins from muscle fibres by eliminating the need for cutting fibres. The development of a common lysis approach for extracting proteins and DNA from the same sample in one tube resulted in the following benefits: precious sample was spared; the extraction approach was simplified; the number of fibres processed was increased; and the costs of protein and DNA extraction was reduced. By developing this method, I was able to obtain enough sample for dot blot fibre typing, and by pooling samples I was able to extract enough DNA and protein from pooled muscle fibre fragments for the simultaneous measurement of DNAm using the EPIC array and protein expression using proteomics.

The aims of this chapter were:

1. To optimise digestion of single muscle fibres in one tube for DNA and protein extraction.

2. To optimise a dot blot protocol for accurate high capacity fibre typing.

3. To simultaneously isolate DNA and proteins from pooled skeletal muscle fibre fragments without the need to cut fibres.

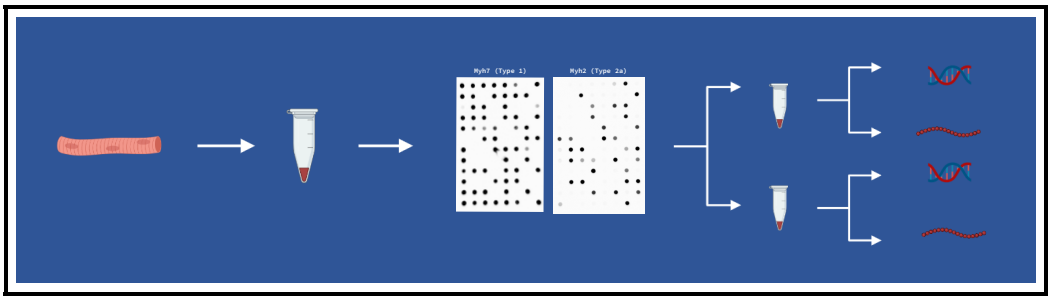


Figure 4.1: Method Schematic

A schematic of the proposed method for simultaneous preparation of protein and DNA from skeletal muscle fibre fragments.

4.2 Methods & Results

4.2.1 Sample Lysis and Digestion

Rationale

The choice of lysis buffer should ideally enable the extraction of both proteins and DNA without the need for cutting fibres. This buffer ideally enables fibre type identification, aids in protein digestion and stabilises DNA protecting it from degradation. I tested three previously reported lysis buffers used for fibre digestion, typing and subsequent analysis.

Method

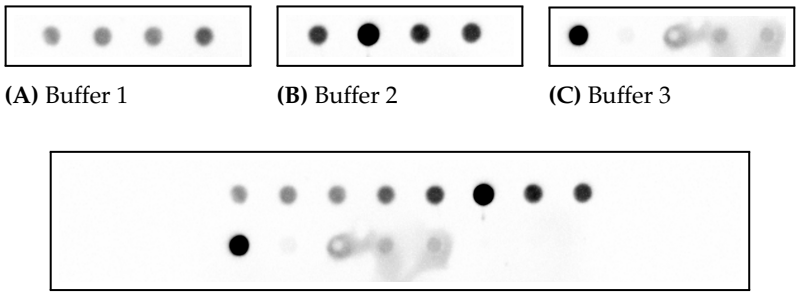
The following three buffers were tested and were slightly modified; buffer one [1% SDS, 6 / EDTA, 0.06 TRIS (pH 6.8), 15% glycerol] [3], buffer two [4% SDS, 0.100 Tris-HCl,100mM DTT][210] and buffer three [4% SDS, 0.125 Tris-HCl, 10% glycerol, 4 urea, 10%] [268] (without 2-mercaptoethanol and bromophenol blue). Small skeletal muscle fibre bundles

(approximately 5 fibres) were digested in 50 µL of buffer and agitated on an orbital plate shaker for four hours. Digested fibres were frozen at -80°C, thawed and passed through a BioDot dot blot apparatus (BioRad, California, United States) onto a nitrocellulose membrane according to manufacturers instructions. Membranes were incubated in primary antibody for MYH2, a mouse IgG secondary antibody and imaged using a chemiluminescence substrate in a Biorad MP imaging system.

Results

A minimal SDS/Tris buffer is adequate for fibre digestion and dot blotting

The resulting dot blots in Figure 4.2 show that all three buffers were able to digest and release proteins for dot blot analysis of MYH2. Buffer two, the buffer with fewest chemicals, appeared to perform better than buffers one and three, although this could reflect a higher amount of MYH2 in these fibres. Buffer two has recently been used in proteomic experiments on pooled skeletal muscle fibre types [210] where Sodium Dodecyl Sulfate (SDS) and Dithiothreitol (DTT) were effectively removed for compatibility with mass spectrometry (MS) proteomics. Common DNA extraction protocols rely on a SDS Trishydroxymethylaminomethane (TRIS) buffer for cell lysis, while the addition of DTT denatures proteins and stabilises DNA. As buffer two performed well on testing, contains the least chemicals and has key reagents for DNA extraction, it was chosen as the lysis buffer of choice for the remaining experiments.



(A) Buffer 1 **(B)** Buffer 2 **(C)** Buffer 3

(D) Un-cropped membrane of the three buffer test

Figure 4.2: Buffer Test

Small muscle fibre bundles were digested with (A) buffer 1 [1% SDS,6mg/ml EDTA, 0.06M Tris (pH 6.8), 15% glycerol], (B) buffer 2 [4% SDS, 0.100M Tris-HCL,100mM DTT], (C) buffer 3 [4% SDS, 0.125M Tris-HCL, 10% glycerol, 4M urea, 10%] and spotted onto a nitrocellulose membrane using a dot blot apparatus. The membrane was probed with a MYH2 antibody and quantified using a ChemiDoc MP. The uncropped membrane can be seen in (D).

4.2.2 Dot Blot Fibre Typing

Rationale

The success of cell-type-specific experiments relies on accurate methods to determine and isolate cell types of interest. Christiansen and others [268] developed a simple method for typing single skeletal muscle fibres by spotting 1 µL of digested fibre fragment directly on a Polyvinylidene Fluoride (PVDF) membrane. This method is indeed rapid, however a dot blot apparatus (or manifold) is routinely used in studies testing the absence or presence of a protein in a sample. The advantage of a vacuum manifold is that it can help concentrate samples and make protein loading consistent and uniform. The application of a vacuum maximises sample binding and reduces sample contamination between wells. As previous studies have not described using a dot blot apparatus, I tested the application of the device for determination of the absence or presence of MYH2 and MYH7. Furthermore, I compared PVDF verses nitrocellulose membranes, tested different protein loading concentrations and tested different antibodies reported for typing Type IIx muscle fibres in order to determine the most robust fibre typing protocol.

Method

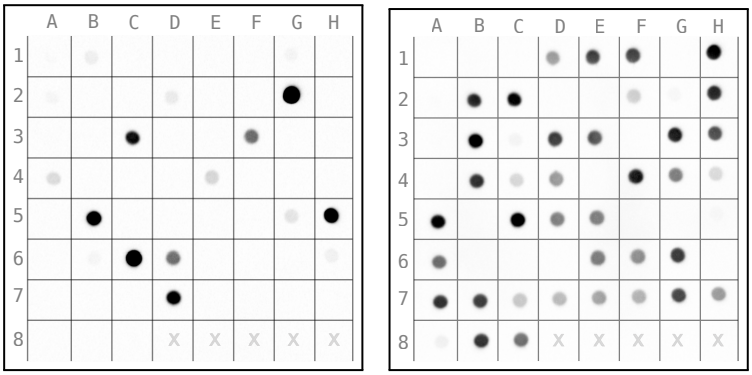
I isolated single muscle fibres under a light microscope in a petri dish on a ceramic plate resting on ice. Fibres were placed in a single PCR tube and 10 µL of lysis buffer was added. Fibres were digested by pulse vortexing for up to 5 minutes. Digested fibres were frozen and thawed for one cycle to assist fibre digestion. For dot blot apparatus experiments 2 µL of sample was added to 50 µL of lysis buffer, mixed and loaded into each well of a dot blot apparatus. Sample was allowed to filter through by gravity onto two nitrocellulose membranes. Membranes were allowed to dry, blocked for 5 mins in TBS-T with 5% milk and then incubated with antibodies for MYH2 or MYH7. Images were acquired using a chemiluminscent substrate and a ChemicDoc MP system.

Results

Using a dot blot apparatus results in consistent and reproducible fibre typing results

Using a dot blot apparatus produces defined and consistent protein dots. As seen in Figure 4.3, the absence and presence of MYH7 and MYH2 can be identified using a

vacuum manifold. The manual dot blot method resulted in variable results, with less defined and clear protein dots (Figure A.1). The use of a dot blot apparatus is not as efficient as directly spotting lysate onto a membrane as it requires more time for sample filtration. This might be a drawback for some studies, however in studies where accuracy is a high priority, a dot blot apparatus is a good option. The dot blot apparatus was used for the remaining experiments as accuracy of typing was the highest priority.



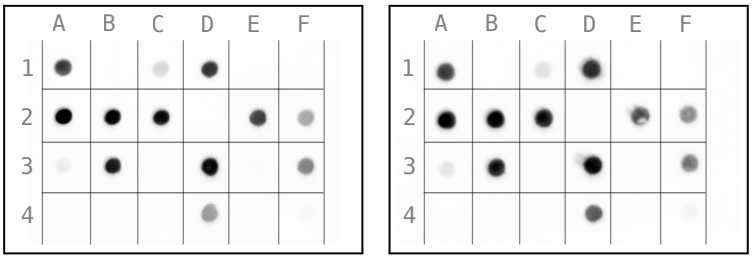
(A) MYH7: Dot Blot Apparatus **(B)** MYH2: Dot Blot Apparatus

Figure 4.3: Dot Blot Apparatus Test

One fifth of the same single muscle fibre fragments were spotted onto nitrocellulose membranes using a dotblot apparatus. (A) shows uniform and consistent identification of MYH7. (B) shows uniform and consistent identification of MYH2. Positions labelled with an X indicate that no fibre was loaded in this well.

Both nitrocellulose and PVDF membranes perform well in a dot blot apparatus

Manufacturers often report that PVDF membranes are able to capture a higher concentration of proteins whereas nitrocellulose membranes have higher resolution. I compared dot blotting onto a PVDF membrane to dot blotting onto a nitrocellulose membrane to determine which is best for myosin identification. One fifth of single muscle fibre fragments were spotted onto a nitrocellulose or PVDF membrane using a dot blot apparatus. PVDF membranes were preactivated in methanol prior to placement of the membrane in the apparatus. There was little difference between the PVDF membrane and nitrocellulose membrane and both are able to determine absence or presence of MYH2 (Figure 4.4). I opted to use nitrocellulose membranes for further experiments as it is reported that they have less background [294] and are recommended as the membrane of choice when using the BioDot apparatus from Biorad.



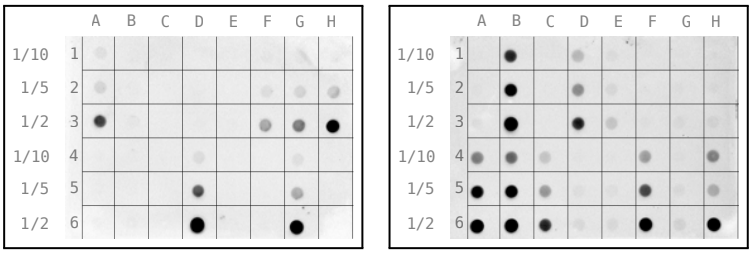
(A) Nitrocellulose **(B)** PVDF

Figure 4.4: PVDF vs Nitrocellulose Comparison

One fifth of the same single muscle fibre fragments were spotted on to either a nitrocellulose or PVDF membrane using a dot blot apparatus. PVDF membranes were preactivated in methanol prior to placement of the membrane in the apparatus. Membranes were incubated with antibodies against MYH2. (A) & (B) both show the same presence or absence of MYH2.

One fifth of a fibre can be used for reliable dot blot fibre typing of skeletal muscle fragments for MYH2 and MYH7

Another important consideration in dot blotting is the amount of protein required for accurate fibre typing. One major challenge when isolating muscle fibres is avoiding contamination from other fibres. Total avoidance of contamination is not feasible, therefore the dot blot method should be sensitive enough to ensure false typing of fibres is reduced. I tested different protein loading concentrations based on proportion of fibre length to determine which length of fibre is optimal for dot blot fibre typing. Single fibre fragments where digested in 50 µL of lysis buffer and either one tenth, one fifth or one half of each sample was dot blotted for both MYH2 and MYH7. As little as one tenth of a fibre was sufficient to determine presence of MYH2, however for MYH7 one tenth of a fibre seems too little for reliable typing (Figure 4.5). The image could be enhanced by contrasting but this may lead to false typing due to background signal. There is variation in muscle fragment length when isolating fibres and low signal may reflect starting protein quantity. It appears (Figure 4.5B) that MYH2 may have unspecific binding when higher protein is loaded. Loading one fifth of a fibre strikes a balance between having enough protein for MYH2 and MYH7 identification and avoiding background signal.



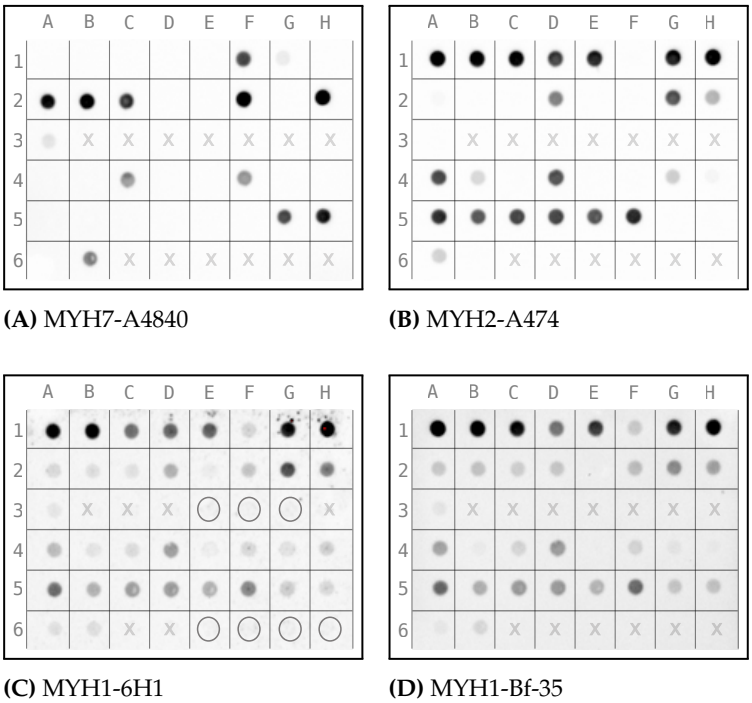
(A) MYH7 **(B)** MYH2

Figure 4.5: Dilution Test

One tenth, one fifth and one half of the same single fibre was spotted onto a nitrocellulose membrane using a dot blot apparatus and then incubated with (A) MYH7 then stripped and reprobed for (B) MYH2.

Accurately typing Type IIx fibres using reported antibodies was difficult

It was previously reported that the blot dot technique could successfully distinguish type I, Type IIa and Type IIx fibres using three antibodies from DSHB (A4-840 - MYH7, A4-74 - MYH2, and 6H1-s - MYH1) [268]. I tested two antibodies reported to distinguish Type IIx muscle fibres (DSHB-6H1 and DSHB-Bf-35). Single muscle fibre fragments were digested and one fifth of a fibre was spotted on to four different membranes and probed with MYH7, MYH2 and the two antibodies reported for Type IIx fibres. Seventeen fibres were loaded in rows one to three and eighteen were loaded in rows four to six. The 6H1 antibody bound to all fibres including dot five in row four which was missing from MYH2 and MYH7, however signal was measured in empty wells (Figure 4.6C). Antibody Bf-35 was detected in all fibres where both MYH2 and MYH7 were present (Figure 4.6D). Two possibilities for these results are: that I did not isolate Type IIx fibres or I was unable to obtain consistent Type IIx typing of fibres. In some cases no dot was visible on typing for MYH2 and MYH7 but was visible after ponceau staining. Typing Type IIx fibres in my hand was unreliable with reported antibodies, so I decided in future experiments to stain each membrane with Ponceau to confirm if dots visible but negative for both MYH2 and MYH7 represent Type IIx fibres.



(A) MYH7-A4840 **(B)** MYH2-A474

(C) MYH1-6H1 **(D)** MYH1-Bf-35

Figure 4.6: Antibody-testing

One fifth of the same single fibre fragment was spotted on to four separate nitrocellulose membranes using a dot blot apparatus. Membranes (A) & (B) show the absence and presence of MYH7 and MYH2 as tested with antibodies A4-840 and A4-74 (respectively) obtained from DSHB. Membranes (C) & (D) show the absence and presence of MYH1 as tested with antibodies 6H1 and BF-35 obtained from DSHB and reported to detect Type IIx fibres. Positions labelled with an X indicate that no fibre was loaded in this well.

4.2.3 Three method comparison

Rationale

The dot blot method has potential limitations: it requires antibodies for determination of fibre type; proteins are not separated by size and therefore antibodies might bind non-specifically to other proteins; and determining myosin isoform content of fibres is not a quantifiable approach. Silver-staining has been a standard method for analysing myosin content of single skeletal muscle fibres [295]. In this method proteins are separated by size and the presence of myosin isoforms is determined by staining with silver nitrate [296]. The advantage of silver-staining is that there is no requirement for antibodies for the different myosin isoforms. The comparison of the dot blot typing and silver-staining was not reported in the original article [268], however I think is warranted. Whilst silver-staining or dot blotting is effective for initial typing of skeletal muscle fibres, a comprehensive assessment of the most abundant proteins in a sample is possible using

MS-based proteomic methods. Recent work has used MS-based proteomics to assess the myosin isoform content of both single muscle fibres and pooled muscle fibre samples [210, 297, 298]. After typing each sample using both the dot blot and silver-staining protocol, I used this same MS-based proteomic approach to assess the purity of the resultant pooled fibre samples.

Method

I isolated sixty single muscle fibres from three different samples under a light microscope on ice. Fibres were digested in a 96-well plate in 50 µL of buffer by agitation for four hours on a plate shaker. Plates were transferred to -80 °C. Upon thawing, 10 µL of each sample was spotted onto a nitrocellulose membrane using the dot blot apparatus and stained for MYH2 and MYH7. After typing each fibre, 10 µL of five MYH2 fibres, five MYH7 fibres and five potentially mixed fibres from two participants were loaded in a 4-10 % gel. Gels were run at 100 V for twenty hours and stained with silver and imaged. In addition, 10 µL of 15 MYH2 fibres, 15 MYH7 fibres and 15 potentially mixed fibres were pooled for three participants for MS-based proteomics.

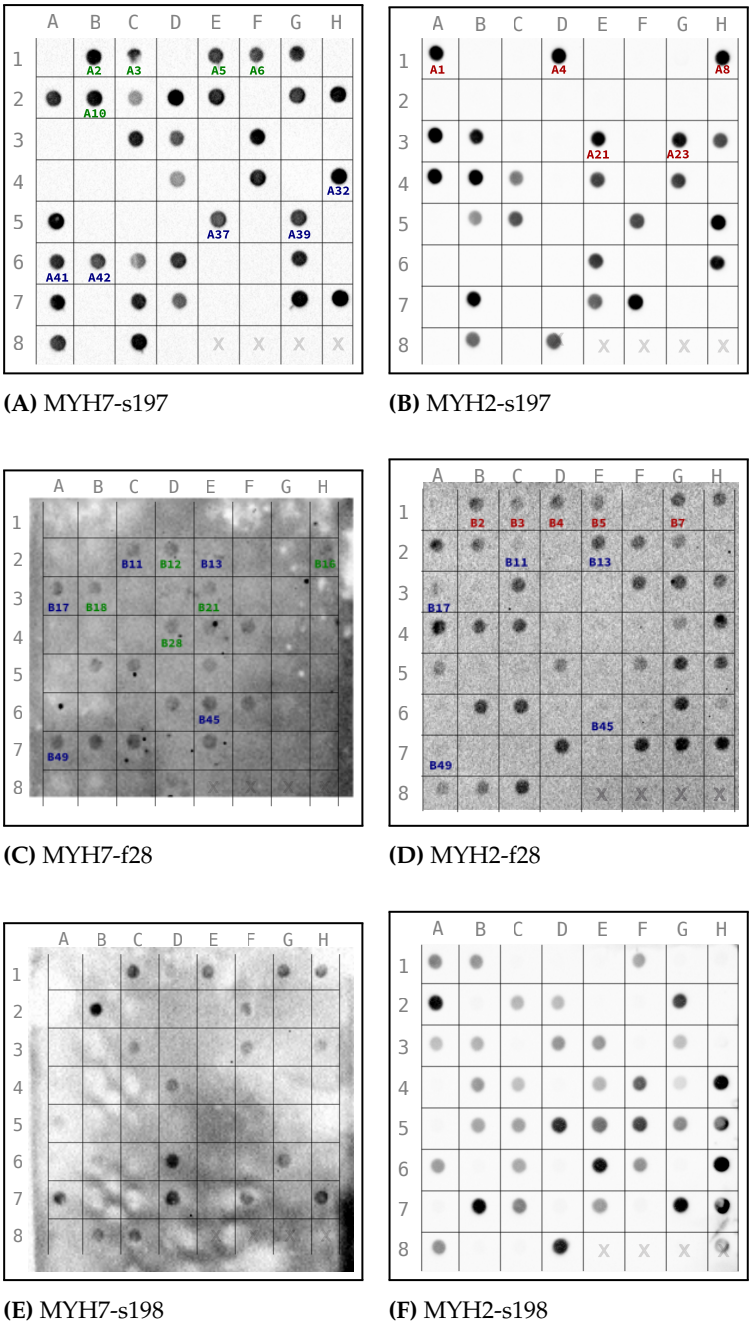
Results

Dot blotting, silver-staining and proteomics all confirm fibre typing of muscle fibre fragments

The dot blots in this experiment were of variable quality as seen in Figure 4.7. This could be due to the over-stripping of protein during the blotting or reblotting, or antibody binding issues. I proceeded with the experiment to assess how an optimal dot blot (Figures 4.7A & 4.7B), a poor dot blot (Figures 4.7C & 4.7D), and a medium quality dot blot (Figures 4.7E & 4.7F) would affect results from the silver-stain and fibre pooling. I chose fifteen fibres to run on a gel for silver-staining from samples s197 and f28 as annotated in Figure 4.7.

Interestingly the results of the dot blots (Figure 4.7) and silver-stain (Figure 4.8) show the same fibre typing. It has been shown that MYH7 migrates faster than MYH2 during electrophoresis [295]. The results of the silver-stain align with the dot blot typing results except for B13 in Figure 4.7B. This was typed as a mixed fibre, but most likely

represents a MYH2 positive, MYH7 negative Type II fibre. It may also be a result of the very low and unreliable signal in Figure 4.7C. Surprisingly, the positive control (a whole muscle lysate) sample shows the same migration pattern as the MYH2 fibres. Even with low quality dot blot results, I was able to correctly type all Type I and Type II fibres with the dot blot and silver-stain method.



(A) MYH7-s197

(B) MYH2-s197

(C) MYH7-f28

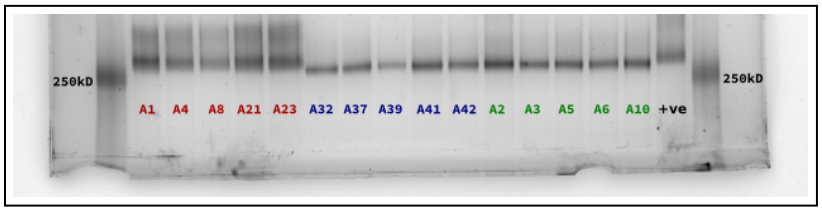
(D) MYH2-f28

(E) MYH7-s198

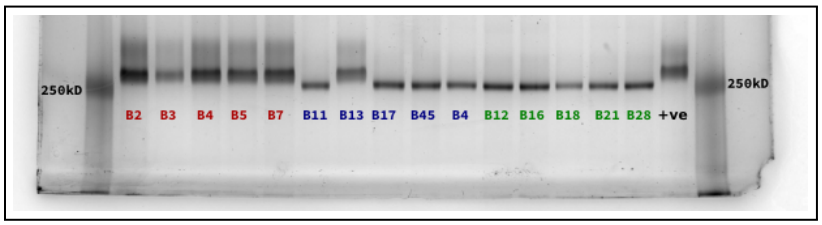
(F) MYH2-s198

Figure 4.7: Dot blot trial

60 single muscle fibre fragments from three individuals were fibre typed for MYH7 and MYH2 using a dot blot apparatus. (A), (B), (E) & (F) show two male participants and (C) & (D) shows a female participant. Indicated fibres in (A), (B), (C) & (D) were also used for silver-staining. Positions labelled with an X indicate that no fibre was loaded in this well.



(A) s197-Silverstain



(B) F28-Silverstain

Figure 4.8: Silver-stain

Five MYH7, MYH2 and potentially mixed fibres from two participants were run for 20 hours on a 4-10% gradient gel. The coloured numbering indicates the same fibre that was typed using the dot blot method. (A) & (B) show that MYH7 fibre fragments migrate faster than MYH2 fragments and typing aligns with the dot blot approach.

After typing single muscle fibres, I pooled 10 µL (one fifth of a fibre) of fifteen type I, fifteen Type II, and fifteen potential mixed fibres for each sample (I pooled 15 µL of five mixed fibres for sample f28 due to insufficient numbers of fibres). These pooled samples were frozen on dry ice and proteins were digested to peptides. Peptides were separated and identified using LC-MS/MS. Figure 4.9 shows the percentage of MYH7, MYH2, MYH1 and MYH4 in each pooled fibre sample. Type I fibres contain over 95 % MYH7 and Type II fibres contain over 80 % MYH2 and interestingly the remaining myosin isoform in Type II was MYH4. As I was unable to identify mixed fibres in samples s197 and s198 (4.7), I pooled an additional 15 Type I fibres for both samples. As confirmation of the sensitivity of the MS-based proteomics the mixed sample from f28 showed a mix of MYH7 and MYH2 which aligns with the results from the silver-staining (Figure 4.8B). The agreement of the dot blot, silver-stain and MS-based proteomics provide strong evidence that the dot blot method is appropriate for typing Type I and Type II skeletal muscle fibres.

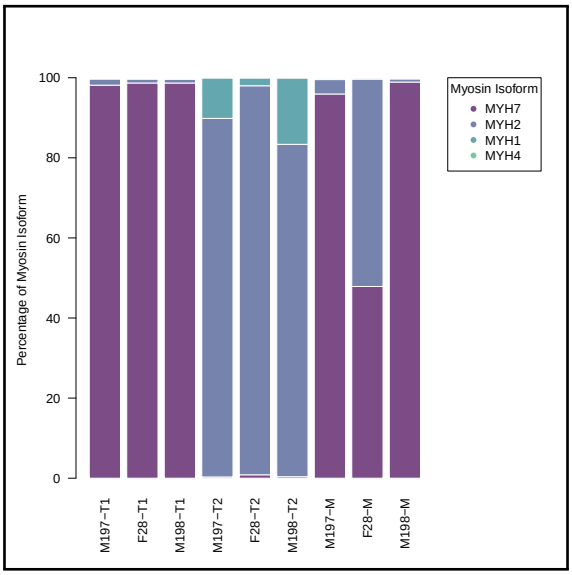


Figure 4.9: Myosin Isoform Content as measured by Mass Spectometry

Figure 4.9 shows the myosin isoform content of 15 pooled MYH2, 15 pooled MYH7 and 15 pooled potentially mixed fibres for three participants. The myosin isoform content aligns with both the dot blot and silver-stain results.

4.2.4 Proteomics

Rationale

One challenge in studying protein in skeletal muscle is that the overwhelming majority of quantifiable protein in a sample belongs to the contractile proteins. As a result, MS-based proteomics can be used to confirm the success of pooling skeletal muscle fibre fragments by quantifying myosin isoform content of each sample. In addition the results could also be used to study fibre type specific differences at baseline and after exercise as previously reported [210]. I investigated whether pooling one fifth of a fibre from fifteen fibres could be used to study protein expression differences between skeletal muscle fibre types.

Method

I assessed the quality of the proteomic results generated from pooling one fifth of fifteen muscle fibres. Proteins were filtered to include only proteins from *Homo sapiens* and proteins identified in at least two condition replicates. MaxLFQ values were log2 transformed, grouped by condition and missing values were imputed assuming they were MNAR. The limma [279] package was used to generate protein-wise linear models using

empirical bayes statistics that identified differentially expressed proteins for each pairwise comparison.

Results

The total number of proteins identified was 1052 with approximately 900 measured in each sample. The number of proteins was consistent across samples and conditions (Figure 4.10). A principal component analysis (PCA) analysis showed that samples clustered along the first principal component (PC) according to Type I and Type II fibres and that this component explained 44 % of the variation (Figure 4.11A). The clustering of samples along Type I and Type II fibres can also be seen in the sample correlation plot (Figure A.4) and missing value heatmap (Figure A.3) confirming a clear separation by fibre type. The PCA rotation plot confirmed that PC1 was driven by proteins associated with fibre type (Figure 4.11B).

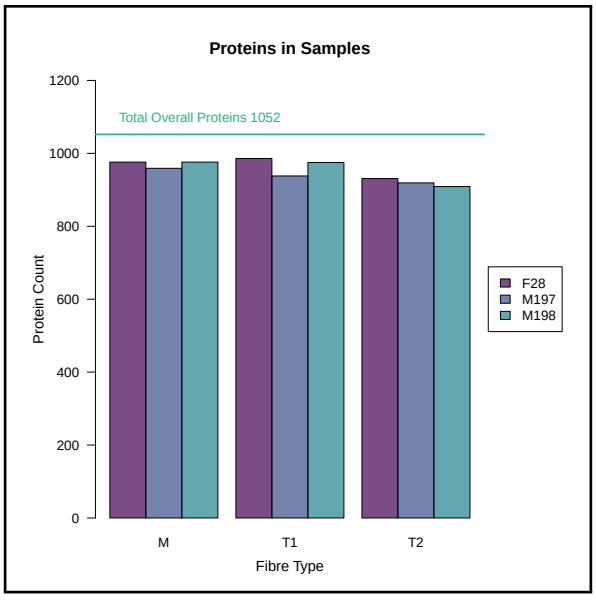
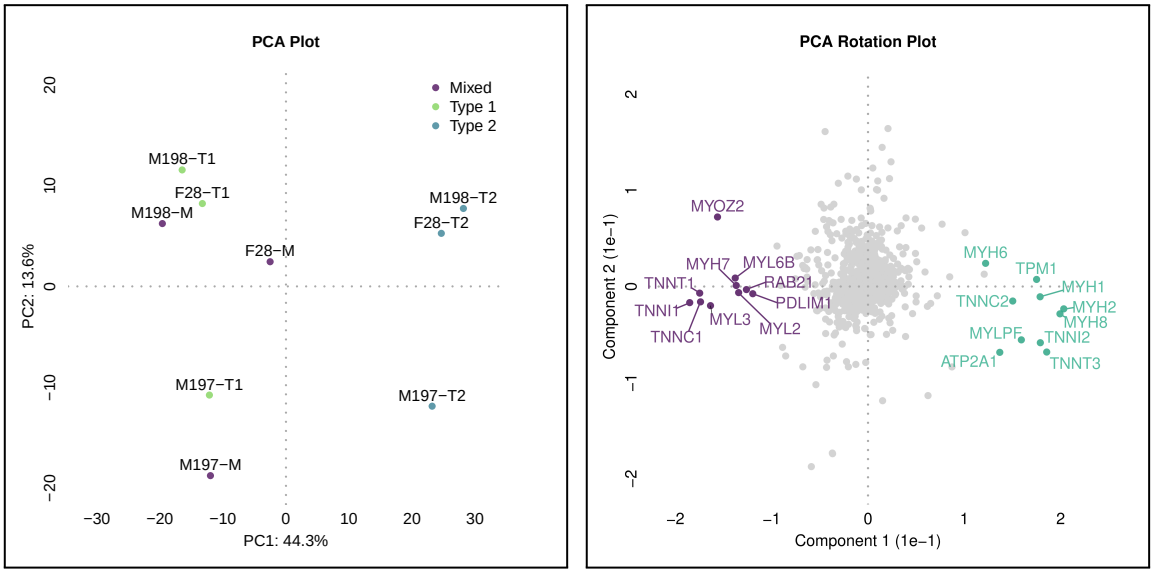


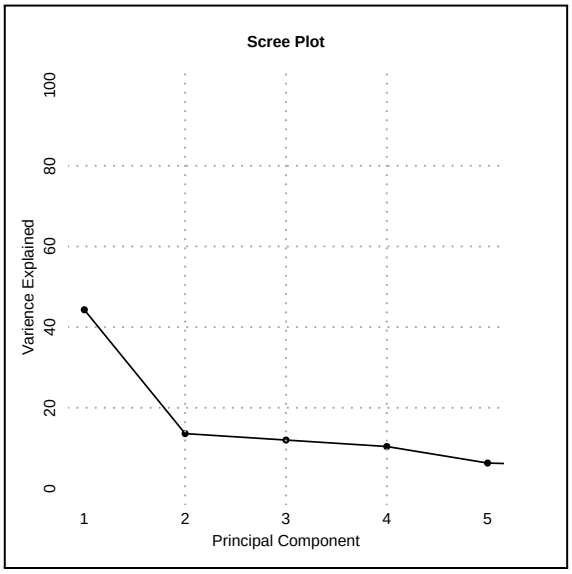
Figure 4.10: Protein Count per Sample

Figure 3.10 shows the number of proteins identified in each sample. Approximately 900 proteins were measured in each sample using MS-based proteomics.



(A) PCA

(B) PCA Rotations

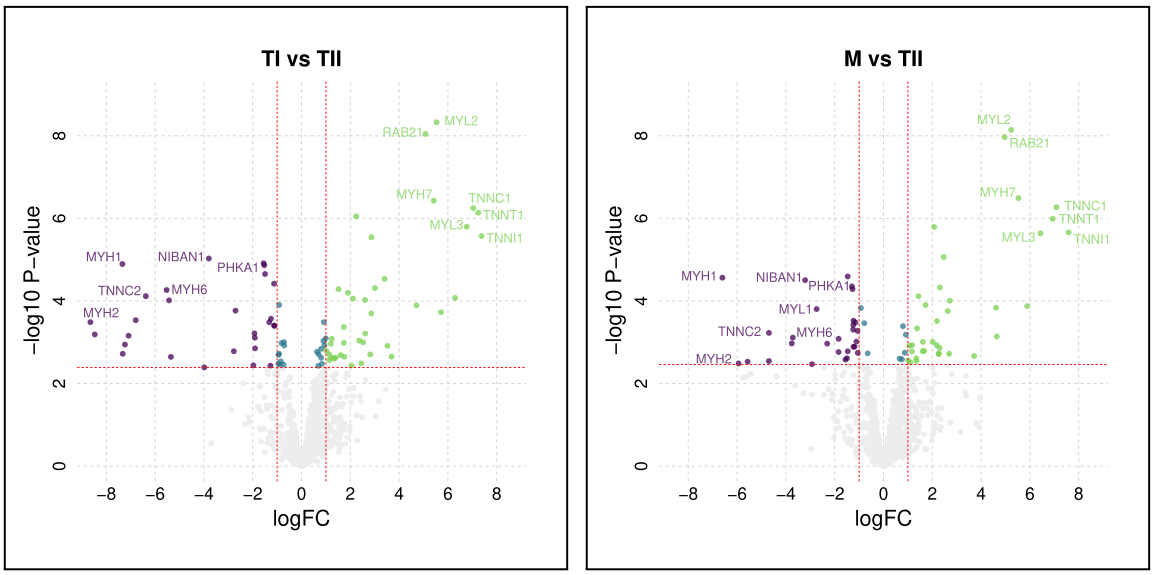


(C) Scree Plot

Figure 4.11: PCA Analysis of Proteome Trial

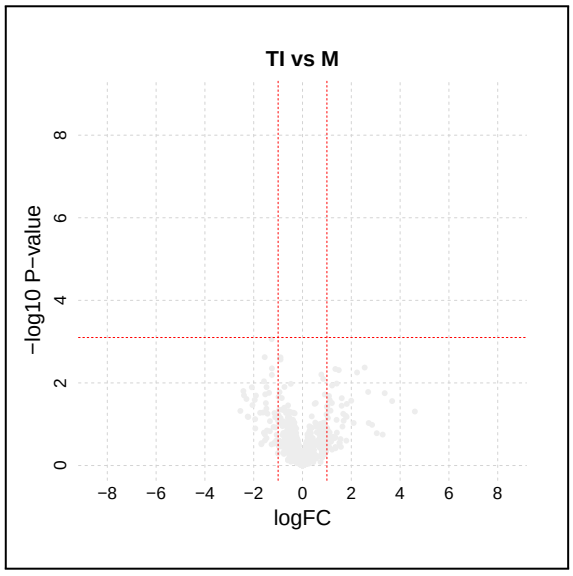
Figure 4.11 is a dimension reduction analysis of the major sources of variation within the dataset. (A) is a PCA plot showing that the first principal component explains 44% of the variation and is clustered according to fibre type. (B) shows the top twenty proteins that explain the variation in PC1. (C) shows a scree plot showing a sharp reduction in variation after PC1.

MYH7 and MYH2 are shown to be differentially expressed between Type I and Type II fibres and between the mixed samples and Type I fibres (figure 4.12). These results align with the myosin isoform content percentages of each sample shown in figure 4.9. It is clear from these results that the mixed fibre samples are mostly Type I fibres with sample f28 having some contamination from Type II fibres. A ranked intensity plot of proteins show MYH7 and MYH2 as some of he most abundant proteins in the samples (figure 4.13). The results from the proteomics highlight the importance of accurate typing and careful pooling of muscle fibre samples for use in downstream experiments. These results provide evidence that MS-based proteomics can be applied to pooled muscle fibre samples using as little as one fifth of a fibre segment. They also show the importance of checking the myosin isoform content of pooled muscle fibre type samples for accurate interpretation of results.



(A) T1 vs T2

(B) M vs T2



(C) M vs T1

Figure 4.12: Volcano Plots

Figure 4.12 shows the differentially expressed proteins between the pooled fibre types as identified using pair-wise comparisons. (A) shows the most significant proteins up regulated in either Type I or Type II fibres. (B) and (C) confirm that the potentially mixed fibre samples are most likely Type I fibres. The red dotted lines indicate an adjusted P-value < 0.05 and a logFC > ± 1.

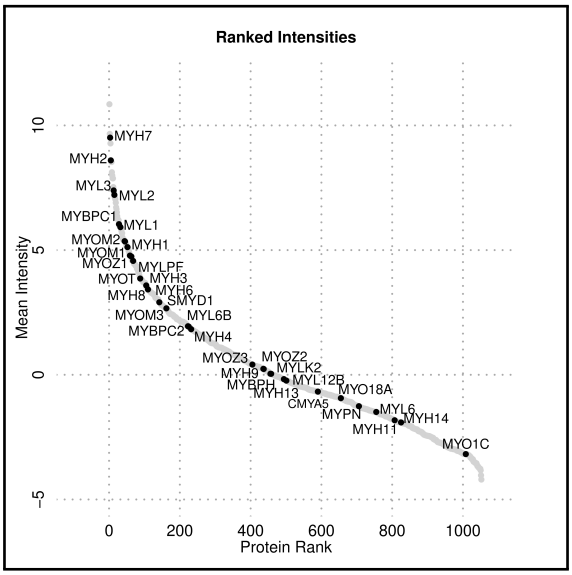


Figure 4.13: Ranked Intensities of Proteins of Proteome Trial

Figure 4.13 shows the ranked intensities of myosin proteins identified in all samples. The most abundant proteins in the samples belong to the major contractile proteins such as MYH7 and MYH2.

4.2.5 DNA extractions

Rationale

The majority of DNAm studies in human skeletal muscle have utilised the infinium EPIC beadchip array. The EPIC array measures methylation at approximately 850,000 CpGs and is a cost effective method to measure DNAm in multiple samples. I was interested in generating DNAm profiles of human skeletal muscle fibre types using the EPIC array whilst also comparing the results to DNAm in whole tissue skeletal muscle samples. The EPIC array requires a minimum of 250 ng of DNA, therefore I aimed to optimise a DNA extraction method from pooled skeletal muscle fibre fragments that could yield 250 ng of DNA. There has been one previous study of DNAm in Type I and Type II skeletal muscle fibres [3]. The approach utilised silicon columns to extract high quality gDNA from pooled Type I and Type II muscle fibre samples. I tested a number of DNA extraction methods to determine the best approach to maximise yield and quality of gDNA.

Method

I tested three general approaches to isolating gDNA: a silicon based column method; an organic phase separation method and a magnetic bead based method. I tested a variety

of silicon column kits as they are reported to be user-friendly and safer methods for DNA extraction. Bundles of twenty skeletal muscle fibres were separated into groups of five fibres and digested in 50 µL of lysis buffer for four hours on a plate shaker. Digested fibres were frozen at -80 °C to assist in cell disruption and lysis. After muscle fibres were thawed to RT, I pooled four wells of five fibres and tested six different DNA extraction protocols on ~twenty pooled fibres, according to the manufacturers instructions with minor modifications. The organic phase separation method was modified from the following protocol [269]. The main modification of each method was to replace the initial lysis buffer with the SDS-DTT lysis buffer. Once DNA was extracted, the concentration was measured using a Qubit flurometer and DNA quality assessed by running gDNA on a 1.5 % agarose gel stained with gelRed..

Results

Qiagen QIAMP DNAmicro kit and PCI DNA methods both yield DNA

I decided to test multiple DNA extraction methods with slight modifications to allow for the use of an SDS-TRIS-DTT buffer. DNA yields of each method are summarised in Table 4.1, where the Qiagen QIAMP DNAmicro kit and PCI methods yielded the most gDNA. On average the PCI method produced higher concentrations of DNA and equalled a DNA yield of approximately 4.5 ng of gDNA per muscle fibre fragment. The other silicon columns (lanes four to seven) and the magnetic beads lanes (ten and eleven) produced significantly less gDNA Table (4.1). The quality of gDNA was high for both the Qiagen and PCI methods (Figure 4.14).

Phenol chloroform isoamyl enables higher yield DNA

On average the PCI method produced higher concentrations of DNA. PCI extractions yielded approximately 4.5 ng of gDNA per muscle fibre fragment, whereas the Qiagen QIAMP DNAmicro kit averaged 3.8 ng per fibre fragment (table 4.1). From initial testing the PCI extraction method is a viable method for extracting high quality gDNA. As the goal of this method was to maximise DNA yield, this method was chosen for further experiments.

Table 4.1: DNA extraction methods

Lane	Extraction Method	DNA yield (total)
Lane 1	DNA Ladder	
Lane 2	Qiagen QIAamp DNAmicro	72 ng
Lane 3	Qiagen QIAamp DNAmicro	81 ng
Lane 4	Qiagen AllPrep DNA/RNA	5 ng
Lane 5	Qiagen AllPrep DNA/RNA	31 ng
Lane 6	NucleoSpin XS (MN)	18 ng
Lane 7	NucleoSpin XS (MN)	24 ng
Lane 8	Phenol/Choloroform/Isoamyl	76 ng
Lane 9	Phenol/Choloroform/Isoamyl	101 ng
Lane 10	NucleoMag Tissue (MN)	40 ng
Lane 11	NucleoMag Tissue (MN)	40 ng

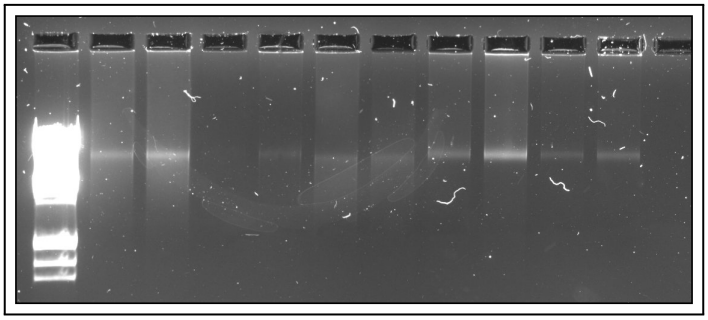


Figure 4.14: DNA extraction tests

DNA extracted from 20 muscle fibre fragments using five methods were loaded on a 1.5% agarose gel and run for 45 minutes. Lanes 2-3 are DNA samples extracted using the Qiagen QIAmp DNAmicro kit/ Lanes 8-9 are DNA samples extracted using the PCI methods. These two methods yielded the most gDNA.

PCI alcohol is a cost effective and viable approach to isolate gDNA from pooled skeletal muscle fibre fragments

In the previous experiment 20 muscle fibres were digested in groups of five and pooled in only 200 µL of total volume. To scale this method up to 50 fibres would require a large volume of starting lysis buffer. Therefore I next tested the PCI method for extracting gDNA where a single fibre was digested in a tube in 10 µL of lysis buffer. After, lysis and digestion, 2 µL of fibre lysate was dot blotted and typed for MYH7 and MYH2. The remaining 8 µL was used for pooling and DNA extractions using the PCI method.

Between 45 and 50 fibres were pooled together resulting in approximately 500 µL of total liquid used for DNA extraction. Carrier RNA is often added to increase DNA yield and this was added to one sample [3]. From 45-50 fibres the PCI method yielded on average 212 ng, equivalent to an average 4.5 ng of DNA per fibre fragment (table 4.2). This is the same yield as when testing DNA extractions using 20 fibres, in which five fibres were digested per tube. Figure 4.15 confirms that the gDNA was of high quality. Interestingly, the sample in lane six included carrier RNA, and in this experiment it did not produce higher yields of gDNA (Table 4.2) but smearing on an agarose gel was observed instead (Figure 4.15). These results show that the PCI method is consistent at producing high quality gDNA from pooled single skeletal muscle fibre fragments.

Table 4.2: PCI DNA extractions

Lane	Starting Material	DNA yield (total)
Lane 1	DNA Ladder	
Lane 2	50 Type I fibres	307 ng
Lane 3	50 Type I fibres	173 ng
Lane 4	50 Type II fibres	230 ng
Lane 5	50 Type II fibres	216 ng
Lane 6	50 Type I fibres	196 ng
Lane 7	45 fibres	193 ng
Lane 8	45 fibres	168 ng

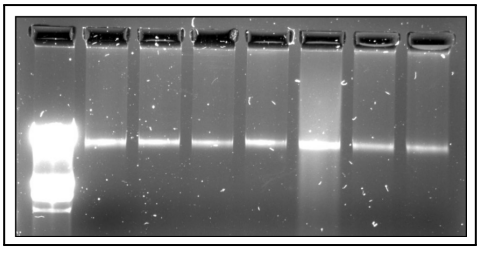


Figure 4.15: DNA extractions PCI pooled single fibres

DNA extracted from 50 muscle fibre fragments using the PCI method were loaded on a 1.5% agarose gel and run for 45 minutes. From 45-50 fibres the PCI method yielded on average 212 ng, equivalent to an average 4.5 ng of DNA per fibre fragment.

4.3 Discussion

This chapter presents a modified method for the simultaneous extraction of DNA and proteins from human skeletal muscle fibre fragments. This work combined and optimised previous methods used to study DNA or proteins in skeletal muscle fibres to provide a flexible approach that can make the most of precious human muscle tissue samples [3, 268, 210]. I demonstrated that by pooling Type I (TI) and Type II (TII) muscle fibre fragments, I obtained enough DNA and protein to run DNAm analyses with the EPIC array and study protein expression from pooled human skeletal muscle fibre samples (Table 4.2 & Figure 4.10).

Previous studies successfully extracted gDNA from pooled skeletal muscle fibres by cutting fibres and using one piece for fibre type identification and another piece for DNA extraction [3]. I have developed the following protocol for the preparation of DNA and proteins from skeletal muscle fibre fragments. Single fibre fragments are digested in 10 µL of an SDS-TRIS buffer in a small PCR tube. One fifth of a fibre is used for accurate fibre typing using a dot blot apparatus. After fibre typing, one fifth of a fibre is pooled for proteomic analysis of pooled fibre type samples. The remaining three fifths of each single fibre can be used to extract high quality gDNA using a PCI protocol.

I have successfully increased the yield of gDNA by using a dot blot apparatus to efficiently type more fibres followed by PCI extraction of DNA. Three fifths of a fibre are expected to yield, depending on fibre length, 3.5 ng of gDNA on average. Therefore, pooling between 65-85 single fibre fragments should yield 250 ng of gDNA. Compared to previous work [3] this increase in yield would enable us to utilise the commonly used Infinium Human BeadChip arrays for probing DNAm and to extend the study of DNAm in skeletal muscle fibre types to multiple samples.

It has previously been reported that typing Type IIx muscle fibres through a process of elimination using the dot blot method is possible [268], however when using a dot blot apparatus we were unable to determine Type IIx fibres using the same approach. Christiansen and others [268] did highlight that antibodies used during dot blot fibre typing are a limitation due to cross reactivity. At the outset of this work I intended to study Type IIx fibres from multiple samples. Due to the inability to accurately identify

Type IIx fibres, I proceeded to characterise Type I slow and Type II fast fibre populations in the following chapters.

Building on the work of Christiansen and others [268], I compared the dot blot method to both silver-staining and LC-MS/MS and showed that all three methods showed agreement when typing Type I and Type II fibres (Figures 4.7, 4.8 and 4.9). This is consistent with Deshmukh et al. who also utilised the dot blot method and proteomics to study differences in Type I and Type II fibres [210].

Using both a recently reported fibre typing method [268] and a proteomic study on pooled skeletal muscle fibres [210], I developed a one tube digestion protocol for the preparation of both DNA and proteins from skeletal muscle fibre fragments. This method would allow researchers to investigate both epigenetic and protein differences in skeletal muscle fibre types from the same biological sample. There is also the potential to adapt this method to extract DNA, RNA and proteins from the same sample to study role of DNAm in the control of both RNA and protein expression in skeletal muscle fibres.

In the subsequent chapters I apply this method to investigate DNAm differences between Type I and Type II human skeletal muscle fibres in twelve healthy adults. I also investigate changes in the methylome and proteome after 12 weeks of High Intensity Interval Training.

5 Results 2: Fibre-Type-Specific and Whole Muscle DNA Methylation differences in human skeletal muscle at baseline

5.1 Introduction

DNA methylation is the chemical addition of a methyl group to the fifth carbon of cytosine nucleotides [35] and occurs the majority of the time at CpG dinucleotides [36, 37]. DNAm can be measured using a number of platforms such as sequencing or arrays. Due to the cost effectiveness of DNAm arrays and their widespread use, I studied DNAm in human skeletal muscle samples using the Infinium MethylationEPIC Beadchip. Recently the EPIC array has been updated as the new EPICv2 beadchip which contains more than 900,000 CpG methylation sites [299].

Skeletal muscle is a heterogeneous organ comprised of multiple cell types, each with its own specific function. The muscle fibre is a major component of muscle tissue and is categorised into three specific types in humans: Type I, Type IIa and Type IIx [59]. To date, studies investigating DNAm in human muscle have predominantly focused on assessing global DNAm of whole tissue samples. Due to the importance of cell-type on DNAm signatures, studying whole tissue samples often requires adjustments for cell-type proportions to gain meaningful biological insights [300]. Begue and others reported DNAm differences between muscle fibre-types using one pooled Type I fibre and one pooled Type II fibre mixed from eight male samples [3]. The overarching aim of this chapter was to compare DNAm profiles between Type I and Type II muscle fibres and whole muscle in both males and females. We extended the study of human skeletal muscle fibre-type-specific DNAm by utilising our modified fibre-typing and DNA/protein extraction method. The application of this method enabled us to assess DNAm using the new EPICv2 Infinium beadchip in 12 individual pooled Type I and Type II muscle

fibre samples from both males (n=7) and females (n=5). We generated fibre-type-specific DNAm profiles from human skeletal muscle that provides a valuable resource for researchers studying DNAm in human skeletal muscle.

The specific aims of this chapter were:

1. To isolate, type and extract ~250ng of DNA and protein from Type I and Type II muscle fibres from 12 human skeletal muscle samples.

2. To assess the most appropriate pre-processing and normalisation method for the new EPICv2 array in the context of human skeletal muscle samples.

3. To compare DNAm of pooled Type I, pooled Type II and whole muscle (WM) samples using the new EPICv2 from multiple samples.

5.2 Methods

5.2.1 Participant Summary

Participant Characteristics

15 apparently healthy participants (males = 8, females = 7) were recruited as part of the ongoing Gene SMART study [260]. At baseline all 15 participants muscle samples were used for whole muscle DNA extraction and a total of 12 participant's (males = 7, females = 5) muscle samples were used to isolate fibre-type specific DNA. At the time of recruitment 34 was the median participant's age (range = 25-42) and mean was body mass index (BMI) 24.6 kg.m^2 (range = 18.0-28.7 kg.m^2). Physiological fitness was measured at baseline: mean VO_{2peak} 39.1 ml.min.kg (range = 27.3-55.7 ml.min.kg); mean lactate threshold 153.5 W (range = 107.5-249.5 W); and mean aerobic W_{peak} 217 W (range = 162.5-351.0 W). Participant characteristics and fitness measures at baseline are summarised in Table 5.1.

Table 5.1: Participant Baseline Characteristics

Participants		**n=15**
Sex	Male	8
	Female	7
Age	Median	34
	Min-Max	25-42
BMI	Mean (SD)	24.6 (3.0)
	Min-Max	18.0-28.7
VO2max	Mean (SD)	39.1 (8.7)
	Min-Max	27.3-55.7
Lactate Threshold	Mean (SD)	153.5 (42.8)
	Min-Max	107.5-249.5
Aerobic Peak Power	Mean (SD)	217.0 (55.5)
	Min-Max	162.5-351.0

5.2.2 DNA, protein and sample preparation

Dot blot fibre typing of single fibre fragments

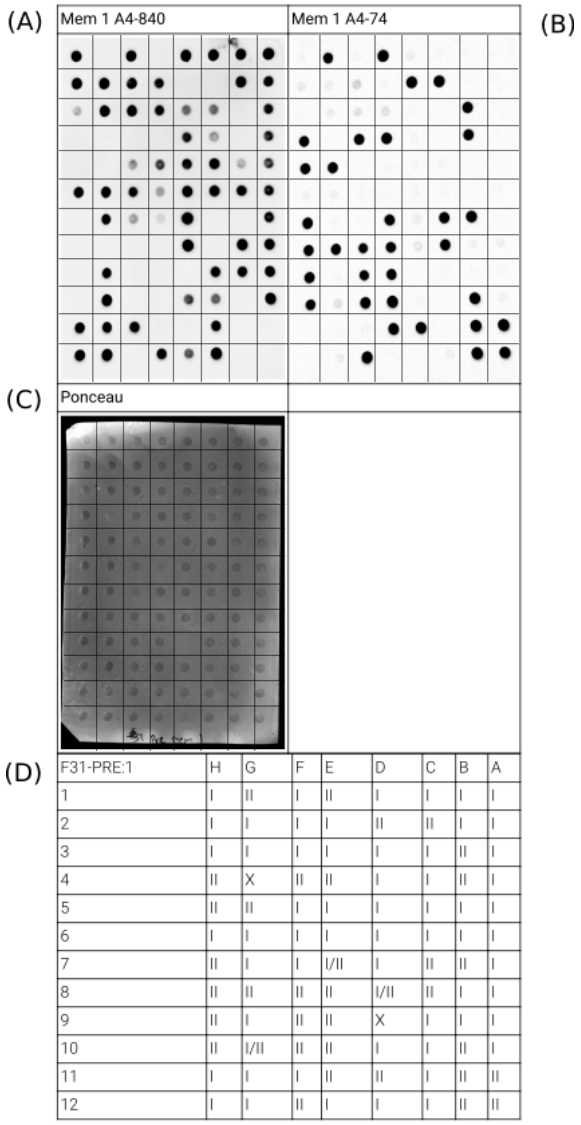
A total of 288 fibres were isolated under a light microscope for each muscle sample and dot blotted using a 96 well dotblot apparatus. Figure 5.1 is a representative blot of 96 fibres isolated from one participant at baseline. Fibres were typed for MYH7 (Type I fibres), MYH2 (Type II fibres) and ponceau stained. Membranes were imaged using a MP Biorad ChemiDoc and contrasted depending on signal intensity. Fibres were identified as Type I if they contained only MYH7, Type II if they contained only MYH2, contaminated if they contained both MYH7 and MYH2 or as Type IIx fibres if they stained positive for ponceau and were negative for both MYH7/MYH2. As seen in Figure 5.1 some fibres were classed as MYH7 positive and MYH2 negative even with small amounts of MYH2 signal present as signal was considered to be background staining. This decision was made as a result of previous experiments outlined in the previous chapter. Fibres that were negative in all three blots were labelled missing fibres. Missing fibres were marked with a X upon typing as shown in Figure 5.1D. The cause of missing fibres is likely due to no single fibre being placed in each tube or the fibre remaining stuck to the tube wall.

The total number of fibres isolated across all samples are summarised in Table 5.2. At baseline a total of 3,218 fibres were isolated and typed. Across 12 participants an average of 113 Type I fibres (min-max = 65-179) were pooled per sample, and an average of 119 (min-max = 66-155) Type II fibres were pooled per sample resulting in a total of 1,353 Type I fibres and 1,427 Type II fibres being used for DNA and protein extractions. On average 23 (min-max = 0-68) Type IIx fibres were identified per sample resulting in a total of 272 Type IIx fibres across all samples. One IIx fibre sample was pooled together for DNA extractions due to low numbers of Type IIx fibres identified using the ponceau approach. On average 14 contaminated fibres (min-max = 2-43) were identified per sample for a total of 166 contaminated fibres. Contaminated fibres were not used in the following study as it was difficult to determine if these fibres represent true hybrid fibres or were contaminated during the isolation procedure.

For the purpose of this chapter we only show DNAm results, whereas the proteins were analysed in the next chapter of this thesis.

Table 5.2: Fibre Numbers

Participants (n=12)	Pre	Post	Total
Type I Fibres	1353	1426	
Type II Fibres	1427	1357	
Type IIx Fibres	272	199	
Contaminated Fibres	166	145	
Missing Fibres	62	33	
Total Fibres	3218	3127	6345



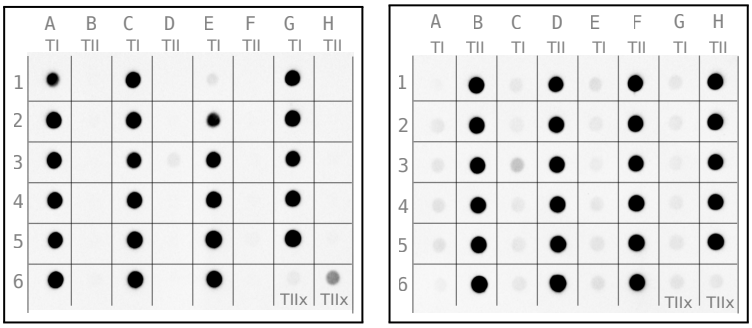
F31-PRE:1	H	G	F	E	D	C	B	A
1	I	II	I	II	I	I	I	I
2	I	I	I	I	II	II	I	I
3	I	I	I	I	I	I	II	I
4	II	X	II	II	I	I	II	I
5	II	II	I	I	I	I	I	I
6	I	I	I	I	I	I	I	I
7	II	I	I	I/II	I	II	II	I
8	II	II	II	II	I/II	II	I	I
9	II	I	II	II	X	I	I	I
10	II	I/II	II	II	I	I	II	I
11	I	I	I	II	II	I	II	II
12	I	I	II	I	I	I	II	II

Figure 5.1: Representative Dot Blot

A representative blot of 96 single muscle fibre fragments dot blotted using a dot blot apparatus. Samples were blotted on one membrane and typed for MYH7 (A) stripped and reprobed for MYH2 (B) and finally stained with ponceau (C). (D) Fibres were typed as either type I (I = MYH7+/MYH2-), type II (II = MYH2+/MYH7-), contaminated (I/II = MYH7+/MYH2+), potential type IIx (IIx = MYH7-/MYH2-/Ponceau+) or missing (X = MYH7-/MYH2-/Ponceau-).

Fibre type purity of pooled samples

After single muscle fibre fragments were typed, one fifth (2 µL) of each fibre was pooled for mass spectrometry analysis of proteins. This resulted in between 132-358 µL of protein lysate per sample. 10 µL of each pooled sample was dot blotted to validate if sample pooling was accurate and 100 µL was sent for mass-spectrometry analysis of proteins. As seen in Figure 5.2 columns A, C, E and G represent pooled Type I fibres and were determined to be MYH7 Type I fibres. Columns B, D, F, and H represent pooled Type II fibres and were determined to be MYH2 positive Type II fibres. Dot 6 in both column G and column H are the pooled Type IIx fibres from baseline and after 12 weeks exercise (see chapter 6), interestingly these fibres show a reduced staining for MYH7 and MYH2. Dot one of column E shows a low amount of both MYH7 and MYH2 expression. The varying intensity of signal could be the result of varying fibre length between muscle samples and therefore a result of lower protein concentrations.



(A) MYH7: Pooled Samples **(B)** MYH2: Pooled Samples

Figure 5.2: Pooled Samples Dot Blot

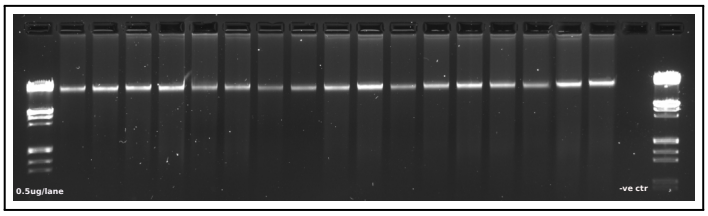
Pooled fibre-type specific samples were dot blotted for MYH7 and MYH2 to assess sample purity. (A) shows MYH7 and indicates that Type I fibres are MYH7 positive and (B) shows MYH2 and indicates that Type II fibres are MYH2 positive.

DNA extractions and quality from pooled fibre fragments

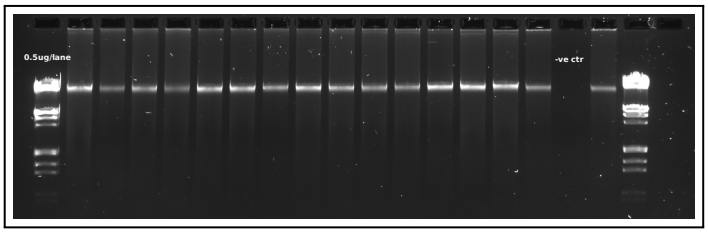
After typing each single fibre fragment, four fifths (8 µL) of each fibre was pooled as a Type I or Type II sample. gDNA was subsequently extracted using a modified PCI method. On average 3.2 ng of gDNA was isolated per single fibre fragment (min-max = 1-7.7 ng) which was dependent on the length of fibres. On average 404 ng of gDNA was extracted from pooled Type I samples (min-max = 128-900 ng) and 368 ng of gDNA was extracted from Type II fibres (min-max = 66-816 ng) as summarised in table 5.3. Less than 250 ng of total gDNA was obtained from 13 out of a total of 48 pooled fibre samples. Five of these samples had less than 200 ng of gDNA and an additional 96 fibres from each of these samples were isolated to obtain more gDNA. At the end of all DNA extractions total DNA concentrations were measured for the 48 samples: three samples had between 70-190 ng; seven had between 200-250 ng; and 38 samples had greater than 250 ng of gDNA. We proceeded to send gDNA for all samples for DNAm assessment with the EPICv2 as it was recently reported that as little as 40 ng of DNA can be used with the EPIC array [301]. To assess the quality of gDNA we measured A260/A280 values and ran agarose gels for all samples with more than 250 ng of gDNA. The average A260/A280 value for pooled muscle fibre samples was 1.7 (min-max = 1.54-1.89) as shown in table 5.3. The gDNA from pooled muscle fibres was of high quality when run on 1.5% agarose gels and stained with gelRed (Figure 5.3).

Table 5.3: DNA concentrations Pooled Fibres

	Type I	Type II	Combined
DNA (ng)	404	368	386
A260/A280	1.7	1.7	1.7



(A) Agarose Gel 1 of Pooled Fibre DNA samples



(B) Agarose Gel 2 of Pooled Fibre DNA samples

Figure 5.3: Pooled Fibre Type DNA

gDNA extracted from pooled muscle fibre fragments were loaded on 1.5% agarose gels and run for 45 minutes. (A) & (B) show all pooled fibre samples that had yield > 250 ng of gDNA. gDNA was of high quality as seen by one clear band on agarose gels. The first and last lane on each gel was a DNA ladder and the missing lanes on both gels were negative controls loaded without gDNA

DNA extractions and quality from Whole Muscle

DNA was extracted from whole muscle from a total of 15 participants at baseline, after four weeks of HIIT and at the end of the 12 week HIIT program (total samples = 43 due to one drop out after three weeks of training). Between 10-12 mg of whole muscle was used to extract both gDNA and RNA using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen, Venlo, Netherlands). gDNA quality was measured using a Nanodrop One (ThermoScientific, Massachusetts, USA) and the average A260/A280 value was 1.84 (min-max = 1.66-1.93). 75 ng of gDNA was run on a 1.5% agarose gel to assess DNA quality. A representative gel is shown in Figure 5.4 and gDNA was of high quality.

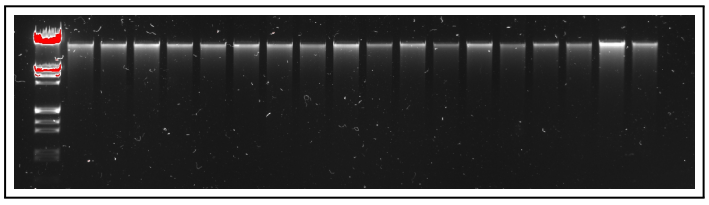


Figure 5.4: Whole Muscle DNA

A representative agarose of gDNA extracted from whole muscle samples were loaded on a 1.5% agarose gel and run for 45 minutes. gDNA was of high quality as seen by one clear band on agarose gels. The first lane on the gel was a DNA ladder.

5.2.3 Sample overview

A total of 96 samples were used to assess DNAm in human skeletal muscle using the Infinium MethylationEPIC Beadchip v2.0 (Illumina). 250 ng of gDNA was bisulfite converted in one plate and samples were hybridised to the EPICv2 arrays using a randomised stratified approach across array and position. A summary of the 96 samples are given in table 5.4. As there were five positions left on the 96 plate, three whole muscle replicates and one each of a pooled Type I and Type II fibre sample were included for additional analyses.

DNAm values from the Infinium arrays are pre-processed and normalised to remove unwanted variation [302]. However, the use of normalisation methods can shift methylation values away from their true values [303] and there is a risk that biological variation is removed [304]. To overcome this I utilised some spare positions on the array to include replicates to guide DNAm pre-processing and normalisation of values from the new EPICv2 in human skeletal muscle samples.

Table 5.4: Samples Assessed with the EPICv2

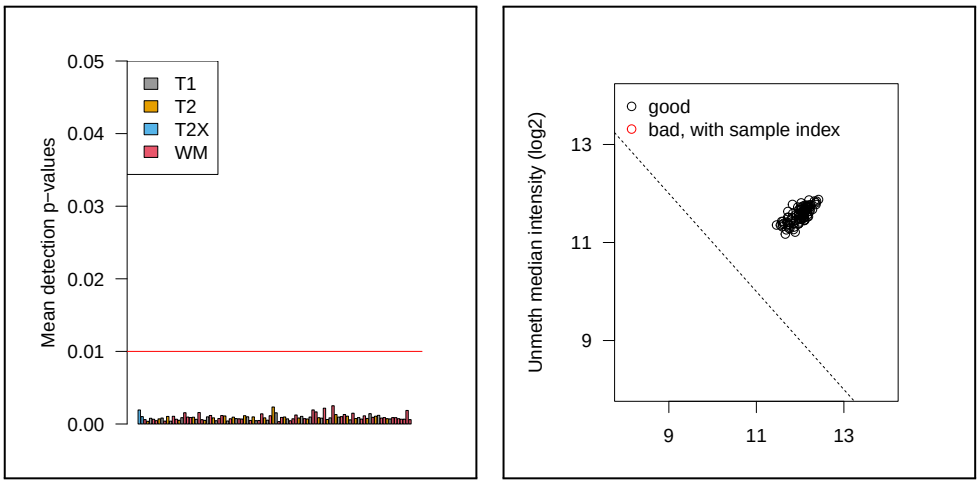
Sample Group	Baseline	4 Weeks	Post	
Type I Fibres	12	0	11	
Type II Fibres	12	0	11	
Pooled Type IIx	1	0	1	
Pooled Type I	1	0	0	
Pooled Type II	1	0	0	
Whole Muscle	15	14	14	
Replicates	3	0	0	
Total Samples	45	14	37	Total 96

5.3 Results

5.3.1 DNA methylation preprocessing and sample quality control

Sample QC

To assess sample quality, a number of quality control (QC) measures were performed on all 96 samples. DNAm values were produced by the human genomics facility at Erasmus University Medical Center and sample and probe based quality control was provided along with raw IDAT files. Further QC was performed in house to assess the quality of DNAm data produced from human skeletal muscle fibre samples using the new EPICv2 platform. All samples had < 0.01 mean detection P values (Figure 5.5A), and log median intensities of the methylated vs unmethylated channels >10.5 (Figure 5.5B), indicating all samples were reliable. To check for potential sample swaps, the sex of samples were predicted based on the difference in log2 median intensity for the X and Y chromosome probes using minfi. Figure B.1 indicates that no samples were mismatched for sex. Based on the results from sample QC no samples were flagged for exclusion. The samples with starting DNA of < 250 ng showed reliable quality control metrics and were included in the following analysis.



(A) Sample mean detection P values

(B) Methylated vs Unmethylated Plot

Figure 5.5: Detection P and QC plots

(A) Mean detection P values for each sample were calculated and plotted. All samples had mean detection P values < 0.01. (B) Log 2 median methylated vs unmethylated values were calculated and plotted for each sample. All samples had > 10.5 Log 2 median methylated vs unmethylated values indicating good quality samples.

Pre-processing comparisons for the EPICv2

It is standard practice to preprocess DNAm data generated using Infinium Methylation BeadChips to minimise unwanted technical variation due to the array design. Beta values are calculated as:

$$\beta = M/(U + M + 100) \tag{5.1}$$

where:

$$M = \text{intensity of the methylated allele}$$

$$U = \text{intensity of the unmethylated allele}$$

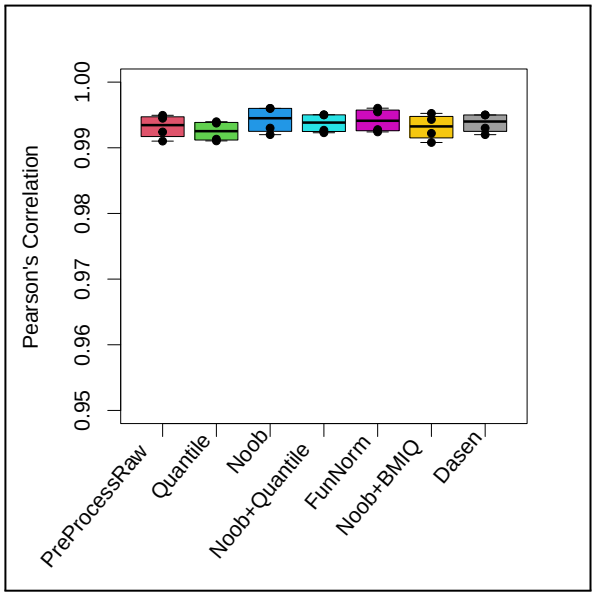
A number of preprocessing workflows have been established to correct experimental variation, however the selection of preprocessing method is not trivial and can lead to loss of methylation signal [305]. Beta values were filtered and normalised using seven published normalisation techniques. To assess the various normalisation methods, five replicate samples were used to assess the agreement of DNAm signals after each preprocessing method. One limitation in using only five replicate samples is that it is recommended to assess agreement in DNAm signal in a minimum of 30 replicate pairs [306]. As this is the first study (to my knowledge) to analyse DNAm measured by the EPICv2 in skeletal muscle and skeletal muscle fibre types, I used the following results to guide sample preprocessing.

Absolute beta value differences between the replicates for each normalisation technique are summarised in table 5.6 along with the percentage of CpGs exceeding a $|\Delta\beta|$ of 0.05 and 0.10. The best performing normalisation method based on $|\Delta\beta|$ differences was the Dasen normalisation method provided in the wateRmelon package in R. Dasen normalisation had a mean $|\Delta\beta|$ of 0.018 and the lowest percentage of CpGs exceeding $|\Delta\beta|$ of 0.05 and 0.10 (11.6% and 1.17% respectively). The worst performing normalisation methods based on $|\Delta\beta|$ were Noob+BMIQ as provided in the ChAMP package and preprocessRaw in the minfi package. Overall these values are higher than previously

reported [307], which may be a result of the replicate samples in this study being placed across array and position. Figure 5.6 shows correlations between duplicate pairs with all normalisation methods showing a similar correlation of > 0.99.

Table 5.6: Absolute Beta differences

	Min	Q1	Mean	Q3	Max	$\|\Delta\beta\|$ > 0.10	$\|\Delta\beta\|$ > 0.05
Raw	0.020	0.021	0.023	0.023	0.028	2.02	17.24
Quantile	0.018	0.018	0.020	0.022	0.022	1.63	13.34
Noob	0.014	0.015	0.017	0.190	0.020	2.44	13.85
Noob + Quantile	0.016	0.016	0.018	0.019	0.020	2.28	13.67
FunNorm	0.014	0.015	0.017	0.019	0.019	2.21	13.30
Noob + BMIQ	0.016	0.019	0.020	0.022	0.025	3.59	18.44
Dasen	0.016	0.016	0.018	0.020	0.021	1.17	11.66



(A) Duplicate Pearson's correlation between normalisation methods

Figure 5.6: Normalisation method Pearson's correlations

Seven published normalisation methods were assessed for duplicate pair correlations. Pearson's correlation values were calculated for all replicate pairs and all normalisation methods had similar correlation of > 0.99.

Density plots of normalised and filtered beta values are shown in Figure B.3 each showing the typical bimodal distribution of beta values around 0 and 1. Compared to the preprocessRaw normalisation the other methods show a smoother correction of beta

values. The noob plus BMIQ normalisation as recommended in the ChAMP package displayed a tighter normalisation for beta values between 0.2 and 0.8 as previously reported [307]. As values between 0.2-0.8 are likely to be of interest between the different cell types in a sample, we checked which variables explain the most data variation using principal component regression. PC regression was used to assess the association of the top 20 principal components and the covariates cell type, slide, array, sex, participant, ethnicity, timepoint and age. Normalisation methods have different effects on the covariates association with the top 20 principal components (Figures B.4). Dasen normalisation retained the most significant association between cell type and the first principal component whilst minimising associations of slide and array with the top five principal components (Figure B.4G) . The Dasen normalisation method was chosen as the normalisation approach due to the lower $|\Delta\beta|$ differences and PC regression results.

5.3.2 Probe Filtering and Quality Control

Probe Filtering

A number of probe filtering steps were implemented to remove unreliable and unwanted variation from the data prior to sample analysis. Probe filtering is summarised in table 5.7. Detection P-value (detp) values were generated to assess probe reliability using minfi and 24,115 probes with detp > 0.01 in one or more samples were removed. A major source of variation in DNAm data is driven by the sex chromosomes as seen in Figure B.5, therefore 22,779 probes on sex chromosomes were removed prior to analysis. Non CpG and single nucleotide polymorphism (SNP) related probes were removed. Cross reactive probes have been identified for the EPICv1 [276] and 450K [277] arrays and we removed probes remaining from a list provided from Erasmus MC. Probes flagged by Illumina as underperforming, inaccurate and masked on the EPICv2 were removed. In total 91,730 probes were removed prior to differential methylation analysis.

Table 5.7: Probe Filtering

Filter Step	Probe Number	Dropped Probes
Full Probe	930075	0
Detection P > 0.01	905960	24115
Bead count < 3	905522	438
Sex Chromosomes	882743	22779
Non CpG	879450	3293
Underperforming Probes	849474	29976
Innaccurate Probes	849355	119
Masked Probes	849354	1
Cross Reactive Probes	849354	0
SNPs	838705	10649
Final Probes Remaining	838705	
Probes Removed		91370

Batch Assessment

Batch effects due to plate, array and position are major sources of variation when using Illumina BeadChips. All samples in this study were bisulfite converted in one 96 well plate to remove plate as a source of variation. PCA was used to assess the influence of array and position on the DNAm data. In B.2A and B.2B samples from each array and position were present in each quadrant of the PCA plot. Both array and position were not major sources of variation on PC regression, particularly when using Dasen normalisation (Figure B.4G). As a result batch correction for array and position was not performed.

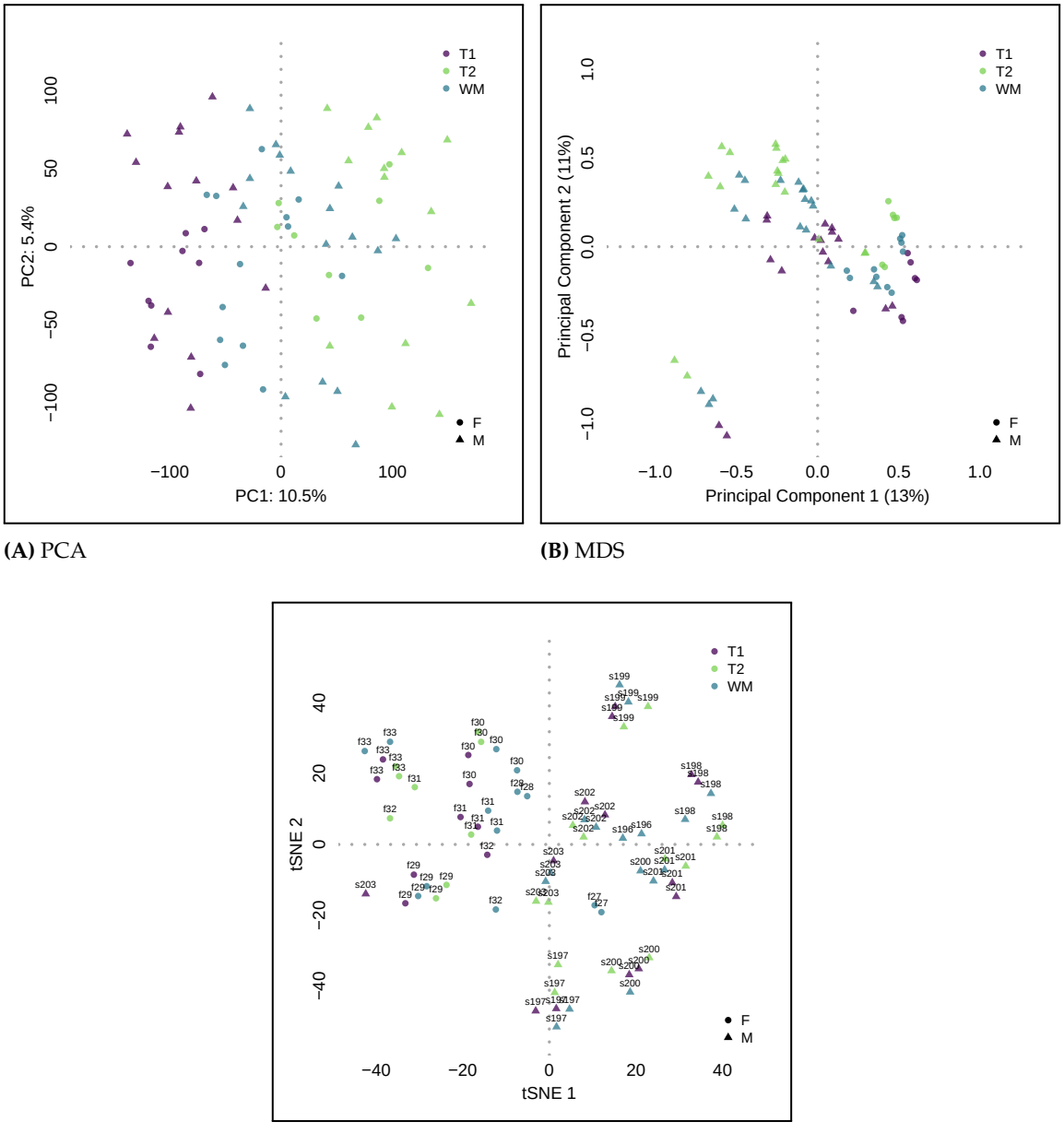
5.3.3 DNA methylation analysis revealed widespread differences between muscle fibres

Genome-wide DNAm analysis was performed on 838,705 DNAm sites probed using the Methylation EPICv2 Beadchip that passed quality control. A total of 74 samples were kept for differential methylation analysis (DMA). Four-week whole muscle samples were excluded from the dataset and replicate samples were averaged before DMA. There were very small mean global DNA methylation differences between TI, TII and WM samples.

A welches t test revealed a very small (0.2%) hypermethylation of TI samples compared to TII samples ($P < 0.001$), and WM samples overall methylation fell in between TI and TII fibres.

Dimension Reduction Analysis

PCA, multidimensional scaling (MDS) and t-distributed stochastic neighbor embedding (t-SNE) were used to assess variation in the data. PCA finds the projection of the data that explains the most variation in the data. PCA analysis showed that 10.5% of variation in the data was explained by the first PC and that samples clustered according to TI and TII fibres with WM dispersed through the middle (Figure 5.7A). MDS analysis aims to maintain some of the original relationship between data points, and showed that seven samples clustered together in the lower left quadrant of Figure 5.7B. Further investigation of the samples showed that six of these seven samples were from the same participant (s201) who had a unique ethnicity (Figure 5.7C). The final sample in this cluster was attributed to a different participant (s200) and can also be seen in Figure 5.7C clustering with participant s201. On further investigation Pearson's correlation analysis between s200-12wk and s201 and s200 revealed that s200-12wk had higher correlation values with s201 than s200 (Figure B.1). As a result s200-12wk was removed from downstream analysis and was deemed to be caused by a sample handling error. t-SNE dimension reduction aims to maintain original distances between samples as much as possible. t-SNE analysis shows a strong clustering by participant, as each participant provided multiple samples in the data (Figure 5.7C). Dimension reduction analysis indicated the need to adjust regression models by sex and participant and that ethnicity may be a source of variation within the data.



(A) PCA

(B) MDS

(C) tSNE

Figure 5.7: Dimension Reduction Plots on DNAm data

(A) PCA analysis shows samples separating according to fibre type along the first PC which explains 10.5% of the variation in the data. (B) MDS plots reveal that seven samples cluster away from the remaining samples. (C) tSNE indicates that samples from the same participant cluster together. T1 = Type I muscle fibres, T2 = Type II muscle fibres and WM = whole muscle samples. M = male and F = female.

Differential methylation of probes (DMPs) Type I vs Type II

To estimate differences in methylation of CpGs between TI and TII muscle fibres at baseline, mixed effects linear models were performed using limma in the R statistical environment on Dasen normalised M values.

A comparison of TI and TII samples at baseline (n = 12 vs 12) revealed 93,465 glspl{DMP} after Benjamani-Hochberg (BH) correction (FDR p < 0.005). After applying a more stringent FDR threshold of < 0.001, 68,026 CpGs were identified as differentially methylated between TI and TII fibres. DMA was conducted on the more statistically valid M values, however linear models were also applied to beta values to produce more interpretable beta value changes. Of the 67,587 DMPs passing FDR < 0.001, 11,468 showed a beta-change of greater than 10% (> 0.1 or < -0.1). Among the DMPs with greater than 10% beta-change 8,633 were hypermethylated and 2,835 were hypomethylated in TI fibres compared with TII fibres. Numbers of DMPs passing each BH-P-value are provided in Table 5.8.

The DMPs <0.001 with 10% beta value changes were predominantly located in the opensea CpG context with 78.4% located in CpG openseas, 11.8% in CpG shores, 7.4% in CpG shelves and 2.4% in CpG islands (Figure 5.9). DMPs were distributed along all autosomal chromosomes with the highest proportion of CpGs located on chromosome 1 (Figure 5.10). The list of all CpGs and statistics generated using limma can be provided on request. The top 50 hypomethylated and the top 50 hypermethylated DMPs in TI fibres compared with TII fibres are summarised in tables 5.9 and 5.10 respectively. In TI compared with TII fibres the top hypomethylated CpGs were annotated to PDLIM1, KALRN, and TNNI1 and the top hypermethylated CpGs were annotated to TNNT3, GREB1 and TPM1 (Figure 5.8).

Table 5.8: DMPs in TI-pre compared to TII-pre fibres

	TI-Pre	**Vs**	**TII-Pre**
BH P-value Threshold	Hyper		Hypo
BH P < 0.001 > 10 % BetaDiff	8633		2835
BH P < 0.001	48344		19682
BH P < 0.005	63059		30406
BH P < 0.05	99257		71772

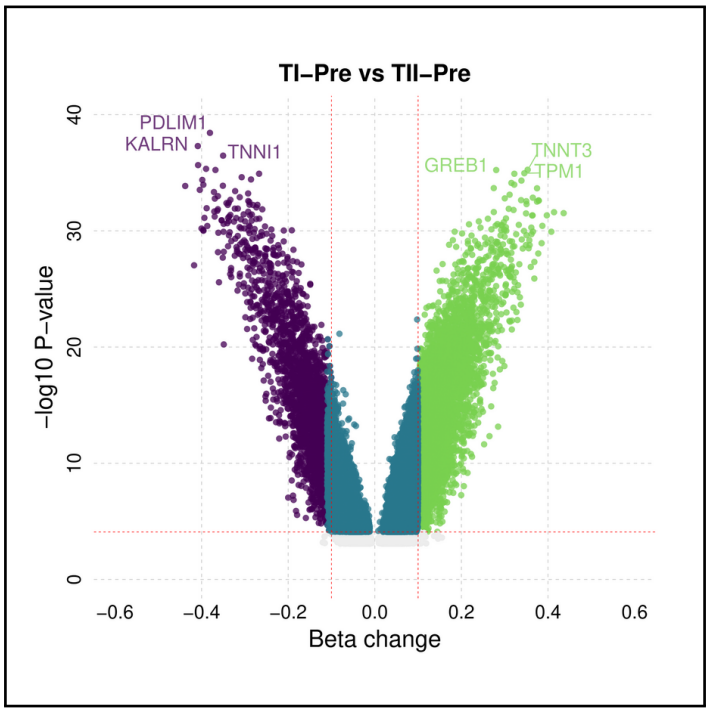


Figure 5.8: Volcano Plot T1vsT2

Volcano plot of the top 100,000 DMPs with hypomethylated (purple) and hypermethylated (green) genes identified in TI vs TII fibres. Beta value change is on the x-axis and -log10 P-value on the y-axis. The horizontal line represents the cutoff of the adjusted (P-value <0.001) and the vertical lines represent the cutoff of the beta change 10 %.

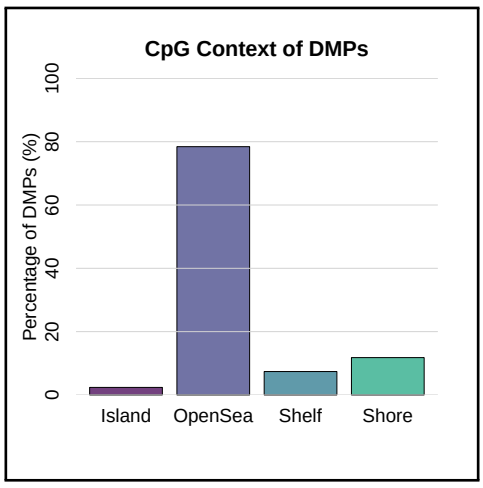


Figure 5.9: CpG Context

CpG context of DMPs with > 10% fold change. Dark purple represents Islands, light purble represents OpenSea, dark green represents Shelf and light green represent Shore CpG contexts.

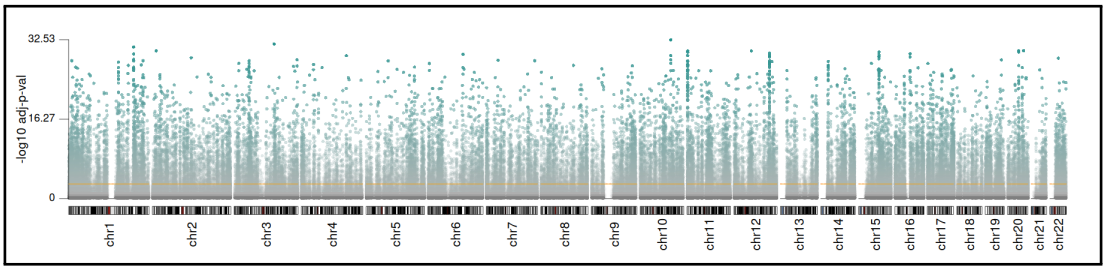


Figure 5.10: Manhatten Plot of DMPs

A Manhattan plot showing all DMPs across autosomes; the orange line represents an adj-P-value = 0.001.

Table 5.9: T1 vs T2 Top50 Hypo-DMPs

	GeneName	logFC	AveExp	t	p.value	adj.p.val	B	logFC_beta	t_beta	AveExpr_beta	Chromosome	ChromStat	CGI
cg08799394_TC21	PDLIM1	-2.38	-0.48	-26.01	3.70E-39	3.10E-33	77.41	-0.38	-25.65	0.43	chr10	TssFlnkU	Shore
cg01718380_BC21	KALRN	-2.69	-0.68	-25.01	5.04E-38	2.11E-32	74.98	-0.41	-23.54	0.40	chr3	Quies	Shelf
cg20562104_BC21	TNNI1	-2.33	0.78	-24.31	3.38E-37	9.46E-32	73.20	-0.35	-25.81	0.62	chr1	Quies	OpenSea
cg15791478_TC21		-2.63	-0.60	-23.63	2.20E-36	4.62E-31	71.44	-0.41	-23.43	0.41	chr20	EnhA1	OpenSea
cg23159950_BC21	MYH7B	-2.57	-0.75	-23.37	4.60E-36	6.11E-31	70.75	-0.39	-22.12	0.39	chr20	Quies	Shelf
cg02304314_BC21	FLJ12825	-2.25	0.22	-23.29	5.76E-36	6.11E-31	70.54	-0.37	-23.38	0.53	chr12	ReprPC	Shelf
cg25091827_BC21	MYH7B	-1.92	1.17	-23.02	1.21E-35	9.29E-31	69.83	-0.27	-24.07	0.68	chr20	Quies	Island
cg23708673_BC21	ATP2A2	-2.09	0.92	-22.78	2.43E-35	1.70E-30	69.18	-0.31	-24.01	0.64	chr12	Tx	OpenSea
cg07648215_BC21	MYLK3	-1.91	0.92	-22.63	3.75E-35	2.42E-30	68.77	-0.29	-22.05	0.64	chr16	Quies	OpenSea
cg13048191_BC21	TNNI1	-2.84	1.02	-22.59	4.16E-35	2.49E-30	68.67	-0.40	-24.15	0.65	chr1	Quies	OpenSea
cg07209719_BC21	PDLIM1	-2.50	-0.49	-22.27	1.08E-34	5.01E-30	67.77	-0.40	-21.58	0.42	chr10	TssFlnk	Shore
cg19755714_TC21	MAML3	-2.20	0.49	-22.20	1.31E-34	5.76E-30	67.59	-0.35	-22.69	0.58	chr4	EnhA1	OpenSea
cg18447962_TC21	ATP2A2	-2.79	0.29	-22.18	1.37E-34	5.76E-30	67.54	-0.44	-23.32	0.54	chr12	EnhWk	Shore
cg06570416_BC21	ATP2A2	-2.72	0.77	-21.92	2.97E-34	1.08E-29	66.81	-0.40	-23.71	0.61	chr12	Quies	Shore
cg03315837_BC21	MYLK3	-2.26	1.10	-21.84	3.80E-34	1.33E-29	66.58	-0.32	-23.37	0.67	chr16	Quies	OpenSea
cg03406283_TC21		-2.48	0.76	-21.78	4.57E-34	1.53E-29	66.40	-0.37	-21.50	0.61	chr2	ReprPCWk	OpenSea
cg00783213_BC21	TPST2	-1.80	0.55	-21.68	6.06E-34	1.95E-29	66.14	-0.29	-22.44	0.59	chr22	TxWk	OpenSea
cg13435855_BC21	TNNI1	-2.28	0.45	-21.50	1.03E-33	3.08E-29	65.63	-0.36	-22.29	0.57	chr1	Quies	OpenSea
cg15415313_BC21	PDLIM1	-2.19	1.13	-21.43	1.27E-33	3.66E-29	65.44	-0.31	-21.44	0.67	chr10	Quies	OpenSea
cg17799456_BC21	EEPD1	-2.46	-0.90	-21.34	1.67E-33	4.39E-29	65.17	-0.37	-20.76	0.36	chr7	EnhA1	OpenSea
cg08901951_BC21	MYL2	-2.01	0.07	-21.26	2.14E-33	5.44E-29	64.94	-0.33	-21.72	0.51	chr12	Quies	OpenSea
cg00313533_TC21	MYL3	-1.97	0.94	-21.18	2.69E-33	6.09E-29	64.72	-0.29	-21.76	0.65	chr3	Quies	OpenSea
cg11418783_BC21	LAD1	-1.84	0.64	-21.11	3.33E-33	6.98E-29	64.52	-0.29	-21.90	0.60	chr1	ReprPC	Shore
cg08240074_TC21	MYH7	-2.26	0.83	-21.07	3.82E-33	7.81E-29	64.39	-0.34	-22.99	0.63	chr14	Quies	OpenSea
cg19926753_BC21	VSIG10	-2.16	0.59	-21.05	4.08E-33	8.16E-29	64.32	-0.34	-21.74	0.59	chr12	Quies	Shelf
cg19509829_BC11	ATP2A2	-2.59	1.86	-20.94	5.59E-33	1.07E-28	64.02	-0.29	-20.37	0.76	chr12	Tx	OpenSea
cg15155209_TC21	SLC6A6	-1.93	-1.28	-20.93	5.78E-33	1.07E-28	63.99	-0.27	-19.99	0.30	chr3	EnhA1	OpenSea
cg16514487_BC21	ATP2A2	-2.39	1.16	-20.93	5.86E-33	1.07E-28	63.98	-0.34	-21.39	0.67	chr12	Quies	Shore
cg16289421_TC21	TPM3	-1.68	-0.20	-20.92	5.97E-33	1.07E-28	63.96	-0.28	-21.47	0.47	chr1	EnhA1	Shelf
cg03978573_BC21		-1.66	0.42	-20.77	9.54E-33	1.67E-28	63.52	-0.27	-21.70	0.57	chr12	EnhA1	OpenSea

Table 5.9: T1 vs T2 Top50 Hypo-DMPs(continued)

	GeneName	logFC	AveExp	t	p.value	adj.p.val	B	logFC_beta	t_beta	AveExpr_beta	Chromosome	ChromStat	CGI
cg14529469_BC21	MYL3	-2.53	0.64	-20.71	1.13E-32	1.93E-28	63.36	-0.39	-21.69	0.60	chr3	Quies	OpenSea
cg00100361_TC21	ATP2A2	-2.13	0.95	-20.69	1.21E-32	2.01E-28	63.29	-0.32	-20.16	0.65	chr12	TxWk	OpenSea
cg15957369_BC21	PDLIM1	-1.96	0.24	-20.69	1.22E-32	2.01E-28	63.28	-0.32	-21.03	0.54	chr10	Quies	OpenSea
cg13131270_BC21	RPA1	-2.38	0.77	-20.65	1.37E-32	2.17E-28	63.17	-0.36	-22.10	0.62	chr17	Tx	OpenSea
cg15412604_TC21	ATP2A2	-1.69	0.85	-20.64	1.41E-32	2.19E-28	63.15	-0.26	-20.99	0.64	chr12	Tx	OpenSea
cg06455211_TC21	TBC1D1	-2.23	-0.18	-20.55	1.87E-32	2.80E-28	62.87	-0.36	-20.70	0.47	chr4	Quies	OpenSea
cg18076458_BC21		-2.34	0.45	-20.50	2.16E-32	3.18E-28	62.74	-0.37	-21.27	0.57	chr12	EnhA1	OpenSea
cg23807570_TC21	ATP2A2	-1.93	0.93	-20.49	2.25E-32	3.26E-28	62.70	-0.29	-21.20	0.65	chr12	Tx	OpenSea
cg14659008_TC21		-2.18	1.24	-20.32	3.85E-32	5.13E-28	62.18	-0.30	-19.50	0.69	chr8	Quies	Shelf
cg06099173_TC21	MYH7	-1.77	0.57	-20.31	3.95E-32	5.17E-28	62.16	-0.28	-21.29	0.59	chr14	EnhA1	OpenSea
cg18600110_BC21	MYH7B	-2.39	1.16	-20.29	4.27E-32	5.50E-28	62.09	-0.33	-19.71	0.67	chr20	Quies	Shore
cg13081262_BC21	MYH7	-1.99	1.04	-20.26	4.55E-32	5.75E-28	62.02	-0.29	-21.44	0.66	chr14	Quies	OpenSea
cg10239861_BC21	GARNL3	-2.28	-0.09	-20.18	5.88E-32	6.96E-28	61.78	-0.37	-20.88	0.48	chr9	Quies	OpenSea
cg00533834_BC21	MYH7	-2.08	0.28	-20.18	5.89E-32	6.96E-28	61.78	-0.34	-20.88	0.54	chr14	Quies	OpenSea
cg13314400_TC11	ATP2A2	-2.84	1.08	-20.10	7.60E-32	8.81E-28	61.54	-0.39	-21.46	0.66	chr12	TssFlnk	Shelf
cg10557930_BC21	MYH7	-2.05	-0.08	-20.10	7.67E-32	8.81E-28	61.53	-0.34	-20.50	0.49	chr14	Quies	OpenSea
cg15439078_BC21	MYL3	-2.11	0.78	-20.09	7.82E-32	8.87E-28	61.51	-0.32	-21.50	0.62	chr3	Quies	OpenSea
cg22460923_BC21	ALPL	-1.93	-1.35	-20.05	8.87E-32	9.79E-28	61.39	-0.27	-20.12	0.29	chr1	EnhWk	OpenSea
cg26270038_BC21	MYH6	-1.74	0.75	-20.03	9.58E-32	1.03E-27	61.31	-0.27	-20.49	0.62	chr14	Quies	OpenSea
cg10123201_TC21	MYL3	-1.90	0.44	-20.00	1.03E-31	1.10E-27	61.24	-0.31	-20.83	0.57	chr3	Quies	OpenSea

Table 5.10: T1 vs T2 Top50 Hyper-DMPs

	GeneName	logFC	AveExp	t	p.value	adj.p.val	B	logFC_beta	t_beta	AveExpr_beta	Chromosome	ChromStat	CGI
cg14600987_BC21	TNNT3	2.29	-0.69	23.30	5.51E-36	6.11E-31	70.58	0.35	23.55	0.39	chr11	Quies	Shelf
cg22495058_BC21	GREB1	1.76	-0.59	23.28	5.83E-36	6.11E-31	70.53	0.28	23.76	0.40	chr2	Quies	Shore
cg25691442_TC21	TPM1	2.40	-1.02	23.06	1.10E-35	9.29E-31	69.93	0.35	21.89	0.35	chr15	EnhA1	OpenSea
cg19642007_BC21	TNNT3	2.07	0.78	23.02	1.22E-35	9.29E-31	69.83	0.32	23.54	0.62	chr11	Quies	Shelf
cg18378755_TC21	MICAL1;ZBTB24	2.13	0.62	22.55	4.75E-35	2.66E-30	68.55	0.34	23.57	0.60	chr6	Tx	OpenSea
cg21221455_BC21	TPM1;TPM1-AS	2.01	0.73	22.39	7.52E-35	3.94E-30	68.11	0.32	22.59	0.61	chr15	TssFlnkU	Shore
cg12038298_TC21	TNNT3	2.18	1.05	22.27	1.06E-34	5.01E-30	67.79	0.32	21.76	0.66	chr11	Quies	Shore
cg18176835_BC21	TLN2	1.63	-0.02	22.04	2.10E-34	8.26E-30	67.14	0.27	22.93	0.50	chr15	TxWk	OpenSea
cg20491601_BC21	TPM1	2.58	-0.95	22.03	2.17E-34	8.26E-30	67.11	0.37	22.01	0.36	chr15	EnhA1	Shelf
cg04578549_BC21	ENSG00000259727	1.93	0.52	21.60	7.72E-34	2.40E-29	65.91	0.31	21.62	0.58	chr15	Quies	OpenSea
cg04344361_TC21	XXYLT1	2.28	0.45	21.40	1.39E-33	3.89E-29	65.35	0.36	22.08	0.57	chr3	Quies	OpenSea
cg19387862_TC21	SPIB	1.80	0.01	21.36	1.60E-33	4.32E-29	65.22	0.30	21.68	0.50	chr19	ReprPC	Shore
cg01671575_BC21	PGBD5	2.05	0.23	21.24	2.28E-33	5.62E-29	64.88	0.34	21.48	0.54	chr1	Quies	OpenSea
cg06455375_BC21	TNNT3	2.40	0.48	21.23	2.37E-33	5.68E-29	64.84	0.38	21.74	0.57	chr11	Quies	Island
cg21108554_BC21	REPIN1-AS1	2.32	0.20	21.20	2.54E-33	5.93E-29	64.77	0.38	21.67	0.53	chr7	Quies	Shelf
cg20488341_TC21		2.13	0.15	21.17	2.80E-33	6.19E-29	64.68	0.35	21.90	0.52	chr1	Quies	OpenSea

Table 5.10: T1 vs T2 Top50 Hyper-DMPs(continued)

	GeneName	logFC	AveExp	t	p.value	adj.p.val	B	logFC_beta	t_beta	AveExpr_beta	Chromosome	ChromStat	CGI
cg12542203_TC21		2.55	0.94	21.14	3.03E-33	6.52E-29	64.61	0.37	21.41	0.64	chr5	Quies	OpenSea
cg05258834_BC21	TNNI2	2.33	0.82	20.98	4.99E-33	9.74E-29	64.13	0.35	21.64	0.62	chr11	TxWk	OpenSea
cg20273012_BC21	BMF	1.89	-0.35	20.68	1.26E-32	2.04E-28	63.25	0.31	21.16	0.44	chr15	EnhA1	OpenSea
cg01712812_BC21	LYRM4	2.30	-1.11	20.60	1.60E-32	2.44E-28	63.02	0.32	21.81	0.33	chr6	EnhA1	Shore
cg01821149_TC21	TNNT3	1.66	0.31	20.47	2.41E-32	3.43E-28	62.63	0.27	21.09	0.55	chr11	Quies	Island
cg13133304_BC11	TNNI2	3.04	1.05	20.46	2.48E-32	3.47E-28	62.60	0.41	23.45	0.64	chr11	TxWk	OpenSea
cg02556649_BC11	TNNT3	3.07	0.89	20.40	2.99E-32	4.12E-28	62.43	0.44	21.01	0.62	chr11	TxWk	Shore
cg01808320_BC21	CNIH3	2.33	0.86	20.38	3.15E-32	4.26E-28	62.38	0.35	21.52	0.63	chr1	Quies	OpenSea
cg18877271_TC11	TNNT3	2.54	0.68	20.26	4.65E-32	5.75E-28	62.01	0.40	20.19	0.60	chr11	Quies	Shore
cg09993645_BC21	NIBAN1	2.24	0.21	20.26	4.66E-32	5.75E-28	62.00	0.36	20.16	0.53	chr1	Quies	OpenSea
cg27651665_BC21	XXYLT1	2.27	0.48	20.25	4.73E-32	5.76E-28	61.99	0.36	20.59	0.58	chr3	Quies	OpenSea
cg02763146_BC21	EIF4G1;SNORD66	1.81	0.03	20.07	8.49E-32	9.49E-28	61.43	0.30	20.42	0.50	chr3	Tx	Shelf
cg04052013_TC21	TNNT3	1.64	0.73	20.05	9.05E-32	9.86E-28	61.37	0.26	20.38	0.62	chr11	Quies	Island
cg19501518_TC21	MICAL1;ZBTB24	1.43	0.30	19.88	1.53E-31	1.51E-27	60.87	0.24	20.77	0.55	chr6	Tx	OpenSea
cg10628683_BC21	TNNC2	1.81	-0.33	19.83	1.79E-31	1.72E-27	60.72	0.30	20.06	0.45	chr20	EnhBiv	Shore
cg21704222_BC21	ATXN1L	1.69	-0.02	19.71	2.60E-31	2.42E-27	60.36	0.28	20.34	0.50	chr16	EnhA2	OpenSea
cg03937290_TC21	ENSG00000237429	1.28	-0.11	19.70	2.72E-31	2.50E-27	60.32	0.22	20.82	0.48	chr1	TssFlnkU	Shore
cg25406657_TC21	TPM1;TPM1-AS	2.01	-0.01	19.66	3.05E-31	2.78E-27	60.21	0.33	19.86	0.50	chr15	TssFlnkU	Shore
cg06441714_BC21	CX3CR1	1.74	-0.01	19.65	3.19E-31	2.87E-27	60.17	0.29	20.00	0.50	chr3	Quies	OpenSea
cg18258887_BC21	MIR583HG	2.16	-1.27	19.64	3.23E-31	2.88E-27	60.15	0.29	18.96	0.31	chr5	Quies	OpenSea
cg20491284_BC21	TPM1	1.68	-0.31	19.61	3.56E-31	3.11E-27	60.06	0.28	19.66	0.45	chr15	TssBiv	Shore
cg18877271_TC12	TNNT3	2.54	-0.71	19.60	3.71E-31	3.21E-27	60.02	0.39	19.99	0.39	chr11	Quies	Shore
cg26371931_TC21	SLC16A5	2.00	0.89	19.52	4.81E-31	3.96E-27	59.77	0.30	19.08	0.64	chr17	TxWk	OpenSea
cg27412039_TC11	TNNT3	3.44	1.83	19.49	5.24E-31	4.24E-27	59.69	0.36	21.93	0.73	chr11	Quies	Island
cg09578829_TC21	TBC1D1	1.45	-0.18	19.49	5.26E-31	4.24E-27	59.68	0.24	20.19	0.47	chr4	Quies	OpenSea
cg09419354_BC21	LYRM4	2.08	-1.38	19.38	7.45E-31	5.73E-27	59.35	0.27	20.03	0.29	chr6	EnhA1	Shore
cg00557924_TC21	DRAIC	2.01	1.04	19.38	7.63E-31	5.76E-27	59.33	0.30	18.14	0.66	chr15	Quies	OpenSea
cg03418649_TC21	TNNT3	2.00	1.07	19.34	8.51E-31	6.32E-27	59.22	0.29	18.51	0.67	chr11	Quies	Shore
cg01667636_TC21	WBP1L	1.80	0.16	19.31	9.56E-31	6.85E-27	59.11	0.30	19.63	0.53	chr10	Tx	OpenSea
cg16205293_BC21		1.84	0.19	19.30	9.85E-31	6.94E-27	59.08	0.30	19.56	0.53	chr11	Quies	OpenSea
cg10423149_BC21	PTPRG	2.12	0.34	19.29	1.02E-30	7.11E-27	59.05	0.34	19.62	0.55	chr3	Quies	OpenSea
cg04563870_BC21	TNNT3	2.25	0.22	19.25	1.16E-30	8.05E-27	58.92	0.37	19.80	0.53	chr11	Quies	OpenSea
cg25171685_BC21	TPM1	1.76	-2.11	19.24	1.18E-30	8.09E-27	58.91	0.18	17.59	0.20	chr15	TssBiv	Island
cg17313218_BC21	LSP1	2.57	0.36	19.24	1.19E-30	8.09E-27	58.91	0.41	19.83	0.55	chr11	Quies	OpenSea

Gene Ontology, Kegg and Reactome overrepresentation analysis

Gene ontology overrepresentation analysis (GO-ORA) was conducted to assess the relationship between DMPs. Cluster profiler was used for GO-ORA using the 11,491 significant DMPs with > 10% beta value changes. 274 enriched terms were identified with a BH-P-value cutoff of < 0.005, with 175 of these terms passing a stringent < 0.001 BH-P-value cutoff. There were significant overlaps between DMPs and muscle related biological process (BP) (121 enriched terms < 0.001), cellular component (CC) (34 enriched terms <0.001) and molecular function (MF) (20 enriched terms <0.001) gene ontology (GO) terms (Figures B.6 & B.7 & B.8). As expected between TI and TII fibres the main enriched gene sets for BP, CC and MF are involved in actin filament organisation, actin cytoskeleton and actin binding. Additionally GTPase regulator activity and GTPase mediated signal transduction were overrepresented in TI vs TII DMPs. This was further confirmed by Reactome pathway overrepresentation analysis with the top five pathways relating to muscle contraction, RHO GTPase cycle pathway and cardiac conduction (Figures B.9. A total of 19 Reactome pathways were identified with a P-value cutoff of <0.005 of which 14 pathways passed a P-value cutoff <0.001. A total of 25 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways with a P-value cutoff <0.005 were enriched, with 20 pathways passing a P-value cutoff <0.001 (Figure B.10). Focal adhesion, regulation of actin cytoskeleton and motor proteins were some of the top enriched KEGG pathways.

A consideration with ORA analysis in methylation studies is the bias introduced by multiple CpGs annotated to the same gene and CpGs annotated to multiple genes [308]. I modified code from the missMethyl package for the new EPICv2 to conduct bias corrected ORA on GO, KEGG and reactome terms [287]. I identified a reduced number of GO terms after adjusting for bias, with five BP, six CC and one MF GO terms with BH adj-P-vals < 0.001. These terms were more specific to muscle related processes and GTPase activity was no longer present (Figure 5.11). Likewise, there were a reduction in the number of reactome terms identified after bias adjustment, with muscle contraction and striated muscle contraction reactome terms passing a BH-adj-P-value < 0.001 (Figure 5.12). No KEGG pathways were identified with a BH-adj-P-value < 0.005, with eight pathways with a BH-adj-P-val < 0.05, with focal adhesion and c-type lectin signalling the top two KEGG pathways (Figure 5.13).

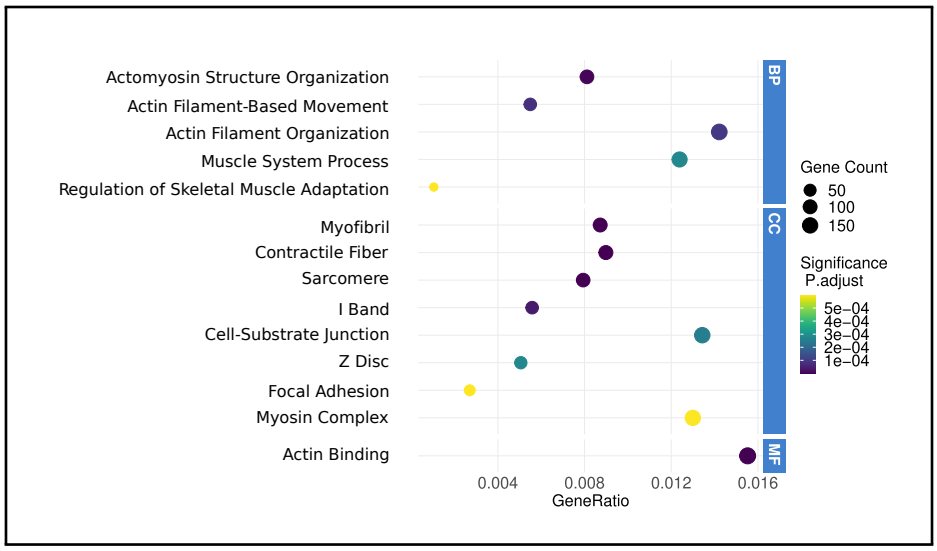


Figure 5.11: ORA of Gene Ontology Terms from TI vs TII DMPs

Gene ontology overrepresentation analysis of the differentially methylated DMPs between TI and TII fibres as calculated by the missMetyl method. Figure 5.11 shows the top GO terms from BP, CC and MF as indicated by the solid blue bar. The circle size represents gene numbers, and the colour represents the adj-P-value.

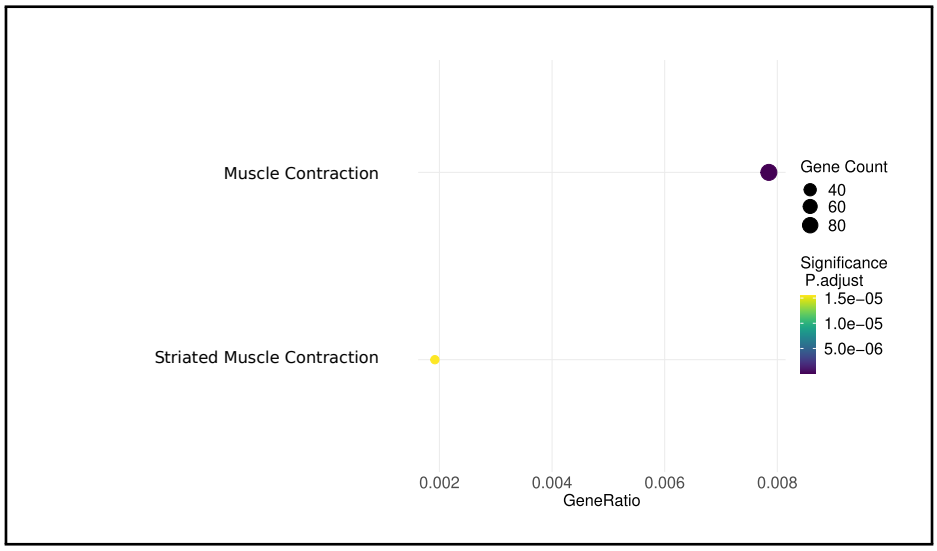


Figure 5.12: ORA of Reactome Terms from TI vs TII DMPs

Reactome term overrepresentation analysis of the differentially methylated DMPs between TI and TII fibres as calculated by the missMetyl method. The circle size represents gene numbers, and the colour represents the adj-P-value.

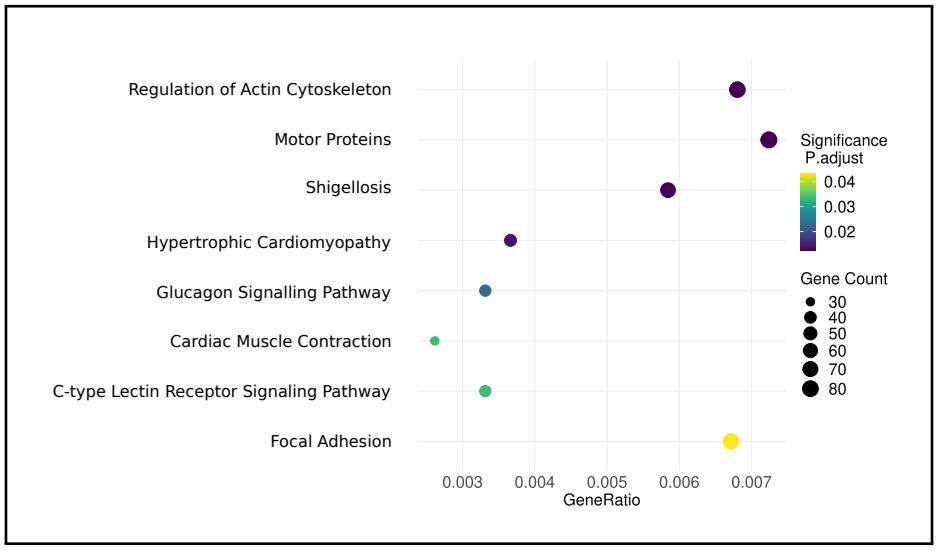


Figure 5.13: ORA of KEGG Pathways from TI vs TII DMPs

Kegg Pathways overrepresentation analysis of the differentially methylated DMPs between TI and TII fibres as calculated by the missMetyl method. The circle size represents gene numbers, and the colour represents the adj-P-value.

Mitochondrial Genes DMPs

A major difference between muscle fibre types are mitochondrial content [309] and evidence suggests that mitochondrial specialisation between fibre types also results in functional differences within mitochondrial processes [309]. To determine differential methylation in mitochondrial genes between TI and TII fibres, the crossover of DMPs with MitoCarta 3.0 [310] mitochondrial genes was assessed. 242 DMPs out of 11,469 DMPs (< 0.001 and > 10% beta difference) were annotated to genes in the MitoCarta gene list (Figure 5.14) with 133 unique MitoCarta genes represented. The top 25 hypomethylated and hypermethylated mitochondrial DMPs (mitoDMPs) are summarised in tables 5.11 and 5.12 respectively. The full set of mitoDMPs can be provided on request. ORA of MitoCarta pathways revealed no enriched terms from the mitoDMPs and GO-ORA of the 242 mitoDMPs revealed no enriched terms. Heatmap clustering showed a distinct pattern in the top 25 hypomethylated and hypermethylated mitoDMPs (Figure 5.15). Multiple top TI verses TII hypermethylated mitoDMPs were annotated to the LYRM4 and CDK5RAP1 genes (Figure 5.14, and table 5.12).

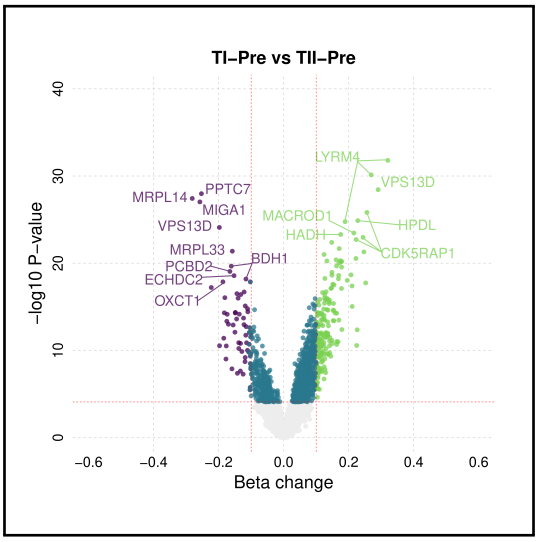


Figure 5.14: Volcano plot of mitochondrial DMPs in TI vs TII fibres

Volcano plot of the mitoDMPs with hypomethylated (purple) and hypermethylated (green) genes identified in TI vs TII fibres. Beta value change is on the x-axis and -log10 P-value on the y-axis. The horizontal line represents the cutoff of the adjusted (P-value <0.001) and the vertical lines represent the cutoff of the beta change 10 %.

Table 5.11: T1 vs T2 Top25 Hypo-mitoDMPs

	GeneName	logFC	AveExp	t	p.value	adj.p.val	B	logFC_beta	t_beta	AveExpr_beta	Chromosome	ChromStat	CGI
cg18081586_BC21	PPTC7	-1.72	1.11	-17.88	1.06E-28	3.85E-25	54.58	-0.25	-18.75	0.67	chr12	EnhWk	OpenSea
cg26371345_BC21	MRPL14	-1.69	0.05	-17.51	3.74E-28	1.22E-24	53.37	-0.28	-17.67	0.51	chr6	EnhWk	Shelf
cg27071127_BC21	MIGA1	-1.61	0.70	-17.25	9.31E-28	2.71E-24	52.49	-0.26	-17.36	0.61	chr1	TxWk	Shelf
cg17494034_TC21	VPS13D	-1.19	-0.46	-15.37	7.92E-25	1.22E-21	45.95	-0.20	-15.80	0.42	chr1	EnhA1	OpenSea
cg05938007_TC21	MRPL33	-0.95	-0.47	-13.73	4.17E-22	3.58E-19	39.84	-0.16	-14.08	0.42	chr2	TssFlnkU	Shore
cg06063407_TC21	BDH1	-1.19	1.49	-12.73	2.16E-20	1.32E-17	35.97	-0.16	-12.19	0.73	chr3	Quies	OpenSea
cg13081156_TC21	PCBD2	-1.19	1.36	-12.39	8.78E-20	4.75E-17	34.60	-0.17	-12.21	0.71	chr5	Quies	OpenSea
cg17471757_BC21	ECHDC2	-0.89	0.11	-12.11	2.68E-19	1.30E-16	33.51	-0.15	-12.50	0.52	chr1	Quies	OpenSea
cg09177982_TC21	BDH1	-0.69	-0.48	-11.90	6.51E-19	2.93E-16	32.64	-0.12	-12.56	0.42	chr3	TxWk	OpenSea
cg11961656_BC21	OXCT1	-1.22	-0.95	-11.72	1.36E-18	5.73E-16	31.91	-0.19	-11.66	0.35	chr5	Quies	OpenSea
cg23388348_TC21	MRPL14	-0.99	2.23	-11.72	1.38E-18	5.80E-16	31.90	-0.10	-11.13	0.82	chr6	TxWk	OpenSea
cg02428556_BC21	HSD17B4	-1.42	0.76	-11.36	6.14E-18	2.25E-15	30.44	-0.22	-11.02	0.62	chr5	TxWk	OpenSea
cg10521049_BC21	COX19	-0.85	-1.26	-11.08	1.99E-17	6.49E-15	29.28	-0.12	-11.48	0.30	chr7	EnhA1	Shelf
cg24207616_TC21	OGDH	-0.85	-0.18	-10.98	3.09E-17	9.65E-15	28.85	-0.14	-11.18	0.47	chr7	Tx	OpenSea
cg16397905_TC21	MECR	-0.80	0.61	-10.92	3.91E-17	1.19E-14	28.62	-0.13	-11.36	0.60	chr1	TxWk	OpenSea
cg03098228_TC21	IMMT	-1.07	-0.23	-10.72	9.00E-17	2.55E-14	27.80	-0.18	-10.82	0.46	chr2	TxWk	OpenSea
cg12577755_BC21	MRPS30	-0.90	0.90	-10.69	1.03E-16	2.86E-14	27.67	-0.14	-10.62	0.65	chr5	Tx	OpenSea
cg14557954_TC21	PTCD3	-0.74	-0.79	-10.22	7.96E-16	1.80E-13	25.66	-0.12	-10.32	0.37	chr2	Tx	OpenSea
cg06411730_BC21	ELAC2	-0.75	-1.11	-10.04	1.67E-15	3.50E-13	24.93	-0.11	-10.23	0.32	chr17	Tx	OpenSea
cg07717466_BC21	VWA8	-0.65	0.29	-9.86	3.72E-15	7.15E-13	24.14	-0.11	-10.28	0.55	chr13	Quies	OpenSea
cg24122311_BC21	ACAT1	-0.90	-0.48	-9.82	4.38E-15	8.29E-13	23.98	-0.15	-9.73	0.42	chr11	EnhWk	Shelf

Table 5.11: T1 vs T2 Top25 Hypo-mitoDMPs(continued)

	GeneName	logFC	AveExp	t	p.value	adj.p.val	B	logFC_beta	t_beta	AveExpr_beta	Chromosome	ChromStat	CGI
cg21817750_BC21	MRPS27	-1.14	0.67	-9.79	4.90E-15	9.18E-13	23.87	-0.18	-9.33	0.61	chr5	TxWk	OpenSea
cg12719649_BC21	SND1	-0.89	0.36	-9.79	5.07E-15	9.45E-13	23.84	-0.15	-9.94	0.56	chr7	Quies	OpenSea
cg09717402_TC21	LAP3	-1.15	1.86	-9.71	6.99E-15	1.26E-12	23.52	-0.14	-9.15	0.78	chr4	Tx	OpenSea
cg21817750_BC23	MRPS27	-1.07	0.63	-9.71	7.15E-15	1.29E-12	23.50	-0.17	-9.36	0.60	chr5	TxWk	OpenSea

Table 5.12: T1 vs T2 Top25 Hyper-mitoDMPs

	GeneName	logFC	AveExp	t	p.value	adj.p.val	B	logFC_beta	t_beta	AveExpr_beta	Chromosome	ChromStat	CGI
cg01712812_BC21	LYRM4	2.30	-1.11	20.60	1.60E-32	2.44E-28	63.02	0.32	21.81	0.33	chr6	EnhA1	Shore
cg09419354_BC21	LYRM4	2.08	-1.38	19.38	7.45E-31	5.73E-27	59.35	0.27	20.03	0.29	chr6	EnhA1	Shore
cg00246695_BC21	VPS13D	1.75	0.25	18.20	3.69E-29	1.56E-25	55.60	0.29	18.55	0.54	chr1	EnhA1	OpenSea
cg25168947_BC21	CDK5RAP1	1.53	0.17	16.46	1.51E-26	3.33E-23	49.79	0.26	16.75	0.53	chr20	Tx	OpenSea
cg03270777_BC21	HPDL	1.58	1.16	15.86	1.32E-25	2.41E-22	47.69	0.23	15.18	0.68	chr1	TxWk	Shore
cg04360434_BC21	LYRM4	1.63	-1.77	15.79	1.70E-25	2.99E-22	47.44	0.19	15.64	0.24	chr6	EnhA1	Shore
cg16443197_BC21	MACROD1	1.27	-0.04	14.98	3.45E-24	4.59E-21	44.51	0.22	15.29	0.49	chr11	ReprPC	Shelf
cg05599982_BC21	HADH	1.03	-0.01	14.87	5.05E-24	6.51E-21	44.14	0.18	15.51	0.50	chr4	TxWk	OpenSea
cg25168946_BC21	CDK5RAP1	1.45	0.19	14.67	1.09E-23	1.31E-20	43.40	0.24	14.85	0.53	chr20	Tx	OpenSea
cg25168944_TC21	CDK5RAP1	1.32	0.07	14.52	1.94E-23	2.21E-20	42.83	0.22	14.76	0.51	chr20	Tx	OpenSea
cg17373573_TC21	QDPR	0.96	-0.85	14.32	4.14E-23	4.42E-20	42.09	0.15	14.49	0.36	chr4	TxWk	OpenSea
cg05756685_TC21	VPS13D	1.02	-0.38	13.90	2.09E-22	1.92E-19	40.51	0.17	14.21	0.44	chr1	Quies	OpenSea
cg13609939_BC21	OAT	1.48	0.27	13.68	5.03E-22	4.22E-19	39.65	0.25	13.56	0.54	chr10	TxWk	Shore
cg00705568_BC21	ALDH1L1	0.72	-0.09	13.50	1.01E-21	8.00E-19	38.98	0.12	14.70	0.48	chr3	Quies	OpenSea
cg15922821_TC21	NT5DC3	1.51	1.18	13.24	2.87E-21	2.09E-18	37.95	0.22	12.81	0.69	chr12	Quies	OpenSea
cg05185262_BC21	CYB5R3	1.04	0.98	13.15	4.10E-21	2.91E-18	37.60	0.16	13.31	0.66	chr22	Quies	Shelf
cg17823326_TC21	NUBPL	1.03	-0.20	13.09	5.12E-21	3.56E-18	37.38	0.17	13.45	0.47	chr14	Quies	OpenSea
cg20341619_TC21	GATM	0.82	-0.58	13.07	5.56E-21	3.84E-18	37.30	0.13	13.47	0.40	chr15	ReprPCWk	OpenSea
cg05049036_BC21	SLC25A26	1.08	0.54	13.04	6.24E-21	4.23E-18	37.19	0.18	13.19	0.59	chr3	Quies	OpenSea
cg01282639_BC21	BCL2	1.17	1.02	13.01	7.15E-21	4.80E-18	37.06	0.18	13.44	0.66	chr18	TxWk	OpenSea
cg02434880_TC21	C6orf136	1.00	0.07	12.94	9.28E-21	6.08E-18	36.80	0.17	13.27	0.51	chr6	Tx	OpenSea
cg22183331_BC21	SCO1	1.03	0.83	12.74	2.09E-20	1.28E-17	36.01	0.16	12.90	0.64	chr17	TxWk	OpenSea
cg04097641_BC21	NDUFA10	1.04	-0.08	12.67	2.77E-20	1.65E-17	35.73	0.18	12.89	0.49	chr2	ZNF/Rpts	OpenSea
cg05196406_BC21	ALDH4A1	0.98	0.97	12.33	1.10E-19	5.84E-17	34.38	0.15	12.24	0.66	chr1	Quies	OpenSea
cg02039987_BC21	GPD2	1.29	-0.55	12.17	2.17E-19	1.08E-16	33.71	0.21	12.47	0.41	chr2	Quies	Shore

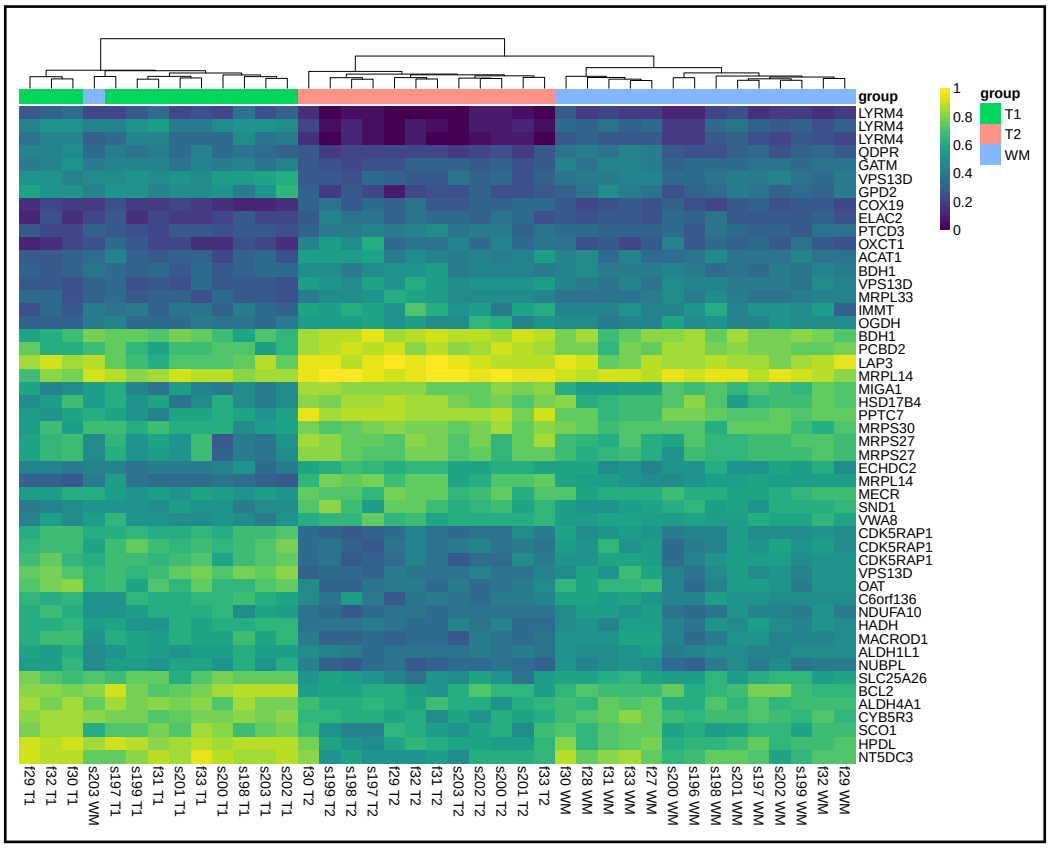


Figure 5.15: Heatmap of top 50 mitoDMPs

Heat map hierarchical clustering analysis of the top 25 hypomethylated and top 25 hypermethylated mitoDMPs. The color-coded scale (purple = hypomethylation and yellow = hypermethylation) are the min-max normalised beta-values. Hierarchical clustering of rows and columns represents beta values of the CpG sites and samples. Type I (green), Type II (red) and Whole Muscle (blue) samples are colour coded and indicated with appropriate bars.

Differential methylation of regions (DMRs) Type I verses Type II

To determine differences in methylation across proximal positions in the genome, I investigated DMRs in TI compared with TII muscle fibres using the DMRcate package in R. A region was defined as a section of the genome containing a minimum of four proximal CpGs. A total of 5778 DMRs with a Fisher's multiple comparison statistic <0.001 were identified using the 67,587 DMPs as probe inputs. The 5778 DMRs were annotated to 3996 unique overlapping genes, with 4203 DMRs showing a mean hypermethylation and 1586 showing a mean hypomethylation in TI compared with TII fibres. The top 25 hypo-DMRs and top 25 hyper-DMRs are summarised in tables 5.13 and 5.14 respectively. DMRs were present across all autosomal chromosomes with the most DMRs located on chromosome 1 followed by chromosome 17 5.16.

Table 5.13: T1 vs T2 Top50 Hypo-DMRs

chromosome	start	end	width	strand	no.cpgs	minSmoothedFdr	Stouffer	HMFDR	Fisher	maxdiff	meandiff	Overlapping Genes
chr14	23431788	23436874	5087	*	14	0.00E+00	3.82E-281	7.54E-28	1.21E-279	-2.26	-1.53	MYH7
chr16	46747163	46749105	1943	*	12	0.00E+00	2.34E-162	2.90E-29	1.81E-176	-1.91	-0.98	MYLK3
chr1	201420503	201422973	2471	*	13	0.00E+00	2.18E-138	3.84E-28	9.25E-174	-2.64	-1.30	TNNI1
chr1	201398760	201400521	1762	*	13	0.00E+00	2.66E-139	9.07E-28	5.46E-149	-1.84	-0.87	"AC119427.1, LAD1"
chr3	46863152	46864894	1743	*	7	0.00E+00	2.74E-136	9.59E-28	1.39E-145	-2.53	-1.75	MYL3
chr20	34979500	34980083	584	*	6	0.00E+00	6.14E-144	5.57E-30	1.44E-140	-2.70	-2.06	MYH7B
chr1	154178123	154180320	2198	*	7	0.00E+00	1.87E-138	6.94E-28	3.19E-138	-1.96	-1.23	TPM3
chr1	154191577	154194960	3384	*	13	9.09E-274	1.95E-90	1.15E-25	1.53E-125	-1.89	-0.69	"TPM3, MIR190B"
chr13	20720919	20722837	1919	*	10	1.91E-174	8.95E-113	6.66E-22	5.07E-118	-1.94	-1.21	IL17D
chr10	95288894	95291855	2962	*	16	0.00E+00	4.02E-53	4.96E-32	1.30E-104	-2.50	-0.55	PDLIM1
chr1	118988308	118993440	5133	*	31	3.76E-138	1.30E-91	1.67E-09	7.47E-103	-0.60	-0.37	TBX15
chr12	110279415	110281235	1821	*	13	0.00E+00	1.37E-40	4.72E-29	8.55E-100	-2.79	-0.86	ATP2A2
chr2	181984881	181985972	1092	*	8	6.76E-279	8.29E-82	1.22E-24	9.07E-99	-1.95	-1.01	PPP1R1C
chr12	110919343	110921157	1815	*	14	1.79E-213	6.55E-50	7.61E-28	1.70E-94	-2.01	-0.51	MYL2
chr12	110934344	110937651	3308	*	11	4.26E-181	3.93E-71	4.05E-23	2.14E-93	-1.78	-0.80	LINC01405
chr12	71440057	71443394	3338	*	11	1.42E-188	5.71E-54	2.60E-25	4.93E-92	-1.92	-0.78	"LGR5, TSPAN8"
chr14	23407614	23410104	2491	*	9	4.31E-194	1.11E-70	9.27E-27	6.31E-87	-1.74	-0.65	MYH6
chr12	110353384	110354893	1510	*	5	7.34E-181	2.81E-87	1.49E-24	2.68E-86	-1.80	-1.36	
chr16	66918237	66920369	2133	*	12	2.21E-209	4.66E-55	2.30E-21	7.86E-84	-1.19	-0.53	CDH16
chr3	46833406	46835071	1666	*	15	5.68E-212	1.00E-55	1.69E-19	2.13E-83	-1.92	-0.71	PRSS42P
chr12	110326382	110327493	1112	*	5	1.71E-272	1.55E-67	8.42E-30	8.86E-83	-2.09	-1.31	ATP2A2
chr6	31621815	31625086	3272	*	33	6.04E-256	1.11E-32	5.97E-20	2.62E-81	-1.64	-0.16	"PRRC2A, SNORA38"
chr1	170664731	170667427	2697	*	11	4.37E-140	2.66E-61	1.50E-17	1.32E-78	-1.77	-0.82	"PRRX1, Z97200.1"
chr4	110628723	110633228	4506	*	19	2.01E-60	4.06E-81	3.23E-10	1.41E-78	-0.78	-0.49	PITX2
chr2	222299739	222305245	5507	*	34	1.64E-60	3.03E-77	1.45E-07	6.37E-78	-0.73	-0.35	CCDC140
chr1	230824734	230825863	1130	*	5	1.04E-158	4.45E-63	1.95E-22	6.51E-77	-1.51	-1.14	

Table 5.13: T1 vs T2 Top50 Hypo-DMRs(continued)

chromosome	start	end	width	strand	no.cpgs	minSmoothedFdr	Stouffer	HMFDR	Fisher	maxdiff	meandiff	Overlapping Genes
chr12	106138433	106141645	3213	*	19	3.57E-147	5.45E-51	7.76E-17	6.75E-77	-1.44	-0.57	NUAK1
chr14	23414073	23415334	1262	*	4	5.82E-143	1.01E-77	2.78E-27	7.81E-77	-2.08	-1.39	MYH7
chr12	110246957	110248326	1370	*	7	1.28E-150	7.35E-64	2.22E-27	1.40E-76	-2.34	-1.12	
chr12	12520090	12521432	1343	*	8	6.50E-202	2.47E-54	3.13E-25	9.95E-76	-1.86	-0.78	DUSP16
chr12	122871014	122873209	2196	*	7	5.31E-112	1.17E-62	2.87E-22	1.82E-74	-1.78	-0.92	"VPS37B, AC027290.2"
chr16	66921647	66922338	692	*	6	6.85E-174	1.69E-72	5.35E-23	2.55E-74	-1.55	-0.99	RRAD
chr12	110344837	110346373	1537	*	6	1.51E-190	2.11E-58	5.44E-23	1.05E-73	-1.43	-0.85	ATP2A2
chr3	147404527	147408975	4449	*	38	1.02E-67	2.21E-62	3.46E-06	1.56E-72	-0.51	-0.28	"ZIC1, ZIC4"
chr6	33276974	33280936	3963	*	69	4.87E-86	2.52E-48	8.34E-06	2.53E-71	0.59	0.00	"B3GALT4, WDR46"
chr12	110348924	110351381	2458	*	6	3.91E-150	2.25E-50	1.61E-27	6.98E-69	-2.40	-0.95	ATP2A2
chr20	57711334	57715281	3948	*	20	2.05E-191	5.12E-46	2.44E-24	7.63E-69	-1.72	-0.09	"NKILA, PMEPA1"
chr1	183304970	183306856	1887	*	8	1.13E-161	3.61E-49	2.26E-25	5.28E-67	-1.90	-0.78	NMNAT2
chr17	7435570	7440214	4645	*	25	2.38E-93	1.65E-56	3.54E-10	2.30E-65	-0.99	-0.39	AC113189.1-4, AC113189.12, TMEM102, FGF11
chr10	101235682	101239004	3323	*	24	1.63E-52	3.57E-71	1.83E-06	3.44E-65	-0.53	-0.35	LBX1-AS1
chr13	109281156	109281939	784	*	6	6.80E-147	3.83E-59	2.49E-19	3.89E-65	-2.02	-1.09	
chr12	54052234	54054984	2751	*	28	2.60E-89	5.05E-57	4.15E-10	4.03E-64	-1.00	-0.31	HOXC4
chr15	60370288	60373176	2889	*	7	2.77E-118	3.02E-56	1.58E-21	2.79E-63	-1.59	-0.80	ANXA2
chr16	1531013	1534514	3502	*	20	2.05E-56	1.49E-59	6.94E-09	3.74E-62	-0.74	-0.37	"TMEM204, AL031719.1, IFT140"
chr3	45007277	45008115	839	*	4	1.39E-155	5.62E-49	1.41E-26	3.84E-60	-1.73	-1.26	"EXOSC7, CLEC3B"
chr3	147422336	147424627	2292	*	18	2.10E-82	6.72E-57	1.29E-06	8.69E-60	-0.57	-0.39	ZIC1
chr21	46219052	46220306	1255	*	7	8.93E-135	2.16E-34	2.09E-20	2.31E-58	-1.61	-0.66	"AP001469.1, LSS"
chr21	34122881	34123943	1063	*	5	4.78E-177	7.71E-34	1.22E-24	9.61E-57	-2.18	-1.13	"MRPS6, AP000317.2"
chr13	110507216	110508928	1713	*	6	1.65E-154	1.52E-38	2.24E-20	8.75E-56	-2.62	-1.07	"COL4A2, COL4A2-AS1"
chr6	100453494	100458506	5013	*	23	3.64E-60	1.03E-58	1.61E-05	2.39E-55	-0.53	-0.36	SIM1

Table 5.14: T1 vs T2 Top50 Hyper-DMRs

chromosome	start	end	width	strand	no.cpgs	minSmoothedFdr	Stouffer	HMFDR	Fisher	maxdiff	meandiff	Overlapping Genes
chr11	1837236	1841246	4011	*	20	0.00E+00	1.52E-288	1.51E-27	0.00E+00	3.04	1.49	"SYT8, TNNI2"
chr17	82229594	82240021	10428	*	52	2.66E-181	9.42E-279	1.27E-13	8.24E-296	1.40	0.66	"SLC16A3, MIR6787, CSNK1D"
chr11	1917703	1921783	4081	*	23	0.00E+00	1.31E-190	1.72E-25	2.55E-248	2.49	1.03	TNNT3
chr5	177088531	177098353	9823	*	48	9.43E-141	8.65E-226	3.56E-12	9.10E-234	1.02	0.50	FGFR4
chr11	1930583	1935061	4479	*	14	0.00E+00	3.40E-183	7.61E-30	2.42E-205	2.54	1.32	TNNT3
chr4	1207526	1212361	4836	*	34	0.00E+00	2.18E-193	2.22E-12	2.29E-203	1.54	0.72	"CTBP1-AS, SPON2, CTBP1"
chr15	63041276	63043102	1827	*	15	0.00E+00	6.00E-101	2.57E-26	1.68E-169	2.29	0.97	TPM1
chr22	46101822	46114659	12838	*	52	2.50E-85	1.45E-149	6.23E-12	9.20E-167	1.09	0.38	"MIRLET7BHG, MIRLET7A3, MIR4763, MIRLET7B"
chr11	1926370	1927899	1530	*	9	1.09E-240	3.14E-143	4.34E-28	1.40E-145	3.44	1.75	TNNT3
chr13	114108812	114113357	4546	*	16	5.50E-161	9.12E-142	5.71E-17	2.88E-140	1.72	0.87	"RASA3, RASA3-IT1"
chr1	109621924	109626771	4848	*	23	1.96E-144	4.15E-98	2.91E-13	2.12E-125	0.83	0.42	AMPD2
chr11	1936263	1938862	2600	*	6	4.34E-214	2.85E-119	5.91E-27	4.70E-117	1.64	1.39	TNNT3

Table 5.14: T1 vs T2 Top50 Hyper-DMRs(continued)

chromosome	start	end	width	strand	no.cpgs	minSmoothedFdr	Stouffer	HMFDR	Fisher	maxdiff	meandiff	Overlapping Genes
chr6	109465874	109467424	1551	*	7	0.00E+00	2.78E-101	1.86E-29	1.16E-115	2.13	1.01	"MICAL1, ZBTB24"
chr15	63046948	63050132	3185	*	20	0.00E+00	5.03E-49	7.87E-29	4.62E-115	2.01	0.52	"TPM1, TPM1-AS"
chr11	1924333	1924737	405	*	5	0.00E+00	2.41E-113	1.93E-27	7.41E-111	3.07	2.22	TNNT3
chr12	6373650	6377946	4297	*	24	5.36E-229	1.02E-69	8.22E-24	7.58E-110	1.87	0.49	"LTBR, SCNN1A"
chr11	44603036	44605398	2363	*	12	1.39E-242	8.00E-84	5.73E-23	2.61E-105	1.27	0.63	"CD82, AC010768.1"
chr17	82106561	82109269	2709	*	12	2.10E-176	2.24E-70	5.70E-25	3.78E-104	1.67	0.72	CCDC57
chr2	231392269	231396526	4258	*	21	6.44E-137	1.44E-70	7.75E-21	7.10E-103	-1.43	0.10	"AC017104.1, B3GNT7"
chr20	45826733	45827844	1112	*	7	9.00E-224	1.32E-93	1.29E-22	2.58E-97	1.88	1.14	TNNC2
chr19	15263531	15266054	2524	*	15	2.93E-173	3.77E-78	3.03E-17	3.00E-94	1.01	0.48	BRD4
chr12	4443862	4446982	3121	*	15	2.60E-136	1.03E-94	8.43E-13	3.10E-94	1.06	0.66	FGF6
chr15	62843130	62845480	2351	*	10	3.85E-268	1.46E-65	8.25E-29	1.12E-93	1.63	0.71	"TLN2, AC103740.1"
chr17	40349347	40355614	6268	*	22	1.56E-135	5.33E-55	1.00E-14	2.71E-91	1.00	0.37	"RARA, AC080112.3"
chr17	17561538	17562367	830	*	9	1.58E-169	1.87E-91	2.37E-16	8.53E-89	0.74	0.61	PEMT
chr4	8391190	8395086	3897	*	16	1.38E-135	2.24E-73	3.46E-21	3.03E-88	-1.48	0.09	"RNA5SP152, ACOX3"
chr20	45823027	45824567	1541	*	7	9.84E-202	7.06E-74	1.21E-26	2.97E-87	1.96	1.11	TNNC2
chr13	114037071	114040243	3173	*	16	1.50E-81	3.83E-83	5.11E-13	3.00E-87	1.10	0.57	RASA3
chr6	5132652	5133925	1274	*	8	8.73E-282	5.61E-61	1.87E-27	4.23E-87	2.30	0.91	LYRM4
chr15	93070223	93074906	4684	*	28	3.59E-104	3.36E-69	2.32E-15	1.56E-85	1.08	0.40	RGMA
chr6	33162918	33167675	4758	*	47	1.49E-75	4.20E-67	1.81E-08	4.50E-84	0.52	0.24	COL11A2
chr16	28900323	28900847	525	*	5	6.12E-200	1.08E-85	2.61E-23	1.06E-83	1.71	1.28	"AC009093.11, ATP2A1"
chr9	134052597	134054548	1952	*	11	1.06E-170	5.47E-79	1.09E-15	1.10E-83	0.86	0.56	"AL445931.1, BRD3"
chr10	86967301	86971537	4237	*	25	8.98E-88	4.03E-63	1.41E-11	3.11E-83	1.05	0.39	"ADIRF, AGAP11, MMRN2
chr14	51394715	51397712	2998	*	10	4.28E-175	3.33E-60	3.26E-23	7.00E-82	1.84	0.84	LINC02310
chr17	81385205	81388556	3352	*	20	6.36E-132	3.94E-69	8.82E-12	3.64E-81	0.79	0.39	AC110285.1, AC110285.2
chr1	223142909	223144507	1599	*	16	5.81E-216	1.46E-43	5.39E-20	8.37E-80	1.42	0.51	"AL359979.2, TLR5"
chr12	6539603	6540827	1225	*	8	1.77E-143	5.78E-77	6.92E-19	5.82E-79	1.48	0.85	IFFO1
chr2	45256003	45257840	1838	*	8	7.65E-156	9.13E-67	1.13E-20	7.89E-79	1.58	0.87	LINC01121
chr4	37908629	37908651	23	*	4	1.80E-184	4.72E-80	1.70E-26	1.82E-78	2.05	1.81	TBC1D1
chr7	47581191	47583870	2680	*	17	2.32E-155	4.23E-50	2.20E-15	2.85E-77	0.93	0.31	TNS3
chr7	1708444	1710037	1594	*	13	4.86E-139	3.95E-65	1.07E-11	7.88E-77	0.82	0.49	ELFN1
chr8	143518474	143520314	1841	*	10	4.40E-135	1.20E-71	5.90E-13	4.62E-76	1.52	0.78	ZC3H3
chr16	28888827	28889726	900	*	4	2.92E-205	8.81E-77	5.80E-24	8.85E-76	1.79	1.45	"AC009093.11, ATP2A1"
chr7	4711979	4713370	1392	*	11	2.34E-175	6.91E-55	4.14E-18	2.58E-75	1.63	0.74	FOXK1
chr17	17816498	17821372	4875	*	18	1.25E-136	7.32E-55	1.64E-17	8.12E-74	1.24	0.44	SREBF1
chr16	70683998	70689697	5700	*	22	9.05E-80	4.63E-52	8.96E-10	9.42E-74	0.70	0.28	"MTSS2, VAC14"
chr17	80961374	80964543	3170	*	18	8.27E-86	4.23E-59	3.89E-11	1.88E-73	0.95	0.39	"RPTOR, AC127496.4"
chr6	30682959	30690898	7940	*	83	6.09E-98	3.12E-41	1.42E-11	2.27E-73	0.75	0.10	"PPP1R18, NRM"
chr9	136052057	136055265	3209	*	28	1.24E-203	2.68E-35	1.34E-20	2.94E-73	1.08	0.21	NACC2

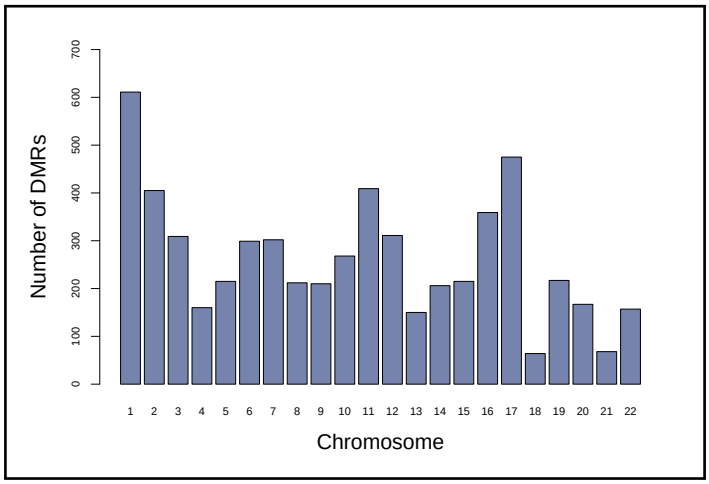


Figure 5.16: Chromosomal distribution of DMRs

Chromosomal distribution of DMRs identified with DMRcate. All autosomal chromosomes are displayed along the x axis with DMR count on the y axis

Gene Ontology and Reactome pathway overrepresentation analysis of DMRs

GO-ORA was conducted to assess the relationship between DMRs identified in TI vs TII fibres. Cluster profiler was used for GO-ORA using the 5,778 significant DMRs. 311 enriched terms were identified with a BH P-value cutoff of < 0.005, with 189 of these terms passing a stringent < 0.001 BH P-value cutoff. There were significant overlaps between DMRs and muscle related biological processes (147 enriched terms < 0.001), cellular component (32 enriched terms <0.001) and molecular function (10 enriched terms <0.001) GO terms (Figures B.11 B.12 B.13). As expected the main enriched gene sets for BP, CC and MF were similar to those identified using DMPs such as actin filament organisation, muscle contraction and GTPase regulator activity. A total of 10 Reactome pathways were identified with a P-value cutoff of <0.005 of which 1 pathway passed a P-value cutoff <0.001. Muscle contraction was the identified reactome pathway with equal numbers of genes hypermethylated or hypomethylated between TI and TII muscle fibres in key contractile proteins. A total of 36 KEGG pathways with a P-value cutoff <0.005 were enriched, with 23 pathways passing a P-value cutoff <0.001 (Figure B.14). There was overlap between kegg pathways identified using DMPs and DMRs (Figures B.14 & B.10) with focal adhesion and axon guidance identified in both. Notably both RAP1 signalling and adrenergic signalling in cardiomyocytes were in the top five enriched KEGG pathways (Figure B.14). Once again adjusting for bias using a modified missMethyl code

resulted in a reduction of the number of GO, KEGG and reactome terms identified at a BH P-value <0.001. A total of 37 GO terms were identified with the top of each GO ontology related to muscle processes (Figure 5.17). A similar set of KEGG pathways were identified using the bias adjusted method, however none of which had BH-P-values < 0.005 (Figure 5.18). Muscle contraction (BH-P-value < 0.001) and striated muscle contraction (BH-P-value < 0.005) were the only reactome terms identified between TI and TII muscle fibres (Figure 5.19).

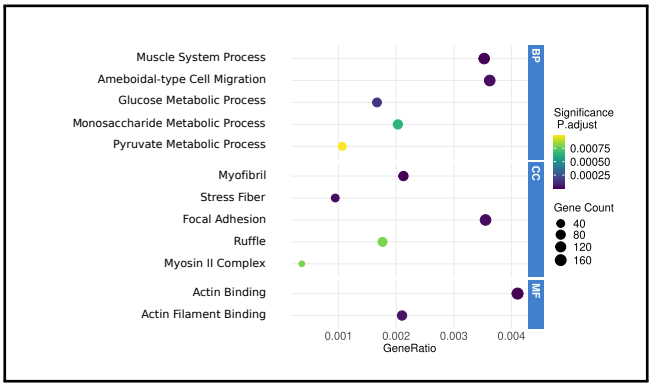


Figure 5.17: Gene ontology ORA of DMRs between TI and TII fibres

Gene ontology overrepresentation analysis of the differentially methylated DMRs between TI and TII fibres as calculated by the missMetyl method. The circle size represents gene numbers, and the colour represents the adj-P-value.

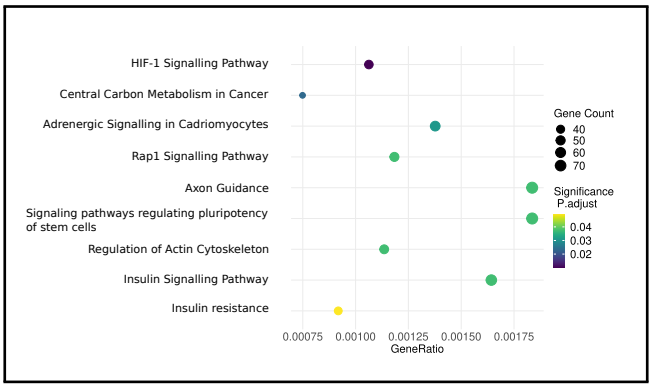


Figure 5.18: KEGG ORA of DMRs between TI and TII fibres

Gene ontology overrepresentation analysis of the differentially methylated DMRs between TI and TII fibres as calculated by the missMetyl method. The circle size represents gene numbers, and the colour represents the adj-P-value.

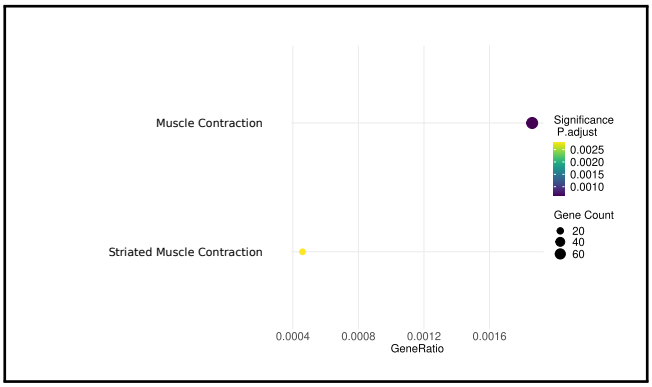


Figure 5.19: Reactome ORA of DMRs between TI and TII fibres

Gene ontology overrepresentation analysis of the differentially methylated DMRs between TI and TII fibres as calculated by the missMetyl method. The circle size represents gene numbers, and the colour represents the adj-P-value.

Whole muscle comparisons with Type I and Type II muscle fibres

To understand the differences between the two fibre types and whole muscle, I compared the pre Type I fibres with the pre whole muscle samples followed by the pre Type II fibres with the pre whole muscle samples. TI pre verses WM pre comparisons revealed 12,809 CpGs with BH-P-value < 0.001 of which 1,584 had > 10% beta change (hypermethylated = 1001 and hypomethylated = 583). TII pre verses WM pre comparisons revealed 4897 CpGs with < 0.001 and only 1040 with > 10% beta change (hyper 454 and hypo 586). The top three hypomethylated DMPs were KALRN , MYH7B and PDLIM1 and the top three hypermethylated DMPs were TNNT3, TNNI2 and TNNT3 from the TI verses WM comparison. The top three hypomethylated DMPs were GREB1, LYRM4 and TNNT3 and the top three hypermethylated DMPs were MYH7B, TNNI1 and ATP2A2 from the TII verses WM comparison. TI verses TII hypermethylated CpGs overlapped with hypermethylated TI verses WM CpGs and TI vs TII hypomethylated CpGs overlapped with the TI verses WM hypomethylated CpGs (Figure 5.20). In contrast TI verses TII hypermethylated CpGs overlapped with T2vsWM hypomethylated CpGs and TI verses TII hypomethylated CpGs overlapped with the TII verses WM hypomethylated CpGs (Figure 5.21).

To visualise differences between TI, TII and WM samples, I performed hierarchical heatmap clustering of the top 50 hypomethylated and top 50 hypermethylated CpGs between TI and TII fibres. The TI, TII and WM samples at baseline separate into three clusters (Figure 5.22). There is a distinct difference between TI and TII fibres as expected

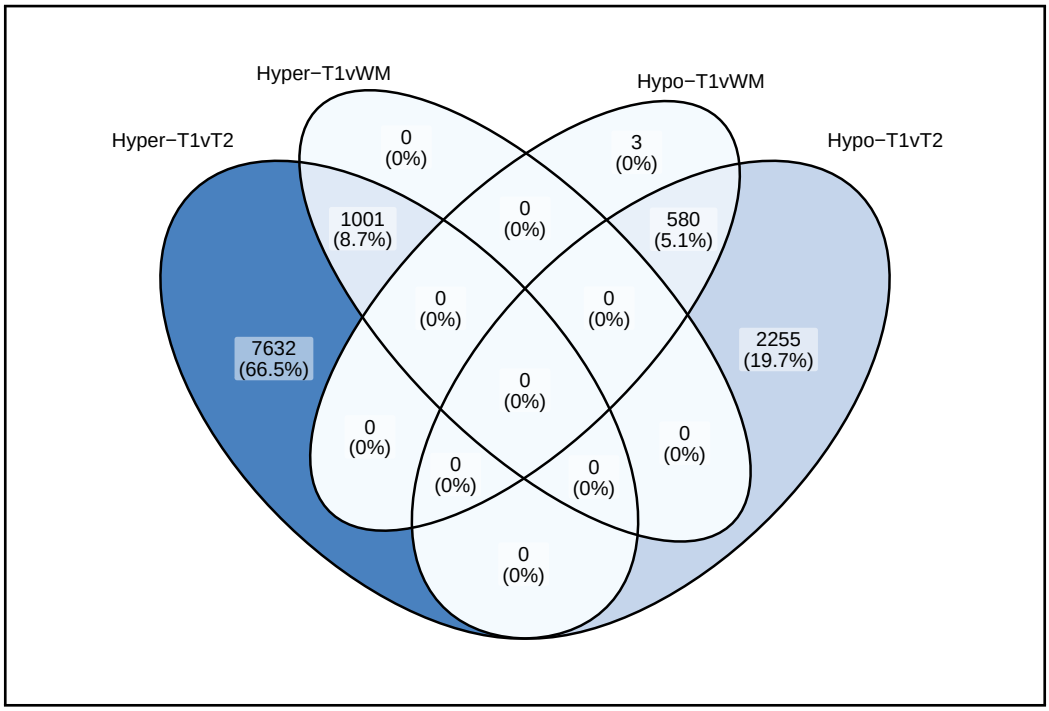


Figure 5.20: Overlaps of TI-WM DMPs with TI-TII DMPs

The overlaps between DMPs identified in TI verses TII and DMPs identified in TI verses WM. DMPs counts displayed as numbers with overall percentages of DMPs present in each overlap.

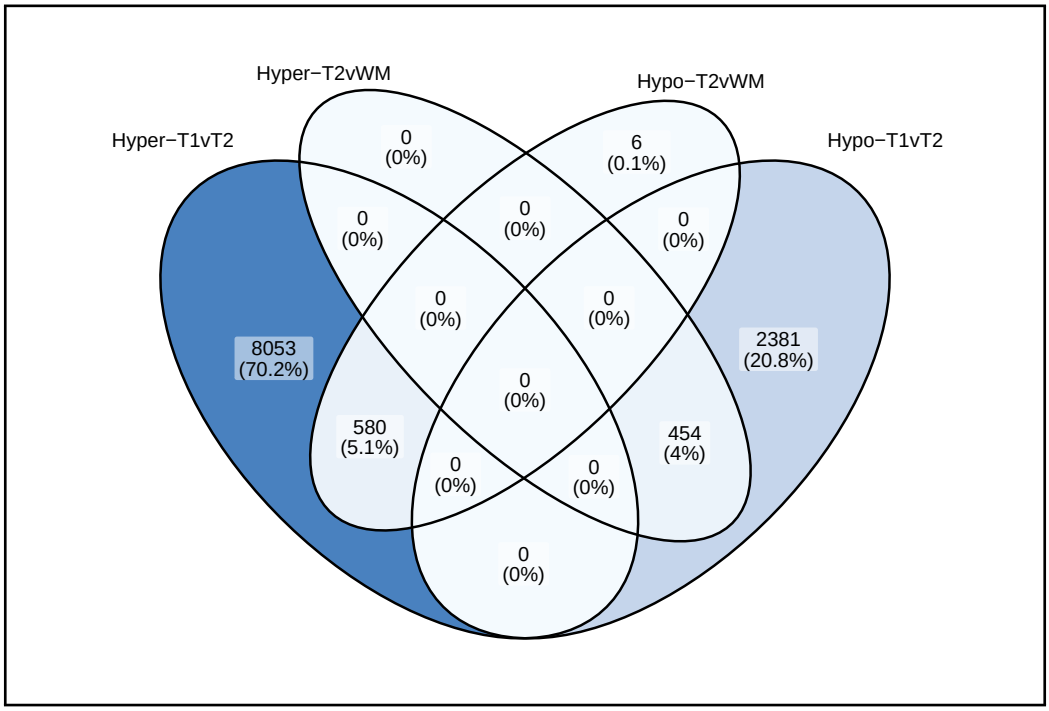


Figure 5.21: Overlaps of TII-WM DMPs with TI-T2 DMPs

The overlaps between DMPs identified in TI verses TII and DMPs identified in TII verses WM. DMPs counts displayed as numbers with overall percentages of DMPs present in each overlap.

with WM samples showing a mixed pattern. Plots of beta values from the top DMPs annotated to MYH7, MYH2, TNNT3, and TNNI1 (Figure 5.23) show a clear methylation difference between TI and TII fibres with WM having beta-values in the middle. Beta values of the top DMP in MYH1 show the same pattern with lower methylation values in TII fibres compared with TI fibres (Figure 5.23E). This confirms that the TII fibres are likely a population of TIIa and TIIx muscle fibres.

The results of WM vs fibre type comparisons suggest that a comparison of TI or TII verses WM is largely driven by muscle fibre type.

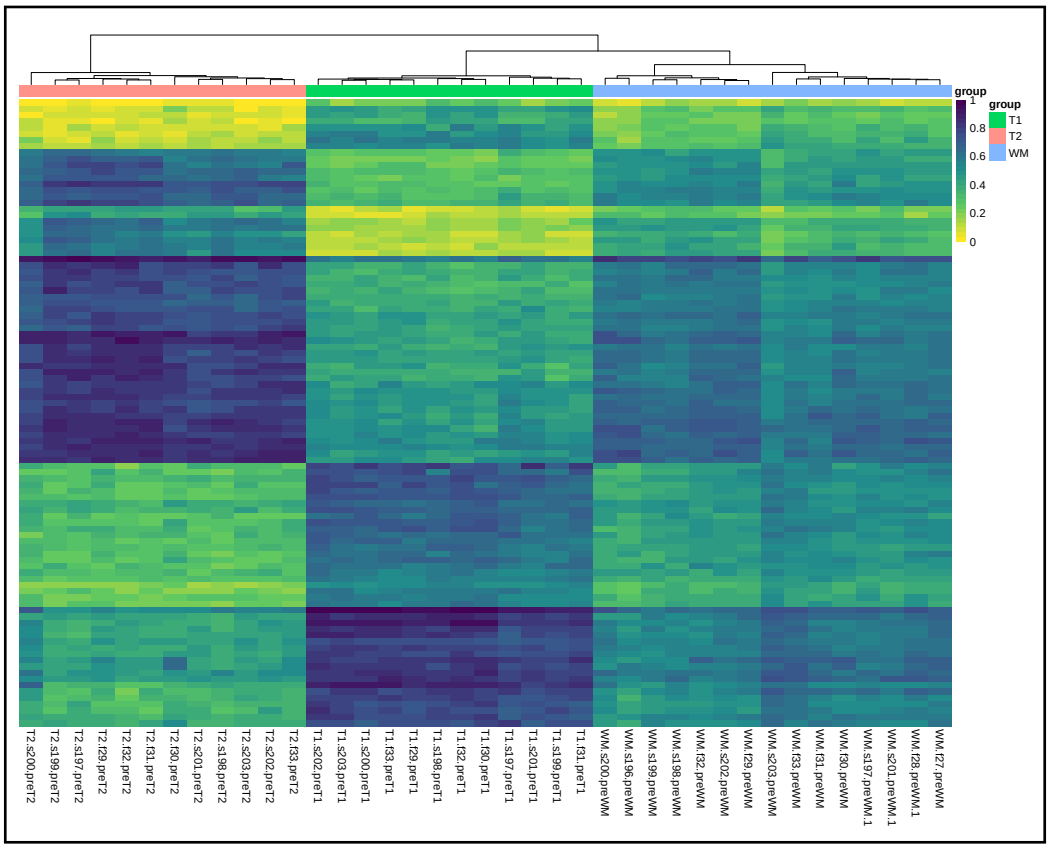
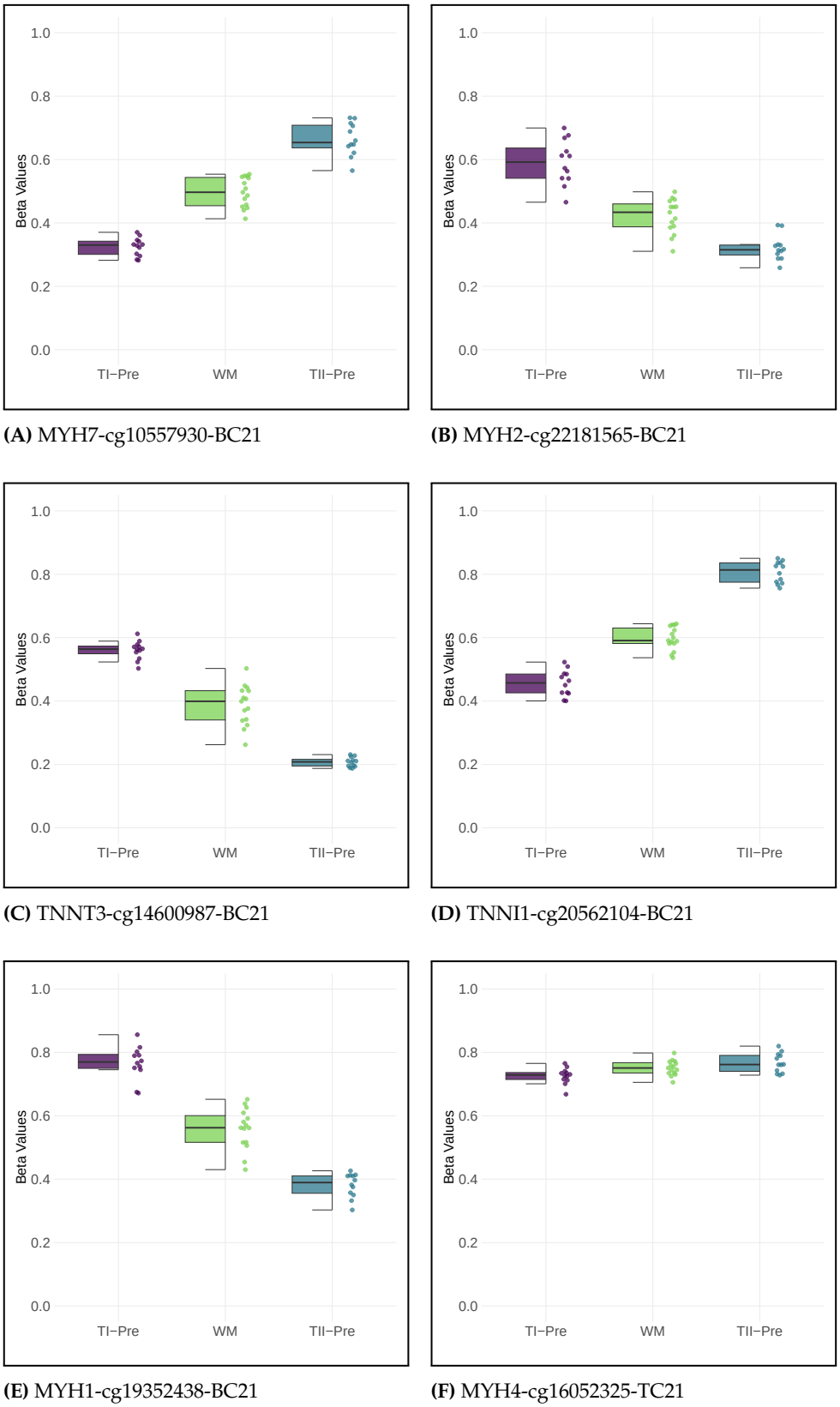


Figure 5.22: Heatmap of top 100 DMPs

Heatmap hierarchical clustering analysis of the top 50 hypomethylated and top 50 hypermethylated DMPs. The colour-coded scale (purple = hypomethylation and yellow = hypermethylation) are the min-max normalised beta-values. Hierarchical clustering of rows and columns represents beta values of the CpG sites and samples. Type I (green), Type II (red) and Whole Muscle (blue) samples are colour coded and indicated with appropriate bars.



(A) MYH7-cg10557930-BC21

(B) MYH2-cg22181565-BC21

(C) TNNT3-cg14600987-BC21

(D) TNNI1-cg20562104-BC21

(E) MYH1-cg19352438-BC21

(F) MYH4-cg16052325-TC21

Figure 5.23: CpG Plots of top genes between TI and TII fibres

Boxplots of methylation beta level at single CpG sites in TI, TII and WM samples at baseline. (A) MYH7, (B) MYH2, (C) TNNT3, (D) TNNI1, (E) MYH1 & (F) MYH4. The y axis shows beta values with sample group separated on the x axis.

5.4 Discussion

Here, we successfully measured DNAm, using a method I developed (Chapter 3), in pooled Type I and pooled Type II skeletal muscle fibres from 12 human participants. I have extended DNAm analysis of Type I and Type II fibres from a previous study [3] with a sample size n=2 (males) to a sample size of n=24 and included both male and female samples. There were significant DNAm differences between Type I and Type II human skeletal muscle fibres as assessed using the latest Infinium EPICv2 beadchip. We identified 11,469 differentially methylated positions at a stringent BH adjusted P-value < 0.001 with > 10% beta fold change (Figure 5.8) and 5,778 differentially methlyated regions at a stringent BH adjusted P-value < 0.001 (Tables 5.13 & 5.14) annotated to 3,996 unique genes between Type I fibres compared to Type II fibres. Overall more of these DMPs and DMRs were hypermethylated in Type I fibres compared with Type II fibres. Differentially methylated positions and regions were identified in key muscle genes involved in actin regulation, muscle contraction and RAP1 signalling (Figures B.11, B.9 & B.14). We also identified differential methylation between Type I and Type II fibre in mitochondrial related genes (Figure 5.14). Further, this is the first reported study utilising the newest EPIC beadchip in skeletal muscle tissue and skeletal muscle fibre types. We assessed seven different DNAm normalisation methods to ensure biological variation was not removed and technical variation was minimised. We showed that the Dasen normalisation method from the wateRmelon R package was an appropriate normalisation strategy.

I found an overall hypermethylation of Type I fibres compared to TII fibres with 78.4% of all DMPs located within gene bodies (Figure 5.9), this is in agreement with results reported previously from two muscle fibre samples assessed using a different method RRBS [3]. I also identified significant DMPs in key myosin genes (Table 5.9 slow; MYH7B, MYH7, MYL2 & Table 5.10 fast; MYH2, MYH8, MYH3) and I report the same hypermethylation in PFKM and SERCA1 (ATP2A1) in Type I fibres compared with Type II fibres as previously reported by Begue and others [3]. I also confirmed these results using differential methylation analysis of regions defined as sections of the genome with ≥ 4 CpG sites. I assessed the functional role of identified DMPs and DMRs by conducting ORA of GO, KEGG and REACTOME terms. Enriched terms were associated with actin cytoskeleton and skeletal muscle contraction.

Not only are there mitochondrial content and volume differences between Type I and Type II muscle fibres, there is also strong evidence that fibre types differ in mitochondrial specialisation [309]. I compared the list of DMPs (BH-P-values < 0.001 and > 10% beta differences) with the genes in the MitoCarta gene list to determine if there are methylation differences in key mitochondrial genes between Type I and Type II muscle fibres. I identified 133 MitoCarta genes with at least one DMP. Multiple mitoDMPs were annotated to LYRM4 a member of the leucine/tyrosin/arginine motif family of proteins. These LYRM4 mitoDMPs formed a hypomethylated cluster in Type II fibres in comparison to Type I fibres.

In this chapter I assessed DNAm in Type I and Type II human skeletal muscle fibres from multiple individuals. Previous work pioneered DNAm assessment in muscle fibre types, however only included male samples [3]. I extended this work to include females and provide a comprehensive DNAm assessment of Type I vs Type II muscle fibres. I expect the DMPs and DMRs reported here to be a useful resource for researchers investigating skeletal muscle fibre types, and skeletal muscle biology. I have also provided evidence that DNAm can be assessed from multiple samples using the Infinium MethylationEPIC Beadchip. Future research could utilise this platform to further study DNAm in human skeletal muscle fibres in disease and during ageing. In the following chapters I highlight the use of this method to study the link between DNAm and protein expression and fibre-type DNAm changes as a result of 12 weeks HIIT.

6 Results 3: Proteomic differences between Type I and Type II fibres and comparison with DNAm

6.1 Introduction

Genes are considered regions of the genome that encode useful information such as proteins (protein-coding-regions) and non-coding RNAs (non-coding-regions). DNAm has been reported in both regions of DNA and can influence the expression of both coding and non-coding regions [311]. For example, there is evidence that heavy methylation in CpG rich regions of genes inhibits transcription [311]. The overall flow of information from DNA through RNA to protein and the study of the interaction of these multiple omics layers is a major interest of the life sciences [312].

Recent work from our laboratory and others have attempted to integrate these diverse omics layers in whole skeletal muscle [159, 313, 263, 167, 181]. Further, the study of the proteome and transcriptome in single muscle fibres have revealed interesting insights in to muscle fibre heterogeneity [211]. DNA and RNA are routinely separated from the same piece of skeletal muscle, however the extraction of DNA and protein are mostly obtained from different pieces of muscle. Using the method I developed in chapter 4, I obtained DNA and protein extracted from the same pool of fibres. I was able to confirm at the protein level the success of pooling MYH7 and MYH2 muscle fibres, I further characterised protein differences between pooled Type I and pooled Type II muscle fibres and I compared the fibre-type-specific proteome data to the fibre-type-specific methylome data. Most DNAm studies have reported DNAm differences and attempted to link this to either RNA or protein changes by curating external data sources. This makes it difficult to control for the variable fibre type proportions of samples studied. The data presented here provide a unique dataset of methylation and protein differences of Type I and Type II muscle fibres.

This chapter addresses the following aims

1. To quantify the distribution of myosin heavy chain isoforms in each sample to assess fibre-type purity and to extend this to myosin light chain isoforms.

2. To assess protein expression differences of pooled Type I and pooled Type II fibres from both male and female samples using MS-based proteomics.

3. To compare DNAm data with protein expression data generated from the pooled Type I and Type II muscle fibre samples.

6.2 Methods

6.2.1 Sample Preparations

Proteomic sample preparation and data were generated from 100 µL of protein lysate from 12 baseline Type I, 12 baseline Type II and two baseline Type IIx pooled fibre samples. Proteins were measured by liquid chromatography with tandem mass spectrometry at Monash Proteomics and Metabolomics Platform at Monash University, Melbourne, Australia. ~1 µg of peptide was injected for LC-MS/MS. Raw LC-MS/MS data files were analysed using Fragpipe (V 19.1) and the MSfragger (V 3.5) search engine [289] to obtain protein identifications. The picky library was used for protein search and identifications [290]. For detailed methods please see the materials and methods.

6.2.2 Sample QC

Quality control and statistical analysis was performed using the R software. First contaminant proteins were removed for all non human species. All proteins identified in at least four Type I four Type II and both Type IIx fibres were kept so that proteins with inconsistent identification were removed. The MaxLFQ values were log2 transformed, grouped by condition and missing values were imputed assuming they were Missing not at random. The total number of proteins identified after protein filtering was on average 1226 per sample (Figure 6.1). The number of proteins was consistent across samples and conditions except for one TII pre sample, which had a reduced number of identified proteins. To visualise the quantitative patterns in the proteomics data I generated heatmap

clustering on missing values and Pearson correlations. Sample s197-pre-T2 clusters separately to the other fibres on the missing value heatmap due to the lower number of identified proteins (Figure C.1). Aside from this value the samples cluster according to fibre type in the missing value heatmap (Figure C.1). Due to the amount of missing values in s197-pre-T2 this sample was excluded from the differential protein assessment, however I proceeded to quantify myosin isoform content of this sample. Pearson's correlation coefficients were computed and a heatmap of results generated. Samples cluster and show distinct patterning overall in a Type I or Type II group. Interestingly, samples appear to have a degree of separation according to sex (Figure 6.2). Coefficient of variation (CV) are often used as a measure of data reliability with results > 30% considered signs of problematic data [314]. CVs for each sample group were in the range of 20% (Figure C.2), I note that overall data quality should not be judged solely on CV values and should be balanced with other QC measures [315]. Considering the QCs measures, the sample preparation and number of identified proteins were of good quality.

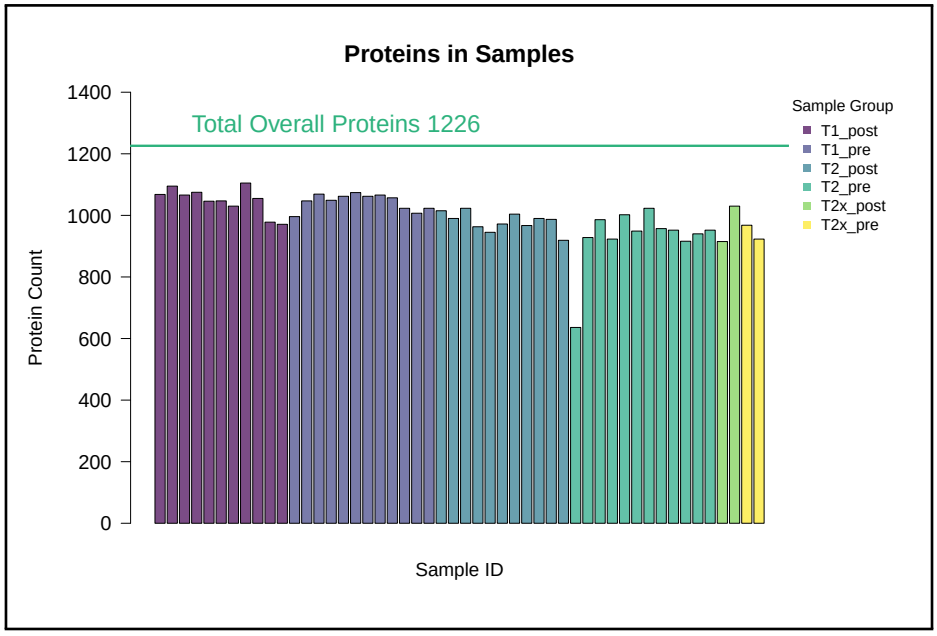


Figure 6.1: Protein Counts of TI, TII and TIIx Samples

Figure 6.1 shows the number of proteins identified in each sample. Approximately 1000 proteins were measured in each sample using MS-based proteomics.

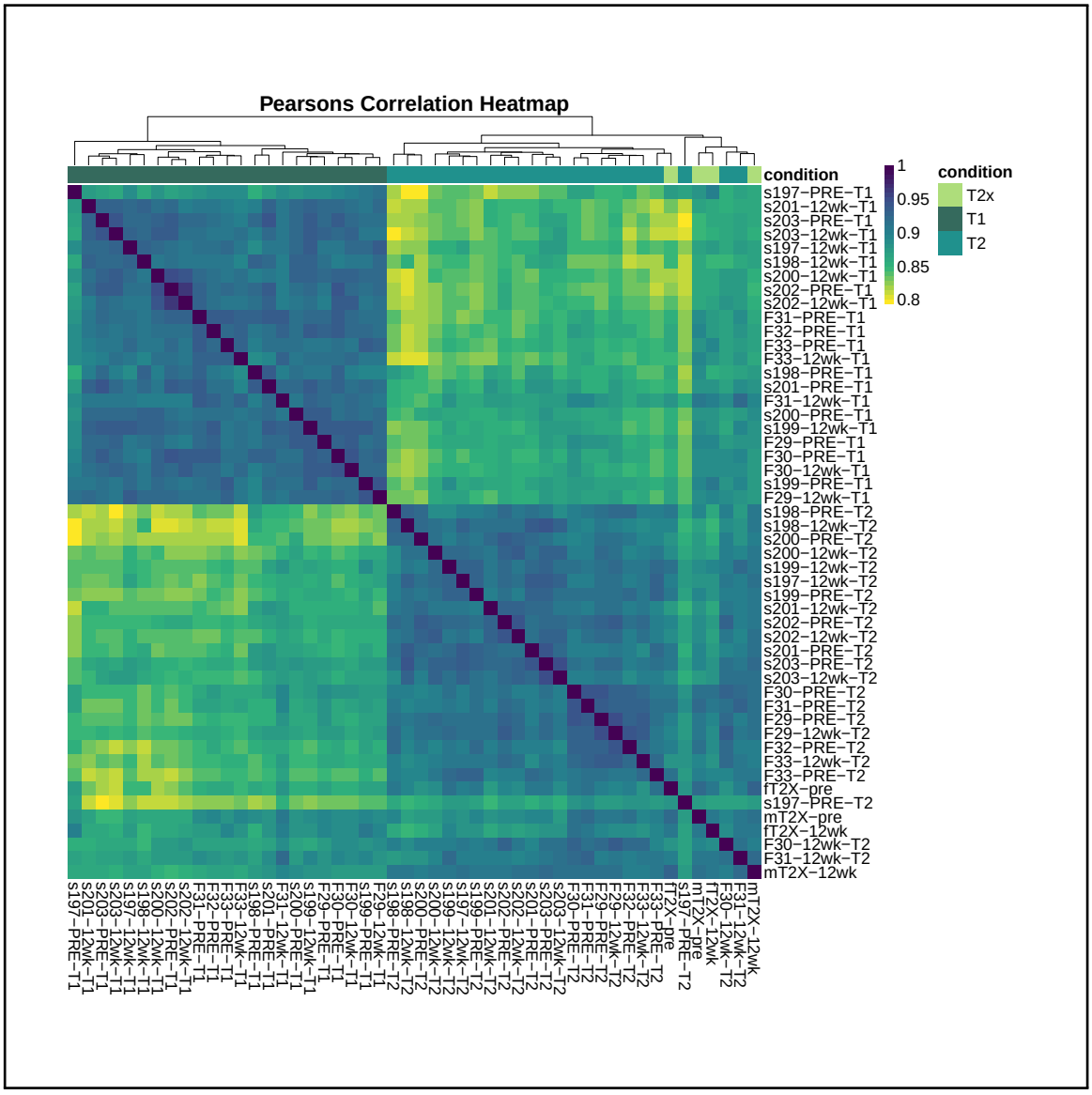


Figure 6.2: Sample Correlation Matrix of TI, TII and TIIx fibres

Sample correlation matrix of TI, TII and TIIx samples. The colour scale indicates the spearman correlation between samples based on protein values. Type I (dark green), Type II (lightgreen) and TIIx (aqua) samples are colour coded and indicated with appropriate bars.

6.3 Results

6.3.1 Proteomic Analysis

Dimension Reduction Analysis

PCA, MDS and t-SNE were used to assess variation in the data. PCA analysis showed that of variation in the data was explained by the first principal component and that samples clustered according to TI and TII fibres with TIIx fibres similar to TII fibres (Figure 6.3). The top proteins driving the separation along PC1 are myosin and troponin proteins (Figure C.3A). MDS analysis showed the same clustering pattern as the PCA (Figure C.3B). t-SNE analysis also confirmed samples clustering according to fibre type (Figure C.3C). PCA, MDS and t-SNE analysis also confirmed that sex is a major source of variation in the data, with the second component in each plot showing strong clustering based on sex, however not all of component two could be explained by sex alone. In each plot there appeared to be a small influence of the paired design of the experiment. As a result, I decided to include both sex and a blocking factor on participant ID into the linear model in a similar manner to the methylation analysis.

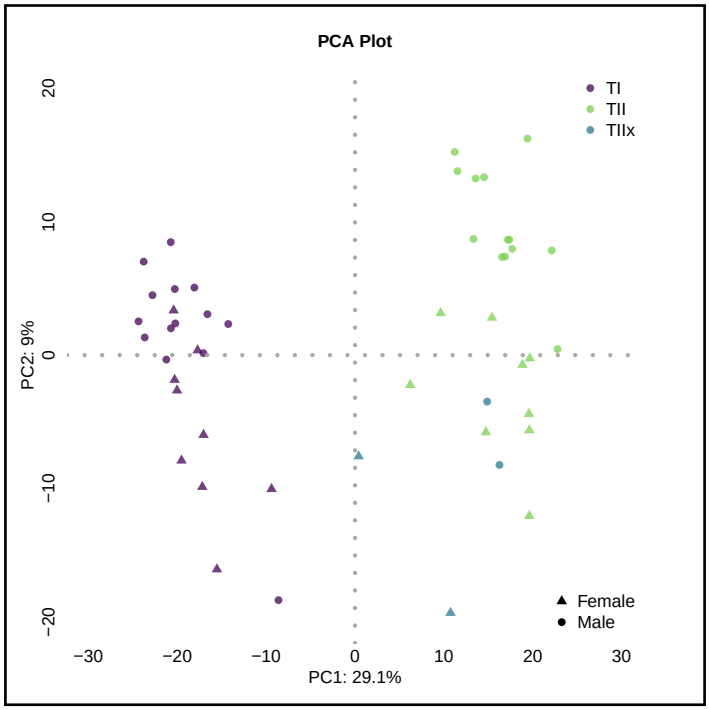


Figure 6.3: PCA Plot

Figure 6.3 is a PCA plot showing that the first principal component explains 29.1% of the variation and is clustered according to fibre type. TI = Type I muscle fibres, TII = Type II muscle fibres and TIIx = Type IIx muscle fibres. Triangles indicate females and circles indicate males.

Myosin Isoform Content of pooled samples

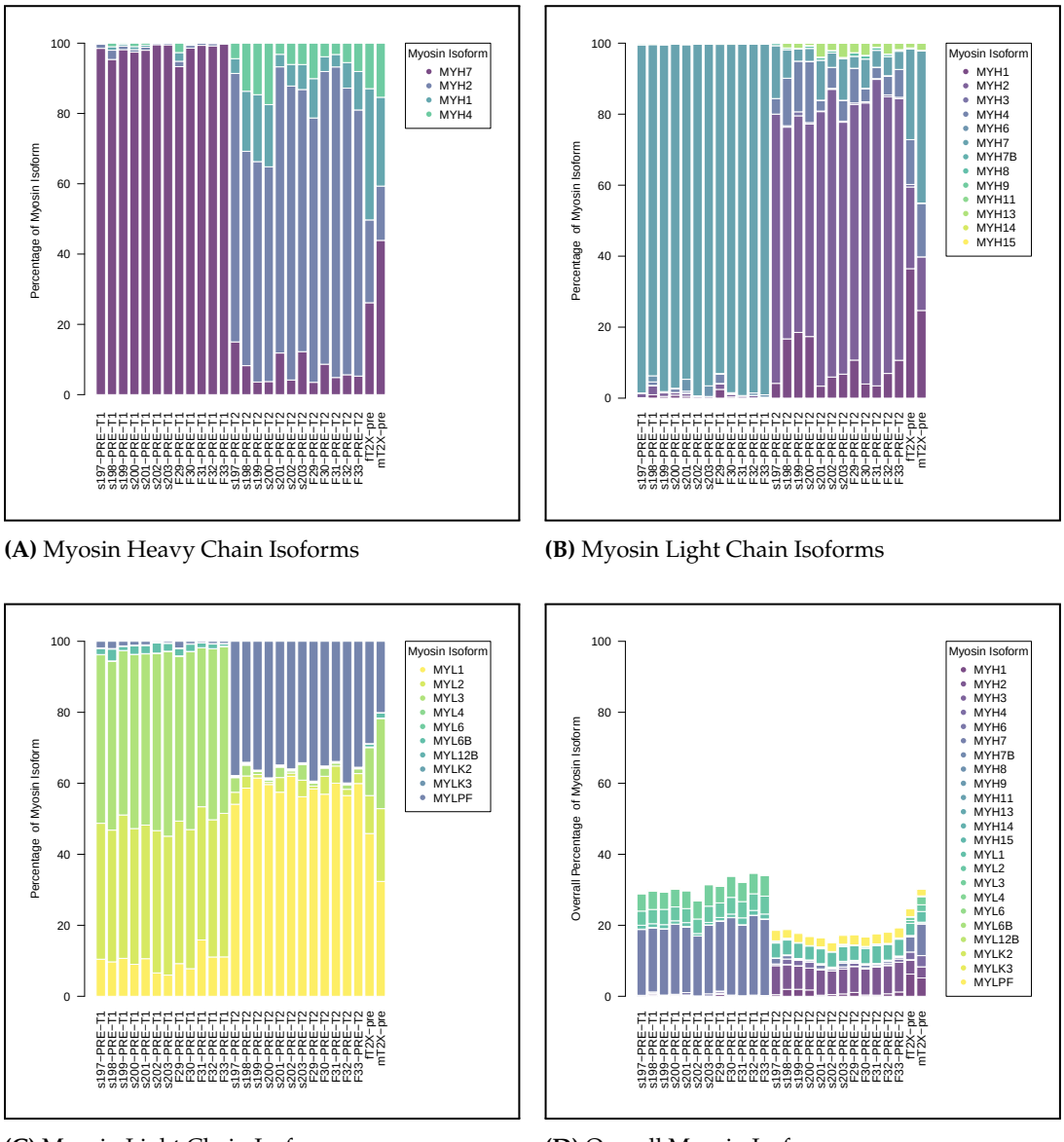
Myosin proteins were assessed to characterise the pooled samples. I identified a total of 11 myosin class two proteins in the samples of which MYH7, MYH2, MYH1 and MYH4 were found to be highly abundant (Figure C.4). I also identified 10 myosin light chain (MLC) proteins of which MLC1, MLC2, MLC3 and MLCPF were identified as some of the most abundant proteins (Figure C.4).

In human skeletal muscle MYH7, MYH2, MYH1 and MYH4 are normally used to type skeletal muscle fibres [59]. Previously, proteomic studies of pooled muscle fibres or single muscle fibres have quantified the percentages of MYH7, MYH2, MYH4 and MYH1 in Type I, Type II and Type IIx muscle fibres [210, 209, 203]. We performed the same assessment in 12 pooled Type I and 12 Type II fibres and on two Type IIx fibres (as we hypothesised were Type IIx fibres due to being MYH7 and MYH2 negative but ponceau positive fibres) at baseline. We show in agreement with previous studies that MYH7 positive Type I fibres express mostly MYH7 (on average 98% MYH7 expression), and MYH2 positive fibres express mostly MYH2 (on average 75 % MYH2) (Figure 6.4A). We reported difficulties in Typing Type IIx fibres using the reported antibodies, and interestingly the Type II fibres show a high proportion of MYH1 and MYH4 with on average 9% and 8% expression respectively. What we speculated as Type IIx fibres showed a mixed percentage of MYH7, MYH2, MYH4 and MYH1 (Figure 6.4A). These calculations of myosin percentages only considered the four reported MYHs (MYH7, MYH2, MYH1, and MYH4). To further characterise the samples I quantified all 11 identified myosin heavy chain isoforms in the pooled fibre samples. The MYH isoform percentages remained the same indicating that MYH7, MYH2, MYH4 and MYH1 are the predominant MYH in these fibre types, interestingly Type II fibres also contained small amounts of MYH13 (Figure 6.4B).

The myosin molecule in skeletal muscle is composed of both heavy chain and light chains the combinations of which may determine the speed of muscle shortening [316]. The pattern of MYL's in skeletal muscle is less frequently reported in proteomic studies of skeletal muscle fibre types. We quantified the same percentage breakdown of MLCs as we did for MYHs, in Type I fibres MYL2 and MYL3 constitute 39% and 48% of the MYL isoforms with a small 10% contribution of MYL1, in comparison Type II fibres were composed of 58% MYL1 and 36% of MYLPF (Figure 6.4C).

The former analysis was conducted within the MYH and MLC proteins, to determine the overall percentage of proteins attributed to MYH and MLC we calculated the percentages of each protein out of all proteins identified in the sample. In Type I fibres on average 30% of total signal was attributed to MYH and MLC with 20% and 10% to MYH7 and MYL2/MYL3 respectively (Figure 6.4D). In Type II fibres on average only 20% of the signal was attributed to MYH and MLC with 12% to MYH2 and 8% to MYL1/MYLPF (Figure 6.4D).

The results of the myosin assessment in the samples are in agreement with the pooled dot blot results presented in Figure 5.2. It is important to note that by not using a dedicated antibody for MYH1 and Type IIx fibres, the Type II pooled fibres are most likely fast Type IIa/IIx fibres. We confirm here by LC-MS/MS proteomics that we have been able to successfully separate and isolate proteins from pooled Type I and Type II muscle fibres and that these samples are a relatively pure representation of each fibre type as determined by MYH and MYL patterns.



(A) Myosin Heavy Chain Isoforms

(B) Myosin Light Chain Isoforms

(C) Myosin Light Chain Isoforms

(D) Overall Myosin Isoforms

Figure 6.4: Overview of Myosin Isoforms

Figure 6.4 shows the myosin isoform content of Pre Type I and Type II muscle fibres. (A) Is the quantity of MYH7, MYH2, MYH1 and MYH4 as a percentage, (B) is the quantity of all MYHs identified, (C) is the quantity of all MYL, and (D) is the quantity of all MYH and MYL as a percentage of all proteins in the sample.

Differential expression of proteins in Type I vs Type II Fibres

To estimate differences in protein expression between TI and TII muscle fibres at baseline, mixed effects linear models were performed using limma in the R statistical environment on normalised, imputed and log2 transformed MaxLFQ values. Sample s197-pre-T2 was removed from the analysis due to the reduced number of proteins identified.

A comparison of TI and TII samples at baseline (n = 12 vs 11) revealed 428 DEPs after Benjamani-Hochberg correction at FDR p < 0.005. After applying a FDR threshold of < 0.001, 339 proteins were identified as differentially expressed between TI and TII fibres. Of the proteins passing FDR < 0.001, 154 showed a log-fold-change of > 1 or < -1. Among these proteins 109 were up-regulated and 34 were down-regulated in TI fibres compared with TII fibres. The list of all proteins and statistics generated using limma will be provided on request. The differentially expressed proteins in TI fibres compared with TII fibres are summarised in tables 6.2. In TI compared with TII fibres the top down-regulated proteins were MYH2, TNNC2, TNNI2, TNNT3, MYL4 and MYLPF, and the top up-regulated proteins were ATP2A2, MYH7, TPM3, TNNT1, MYL3 and TNNC1 (Figure 6.5). This result aligns with the myosin isoform comparisons and the most differentially expressed proteins tend to come from the most abundant proteins in the samples. We also observed that PDLIM1 was up-regulated in TI fibres compared with TII fibres as previously reported [210]. We observed higher expression of PDLIM7 in TII fibres compared with TI fibres, a previous single fibre proteomic study reported PDLIM7 to be a TIIx specific protein [209], suggesting that the TII fibre population is likely a combination of TIIa and TIIx fibres.

Table 6.1: Differentially expressed Proteins in TI-pre compared to TII-pre fibres

	TI-Pre	**Vs**	**TII-Pre**
BH P-value Threshold	Up	Non-DEP	Down
BH P < 0.001	283	888	55
BH P < 0.005	363	801	62
BH P < 0.05	522	622	82

Table 6.2: T1 vs T2 Top50 DEPs

	logFC	AveExpr	t	P.Value	adj.P.Val	B
ATP2A2	5.74	3.18	22.84	1.76E-26	2.22E-23	50.17
MYH7	5.44	8.94	22.52	3.17E-26	2.22E-23	49.59
TNNT1	6.35	5.18	20.70	1.02E-24	4.78E-22	46.17
TPM3	5.97	6.67	20.47	1.63E-24	5.26E-22	45.71
CYB5R1	2.02	2.39	20.40	1.88E-24	5.26E-22	45.57
TNNC1	5.94	5.22	19.90	5.12E-24	1.19E-21	44.58
MYL3	6.05	6.79	19.80	6.40E-24	1.28E-21	44.36
TNNI1	6.41	4.17	18.56	8.68E-23	1.52E-20	41.77
MYL6B	5.41	2.25	17.70	5.71E-22	8.61E-20	39.90
MYL2	5.21	6.81	17.67	6.15E-22	8.61E-20	39.83
PDLIM1	5.20	-0.37	17.18	1.88E-21	2.39E-19	38.72
MYH2	-5.15	7.08	-16.45	1.03E-20	1.20E-18	37.02
TNNC2	-4.63	4.71	-16.25	1.65E-20	1.78E-18	36.55
TNNT3	-4.45	8.14	-15.79	5.00E-20	5.00E-18	35.45
MYLPF	-4.40	5.96	-15.33	1.53E-19	1.43E-17	34.33
MYOZ2	5.64	0.38	15.06	3.07E-19	2.68E-17	33.63
TNNI2	-4.59	5.73	-14.37	1.76E-18	1.45E-16	31.89
MYL4	-4.72	-1.24	-13.68	1.09E-17	8.45E-16	30.07
DPYSL3	3.68	-1.17	13.31	2.99E-17	2.21E-15	29.05
MYBPC2	-4.83	1.95	-12.74	1.44E-16	1.01E-14	27.48
CASQ2	2.30	3.01	12.44	3.28E-16	2.14E-14	26.66
MYL1	-1.57	8.09	-12.44	3.36E-16	2.14E-14	26.63
TFAM	2.98	-2.14	12.35	4.23E-16	2.58E-14	26.40
ATP2A1	-1.07	5.35	-11.20	1.23E-14	7.15E-13	23.03
CPT1B	1.21	1.18	11.04	1.96E-14	1.10E-12	22.55
ECI2	1.25	1.25	10.54	9.09E-14	4.89E-12	21.02

Table 6.2: T1 vs T2 Top50 DEPs (continued)

	logFC	AveExpr	t	P.Value	adj.P.Val	B
HSD17B12	2.06	-2.37	10.26	2.14E-13	1.11E-11	20.16
LDHB	2.20	1.20	10.06	4.03E-13	2.02E-11	19.52
BTF3	2.60	-2.13	9.82	8.49E-13	4.10E-11	18.78
PDLIM7	-1.26	1.25	-9.49	2.46E-12	1.15E-10	17.71
PHKG1	-0.99	1.07	-9.48	2.57E-12	1.16E-10	17.67
IDI1	2.12	-2.39	9.44	2.87E-12	1.26E-10	17.55
MYOT	0.96	4.65	9.29	4.68E-12	1.99E-10	17.07
LMOD2	1.34	-1.02	9.15	7.31E-12	3.01E-10	16.62
ALDH7A1	2.47	-1.85	8.90	1.66E-11	6.62E-10	15.80
PHKA1	-0.99	1.17	-8.73	2.90E-11	1.13E-09	15.24
PHKB	-1.08	0.43	-8.71	3.11E-11	1.18E-09	15.17
COQ8A	1.10	2.64	8.65	3.83E-11	1.41E-09	14.96
TPM4	2.37	-2.12	8.55	5.30E-11	1.90E-09	14.64
MAOB	1.02	1.77	8.49	6.47E-11	2.26E-09	14.44
GPD2	-1.25	-0.70	-8.46	7.19E-11	2.45E-09	14.33
VDAC1	0.82	5.21	8.39	8.87E-11	2.96E-09	14.12
MGST3	0.83	3.44	8.31	1.17E-10	3.82E-09	13.84
ACAA2	1.09	1.76	8.25	1.41E-10	4.50E-09	13.65
SMYD1	0.78	3.11	8.22	1.56E-10	4.86E-09	13.55
MYL12B	0.84	-0.01	8.22	1.60E-10	4.86E-09	13.53
MYOM3	1.47	3.16	8.13	2.15E-10	6.42E-09	13.23
ME2	2.96	-1.91	8.08	2.57E-10	7.50E-09	13.05
NIBAN1	-2.46	-1.93	-7.98	3.58E-10	1.02E-08	12.72

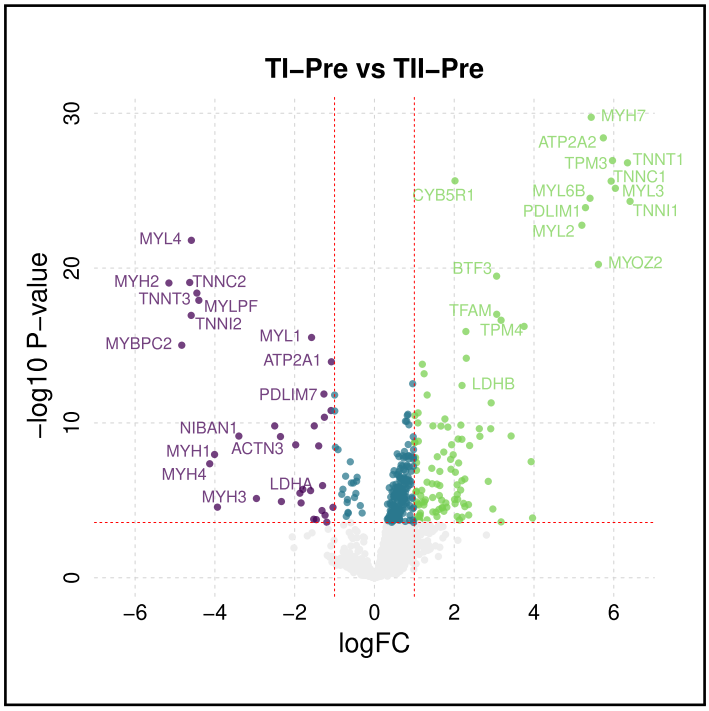


Figure 6.5: Volcano Plot T1vsT2

Volcano plot of the differentially expressed proteins between TI-pre and TII-pre samples, with log2 FC on the x-axis and -log10 P-value on the y-axis. Purple indicates down-regulated proteins and green up-regulated proteins in TI compared with TII fibres. The horizontal line represents the cutoff of the corresponding adjP-value <0.001 and the vertical lines represent the cutoff of the log2FC 1 and -1.

Gene Ontology, KEGG and Reactome Pathway Analysis

To determine pathways and processes that are over-represented in the DEPs we conducted ORA and FCS. Cluster profiler was used for ORA and fGSEA was used for FCS.

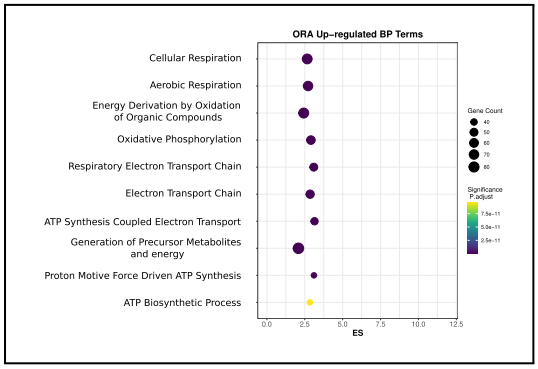
GENE ONTOLOGY

Using the 283 up-regulated DEPs with > 0 logFC, we identified 84 enriched GO terms with a BH-P-value cutoff of < 0.005, with 70 of these terms passing a < 0.001 BH-P-value cutoff. There were overlaps between DEPs and muscle related BPs (38 enriched terms < 0.001), CCs (18 enriched terms <0.001) and MFs (14 enriched terms <0.001) GO terms (Figures 6.6 & C.5 & C.6). Between TI and TII fibres the main up-regulated enriched gene sets for BP, CC and MF were related to aerobic respiration, oxidative phosphorylation, mitochondria and the electron transport chain (Figures 6.6A, C.5A & C.6A). Using the 56 down-regulated DEPs with < 0 logFC, we identified 46 enriched terms with a BH-P-value cutoff of < 0.005, with 27 of these terms passing a < 0.001 BH-P-value cutoff. There were overlaps between DEPs and muscle related biological processes (15 enriched terms < 0.001), cellular component (7 enriched terms <0.001) and molecular function (5 enriched terms <0.001) GO terms (Figures 6.6 & C.5 & C.6). Between TI and TII fibres the main down-regulated enriched gene sets for BP, CC and MF were related to muscle contraction, carbohydrate metabolism, the myosin complex and the calcium complex (Figures 6.6B, C.5B & C.6B).

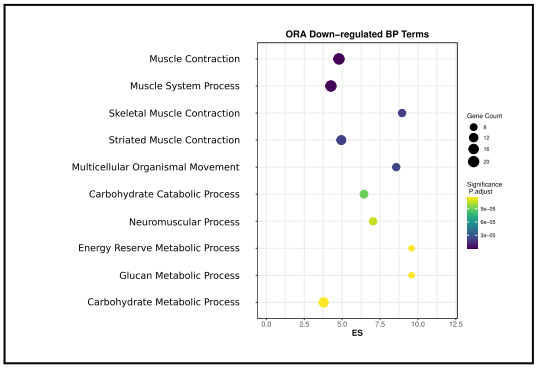
We also conducted FCS analysis using all DEPs ranked with the following equation ((-log10 * pvalue) * sign(logFC)). We identified 86 enriched GO pathways with a BH-P-value cutoff of < 0.05. 70 GO terms were up-regulated in TI fibres compared to Type II fibres. 16 GO terms were down-regulated in TI fibres compared to Type II fibres. There were overlaps between DEPs and muscle related biological processes (6 enriched terms < 0.001), cellular component (10 enriched terms <0.001) and molecular function (3 enriched terms <0.001) GO terms (Figures 6.6C & C.5C & C.6C). Between TI and TII fibres the main up-regulated enriched gene sets for BP, CC and MF were related to aerobic respiration, oxidative phosphorylation, the electron transport chain, mitochondria and oxidoreductase activity. 16 GO terms were down-regulated in TI fibres compared to TII fibres with. There were overlaps between DEPs and muscle related biological processes (1 enriched term < 0.001). There were no enriched cellular component and molecular function GO

terms at adj-P-value < 0.001 or 0.005, we did identify 3 cellular component and 2 molecular function pathways at adj-P-value < 0.05 (Figures 6.6D & C.5D & C.6D). Between TI and TII fibres the main down-regulated enriched gene sets for BP, CC and MF were related to glycolytic and carbohydrate processes and voltage gated calcium complexes.

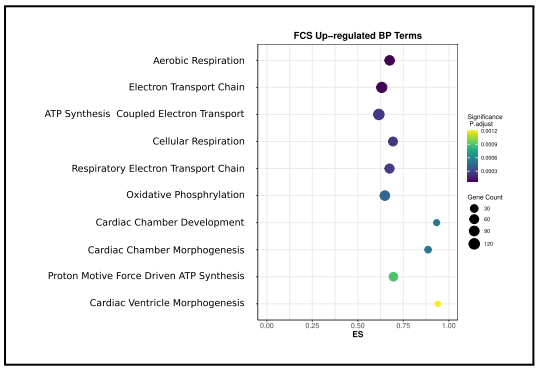
We determined the overlap of enriched pathways up-regulated and down-regulated from both ORA and FCS in TI compared with TII fibres (Figure C.7). A total of 46 up-regulated GO terms were identified using both methods, the up-regulated terms in TI fibres compared to TII fibres were involved in aerobic respiration and mitochondrial function. Seven GO terms were down-regulated and were related to carbohydrate metabolism and glycolytic processes.



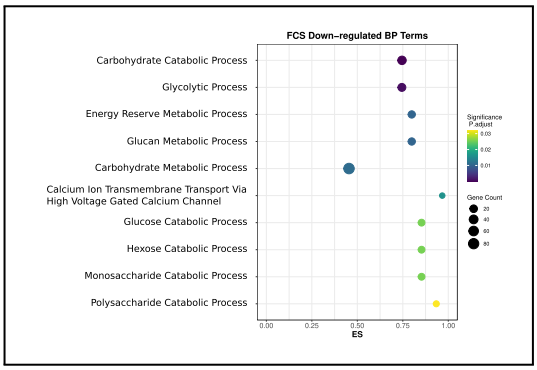
(A) ORA of up-regulated GO-BP terms



(B) ORA of down-regulated GO-BP terms



(C) FCS of up-regulated GO-BP terms



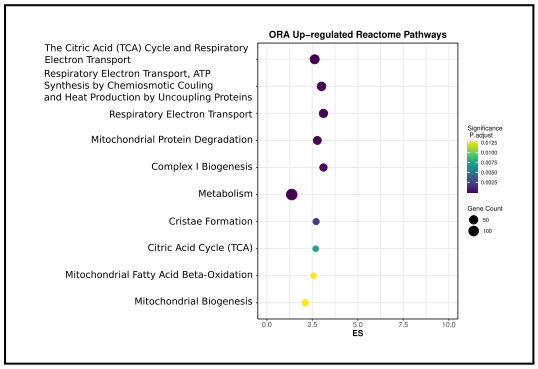
(D) FCS of down-regulated GO-BP terms

Figure 6.6: BP enrichment analysis of DEPs

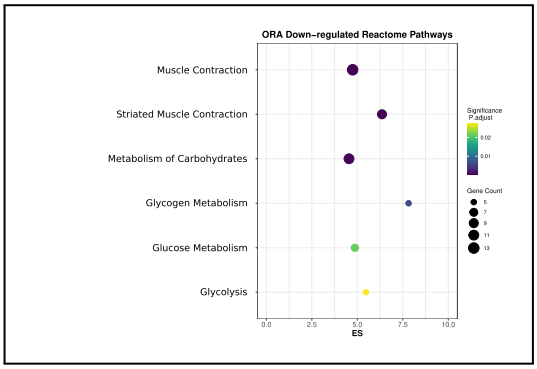
Gene ontology gene set enrichment analysis of the DEPs between TI and TII fibres. (A) & (B) are the up regulated and down-regulated biological process GO terms identified using ORA. (C) & (D) are the up-regulated and down-regulated biological processes GO terms identified using FCS. The circle size represents gene numbers, and the color represents the adj-P-value and the enrichment scores are on the x axis.

REACTOME pathways

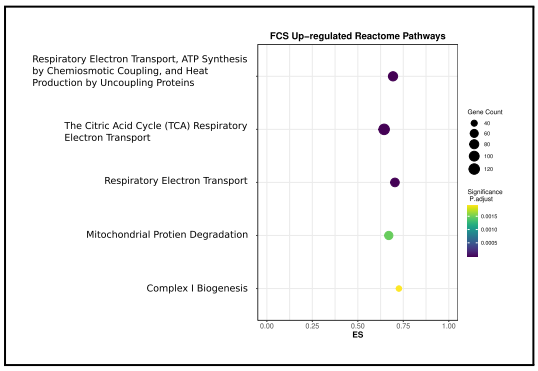
Using the 283 up-regulated DEPs with > 0 logFC, we identified 14 enriched reactome terms with a relaxed BH-P-value cutoff of < 0.05, with 7 and 6 of these terms passing a < 0.005 and < 0.001 BH-P-value cutoff respectively. Using the 56 down-regulated DEPs with < 0 logFC, we identified 6 enriched reactome terms with a BH-P-value cutoff of < 0.05, with 3 of these terms passing a < 0.001 BH-P-value cutoff. There were overlaps between up-regulated DEPs and TCA cycle, respiratory electron transport, mitochondrial protein degradation, mitochondrial metabolism and metabolism reactome pathways (Figure 6.7A). Down-regulated DEPs were enriched for muscle contraction, glucose metabolism, glycogen metabolism and glycolysis reactome pathways (Figure 6.7B). We again conducted FCS analysis using all DEPs. We identified 5 down-regulated reactome pathways with a BH-P-value cutoff of < 0.05 and 3 down-regulated reactome pathways enriched in TI fibres compared to TII fibres. The up-regulated reactome pathways were involved complex I biogenesis, mitochondrial protein degradation and the respiratory electron transport (Figure 6.7C), whilst the down-regulated pathways were glycogenolysis, glycolysis and metabolism of carbohydrates (Figure 6.7D). Once again we identified those reactome pathways that were identified using both ORA and FCS. The top up-regulated overlapping reactome terms were the citric acid cycle, respiratory electron transport, and the down-regulated overlapping terms were metabolism of carbohydrates and glycolysis (Figure C.8).



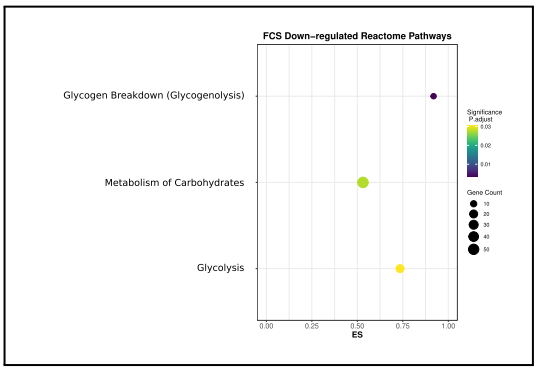
(A) ORA of upregulated REACTOME terms



(B) ORA of downregulated REACTOME terms



(C) FCS of REACTOME terms



(D) FCS of REACTOME terms

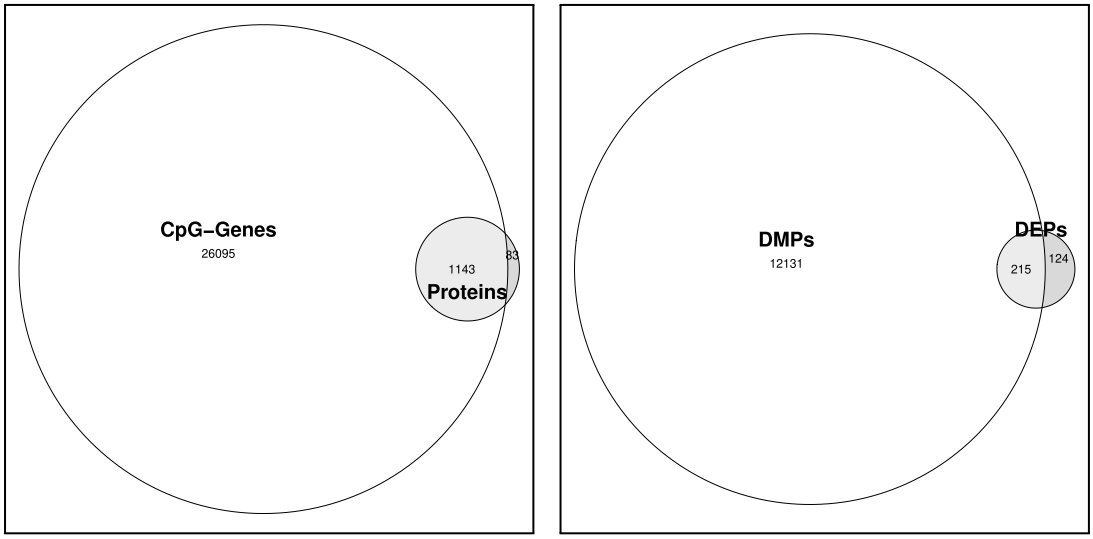
Figure 6.7: Reactome enrichment analysis of DEPs

Reactome gene set enrichment analysis of the DEPs between TI and TII fibres. (A) & (B) are the up regulated and down-regulated reactome terms identified using ORA. (C) & (D) are the up-regulated and down-regulated reactome terms identified using FCS. The circle size represents gene numbers, and the color represents the adj-P-value and the enrichment scores are on the x axis.

6.3.2 Comparisons of Proteome Methylome

Overlap of Methylome and Proteome

The second major purpose for simultaneously measuring protein and methylation in the samples was to compare and contrast the differential methylation results with the differential protein expression results. Firstly we took the 1,226 proteins identified and checked how many CpG probes on the new EPICv2 were annotated to the gene of each protein. Of the 1,226 proteins 1,143 had at least one CpG annotated to them (Figure 6.8A). Of the 339 DEPs with < 0.001 BH-P-value, 215 overlapped with at least one DMP (BH-P-value < 0.001 & logFC > 0.1 or < -0.1) (Figure 6.8B). To identify the relationship between differential methylation and differential expression of proteins, I took the overlap of up-regulated DMPs and DEPs and down-regulated DMPs and DEPs and found the overlapping genes/proteins. There were no clear relationship between DMP methylation direction and protein regulation considering just overlaps (Figure 6.9A). Comparing the overlaps of DMRs to DEPs resulted in a similar overlap of up regulated DEPs with either hypermethylated or hypomethylated DMRs (Figure 6.9B).

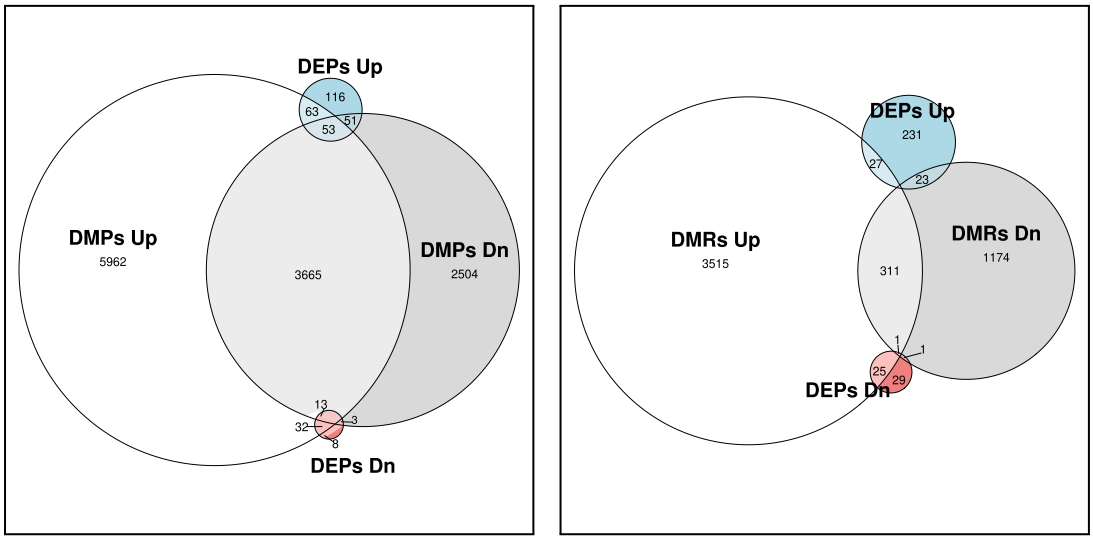


(A) Overlap of CpGs and Identified Proteins

(B) Overlap of DMPs to DEPs

Figure 6.8: Venn Diagrams of Overlaps

(A) Overlaps of CpG Genes with proteins identified using LC-MS/MS. and (B) Overlaps of DMPs and DEPs.



(A) Overlap of DMPs to DEPs

(B) Overlap of DMRs to DEPs

Figure 6.9: Venn Diagrams of Overlaps

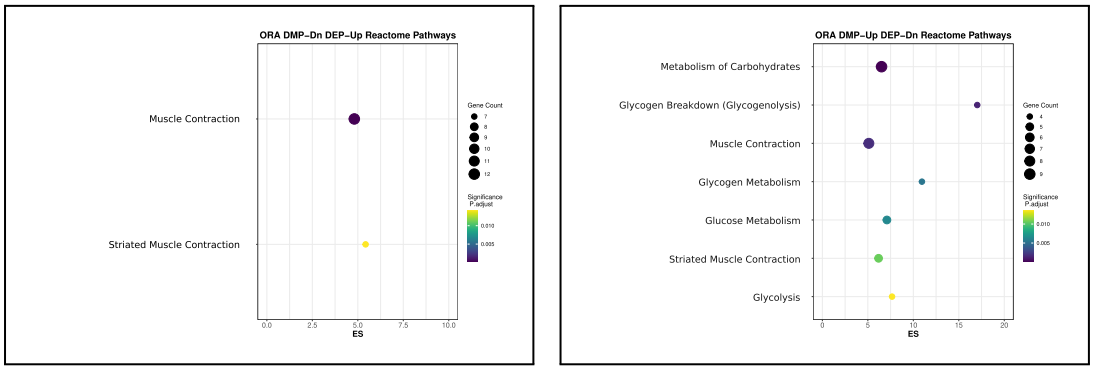
(A) Overlaps of Up and Down Regulated DEPs with Hypermethylated or Hypomethylated DMPs and (B) Overlaps of Up and Down Regulated DEPs with Hypermethylated or Hypomethylated DMRs.

Enrichment analysis of overlaps

I conducted ORA on Reactome pathways and GO-terms to identify the classes of genes over-represented in each of the overlapping categories; DMP-UP/DEP-UP, DMP-UP/DEP-DN, DMP-DN/DEP-UP and DMP-DN/DEP-DN. I identified no over-represented reactome terms using either the 63 genes/proteins from the DMP-UP/DEP-UP or the 3 genes/proteins from the DMP-DN/DEP-DN overlaps with a BH-adj-P-value < 0.005. After ORA analysis using the 32 DMP-DN/DEP-UP overlapping genes/proteins, I identified that glycogen metabolism, carbohydrate metabolism and muscle contraction reactome terms were overrepresented, and using the 51 DMP-DN/DEP-UP genes/proteins I identified only muscle contraction and skeletal muscle contraction reactome terms over-represented. I provide evidence that key genes/proteins involved in substrate metabolism have higher levels of DNA methylation and lower levels of protein expression in TI compared with TII skeletal muscle fibres.

Correlations between Methylome and Proteome

To measure the link between DNA methylation status and protein expression in TI and TII muscle fibres, I assessed the relationship between DMPs, DMRs and protein expres-



(A) ORA Reactome Terms Overlaps **(B)** ORA Reactome Terms Overlaps

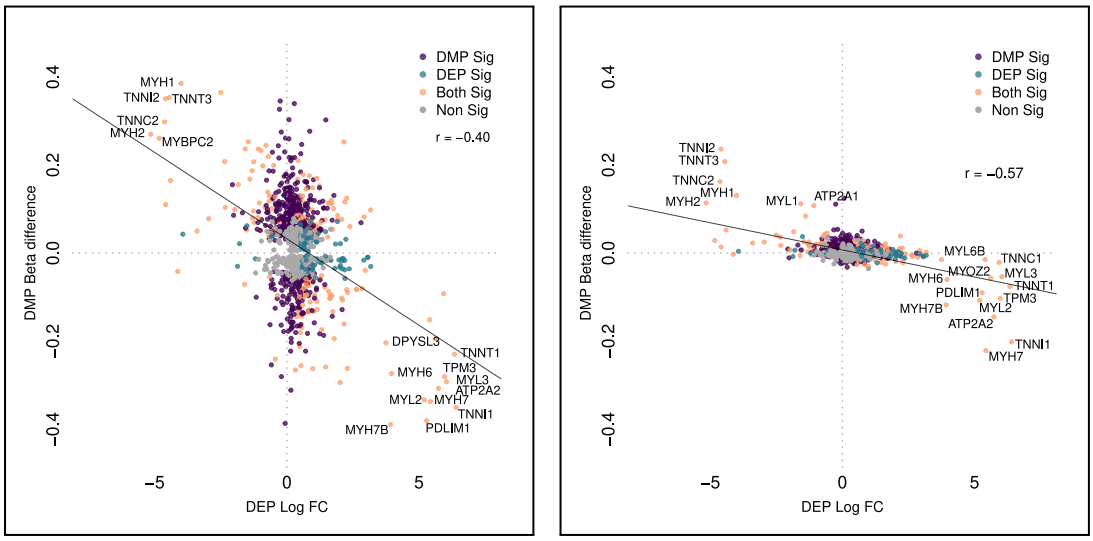
Figure 6.10: ORA of overlapping DMPs and DEPs

Reactome gene set enrichment analysis of the overlaps of DMPs and DEPs between TI and TII fibres. (A) are the reactome terms identified with ORA from the overlapping hypomethylated DMPs and the up-regulated DEPs & (B) are the reactome terms identified with ORA from the overlapping hypermethylated DMPs and the down-regulated DEPs .The circle size represents gene numbers, and the color represents the adj-P-value and the enrichment scores are on the x axis.

sion by computing Pearson's correlation coefficients. I matched the 1,143 identified proteins from my proteome data to the top CpG annotated to the gene of these proteins from my TI vs TII comparison. I then plotted the most differentially methylated position (highest beta logFC DMP) for each corresponding protein against the logFC of protein expression from my TI vs TII comparison (Figure 6.11A). As seen in Figure 6.11A the top hypermethylated DMPs between TI and TII fibres were annotated to key contractile genes such as MYH2, TNNI2, TNNT3 and these had a lower level of protein expression. Conversely the top hypomethylated DMPs in TI to TII fibres were annotated to key contractile genes such as MYH7, MYL2, TNNI1 and MYL3 and had a higher level of protein expression. There was a moderate negative Pearson's correlation between the beta logFC of the top DMPs of each gene and protein expression, r(1141) = -0.40, p = <0.001. On the EPICv2 array each multiple probes can be annotated to the same gene. We calculated the mean beta logFC across all CpGs annotated in each of the 1,143 identified proteins. We plotted the mean beta logFC across each gene to the logFC of protein expression of the matched gene in TI compared with TII fibres (Figure 6.11B). The overall level of methylation as measured by beta value logFC decreased across the gene compared to the top probes as expected but the logFC of beta values tended to zero across the entire gene. The same relationship was observed between the top mean hyper and hypomethylated genes and the top under and over-expressed proteins in TI compared with TII fibres, with

TNNI2, TNNT3 MYH1 and MYH2 hypermethylated genes having less logfc protein expression and the opposite for hypomethylated genes MYH7, TNNI1 MY7B and MYL2. The reduction in noise of methylation differences across the entire gene resulted in a stronger Pearson's correlation between differential methylation and protein expression, r(1141) = -0.57, p = < 0.001.

The previous comparison was made using all probes annotated to each gene, we were curios if there was a relationship between differentially methylated regions and protein expression in these genes. We calculated DMRs between TI and TII fibres using at least 4 consecutive CpGs that were differentially methylated. We assessed the relationship between the top DMR based on mean difference in methylation and logFC protein expression for 209 overlapping DMR/DEPs. We observed the same pattern and genes for the probe/protein analysis with the top hypermethylated DMRs having less protein expression and the top hypomethylated DMRs having more protein expression (Figure 6.12A). The same pattern emerged when we took the average mean methylation differences across all DMRs identified in each of the 209 gene/protein overlaps (Figure 6.12B). As the methylation values were taken across multiple CpGs there was an increase in the Pearson correlation for the topDMR methylation vs logFC protein expression and meanDMR methylation vs logFC protein expression, with a modest negative Pearson correlation of r(207) = -0.62, p = < 0.001 and r(207) = -0.63, p = < 0.001 respectively. There appears to be a relationship between methylation differences and protein expression differences in TI and TII muscle fibres, particularly in key genes/proteins involved in skeletal muscle fibre type specification.

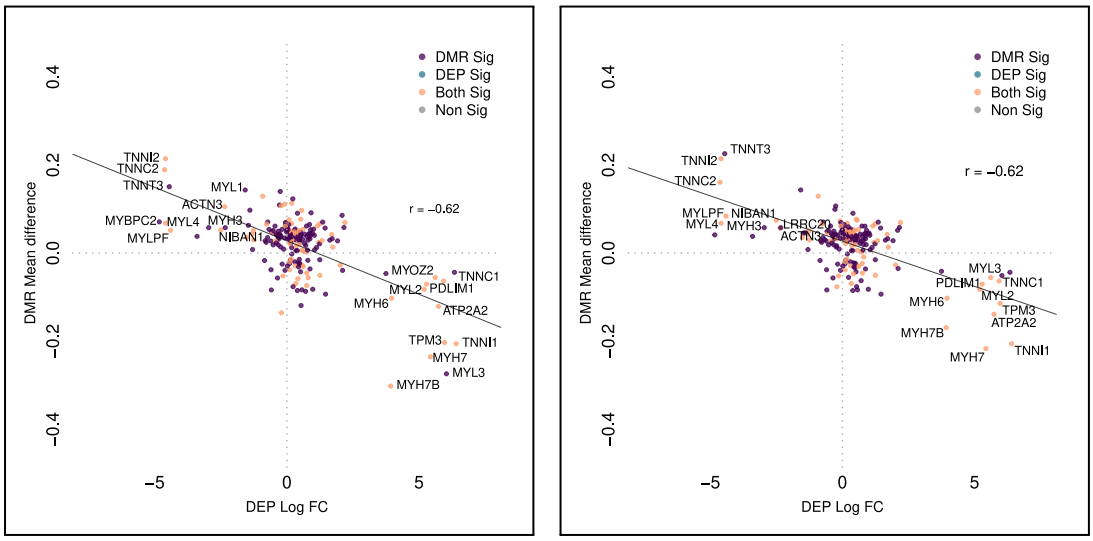


(A) DMPs Top Beta-diff vs LogFC DEPs

(B) Mean DMPs Beta-diff vs LogFC DEPs

Figure 6.11: DMPs Betdiff vs DEP LogFC

Figure 6.11 shows the relationship between beta differences and protein logFC for all identified proteins. The y axis is the beta difference between TI and TII fibres and the x axis is the logFC difference of proteins between TI and TII fibres. (A) shows the relationship between the top CpG in each identified protein and protein logFC. (B) shows the average beta difference of all CpGs across the identified protein and protein logFC. Purple dots indicate DMPs, aqua dots indicate DEPs, orange indicate that the protein was both a DMP and DEP and grey dots indicate that the protein was neither a DMP or DEP.



(A) DMRs Top Meandiff vs LogFC DEPs

(B) Average DMRs Meandiff vs LogFC DEPs

Figure 6.12: DMRs Meandiff vs DEP LogFC

Figure 6.11 shows the relationship between beta differences and protein logFC for all identified proteins. The y axis is the beta difference between TI and TII fibres and the x axis is the logFC difference of proteins between TI and TII fibres. (A) shows the relationship between the top DMR in each identified protein and protein logFC. (B) shows the average beta difference of all DMRs across the identified protein and protein logFC. Purple dots indicate DMPs, aqua dots indicate DEPs, orange indicate that the protein was both a DMR and DEP and grey dots indicate that the protein was neither a DMR or DEP.

Visualisations of methylome across DEP genes

From the data presented in the previous section it is clear that methylation levels between TI and TII skeletal muscle fibres are related to protein expression differences between the two fibre types. However, the methylation levels at single CpGs can be considerably more than taking an average across the gene. There is a trend that methylation differences across the gene have a relationship with protein expression of these genes. From Figures 6.11 & 6.12 we see that hypermethylation is correlated with decreased protein expression and hypomethylation with increased protein expression, however this is not always the case. The relationship between methylation and protein expression between TI and TII fibres varies with some regions and genes having hypermethylation differences yet higher protein expression in TI compared with TII muscle fibres. To visualise the underlying pattern of CpG methylation in genes from two DEPs, I generated human genome tracks of MYH7 and MYH2.

Figure 6.13 shows the Myosin 7 gene which was determined to be hypomethylated with increased protein expression in TI compared with TII fibres. We took every EPICv2 CpG probe annotated to the MYH7 gene and plotted their location and the methylation beta values across the gene. Comparing the DNAm pattern across MYH7 and MYH2 there is a distinct region at the start of each gene where there is hypomethylation (Figures 6.13 &C.9). Across the gene body there was a hypomethylation of MYH7 in TI compared with TII fibres as shown by plotting beta values for each sample along all CpGs annotated to MYH7 (Figure 6.13). In MYH2 the pattern of methylation across the gene was the reverse of the MYH7 gene with a distinct hypermethylation across the gene in TI compared with TII fibres.

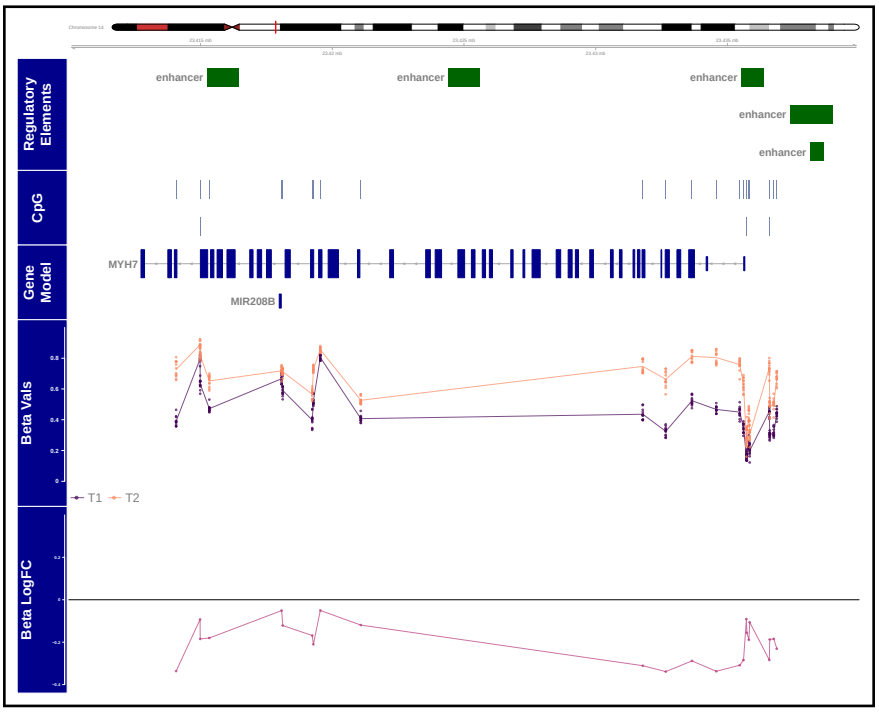


Figure 6.13: MYH7 Genomic Visualisation

Figure 6.13 is a visualisation of methylation values as measured with the EPICv2 across the myosin 7 gene. The genome visualisation shows all the CpGs (blue bars) across the MYH7 gene on chromosome 14. Enhancers are indicated by green blocks, the gene model is indicated by darkblue genetracks. The orange line represents the beta value of TII fibres from each of the 12 samples and the purple line represents the beta value of TI fibres from each of the 12 samples. The pink line represents the beta difference between TI and TII fibres at all CpGs across the genomic region.

6.4 Discussion

In this chapter I measured protein expression differences between TI and TII pooled skeletal muscle fibres from both males and females. I was able to compare my previous DNA methylation data with the proteomic data to assess the correlation of DNA methylation and protein expression in the same sample.

I was able to assess the quality of fibre pooling by measuring the percentage of MYH7, MYH2, MYH1 and MYH4 in the samples. A 80% cut off for MYH7 and MYH2 in a sample is commonly used to define a relatively pure fibre sample [209]. The pooled Type I samples showed on average > 95% MYH7 isoform and the pooled Type II samples showed on average >90% MYH2/MYH4/MYH1 isoforms (Figure 6.4), these results confirm that we were able to isolate and pool relatively pure Type I and Type II fibre populations. It also confirms that the Type II sample is representative of a Type IIa/IIx population. These results also are in agreement with previously reported proteomic studies of pooled skeletal muscle fibres [210]. I also extended the study of myosin isoforms in muscle fibre types to show that myosin light chains are an important protein family in skeletal muscle fibre types and that overall the myosin heavy chain and light chain proteins comprise between 20-40% of the proteome signal as measured by mass spectrometry.

I identified 339 differentially expressed proteins at a BH-adj-P-value < 0.001 in Type I compared with Type II fibres. 154 of these proteins showed over-expression with a logFC of > 1 and 34 showed under-expression with a logFC < -1 in Type I fibres. The over-expressed proteins in Type I fibres were over represented in aerobic energy metabolism GO terms and Reactome pathways such as cellular and aerobic respiration, respiratory electron transport and mitochondrial function whereas the down-regulated proteins were involved in glycogen and carbohydrate metabolism (Figures 6.6, C.5, C.6 & 6.7). I then compared DNAm differences to protein expression differences between Type I and Type II muscle fibres. There was a modest negative correlation between methylation differences and protein expression differences in Type I and Type II fibres suggesting DNAm plays a role in downstream protein expression differences in skeletal muscle fibres (Figures 6.11 & 6.12).

Here we provide further evidence that the main proteins differentially expressed between Type I and Type II muscle fibres are major contractile proteins MYH7, TNN1, TNNI, MYL3, MYL2 (TI fibres), MYH2, TNNC2, TNNT3, MYL4 and MYLPF (TII fibres) (Figure 6.5) as previously reported in both single fibre [209] and pooled fibre-type proteomics analyses [210]. We also confirm the findings that PDLIM1 has higher expression in Type I fibres [210], and interestingly we show PDLIM7 has higher expression in Type II fibres, but this may confirm that these Type II fibres are indeed a mixture of Type IIa and Type IIx muscle fibres as it was previously reported that PDLIM7 has higher protein expression levels only in Type I vs Type IIx comparisons [209, 210]. We also show that DEPs between Type I and Type II fibres are enriched in substrate metabolism and skeletal muscle contractile GO terms and reactome pathways as reported by Deshmukh and others [210]. We provide further evidence that Type I muscle fibres are enriched for proteins involved in aerobic and fatty-acid based metabolism and that Type II fibres are enriched for proteins involved in glycolysis, glucose metabolism and carbohydrate metabolism (Figure 6.6, C.5, C.6 & 6.7).

I have compared methylation differences to protein differences in the same pooled Type I and Type II skeletal muscle fibre samples. I show evidence of overlaps between hypermethylated DMPs/DMRs and down-regulated DEPs and overlaps between hypomethylated DMPs and up-regulated DEPs identified between Type I and Type II skeletal muscle fibres (Figures 6.9). The overlapping hypermethylated genes/down-regulated proteins in Type I fibres were enriched for glycogen and carbohydrate metabolism reactome pathways and the overlapping hypomethylated genes/up-regulated proteins were enriched for muscle contraction based reactome pathways (Figures 6.10). These findings are in line with the long held classification scheme that Type I fibres are oxidative and Type II fibres populations are glycolytic in nature [59]. This analysis provides evidence that Type I and Type II fibres are differentially methylated in key genes important for either aerobic/anaerobic metabolic systems and that these methylation differences are reflected in protein expression differences. We also show that the direction of methylation differences between Type I and Type II fibres is not always reflective of the direction of protein expression (Figures 6.11 & 6.12).

The combined study of protein and DNAm differences in the same skeletal muscle fibre-type samples provide a unique look at the connection between DNAm and protein

expression using a systems-biology approach. This data provides a resource of methylation and protein values in Type I and Type II muscle fibres that we hope can serve as a reference for researchers studying skeletal muscle biology. In the next chapter we extend this approach to assess the DNA methylation and protein changes between Type I and Type II muscle fibres after 12 weeks HIIT.

7 Results 4: Fibre-Type-Specific Methylome and Proteome Changes in Response to 12 Weeks of High Intensity Interval Training

7.1 Introduction

Skeletal muscle has a unique ability to respond to physical activity and exercise leads to a number of adaptations within skeletal muscle at both physiological and cellular levels [93, 317]. Evidence suggests that Type I slow and Type II fast muscle fibres undergo specific fibre-type changes in response to different types of exercise [258]. For example, resistance exercise promotes fibre-specific size increases [257], and aerobic exercise may result in shifts in fibre-type proportions [258]. The response of muscle to different aerobic training interventions has been extensively assessed [318]. High intensity interval training (HIIT), exercise that alternates between intense bouts and recovery periods, has been associated with improved mitochondrial function, insulin sensitivity and fitness [319]. The intensity of aerobic exercise can influence the recruitment patterns of skeletal muscle fibre types [320], and HIIT has been shown to decrease the quantity of Type IIx fibres [219] and cause enzymatic changes in Type IIa fibres [259]. Considering that intensity of aerobic training causes fibre-type-specific responses, it is important to study the response of skeletal muscle to HIIT at the fibre-type level.

To date studies have investigated DNAm changes due to both resistance and aerobic exercise interventions in whole skeletal muscle [321, 322, 181, 263, 177]. Our group has previously studied the effect of 4 weeks or 12 weeks HIIT on DNAm in whole human skeletal muscle samples in either a cohort of males [263], or a cohort of males and females [181]. To understand the relationship between methylation and protein regula-

tion after HIIT our group studied both the proteome and methylome in whole human skeletal muscle samples [181, 263]. Recent work has shown that the proteome responds differently to aerobic exercise (including interval type training) in Type I and Type II muscle fibres [210, 323]. However, DNAm pre and post exercise has not yet been investigated on a fibre-type level. No studies to date have investigated the effect of HIIT on both the methylome and proteome in a fibre-type specific manner. My method enabled the comparison of the differential shifts in methylation and protein regulation in the same pooled human skeletal muscle fibre samples before and after HIIT.

This chapter addresses the following aims

1. To identify differential shifts in DNAm in Type I and Type II muscle fibres before and after 12 weeks of HIIT in male and females.

2. To identify changes in the proteome in Type I and Type II muscle fibres before and after 12 weeks of HIIT in male and females.

7.2 Methods

Participants were recruited to train three times a week for 12 weeks on a Velotron, to study the effect of 12 weeks HIIT on DNAm and protein regulation in TI and TII muscle fibres. Training was comprised of 6-14 two-minute intervals interspersed with one minute active recovery. Intervals were performed above lactate threshold (LT) with active recovery periods performed at 25-60 W depending on participants pre-testing results. Graded exercise tests (GXT) were performed to measure peak VO_{2peak}, peak aerobic power (W_{peak}) and LT before, at four-weeks, at eight-weeks and after the intervention. Training intensities were reset every four weeks based upon results from the four-week and eight-week GXTs. Participants were provided with a 48-hour controlled diet before each biopsy. A schematic of the intervention is provided in Figure 7.1. For full intervention details please see the Participant, Methods and Ethics section.

Overall 15 (8 = male & 7 = female) participants were recruited to this study. Of the 15 participants one participant dropped out after 3 weeks of HIIT. The 14 remaining participants completed all training sessions. Due to unexpected circumstances between

the final biopsy and follow up testing, two participants (acute injury & COVID infection) were unable to attend post testing sessions. We were able to successfully isolate pooled fibre samples from 12 participants (7 = male & 5 = female) that have been used for the assessment of methylome and proteome differences between TI and TII fibres after 12 weeks HIIT. Physiological adaptations to exercise have been calculated using all available data from participants enlisted in the overall study.

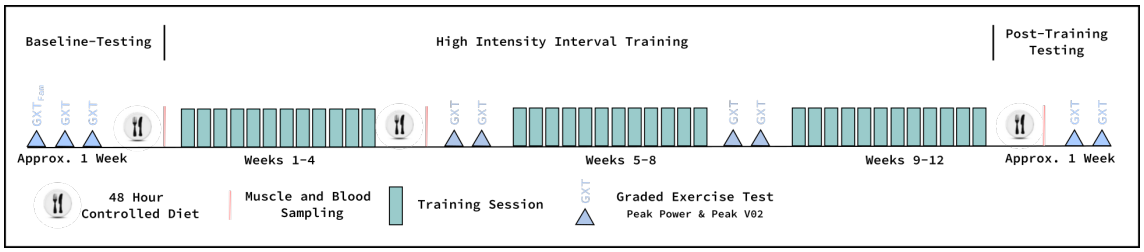


Figure 7.1: Method Schematic

An overview of the 12 week HIIT study design.

7.2.1 Participant Characteristics

As a reminder, 15 participants (8 = male & 7 = female) muscle samples were used for whole muscle DNA extraction and a total of 12 participants (7 = male & 5 = female) muscle samples were used to isolate fibre-type specific DNA. At the time of recruitment participants median age was 34 (range = 25-42). The breakdown of participant sex and age are provided again in Table 7.1.

Table 7.1: Participant Characteristics

Participants		**n=15**
Sex	Male	8
	Female	7
Age	Median	34
	Min-Max	25-42

7.3 Results

7.3.1 Physiological Adaptations

Table 7.2 shows the mean and standard deviation of Pre and Post VO_{2peak}, LT and W_{Peak}.

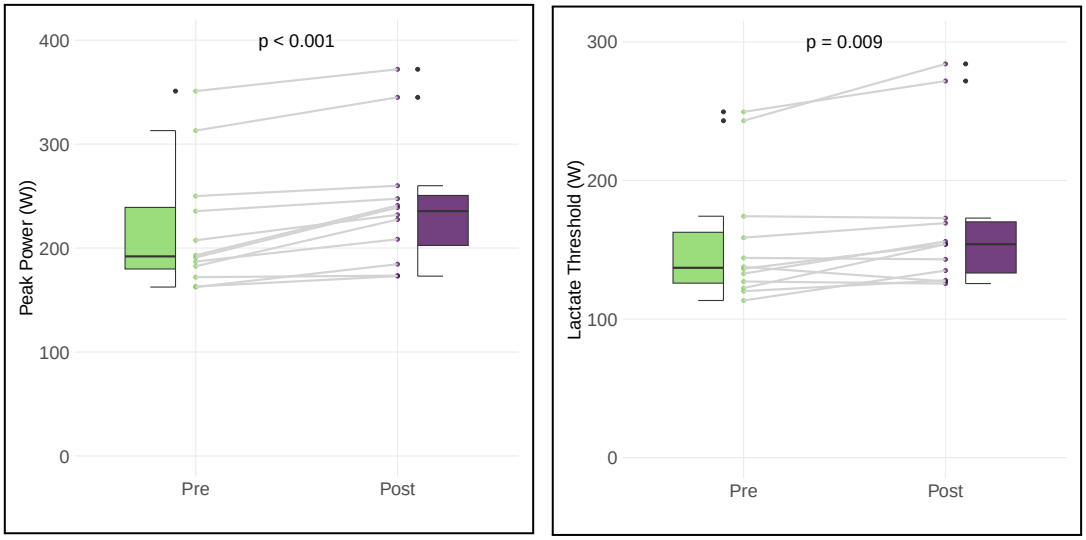
Table 7.2: Intervention Results

Timepoint	Pre	Post	P-Val
	Mean (SD)	Mean (SD)	
VO_{2peak}	38.0 ml.min.kg^{-1}(7.4)	44.4 ml.min.kg^{-1}(8.3)	P=0.005
Lactate Threshold	155.0 W (45.9)	168.4 W (53.6)	P=0.009
W_{Peak}	217.3 W (60.3)	242.0 W (61.8)	P<0.001

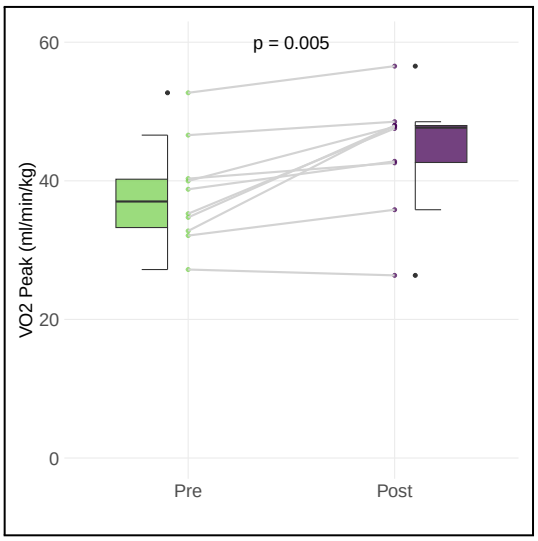
Figure 7.2A shows the difference in W_{Peak} due to 12 weeks HIIT. Participants improved W_{Peak} on average by 25 Watts after 12 weeks HIIT with a 95% confidence interval of 15 to 35 Watts. This increase represents an average improvement of 12% on pre HIIT W_{Peak}. The p-value for a paired two sided t-test of the null hypothesis that the true average was zero Watts was P<0.001.

Figure 7.2B shows the difference in power at LT due to 12 weeks HIIT. Participants improved power at LT on average by 12 Watts after 12 weeks HIIT with a 95% confidence interval of 4 to 21 Watts. This represents an average improvement of 7% on pre HIIT power at LT. The p-value for a paired two sided t-test of the null hypothesis that the true average was zero Watts was P=0.009.

Figure 7.2C shows the difference in VO_{2peak} due to 12 weeks HIIT. Participants improved VO_{2peak} on average by 6.3 ml/min/kg after 12 weeks HIIT, with a 95% confidence interval of 2.4 to 10.3 ml/min/kg. This represents an improvement of 15.7% on pre VO_{2peak}. The p-value for a paired two sided t-test of the null hypothesis that the true average was zero ml/min/kg oxygen uptake was P = 0.005.



(A) $W_{peak}(n = 12)$ **(B)** LT (n=12)



(C) VO_{2peak}

Figure 7.2: Physiological Adaptations Pre vs Post

Peak aerobic power (n = 12), power at lactate threshold (n = 12) and peak rate of oxygen consumption (n = 10) before and after 12 weeks HIIT. Green indicates pre sampes and purple indicates post samples.

7.3.2 Fibre-type proportions before and after exercise

To assess if HIIT can cause a shift in fibre type, I compared the proportions of TI and TII fibres before and after the intervention. We calculated the proportions of TI and TII fibres in the samples with the following equation TI-proportion = number of TI fibres / ((number of TI fibres) + (number of TII fibres)). TIIx fibres were excluded as we determined that our ponceau +ve MYH7/MYH2 -ve fibres are likely due to insufficient antibody binding

and not a true TIIx population. Likewise hybrid fibres were not included as we cannot be certain that they do not represent contaminated fibres.

Table 7.3 summarises the average proportion of fibres before and after the intervention. We observed no change in fibre type proportion after HIIT. There is no statistical support in a true shift in the mean for TI (p = 0.37) or TII (p = 0.37) fibre proportions after HIIT using a Wilcoxon non-parametric rank test. Figure 7.3 shows individual fibre-type proportions before and after exercise confirming that there is variation in fibre-types after 12 weeks HIIT.

Table 7.3: Fibre Numbers

Proportions	Pre	Post	P-Val
Type I Fibres	48.4	51.6	P=0.37
Type II Fibres	51.6	48.4	P=0.37

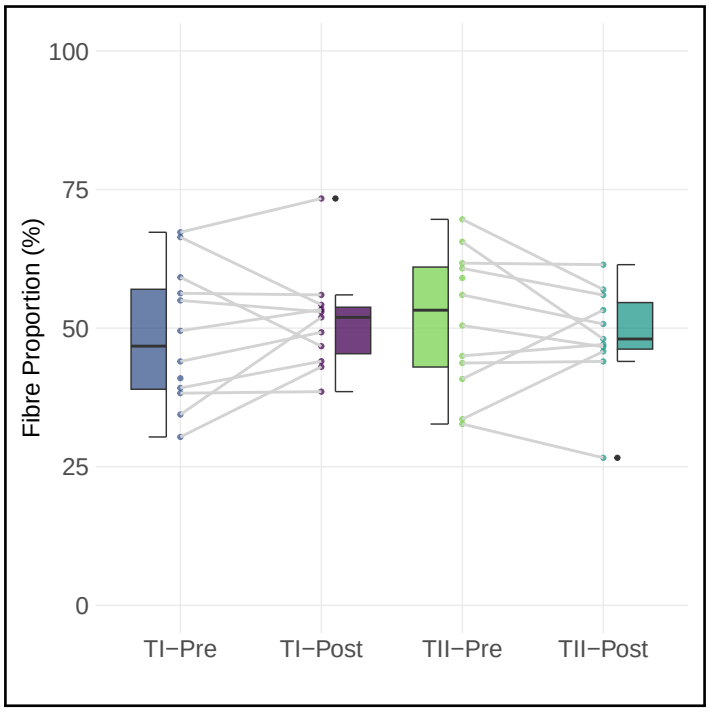


Figure 7.3: Fibre Type Proportions

The percentage of Type I and Type II fibres of each participant before and after HIIT.

7.3.3 Myosin isoform Content of pooled samples

We confirmed the purity of sample pooling by assessing myosin heavy chain content in the 12 weeks pooled Type I and Type II fibres. We show in agreement with our previous chapter that MYH7 positive Type I fibres express mostly MYH7 (on average 98% MYH7 expression), and MYH2 positive fibres express mostly MYH2 (on average 76% MYH2) Our Type II fibres show a high proportion of MYH1 and MYH4 with on average 8% and 6% expression respectively (Figure 7.4). The two Type IIx pooled post samples also show the same pattern as the pre samples, indicating a mixed fibre group. Sample F31-12wk-T2 had a higher than expected MYH7 content and was deemed contaminated and removed from the subsequent analyses. Sample F31-12wk-T1 also had a higher than expected quantity of MYH2, we noted an error with fibre labelling during fibre type pooling of this sample. This sample had >90% MYH7 and was included as a Type I fibre in the analyses.

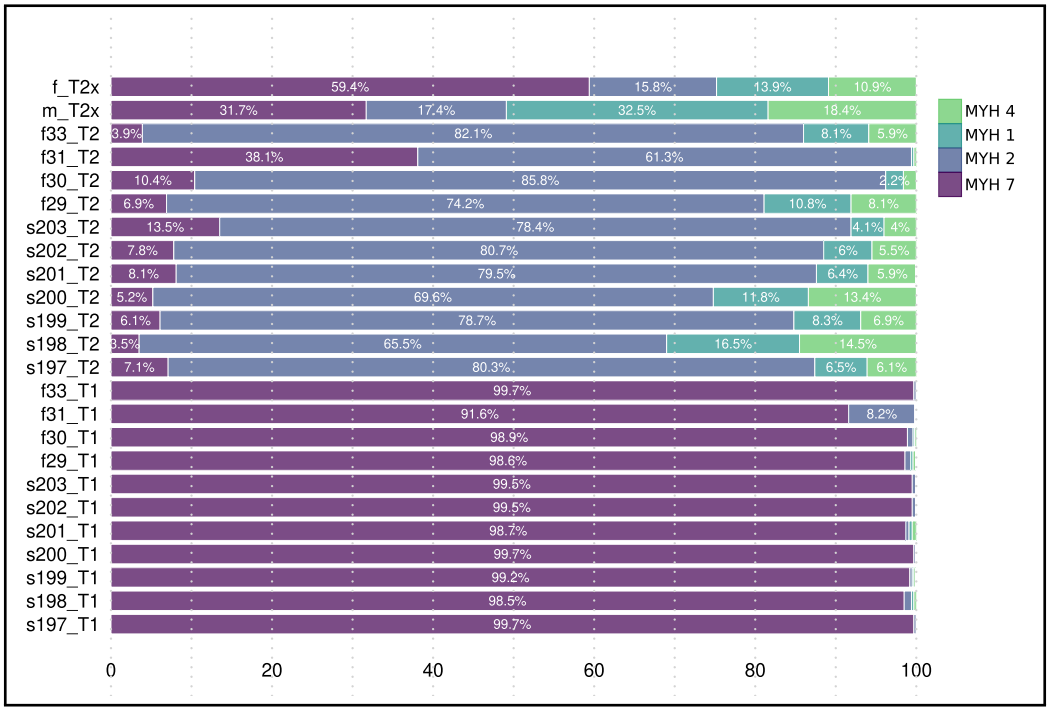


Figure 7.4: Myosin Isoform Content Post Samples

The proportions of MYH7, MYH2, MYH1 and MYH4 identified in post HIIT samples. Percentages were calculated as a myosin isoform intensity divided by the quantity of MYH7, MYH2, MYH1 and MYH4.

7.3.4 DNAm analysis

I used linear-mixed-effect-models in limma, to estimate differences in methylation of CpGs between TI and TII muscle fibres and between whole muscle samples before and after HIIT.

A comparison within TI fibres pre and post HIIT (n=12vs11) revealed no DMPs at a BH-P-value of < 0.005 and two CpGs that had a BH-P-value <0.05 (Table 7.4). NPDC1 (P = 0.009) and SENP1 (P = 0.02) were both identified as hypermethylated in TI fibres after HIIT at a relaxed BH-P-value <0.05. The comparisons of TII-post vs TII-pre fibres (n=12vs10) (Table 7.5) or WM-post vs WM-pre (Table 7.6) revealed no differentially methylated CpGs at a BH-P-value < 0.05.

In chapter 4, I presented data showing numerous CpGs with beta value differences between 10-40% between TI and TII fibres (Figure 5.8). In comparison, beta values differences within TI and whitn TII samples after HIIT were between 0-5% (Figures 7.5). In TII-pre vs TII-post fibres 1158 CpGs showed a larger than 5% beta change whereas TI-pre vs TI-post and WM-pre vs WM-post had 596 and 442 CpGs with greater than or less than 5% beta change (Figures 7.5).

Table 7.4: Differentially expressed CpGs in TI-pre compared to TI-post fibres

	TI-Pre	**Vs**	**TI-Post**
BH P-value Threshold	Up	Non-DMP	Down
BH P < 0.001	0	838705	0
BH P < 0.005	0	838705	0
BH P < 0.05	2	838703	0

Table 7.5: Differentially expressed CpGs in TII-pre compared to TII-post fibres

	TII-Pre	**Vs**	**TII-Post**
BH P-value Threshold	Up	Non-DMP	Down
BH P < 0.001	0	838705	0
BH P < 0.005	0	838705	0
BH P < 0.05	0	838703	0

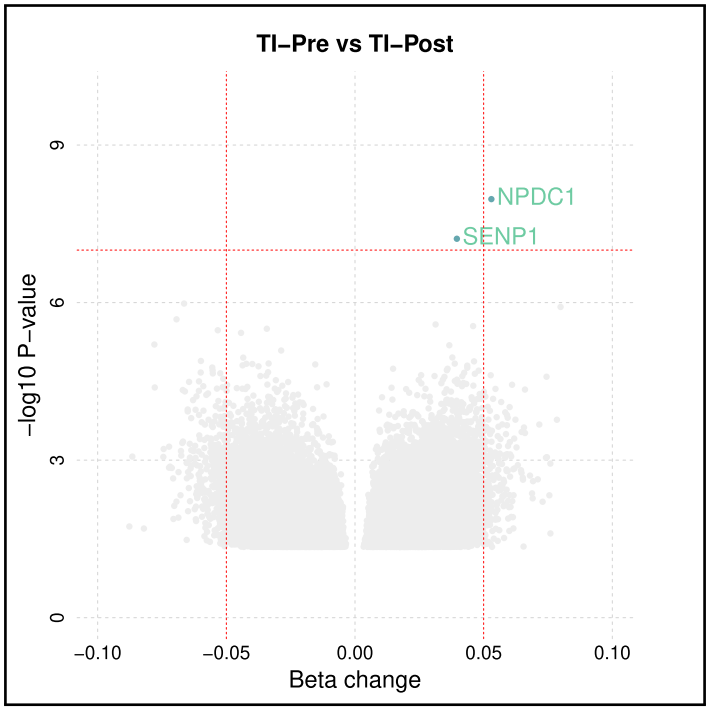


Figure 7.5: TIpre vs TI-post

Volcano plot of the top 50000 CpGs with beta change on the x-axis and -log10 P-value on the y-axis. The horizontal line represents the cutoff of the corresponding adj-P-value <0.05 and the vertical lines represent the cutoff of the Beta Change 0.05 and -0.05.

Table 7.6: Differentially expressed CpGs in WM-pre compared to WM-post fibres

	WM-Pre	**Vs**	**WM-Post**
Significance Threshold	Up	Non-DMP	Down
BH P < 0.001	0	838705	0
BH P < 0.005	0	838705	0
BH P < 0.05	0	838703	0

7.3.5 Proteomic Analysis

Differentially Expressed Proteins

To estimate differences in protein expression between TI and TII muscle fibres before and after 12 weeks HIIT, mixed effects linear models were performed using limma in the R statistical environment on normalised, imputed and log2 transformed MaxLFQ values. Sample s197-pre-T2 was removed from the analysis due to the reduced number of proteins identified.

A comparison of TI-pre fibres and TI-post fibres (n=12vs11) revealed no DEPs at a BH-P-value of < 0.05 (Table 7.7). The comparison of TII-post vs TII-pre fibres (Table 7.8) revealed 123 DEPs at a BH-P-value < 0.05. 120 of these DEPs were up-regulated after the intervention with four DEPs having a BH-P-value < 0.005. Only three DEPs were down-regulated post 12 weeks HIIT in TII fibres, none of which had BH-P-Value < 0.005. The top up-regulated proteins in TII fibres after HIIT were BCS1L, LDHD and CKMT2. The three down-regulated proteins were PGM5, PDLIM3 and HNRNPU (Figure 7.6). All statistics generated with limma for the pre-post protein analysis can be provided on request.

Table 7.7: Differentially expressed DEPs in TI-pre compared to TI-post fibres

	TI-Post	**Vs**	**TI-Pre**
Significance Threshold	Up	Non-Sig	Down
BH P < 0.001	0	1226	0
BH P < 0.005	0	1226	0
BH P < 0.05	0	1226	0

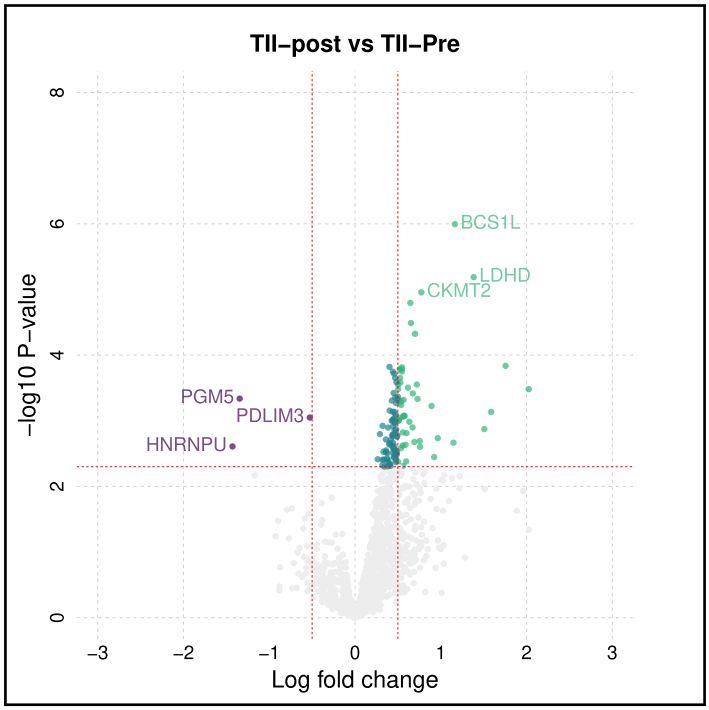


Figure 7.6: Volcano Plot T1vsT2

Volcano plot of the differentially expressed proteins with log2 FC on the x-axis and -log10 P-value on the y-axis. The horizontal line represents the cutoff of the corresponding adj-P-value <0.05 and the vertical lines represent the cutoff of the log2 FC 0.5 and -0.5.

Table 7.8: Differentially expressed DEPs in TII-pre compared to TII-post fibres

	TII-Post	Vs	TII-Pre
Significance Threshold	Up	Non-Sig	Down
BH P < 0.001	0	1226	0
BH P < 0.005	4	1222	0
BH P < 0.05	120	1103	3

Up-regulated proteins are overrepresented in key metabolic functions

The 120 up-regulated DEPs are mitochondria related proteins, such as proteins involved in mitochondrial complexes, carriers and chaperones (Table 7.9). All 120 up-regulated DEPs in TII-post compared with TII-pre fibres were over represented in key metabolic reactome pathways as measured by ORA (Figure 7.7). I identified 19 reactome pathways passing a < 0.05 BH-P-Value cutoff, with 16 and 12 passing <0.005 and <0.001 BH-P-value respectively. The top three reactome pathways were related to the TCA cycle and electron transport chain. To confirm ORA results, I conducted FCS using all proteins and statistics to identify reactome pathways. The same pathways were over-represented among proteins as those identified using ORA (appendix Figure D.1). All reactome pathways and statistics can be provided on request.

Table 7.9: T2post vs T2pre DEPs

	logFC	AveExpr	t	P.Value	adj.P.Val	B
BCS1L	1.16	-1.86	5.71	1.01E-06	0.001	5.49
LDHD	1.38	-1.46	5.15	6.49E-06	0.004	3.78
CKMT2	0.77	6.19	4.99	1.11E-05	0.005	3.29
MT-CO2	0.65	4.79	4.87	1.60E-05	0.005	2.95
COX5A	0.65	3.73	4.65	3.25E-05	0.008	2.30
EPHX1	0.70	1.53	4.53	4.73E-05	0.010	1.96
TTR	1.76	-1.08	4.18	1.46E-04	0.017	0.93
MTCH2	0.40	0.25	4.16	1.51E-04	0.017	0.90
NDUFA6	0.55	1.91	4.16	1.54E-04	0.017	0.88
UQCRFS1	0.52	3.35	4.14	1.63E-04	0.017	0.82

Table 7.9: T2post vs T2pre DEPs(continued)

	logFC	AveExpr	t	P.Value	adj.P.Val	B
HSPD1	0.54	2.16	4.14	1.64E-04	0.017	0.82
MT-ND4	0.55	1.00	4.11	1.78E-04	0.017	0.75
TUFM	0.44	3.04	4.11	1.79E-04	0.017	0.74
NDUFS2	0.46	2.50	4.09	1.91E-04	0.017	0.68
NDUFV1	0.47	2.89	4.04	2.22E-04	0.017	0.54
IDH3A	0.52	1.61	4.04	2.23E-04	0.017	0.54
PTGES2	0.48	1.09	3.99	2.58E-04	0.017	0.41
NDUFS8	0.53	2.25	3.99	2.62E-04	0.017	0.39
TIMM44	0.72	-1.10	3.96	2.80E-04	0.017	0.33
NDUFA8	0.50	2.83	3.96	2.81E-04	0.017	0.33
FBP2	0.62	2.75	3.93	3.12E-04	0.018	0.23
NDUFS3	0.50	2.49	3.91	3.29E-04	0.018	0.19
NDUFB11	2.02	1.58	3.91	3.29E-04	0.018	0.18
UQCRC1	0.45	4.90	3.86	3.77E-04	0.018	0.06
XPO1	0.67	-3.34	3.86	3.84E-04	0.018	0.04

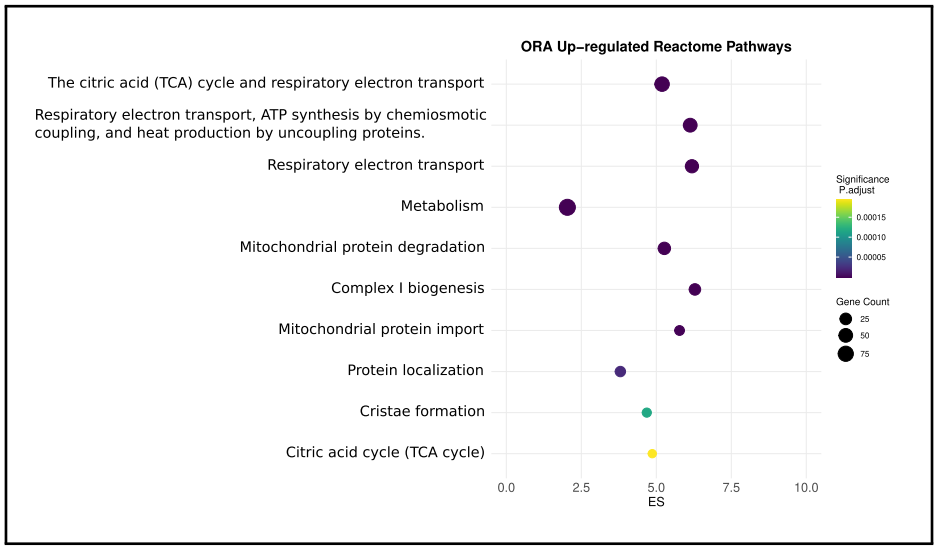


Figure 7.7: ORA Reactome

Top 10 reactome terms as measured by over representation analysis of the 120 up-regulated DEPs. The circle size represents gene numbers, and the coluor represents the adj-P-value.

TII fibres shift toward TI fibres in key mitochondrial proteins

Using the 120 up-regulated proteins, I generated a heatmap of the log2 imputed data values. The TII-post samples clustered with higher similarity to the TI-pre and TI-post samples (Figure 7.8) suggesting that the expression pattern of the 120 DEPs was more similar to TI fibres after 12 weeks of HIIT. To further assess the up-regulated DEPs in TII-post fibres, I visualised the distribution or log2 intensities in TI and TII samples before and after HIIT. Log2 intensities of BCS1L and LDHD in TII-post fibres had similar values as TI-pre and TII-post samples compared with TII-pre samples (Figure 7.9).

There was a larger variation in log2 intensities only within the TII-pre samples (Figure 7.9). To find an explanation for the observed results I calculated Pearson's correlation statistics for TII-pre log2-intensities of both BCS1L and LDHD and LT, W_{peak}, VO_{2peak}, TI and TII fibre-type proportions and overall sample protein count. There was only evidence for a correlation between LDHD log2 intensity and protein count of samples (r(10) = 0.59, p = 0.04).

Next I checked the number of imputed values for both LDHD and BCS1L amongst the four sample groups. Interestingly in both proteins only TII-pre samples had missing values that were imputed with seven and samples missing for BCS1L (Figure D.2). The imputed samples were amongst the lowest log2 intensity values (Figure D.2), as I opted

for a MNAR approach to imputation which drew random samples from a down shifted Gaussian distribution. To confirm if imputation effected correlations, I calculated Pearson's correlations using only the non-missing TII-pre samples. I did not find evidence for any correlations when using non-imputed values.

On the consideration of the evidence, the choice to impute protein values using a MNAR strategy was warranted, and the missing values likely represent low abundance of BCS1L in these samples.

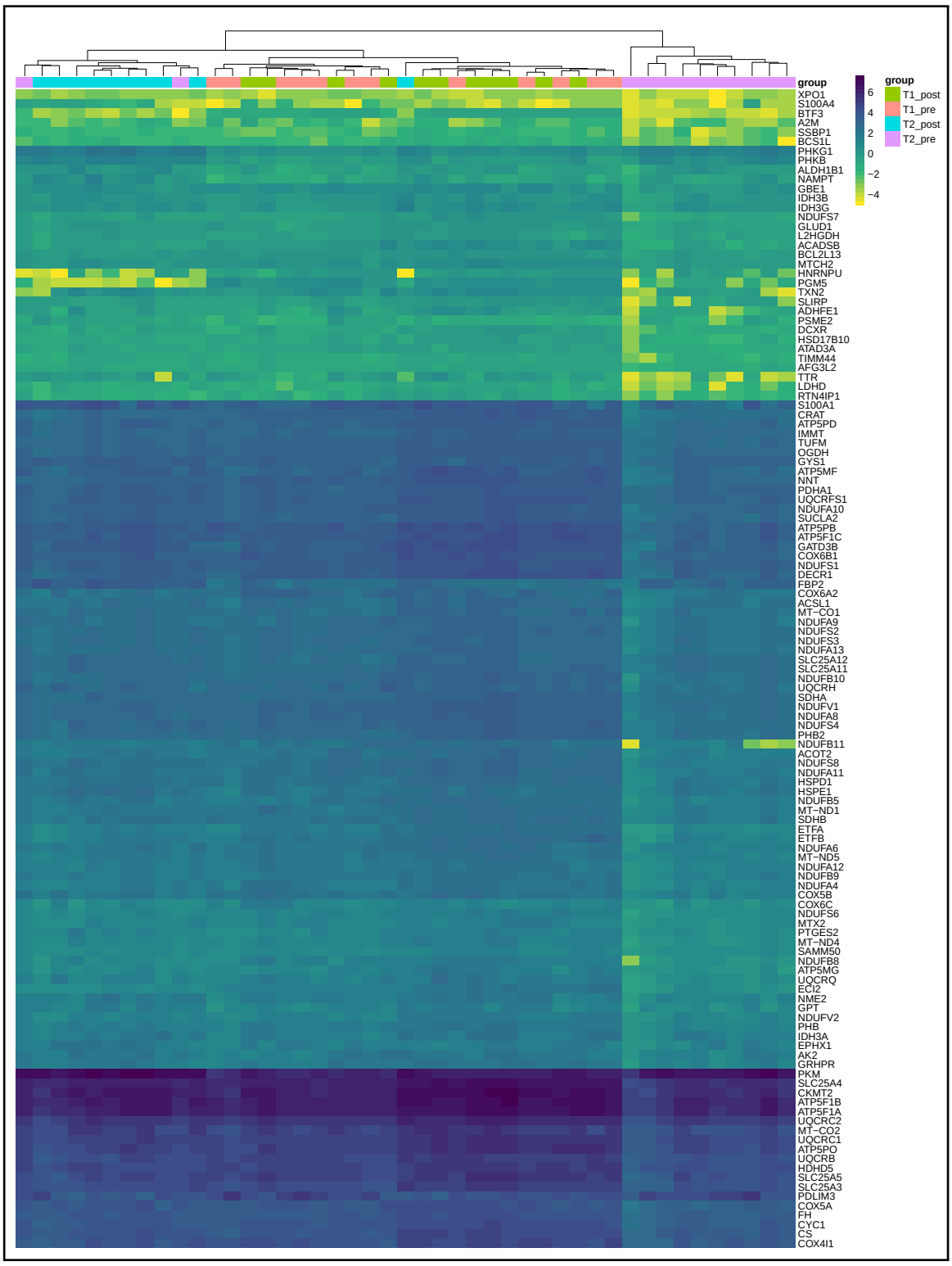
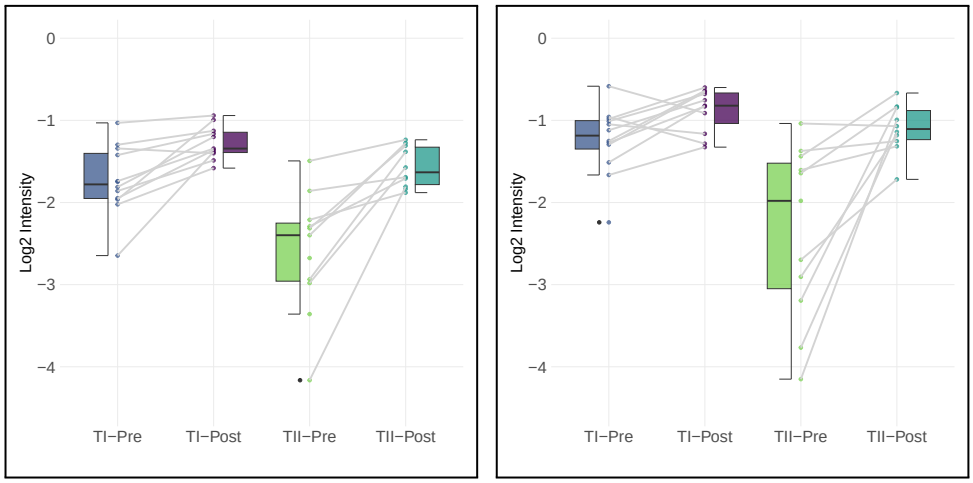


Figure 7.8: Heatmap of 120 DEPs in TII-post vs TII pre Fibres

Heatmap hierarchical clustering analysis of the top 120 up-regulated DEPs in TII-pre vs TII-post samples. The colour-coded scale (purple = up-regulated and yellow = down-regulated) are the min-max normalised log2 maxLFQ intensities. Type I post (green), Type I pre (red), TII post (blue) and T2 pre (purple) samples are colour coded and indicated with appropriate bars.



(A) BCS1L **(B)** LDHD

Figure 7.9: Log2 intensities of BCS1L and LDHD

Log2 intensities of BCS1L and LDHD in Type I and Type II fibres before and after 12 weeks HIIT.

7.4 Discussion

In this chapter I measured both DNAm and protein differences in TI and TII fibres before and after 12 weeks HIIT in both males and females. To my knowledge this is the first assessment of DNAm before and after an exercise intervention in Type I and integrate the Type II muscle fibres. 12 weeks HIIT resulted in changes in LT, W_{peak} and VO_{2peak} at the group level (Figure 7.2). There were no changes in fibre-type proportions as a result of 12 weeks HIIT. The purity of pooled TI and TII-post samples was high, with an exception of one sample labelled as contaminated which was removed form the analysis (Figure 7.3). I found two hypermethylated CpGs in TI-post compared with TI-pre samples at a relaxed BH-P-value <0.05 and comparisons of TII and WM post vs pre samples revealed no DMPs. There were no DEPs with a BH-P-Value < 0.05 when comparing TI-post and TI-pre fibres. In this chapter, I report 123 differentially regulated proteins in TII-post compared with TII-pre fibres at a BH-P-value < 0.05 (Table 6.2). 120 of these proteins were up-regulated after 12 weeks of HIIT in the TII fibre population. These 120 proteins were over-represented in key metabolic pathways and mitochondrial proteins (Figure 7.7).

The results presented here show 12 weeks HIIT leads to improvements in VO_{2peak}, W_{peak} and LT, in agreement with previous studies conducted in our laboratory [324].

We utilised a within participant design with no separate control group, however other studies that have included a control group have also shown improvements in VO_{2peak} after 12 weeks HIIT cycling [325]. The fibre-type proportion changes with 12 weeks HIIT show higher individual variation (Figure 7.3). Fibre type proportions were calculated using the data from dot blot fibre typing of 288 single muscle fibres before and after HIIT. It has been demonstrated that fibre type transformations occur in Type II populations of fibres and not in TI fibres [326]. Our laboratory previously reported no changes with fibre type proportions with 12 weeks HIIT [324]. The results reported here are likely variation from different muscle bundles used for fibre isolations. Future experiments will be performed to compare the fibre-typing results presented here with traditional fibre typing using immunohistochemistry.

I identified NPDC1 and SENP1 as the only differentially methylated CpGs after HIIT specifically in TI-post compared with TI-pre fibres. NPDC1 is a neural factor involved in proliferation and differentiation of neuronal cells [327]. Conflicting results have been reported on the presence of NPDC1 in skeletal muscle [328, 329], but NPDC1 mRNA expression has been reported in both endothelial cells and adipocytes [330]. It is interesting to speculate that the difference in methylation of NPDC1 between Type I fibres before and after exercise might indicate the presence of endothelial cells which are more numerous along Type I fibres after exercise. It is also possible that the NPDC1 is a potential false positive as I only identified one CpG differentially methylated at a relaxed BH adjusted-P-value. I did not identify NPDC1 in the proteomic experiment and follow up experiments could be conducted to confirm NPDC1 protein levels. SENP1 on the other hand has been reported to lead to deSUMOylation of KLF5 influencing lipid metabolism in skeletal muscle cells [331]. Recent evidence suggests that accumulated lactate inhibits SENP1, influencing the cell cycle and proliferation [332]. I identified only one CpG within SENP1 that was hypermethylated in TI-post compared with TI-pre fibres ($p = 0.03$, betadiff = 0.39). Although it would be interesting to speculate on changes in SENP1 with exercise, these results provide little support of methylation differences at this gene. I did not find support that 12 weeks of HIIT alters the methylome within TI or within TII muscle fibres. These findings align with results presented from our laboratory that compared whole muscle before and after 12 weeks HIIT [263].

At the protein level, I identified differential regulation of proteins after 12 weeks of HIIT. Interestingly, there were no DEPs after HIIT within TI fibres, rather there were 120 up-regulated and 3 down regulated proteins within TII fibres after 12 weeks HIIT. A recent MS-based analysis of TI and TII fibres after 12 weeks of aerobic training showed both up-regulation and down-regulation of proteins in both TI-post vs TI-pre and TII-post and TII-pre comparisons [210]. Deshmukh and others report that a greater number of proteins are differentially regulated in Type I fibres compared with TII fibres [210] which is in stark contrast to the results presented in this chapter. However, in agreement with their conclusion, these findings suggest that TI and TII fibres respond differently to 12 weeks aerobic training. Exercise performed above 80% of maximal effort is likely to activate more Type II fibres than Type I fibres [320]. The HIIT in this study is a combination of maximal efforts closer to V02max than the aerobic exercise performed in the Deshmukh et. al. cohort. This is a possible explanation of the differences observed between Deshmukh et. al. [333] and the findings presented here.

In this chapter I present the first fibre-type-specific DNAm data combined with fibre-type-specific proteome data of TI and TII fibres before and after 12 weeks HIIT. This data provides novel insights into fibre-type response to 12 weeks of HIIT and extends DNAm studies in skeletal muscle to Type I and Type II muscle fibres, overcoming often cited limitations in studying whole skeletal muscle samples. The results suggest that 12 weeks HIIT leads to changes in key metabolic proteins in TII muscle fibres and that these changes are not reflected in changes in DNAm.

8 General Discussion

8.1 Overview

In this thesis, I have developed and implemented a method to study DNAm and protein regulation simultaneously from pooled Type I and Type II human skeletal muscle fibre samples (Chapter 3). This method enabled the genome-wide assessment of DNAm using the widely adopted new version of the Infinium MethylationEPIC Beadchip, which captures up to 950,000 CpGs across the genome, and the subsequent comparison of DNAm values to protein regulation using mass spectrometry proteomics. It was previously shown using one Type I and one Type II pooled fibre sample that there are DNAm differences between Type I and Type II fibres [3]. In this work, I extended the assessment of DNAm to a total of 12 participants (including both males and females). I report extensive baseline methylation differences between Type I and Type II fibres with 11,468 CpGs having greater than 10% beta methylation differences at a stringent adjusted P-value ($<$ 0.001) (Chapter 4) in agreement with the previous work [3].

Recent work has studied protein regulation in pooled Type I and Type II muscle fibres using MS-based proteomics [210, 323]. Using the method developed in this thesis, I found the following relationships in key contractile and metabolic genes: DNA hypermethylation was correlated with down-regulation of proteins; and DNA hypomethylation was correlated with up-regulation of proteins (Chapter 5). The capacity of muscle to respond beneficially to exercise is of interest to many in the field of muscle biology. To the best of my knowledge no study to date has examined fibre-type-specific DNAm changes as a result of an exercise intervention in human skeletal muscle. Therefore using my method, I studied the fibre-type changes in DNAm and protein regulation as a result of 12 weeks HIIT. I report no DNAm differences between Type I and Type II fibres at a stringent or relaxed adjusted P-value, however there were 120 up-regulated and 3 down-regulated proteins solely in Type II muscle fibres after 12 weeks of HIIT at the relaxed adjusted P-value threshold $<$ 0.05 (Chapter 6). These proteins are involved in key metabolic structures and functions and there is a potential remodelling of mitochondrial networks in Type II fibres as a result of our HIIT intervention.

8.2 DNA AND PROTEIN CAN BE SIMULTANEOUSLY PREPARED IN ONE TUBE FROM HUMAN SINGLE MUSCLE FRAGMENTS

DNAm values measured in complex whole tissues are significantly influenced by the cell composition of the sample [334]. In skeletal muscle the variable percentages of muscle fibre types between samples are often cited as a major limitation to interpretation of results and fibre type DNAm signatures are needed for future research [214]. To enable the investigation of fibre-type DNAm differences in skeletal muscle, I developed a one tube digestion protocol for the preparation of both DNA and proteins from the same skeletal muscle fibre fragments that yields enough DNA and protein for use with the Infinium MethylationEPIC Beadchip and MS proteomics.

The protocol was developed by combining and modifying methods developed for studying either DNAm or proteins [3, 210] in pooled skeletal muscle fibres. I chose a buffer with both DNA and protein stabilising properties that could be used for DNA extraction and MS-based proteomics. I then adapted a recent dot-blot protocol [268], for use with a dot-blot-apparatus to standardise the single fibre-digestion and typing protocol to improve typing accuracy on a larger scale. I was able to extract ~ 200 ng of DNA per sample by pooling three fifths of 50 single fibre fragments (Table 4.2) in agreement with previous methods [3]. By pooling one fifth of 15 fibres, I was able to measure ~ 1000 proteins using MS-proteomics and confirm that the Type I and Type II fibre pools were of high purity (Figure 4.9). The method described in this thesis used a PCI phase separation instead of silicon column separation of DNA as previously developed by Begue and others [3]. Using the PCI approach enabled a higher volume of liquid to be processed and achieved a higher yield of DNA compared with silicon columns. This was also achieved using a simple SDS-DTT based buffer that has been successfully used in proteomic studies of pooled fibre-types [210]. Overall this method enables typing of fibres followed by preparation of both DNA and proteins from Type I and Type II muscle fibres from a single tube digestion.

Limitations & Strengths

Initially this thesis aimed to assess DNAm from Type I, Type IIa and Type IIx muscle fibres. A major hurdle in the assessment of Type IIx fibres was the inability to re-

produce accurate and clear typing of MYH1 positive Type IIx fibres using the dot blot method. I adapted a newer dot plot protocol for fibre typing [268], in contrast, Begue and others [3] used protein separation on polyacrylamide gels followed by silver-staining and myosin migration distance for fibre typing [335]. This enabled them to group fibres as Type I and Type IIa samples. In the dot blot method it is difficult to determine mixed fibres from contaminated fibres, and to determine Type IIx from Type IIa fibres. The dot blot method does however enable higher throughput and cheaper typing of many fibres. An additional method to type single muscle fibres has recently been developed using immunofluorescence and microscopy, called Thrifty [336]. This method is reported to be more efficient than dot blotting and enables typing of contaminated and mixed fibres. This method is a useful approach to typing fibres and overcomes limitations with the dot blot method, however it still requires the cutting of fibres. The key advantage in the approach presented in this thesis is that there is no need to cut fibres. Additionally once fibres are digested and a small sample taken for typing the remaining sample is already in a stabilising buffer.

Another drawback of the presented method is the use of PCI during extraction. PCI requires a higher level of expertise in order to remove DNA and uses more hazardous reagents. However, PCI is cheaper and in this protocol enabled the one tube digestion of fibres without the need to cut and move fibres from tube to tube. This resulted in an increased ability to process a large number of fibres for the assessment of DNAm with the EPIC array. It has been reported in a variety of contexts that phenol chloroform extraction yields more DNA than silica based membranes [337, 338]. I observed that DNA yield from PCI extractions were higher than silica membranes for skeletal muscle fibres. Furthermore the volume of liquid sample that can be processed is higher from PCI than for column based separation. This enables the efficient isolation of DNA from hundreds of pooled single muscle fibre fragments. The use of the PCI method in this thesis enabled the extraction of enough high quality DNA to measure DNAm on the epic array and minimised sample preparation costs.

Application and potential

Overall the strengths of this method are the minimisation of sample handling errors and reduction in sample preparation time as there is no need to cut fibres. As a result,

this approach enables the isolation and preparation of a larger number of fibres. The novelty of this method is the simultaneous preparation of both DNA and protein from the same pooled fibre sample without the need to cut fibres. There is potential to adapt the method for RNA extractions and to automate the entire extraction process in one plate. This protocol was efficient at preparing DNA and protein from multiple samples and provided a method to assess DNAm in human skeletal muscle fibre types in a variety of contexts.

8.3 Type I and Type II muscle fibres are differentially methylated in key contractile and metabolic genes

The fibre-type composition of muscle samples is often listed as a limitation or disregarded in DNAm studies. As cell-type is a major factor when studying DNAm [339] and understanding cell-type populations and their proportions in samples is crucial when designing DNAm studies [21]. I investigated DNAm in Type I and Type II skeletal muscle fibres. I report considerable baseline DNAm differences between Type I and Type II fibres and show evidence that DNAm patterns in WM are driven largely by Type I and Type II fibres. These data indicate that fibre-type should be a consideration in DNAm studies using whole muscle samples.

Using the modified DNA and protein extraction protocol, I obtained high quality DNA from Type I and Type II human samples. I measured DNAm using the new Infinium MethylationEPIC Beadchip, which covers close to one million CpGs sites, and assessed the differences between Type I and Type II fibres. The comparison of Type I and Type II pre fibres revealed that the majority of DMPs were in OpenSea regions and more CpGs were hypermethylated in Type I fibres in agreement with previous data [3]. I confirmed that MYH7, MYH2, TPM1, TNNI1, ATP2A2, ATP2A1 and PFKM are differentially methylated between the fibres types as previously reported [3]. A recent study comparing DNAm in mouse slow and fast muscle fibres reports similar findings to those presented here [213]. Commonly one DMP is of less interest to researchers than finding DMRs [281]. I extended on the previous work by generating a list of DMRs with at least 4 CpG sites that were differentially methylated between Type I and Type II muscle fibres (Tables 5.13 & 5.14). DMRs were also found in the key contractile genes

known to be differentially regulated in muscle fibres. ORA of both the list of DMPs and DMRs showed that methylation differences were over represented in key contractile and metabolic genes providing robust evidence in support of previous work. These findings are in agreement with similar experiments assessing protein differences in Type I and Type II skeletal muscle fibres [210, 209]. In the next section of the discussion the relationship between DNAm and protein will be expanded.

Due to well known differences in metabolic properties of muscle fibre types [183], I investigated which Type I vs Type II DMPs were present in the MitoCarta gene list [310]. MitoCarta is a curated list of 1136 human genes known to be localised to or relevant to mitochondria. I identified 133 MitoCarta genes with at least one DMP (Figure 5.14). LYRM4 was annotated to multiple DMPs and was also identified as a DMR with a hypermethylation in Type I compared with Type II fibres. Multiple LYRM4 CpG sites showed a strong cluster that distinguishes Type I and Type II fibres. LYRM4 is a member of the leucine/tyrosin/arginine motif family (LYRM) which are found in the mitochondria [340]. LYRM4 encodes the ISD11 protein and is important in stabilising NFS1 and plays an important role in iron sulfide (Fe-S) cluster biogenesis [341, 342]. LYRMs and Fe-S clusters have been suggested to coordinate cellular energy status [341, 343]. The role of LYRMs have not been extensively studied in skeletal muscle, however LYRM1 in muscle may have a role in glucose homeostasis [344], and mutations in LYRM7 have been reported to cause exercise intolerance [345]. Recently our group reported LYRM7 as a protein with large inter-individual variation as a result of 12 weeks HIIT [263], however it would be interesting to understand the role of fibre-type proportions in such analyses of whole muscle. Future experiments could investigate differences in LYRM proteins between muscle fibre types and assess the potential role of these proteins in the coordination of energy utilisation in human muscle.

Limitations & Strengths

One of the major differences between array platforms and WGBS is the number of CpG sites assessed across the genome. Begue and others [3] used RRBS to measure DNAm and reported a total 5x coverage across ~17,000,000 CpGs. I used the methylation EPICv2 array to asses DNAm across ~900,000 CpGs. The number I assessed is a small fraction of the ~28.3 million CpG sites in the human genome [346] and those measured

in muscle fibre types previously. There are likely interesting overall methylation differences across many more sites than assessed in this work. RRBS and WGBS are routinely measured at less than 100x coverage, and to achieve the same level of precision as an array when conducting RRBS or WGBS a sequencing coverage of 100x is needed [347]. Although this work measured fewer DNAm sites across the genome, the use of the EPIC array increased the precision of measurements on previous work.

An additional strength of this work is the substantial increase in sample size from n=2 to n=24. Both previous RRBS studies of skeletal muscle fibre types utilised a final sample of one Type I and one Type II fibre pooled from either multiple human [3] or mouse donors [213]. There is potential for DNAm analyses using fishers exact tests conducted on n=2 samples to over-inflate DNAm differences as well as not accounting for biological variation [348]. Begue and others reported 60-70% methylation differences in key myosin and tropomyosin genes in Type I verses Type II fibres, in contrast I report differences of only 30-40% in these genes in Type I verses Type II fibres. I used empirical bayes adjusted statistics that have been suggested as strong methods to control for false discoveries [348]. The work in piloting DNAm studies in human skeletal muscle by Begue and others [3] has been instrumental in the generation of this work. Building upon this work, I have been able to generate a more comprehensive DNAm dataset in human skeletal muscle fibre types that overcome earlier limitations in sample sizes, precision and choice of statistics.

Epigenetics and DNAm has been associated with more than just cell type. Ageing is associated with epigenetic changes [349]; epigenetic alterations has been listed as a hallmark of ageing [350, 351]; DNAm robustly predicts age of samples [352]. Furthermore, maintaining strength and muscle health is important for healthy ageing [353], with reduced muscle mass during ageing largely driven by a reduction in size of Type II muscle fibres [354]. Due to the ageing population across much of the world, determining the contribution of methylation to the ageing phenotype is underway [355]. Our laboratory has recently conducted a large meta-analysis of DNAm changes with age in skeletal muscle, and reported extensive changes in genes involved in muscle structure and development [159]. In this study samples were collected from apparently healthy adults between the ages of 18 and 45, and the methods developed provide a foundation for future research in fibre-type specific DNAm in ageing.

Application and potential

Skeletal muscle displays a number of sex differences: differences in mitochondria, variation in responses to anabolic/catabolic stimuli, and variability in fibre type proportions [356]. However, females remain an understudied population in the field of muscle biology. Our laboratory recently investigated sex differences in DNAm in human skeletal muscle [167], and found numerous genes to be differentially methylated between the sexes. Furthermore, they reported a relationship between DNAm differences and fibre-type proportion differences, suggesting that fibre-type proportion differences drive the DNAm sex differences [167]. In this thesis, I studied DNAm differences between Type I and Type II fibres in both males and females. Although a sex comparison between the fibre types is not included in this thesis, we intend to conduct sex difference analyses using this data set. Overall a major potential of the work presented here is the inclusion of both sexes which can be leveraged in future investigations. This work also included both male and female, and did not exclude ethinicities as it was hypothesised that muscle fibre types are more similar at the level of DNAm and we wanted to generate DNAm data more representative of the healthy adult population. In Chapter 4, I reported considerable DNAm differences between Type I and Type II fibres. A comparison with WM samples showed that a major source of signal in WM is attributed to fibre types, providing further evidence that fibre-type should be a consideration in DNAm studies using whole muscle samples.

8.4 There are relationships between methylation direction and protein regulation in Type I and Type II muscle fibres

A major role of DNAm is in the control of gene expression and DNAm in different genomic contexts, can either induce or repress gene expression [357]. There is a growing interest in using a system biology approach to study the link between DNAm and gene/protein regulation [312]. Normally the study of the methylome, transcriptome and proteome are conducted using different datasets generated from different studies or they utilise different samples from the same individual. I applied my method to study protein

differences between Type I and Type II fibres extracted simultaneously from the exact same samples as presented in Chapter 5.

I measured protein abundance using LC-MS/MS and assessed the differences between Type I and Type II fibres. I identified 339 DEPs with a BH-P-Value < 0.001 between Type I and Type II pre fibres. I confirmed that MYH7, MYH2, TNNI1, TNNT1, TNNT3m ATP2A2 and ATP2A1 were all differentially regulated between the fibres types (Table 6.2). These DEPs are the same as those listed in the previous methylation section and the direction of protein abundance was in the opposite direction to methylation. These results link those presented in previous methylome and proteome studies of human skeletal muscle fibre types [3, 210, 323], and provide evidence of the relationship between methylation levels and protein abundance. A comparison of the beta differences of identified DMPs to the protein abundance logFC of identified DEPs showed that there was a moderate correlation (R=-0.4 to -0.57) between hypermethylation and down regulation of key muscle proteins (and vice versa) (Figure 6.11). When I compared the identified DMRs from Chapter 4 with the DEPs, there was an increase in the strength of the relationship between hypermethylation of a DMR and protein down regulation (R=-0.62 to -0.63) (Figure 6.12). Key metabolism reactome pathways such as glycolysis, metabolism of carbohydrates, glycogen metabolism and glycogen breakdown were identified using ORA of the overlapping hypermethylated DMP and down regulated DEP from Type I compared with Type II fibres (Figure 6.10). It is well reported that Type II muscle fibres rely more on glycolytic energy pathways than Type I fibres (reviewed here) [59, 317, 358]. I provide evidence of relationships between DNAm and proteins and suggest there is a distinct role of DNAm in energy metabolism in muscle fibre types. The role being that there is a preparedness of hypomethylation in glycolytic pathway genes in Type II muscle fibres compared to Type I fibres.

Assessing the myosin isoform content (myosin purity) of pooled muscle fibre type samples was an additional goal of measuring protein abundance. Previous pooled fibre and single fibre proteomics reported a purity of MYH7 Type I fibres of ~90% and purity of MYH2 Type II fibres of ~80% [203, 210]. Whereas, a more recent proteomic study on pooled Type I and Type II fibres showed a large variability in myosin isoform content of samples, in particular, within MYH7 Type I fibres there was only up to 70% MYH7 content [323]. I present results showing that Type I fibres had >95% MYH7 content, and

Type II fibres had >80% MYH2/MYH1 content (Figure 6.4A). This is in agreement with the previous work [298, 210]. A recent paper (biorxiv) looking at both the proteome and transcriptome of single muscle fibres confirm the presence of highly pure Type I fibres based on MYH7 content and they showed that the majority of fibres (both TI and TII) express some MYH7. The also show evidence that 100% pure MYH2 fibres were not present [211]. Finally, they also report that at the protein level, a pure Type IIx fibre was not identifiable and that most Type IIx fibres were instead Type IIa/IIx hybrids. This may be why I was unable to successfully type Type IIx fibres. A comparison of the dot blot method and proteomics revealed a high agreement between the two typing strategies for typing single fibres [211]. The results of the myosin proteomic analysis of pooled Type I and Type II fibres in this work are in agreement and show that the modified dot blot method is effective at preparing pure Type I and Type II muscle fibre populations. This provides a level of confidence that the DNAm data generated in this thesis are truly representative of the two major muscle fibre types.

I identified MYH7, MYH2, MYH1 and MYH4 using proteomics in agreement with previous studies assessing human skeletal muscle fibres at both the single fibre and pooled fibre level [203, 210]. MYH4 has been reported to have little to no expression in human skeletal muscle [59], however in Type II fibres I identified between 0-20% MYH4 content. These values are the same as have been reported in single fibre proteomics in human quadriceps muscle [203]. Additionally, these results are also supported by western blot analysis of whole skeletal muscle that shows an increase in MYH4 content with increasing levels of Type II fibre content [359]. The finding of MYH4 in Type II fibres is difficult to interpret due to the high sequence similarity between MYH4 and MYH1 with only one amino acid difference between the two isoforms, additionally, MYH1, MYH4 and MYH2 are part of the fast myosin complex on chromosome 17 in humans [360] and in mouse muscle this locus is under the influence of a super enhancer [361]. Further complicating proteomic studies of human skeletal muscle fibres, a recent protocol has shown that in skeletal muscle fibres protein quantification technique can drastically change the reported values of MYH4 [362]. However further studies are warranted in human skeletal muscle and I suggest the application of methods studying all three biological layers from the same sample may help unravel the biology of fast, slow and mixed fibre populations.

Limitations & Strengths

It is well known that measuring proteins in muscle samples is difficult due to contractile proteins making up the bulk of the sample [333]. Due to this it can be difficult to assess less abundant proteins. In Chapter 4, I identified LYRM4 as a potential gene of interest due to DNAm differences between Type I and Type II fibres. ISD11 the protein encoded by LYRM4 has been identified in human skeletal muscle [342], however was absent in the proteome data. The most obvious deficiency in this data is the lack of transcriptomic data. There is evidence that relationships exist between methylation and protein regulation, however I identified a number of protein abundance differences between Type I and Type II fibres that did not have differential methylation. It would be of interest to study transcriptomic differences across the genes to assess the coordination of DNAm, RNA transcription and protein abundance. Although RNA was not extracted, the modified method described in this thesis could be readily adapted to extract DNA, RNA and protein.

Skeletal muscle is not only made up of muscle fibres but rather it is estimated that only between 50-70% of nuclei come from myonuclei [109, 363]. It can not be excluded that all other cell types are removed from the fibres upon isolation. Single fibre imaging has shown that satellite cells and and endothelial cells have the potential to line single skeletal muscle fibres [364]. As a result, there is potential for DNA extractions from pooled muscle fibre types to include nuclei other than myonuclei. Some genes had methylation values with a different direction to protein levels (Figures 6.11 & 6.12) which could be explained by contamination of muscle fibres with nuclei from other cells types. If the number of nuclei attributed to myonuclei is closer to 50%, there is a significant amount of DNA from other cell types that could influence DNAm values. This combined with the larger proportion of proteins coming from the contractile fibres might influence the reported correlations between DMP/DMR values and protein values. Future work should extend on the results presented here by considering the additional cell types resident in skeletal muscle or investigating the potential of adapting the method here to prepare single myonuclei from pooled muscle fibre types. A major limitation of studying pooled fibres is the inability to assess fibre type heterogeneity within a fibre population. Investigations on single muscle fibres suggest considerable heterogeneity in fibres within a particular fibre type [209, 211]. Recent work assessed the proteome and

transcriptome of single human skeletal muscle fibres and they reported that the major source of variation was not attributed to MYH but rather ribosomal differences [211]. In contrast our pooled fibre samples clearly separated along the first principal component in both the methylome (after removing sex chromosomes) and proteome based upon typing for MYH7 and MYH2. Missing data is one of the major challenges of single cell studies [365]. An advantage of choosing a pooled fibre approach is that it increases biological yield of DNA and protein and has the benefit of generally increasing measurement precision [366].

Application and potential

A strength of the DNAm dataset is that unlike proteins, DNA is extracted from the many myonuclei within the muscle fibre and there is the potential to observe differences in important genes that would be difficult to identify in MS-based proteomics. Further I have generated a proteome data set that was successfully used to assess fibre-type purity and can be used to assess the relationship between DNAm and protein in some of the most abundant skeletal muscle proteins. We hope that these resources are useful to investigate fibre-type specific muscle biology.

8.5 Type II muscle fibres shift towards Type I fibres in key metabolic proteins after 12 weeks HIIT

There has been considerable interest in identifying changes in DNAm due to environmental stimuli [367]. A number of studies have reported changes in DNAm as a result of exercise (reviewed in [368]). Further, it has been speculated that DNAm may encode a memory of previous exercise within skeletal muscle [369]. However DNAm studies investigating changes with exercise have been limited to using whole skeletal muscle samples. Due to the influence of fibre-type on DNAm in WM samples, I assessed within fibre DNAm changes before and after 12 weeks HIIT.

12 weeks of HIIT led to improvements in VO_{2peak}, LT and W_{peak} at the group level as previously reported in 12 weeks HIIT interventions [370, 324, 371, 325]. I conducted within fibre comparisons of either Type I pre to Type I post, Type II pre to Type II post fibres or WM pre to WM post samples. I identified two DMPs in Type II fibres after 12 weeks HIIT, which were annotated to NPDC1 and SENP1 both which had BH-P-vals of 0.009 and 0.02 respectively. Both the Type II pre to post and WM pre to post comparisons revealed no DMPs with a BH-P-val < 0.05. These results are in agreement with those previously reported from our laboratory, where both 4 weeks and 12 weeks HIIT resulted in minimal DNAm changes [263, 181]. In contrast, 12 weeks of resistance training interventions have been reported to induce DNAm changes in whole skeletal muscle [177, 178, 175] and, interestingly, after a period of detraining, participants had greater changes in DNAm after a retraining period [177, 178]. The discrepancies in HIIT verses resistance training could be that resistance training promotes greater DNAm muscle changes than aerobic exercise, however to date there has been no direct comparison of DNAm induced changes as a result of resistance or aerobic interventions. The length of intervention is a possible explanation of reporting no DNAm changes after 12 weeks of HIIT. Previous studies have reported DNAm changes as a result of longer aerobic interventions: 6 months aerobic training altered DNAm in a diabetes context [170] and 5 months of aerobic training altered DNAm in the context of cancer [322]. Our laboratory provide evidence that DNAm changes in skeletal muscle are associated with VO2max [181, 372] and there is growing evidence that long-term physical activity is associated

with altered DNAm [373, 374, 372, 156] indicating that potential DNAm changes with exercise manifest after years of training.

Interestingly, I did identify 120 proteins up-regulated after 12 weeks HIIT only within Type II fibres (Figure 7.6). The proteins were predominantly mitochondrial related and ORA revealed that they were involved in the citric acid cycle, metabolism and complex one biogenesis. Heatmap hierarchical clustering of the 120 up-regulated proteins suggest that within these proteins Type II post fibres appear more similar to both Type I pre and post fibres (Figure 7.8). These data suggest a potential mitochondrial remodelling of Type II fibres in response to 12 weeks HIIT, and follow up studies are planned to assess differences in mitochondrial networks in Type I and Type II single muscle fibres before and after 12 weeks HIIT.

A recent proteomics assessment of Type I and Type II fibres after moderate endurance exercise also revealed fibre-type proteomic changes [210]. The moderate exercise intervention resulted in changes in both Type I and Type II fibres in contrast to my results which showed a preferential adaptation only in Type II fibres with HIIT. The intervention performed in their study was reported as 12 weeks aerobic training consisting of four one hour sessions per week [375, 210], and the present work assessed 12 weeks of HIIT. A recent preprint paper also assessed protein differences before and after exercise training [323]. They assessed both moderate training and sprint interval training and reported no fibre-type specific changes in the proteome before and after either intervention [323]. Reisman and others [323] suggest that the discrepancy could be due to the statistical method used, as they reported using empirical Bayes. Interestingly, I also utilised empirical Bayes statistics and provide evidence of up-regulation of proteins after HIIT. A comparison of the pooled fibre samples from each of the studies and those presented in the thesis indicate that the discrepancies could be due to pooled fibre sample purity. Pooled fibre samples in the pre-print paper show highly variable myosin isoform content of Type I and Type II fibres [323], in contrast both my results and those of Deshmukh and others [210] show high purity of pooled fibre samples. This highlights the importance in fibre-typing accuracy and shows that a method that reduces user error such as the one presented in this thesis is highly valuable in the study of fibre type changes with exercise. The results here provide evidence that our 12 week intervention has potential to

shift Type II fibre metabolism towards Type I fibres in the proteome, however it does not appear that this is driven by changes in the methylome.

Limitations & Strengths

The DNAm results described here for the WM pre and post comparison are unsurprising as previous samples sizes were larger in those studies. Prior to this project, power calculations indicated a total of 20 participants were required to detect DNAm differences of 10% in both the fibre-type specific and pre and post analyses. I observed no DNAm differences >10% in any group observation after 12 weeks HIIT. There is potential for DNAm changes between 0-5% after exercise to be large enough to effect protein regulation. With a larger sample size or a sample size approaching 20 participants we might achieve enough power to identify such DNAm changes with exercise.

Another limitation is that proteomic studies of muscle are dominated by a small number of highly abundant proteins. The identification of lowly abundant proteins is more difficult and I showed that Type II pre fibres had a number of up-regulated proteins that had a number of missing values. The assumption is that these values are below the detection threshold but still present in the sample. However, I will conduct validation experiments on our key up-regulated proteins to confirm if the claim that exercise increases mitochondrial proteins in muscle fibre types holds true or if it is a potential consequence of missing data.

This study was significantly impacted by COVID19. A number of participants were infected with COVID19 shortly before, during or shortly after the training intervention. In consultation with our qualified physician, we decided to continue training participants if they were able to return to training after a mandatory seven day home isolation. Of the fifteen recruited participants, only one participant was unable to continue training due an interruption of > 10 days, and two participants developed COVID19 during the intervention phase and successfully returned to training within 10 days. Although a limitation, these disruptions also provide strength to this study, as they represent a more realistic situation where people's training can be interrupted for various reasons. Despite these disruptions, I report changes in Type II muscle fibres and report that the 12 week intervention elicited changes to physical performance.

The exercise study protocol implemented in this thesis was part of the larger Gene SMART study [260]. Baseline testing and familiarisation sessions were conducted prior to the baseline biopsies and a 48 hour washout period was implemented between the final GXT and muscle sampling. There is evidence that gene transcription is altered 48-96 hours after an acute aerobic exercise session [376, 377] and although there is a paucity of research into the effect of acute exercise on the proteome and methylome 48 hours or longer after exercise it could be expected that the proteome is perturbed even 48 hours after an acute exercise session. This may result in difficulties in identifying DEPs in this study as baseline testing was performed prior to biopsies. Despite this I report changes in protein levels in Type II fibres after HIIT training and the changes could be considered as a result of the HIIT program and not purely as a result of an acute exercise session/GXT considering both pre and post biopsies were performed at least 48 hours after the last exercise. Nonetheless the timing of biopsies after GXT sessions should be carefully considered when interpreting the findings in this chapter as they may counfound the results.

Application and potential

I hypothesised that 12 weeks HIIT training would shift Type II fibres due to the preferential recruitment of the Type II fibre population with the increased load. The intervention employed in this study was 12 weeks HIIT above LT with sessions lasting 20-45 minutes, this fits the results presented as it could be assumed that the higher load recruits Type II fibres and the sessions lasting > 30 minutes would induce a metabolic shift towards a Type I fibre. These results support the hypothesis that investigations in whole skeletal muscle may mask interesting underlying fibre-type specific biological changes in muscle. Further, chapter 7 shows the importance of developing accurate methods for preparing skeletal muscle fibres when studying fibre-type changes using pooled muscle fibre types. Although I present here a brief investigation of fibre-type specific changes in skeletal muscle after exercise, the study of fibre-type specific changes should be considered when studying other factors, such as ageing, sex or disease.

8.6 Future work

Future work is planned for a more comprehensive investigation of sex differences in Type I and Type II skeletal muscle fibres using this data, as well as further investigations in fibre type methylation and protein responses to exercise. A number of validation experiments are required to assess the presented findings, plans are made to confirm reported differences between Type I and Type II fibres using single fibre imaging and single fibre western blot.

There are a number of potential adaptations of the method presented: it could be used for preparation of RNA; once fibres are isolated the workflow could be automated; the digested fibres could easily be used for single-fibre DNAm analysis; and there is potential for the one tube digestion approach to be used for single-nuclei isolations from pooled muscle fibre types. Overall fibre-type has a large effect on DNAm in human skeletal muscle and should not be disregarded. Therefore, this work provides the foundation for the following. This study was limited to healthy young adults and future work should study fibre-type DNAm differences in the context of ageing. There is potential to use this work to generate DNAm profiles of Type I and Type II skeletal muscle fibres which can be used to develop fibre-type correction algorithms for whole muscle DNAm studies. Finally, due to the up-scaling potential of the method developed in this thesis, this approach could be applied to studies with larger samples sizes or longer interventions to study DNAm adaptations of skeletal muscle in a variety of contexts.

8.7 Conclusion

At the outset of this thesis the overarching aim was to to establish a method in which we can extract DNA and proteins from the same pooled fibre samples with high purity and yield. I then aimed to extend the assessment of fibre-type-specific DNAm to: include multiple samples; include both males and females; and to use the EPIC array. To achieve this I aimed to adapt methods to increase the yield of DNA and increase the throughput of fibres. Once equipped with an appropriate method I also aimed to study fibre-type-specific DNAm after 12 weeks of HIIT. During the participant recruitment and training phase of this thesis conflicting evidence emerged of the purity of pooled Type I and Type II muscle fibres, as a result I aimed to simultaneously isolate enough protein to check sample purity to ensure robustness of fibre-type specific DNAm results. This also had the additional benefit of providing a link between DNAm and fibre phenotype.

To this end, here I present a modified method for the isolation of DNA and protein from pooled Type I and Type II muscle fibres. A major outcome of this work is the comprehensive assessment of DNAm differences between Type I and Type II muscle fibres and the simultaneous study of proteins. This approach has enabled for the first time an in-depth study of the methylome of Type I and Type II fibres whilst providing confirmation of the purity of pooled samples. I intend to make both the methylation values and protein values available as a resource for others to use freely. I have taken care in this thesis to present results without referring to statistical significance and to outline the number of DMPs and DEPs identified at varying levels of adjusted P value cutoffs. All test statistics can be provided on request and depending on the specific research question an assessment of appropriate cutoffs can be made. I decided to focus on the most stringent adjusted P-value cutoff and those values with greater than 10% beta differences for the methylation assessment between Type I and Type II fibres in order to reduce the size of the results presented. This has resulted in a high level of confidence in the results presented. Overall the datasets generated here provide robust DNAm profiles of Type I and Type II muscle fibres. I expect the data provided in this work to be useful to the muscle biology field.

Bibliography

[1] Haeckel EHPA, Gifford B. The story of the development of a youth : letters to his parents 1852-1856. Harper; 1923.

[2] Coleman W. Cell, Nucleus, and Inheritance: An Historical Study. Proceedings of the American Philosophical Society. 1965;109(3):124-58. Available from: http://www.jstor.org/stable/986128.

[3] Begue G, Raue U, Jemiolo B, Trappe S. Dna Methylation Assessment From Human Slow- and Fast-Twitch Skeletal Muscle Fibers. Journal of Applied Physiology. 2017;122(4):952-67. Available from: http://dx.doi.org/10.1152/japplphysiol.00867.2016.

[4] Haeckel E. Generelle Morphologie der Organismen. 1866:287-8.

[5] Flemming W. Zellsubstanz, Kern und Zelltheilung. 1882.

[6] Miescher F. Ueber die chemische Zusammensetzung der Eiterzellen. Med -Chem Unters. 1871;4:441-60.

[7] Kossel A. Ueber Einen Peptonartigen Bestandtheil Des Zellkerns. bchm. 1884;8(6):511-5. Available from: http://dx.doi.org/10.1515/bchm1.1884.8.6.511.

[8] Altmann R. Über nucleinsäuren. Anat Physiol. 1889:524-36.

[9] Kossel A. Untersuchungen über die Nucleine und ihre Spaltungsproduckte. 1881.

[10] Johannsen W. Elemente der exakten erblichkeitslehre. Deutsche wesentlich erweiterte ausgabe in fünfundzwanzig vorlesungen. Jena, G Fischer.; 1909.

[11] Morgan TH, Sturtevant AH, Muller HJ, Bridges CB. The mechanism of Mendelian heredity. H. Holt and Company; 1923.

[12] Avery OT, MacLeod CM, McCarty M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types. Journal of Experimental Medicine. 1944;79(2):137-58. Available from: http://dx.doi.org/10.1084/jem.79.2.137.

[13] Watson JD, Crick FHC. Molecular Structure of Nucleic Acids: a Structure for Deoxyribose Nucleic Acid. Nature. 1953;171(4356):737-8. Available from: http://dx.doi.org/10.1038/171737a0.

[14] Franklin RE, Gosling RG. Molecular Configuration in Sodium Thymonucleate. Nature. 1953;171(4356):740-1. Available from: http://dx.doi.org/10.1038/171740a0.

[15] Wilkins MHF, Stokes AR, Wilson HR. Molecular Structure of Nucleic Acids: Molecular Structure of Deoxypentose Nucleic Acids. Nature. 1953;171(4356):738-40. Available from: http://dx.doi.org/10.1038/171738a0.

[16] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The Sequence of the Human Genome. Science. 2001;291(5507):1304-51. Available from: http://dx.doi.org/10.1126/science.1058040.

[17] Consortium IHGS, Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. Initial Sequencing and Analysis of the Human Genome. Nature. 2001;409(6822):860-921. Available from: http://dx.doi.org/10.1038/35057062.

[18] Haig D. Commentary: the Epidemiology of Epigenetics: Figure 1. International Journal of Epidemiology. 2011;41(1):13-6. Available from: http://dx.doi.org/10.1093/ije/dyr183.

[19] Waddington CH. The Epigenotype. International Journal of Epidemiology. 1942;41(1):10-3. Available from: http://dx.doi.org/10.1093/ije/dyr184.

[20] Nanney DL. Epigenetic Control Systems. Proceedings of the National Academy of Sciences. 1958;44(7):712-7. Available from: http://dx.doi.org/10.1073/pnas.44.7.712.

[21] Lappalainen T, Greally JM. Associating Cellular Epigenetic Models With Human Phenotypes. Nature Reviews Genetics. 2017;18(7):441-51. Available from: http://dx.doi.org/10.1038/nrg.2017.32.

[22] Bird A. Perceptions of Epigenetics. Nature. 2007;447(7143):396-8. Available from: http://dx.doi.org/10.1038/nature05913.

[23] Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-Scale Dna Methylation Maps of Pluripotent and Differentiated Cells. Nature. 2008;454(7205):766-70. Available from: http://dx.doi.org/10.1038/nature07107.

[24] Felsenfeld G. A Brief History of Epigenetics. Cold Spring Harbor Perspectives in Biology. 2014;6(1):a018200-0. Available from: http://dx.doi.org/10.1101/cshperspect.a018200.

[25] Rivera CM, Ren B. Mapping Human Epigenomes. Cell. 2013;155(1):39-55. Available from: http://dx.doi.org/10.1016/j.cell.2013.09.011.

[26] Kornberg RD. Chromatin Structure: a Repeating Unit of Histones and Dna. Science. 1974;184(4139):868-71. Available from: http://dx.doi.org/10.1126/science.184.4139.868.

[27] Noll M. Internal Structure of the Chromatin Subunit. Nucleic Acids Research. 1974;1(11):1573-8. Available from: http://dx.doi.org/10.1093/nar/1.11.1573.

[28] Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal Structure of the Nucleosome Core Particle At 2.8 Å Resolution. Nature. 1997;389(6648):251-60. Available from: http://dx.doi.org/10.1038/38444.

[29] Peterson CL, Laniel MA. Histones and Histone Modifications. Current Biology. 2004;14(14):R546-51. Available from: http://dx.doi.org/10.1016/j.cub.2004.07.007.

[30] Allfrey VG, Faulkner R, Mirsky AE. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proceedings of the National Academy of Sciences. 1964;51(5):786-94. Available from: http://dx.doi.org/10.1073/pnas.51.5.786.

[31] Tamaru H. Confining Euchromatin/heterochromatin Territory: jumonji Crosses the Line. Genes amp; Development. 2010;24(14):1465-78. Available from: http://dx.doi.org/10.1101/gad.1941010.

[32] Eddy SR. Non-Coding Rna Genes and the Modern Rna World. Nature Reviews Genetics. 2001;2(12):919-29. Available from: http://dx.doi.org/10.1038/35103511.

[33] Kaikkonen MU, Lam MTY, Glass CK. Non-Coding Rnas As Regulators of Gene Expression and Epigenetics. Cardiovascular Research. 2011;90(3):430-40. Available from: http://dx.doi.org/10.1093/cvr/cvr097.

[34] Gibney ER, Nolan CM. Epigenetics and Gene Expression. Heredity. 2010;105(1):4-13. Available from: http://dx.doi.org/10.1038/hdy.2010.54.

[35] Hotchkiss RD. The Quantitative Separation of Purines, Pyrimidines, and Nucleosides By Paper Chromatography. Journal of Biological Chemistry. 1948;175(1):315-32. Available from: http://dx.doi.org/10.1016/S0021-9258(18)57261-6.

[36] Doskočil J, Sorm F. Distribution of 5-methylcytosine in Pyrimidine Sequences of Deoxyribonucleic Acids. Biochimica et Biophysica Acta. 1962;55(6):953-9. Available from: http://dx.doi.org/10.1016/0006-3002(62)90909-5.

[37] Gruenbaum Y, Stein R, Cedar H, Razin A. Methylation of Cpg Sequences in Eukaryotic Dna. FEBS Letters. 1981;124(1):67-71. Available from: http://dx.doi.org/10.1016/0014-5793(81)80055-5.

[38] Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, et al. Amount and Distribution of 5-methylcytosine in Human Dna From Different Types of Tissues Or Cells. Nucleic Acids Research. 1982;10(8):2709-21. Available from: http://dx.doi.org/10.1093/nar/10.8.2709.

[39] Moore LD, Le T, Fan G. Dna Methylation and Its Basic Function. Neuropsychopharmacology. 2012;38(1):23-38. Available from: http://dx.doi.org/10.1038/npp.2012.112.

[40] Bestor TH. The Dna Methyltransferases of Mammals. Human Molecular Genetics. 2000;9(16):2395-402. Available from: http://dx.doi.org/10.1093/hmg/9.16.2395.

[41] Smith ZD, Meissner A. Dna Methylation: Roles in Mammalian Development. Nature Reviews Genetics. 2013;14(3):204-20. Available from: http://dx.doi.org/10.1038/nrg3354.

[42] Bird AP. Cpg-Rich Islands and the Function of Dna Methylation. Nature. 1986;321(6067):209-13. Available from: http://dx.doi.org/10.1038/321209a0.

[43] Bird AP. Cpg Islands As Gene Markers in the Vertebrate Nucleus. Trends in Genetics. 1987;3(nil):342-7. Available from: http://dx.doi.org/10.1016/0168-9525(87)90294-0.

[44] Robertson KD. Dna Methylation and Human Disease. Nature Reviews Genetics. 2005;6(8):597-610. Available from: http://dx.doi.org/10.1038/nrg1655.

[45] Jones PA. Functions of Dna Methylation: Islands, Start Sites, Gene Bodies and Beyond. Nature Reviews Genetics. 2012;13(7):484-92. Available from: http://dx.doi.org/10.1038/nrg3230.

[46] Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, et al. Dna Methylation Profiling of Human Chromosomes 6, 20 and 22. Nature Genetics. 2006;38(12):1378-85. Available from: http://dx.doi.org/10.1038/ng1909.

[47] Reik W. Stability and Flexibility of Epigenetic Gene Regulation in Mammalian Development. Nature. 2007;447(7143):425-32. Available from: http://dx.doi.org/10.1038/nature05918.

[48] Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, et al. A Bayesian Deconvolution Strategy for Immunoprecipitation-Based Dna Methylome Analysis. Nature Biotechnology. 2008;26(7):779-85. Available from: http://dx.doi.org/10.1038/nbt1414.

[49] Loyfer N, Magenheim J, Peretz A, Cann G, Bredno J, Klochendler A, et al. A Dna Methylation Atlas of Normal Human Cell Types. Nature. 2023;nil(nil):nil. Available from: http://dx.doi.org/10.1038/s41586-022-05580-6.

[50] Lewis J, Bird A. Dna Methylation and Chromatin Structure. FEBS Letters. 1991;285(2):155-9. Available from: http://dx.doi.org/10.1016/0014-5793(91)80795-5.

[51] Cedar H, Bergman Y. Linking Dna Methylation and Histone Modification: Patterns and Paradigms. Nature Reviews Genetics. 2009;10(5):295-304. Available from: http://dx.doi.org/10.1038/nrg2540.

[52] Hooke R. Extracts from Micrographia: or, some physiological descriptions of minute bodies made by magnifying glasses with observations and inquiries there-

upon. 5. Clay; 1894.

[53] Ribatti D. An Historical Note on the Cell Theory. Experimental Cell Research. 2018;364(1):1-4. Available from: http://dx.doi.org/10.1016/j.yexcr.2018.01.038.

[54] Breschi A, Muñoz-Aguirre M, Wucher V, Davis CA, Garrido-Martín D, Djebali S, et al. A Limited Set of Transcriptional Programs Define Major Cell Types. Genome Research. 2020;30(7):1047-59. Available from: http://dx.doi.org/10.1101/gr.263186.120.

[55] Graf T, Enver T. Forcing Cells To Change Lineages. Nature. 2009;462(7273):587-94. Available from: http://dx.doi.org/10.1038/nature08533.

[56] Morris SA. The Evolving Concept of Cell Identity in the Single Cell Era. Development. 2019;146(12):nil. Available from: http://dx.doi.org/10.1242/dev.169748.

[57] Trapnell C. Defining Cell Types and States With Single-Cell Genomics. Genome Research. 2015;25(10):1491-8. Available from: http://dx.doi.org/10.1101/gr.190595.115.

[58] Buettner F, Natarajan KN, Casale FP, Proserpio V, Scialdone A, Theis FJ, et al. Computational Analysis of Cell-To-Cell Heterogeneity in Single-Cell Rna-Sequencing Data Reveals Hidden Subpopulations of Cells. Nature Biotechnology. 2015;33(2):155-60. Available from: http://dx.doi.org/10.1038/nbt.3102.

[59] Schiaffino S, Reggiani C. Fiber Types in Mammalian Skeletal Muscles. Physiological Reviews. 2011;91(4):1447-531. Available from: http://dx.doi.org/10.1152/physrev.00031.2010.

[60] Bernstein BE, Meissner A, Lander ES. The Mammalian Epigenome. Cell. 2007;128(4):669-81. Available from: http://dx.doi.org/10.1016/j.cell.2007.01.033.

[61] Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, et al. The Nih Roadmap Epigenomics Mapping Consortium. Nature Biotechnology. 2010;28(10):1045-8. Available from: http://dx.doi.org/10.1038/nbt1010-1045.

[62] Smallwood SA, Kelsey G. De Novo Dna Methylation: a Germ Cell Perspective. Trends in Genetics. 2012;28(1):33-42. Available from: http://dx.doi.org/10.1016/j.tig.2011.09.004.

[63] Ji H, Ehrlich LIR, Seita J, Murakami P, Doi A, Lindau P, et al. Comprehensive Methylome Map of Lineage Commitment From Haematopoietic Progenitors. Nature. 2010;467(7313):338-42. Available from: http://dx.doi.org/10.1038/nature09367.

[64] Shipony Z, Mukamel Z, Cohen NM, Landan G, Chomsky E, Zeliger SR, et al. Dynamic and Static Maintenance of Epigenetic Memory in Pluripotent and Somatic Cells. Nature. 2014;513(7516):115-9. Available from: http://dx.doi.org/10.1038/nature13458.

[65] Kim M, Costello J. Dna Methylation: an Epigenetic Mark of Cellular Memory. Experimental amp; Molecular Medicine. 2017;49(4):e322-2. Available from: http://dx.doi.org/10.1038/emm.2017.10.

[66] Bird A. Dna Methylation Patterns and Epigenetic Memory. Genes & Development. 2002;16(1):6-21. Available from: http://dx.doi.org/10.1101/gad.947102.

[67] Suzuki MM, Bird A. Dna Methylation Landscapes: Provocative Insights From Epigenomics. Nature Reviews Genetics. 2008;9(6):465-76. Available from: http://dx.doi.org/10.1038/nrg2341.

[68] Zhu T, Liu J, Beck S, Pan S, Capper D, Lechner M, et al. A Pan-Tissue Dna Methylation Atlas Enables in Silico Decomposition of Human Tissue Methylomes At Cell-Type Resolution. Nature Methods. 2022;19(3):296-306. Available from: http://dx.doi.org/10.1038/s41592-022-01412-7.

[69] Cruz-Jentoft AJ, Baeyens JP, Bauer JM, Boirie Y, Cederholm T, Landi F, et al. Sarcopenia: European Consensus on Definition and Diagnosis: Report of the European Working Group on Sarcopenia in Older People. Age and Ageing. 2010;39(4):412-23. Available from: http://dx.doi.org/10.1093/ageing/afq034.

[70] Faulkner JA, Larkin LM, Claflin DR, Brooks SV. Age-Related Changes in the Structure and Function of Skeletal Muscles. Clinical and Experimental Pharmacology

and Physiology. 2007;34(11):1091-6. Available from: http://dx.doi.org/10.1111/j.1440-1681.2007.04752.x.

[71] Briggs AM, Woolf AD, Dreinhöfer K, Homb N, Hoy DG, Kopansky-Giles D, et al. Reducing the Global Burden of Musculoskeletal Conditions. Bulletin of the World Health Organization. 2018;96(5):366-8. Available from: http://dx.doi.org/10.2471/blt.17.204891.

[72] Vos T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. 2017;390(10100):1211-59.

[73] Wosczyna MN, Rando TA. A Muscle Stem Cell Support Group: Coordinated Cellular Responses in Muscle Regeneration. Developmental Cell. 2018;46(2):135-43. Available from: http://dx.doi.org/10.1016/j.devcel.2018.06.018.

[74] Mauro A. Satellite Cell of Skeletal Muscle Fibers. The Journal of Biophysical and Biochemical Cytology. 1961;9(2):493-5. Available from: http://dx.doi.org/10.1083/jcb.9.2.493.

[75] Moss FP, Leblond CP. Satellite Cells As the Source of Nuclei in Muscles of Growing Rats. The Anatomical Record. 1971;170(4):421-35. Available from: http://dx.doi.org/10.1002/ar.1091700405.

[76] Collins CA, Olsen I, Zammit PS, Heslop L, Petrie A, Partridge TA, et al. Stem Cell Function, Self-Renewal, and Behavioral Heterogeneity of Cells From the Adult Muscle Satellite Cell Niche. Cell. 2005;122(2):289-301. Available from: http://dx.doi.org/10.1016/j.cell.2005.05.010.

[77] Kuang S, Kuroda K, Grand FL, Rudnicki MA. Asymmetric Self-Renewal and Commitment of Satellite Stem Cells in Muscle. Cell. 2007;129(5):999-1010. Available from: http://dx.doi.org/10.1016/j.cell.2007.03.044.

[78] Pallafacchina G, Blaauw B, Schiaffino S. Role of Satellite Cells in Muscle Growth and Maintenance of Muscle Mass. Nutrition, Metabolism and Cardiovascular Diseases. 2013;23(nil):S12-8. Available from: http://dx.doi.org/10.1016/j.numecd.2012.02.002.

[79] Chargé SBP, Rudnicki MA. Cellular and Molecular Regulation of Muscle Regeneration. Physiological Reviews. 2004;84(1):209-38. Available from: http://dx.doi.org/10.1152/physrev.00019.2003.

[80] Harms CA. Effect of Skeletal Muscle Demand on Cardiovascular Function. Medicine & Science in Sports & Exercise. 2000;nil(nil):94. Available from: http://dx.doi.org/10.1097/00005768-200001000-00015.

[81] Saltin B, Rådegran G, Koskolou MD, Roach RC. Skeletal Muscle Blood Flow in Humans and Its Regulation During Exercise. Acta Physiologica Scandinavica. 1998;162(3):421-36. Available from: http://dx.doi.org/10.1046/j.1365-201x.1998.0293e.x.

[82] Andersen P. Capillary Density in Skeletal Muscle of Man. Acta Physiologica Scandinavica. 1975;95(2):203-5. Available from: http://dx.doi.org/10.1111/j.1748-1716.1975.tb10043.x.

[83] Andersen P, Henriksson J. Capillary Supply of the Quadriceps Femoris Muscle of Man: Adaptive Response To Exercise. The Journal of Physiology. 1977;270(3):677-90. Available from: http://dx.doi.org/10.1113/jphysiol.1977.sp011975.

[84] Verma M, Asakura Y, Murakonda BSR, Pengo T, Latroche C, Chazaud B, et al. Muscle Satellite Cell Cross-Talk With a Vascular Niche Maintains Quiescence Via Vegf and Notch Signaling. Cell Stem Cell. 2018;23(4):530-43.e9. Available from: http://dx.doi.org/10.1016/j.stem.2018.09.007.

[85] Zurlo F, Larson K, Bogardus C, Ravussin E. Skeletal Muscle Metabolism Is a Major Determinant of Resting Energy Expenditure. Journal of Clinical Investigation. 1990;86(5):1423-7. Available from: http://dx.doi.org/10.1172/JCI114857.

[86] Wolfe RR. The Underappreciated Role of Muscle in Health and Disease. The American Journal of Clinical Nutrition. 2006;84(3):475-82. Available from: http://dx.doi.org/10.1093/ajcn/84.3.475.

[87] Argilés JM, Campos N, Lopez-Pedrosa JM, Rueda R, Rodriguez-Mañas L. Skeletal Muscle Regulates Metabolism Via Interorgan Crosstalk: Roles in Health and Disease. Journal of the American Medical Directors Association. 2016;17(9):789-96. Available from: http://dx.doi.org/10.1016/j.jamda.2016.04.019.

[88] Kim G, Kim JH. Impact of Skeletal Muscle Mass on Metabolic Health. Endocrinology and Metabolism. 2020;35(1):1. Available from: http://dx.doi.org/10.3803/EnM.2020.35.1.1.

[89] Chew NWS, Ng CH, Tan DJH, Kong G, Lin C, Chin YH, et al. The Global Burden of Metabolic Disease: Data From 2000 To 2019. Cell Metabolism. 2023;35(3):414-28.e3. Available from: http://dx.doi.org/10.1016/j.cmet.2023.02.003.

[90] Cornier MA, Dabelea D, Hernandez TL, Lindstrom RC, Steig AJ, Stob NR, et al. The Metabolic Syndrome. Endocrine Reviews. 2008;29(7):777-822. Available from: http://dx.doi.org/10.1210/er.2008-0024.

[91] Stump CS, Henriksen EJ, Wei Y, Sowers JR. The Metabolic Syndrome: Role of Skeletal Muscle Metabolism. Annals of Medicine. 2006;38(6):389-402. Available from: http://dx.doi.org/10.1080/07853890600888413.

[92] Leduc-Gaudet JP, Hussain SNA, Barreiro E, Gouspillou G. Mitochondrial Dynamics and Mitophagy in Skeletal Muscle Health and Aging. International Journal of Molecular Sciences. 2021;22(15):8179. Available from: http://dx.doi.org/10.3390/ijms22158179.

[93] Hamilton MT, Booth FW. Skeletal Muscle Adaptation To Exercise: a Century of Progress. Journal of Applied Physiology. 2000;88(1):327-31. Available from: http://dx.doi.org/10.1152/jappl.2000.88.1.327.

[94] DeFreitas JM, Beck TW, Stock MS, Dillon MA, Kasishke PR. An Examination of the Time Course of Training-Induced Skeletal Muscle Hypertrophy. European Journal of Applied Physiology. 2011;111(11):2785-90. Available from: http://dx.doi.org/10.1007/s00421-011-1905-4.

[95] Konopka AR, Harber MP. Skeletal Muscle Hypertrophy After Aerobic Exercise Training. Exercise and Sport Sciences Reviews. 2014;42(2):53-61. Available from: http://dx.doi.org/10.1249/JES.0000000000000007.

[96] Granata C, Jamnick NA, Bishop DJ. Training-Induced Changes in Mitochondrial Content and Respiratory Function in Human Skeletal Muscle. Sports Medicine. 2018;48(8):1809-28. Available from: http://dx.doi.org/10.1007/s40279-018-0936-y.

[97] Abe T, Kojima K, Kearns CF, Yohena H, Fukuda J. Whole Body Muscle Hypertrophy From Resistance Training: Distribution and Total Mass. British Journal of Sports Medicine. 2003;37(6):543-5. Available from: http://dx.doi.org/10.1136/bjsm.37.6.543.

[98] Burgomaster KA, Hughes SC, Heigenhauser GJF, Bradwell SN, Gibala MJ. Six Sessions of Sprint Interval Training Increases Muscle Oxidative Potential and Cycle Endurance Capacity in Humans. Journal of Applied Physiology. 2005;98(6):1985-90. Available from: http://dx.doi.org/10.1152/japplphysiol.01095.2004.

[99] Pedersen BK, Saltin B. Exercise As Medicine - Evidence for Prescribing Exercise As Therapy in 26 Different Chronic Diseases. Scandinavian Journal of Medicine & Science in Sports. 2015;25(nil):1-72. Available from: http://dx.doi.org/10.1111/sms.12581.

[100] Ubaida-Mohien C, Gonzalez-Freire M, Lyashkov A, Moaddel R, Chia CW, Simonsick EM, et al. Physical Activity Associated Proteomics of Skeletal Muscle: Being Physically Active in Daily Life May Protect Skeletal Muscle From Aging. Frontiers in Physiology. 2019;10(nil):nil. Available from: http://dx.doi.org/10.3389/fphys.2019.00312.

[101] Pillon NJ, Gabriel BM, Dollet L, Smith JAB, Puig LS, Botella J, et al. Transcriptomic Profiling of Skeletal Muscle Adaptations To Exercise and Inactivity. Nature Communications. 2020;11(1):470. Available from: http://dx.doi.org/10.1038/s41467-019-13869-w.

[102] Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of Physical Inactivity on Major Non-Communicable Diseases Worldwide: an Analysis of Burden of Disease and Life Expectancy. The Lancet. 2012;380(9838):219-29. Available from: http://dx.doi.org/10.1016/s0140-6736(12)61031-9.

[103] Giordani L, He GJ, Negroni E, Sakai H, Law JYC, Siu MM, et al. High-Dimensional Single-Cell Cartography Reveals Novel Skeletal Muscle-Resident Cell Populations. Molecular Cell. 2019;74(3):609-21.e6. Available from: http://dx.doi.org/10.1016/j.molcel.2019.02.026.

[104] Dell'Orso S, Juan AH, Ko KD, Naz F, Gutierrez-Cruz G, Feng X, et al. Single-Cell Analysis of Adult Skeletal Muscle Stem Cells in Homeostatic and Regenerative Conditions. Development. 2019;nil(nil):nil. Available from: http://dx.doi.org/10.1242/dev.174177.

[105] Micheli AJD, Laurilliard EJ, Heinke CL, Ravichandran H, Fraczek P, Soueid-Baumgarten S, et al. Single-Cell Analysis of the Muscle Stem Cell Hierarchy Identifies Heterotypic Communication Signals Involved in Skeletal Muscle Regeneration. Cell Reports. 2020;30(10):3583-95.e5. Available from: http://dx.doi.org/10.1016/j.celrep.2020.02.067.

[106] Rubenstein AB, Smith GR, Raue U, Begue G, Minchev K, Ruf-Zamojski F, et al. Single-Cell Transcriptional Profiles in Human Skeletal Muscle. Scientific Reports. 2020;10(1):229. Available from: http://dx.doi.org/10.1038/s41598-019-57110-6.

[107] Micheli AJD, Spector JA, Elemento O, Cosgrove BD. A Reference Single-Cell Transcriptomic Atlas of Human Skeletal Muscle Tissue Reveals Bifurcated Muscle Stem Cell Populations. Skeletal Muscle. 2020;10(1):19. Available from: http://dx.doi.org/10.1186/s13395-020-00236-3.

[108] Buckingham M, Bajard L, Chang T, Daubas P, Hadchouel J, Meilhac S, et al. The Formation of Skeletal Muscle: From Somite To Limb. Journal of Anatomy. 2003;202(1):59-68. Available from: http://dx.doi.org/10.1046/j.1469-7580.2003.00139.x.

[109] Schmalbruch H, Hellhammer U. The Number of Nuclei in Adult Rat Muscles With Special Reference To Satellite Cells. The Anatomical Record. 1977;189(2):169-75. Available from: http://dx.doi.org/10.1002/ar.1091890204.

[110] Huxley HE. The Mechanism of Muscular Contraction. Science. 1969;164(3886):1356-66. Available from: http://dx.doi.org/10.1126/science.164.3886.1356.

[111] Gillies AR, Lieber RL. Structure and Function of the Skeletal Muscle Extracellular Matrix. Muscle & Nerve. 2011;44(3):318-31. Available from: http://dx.doi.org/10.1002/mus.22094.

[112] Wakelam MJ. The Fusion of Myoblasts. Biochemical Journal. 1985;228(1):1-12. Available from: http://dx.doi.org/10.1042/bj2280001.

[113] Chapman MA, Meza R, Lieber RL. Skeletal Muscle Fibroblasts in Health and Disease. Differentiation. 2016;92(3):108-15. Available from: http://dx.doi.org/10.1016/j.diff.2016.05.007.

[114] Kühl U, Öcalan M, Timpl R, Mayne R, Hay E, von der Mark K. Role of Muscle Fibroblasts in the Deposition of Type-Iv Collagen in the Basal Lamina of Myotubes. Differentiation. 1984;28(2):164-72. Available from: http://dx.doi.org/10.1111/j.1432-0436.1984.tb00279.x.

[115] Montarras D, Morgan J, Collins C, Relaix F, Zaffran S, Cumano A, et al. Direct Isolation of Satellite Cells for Skeletal Muscle Regeneration. Science. 2005;309(5743):2064-7. Available from: http://dx.doi.org/10.1126/science.1114758.

[116] Seale P, Sabourin LA, Girgis-Gabardo A, Mansouri A, Gruss P, Rudnicki MA. Pax7 Is Required for the Specification of Myogenic Satellite Cells. Cell. 2000;102(6):777-86. Available from: http://dx.doi.org/10.1016/s0092-8674(00)00066-0.

[117] Lepper C, Partridge TA, Fan CM. An Absolute Requirement for Pax7-positive Satellite Cells in Acute Injury-Induced Skeletal Muscle Regeneration. Development. 2011;138(17):3639-46. Available from: http://dx.doi.org/10.1242/dev.067595.

[118] Friedenstein A, Piatetzky-Shapiro I, Petrakova K. Osteogenesis in transplants of bone marrow cells. Development. 1966;16(3):381-90.

[119] Méndez-Ferrer S, Michurina TV, Ferraro F, Mazloom AR, MacArthur BD, Lira SA, et al. Mesenchymal and Haematopoietic Stem Cells Form a Unique Bone Marrow Niche. Nature. 2010;466(7308):829-34. Available from: http://dx.doi.org/10.1038/nature09262.

[120] hao Wang Y, ri Wang D, chen Guo Y, yuan Liu J, Pan J. The Application of Bone Marrow Mesenchymal Stem Cells and Biomaterials in Skeletal Muscle Regeneration. Regenerative Therapy. 2020;15(nil):285-94. Available from: http://dx.doi.org/10.1016/j.reth.2020.11.002.

[121] Zheng B, Cao B, Crisan M, Sun B, Li G, Logar A, et al. Prospective Identification of Myogenic Endothelial Cells in Human Skeletal Muscle. Nature Biotechnology. 2007;25(9):1025-34. Available from: http://dx.doi.org/10.1038/nbt1334.

[122] Nesti LJ, Jackson WM, Shanti RM, Koehler SM, Aragon AB, Bailey JR, et al. Differentiation Potential of Multipotent Progenitor Cells Derived From War-Traumatized Muscle Tissue. The Journal of Bone and Joint Surgery-American Volume. 2008;90(11):2390-8. Available from: http://dx.doi.org/10.2106/jbjs.h.00049.

[123] Jackson WM, Nesti LJ, Tuan RS. Potential Therapeutic Applications of Muscle-Derived Mesenchymal Stem and Progenitor Cells. Expert Opinion on Biological Therapy. 2010;10(4):505-17. Available from: http://dx.doi.org/10.1517/14712591003610606.

[124] Joe AWB, Yi L, Natarajan A, Grand FL, So L, Wang J, et al. Muscle Injury Activates Resident Fibro/adipogenic Progenitors That Facilitate Myogenesis. Nature Cell Biology. 2010;12(2):153-63. Available from: http://dx.doi.org/10.1038/ncb2015.

[125] Biferali B, Proietti D, Mozzetta C, Madaro L. Fibro-Adipogenic Progenitors Cross-Talk in Skeletal Muscle: the Social Network. Frontiers in Physiology. 2019;10(nil):nil. Available from: http://dx.doi.org/10.3389/fphys.2019.01074.

[126] Li S, Karri D, Sanchez-Ortiz E, Jaichander P, Bassel-Duby R, Liu N, et al. Sema3a-Nrp1 Signaling Mediates Fast-Twitch Myofiber Specificity of Tw2+ Cells. Developmental Cell. 2019;51(1):89-98.e4. Available from: http://dx.doi.org/10.1016/j.devcel.2019.08.002.

[127] Liu N, Garry GA, Li S, Bezprozvannaya S, Sanchez-Ortiz E, Chen B, et al. A Twist2-dependent Progenitor Cell Contributes To Adult Skeletal Muscle. Nature Cell Biology. 2017;19(3):202-13. Available from: http://dx.doi.org/10.1038/ncb3477.

[128] Schmalbruch H. Skeletal muscle. vol. 2. Springer Science & Business Media; 1985.

[129] Gustafsson T, Puntschart A, Kaijser L, Jansson E, Sundberg CJ. Exercise-Induced Expression of Angiogenesis-Related Transcription and Growth Factors in Human Skeletal Muscle. American Journal of Physiology-Heart and Circulatory

Physiology. 1999;276(2):H679-85. Available from: http://dx.doi.org/10.1152/ajpheart.1999.276.2.h679.

[130] Latroche C, Gitiaux C, Chrétien F, Desguerre I, Mounier R, Chazaud B. Skeletal Muscle Microvasculature: a Highly Dynamic Lifeline. Physiology. 2015;30(6):417-27. Available from: http://dx.doi.org/10.1152/physiol.00026.2015.

[131] Armulik A, Genové G, Betsholtz C. Pericytes: Developmental, Physiological, and Pathological Perspectives, Problems, and Promises. Developmental Cell. 2011;21(2):193-215. Available from: http://dx.doi.org/10.1016/j.devcel.2011.07.001.

[132] Tilton RG, Kilo C, Williamson JR. Pericyte-Endothelial Relationships in Cardiac and Skeletal Muscle Capillaries. Microvascular Research. 1979;18(3):325-35. Available from: http://dx.doi.org/10.1016/0026-2862(79)90041-4.

[133] Dellavalle A, Sampaolesi M, Tonlorenzi R, Tagliafico E, Sacchetti B, Perani L, et al. Pericytes of Human Skeletal Muscle Are Myogenic Precursors Distinct From Satellite Cells. Nature Cell Biology. 2007;9(3):255-67. Available from: http://dx.doi.org/10.1038/ncb1542.

[134] Dellavalle A, Maroli G, Covarello D, Azzoni E, Innocenzi A, Perani L, et al. Pericytes Resident in Postnatal Skeletal Muscle Differentiate Into Muscle Fibres and Generate Satellite Cells. Nature Communications. 2011;2(1):499. Available from: http://dx.doi.org/10.1038/ncomms1508.

[135] Kostallari E, Baba-Amer Y, Alonso-Martin S, Ngoh P, Relaix F, Lafuste P, et al. Pericytes in the Myovascular Niche Promote Post-Natal Myofiber Growth and Satellite Cell Quiescence. Development. 2015;nil(nil):nil. Available from: http://dx.doi.org/10.1242/dev.115386.

[136] Caputa G, Castoldi A, Pearce EJ. Metabolic Adaptations of Tissue-Resident Immune Cells. Nature Immunology. 2019;20(7):793-801. Available from: http://dx.doi.org/10.1038/s41590-019-0407-0.

[137] Pimorady-Esfahani A, Grounds MD, McMenamin PG. Macrophages and Dendritic Cells in Normal and Regenerating Murine Skeletal Muscle. Muscle &

Nerve. 1997;20(2):158-66. Available from: http://dx.doi.org/10.1002/(sici)1097-4598(199702)20:2<158::aid-mus4>3.0.co;2-b.

[138] Mackey AL, Kjaer M. The Breaking and Making of Healthy Adult Human Skeletal Muscle in Vivo. Skeletal Muscle. 2017;7(1):24. Available from: http://dx.doi.org/10.1186/s13395-017-0142-x.

[139] Kosmac K, Peck B, Walton R, Mula J, Kern P, Bamman M, et al. Immunohistochemical Identification of Human Skeletal Muscle Macrophages. BIO-PROTOCOL. 2018;8(12):nil. Available from: http://dx.doi.org/10.21769/bioprotoc.2883.

[140] Saclier M, Yacoub-Youssef H, Mackey AL, Arnold L, Ardjoune H, Magnan M, et al. Differentially Activated Macrophages Orchestrate Myogenic Precursor Cell Fate During Human Skeletal Muscle Regeneration. Stem Cells. 2013;31(2):384-96. Available from: http://dx.doi.org/10.1002/stem.1288.

[141] Shang M, Cappellesso F, Amorim R, Serneels J, Virga F, Eelen G, et al. Macrophage-Derived Glutamine Boosts Satellite Cells and Muscle Regeneration. Nature. 2020;587(7835):626-31. Available from: http://dx.doi.org/10.1038/s41586-020-2857-9.

[142] Latroche C, Weiss-Gayet M, Muller L, Gitiaux C, Leblanc P, Liot S, et al. Coupling Between Myogenesis and Angiogenesis During Skeletal Muscle Regeneration Is Stimulated By Restorative Macrophages. Stem Cell Reports. 2017;9(6):2018-33. Available from: http://dx.doi.org/10.1016/j.stemcr.2017.10.027.

[143] Taylor SM, Jones PA. Multiple New Phenotypes Induced in and 3t3 Cells Treated With 5-azacytidine. Cell. 1979;17(4):771-9. Available from: http://dx.doi.org/10.1016/0092-8674(79)90317-9.

[144] Jones PA, Taylor SM. Cellular Differentiation, Cytidine Analogs and Dna Methylation. Cell. 1980;20(1):85-93. Available from: http://dx.doi.org/10.1016/0092-8674(80)90237-8.

[145] Davis RL, Weintraub H, Lassar AB. Expression of a Single Transfected Cdna Converts Fibroblasts To Myoblasts. Cell. 1987;51(6):987-1000. Available from: http://dx.doi.org/10.1016/0092-8674(87)90585-X.

[146] Bar-Nur O, Gerli MFM, Stefano BD, Almada AE, Galvin A, Coffey A, et al. Direct Reprogramming of Mouse Fibroblasts Into Functional Skeletal Muscle Progenitors. Stem Cell Reports. 2018;10(5):1505-21. Available from: http://dx.doi.org/10.1016/j.stemcr.2018.04.009.

[147] Qabrati X, Kim I, Ghosh A, Bundschuh N, Noé F, Palmer AS, et al. Transgene-Free Direct Conversion of Murine Fibroblasts Into Functional Muscle Stem Cells. npj Regenerative Medicine. 2023;8(1):43. Available from: http://dx.doi.org/10.1038/s41536-023-00317-z.

[148] Carrió E, Díez-Villanueva A, Lois S, Mallona I, Cases I, Forn M, et al. Deconstruction of Dna Methylation Patterns During Myogenesis Reveals Specific Epigenetic Events in the Establishment of the Skeletal Muscle Lineage. Stem Cells. 2015;33(6):2025-36. Available from: http://dx.doi.org/10.1002/stem.1998.

[149] Brunk BP, Goldhamer DJ, Charles P Emerson J. Regulated Demethylation of the Myod Distal Enhancer During Skeletal Myogenesis. Developmental Biology. 1996;177(2):490-503. Available from: http://dx.doi.org/10.1006/dbio.1996.0180.

[150] Terruzzi I, Senesi P, Montesano A, Torre AL, Alberti G, Benedini S, et al. Genetic Polymorphisms of the Enzymes Involved in Dna Methylation and Synthesis in Elite Athletes. Physiological Genomics. 2011;43(16):965-73. Available from: http://dx.doi.org/10.1152/physiolgenomics.00040.2010.

[151] Montesano A, Luzi L, Senesi P, Terruzzi I. Modulation of Cell Cycle Progression By 5-azacytidine Is Associated With Early Myogenesis Induction in Murine Myoblasts. International Journal of Biological Sciences. 2013;9(4):391-402. Available from: http://dx.doi.org/10.7150/ijbs.4729.

[152] Laker RC, Ryall JG. Dna Methylation in Skeletal Muscle Stem Cell Specification, Proliferation, and Differentiation. Stem Cells International. 2016;2016(nil):1-9. Available from: http://dx.doi.org/10.1155/2016/5725927.

[153] Yang Y, Fan X, Yan J, Chen M, Zhu M, Tang Y, et al. A Comprehensive Epigenome Atlas Reveals Dna Methylation Regulating Skeletal Muscle Development. Nucleic

Acids Research. 2021;49(3):1313-29. Available from: http://dx.doi.org/10.1093/nar/gkaa1203.

[154] Zykovich A, Hubbard A, Flynn JM, Tarnopolsky M, Fraga MF, Kerksick C, et al. Genome-wide Dna Methylation Changes With Age in Disease-free Human Skeletal Muscle. Aging Cell. 2013;13(2):360-6. Available from: http://dx.doi.org/10.1111/acel.12180.

[155] Day K, Waite LL, Thalacker-Mercer A, West A, Bamman MM, Brooks JD, et al. Differential Dna Methylation With Age Displays Both Common and Dynamic Features Across Human Tissues That Are Influenced By Cpg Landscape. Genome Biology. 2013;14(9):R102. Available from: http://dx.doi.org/10.1186/gb-2013-14-9-r102.

[156] Turner DC, Gorski PP, Maasar MF, Seaborne RA, Baumert P, Brown AD, et al. Dna Methylation Across the Genome in Aged Human Skeletal Muscle Tissue and Muscle-Derived Cells: the Role of Hox Genes and Physical Activity. Scientific Reports. 2020;10(1):15360. Available from: http://dx.doi.org/10.1038/s41598-020-72730-z.

[157] Bigot A, Duddy WJ, Ouandaogo ZG, Negroni E, Mariot V, Ghimbovschi S, et al. Age-Associated Methylation Suppresses Spry1 , Leading To a Failure of Re-Quiescence and Loss of the Reserve Stem Cell Pool in Elderly Muscle. Cell Reports. 2015;13(6):1172-82. Available from: http://dx.doi.org/10.1016/j.celrep.2015.09.067.

[158] Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. From the Cover: Epigenetic Differences Arise During the Lifetime of Monozygotic Twins. Proceedings of the National Academy of Sciences. 2005;102(30):10604-9. Available from: http://dx.doi.org/10.1073/pnas.0500398102.

[159] Voisin S, Jacques M, Landen S, Harvey NR, Haupt LM, Griffiths LR, et al. Meta-analysis of Genome-wide Dna Methylation and Integrative Omics of Age in Human Skeletal Muscle. Journal of Cachexia, Sarcopenia and Muscle. 2021;12(4):1064-78. Available from: http://dx.doi.org/10.1002/jcsm.12741.

[160] Gensous N, Bacalini MG, Franceschi C, Meskers CGM, Maier AB, Garagnani P. Age-Related Dna Methylation Changes: Potential Impact on Skeletal Muscle Aging in Humans. Frontiers in Physiology. 2019;10(nil):nil. Available from: http://dx.doi.org/10.3389/fphys.2019.00996.

[161] Janssen I, Heymsfield SB, Wang Z, Ross R. Skeletal Muscle Mass and Distribution in 468 Men and Women Aged 18-88 Yr. Journal of Applied Physiology. 2000;89(1):81-8. Available from: http://dx.doi.org/10.1152/jappl.2000.89.1.81.

[162] Miotto PM, McGlory C, Holloway TM, Phillips SM, Holloway GP. Sex Differences in Mitochondrial Respiratory Function in Human Skeletal Muscle. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology. 2018;314(6):R909-15. Available from: http://dx.doi.org/10.1152/ajpregu.00025.2018.

[163] Wüst RCI, Morse CI, Haan AD, Jones DA, Degens H. Sex Differences in Contractile Properties and Fatigue Resistance of Human Skeletal Muscle. Experimental Physiology. 2008;93(7):843-50. Available from: http://dx.doi.org/10.1113/expphysiol.2007.041764.

[164] Lindholm ME, Huss M, Solnestam BW, Kjellqvist S, Lundeberg J, Sundberg CJ. The Human Skeletal Muscle Transcriptome: Sex Differences, Alternative Splicing, and Tissue Homogeneity Assessed With Rna Sequencing. The FASEB Journal. 2014;28(10):4571-81. Available from: http://dx.doi.org/10.1096/fj.14-255000.

[165] Welle S, Tawil R, Thornton CA. Sex-Related Differences in Gene Expression in Human Skeletal Muscle. PLoS ONE. 2008;3(1):e1385. Available from: http://dx.doi.org/10.1371/journal.pone.0001385.

[166] Davegårdh C, Wedin EH, Broholm C, Henriksen TI, Pedersen M, Pedersen BK, et al. Sex Influences Dna Methylation and Gene Expression in Human Skeletal Muscle Myoblasts and Myotubes. Stem Cell Research amp; Therapy. 2019;10(1):26. Available from: http://dx.doi.org/10.1186/s13287-018-1118-4.

[167] Landen S, Jacques M, Hiam D, Alvarez-Romero J, Harvey NR, Haupt LM, et al. Skeletal Muscle Methylome and Transcriptome Integration Reveals Profound Sex Differences Related To Muscle Function and Substrate Metabolism. Clin-

ical Epigenetics. 2021;13(1):202. Available from: http://dx.doi.org/10.1186/s13148-021-01188-1.

[168] Yousefi PD, Suderman M, Langdon R, Whitehurst O, Smith GD, Relton CL. Dna Methylation-Based Predictors of Health: Applications and Statistical Considerations. Nature Reviews Genetics. 2022;23(6):369-83. Available from: http://dx.doi.org/10.1038/s41576-022-00465-w.

[169] Davegårdh C, García-Calzón S, Bacos K, Ling C. Dna Methylation in the Pathogenesis of Type 2 Diabetes in Humans. Molecular Metabolism. 2018;14(nil):12-25. Available from: http://dx.doi.org/10.1016/j.molmet.2018.01.022.

[170] Nitert MD, Dayeh T, Volkov P, Elgzyri T, Hall E, Nilsson E, et al. Impact of an Exercise Intervention on Dna Methylation in Skeletal Muscle From First-Degree Relatives of Patients With Type 2 Diabetes. Diabetes. 2012;61(12):3322-32. Available from: http://dx.doi.org/10.2337/db11-1653.

[171] Ribel-Madsen R, Fraga MF, Jacobsen S, Bork-Jensen J, Lara E, Calvanese V, et al. Genome-Wide Analysis of Dna Methylation Differences in Muscle and Fat From Monozygotic Twins Discordant for Type 2 Diabetes. PLoS ONE. 2012;7(12):e51302. Available from: http://dx.doi.org/10.1371/journal.pone.0051302.

[172] Barrès R, Osler ME, Yan J, Rune A, Fritz T, Caidahl K, et al. Non-Cpg Methylation of the PGC-1α Promoter Through Dnmt3b Controls Mitochondrial Density. Cell Metabolism. 2009;10(3):189-98. Available from: http://dx.doi.org/10.1016/j.cmet.2009.07.011.

[173] Jacobsen SC, Brøns C, Bork-Jensen J, Ribel-Madsen R, Yang B, Lara E, et al. Effects of Short-Term High-Fat Overfeeding on Genome-Wide Dna Methylation in the Skeletal Muscle of Healthy Young Men. Diabetologia. 2012;55(12):3341-9. Available from: http://dx.doi.org/10.1007/s00125-012-2717-8.

[174] Barrès R, Yan J, Egan B, Treebak JT, Rasmussen M, Fritz T, et al. Acute Exercise Remodels Promoter Methylation in Human Skeletal Muscle. Cell Metabolism. 2012;15(3):405-11. Available from: http://dx.doi.org/10.1016/j.cmet.2012.01.001.

[175] Lindholm ME, Marabita F, Gomez-Cabrero D, Rundqvist H, Ekström TJ, Tegnér J, et al. An Integrative Analysis Reveals Coordinated Reprogramming of the Epigenome and the Transcriptome in Human Skeletal Muscle After Training. Epigenetics. 2014;9(12):1557-69. Available from: http://dx.doi.org/10.4161/15592294.2014.982445.

[176] Robinson MM, Dasari S, Konopka AR, Johnson ML, Manjunatha S, Esponda RR, et al. Enhanced Protein Translation Underlies Improved Metabolic and Physical Adaptations To Different Exercise Training Modes in Young and Old Humans. Cell Metabolism. 2017;25(3):581-92. Available from: http://dx.doi.org/10.1016/j.cmet.2017.02.009.

[177] Seaborne RA, Strauss J, Cocks M, Shepherd S, O'Brien TD, van Someren KA, et al. Human Skeletal Muscle Possesses an Epigenetic Memory of Hypertrophy. Scientific Reports. 2018;8(1):1898. Available from: http://dx.doi.org/10.1038/s41598-018-20287-3.

[178] Blocquiaux S, Ramaekers M, Thienen RV, Nielens H, Delecluse C, Bock KD, et al. Recurrent Training Rejuvenates and Enhances Transcriptome and Methylome Responses in Young and Older Human Muscle. JCSM Rapid Communications. 2021;5(1):10-32. Available from: http://dx.doi.org/10.1002/rco2.52.

[179] Liu D, Sartor MA, Nader GA, Gutmann L, Treutelaar MK, Pistilli EE, et al. Skeletal Muscle Gene Expression in Response To Resistance Exercise: Sex Specific Regulation. BMC Genomics. 2010;11(1):659. Available from: http://dx.doi.org/10.1186/1471-2164-11-659.

[180] Roepstorff C, Schjerling P, Vistisen B, Madsen M, Steffensen CH, Rider MH, et al. Regulation of Oxidative Enzyme Activity and Eukaryotic Elongation Factor 2 in Human Skeletal Muscle: Influence of Gender and Exercise. Acta Physiologica Scandinavica. 2005;184(3):215-24. Available from: http://dx.doi.org/10.1111/j.1365-201X.2005.01442.x.

[181] Landen S, Jacques M, Hiam D, Alvarez-Romero J, Schittenhelm RB, Shah AD, et al. Sex Differences in Muscle Protein Expression and Dna Methylation in Response To Exercise Training. Biology of Sex Differences. 2023;14(1):56. Available from: http://dx.doi.org/10.1186/s13293-023-00539-2.

[182] Hargreaves M, Spriet LL. Skeletal Muscle Energy Metabolism During Exercise. Nature Metabolism. 2020;2(9):817-28. Available from: http://dx.doi.org/10.1038/s42255-020-0251-4.

[183] Bottinelli R, Reggiani C. Human Skeletal Muscle Fibres: Molecular and Functional Diversity. Progress in Biophysics and Molecular Biology. 2000;73(2-4):195-262. Available from: http://dx.doi.org/10.1016/S0079-6107(00)00006-7.

[184] Plotkin DL, Roberts MD, Haun CT, Schoenfeld BJ. Muscle Fiber Type Transitions With Exercise Training: Shifting Perspectives. Sports. 2021;9(9):127. Available from: http://dx.doi.org/10.3390/sports9090127.

[185] Bengtsen M, Winje IM, Eftestøl E, Landskron J, Sun C, Nygård K, et al. Comparing the Epigenetic Landscape in Myonuclei Purified With a Pcm1 Antibody From a Fast/glycolytic and a Slow/oxidative Muscle. PLOS Genetics. 2021;17(11):e1009907. Available from: http://dx.doi.org/10.1371/journal.pgen.1009907.

[186] Borowik AK, Davidyan A, Peelor FF, Voloviceva E, Doidge SM, Bubak MP, et al. Skeletal Muscle Nuclei in Mice Are Not Post-Mitotic. Function. 2022;4(1):nil. Available from: http://dx.doi.org/10.1093/function/zqac059.

[187] Perez K, Ciotlos S, McGirr J, Limbad C, Doi R, Nederveen JP, et al. Single Nuclei Profiling Identifies Cell Specific Markers of Skeletal Muscle Aging, Frailty, and Senescence. Aging. 2022;nil(nil):nil. Available from: http://dx.doi.org/10.18632/aging.204435.

[188] Chalmers GR, Row BS. Common Errors in Textbook Descriptions of Muscle Fiber Size in Nontrained Humans. Sports Biomechanics. 2011;10(3):254-68. Available from: http://dx.doi.org/10.1080/14763141.2011.592211.

[189] Lexell J, Henriksson-Larsén K, Winblad B, Sjöström M. Distribution of Different Fiber Types in Human Skeletal Muscles: Effects of Aging Studied in Whole Muscle Cross Sections. Muscle amp; Nerve. 1983;6(8):588-95. Available from: http://dx.doi.org/10.1002/mus.880060809.

[190] Nuzzo JL. Sex Differences in Skeletal Muscle Fiber Types: a Meta-analysis. Clinical Anatomy. 2023;37(1):81-91. Available from: http://dx.doi.org/10.1002/ca.

24091.

[191] Henriksson-Larsen KB, Lexell J, Sjoestroem M. Distribution of Different Fibre Types in Human Skeletal Muscles. I. Method for the Preparation and Analysis of Cross-Sections of Whole Tibialis Anterior. The Histochemical Journal. 1983;15(2):167-78. Available from: http://dx.doi.org/10.1007/BF01042285.

[192] Larsson L, Moss RL. Maximum Velocity of Shortening in Relation To Myosin Isoform Composition in Single Fibres From Human Skeletal Muscles. The Journal of Physiology. 1993;472(1):595-614. Available from: http://dx.doi.org/10.1113/jphysiol.1993.sp019964.

[193] Bottinelli R, Canepari M, Pellegrino MA, Reggiani C. Force-Velocity Properties of Human Skeletal Muscle Fibres: Myosin Heavy Chain Isoform and Temperature Dependence. The Journal of Physiology. 1996;495(2):573-86. Available from: http://dx.doi.org/10.1113/jphysiol.1996.sp021617.

[194] Howald H, Hoppeler H, Claassen H, Mathieu O, Straub R. Influences of Endurance Training on the Ultrastructural Composition of the Different Muscle Fiber Types in Humans. Pflgers Archiv European Journal of Physiology. 1985;403(4):369-76. Available from: http://dx.doi.org/10.1007/bf00589248.

[195] Ogata T, Yamasaki Y. Ultra-High-Resolution Scanning Electron Microscopy of Mitochondria and Sarcoplasmic Reticulum Arrangement in Human Red, White, and Intermediate Muscle Fibers. The Anatomical Record. 1997;248(2):214-23. Available from: http://dx.doi.org/10.1002/(SICI)1097-0185(199706)248:2<214::AID-AR8>3.0.CO;2-S.

[196] Dahl R, Larsen S, Dohlmann TL, Qvortrup K, Helge JW, Dela F, et al. Three-Dimensional Reconstruction of the Human Skeletal Muscle Mitochondrial Network As a Tool To Assess Mitochondrial Content and Structural Organization. Acta Physiologica. 2014;213(1):145-55. Available from: http://dx.doi.org/10.1111/apha.12289.

[197] Essén B, Jansson E, Henriksson J, Taylor AW, Saltin B. Metabolic Characteristics of Fibre Types in Human Skeletal Muscle. Acta Physiologica Scandinavica.

1975;95(2):153-65. Available from: http://dx.doi.org/10.1111/j.1748-1716.1975.tb10038.x.

[198] Lowry C, Kimmey J, Felder S, Chi M, Kaiser K, Passonneau P, et al. Enzyme patterns in single human muscle fibers. The Journal of Biological Chemistry. 1978;253(22):8269-77.

[199] Naro F, Venturelli M, Monaco L, Toniolo L, Muti E, Milanese C, et al. Skeletal Muscle Fiber Size and Gene Expression in the Oldest-Old With Differing Degrees of Mobility. Frontiers in Physiology. 2019;10(nil):nil. Available from: http://dx.doi.org/10.3389/fphys.2019.00313.

[200] Deschenes MR. Effects of Aging on Muscle Fibre Type and Size. Sports Medicine. 2004;34(12):809-24. Available from: http://dx.doi.org/10.2165/00007256-200434120-00002.

[201] Grosicki GJ, Zepeda CS, Sundberg CW. Single Muscle Fibre Contractile Function With Ageing. The Journal of Physiology. 2022;nil(nil):nil. Available from: http://dx.doi.org/10.1113/JP282298.

[202] Grosicki GJ, Standley RA, Murach KA, Raue U, Minchev K, Coen PM, et al. Improved Single Muscle Fiber Quality in the Oldest-Old. Journal of Applied Physiology. 2016;121(4):878-84. Available from: http://dx.doi.org/10.1152/japplphysiol.00479.2016.

[203] Murgia M, Toniolo L, Nagaraj N, Ciciliot S, Vindigni V, Schiaffino S, et al. Single Muscle Fiber Proteomics Reveals Fiber-Type-Specific Features of Human Muscle Aging. Cell Reports. 2017;19(11):2396-409. Available from: http://dx.doi.org/10.1016/j.celrep.2017.05.054.

[204] Staron RS, Hagerman FC, Hikida RS, Murray TF, Hostler DP, Crill MT, et al. Fiber Type Composition of the Vastus Lateralis Muscle of Young Men and Women. Journal of Histochemistry & Cytochemistry. 2000;48(5):623-9. Available from: http://dx.doi.org/10.1177/002215540004800506.

[205] Teigen LE, Sundberg CW, Kelly LJ, Hunter SK, Fitts RH. Ca2+ dependency of Limb Muscle Fiber Contractile Mechanics in Young and Older Adults. American Journal

of Physiology-Cell Physiology. 2020;318(6):C1238-51. Available from: http://dx.doi.org/10.1152/ajpcell.00575.2019.

[206] Horwath O, Moberg M, Larsen FJ, Philp A, Apró W, Ekblom B. Influence of Sex and Fiber Type on the Satellite Cell Pool in Human Skeletal Muscle. Scandinavian Journal of Medicine & Science in Sports. 2020;31(2):303-12. Available from: http://dx.doi.org/10.1111/sms.13848.

[207] Spangenburg EE, Booth FW. Molecular Regulation of Individual Skeletal Muscle Fibre Types. Acta Physiologica Scandinavica. 2003;178(4):413-24. Available from: http://dx.doi.org/10.1046/j.1365-201X.2003.01158.x.

[208] Chemello F, Bean C, Cancellara P, Laveder P, Reggiani C, Lanfranchi G. Microgenomic Analysis in Skeletal Muscle: Expression Signatures of Individual Fast and Slow Myofibers. PLoS ONE. 2011;6(2):e16807. Available from: http://dx.doi.org/10.1371/journal.pone.0016807.

[209] Murgia M, Nogara L, Baraldo M, Reggiani C, Mann M, Schiaffino S. Protein Profile of Fiber Types in Human Skeletal Muscle: a Single-Fiber Proteomics Study. Skeletal Muscle. 2021;11(1):24. Available from: http://dx.doi.org/10.1186/s13395-021-00279-0.

[210] Deshmukh AS, Steenberg DE, Hostrup M, Birk JB, Larsen JK, Santos A, et al. Deep Muscle-Proteomic Analysis of Freeze-Dried Human Muscle Biopsies Reveals Fiber Type-Specific Adaptations To Exercise Training. Nature Communications. 2021;12(1):304. Available from: http://dx.doi.org/10.1038/s41467-020-20556-8.

[211] Moreno-Justicia R, der Stede TV, Stocks B, Laitila J, Seaborne RA, de Loock AV, et al.. Human Skeletal Muscle Fiber Heterogeneity Beyond Myosin Heavy Chains; 2023. Available from: http://dx.doi.org/10.1101/2023.09.07.556665.

[212] Schiaffino S, Reggiani C, Murgia M. Fiber Type Diversity in Skeletal Muscle Explored By Mass Spectrometry-Based Single Fiber Proteomics. Histology and Histopathology. 2020;35(03):239-46. Available from: https://doi.org/10.14670/HH-18-170.

[213] Oe M, Ojima K, Muroya S. Difference in Potential Dna Methylation Impact on Gene Expression Between Fast- and Slow-Type Myofibers. Physiological Genomics. 2021;53(2):69-83. Available from: http://dx.doi.org/10.1152/physiolgenomics.00099.2020.

[214] Taylor DL, Jackson AU, Narisu N, Hemani G, Erdos MR, Chines PS, et al. Integrative Analysis of Gene Expression, Dna Methylation, Physiological Traits, and Genetic Variation in Human Skeletal Muscle. Proceedings of the National Academy of Sciences. 2019;116(22):10883-8. Available from: http://dx.doi.org/10.1073/pnas.1814263116.

[215] Gollnick PD, Armstrong RB, Saubert CW, Piehl K, Saltin B. Enzyme Activity and Fiber Composition in Skeletal Muscle of Untrained and Trained Men. Journal of Applied Physiology. 1972;33(3):312-9. Available from: http://dx.doi.org/10.1152/jappl.1972.33.3.312.

[216] Andersen P, Henriksson J. Training Induced Changes in the Subgroups of Human Type Ii Skeletal Muscle Fibres. Acta Physiologica Scandinavica. 1977;99(1):123-5. Available from: http://dx.doi.org/10.1111/j.1748-1716.1977.tb10361.x.

[217] Jansson E, Sjödin B, Tesch P. Changes in Muscle Fibre Type Distribution in Man After Physical Training: a Sign of Fibre Type Transformation? Acta Physiologica Scandinavica. 1978;104(2):235-7. Available from: http://dx.doi.org/10.1111/j.1748-1716.1978.tb06272.x.

[218] Houston ME, Wilson DM, Green HJ, Thomson JA, Ranney DA. Physiological and Muscle Enzyme Adaptations To Two Different Intensities of Swim Training. European Journal of Applied Physiology and Occupational Physiology. 1981;46(3):283-91. Available from: http://dx.doi.org/10.1007/bf00423404.

[219] Simoneau JA, Lortie G, Boulay MR, Marcotte M, Thibault MC, Bouchard C. Human Skeletal Muscle Fiber Type Alteration With High-Intensity Intermittent Training. European Journal of Applied Physiology and Occupational Physiology. 1985;54(3):250-3. Available from: http://dx.doi.org/10.1007/BF00426141.

[220] Baumann H, Jäggi M, Soland F, Howald H, Schaub MC. Exercise Training Induces Transitions of Myosin Isoform Subunits Within Histochemically Typed Hu-

man Muscle Fibres. Pflügers Archiv - European Journal of Physiology. 1987;409(4-5):349-60. Available from: http://dx.doi.org/10.1007/bf00583788.

[221] Jansson E, Esbjörnsson M, Holm I, Jacobs I. Increase in the Proportion of Fast-Twitch Muscle Fibres By Sprint Training in Males. Acta Physiologica Scandinavica. 1990;140(3):359-63. Available from: http://dx.doi.org/10.1111/j.1748-1716.1990.tb09010.x.

[222] Allemeier CA, Fry AC, Johnson P, Hikida RS, Hagerman FC, Staron RS. Effects of Sprint Cycle Training on Human Skeletal Muscle. Journal of Applied Physiology. 1994;77(5):2385-90. Available from: http://dx.doi.org/10.1152/jappl.1994.77.5.2385.

[223] Costill DL, Coyle EF, Fink WF, Lesmes GR, Witzmann FA. Adaptations in Skeletal Muscle Following Strength Training. Journal of Applied Physiology. 1979;46(1):96-9. Available from: http://dx.doi.org/10.1152/jappl.1979.46.1.96.

[224] MacDougall J, Sale D, Moroz J, Elder G, Sutton J, Howald H. Mitochondrial volume density in human skeletal muscle following heavy resistance training. Medicine and science in sports. 1979;11(2):164-6.

[225] Tesch P, Komi P, Häkkinen K. Enzymatic Adaptations Consequent To Long-Term Strength Training*. International Journal of Sports Medicine. 1987;08(S 1):S66-9. Available from: http://dx.doi.org/10.1055/s-2008-1025706.

[226] Frontera WR, Meredith CN, O'Reilly KP, Knuttgen HG, Evans WJ. Strength Conditioning in Older Men: Skeletal Muscle Hypertrophy and Improved Function. Journal of Applied Physiology. 1988;64(3):1038-44. Available from: http://dx.doi.org/10.1152/jappl.1988.64.3.1038.

[227] Staron RS, Malicky ES, Leonardi MJ, Falkel JE, Hagerman FC, Dudley GA. Muscle Hypertrophy and Fast Fiber Type Conversions in Heavy Resistance-Trained Women. European Journal of Applied Physiology and Occupational Physiology. 1990;60(1):71-9. Available from: http://dx.doi.org/10.1007/bf00572189.

[228] Hather BM, Tesch PA, Buchanan P, Dudley GA. Influence of Eccentric Actions on Skeletal Muscle Adaptations To Resistance Training. Acta Physiologica Scan-

dinavica. 1991;143(2):177-85. Available from: http://dx.doi.org/10.1111/j.1748-1716.1991.tb09219.x.

[229] Staron RS, Leonardi MJ, Karapondo DL, Malicky ES, Falkel JE, Hagerman FC, et al. Strength and Skeletal Muscle Adaptations in Heavy-Resistance-Trained Women After Detraining and Retraining. Journal of Applied Physiology. 1991;70(2):631-40. Available from: http://dx.doi.org/10.1152/jappl.1991.70.2.631.

[230] Adams GR, Hather BM, Baldwin KM, Dudley GA. Skeletal Muscle Myosin Heavy Chain Composition and Resistance Training. Journal of Applied Physiology. 1993;74(2):911-5. Available from: http://dx.doi.org/10.1152/jappl.1993.74.2.911.

[231] Staron RS, Karapondo DL, Kraemer WJ, Fry AC, Gordon SE, Falkel JE, et al. Skeletal Muscle Adaptations During Early Phase of Heavy-Resistance Training in Men and Women. Journal of Applied Physiology. 1994;76(3):1247-55. Available from: http://dx.doi.org/10.1152/jappl.1994.76.3.1247.

[232] Bamman MM, Hill VJ, Adams GR, Haddad F, Wetzstein CJ, Gower BA, et al. Gender Differences in Resistance-Training-Induced Myofiber Hypertrophy Among Older Adults. The Journals of Gerontology: Series A. 2003;58(2):B108-16. Available from: http://dx.doi.org/10.1093/gerona/58.2.b108.

[233] Bickel CS, Cross JM, Bamman MM. Exercise Dosing To Retain Resistance Training Adaptations in Young and Older Adults. Medicine & Science in Sports & Exercise. 2011;43(7):1177-87. Available from: http://dx.doi.org/10.1249/mss.0b013e318207c15d.

[234] Schantz P, Billeter R, Henriksson J, Jansson E. Training-Induced Increase in Myofibrillar Atpase Intermediate Fibers in Human Skeletal Muscle. Muscle & Nerve. 1982;5(8):628-36. Available from: http://dx.doi.org/10.1002/mus.880050807.

[235] Schantz P, Henriksson J. Increases in Myofibrillar Atpase Intermediate Human Skeletal Muscle Fibers in Response To Endurance Training. Muscle & Nerve. 1983;6(8):553-6. Available from: http://dx.doi.org/10.1002/mus.880060803.

[236] Williamson DL, Godard MP, Porter DA, Costill DL, Trappe SW. Progressive Resistance Training Reduces Myosin Heavy Chain Coexpression in Single Muscle Fibers

From Older Men. Journal of Applied Physiology. 2000;88(2):627-33. Available from: http://dx.doi.org/10.1152/jappl.2000.88.2.627.

[237] D'Antona G, Pellegrino MA, Carlizzi CN, Bottinelli R. Deterioration of Contractile Properties of Muscle Fibres in Elderly Subjects Is Modulated By the Level of Physical Activity. European Journal of Applied Physiology. 2007;100(5):603-11. Available from: http://dx.doi.org/10.1007/s00421-007-0402-2.

[238] Gries KJ, Minchev K, Raue U, Grosicki GJ, Begue G, Finch WH, et al. Single-Muscle Fiber Contractile Properties in Lifelong Aerobic Exercising Women. Journal of Applied Physiology. 2019;127(6):1710-9. Available from: http://dx.doi.org/10.1152/japplphysiol.00459.2019.

[239] Trappe S, Williamson D, Godard M, Porter D, Rowden G, Costill D. Effect of Resistance Training on Single Muscle Fiber Contractile Function in Older Men. Journal of Applied Physiology. 2000;89(1):143-52. Available from: http://dx.doi.org/10.1152/jappl.2000.89.1.143.

[240] Trappe S, Godard M, Gallagher P, Carroll C, Rowden G, Porter D. Resistance Training Improves Single Muscle Fiber Contractile Function in Older Women. American Journal of Physiology-Cell Physiology. 2001;281(2):C398-406. Available from: http://dx.doi.org/10.1152/ajpcell.2001.281.2.c398.

[241] Pansarasa O, Rinaldi C, Parente V, Miotti D, Capodaglio P, Bottinelli R. Resistance Training of Long Duration Modulates Force and Unloaded Shortening Velocity of Single Muscle Fibres of Young Women. Journal of Electromyography and Kinesiology. 2009;19(5):e290-300. Available from: http://dx.doi.org/10.1016/j.jelekin.2008.07.007.

[242] Raue U, Slivka D, Minchev K, Trappe S. Improvements in Whole Muscle and Myocellular Function Are Limited With High-Intensity Resistance Training in Octogenarian Women. Journal of Applied Physiology. 2009;106(5):1611-7. Available from: http://dx.doi.org/10.1152/japplphysiol.91587.2008.

[243] Slivka D, Raue U, Hollon C, Minchev K, Trappe S. Single Muscle Fiber Adaptations To Resistance Training in Old (>80 yr) Men: Evidence for Limited Skeletal Muscle Plasticity. American Journal of Physiology-Regulatory, Integrative and Compar-

ative Physiology. 2008;295(1):R273-80. Available from: http://dx.doi.org/10.1152/ajpregu.00093.2008.

[244] Trappe S, Harber M, Creer A, Gallagher P, Slivka D, Minchev K, et al. Single Muscle Fiber Adaptations With Marathon Training. Journal of Applied Physiology. 2006;101(3):721-7. Available from: http://dx.doi.org/10.1152/japplphysiol.01595.2005.

[245] Harber MP, Konopka AR, Douglass MD, Minchev K, Kaminsky LA, Trappe TA, et al. Aerobic Exercise Training Improves Whole Muscle and Single Myofiber Size and Function in Older Women. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology. 2009;297(5):R1452-9. Available from: http://dx.doi.org/10.1152/ajpregu.00354.2009.

[246] Harber MP, Konopka AR, Undem MK, Hinkley JM, Minchev K, Kaminsky LA, et al. Aerobic Exercise Training Induces Skeletal Muscle Hypertrophy and Age-Dependent Adaptations in Myofiber Function in Young and Older Men. Journal of Applied Physiology. 2012;113(9):1495-504. Available from: http://dx.doi.org/10.1152/japplphysiol.00786.2012.

[247] Esbjörnsson-Liljedahl M, Sundberg CJ, Norman B, Jansson E. Metabolic Response in Type I and Type II Muscle Fibers During a 30-s Cycle Sprint in Men and Women. Journal of Applied Physiology. 1999;87(4):1326-32. Available from: http://dx.doi.org/10.1152/jappl.1999.87.4.1326.

[248] Karatzaferi C, de Haan A, van Mechelen W, Sargeant AJ. Metabolic Changes in Single Human Muscle Fibres During Brief Maximal Exercise. Experimental Physiology. 2001;86(3):411-5. Available from: http://dx.doi.org/10.1113/eph8602223.

[249] Vigh-Larsen JF, Ørtenblad N, Andersen OE, Thorsteinsson H, Kristiansen TH, Bilde S, et al. Fibre Type- and Localisation-specific Muscle Glycogen Utilisation During Repeated High-intensity Intermittent Exercise. The Journal of Physiology. 2022;600(21):4713-30. Available from: http://dx.doi.org/10.1113/JP283225.

[250] Gollnick PD, Armstrong RB, Sembrowich WL, Shepherd RE, Saltin B. Glycogen Depletion Pattern in Human Skeletal Muscle Fibers After Heavy Exercise. Journal

of Applied Physiology. 1973;34(5):615-8. Available from: http://dx.doi.org/10.1152/jappl.1973.34.5.615.

[251] Jemiolo B, Trappe S. Single Muscle Fiber Gene Expression in Human Skeletal Muscle: Validation of Internal Control With Exercise. Biochemical and Biophysical Research Communications. 2004;320(3):1043-50. Available from: http://dx.doi.org/10.1016/j.bbrc.2004.05.223.

[252] Konopka AR, Trappe TA, Jemiolo B, Trappe SW, Harber MP. Myosin Heavy Chain Plasticity in Aging Skeletal Muscle With Aerobic Exercise Training. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2011;66A(8):835-41. Available from: http://dx.doi.org/10.1093/gerona/glr088.

[253] Yang Y, Jemiolo B, Trappe S. Proteolytic Mrna Expression in Response To Acute Resistance Exercise in Human Single Skeletal Muscle Fibers. Journal of Applied Physiology. 2006;101(5):1442-50. Available from: http://dx.doi.org/10.1152/japplphysiol.00438.2006.

[254] Raue U, Trappe TA, Estrem ST, Qian HR, Helvering LM, Smith RC, et al. Transcriptome Signature of Resistance Exercise Adaptations: Mixed Muscle and Fiber Type Specific Profiles in Young and Old Adults. Journal of Applied Physiology. 2012;112(10):1625-36. Available from: http://dx.doi.org/10.1152/japplphysiol.00435.2011.

[255] Koopman R, Gleeson BG, Gijsen AP, Groen B, Senden JMG, Rennie MJ, et al. Post-Exercise Protein Synthesis Rates Are Only Marginally Higher in Type I Compared With Type II Muscle Fibres Following Resistance-Type Exercise. European Journal of Applied Physiology. 2011;111(8):1871-8. Available from: http://dx.doi.org/10.1007/s00421-010-1808-9.

[256] Feil R, Fraga MF. Epigenetics and the Environment: Emerging Patterns and Implications. Nature Reviews Genetics. 2012;13(2):97-109. Available from: http://dx.doi.org/10.1038/nrg3142.

[257] Bellamy LM, Joanisse S, Grubb A, Mitchell CJ, McKay BR, Phillips SM, et al. The Acute Satellite Cell Response and Skeletal Muscle Hypertrophy Following Resis-

tance Training. PLoS ONE. 2014;9(10):e109739. Available from: http://dx.doi.org/10.1371/journal.pone.0109739.

[258] Qaisar R, Bhaskaran S, Remmen HV. Muscle Fiber Type Diversification During Exercise and Regeneration. Free Radical Biology and Medicine. 2016;98(nil):56-67. Available from: http://dx.doi.org/10.1016/j.freeradbiomed.2016.03.025.

[259] Kohn TA, Essén-Gustavsson B, Myburgh KH. Specific Muscle Adaptations in Type Ii Fibers After High-Intensity Interval Training of Well-Trained Runners. Scandinavian Journal of Medicine & Science in Sports. 2010;21(6):765-72. Available from: http://dx.doi.org/10.1111/j.1600-0838.2010.01136.x.

[260] Yan X, Eynon N, Papadimitriou ID, Kuang J, Munson F, Tirosh O, et al. The Gene Smart Study: Method, Study Design, and Preliminary Findings. BMC Genomics. 2017;18(S8):821. Available from: http://dx.doi.org/10.1186/s12864-017-4186-4.

[261] Bishop D, Jenkins DG, Mackinnon LT. The Relationship Between Plasma Lactate Parameters, Wpeak and 1-h Cycling Performance in Women. Medicineamp; Science in Sports amp; Exercise. 1998;30(8):1270-5. Available from: http://dx.doi.org/10.1097/00005768-199808000-00014.

[262] Newell J, Higgins D, Madden N, Cruickshank J, Einbeck J, McMillan K, et al. Software for Calculating Blood Lactate Endurance Markers. Journal of Sports Sciences. 2007;25(12):1403-9. Available from: http://dx.doi.org/10.1080/02640410601128922.

[263] Jacques M, Landen S, Romero JA, Hiam D, Schittenhelm RB, Hanchapola I, et al. Methylome and Proteome Integration in Human Skeletal Muscle Uncover Group and Individual Responses To High-intensity Interval Training. The FASEB Journal. 2023;37(10):nil. Available from: http://dx.doi.org/10.1096/fj.202300840RR.

[264] Jacques M, Landen S, Romero JA, Yan X, Hiam D, Jones P, et al. Implementation of Multiple Statistical Methods To Estimate Variability and Individual Response To Training. European Journal of Sport Science. 2022;23(4):588-98. Available from: http://dx.doi.org/10.1080/17461391.2022.2048894.

[265] Mifflin M, Jeor SS, Hill L, Scott B, Daugherty S, Koh Y. A New Predictive Equation for Resting Energy Expenditure in Healthy Individuals. The American Journal of Clinical Nutrition. 1990;51(2):241-7. Available from: http://dx.doi.org/10.1093/ajcn/51.2.241.

[266] National Health and Medical Research Council. Australian Dietary Guidelines: Providing the Scientific Evidence for Healthier Australian Diets; 2013. Available from: https://www.nhmrc.gov.au/_files_nhmrc/file/publications/n55_australian_dietary_guidelines1.pdf.

[267] Bergstrom J. Muscle electrolytes in man determined by neutron activation analysis on needle biopsy specimens. Scandinavian Journal of Clinical and Laboratory Investigation (England). 1962;14(Suppl 68).

[268] Christiansen D, MacInnis MJ, Zacharewicz E, Xu H, Frankish BP, Murphy RM. A Fast, Reliable and Sample-Sparing Method To Identify Fibre Types of Single Muscle Fibres. Scientific Reports. 2019;9(1):6473. Available from: http://dx.doi.org/10.1038/s41598-019-42168-z.

[269] Gautam A. 3. In: Phenol-Chloroform DNA Isolation Method. DNA and RNA Isolation Techniques for Non-Experts. Springer International Publishing; 2022. p. 33-9. Available from: http://dx.doi.org/10.1007/978-3-030-94230-4_3.

[270] Buhule OD, Minster RL, Hawley NL, Medvedovic M, Sun G, Viali S, et al. Stratified Randomization Controls Better for Batch Effects in 450k Methylation Analysis: a Cautionary Tale. Frontiers in Genetics. 2014;5(nil):nil. Available from: http://dx.doi.org/10.3389/fgene.2014.00354.

[271] Zindler T, Frieling H, Neyazi A, Bleich S, Friedel E. Simulating Combat: How Batch Correction Can Lead To the Systematic Introduction of False Positive Results in Dna Methylation Microarray Studies. BMC Bioinformatics. 2020;21(1):271. Available from: http://dx.doi.org/10.1186/s12859-020-03559-6.

[272] Bauer M. Cell-Type-Specific Disturbance of Dna Methylation Pattern: a Chance To Get More Benefit From and To Minimize Cohorts for Epigenome-Wide Association Studies. International Journal of Epidemiology. 2018;47(3):917-27. Available from: http://dx.doi.org/10.1093/ije/dyy029.

[273] Mansell G, Gorrie-Stone TJ, Bao Y, Kumari M, Schalkwyk LS, Mill J, et al. Guidance for Dna Methylation Studies: Statistical Insights From the Illumina Epic Array. BMC Genomics. 2019;20(1):366. Available from: http://dx.doi.org/10.1186/s12864-019-5761-7.

[274] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2013. Available from: http://www.R-project.org/.

[275] Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a Flexible and Comprehensive Bioconductor Package for the Analysis of Infinium Dna Methylation Microarrays. Bioinformatics. 2014;30(10):1363-9. Available from: http://dx.doi.org/10.1093/bioinformatics/btu049.

[276] Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical Evaluation of the Illumina Methylationepic Beadchip Microarray for Whole-Genome Dna Methylation Profiling. Genome Biology. 2016;17(1):208. Available from: http://dx.doi.org/10.1186/s13059-016-1066-1.

[277] an Chen Y, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of Cross-Reactive Probes and Polymorphic Cpgs in the Illumina Infinium Humanmethylation450 Microarray. Epigenetics. 2013;8(2):203-9. Available from: http://dx.doi.org/10.4161/epi.23470.

[278] Pidsley R, Wong CCY, Volta M, Lunnon K, Mill J, Schalkwyk LC. A Data-Driven Approach To Preprocessing Illumina 450k Methylation Array Data. BMC Genomics. 2013;14(1):293. Available from: http://dx.doi.org/10.1186/1471-2164-14-293.

[279] Smyth GK. Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. Statistical Applications in Genetics and Molecular Biology. 2004;3(1):1-25. Available from: http://dx.doi.org/10.2202/1544-6115.1027.

[280] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma Powers Differential Expression Analyses for Rna-Sequencing and Microarray Studies. Nucleic Acids Research. 2015;43(7):e47-7. Available from: http://dx.doi.org/10.1093/nar/gkv007.

[281] Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, Lord RV, et al. De Novo Identification of Differentially Methylated Regions in the Human Genome. Epigenetics & Chromatin. 2015;8(1):6. Available from: http://dx.doi.org/10.1186/1756-8935-8-6.

[282] Peters T. DMRcate. Bioconductor; 2017. Available from: https://bioconductor.org/packages/DMRcate.

[283] Chicco D, Agapito G. Nine Quick Tips for Pathway Enrichment Analysis. PLOS Computational Biology. 2022;18(8):e1010348. Available from: http://dx.doi.org/10.1371/journal.pcbi.1010348.

[284] Yu G, Wang LG, Han Y, He QY. Clusterprofiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS: A Journal of Integrative Biology. 2012;16(5):284-7. Available from: http://dx.doi.org/10.1089/omi.2011.0118.

[285] Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. Clusterprofiler 4.0: a Universal Enrichment Tool for Interpreting Omics Data. The Innovation. 2021;2(3):100141. Available from: http://dx.doi.org/10.1016/j.xinn.2021.100141.

[286] Phipson B. Missmethyl. nil. 2017;nil(nil):nil. Available from: https://bioconductor.org/packages/missMethyl.

[287] Phipson B, Maksimovic J, Oshlack A. Missmethyl: an R Package for Analyzing Data From Illumina's Humanmethylation450 Platform. Bioinformatics. 2015;32(2):286-8. Available from: http://dx.doi.org/10.1093/bioinformatics/btv560.

[288] Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: a Practical and Powerful Approach To Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57(1):289-300. Available from: http://dx.doi.org/10.1111/j.2517-6161.1995.tb02031.x.

[289] Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. Msfragger: Ultrafast and Comprehensive Peptide Identification in Mass Spectrometry-Based Proteomics. Nature Methods. 2017;14(5):513-20. Available from: http://dx.doi.org/10.1038/nmeth.4256.

[290] Zauber H, Kirchner M, Selbach M. Picky: a Simple Online Prm and Srm Method Designer for Targeted Proteomics. Nature Methods. 2018;15(3):156-7. Available from: http://dx.doi.org/10.1038/nmeth.4607.

[291] Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis; 2016. Available from: http://dx.doi.org/10.1101/060012.

[292] Pette D, Peuker H, Staron RS. The Impact of Biochemical Methods for Single Muscle Fibre Analysis. Acta Physiologica Scandinavica. 1999;166(4):261-77. Available from: http://dx.doi.org/10.1046/j.1365-201x.1999.00573.x.

[293] Moran S, Arribas C, Esteller M. Validation of a Dna Methylation Microarray for 850,000 Cpg Sites of the Human Genome Enriched in Enhancer Sequences. Epigenomics. 2016;8(3):389-99. Available from: http://dx.doi.org/10.2217/epi.15.114.

[294] Yang PC, Mahmood T. Western Blot: Technique, Theory, and Trouble Shooting. North American Journal of Medical Sciences. 2012;4(9):429. Available from: http://dx.doi.org/10.4103/1947-2714.100998.

[295] Biral D, Betto R, Danieli-Betto D, Salviati G. Myosin Heavy Chain Composition of Single Fibres From Normal Human Muscle. Biochemical Journal. 1988;250(1):307-8. Available from: http://dx.doi.org/10.1042/bj2500307.

[296] Oakley BR, Kirsch DR, Morris NR. A Simplified Ultrasensitive Silver Stain for Detecting Proteins in Polyacrylamide Gels. Analytical Biochemistry. 1980;105(1):361-3. Available from: http://dx.doi.org/10.1016/0003-2697(80)90470-4.

[297] Murgia M, Tan J, Geyer PE, Doll S, Mann M, Klopstock T. Single fiber proteomics of respiratory chain defects in mitochondrial disorders; 2018. Available from: http://dx.doi.org/10.1101/421750.

[298] Murgia M, Nagaraj N, Deshmukh AS, Zeiler M, Cancellara P, Moretti I, et al. Single Muscle Fiber Proteomics Reveals Unexpected Mitochondrial Specialization. EMBO reports. 2015;16(3):387-95. Available from: http://dx.doi.org/10.15252/embr.201439757.

[299] Noguera-Castells A, García-Prieto CA, Álvarez-Errico D, Esteller M. Validation of the New Epic Dna Methylation Microarray (900K Epic v2) for High-Throughput Profiling of the Human Dna Methylome. Epigenetics. 2023;18(1):nil. Available from: http://dx.doi.org/10.1080/15592294.2023.2185742.

[300] Teschendorff AE, Zheng SC. Cell-Type Deconvolution in Epigenome-Wide Association Studies: a Review and Recommendations. Epigenomics. 2017;9(5):757-68. Available from: http://dx.doi.org/10.2217/epi-2016-0153.

[301] Watkins SH, Ho K, Testa C, Falk L, Soule P, Nguyen LV, et al. The Impact of Low Input Dna on the Reliability of Dna Methylation As Measured By the Illumina Infinium Methylationepic Beadchip. Epigenetics. 2022;17(13):2366-76. Available from: http://dx.doi.org/10.1080/15592294.2022.2123898.

[302] Teschendorff AE, Zhuang J, Widschwendter M. Independent Surrogate Variable Analysis To Deconvolve Confounding Factors in Large-Scale Microarray Profiling Studies. Bioinformatics. 2011;27(11):1496-505. Available from: http://dx.doi.org/10.1093/bioinformatics/btr171.

[303] Dedeurwaerder S, Defrance M, Bizet M, Calonne E, Bontempi G, Fuks F. A Comprehensive Overview of Infinium Humanmethylation450 Data Processing. Briefings in Bioinformatics. 2013;15(6):929-41. Available from: http://dx.doi.org/10.1093/bib/bbt054.

[304] Min JL, Hemani G, Smith GD, Relton C, Suderman M. Meffil: Efficient Normalization and Analysis of Very Large Dna Methylation Datasets. Bioinformatics. 2018;34(23):3983-9. Available from: http://dx.doi.org/10.1093/bioinformatics/bty476.

[305] Xu Z, Niu L, Taylor JA. The Enmix Dna Methylation Analysis Pipeline for Illumina Beadchip and Comparisons With Seven Other Preprocessing Pipelines. Clinical Epigenetics. 2021;13(1):216. Available from: http://dx.doi.org/10.1186/s13148-021-01207-1.

[306] Xu Z, Taylor JA. Reliability of Dna Methylation Measures Using Illumina Methylation Beadchip. Epigenetics. 2020;16(5):495-502. Available from: http://dx.doi.org/10.1080/15592294.2020.1805692.

[307] Welsh H, Batalha CMPF, Li W, Mpye KL, Souza-Pinto NC, Naslavsky MS, et al. A Systematic Evaluation of Normalization Methods and Probe Replicability Using Infinium Epic Methylation Data. Clinical Epigenetics. 2023;15(1):41. Available from: http://dx.doi.org/10.1186/s13148-023-01459-z.

[308] Maksimovic J, Oshlack A, Phipson B. Gene Set Enrichment Analysis for Genome-Wide Dna Methylation Data. Genome Biology. 2021;22(1):173. Available from: http://dx.doi.org/10.1186/s13059-021-02388-x.

[309] Picard M, Hepple RT, Burelle Y. Mitochondrial Functional Specialization in Glycolytic and Oxidative Muscle Fibers: Tailoring the Organelle for Optimal Function. American Journal of Physiology-Cell Physiology. 2012;302(4):C629-41. Available from: http://dx.doi.org/10.1152/ajpcell.00368.2011.

[310] Rath S, Sharma R, Gupta R, Ast T, Chan C, Durham TJ, et al. Mitocarta3.0: an Updated Mitochondrial Proteome Now With Sub-Organelle Localization and Pathway Annotations. Nucleic Acids Research. 2020;49(D1):D1541-7. Available from: http://dx.doi.org/10.1093/nar/gkaa1011.

[311] Boyes J, Bird A. Repression of Genes By Dna Methylation Depends on Cpg Density and Promoter Strength: Evidence for Involvement of a Methyl-Cpg Binding Protein. The EMBO Journal. 1992;11(1):327-33. Available from: http://dx.doi.org/10.1002/j.1460-2075.1992.tb05055.x.

[312] Manzoni C, Kia DA, Vandrovcova J, Hardy J, Wood NW, Lewis PA, et al. Genome, Transcriptome and Proteome: the Rise of Omics Data and Their Integration in Biomedical Sciences. Briefings in Bioinformatics. 2016;19(2):286-302. Available from: http://dx.doi.org/10.1093/bib/bbw114.

[313] Rowlands DS, Page RA, Sukala WR, Giri M, Ghimbovschi SD, Hayat I, et al. Multi-Omic Integrated Networks Connect Dna Methylation and Mirna With Skeletal Muscle Plasticity To Chronic Exercise in Type 2 Diabetic Obesity. Physiological Genomics. 2014;46(20):747-65. Available from: http://dx.doi.org/10.1152/physiolgenomics.00024.2014.

[314] Brown CE. 13. In: Coefficient of Variation. Applied Multivariate Statistics in Geohydrology and Related Sciences. Springer Berlin Heidelberg; 1998. p. 155-7. Avail-

able from: http://dx.doi.org/10.1007/978-3-642-80328-4_13.

[315] Ivanov MV, Garibova LA, Postoenko VI, Levitsky LI, Gorshkov MV. On the Excessive Use of Coefficient of Variation As a Metric of Quantitation Quality in Proteomics. PROTEOMICS. 2023;24(1-2):nil. Available from: http://dx.doi.org/10.1002/pmic.202300090.

[316] Lowey S, Waller GS, Trybus KM. Function of Skeletal Muscle Myosin Heavy and Light Chain Isoforms By an in Vitro Motility Assay. Journal of Biological Chemistry. 1993;268(27):20414-8. Available from: http://dx.doi.org/10.1016/S0021-9258(20)80744-3.

[317] Smith JAB, Murach KA, Dyar KA, Zierath JR. Exercise Metabolism and Adaptation in Skeletal Muscle. Nature Reviews Molecular Cell Biology. 2023;24(9):607-32. Available from: http://dx.doi.org/10.1038/s41580-023-00606-x.

[318] MacInnis MJ, Gibala MJ. Physiological Adaptations To Interval Training and the Role of Exercise Intensity. The Journal of Physiology. 2016;595(9):2915-30. Available from: http://dx.doi.org/10.1113/JP273196.

[319] Ruegsegger GN, Pataky MW, Simha S, Robinson MM, Klaus KA, Nair KS. High-Intensity Aerobic, But Not Resistance Or Combined, Exercise Training Improves Both Cardiometabolic Health and Skeletal Muscle Mitochondrial Dynamics. Journal of Applied Physiology. 2023;135(4):763-74. Available from: http://dx.doi.org/10.1152/japplphysiol.00405.2023.

[320] Sale DG. 5 Influence of exercise and training on motor unit activation. Exercise and sport sciences reviews. 1987;15(1):95-152.

[321] Turner DC, Seaborne RA, Sharples AP. Comparative Transcriptome and Methylome Analysis in Human Skeletal Muscle Anabolism, Hypertrophy and Epigenetic Memory. Scientific Reports. 2019;9(1):4251. Available from: http://dx.doi.org/10.1038/s41598-019-40787-0.

[322] Gorski PP, Raastad T, Ullrich M, Turner DC, Hallén J, Savari SI, et al. Aerobic Exercise Training Resets the Human Skeletal Muscle Methylome 10 Years After Breast Cancer Treatment and Survival. The FASEB Journal. 2022;37(1):nil. Available from: http://dx.doi.org/10.1096/fj.202201510RR.

[323] Reisman E, Botella J, Huang C, Schittenhelm RB, Stroud DA, Granata C, et al.. Fibre-specific mitochondrial protein abundance is linked to resting and post-training mitochondrial content in human muscle; 2022. Available from: http://dx.doi.org/10.1101/2022.10.23.512956.

[324] Jacques M, Landen S, Romero JA, Yan X, Garnham A, Hiam D, et al. Individual Physiological and Mitochondrial Responses During 12 Weeks of Intensified Exercise. Physiological Reports. 2021;9(15):nil. Available from: http://dx.doi.org/10.14814/phy2.14962.

[325] Robinson MM, Lowe VJ, Nair KS. Increased Brain Glucose Uptake After 12 Weeks of Aerobic High-Intensity Interval Training in Young and Older Adults. The Journal of Clinical Endocrinology amp; Metabolism. 2017;103(1):221-7. Available from: http://dx.doi.org/10.1210/jc.2017-01571.

[326] Fitzsimons DP, Diffee GM, Herrick RE, Baldwin KM. Effects of Endurance Exercise on Isomyosin Patterns in Fast- and Slow-Twitch Skeletal Muscles. Journal of Applied Physiology. 1990;68(5):1950-5. Available from: http://dx.doi.org/10.1152/jappl.1990.68.5.1950.

[327] Sansal I, Dupont E, Toru D, Evrard C, Rouget P. Npdc-1, a Regulator of Neural Cell Proliferation and Differentiation, Interacts With E2f-1, Reduces Its Binding To Dna and Modulates Its Transcriptional Activity. Oncogene. 2000;19(43):5000-9. Available from: http://dx.doi.org/10.1038/sj.onc.1203843.

[328] Galiana E, Vernier P, Dupont E, Evrard C, Rouget P. Identification of a Neural-Specific Cdna, Npdc-1, Able To Down-Regulate Cell Proliferation and To Suppress Transformation. Proceedings of the National Academy of Sciences. 1995;92(5):1560-4. Available from: http://dx.doi.org/10.1073/pnas.92.5.1560.

[329] Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-Based Map of the Human Proteome. Science. 2015;347(6220):nil. Available from: http://dx.doi.org/10.1126/science.1260419.

[330] Karlsson M, Zhang C, Méar L, Zhong W, Digre A, Katona B, et al. A Single-Cell Type Transcriptomics Map of Human Tissues. Science Advances. 2021;7(31):nil. Available from: http://dx.doi.org/10.1126/sciadv.abh2169.

[331] Oishi Y, Manabe I, Tobe K, Ohsugi M, Kubota T, Fujiu K, et al. Sumoylation of Krüppel-Like Transcription Factor 5 Acts As a Molecular Switch in Transcriptional Programs of Lipid Metabolism Involving PPAR-δ. Nature Medicine. 2008;14(6):656-66. Available from: http://dx.doi.org/10.1038/nm1756.

[332] Liu W, Wang Y, Bozi LHM, Fischer PD, Jedrychowski MP, Xiao H, et al. Lactate Regulates Cell Cycle By Remodelling the Anaphase Promoting Complex. Nature. 2023;616(7958):790-7. Available from: http://dx.doi.org/10.1038/s41586-023-05939-3.

[333] Deshmukh AS, Murgia M, Nagaraj N, Treebak JT, Cox J, Mann M. Deep Proteomics of Mouse Skeletal Muscle Enables Quantitation of Protein Isoforms, Metabolic Pathways, and Transcription Factors*. Molecular amp; Cellular Proteomics. 2015;14(4):841-53. Available from: http://dx.doi.org/10.1074/mcp.M114.044222.

[334] Zheng SC, Beck S, Jaffe AE, Koestler DC, Hansen KD, Houseman AE, et al. Correcting for Cell-Type Heterogeneity in Epigenome-Wide Association Studies: Revisiting Previous Analyses. Nature Methods. 2017;14(3):216-7. Available from: http://dx.doi.org/10.1038/nmeth.4187.

[335] Dickinson JM, Lee JD, Sullivan BE, Harber MP, Trappe SW, Trappe TA. A New Method To Study in Vivo Protein Synthesis in Slow- and Fast-Twitch Muscle Fibers and Initial Measurements in Humans. Journal of Applied Physiology. 2010;108(5):1410-6. Available from: http://dx.doi.org/10.1152/japplphysiol.00905.2009.

[336] Horwath O, Edman S, Andersson A, Larsen FJ, Apró W. Thrifty, a Novel High-throughput Method for Rapid Fibre Type Identification of Isolated Skeletal Muscle Fibres. The Journal of Physiology. 2022;nil(nil):nil. Available from: http://dx.doi.org/10.1113/JP282959.

[337] Tsai WLE, Schedl ME, Maley JM, McCormack JE. More Than Skin and Bones: Comparing Extraction Methods and Alternative Sources of Dna From Avian Museum Specimens. Molecular Ecology Resources. 2019;20(5):1220-7. Available from: http://dx.doi.org/10.1111/1755-0998.13077.

[338] Hykin SM, Bi K, McGuire JA. Fixing Formalin: a Method To Recover Genomic-Scale Dna Sequence Data From Formalin-Fixed Museum Specimens Using High-Throughput Sequencing. PLOS ONE. 2015;10(10):e0141579. Available from: http://dx.doi.org/10.1371/journal.pone.0141579.

[339] Jaffe AE, Irizarry RA. Accounting for Cellular Heterogeneity Is Critical in Epigenome-Wide Association Studies. Genome Biology. 2014;15(2):R31. Available from: http://dx.doi.org/10.1186/gb-2014-15-2-r31.

[340] Ivanova A, Gill-Hille M, Huang S, Branca RM, Kmiec B, Teixeira PF, et al. A Mitochondrial Lyr Protein Is Required for Complex I Assembly. Plant Physiology. 2019;181(4):1632-50. Available from: http://dx.doi.org/10.1104/pp.19.00822.

[341] Boniecki MT, Freibert SA, Mühlenhoff U, Lill R, Cygler M. Structure and Functional Dynamics of the Mitochondrial Fe/s Cluster Synthesis Complex. Nature Communications. 2017;8(1):1287. Available from: http://dx.doi.org/10.1038/s41467-017-01497-1.

[342] Lim SC, Friemel M, Marum JE, Tucker EJ, Bruno DL, Riley LG, et al. Mutations in Lyrm4, Encoding Iron-Sulfur Cluster Biogenesis Factor Isd11, Cause Deficiency of Multiple Respiratory Chain Complexes. Human Molecular Genetics. 2013;22(22):4460-73. Available from: http://dx.doi.org/10.1093/hmg/ddt295.

[343] Angerer H. The Superfamily of Mitochondrial Complex1-lyr Motif-Containing (LYRM) Proteins. Biochemical Society Transactions. 2013;41(5):1335-41. Available from: http://dx.doi.org/10.1042/BST20130116.

[344] Kou C, Cao X, Qin D, Ji C, Zhu J, Zhang C, et al. Over-Expression of Lyrm1 Inhibits Glucose Transport in Rat Skeletal Muscles Via Attenuated Phosphorylation of Pi3k (p85) and Akt. Molecular and Cellular Biochemistry. 2010;348(1-2):149-54. Available from: http://dx.doi.org/10.1007/s11010-010-0649-5.

[345] Alfattal R, Alfarhan M, Algaith AM, Albash B, Elshafie RM, Alshammari A, et al. Lyrm7-associated Mitochondrial Complex iii Deficiency With Non-cavitating Leukoencephalopathy and Stroke-like Episodes. American Journal of Medical Genetics Part A. 2023;191(5):1401-11. Available from: http://dx.doi.org/10.1002/ajmg.a.63143.

[346] Babenko VN, Chadaeva IV, Orlov YL. Genomic Landscape of Cpg Rich Elements in Human. BMC Evolutionary Biology. 2017;17(S1):19. Available from: http://dx.doi.org/10.1186/s12862-016-0864-0.

[347] Zhou L, Ng HK, Drautz-Moses DI, Schuster SC, Beck S, Kim C, et al. Systematic Evaluation of Library Preparation Methods and Sequencing Platforms for High-Throughput Whole Genome Bisulfite Sequencing. Scientific Reports. 2019;9(1):10383. Available from: http://dx.doi.org/10.1038/s41598-019-46875-5.

[348] Robinson MD, Kahraman A, Law CW, Lindsay H, Nowicka M, Weber LM, et al. Statistical Methods for Detecting Differentially Methylated Loci and Regions. Frontiers in Genetics. 2014;5(nil):nil. Available from: http://dx.doi.org/10.3389/fgene.2014.00324.

[349] Zhang W, Qu J, Liu GH, Belmonte JCI. The Ageing Epigenome and Its Rejuvenation. Nature Reviews Molecular Cell Biology. 2020;21(3):137-50. Available from: http://dx.doi.org/10.1038/s41580-019-0204-5.

[350] López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. Hallmarks of Aging: an Expanding Universe. Cell. 2023;186(2):243-78. Available from: http://dx.doi.org/10.1016/j.cell.2022.11.001.

[351] López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The Hallmarks of Aging. Cell. 2013;153(6):1194-217. Available from: http://dx.doi.org/10.1016/j.cell.2013.05.039.

[352] Horvath S, Raj K. Dna Methylation-Based Biomarkers and the Epigenetic Clock Theory of Ageing. Nature Reviews Genetics. 2018;19(6):371-84. Available from: http://dx.doi.org/10.1038/s41576-018-0004-3.

[353] McLeod M, Breen L, Hamilton DL, Philp A. Live Strong and Prosper: the Importance of Skeletal Muscle Strength for Healthy Ageing. Biogerontology. 2016;17(3):497-510. Available from: http://dx.doi.org/10.1007/s10522-015-9631-7.

[354] Nilwik R, Snijders T, Leenders M, Groen BBL, van Kranenburg J, Verdijk LB, et al. The Decline in Skeletal Muscle Mass With Aging Is Mainly Attributed To a Re-

duction in Type Ii Muscle Fiber Size. Experimental Gerontology. 2013;48(5):492-8. Available from: http://dx.doi.org/10.1016/j.exger.2013.02.012.

[355] Seale K, Horvath S, Teschendorff A, Eynon N, Voisin S. Making Sense of the Ageing Methylome. Nature Reviews Genetics. 2022;23(10):585-605. Available from: http://dx.doi.org/10.1038/s41576-022-00477-6.

[356] Rosa-Caldwell ME, Greene NP. Muscle Metabolism and Atrophy: Let's Talk About Sex. Biology of Sex Differences. 2019;10(1):43. Available from: http://dx.doi.org/10.1186/s13293-019-0257-3.

[357] Greenberg MVC, Bourc'his D. The Diverse Roles of Dna Methylation in Mammalian Development and Disease. Nature Reviews Molecular Cell Biology. 2019;20(10):590-607. Available from: http://dx.doi.org/10.1038/s41580-019-0159-6.

[358] Pette D. Metabolic Heterogeneity of Muscle Fibres. Journal of Experimental Biology. 1985;115(1):179-89. Available from: http://dx.doi.org/10.1242/jeb.115.1.179.

[359] Stuart CA, Stone WL, Howell MEA, Brannon MF, Hall HK, Gibson AL, et al. Myosin Content of Individual Human Muscle Fibers Isolated By Laser Capture Microdissection. American Journal of Physiology-Cell Physiology. 2016;310(5):C381-9. Available from: http://dx.doi.org/10.1152/ajpcell.00317.2015.

[360] Shrager JB, Desjardins PR, Burkman JM, Konig SK, Stewart DR, Su L, et al. []. Journal of Muscle Research and Cell Motility. 2000;21(4):345-55. Available from: http://dx.doi.org/10.1023/A:1005635030494.

[361] Santos MD, Backer S, Auradé F, Wong MMK, Wurmser M, Pierre R, et al. A Fast Myosin Super Enhancer Dictates Muscle Fiber Phenotype Through Competitive Interactions With Myosin Genes. Nature Communications. 2022;13(1):1039. Available from: http://dx.doi.org/10.1038/s41467-022-28666-1.

[362] Momenzadeh A, Jiang Y, Kreimer S, Teigen LE, Zepeda CS, Haghani A, et al. A Complete Workflow for High Throughput Human Single Skeletal Muscle Fiber Proteomics. Journal of the American Society for Mass Spectrometry. 2023;34(9):1858-67. Available from: http://dx.doi.org/10.1021/jasms.3c00072.

[363] Santos MD, Backer S, Saintpierre B, Izac B, Andrieu M, Letourneur F, et al. Single-Nucleus Rna-Seq and Fish Identify Coordinated Transcriptional Activity in Mammalian Myofibers. Nature Communications. 2020;11(1):5102. Available from: http://dx.doi.org/10.1038/s41467-020-18789-8.

[364] Gnocchi VF, White RB, Ono Y, Ellis JA, Zammit PS. Further Characterisation of the Molecular Signature of Quiescent and Activated Mouse Muscle Satellite Cells. PLoS ONE. 2009;4(4):e5205. Available from: http://dx.doi.org/10.1371/journal.pone.0005205.

[365] Lähnemann D, Köster J, Szczurek E, McCarthy DJ, Hicks SC, Robinson MD, et al. Eleven Grand Challenges in Single-Cell Data Science. Genome Biology. 2020;21(1):31. Available from: http://dx.doi.org/10.1186/s13059-020-1926-6.

[366] Li Y, Ma L, Wu D, Chen G. Advances in Bulk and Single-Cell Multi-Omics Approaches for Systems Biology and Precision Medicine. Briefings in Bioinformatics. 2021;nil(nil):nil. Available from: http://dx.doi.org/10.1093/bib/bbab024.

[367] Michels KB, Binder AM, Dedeurwaerder S, Epstein CB, Greally JM, Gut I, et al. Recommendations for the Design and Analysis of Epigenome-Wide Association Studies. Nature Methods. 2013;10(10):949-55. Available from: http://dx.doi.org/10.1038/nmeth.2632.

[368] Widmann M, Nieß AM, Munz B. Physical Exercise and Epigenetic Modifications in Skeletal Muscle. Sports Medicine. 2019;49(4):509-23. Available from: http://dx.doi.org/10.1007/s40279-019-01070-4.

[369] Sharples AP, Turner DC. Skeletal Muscle Memory. American Journal of Physiology-Cell Physiology. 2023;324(6):C1274-94. Available from: http://dx.doi.org/10.1152/ajpcell.00099.2023.

[370] Williams CJ, Li Z, Harvey N, Lea RA, Gurd BJ, Bonafiglia JT, et al. Genome Wide Association Study of Response To Interval and Continuous Exercise Training: the Predict-Hiit Study. Journal of Biomedical Science. 2021;28(1):37. Available from: http://dx.doi.org/10.1186/s12929-021-00733-7.

[371] Wen D, Utesch T, Wu J, Robertson S, Liu J, Hu G, et al. Effects of Different Protocols of High Intensity Interval Training for Vo2max Improvements in Adults: a

Meta-Analysis of Randomised Controlled Trials. Journal of Science and Medicine in Sport. 2019;22(8):941-7. Available from: http://dx.doi.org/10.1016/j.jsams.2019.01.013.

[372] Voisin S, Seale K, Jacques M, Landen S, Harvey NR, Haupt LM, et al. Exercise Is Associated With Younger Methylome and Transcriptome Profiles in Human Skeletal Muscle. Aging Cell. 2023;23(1):nil. Available from: http://dx.doi.org/10.1111/acel.13859.

[373] Fiorito G, Caini S, Palli D, Bendinelli B, Saieva C, Ermini I, et al. Dna Methylation-based Biomarkers of Aging Were Slowed Down in a Two-year Diet and Physical Activity Intervention Trial: the Dama Study. Aging Cell. 2021;20(10):nil. Available from: http://dx.doi.org/10.1111/acel.13439.

[374] Jokai M, Torma F, McGreevy KM, Koltai E, Bori Z, Babszki G, et al. Dna Methylation Clock Dnamfitage Shows Regular Exercise Is Associated With Slower Aging and Systemic Adaptation. GeroScience. 2023;45(5):2805-17. Available from: http://dx.doi.org/10.1007/s11357-023-00826-1.

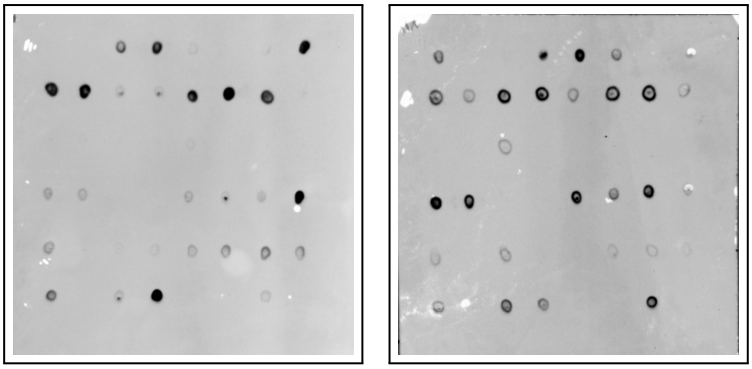
[375] Steenberg DE, Jørgensen NB, Birk JB, Sjøberg KA, Kiens B, Richter EA, et al. Exercise Training Reduces the Insulin-sensitizing Effect of a Single Bout of Exercise in Human Skeletal Muscle. The Journal of Physiology. 2018;597(1):89-103. Available from: http://dx.doi.org/10.1113/JP276735.

[376] Neubauer O, Sabapathy S, Ashton KJ, Desbrow B, Peake JM, Lazarus R, et al. Time Course-Dependent Changes in the Transcriptome of Human Skeletal Muscle During Recovery From Endurance Exercise: From Inflammation To Adaptive Remodeling. Journal of Applied Physiology. 2014;116(3):274-87. Available from: http://dx.doi.org/10.1152/japplphysiol.00909.2013.

[377] Mahoney DJ, Parise G, Melov S, Safdar A, Tarnopolsky MA. Analysis of Global Mrna Expression in Human Skeletal Muscle During Recovery From Endurance Exercise. The FASEB Journal. 2005;19(11):1498-500. Available from: http://dx.doi.org/10.1096/fj.04-3149fje.

9 Appendices

A Results 1



(A) MYH7: Manual Dot Blot **(B)** MYH2: Manual Dot Blot

Figure A.1: Manual Dot Blot

One tenth of a muscle fibre fragment was manually spotted onto a PVDF membrane. (A) shows antibody staining for MYH7 that represents Type I fibres. (B) shows antibody staining for MYH2 that represents Type II fibres.

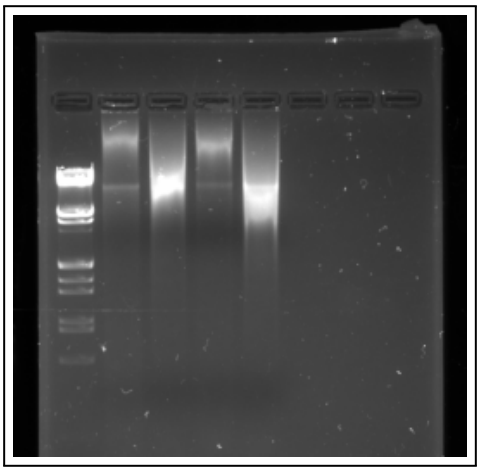


Figure A.2: DNA-extraction

DNA seperated on a 1.5% agarose gel. Lane 1 and 3 are following the protocol from Qiagen and lanes 2 and 4 are following the modified protocol reported in Begue et. al.

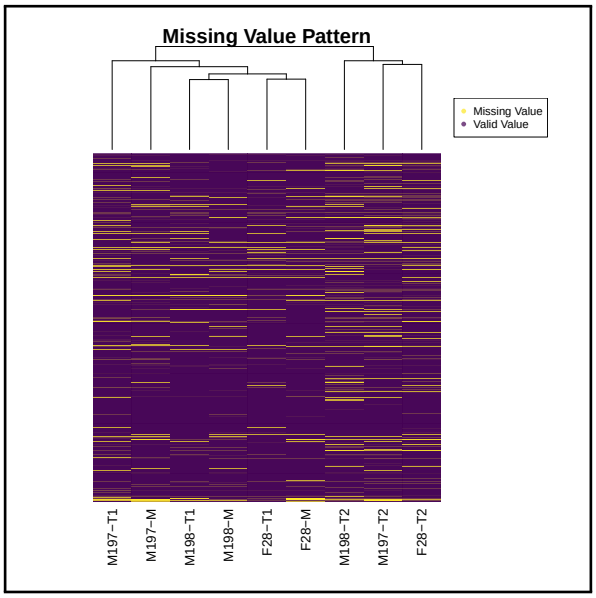


Figure A.3: Missing Value Heatmap

Missing value heatmap of proteins. Yellow bars indicate missing values and purple bars indicate valid values. Samples cluster according to fibre type.

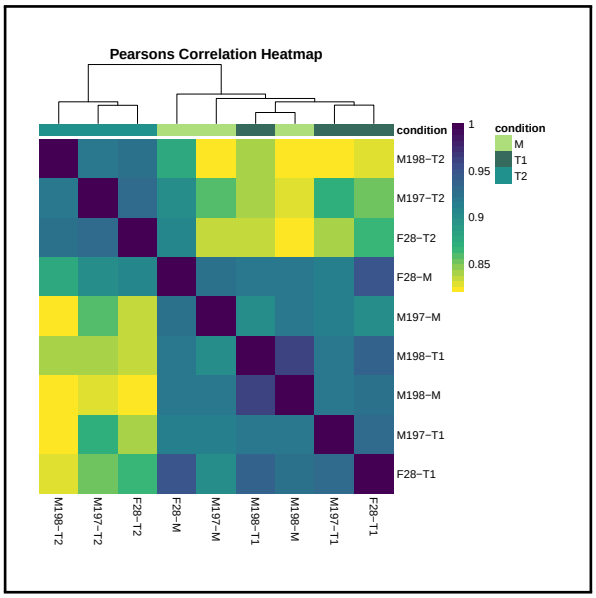


Figure A.4: Pearson's Correlation Heatmap

Sample correlation matrix of Type I muscle fibres (TI), Type II muscle fibres (TII) and mixed muscle fibres (M). The colour scale indicates the spearman correlation between samples based on protein values.

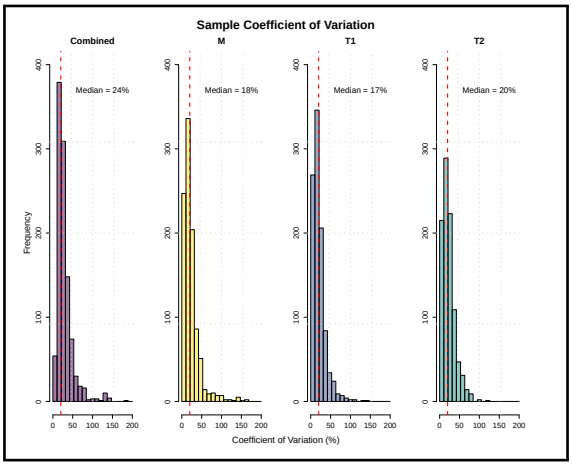


Figure A.5: Coefficient of Variation

Coefficient of variation plots for all samples combined, mixed fibres (M), Type I fibres (T1) and TII fibres (T2).

B RESULTS 2

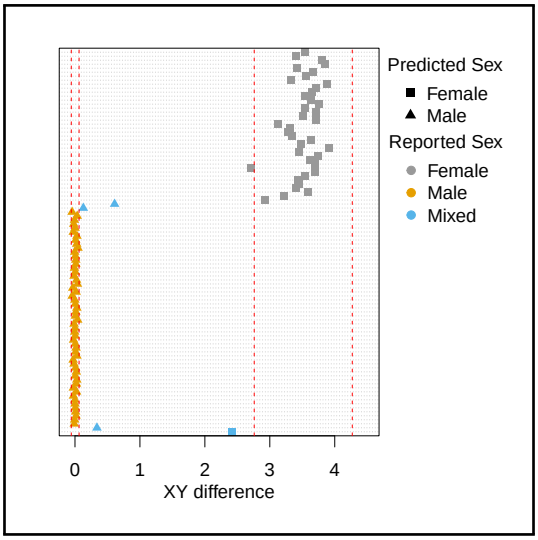
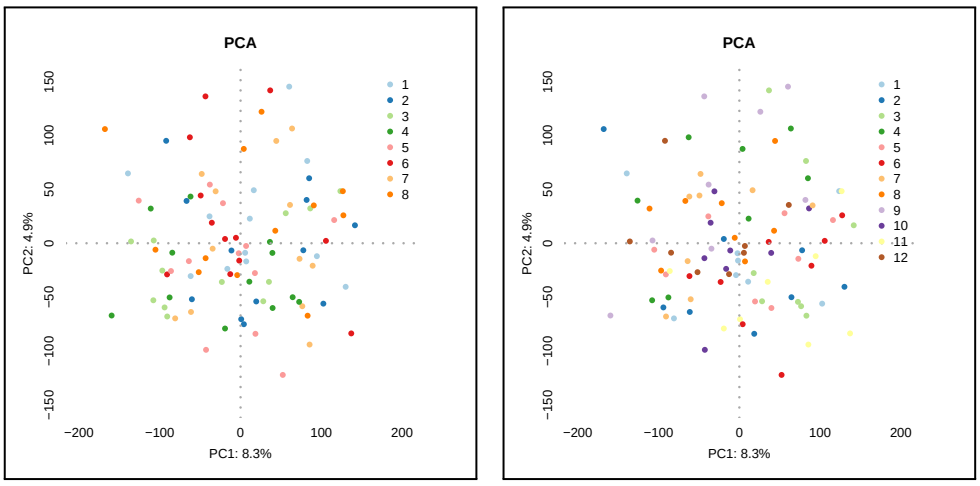


Figure B.1: Sample Sex Prediction

The sex of samples were predicted using the mean difference in X verses Y probes. No sex mismatches were identified. Squares and triangles represent predicted female and male samples respectively. Grey, orange and blue colours represent reported female, male or mixed samples.

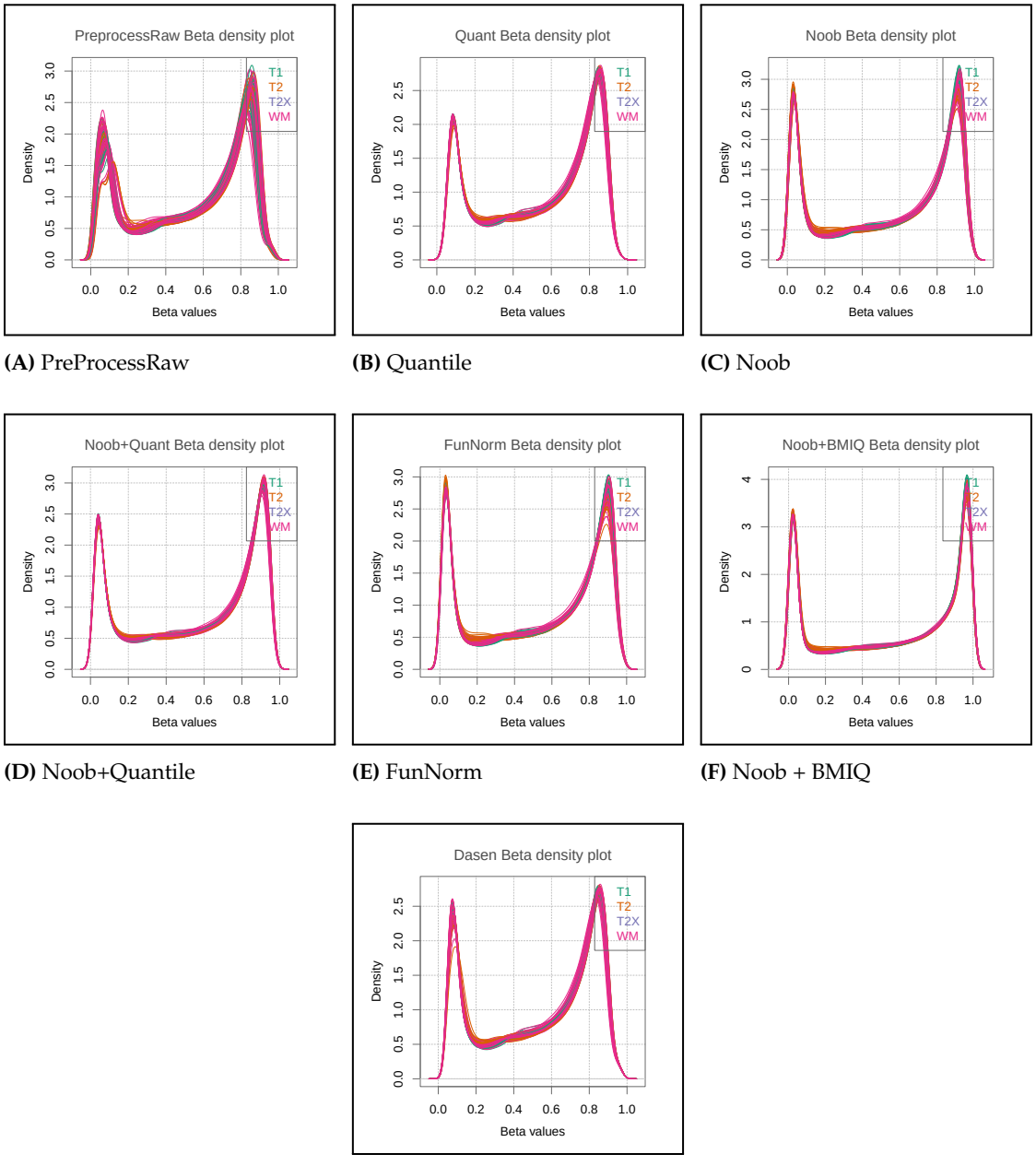


(A) PCA coloured by slide

(B) PCA coloured by position

Figure B.2: Batch Effects PCA

Pricinpal component analysis was performed on filtered M values. PCA plots coloured by slide (A) and position (B).



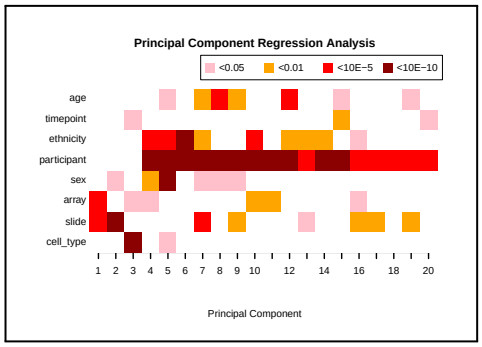
(A) PreProcessRaw **(B)** Quantile **(C)** Noob

(D) Noob+Quantile **(E)** FunNorm **(F)** Noob + BMIQ

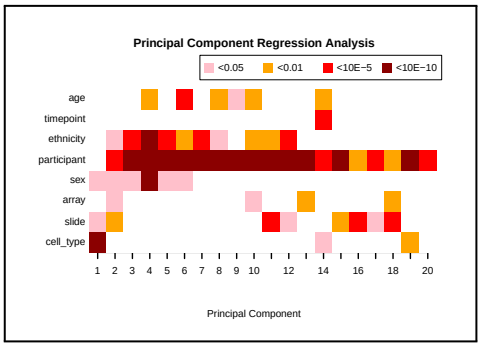
(G) Dasen

Figure B.3: DensityPlots

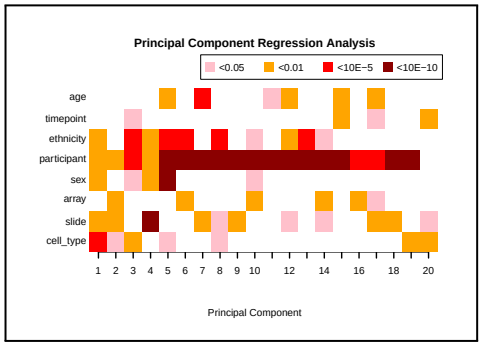
Density plots for each of the seven normalisation methods were produced to assess the distribution of beta values after normalisation. (A) preprocessRaw produced a density plot that had higher variation across all beta values. (B)-(G) show improved density plots compared to (A). All plots show the typical bimodal distribution of beta values. T1 = TI fibres , T2 = T2 fibres, T2X = Type IIx fibres, WM = whole muscle.



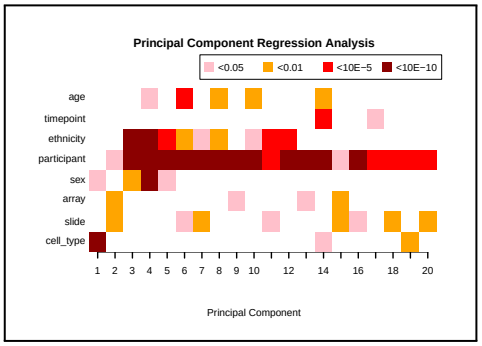
(A) PreProcessRaw



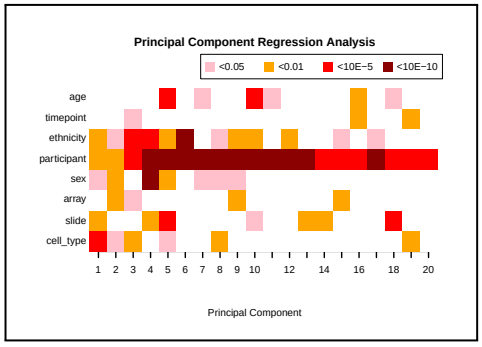
(B) Quantile



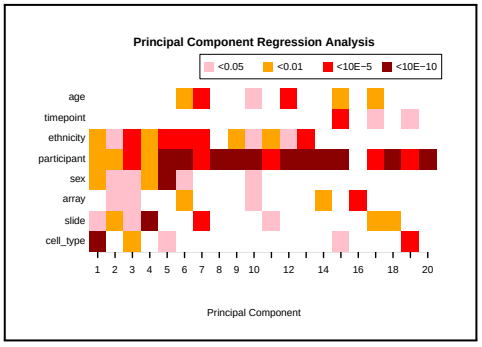
(C) Noob



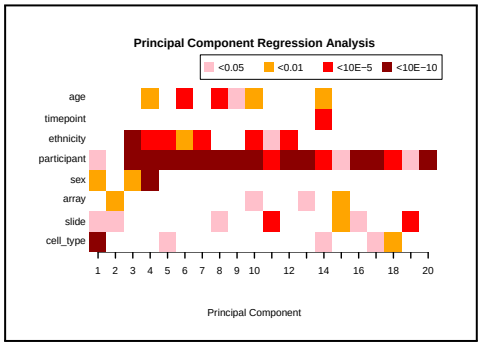
(D) Noob+Quantile



(E) FunNorm



(F) Noob + BMIQ



(G) Dasen

Figure B.4: PCR Regression Plots

PC regression on filtered M-values. Regression analysis shows the associated P values between the top 20 PCs and covariates of interest. The darkest colour is the smallest P value (p < 10E-10).

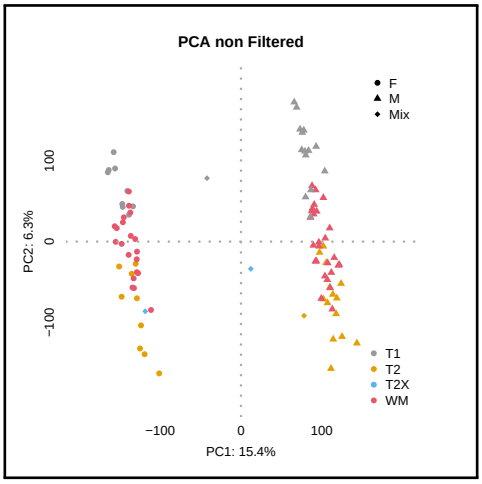


Figure B.5: PCA non filtered

Pricinpal component analysis was performed on non filtered M values. When sex chromosomes are not removed sex is the major source of variation on the first PC.

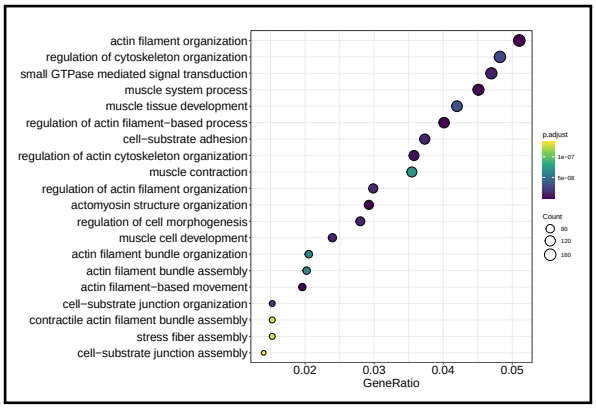


Figure B.6: Gene ontology ORA of Biological Processes

Gene ontology overrepresentation analysis of the differentially methylated DMPs between TI and TII fibres. The circle size represents gene numbers, and the color represents the P-adjusted value.

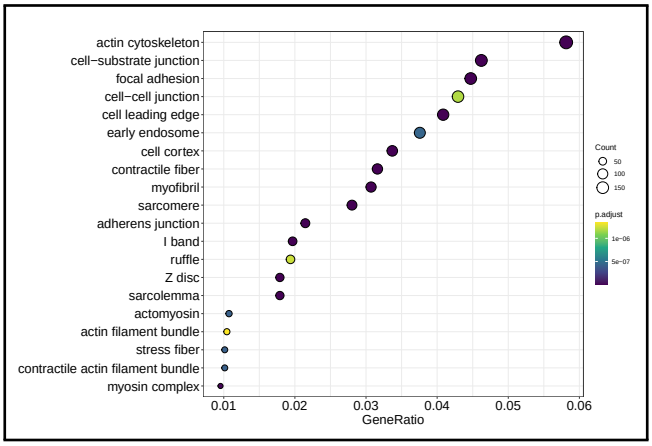


Figure B.7: Gene ontology ORA of Cellular Component

Gene ontology overrepresentation analysis of the differentially methylated DMPs between TI and TII fibres. The circle size represents gene numbers, and the color represents the P-adjusted value.

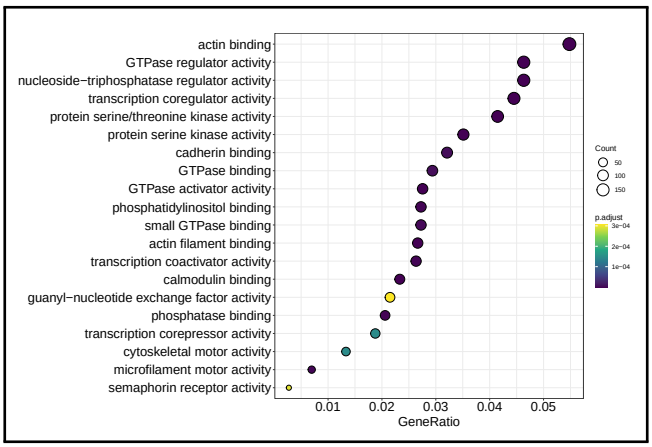


Figure B.8: Gene ontology ORA of Molecular Function

Molecular function Gene ontology overrepresentation analysis of the differentially methylated DMPs between TI and TII fibres. The circle size represents gene numbers, and the color represents the P-adjusted value.

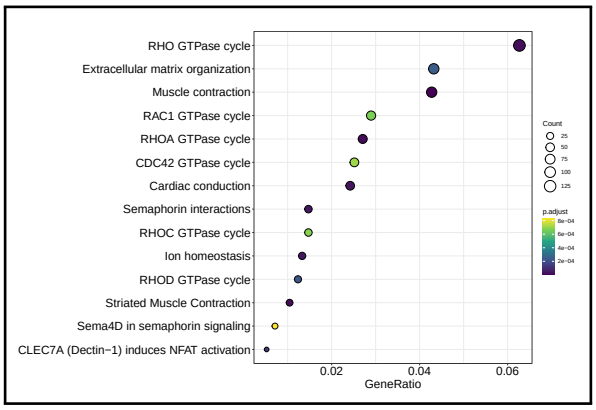


Figure B.9: Reactome Dotplot of DMPs in TI vs TII fibres

Reactome overrepresentation analysis of the differentially methylated DMPs between TI and TII fibres. The circle size represents gene numbers, and the color represents the P-adjusted value.

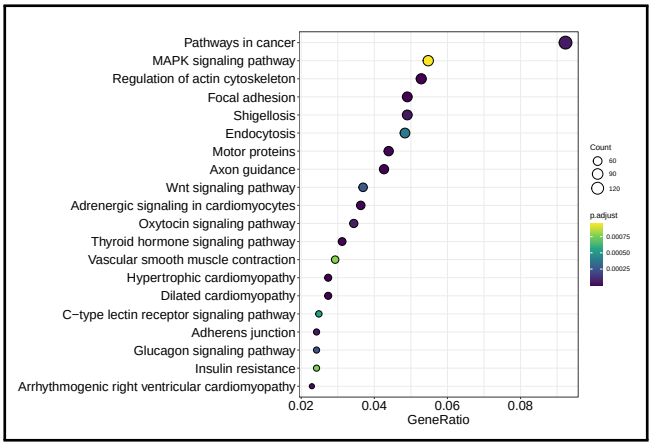


Figure B.10: Kegg Dotplot of DMPs in TI vs TII fibres

KEGG overrepresentation analysis of the differentially methylated DMPs between TI and TII fibres. The circle size represents gene numbers, and the color represents the P-adjusted value.

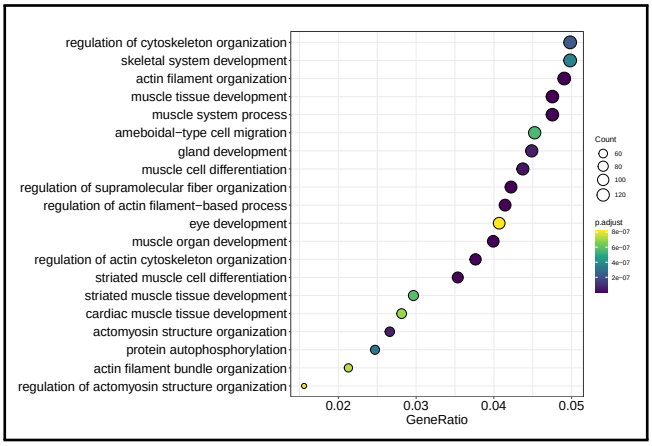


Figure B.11: Gene ontology ORA of Biological Processes

Biological process gene ontology overrepresentation analysis of the differentially methylated DMRs between TI and TII fibres. The circle size represents gene numbers, and the color represents the P-adjusted value.

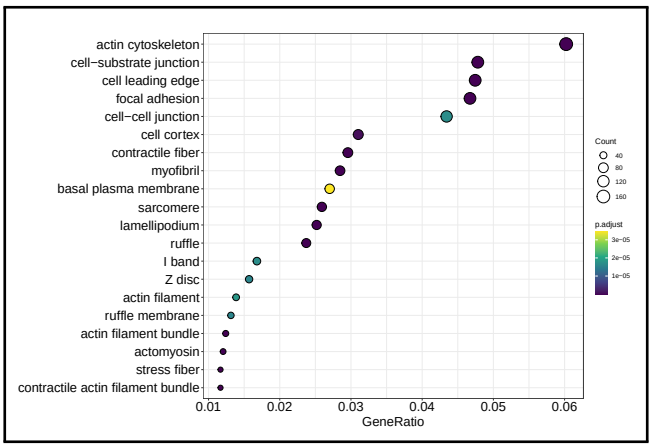


Figure B.12: Gene ontology ORA of Cellular Component

Cellular component gene ontology overrepresentation analysis of the differentially methylated DMRs between TI and TII fibres. The circle size represents gene numbers, and the color represents the P-adjusted value.

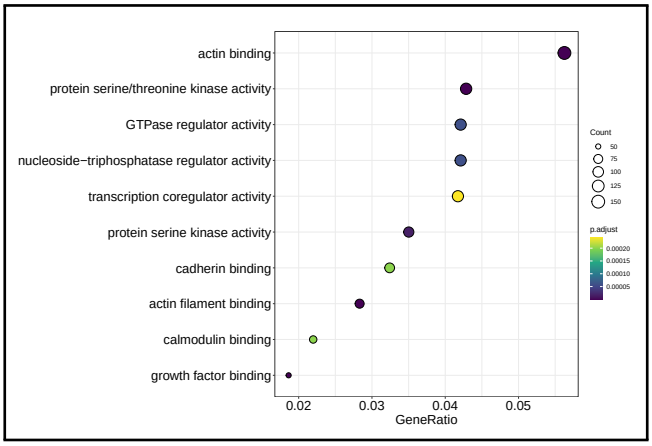


Figure B.13: Gene ontology ORA of Molecular Function

Molecular function gene ontology overrepresentation analysis of the differentially methylated DMRs between TI and TII fibres. The circle size represents gene numbers, and the color represents the P-adjusted value.

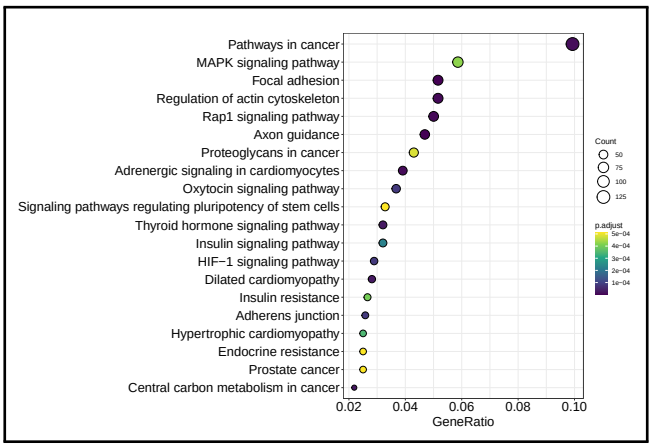


Figure B.14: Reactome of DMRs in TI vs TII fibres

KEGG overrepresentation analysis of the differentially methylated DMRs between TI and TII fibres. The circle size represents gene numbers, and the color represents the P-adjusted value.

Table B.1: Correlations between s200 and s201

	s200_12wk	s200_4w	s200_pre	s201_12wk	s201_4w	s201_pre
s200_12wk	1.000	0.990	0.988	0.993	0.993	0.994
s200_4w	0.990	1.000	0.993	0.989	0.990	0.991
s200_pre	0.988	0.993	1.000	0.986	0.988	0.990
s201_12wk	0.993	0.989	0.986	1.000	0.992	0.992
s201_4w	0.993	0.990	0.988	0.992	1.000	0.993
s201_pre	0.994	0.991	0.990	0.992	0.993	1.000

C Results 3

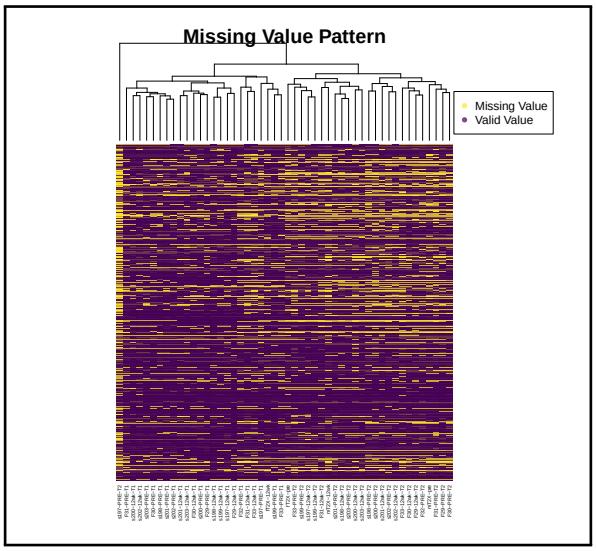
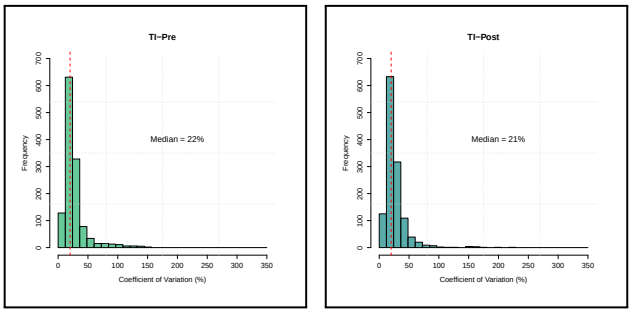


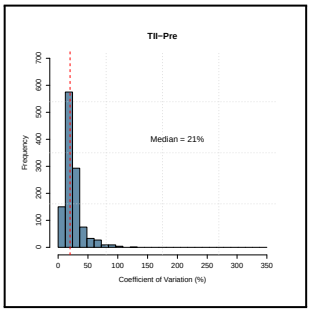
Figure C.1: Missing Value Heatmap Fibre-type Proteomics

Missing value heatmap of fibre-type proteomic values. Yellow bars indicate missing values and purple bars indicate valid values. Samples cluster according to fibre type.

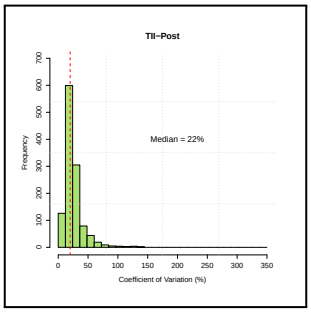


(A) Coefficient of Variation: TI-Pre

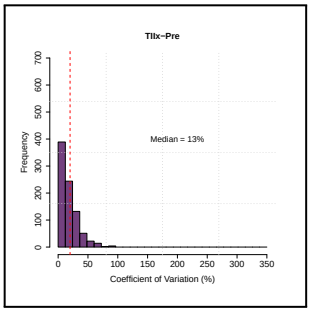
(B) Coefficient of Variation: TI-Post



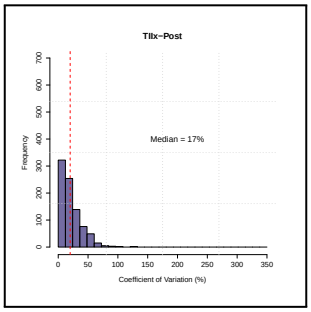
(C) Coefficient of Variation: TII-Pre



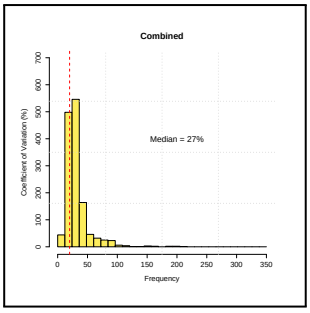
(D) Coefficient of Variation: TII-Post



(E) Coefficient of Variation: TIIX-Pre



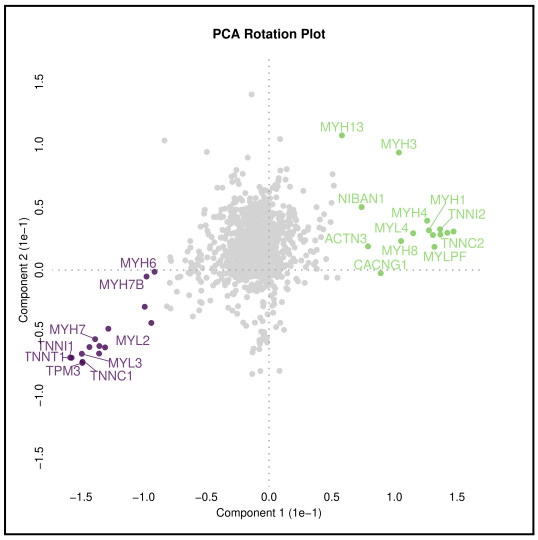
(F) Coefficient of Variation: TIIX-Post



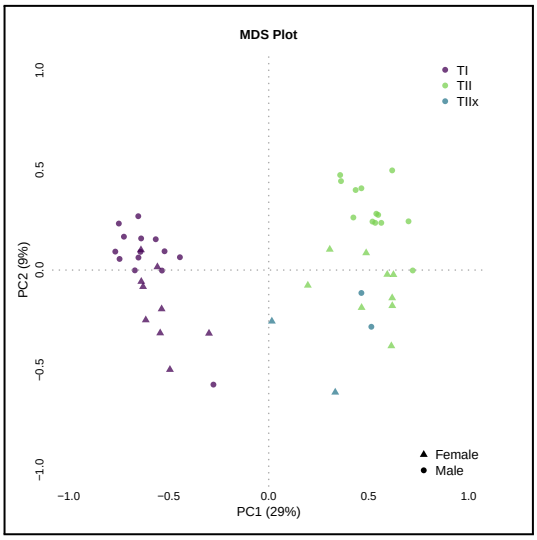
(G) Coefficient of Variation: Combined

Figure C.2: Coefficient of Variation

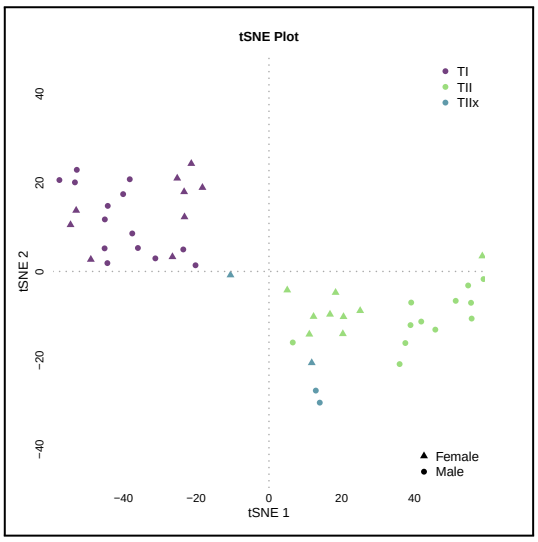
Coefficient of variation of fibre-type specific proteomics samples. (A) Type I pre samples, (B) Type I post samples, (C) Type II pre samples, (D) Type II post samples, (E) Type IIx pre samples, (F) Type IIx post samples and (G) all samples combined.



(A) Rotation Plot



(B) MDS plot



(C) tSNE Plot

Figure C.3: Dimension Reduction Analysis

Figure D.3 is a dimension reduction analysis of the major sources of variation within the dataset. (A) is a PCA rotation plot showing the proteins driving the first PC. (B) is an MDS plot showing clustering along the first component according to fibre type. (C) shows a tSNE plot showing a similar clustering of fibre type samples across tSNE1.

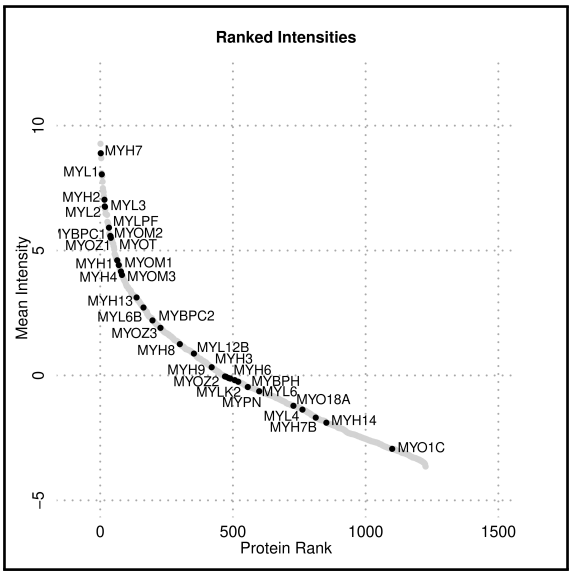
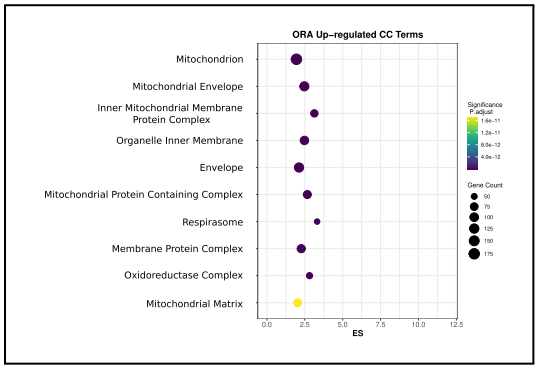
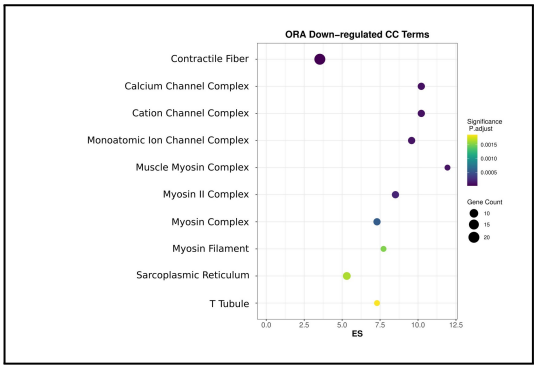


Figure C.4: Ranked Intensities of Proteins

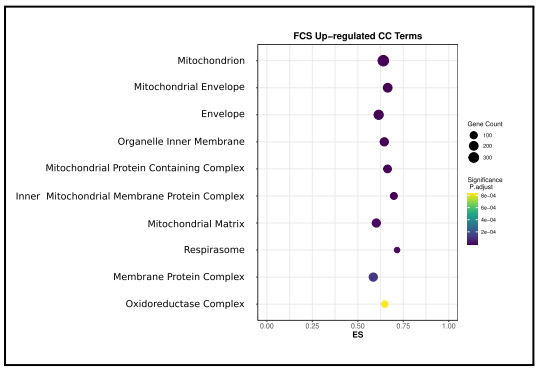
Shows the ranked intensities of myosin proteins identified in all samples. The most abundant proteins in the samples belong to the major contractile proteins such as MYH7 and MYH2.



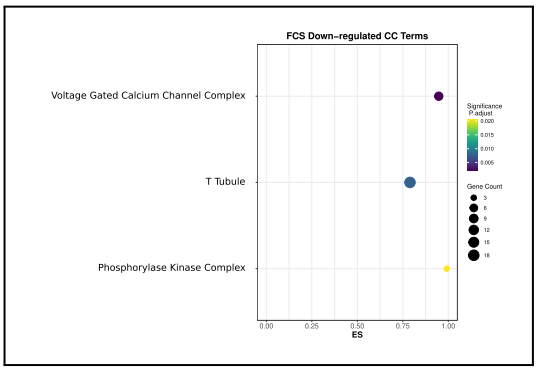
(A) ORA of up-regulated GO-CC terms



(B) ORA of down-regulated GO-CC terms



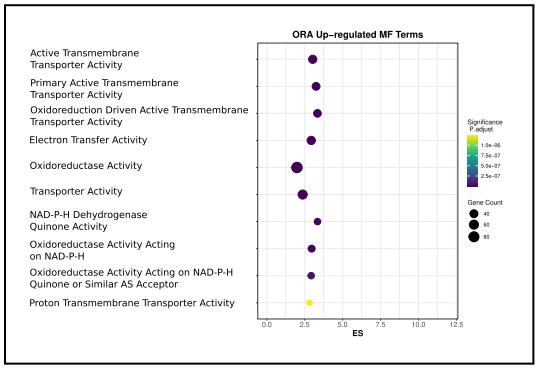
(C) FCS of up-regulated GO-CC terms



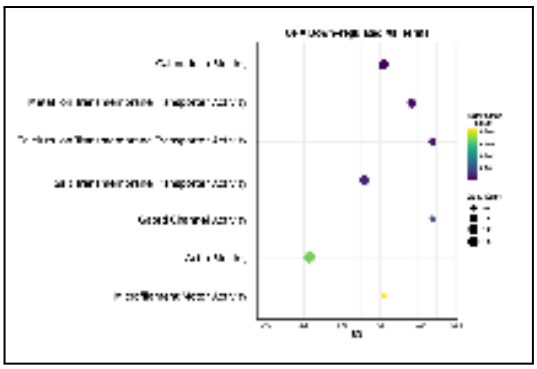
(D) FCS of down-regulated GO-CC terms

Figure C.5: CC enrichment analysis

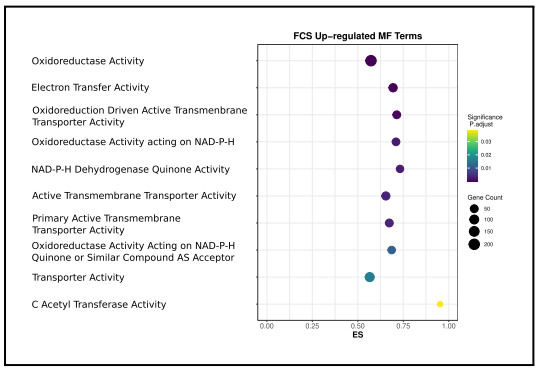
Cellular component gene ontology gene set enrichment analysis of the DEPs between TI and TII fibres. (A) & (B) are the up regulated and down-regulated biological process GO terms identified using ORA. (C) & (D) are the up-regulated and down-regulated biological processes GO terms identified using FCS. The circle size represents gene numbers, and the color represents the adj-P-value and the enrichment scores are on the x axis.



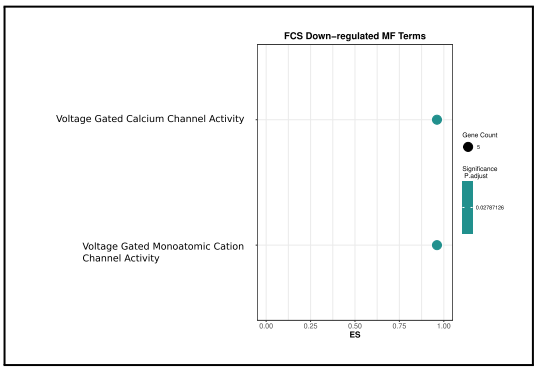
(A) ORA of upregulated GO-MF terms



(B) ORA of downregulated GO-MF terms



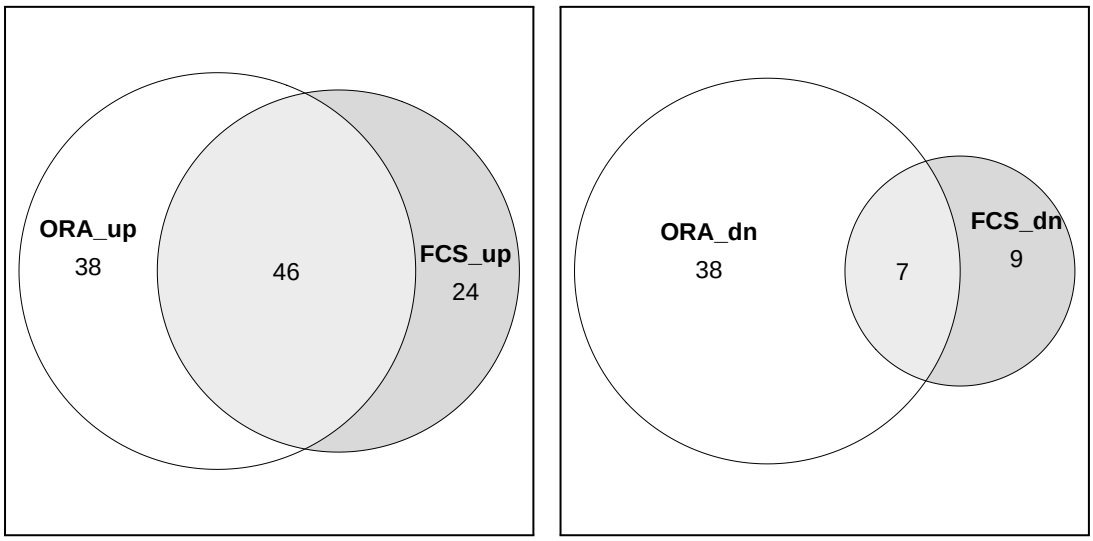
(C) FCS of up-regulated GO-MF terms



(D) FCS of down-regulated GO-MF terms

Figure C.6: MF enrichment analysis

Molecular Function gene ontology gene set enrichment analysis of the DEPs between TI and TII fibres. (A) & (B) are the up regulated and down-regulated biological process GO terms identified using ORA. (C) & (D) are the up-regulated and down-regulated biological processes GO terms identified using FCS. The circle size represents gene numbers, and the color represents the adj-P-value and the enrichment scores are on the x axis.

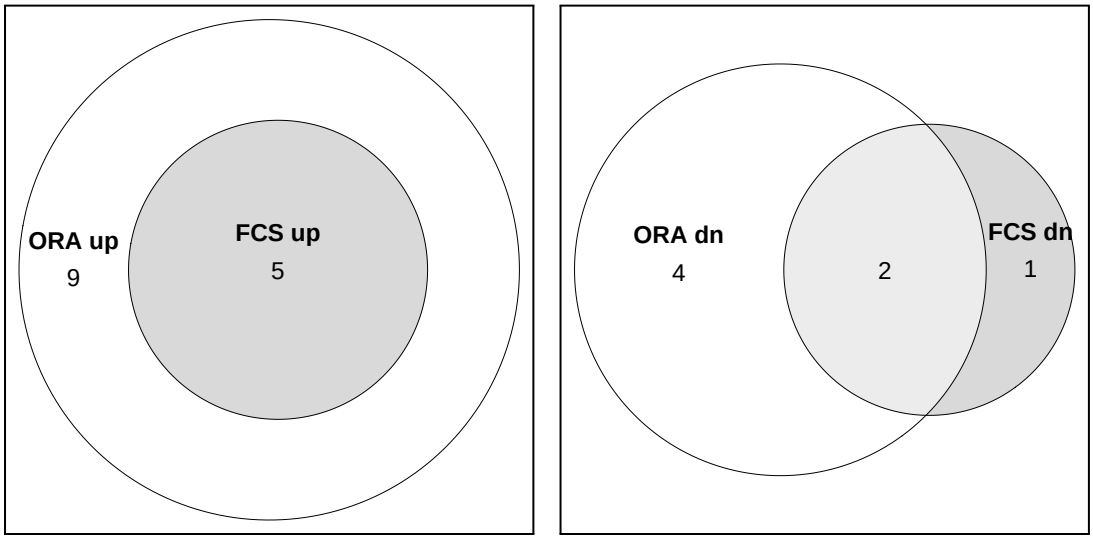


(A) Venn diagrams of upregulated GO terms

(B) Venn diagrams of downregulated GO terms

Figure C.7: Overlap of ORA and FCS GO Terms

Overlaps of ORA and FCS GO terms.



(A)

(B)

Figure C.8: Overlap of ORA and FCS Reactome Terms

Overlaps of ORA and FCS GO terms.

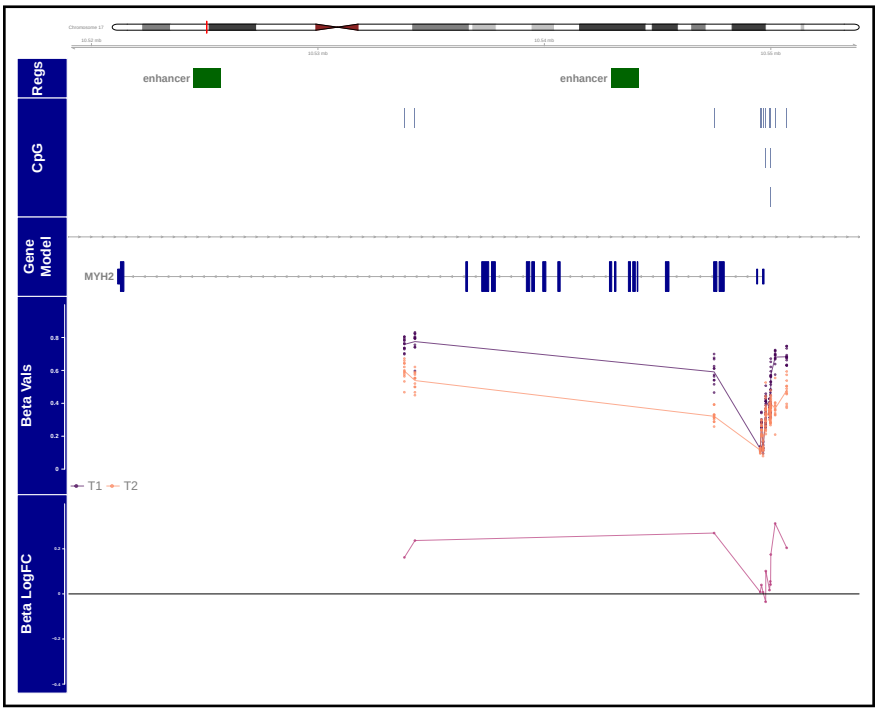


Figure C.9: MYH2-Genomic-Viz

Figure D.9 is a visualisation of methylation values as measured with the EPICv2 across the myosin 2 gene. The genome visualisation shows all the CpGs (blue bars) across the MYH7 gene on chromosome 14. Enhancers are indicated by green blocks, the gene model is indicated by darkblue genetracks. The orange line represents the beta value of TII fibres from each of the 12 samples and the purple line represents the beta value of TI fibres from each of the 12 samples. The pink line represents the beta difference between TI and TII fibres at all CpGs across the genomic region.

D Results 4

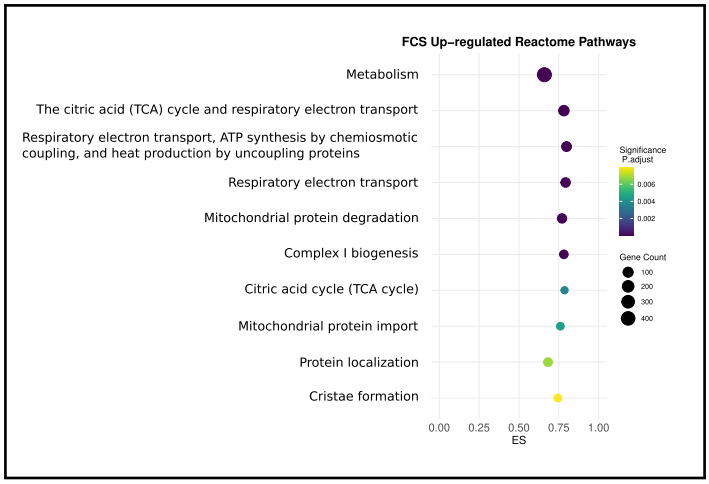
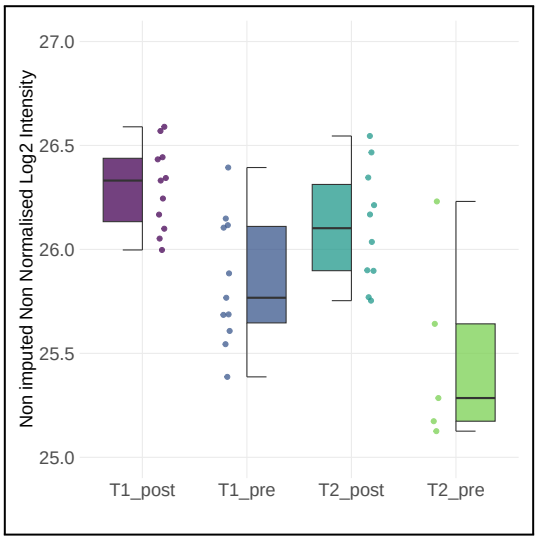
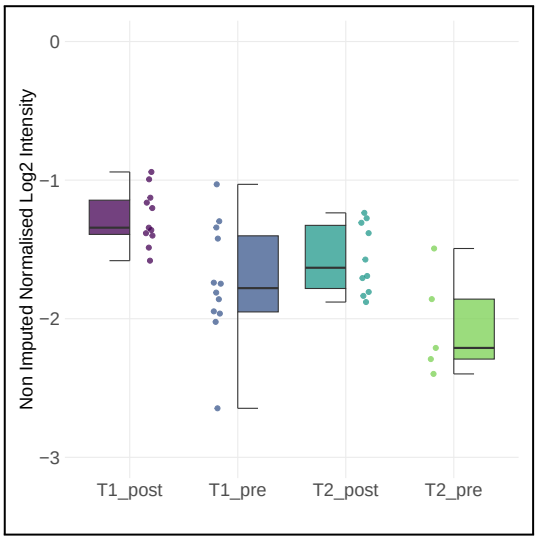


Figure D.1: FCS Reactome Terms of Upregulated DEPs

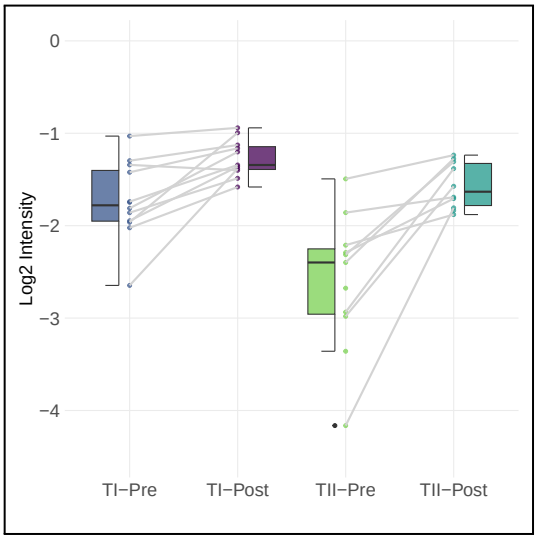
FCS Reactome gene set enrichment analysis of the up-regulated DEPs between TI and TII fibres. The circle size represents gene numbers, and the colour represents the adj-P-value and the enrichment scores are on the x axis.



(A) Non Imputed Non Normalised Transformed



(B) Non Imputed Normalised Transformed



(C) Imputed Normalised Transformed

Figure D.2: BCS1L values before and after imputation

Figure E.3 shows Log2 protein intensities of BCS1L in TI-pre, TI-post, TII-pre and TII-post samples. (A) shows the non-imputed non-normalised log2 intensities, (B) shows the non-imputed normalised log2 intensities and (C) shows all the log2 intensities including those that were imputed.