UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE

Human Development and Health

Volume 1 of 1

Molecular and Epigenetic Changes Associated with Age-Related Skeletal Muscle Loss and Sarcopenia

by

Elie Antoun

Thesis for the degree of Doctor of Philosophy

May 2018

“Creativity Is Intelligence Having Fun.”

- Albert Enstein

Acknowledgments

First of all, I would like to thank my supervisors, Prof. Karen Lillycrop, Prof. Keith Godfrey and Dr Harnish Patel for all their support and guidance throughout my PhD journey. Your support has helped make me a better scientist and I am truly grateful for this great opportunity you have given me.

To Emma, thanks for all your help throughout my PhD. Working together on the first muscle project EpiGen have had has been great, despite the hurdles we have had to work through and negotiate. You have always been there to offer advice on my lab work and have taught me the majority of what I have learnt throughout my time in EpiGen. I want to also thank the whole of the EpiGen group for their advice on the technical aspects of my PhD, helping when things go wrong and helping when things get too busy for me alone, some things could not have been done without your assistance.

To the fencing club, I want to thank them for keeping me sane throughout the past 4 years. Allowing me to hit them with swords after a stressful day at work followed by unwinding after a couple of drinks at the pub. Being around such an energetic group of people has kept reminding me that things are never as bad as they seem and there is always tomorrow.

Finally, to my mum, my dad and my whole family. For always being there when I needed you and being the people I can turn to during the bad days and the good days. Although a lot of what I say, you don't understand, thanks for being interested in my work the past four years, and encouraging me to keep going and push myself further. You are always there for me and I thank you massively for that.

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE

Human Development and Health

Thesis for the degree of Doctor of Philosophy

Molecular and Epigenetic Changes Associated with Age-Related Skeletal Muscle Loss and Sarcopenia

Elie Antoun

Sarcopenia is defined as the age-related loss of muscle mass and function and is common in both men and women over the age of 65, with a worldwide prevalence estimated at between 3-30%. As the population is becoming more aged, the prevalence of sarcopenia in the community is increasing and with that, an increase in the number of adverse physical and metabolic outcomes including frailty, disability, metabolic disorders and osteoporosis. However, the molecular pathways altered during muscle ageing and how they contribute to sarcopenia are poorly understood.

Previous studies have investigated the mechanisms contributing to muscle ageing and the pathogenesis of sarcopenia. However to date, there have been no investigations into the variability in skeletal muscle mass and function in the elderly, using age-matched controls to investigate the pathogenesis of sarcopenia.

Using both genome-wide and candidate gene approaches, skeletal muscle tissue and myoblasts from participants of the Hertfordshire Sarcopenia Study were utilized and the signalling pathways that contribute to the variability in muscle mass and function among community-dwelling older people were investigated. Genome-wide transcriptome analysis highlighted mitochondrial function, DNA damage and myogenesis as the major pathways altered in sarcopenia. RT-PCR validation of the RNAseq data confirmed an association between reduced muscle mass and higher expression of the long non-coding RNA H19, with a concomitant increase in miR-675-3p/5p expression and decreased SMAD1/5 expression, resulting in reduced muscle hypertrophy. Single-cell transcriptomics of the isolated myoblasts from the muscle biopsies showed that many of the pathways altered with respect to muscle mass were intrinsic to the muscle cells. Changes in transcription were also accompanied by genome wide changes in DNA methylation with pathways related to calcium signalling, denervation/muscle atrophy and muscle development being changed and enriched for the polycomb regulator EZH2. Moreover, the epigenetic signatures related to muscle mass and strength were conserved in primary myoblasts after culture. Skeletal muscle myoblasts isolated from sarcopenic muscle also exhibited altered mitochondrial respiration, with reduced ATP production and maximal respiration compared to healthy elderly people, as well as increased levels of senescence in myoblasts as suggested by an increase in $p16^{INK4a}$ expression and a decrease in ANRIL expression.

Taken together, these findings suggest an impairment in the balance between muscle hypertrophy and atrophy in sarcopenia. This impairment is intrinsic to myogenic cells, with increased cell death,

in combination with increased mitochondrial dysfunction, cellular senescence and DNA damage. This may result in the impaired regenerative capacity of skeletal muscle, and together with reduced hypertrophic signalling, a reduced capability to efficiently repair skeletal muscle in the elderly. Understanding these mechanisms may provide better insight into the development of therapeutics for the treatment of sarcopenia, while the muscle-derived myoblasts provide an *in vitro* model system to investigate the efficacy of new treatments. The methylation changes associated with the measures of muscle mass and function may also provide potential biomarkers to identify those at increased risk of developing sarcopenia, allowing early intervention for prevention and treatment.

Table of Contents

ABSTRACT ..... **i**

Table of Contents ..... **iii**

List of Figures ..... **xi**

List of Tables ..... **xiii**

DECLARATION OF AUTHORSHIP ..... **xv**

List of Abbreviations ..... **xvii**

Chapter 1 – General Introduction **5**

1.1. Skeletal Muscle *5*

1.1.1. Structure of Skeletal Muscle 5

1.1.1.1 Specific Functions of Skeletal Muscle Compartments 7

1.1.1.2. Heterogeneity of Skeletal Muscle 8

1.1.2. Skeletal Muscle Physiology 10

1.1.3. Skeletal Muscle Development and Growth 11

1.1.4. Exercise and Adaptation to Training 12

1.1.5. Skeletal Muscle Atrophy 13

1.1.6. Skeletal Muscle Repair 14

1.2. Sarcopenia and Ageing *15*

1.2.1. Epidemiology 15

1.2.2. Clinical Parameters for Diagnosis of Sarcopenia 17

1.2.3. Pathophysiology of Sarcopenia 19

1.2.3.1. Physical Activity 19

1.2.4. Skeletal Muscle Ageing 20

1.2.5. Molecular Mechanisms of Sarcopenia and Ageing 21

1.2.5.1. Cellular Senescence 22

1.2.5.1.1. Mechanisms of Cellular Senescence 23

1.2.5.1.1.1. p16-Rb Pathway 23

1.2.5.1.1.2. p53-p21-Rb Pathway 24

1.2.5.1.2. Cellular Senescence during Ageing 24

1.2.5.1.3. Cellular Senescence in Skeletal Muscle 25

1.2.5.1.4. Senescence-Associated Secretory Phenotype 26

1.2.5.2. Mitochondrial Dysfunction and the Mitochondrial Theory of Ageing 27

1.2.5.3. Inflammaging 29

1.2.5.4. Neuromuscular Ageing 31

1.2.5.5. Hormone Levels and Sensitivity 31

1.2.5.6. Genomic Instability 32

1.2.5.7. Epigenetic Alterations 33

1.2.5.8. Loss of Proteostasis 33
1.2.5.9. Deregulated Nutrient Sensing 34
1.2.5.10. Stem Cell Exhaustion 34
1.2.6. The Skeletal Muscle Transcriptome during Ageing 34
1.2.6.1. Skeletal Muscle Transcriptome during Ageing 35
1.2.7. Genetic Mechanisms of Sarcopenia 36
1.2.8. Current Treatments for Sarcopenia 37

1.3. Epigenetics and Regulation of Gene Expression 38
1.3.1. DNA Methylation 39
1.3.2. Histone Modifications 41
1.3.2.1. Histone Acetylation 41
1.3.2.2. Histone Methylation 41
1.3.3. MicroRNAs 42
1.3.4. The Skeletal Muscle Epigenome 43
1.3.4.1. DNA Methylation during Skeletal Muscle Development 43
1.3.4.2. DNA Methylation during Skeletal Muscle Ageing 44
1.3.4.3. Histone Modification Changes during Skeletal Muscle Ageing 45

1.4. Aims 45

Chapter 2 – Methods 51

2.1. Materials 51

2.2. Methods 55
2.2.1. Cohort details 55
2.2.2. Tissue Collection and Anthropometric Measurements 55
2.2.3. Myoblast Cell Cultures 56
2.2.4. Immuno-magnetic Bead Sorting of Cells 57
2.2.5. RNA extraction 58
2.2.6. DNA Extraction 59
2.2.7. Protein Extraction 60
2.2.8. Total RNA-Seq Library Prep and Sequencing 61
2.2.9. DropSeq Encapsulation, Library Prep and Sequencing 61
2.2.9.1. Manufacturing Microfluidic Droplet Devices 61
2.2.9.2. Cell Suspension Preparation 62
2.2.9.3. Droplet Encapsulation and Breakage 62
2.2.9.4. Reverse Transcription and Exonuclease I Treatment 62
2.2.9.5. Library Prep and Sequencing 64
2.2.10. Methylation Arrays 64
2.2.11. Total RNAseq Bioinformatic Analysis 64
2.2.11.1 QC and pre-processing of reads 64
2.2.11.2 Read alignment to the genome 66

2.2.11.3. Count table generation 66
2.2.11.4. Differential gene expression 66
2.2.11.5. Gene Set Enrichment Analysis 67
2.2.12. Methylation Array Bioinformatic Analysis 67
2.2.12.1 QC and pre-processing of file 67
2.2.12.2. Differential Methylation Analysis 68
2.2.12.3. Pathway analysis 68
2.2.13. DropSeq Bioinformatic Analysis 68
2.2.13.1 QC and pre-processing of files 68
2.2.13.2. Cell/Gene Level QC 69
2.2.13.3. Data Normalization 69
2.2.13.4. Cell Cycle Classification 70
2.2.13.5. Highly Variable Genes and Cell Cluster Identification 70
2.2.13.6. Differential Expression for Marker Gene Identification 70
2.2.13.7. Gene Pathway Analysis 71
2.2.14. cDNA Synthesis 71
2.2.15. Quantitative Real Time PCR (qRT-PCR) 71
2.2.16. gDNA Bisulfite Conversion 72
2.2.17. Quantitative DNA Methylation Analysis 72
2.2.17.1. Pyrosequencing 72
2.2.17.2. Sequenom 72
2.2.18. siRNA Knockdown/miRNA Ectopic Expression 73
2.2.19. Western Blotting 74
2.2.20. Metabolic Flux Assay 75
2.2.21. Immunocytochemistry 76
2.2.22. BrdU Proliferation Assay 76
2.2.23. β-Galactosidase Senescence Assay 77
2.2.24. Statistics 77

Chapter 3 – INK4a/ARF Locus Genes Correlate With Ageing and Inflammation in Human Skeletal Muscle83

3.1. Introduction 83
3.1.1. INK4a/ARF Locus and the Cell Cycle 83
3.1.2. INK4A/ARF Locus and Frailty 85
3.1.3. Aims 86

3.2. Results 86
3.2.1. Cohort Characteristics 86
3.2.2. Expression of the Senescence Marker, $p16^{INK4a}$, and ANRIL is Associated with Age 87
3.2.3. ANRIL Correlates with the Expression of $p16^{INK4a}$ in Myoblasts but Not in Native Muscle Tissue 87
3.2.4. Expression of ANRIL and $p16^{INK4a}$ in Myoblasts Correlates with Expression in Muscle Tissue 89
3.2.5. ANRIL May Regulate the Expression of the Frailty Marker and Pro-Inflammatory Cytokine, IL6 .91

3.2.6. Methylation at the ANRIL Promoter Correlates with ANRIL Expression in Aged Myoblasts 92
3.2.7. Expression of $p16^{INK4a}$ and ANRIL in the myoblasts correlates with measures of muscle mass and function 92

3.3. Discussion *94*
3.3.1. Geroconversion from Quiescence to Senescence in Myoblasts and Skeletal Muscle Tissue 94
3.3.2. DNA Methylation May Regulate the Expression of ANRIL in Skeletal Muscle Myoblasts 95
3.3.3. Older Myoblasts Exhibit the Senescence-Associated Secretory Phenotype 97
3.3.4. Novel Role for ANRIL in Skeletal Muscle 97
3.3.5. Further Considerations 99
3.3.6. Conclusion 99

Chapter 4 – Genome Wide Changes in Skeletal Muscle Associated with the Severity of Sarcopenia..... 103

4.1 Introduction *103*
4.1.1. The effect of ageing on the muscle transcriptome 103
4.1.2. Aims 105

4.2 Results *105*
4.2.1. Cohort Characteristics 105
4.2.2. RNAseq and QC Analysis 105
4.2.3. RNA-seq Alignment Metrics 108
4.2.4. HTSeq Metrics 110
4.2.5. Count Table Analysis 111
4.2.6. Differential Gene Expression 113
.....118
4.2.7. Gene Set Enrichment Analysis 118
4.2.7.1. Phenotypic variables 118
4.2.7.2. Sarcopenia 121

4.3 Discussion *123*
4.3.1 Mitochondrial Theory of Ageing and Sarcopenia 123
4.3.2. Protein Synthesis and Ribosomal Function in Aged Skeletal Muscle 125
4.3.3. Role of H19 in Skeletal Muscle Ageing 126
4.3.4. AGL in Healthy and Diseased Muscle 128
4.3.5. DNA Damage and Repair 129
4.3.6. Inflammation 130
4.3.7. Further Considerations 131
4.3.8. Conclusion 131

Chapter 5 – The long non-coding RNA H19 is associated with low muscle mass in community-dwelling older people – findings from the Hertfordshire Sarcopenia Study (HSS) 137

5.1. Introduction *137*
5.1.1. Mechanisms controlling muscle mass 137

5.1.2. H19 in Skeletal Muscle 137
5.1.3. Aims 139

5.2. Results *139*
5.2.1. Participant characteristics 139
5.2.2. H19 expression in skeletal muscle 140
5.2.3. H19 promoter/ICR methylation 142
5.2.4. Association of miR-675-3p/5p expression with ALMI 143
5.2.5. SMAD1 and SMAD5 are potential targets for miR-675-3p/5p 145
5.2.6. H19-miR-675-3p/5p directly act to affect the expression of SMAD1/5 in human myotubes 146
5.2.7. H19-miR-675-3p/5p may act via the same pathway in myoblasts 147
5.2.8. SMAD1/5 may act to increase myosin heavy chain content and protect from muscle atrophy 147

5.3. Discussion *148*
5.3.1. Role of H19 in Skeletal Muscle Hypertrophy 148
5.3.2. Imprinting and H19 methylation 150
5.3.3. Conclusion 151

Chapter 6 – Cultured Human Myoblasts Provide a Model for the Investigation of Skeletal Muscle Ageing and Interventions 155

6.1. Introduction *155*
6.1.1. Skeletal Muscle Ageing 155
6.1.2. Cellular Models of Skeletal Muscle Function 156
6.1.3. Aims 156

6.2. Results *157*
6.2.1. Cohort Characteristics 157
6.2.2. Beta Values Sample QC 158
6.2.3. Beta Values Probe QC 161
6.2.4. Differentially Methylated Probes 161
6.2.4.1. Differential Methylation in Muscle Tissue 161
6.2.4.2. Differential Methylation in Myoblast Samples 164
6.2.4.3. Myoblasts Retain the Methylation Signature Associated with Muscle Mass and Function 166
6.2.5. Myoblast Functional Characterisation 167
6.2.5.1. Myoblast Proliferation Capacity 167
6.2.5.2. Senescence 167
6.2.5.3. Differentiation Potential of Myoblasts 168
6.2.5.4. Glycolysis and Mitochondrial Respiration of Myoblasts and Myotubes 169
6.2.5.5. Correlations between Myoblast/Myotube Molecular Characteristics 170

6.3. Discussion *172*
6.3.1. Cultured Human Myoblasts as an *in vitro* model 172
6.3.1.1. Muscle-Derived Myoblasts Retain an Epigenetic Memory 173

6.3.1.2. Muscle-Derived Myoblasts Retain the Functional Characteristics of Muscle 174
6.3.2. Further Considerations 176
6.3.3. Conclusions 176

Chapter 7 – Genome-wide methylation analysis human skeletal muscle obtained from older people ... 181

7.1. Introduction 181
7.1.1. Molecular Function of DNA Methylation 181
7.1.2. DNA Methylation in Skeletal Muscle 182
7.1.3. Aims 183

7.2. Results 183
7.2.1. Cohort Characteristics 183
7.2.2. Beta Values QC 184
7.2.3. Beta Values Probe QC 184
7.2.4. Differentially Methylated Probes 188
7.2.4.1. Continuous Analysis 188
7.2.4.2. Categorical Analysis 189
7.2.4.3. Genomic Locations of dmCpGs 190
7.2.4.4. Transcription Factor Enrichment at dmCpG-Associated Genes 191
7.2.5. Differentially Methylated Regions 191
7.2.6. Pathway Analysis 194
7.2.7. Methylation Clock 195

7.3. Discussion 200
7.3.1. DNA Methylation and Skeletal Muscle Mass and Function 200
7.3.2. DNA Methylation and Other Epigenetic Changes 202
7.3.3. Skeletal Muscle DNA Methylation Age 203
7.3.4. Further Considerations 204
7.3.5. Conclusions 204

Chapter 8 – Single Cell RNAseq Shows Cellular Heterogeneity of Skeletal Muscle Myoblasts from Older Individuals 209

8.1. Introduction 209
8.1.1. Satellite Cell Heterogeneity 209
8.1.2. Single-Cell Analysis in Skeletal Muscle 210
8.1.3. Aims 210

8.2. Results 211
8.2.1. Cohort Characteristics 211
8.2.2. Cell/Gene Level QC Metrics 211
8.2.3. Data Normalization 215
8.2.4. Cell Cycle Classification 215
8.2.5. Identification of Highly Variable Genes (HVGs) 216

8.2.6. Cell Clustering 218
8.2.7. Marker Genes between Cell Subpopulations 223
8.2.8. Marker Genes Pathway Analysis 225
8.2.9. Cell Clustering of Myoblasts from Healthy and Sarcopenic Individuals 227
8.2.9.1. Myoblasts from Healthy Individuals 227
8.2.9.2. Myoblasts from Sarcopenic Individuals 228

8.3. Discussion 229
8.3.1. Cellular Heterogeneity in Human Skeletal Muscle Myoblasts 229
8.3.2. Consequences of Myoblast Heterogeneity in Sarcopenia 230
8.3.3. Further Considerations 232
8.3.4. Conclusions 232

Chapter 9 – Discussion 237

9.1. Decreased Regenerative Capacity May Contribute to Decreased Skeletal Muscle Repair 237
9.1.1. Cellular Senescence 238
9.1.2. Mitochondrial Dysfunction and ROS 239
9.1.3. Increased DNA Damage and Cell Death 240
9.1.4. Altered Myoblast Proliferation and Differentiation in Aged Skeletal Muscle 242

9.2. Decreased Hypertrophy and Increased Atrophy in Aged Skeletal Muscle 242
9.2.1. Role of H19 Signalling in Regulating Muscle Hypertrophy 242
9.2.2. Role of HDAC4-Myogenin on Muscle Atrophy 243

9.3. Implications for Sarcopenia in the Population 244

9.4. Future Directions 246

Appendices 249

References 295

List of Figures

Figure 1.1: Overall structure of human skeletal muscle 6
Figure 1.2: EWGSOP characterisation of sarcopenia 16
Figure 1.3: EWGSOP extended definition of sarcopenia 17
Figure 1.4: Main contributors to the pathogenesis of sarcopenia 19
Figure 2.1: Flow chart depicting the recruitment of participants from the Hertfordshire Cohort Study (HCS) to the Hertfordshire Sarcopenia Study (HSS) 56
Figure 2.2: Experimental outline of HSS and HSSe Cohorts 57
Figure 2.3: CD56 staining of myoblast cultures. 58
Figure 2.4: RNA and gDNA extractions. 59
Figure 2.5: DropSeq microfluidics. 61
Figure 2.6: Bioanalyzer traces of DropSeq libraries. 63
Figure 2.7: Independent filtering 65
Figure 2.8: Methylation analysis PCR products 73
Figure 2.9: Results from the transfection controls 74
Figure 2.10: Seahorse XF96 assay optimisation results 75
Figure 3.1: Schematic of the 9p21 susceptibility locus 84
Figure 3.2: Expression of CDKN2A and ANRIL in the myoblasts 88
Figure 3.3: Expression of CDKN2A and ANRIL in the muscle tissue 89
Figure 3.5: Correlation between IL6 and CDKN2A/ANRIL in the myoblasts 90
Figure 3.4: Correlation between CDKN2A and ANRIL in the myoblasts and muscle tissue 90
Figure 3.6: Associations between p16INK4a and ANRIL expression and ALMI and gait speed 93
Figure 4.1: Clinical variables as a function of sarcopenia status 107
Figure 4.2: RNA-seq Alignment Metrics as determined by Picard tools 109
Figure 4.3: Total read metrics 110
Figure 4.4: Count data visualization 111
Figure 4.5: Heatmap of the 1000 most variable genes 112
Figure 4.6: MA plots and volcano plots of differentially expressed genes 117
Figure 4.7: Overlap of the differentially expressed genes 118
Figure 5.1: Validation and replication of the expression of H19 in HSS and HSSe muscle tissue determined by RT-PCR. 140
Figure 5.2: Schematic of the H19 promoter region and ICR 141
Figure 5.3: miR-675-3p/5p expression in skeletal muscle tissue 143
Figure 5.4: SMAD1/5 RNA and protein expression in skeletal muscle and association with miR-675-3p/5p expression 144
Figure 5.5: Effect of siRNA-mediated H19 knockdown and miR-675-3p/5p ectopic expression on SMAD1/5 levels in myotubes 145

Figure 5.6: Effect of siRNA-mediated H19 knockdown and miR-675-3p/5p ectopic expression on SMAD1/5 levels in myoblasts....146
Figure 5.7: Role of H19-miR-675-SMAD1/5 axis in skeletal muscle hypertrophy....148
Figure 6.1: PCA plots of samples for the first two principle components....158
Figure 6.2: Hierarchal clustering of samples based on their Euclidean distances....159
Figure 6.3: Manhattan plots for the differential methylation analysis in muscle and myoblasts....162
Figure 6.4: Q-Q plots of the observed vs expected p values of the differential methylation analysis in muscle ..163
Figure 6.5: Q-Q plots of the observed vs expected p values of the differential methylation analysis in myoblasts 164
Figure 6.6: Regression coefficients in the myoblasts against those in the muscle samples for all probes. 165
Figure 6.7: Regression coefficients in the myoblasts against those in the muscle samples for probes with p-value <0.05....166
Figure 6.8: Senescence-associated b-gal staining for senescent cells....168
Figure 6.9: Myosin Heavy Chain staining of differentiated myotubes at day 2 of differentiation from participants with low and high gait speed....169
Figure 7.1: Hierarchal clustering of samples based on their Euclidean distances....185
Figure 7.2: PCA plots of samples for the first two principle components....186
Figure 7.3: The most significantly differentially methylated....189
Figure 7.4: Distribution of significant CpGs within different genomic regions....190
Figure 7.5: Significant DMRs found in skeletal muscle....193
Figure 7.6: Gene clusters showing significant gene ontology enrichment with genes showing increased methylation....197
Figure 7.7: Gene clusters showing significant gene ontology enrichment with genes showing decreased methylation....198
Figure 7.8: DNA methylation age comparison of skeletal muscle tissue....199
Figure 8.1: Representative density plot and cumulative read fraction (knee) plot to determine actual number of cells sequenced....212
Figure 8.2: Cell QC metric plots....213
Figure 8.3: Expression plot of the 50 genes with the highest expression across the dataset....214
Figure 8.4: Scatter plot of library sizes against size factors....215
Figure 8.5: Cell cycle classification of the myoblasts....216
Figure 8.6: Violin plots of the top 10 HVGs....218
Figure 8.7: t-SNE plot of the cell subpopulations....220
Figure 8.8: Average percentage of cells in each cluster....222
Figure 8.9: t-SNE plots for the cluster analysis of healthy cells and sarcopenic cells individually....226
Representative images of FastQC results, before concatenating the 8 fastq files....252
FastQC results for sample 1, following concatenation, quality and adapter trimming....256
FastQC results for sample 1, following concatenation, quality and adapter trimming....256
FastQC results for sample 1, following concatenation, quality and adapter trimming....256
FastQC results for sample 1, following concatenation, quality and adapter trimming....256

List of Tables

Table 1.1. List of myosin isoforms 7

Table 1.2. Characteristics of the major muscle fibre types in humans 9

Table 2.1: Primers used for DropSeq experiments 78

Table 2.2: List of primers used in RT-PCR and methylation 79

Table 3.1: Linear regression analysis results 91

Table 3.2: Pearson correlation analysis results 91

Table 3.3: DNA methylation analysis of the ANRIL DMR 92

Table 4.1: Characteristics of HSS participants that underwent RNAseq analysis 106

Table 4.2: List of differentially expressed genes (FDR<0.1) 114

Table 4.3: List of differentially expressed genes in pre-sarcopenia vs healthy controls (FDR<0.1) 115

Table 4.4: List of top differentially expressed genes in sarcopenia vs healthy controls (FDR<0.1) 116

Table 4.5: Top 5 positively and negatively enriched hallmark gene sets 120

Table 4.6: Gene set enrichment analysis of differentially expressed genes associated with sarcopenia 122

Table 5.1: Cohort Characteristics 139

Table 5.2: Correlation results of DNA methylation at the H19 promoter and ICR and H19 expression 142

Table 6.1: Cohort characteristics for participants of HSSe used in this study 157

Table 6.2: MADscores for all samples plus replicates 160

Table 6.3: Genomic inflation factors of the p-values for the differential methylation analysis 163

Table 6.4: Spearman rho correlations of regression coefficients between the muscle and myoblasts 167

Table 6.5: Results of the characterisation of the myoblasts and myotubes with respect to ALMI 170

Table 6.6: Results of the characterisation of the myoblasts and myotubes with respect to gait speed 171

Table 6.7: Results of the characterisation of the myoblasts and myotubes with respect to grip strength 171

Table 7.1: Cohort characteristics for participants of HSS, HSSe and combined datasets used in the methylation analysis 183

Table 7.2: Table of madscores (modified z-scores) for all samples 187

Table 7.3: List of enriched TF binding sites 191

Table 7.4: List of DMRs identified 192

Table 7.5: Genes in each individual cluster identified by Cytoscape 196

Table 8.1: Cohort characteristics for participants of HSSe whose myoblast cultures were used for the scRNAseq analysis 211

Table 8.2: List of the top 50 highly variable genes ordered based on their biological variance 217

Table 8.3: Top 50 correlations between the HVGs 219

Table 8.4: Number of cells in each cluster, divided by sarcopenia status 221

Table 8.5: List of the top 20 marker genes for each cell cluster 224

Table 8.6: List of the top 20 marker genes for each cell cluster in the analysis of cells from healthy individuals 227

Table 8.7: List of the top 20 marker genes for each cell cluster in the analysis of cells from sarcopenic individuals 228

DECLARATION OF AUTHORSHIP

Academic Thesis: Declaration Of Authorship

I, Elie Antoun, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Genetic and Epigenetic Changes Associated with Age-Related Skeletal Muscle Loss and Sarcopenia

I confirm that:

This work was done wholly or mainly while in candidature for a research degree at this University;

Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

Where I have consulted the published work of others, this is always clearly attributed;

Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

I have acknowledged all main sources of help;

Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

None of this work has been published before submission

Signed: ……

Date: ……

List of Abbreviations

Abbreviation	Full Term
6mTUG	6 metre Timed Up-and-Go
ALM	Appendicular Lean Mass
ALMI	Appendicular Lean Mass Index
ANRIL	Antisense Non-Coding RNA in the INK4A Locus
AWGS	Asian Working Group on Sarcopenia
BMI	Body Mass Index
BMP	Bone Morphogenetic Protein
BP	Biological Process
BrdU	Bromodeoxyuridine
CC	Cellular Compartment
CDK	Cyclin-Dependent Kinase
cDNA	Complementary DNA
CpG	Cytosine-Phosphate-Guanine
CORT	Cortisol
CPT1B	Carnitine Palmitoyltransferase 1, Muscle Isoform
CSA	Cross-Sectional Area
CTCF	CCCTC-Binding Factor
DDR	DNA Damage Response
DGE	Digital Gene Expression
dmCpG	Differentially Methylated CpG
DMEM	Dulbecco's Modified Eagle Medium
DMR	Differentially Methylated Region
dNTPs	Deoxynucleotide Triphosphates
DXA	Dual X-ray Absorptiometry
ECAR	Extracellular Acidification Rate
ETC	Electron Transport Chain
EWGSOP	European Working on Sarcopenia in Older People
EZH2	Enhancer of Zester Homolog 2
FAO	Fatty Acid Oxidation
FBS	Foetal Bovine Serum
FDR	False Discovery Rate
FFMI	Fat-Free Mass Index
FNIH	Foundation for the National Institutes of Health
GC	Glucocorticoid
GO	Gene Ontology
GUSTO	Growing Up in Singapore Towards Healthy Outcomes
GSEA	Gene-Set Enrichment Analysis
HCS	Hertfordshire Cohort Study
HDAC	Histone Deacetylase
HS	Horse Serum
HSS	Hertfordshire Sarcopenia Study
HSSe	Hertfordshire Sarcopenia Study extension
HVGs	Highly Variable Genes
ICD-10	International Statistical Classification of Diseases and Related Health Problems 10th Revision
ICR	Imprinting Control Region
IGF	Insulin-like Growth Factor

IL	Interleukin
IWGS	International Working Group on Sarcopenia
lncRNA	Long non-coding RNA
MF	Molecular Function
miRNA	Micro RNA
MPC	Myogenic Precursor Cell
mtDNA	Mitochondrial DNA
MuRF1	Muscle RING-Finger Protein 1
myomiRs	Muscle-specific miRNAs
NMJ	Neuromuscular Junction
OCR	Oxygen Consumption Rate
PBS	Phosphate-Buffered Saline
PBMC	Peripheral Blood Mononuclear Cells
PCA	Principle Component Analysis
PRC	Polycomb Repressive Complex
qRT-PCR	Quantitative Real-Time Polymerase Chain Reaction
Rb	Retinoblastoma Protein
RNAseq	RNA sequencing
ROS	Reactive Oxygen Species
SA-β-gal	Senescence-Associated β-Galactosidase
SASP	Senescence-Associated Secretory Phenotype
SC	Satellite Cells
SD	Standard Deviation
scRNAseq	Single-Cell RNAseq
siRNA	Small Interfering RNA
SMAD	Mothers Against Decapentaglegic Homolog
SNP	Single-Nucleotide Polymorphism
TF	Transcription Factor
TSS	Transcription Start Site
UTR	Untranslated Region
VST	Variance Stabilizing Transformation

Chapter 1 –

General Introduction

Chapter 1 – General Introduction

1.1. Skeletal Muscle

Skeletal muscle comprises around 40% of the total body weight and consists of between 50-75% of total body protein, accounting for around 40% of protein turnover and is one of the most plastic and dynamic tissues in the human body. Muscle mass, a function of myofibre number and size, is dependent on a balance between protein synthesis and degradation, processes that are sensitive to external stimuli such as nutritional status, hormone balance, physical activity and injury. As such, the regulation of muscle mass and fibre size reflects muscle protein turnover [1] The main function of skeletal muscle is the production of mechanical energy from chemical energy to generate sufficient force and power to maintain posture and produce movement that influences activity. Skeletal muscle is also the major site of glucose uptake and utilisation in the body accounting for approximately 85% of whole body glucose uptake [2,3]. As such, skeletal muscle also contributes to multiple metabolic processes other than movement

1.1.1. Structure of Skeletal Muscle

The overall structure of human skeletal muscle is shown in figure **1.1**. An individual muscle is surrounded by a layer of connective tissue called the epimysium. Muscle is composed of fascicles, which are groups of muscle fibres surrounded by a layer of connective tissue known as the perimysium. Muscle fibres are composed of individual myofibrils bundled together into groups. A single muscle fibre is surrounded by the sarcolemma (a cell membrane). Associated with the sarcolemma is a complex of proteins, physically connected to the internal myofilament structure, in particular the actin protein [4,5].

Skeletal muscle is organised into a very particular and well-described arrangement of muscle fibres and connective tissue [6]. The size of a muscle is generally determined by the number of fibres present and how large they are, although disease can cause an infiltration of fat and connective tissue that will increase muscle size. Muscle fibres are multinucleated with each nucleus controlling protein synthesis in the surrounding area, referred to as the myonuclear domain [7,8]. Protein expression in adjacent regions of a single muscle fibre are coordinated such that the protein isoform (in particular the myosin isoform) expressed is similar across the length of the fibre. Muscle also contains satellite cells that act as the resident stem cells of skeletal muscle [9], contributing to muscle growth, repair and regeneration [10]. They are located between the sarcolemma (the cell membrane of muscle fibres) and the basal lamina, present in a quiescent state. When activated by myogenic factors, they

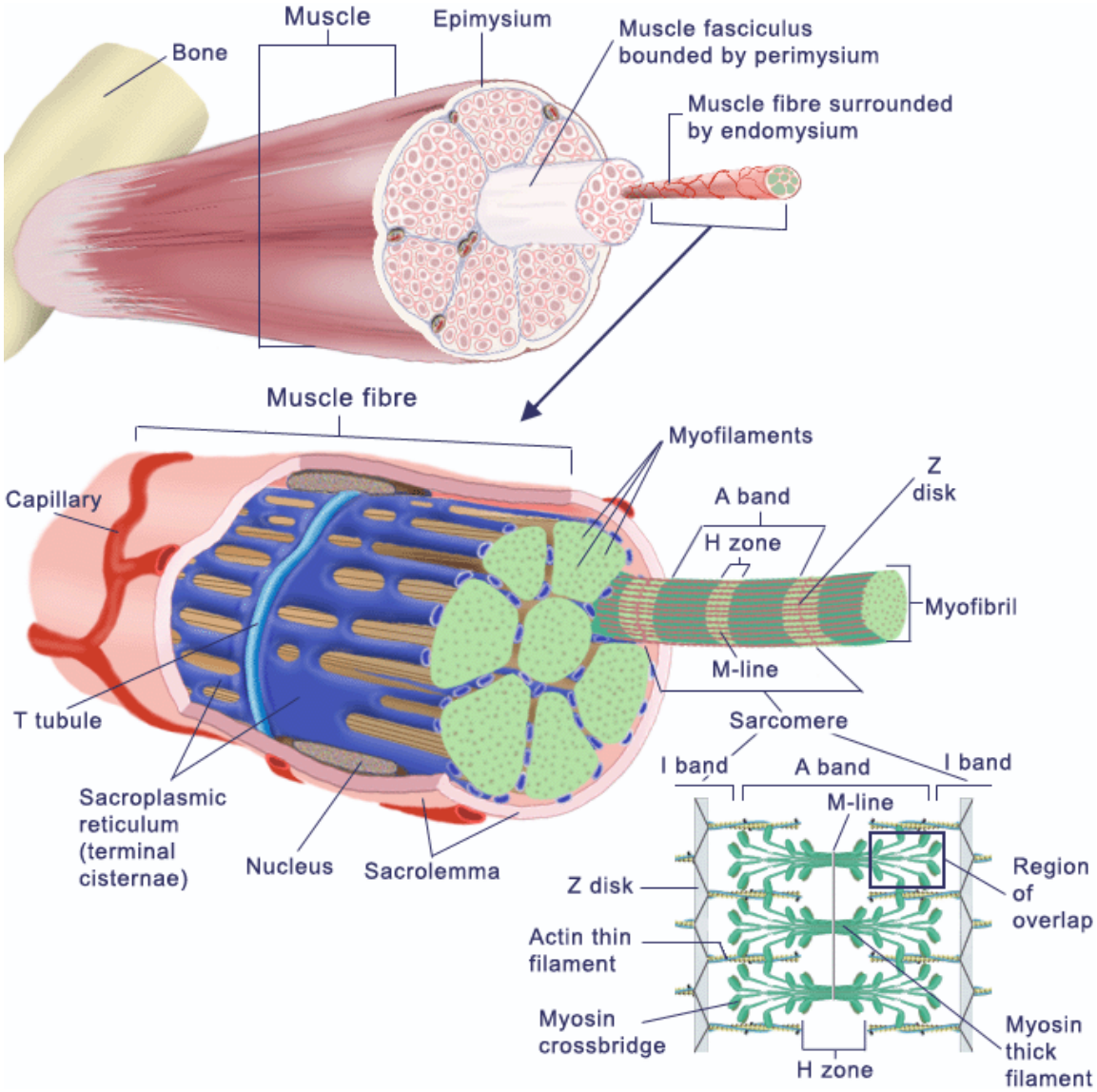


Figure 1.1: Overall structure of human skeletal muscle

The structure of human skeletal muscle, from a whole muscle, down to a muscle fibre and the individual myofilaments. Image modified from figure. 11-2 in Moffett, et al., Human Physiology, 2nd ed., Mosby, 1993, p. 293.

proliferate into myoblasts and differentiate into myotubes, fusing with existing fibres and aiding in muscle growth and repair.

Other than water, muscle fibres are composed of approximately 80% protein (contractile, regulatory and cytoskeletal) and 8% sarcoplasm [11]. The rest of the muscle is composed of connective tissue, vascular tissue, intramuscular adipocytes, non-protein nitrogen (e.g. ATP), carbohydrates (e.g. glycogen) and inorganic minerals (e.g. zinc and iron). An interplay of all the different resident cells/tissues is important in the normal function of skeletal muscle [12]. Each fibre is composed of myofibrils and each myofibril is composed of myofilaments. The two most abundant myofilaments are myosin and actin, making up 70-80% of a single muscle fibre. Actin and myosin filaments are assembled together to form sarcomeres, the most basic contractile unit of skeletal muscle. Myosin

is the main molecular motor and in mammals, a total of 11 myosin genes and their respective proteins have been described (table **1.1**) [13]. The sarcomere consists of several other proteins that contribute to cytoskeletal structure, excitation-contraction coupling, energy production and the generation of force and power, including the troponin complex [14], titin [15], nebulin [16] and desmin [17]. Other elements in the sarcoplasm include the transverse tubule system (T tubules), the sarcoplasmic reticulum and a mitochondrial network, the proportion of each element dependent on the fibre type. Several complexes of protein exist bound to the myofilaments that regulate muscle contraction, including troponin C, I, M and tropomyosin (TnC, TnI, TnM and TM respectively).

Table 1.1. *List of myosin isoforms*

Gene	Protein	Expression pattern
MYH1	MyHC-2X	Fast 2X fibres
MYH1	MyHC-2A	Fast 2A fibres
MYH3	MyHC-emb	Developing muscle
MYH4	MyHC-2B	Fast 2B fibres
MYH6	MyHC-α	Cardiac muscle
MYH7	MyHC-β/slow	Slow muscle and cadiac muscle
MYH8	MyHC-neo	Developing muscle
MYH9	MyHC-9	Non-muscle
MYH10	MyHC-10	Non-muscle
MYH11	MyHC-11	Smooth muscle
MYH12	MYO5A	Non-muscle
MYH13	MyHC-EO	Extraocular muscle
MYH14	MyHC-slow/tonic	Extraocular muscle at the protein level. Cardiac and slow muscle at the transcript level.
MYH15	MyHC-15	Extraocular muscle
MYH16	MYH5/MYH16 pseudogene	Inactivated in human masticatory muscles, only expressed in some mammalian species

1.1.1.1 Specific Functions of Skeletal Muscle Compartments

The T tubule system is part of the sarcolemma and has an important role in the conduction of an action potential to the inside of the cell. Proteins localized in the membrane of the T tubules are sensitive to fluctuations in calcium concentrations, allowing transmission of the action potential to the muscle fibre. The sarcoplasmic reticulum (SR) is important for the storage, release and reuptake of calcium ions during the excitation-contraction coupling [18] and contains two types of proteins important for maintaining calcium homeostasis: the ryanodine receptor (RyR) and the sarcoplasmic/endoplasmic reticulum calcium ATPase (SERCA). The ends of the SR, known as the terminal cisternae, are where the calcium ions are stored. These are in close contact with the T tubules and together form a structure known as the triad [19]. Mitochondria form a three-dimensional

network that generate the energy required for normal muscle function. Various stimuli, such as exercise and stress, as well as the normal ageing process can induce significant changes in the structure and function of the cellular elements of skeletal muscle [20,21]. For example, prolonged endurance exercise can increase the number and size of mitochondria [11], whilst in ageing, fragmentation of the sarcoplasmic reticulum, impairing calcium release and muscle activation can be seen [22]. These processes highlight the profound plasticity of skeletal muscle in humans.

1.1.1.2. Heterogeneity of Skeletal Muscle

The heterogeneity of human skeletal muscle is highlighted by the variability in biochemical, mechanical and metabolic profiles of individual muscle fibres [13,23]. There are four major fibre types, each with distinct myosin heavy chain composition, conferring differing biochemical and mechanical properties (table **1.2**). Type I muscle fibres are slow twitch fibres, expressing the MYH7 myosin isoform. These fibres contain a higher density of mitochondria and mainly use aerobic/oxidative respiration. Due to the slow-twitch nature of type I fibres, the power generated by type I fibres alone is relatively low but type I fibres are fatigue-resistant allowing for endurance activities. Type II fibres are sub-divided into 3 main forms: type IIa, IIb and IIx, depending on the myosin isoform expressed: MYH2, MYH4 and MYH1 respectively. Type II fibres are all fast-twitch fibres to varying degrees, with type IIb fibres being the fastest, IIa the slowest of the fast-twitch fibres and IIx fibres exhibiting an intermediate phenotype between that of IIa and IIb. As a result, IIb fibres fatigue the quickest of all the fibre types and are glycolytic fibres, relying on glycogen for energy production over aerobic/oxidative respiration. Type IIb fibres also have a low mitochondrial content. Conversely, IIa fibres are oxidative-glycolytic fibres, making them fast-twitch, fatigue-resistant fibres, with a relatively high mitochondrial content relative to IIb fibres, but lower than type I fibres. However, in large mammals, including humans, type IIb fibres are not present. This is most likely due to energy conservation as type IIb fibres utilize a much greater proportion of ATP per unit of force produced compared to IIa and IIx fibres [24]. Several other myosin gene isoforms exist (table **1.1**) and are expressed in specific muscle types (e.g. MYH6 in cardiac muscle and muscles of the jaw and MYH15 expressed in extraocular muscles) and during development (MYH8 and MYH3). Under normal conditions, fibres may co-express more than one myosin heavy chain isoform, resulting in polymorphic fibres, the significance of which is not fully understood, but may play a role in adaptation to exercise and other stimuli.

Myofibre content in specific muscle bodies varies according to species and function. For example in humans, the soleus muscle contains predominantly slow-twitch fibres (type I) with a mean of approximately 80.4% (range 64-100%) whilst the gastrocnemius muscle contains 60.2% (range 45-82%) slow-twitch fibres [25]. The gastrocnemius muscle is primarily involved in running and other fast movements of the leg, and so requiring more fast-twitch fibres than the soleus muscle, which is

***Table 1.2**. Characteristics of the major muscle fibre types in humans*

	Slow-twitch	**Fast-twitch**		
	Type I	**Type IIa**	**Type IIx**	**Type IIb**
Myosin isoform expressed	MYH7	MYH2	MYH4	MYH1
Contraction speed	Slow	Moderately fast	Fast	Very Fast
Main respiration method	Oxidative	Oxidative-glycolytic	Glycolytic	Glycolytic
Resistance to fatigue	Very high	High	Intermediate	Low
Mitochondrial density	Very high	High	Intermediate	Low
Oxidative capacity	High	High	Intermediate	Low
Capillary density	High	Intermediate	Low	Low
Major fuel source	Triglycerides	Glycogen, creatine phosphate	Glycogen, creatine phosphate, ATP	Glycogen, ATP

involved in the plantarflexion of the foot, playing a major role in maintaining standing posture. The vastus lateralis muscle is composed of approximately 38% type I fibres at the surface of the muscle and approximately 47% deeper in the tissue [26]. This composition reflects the locomotor function of the vastus lateralis, being involved in movement, and the maintenance of posture when standing.

1.1.2. Skeletal Muscle Physiology

Excitation-contraction coupling is the process by which a stimulus is transmitted from the neuromuscular junction (NMJ) to the triad (the terminal cisternae and T tubules), followed by a release of calcium ions (Ca^{2+}) from the SR and cross-bridge formation between myosin and actin filaments, resulting in the contraction of muscle fibres [27,28]. Ca^{2+} ions are the main signal transducers in skeletal muscle, inducing fibre contraction. When an action potential arrives at the fibre membrane from the NMJ, the voltage-sensing subunit of the dihydropyridine receptor ($Ca_v1.1$) senses the change in voltage, which is propagated by sodium channels as action potentials to the T tubules. $Ca_v1.1$ undergoes a conformational change, allosterically activating the ryanodine receptor, RyR1, in the terminal cisternae of the SR. This causes a large influx of Ca^{2+} ions into the sarcoplasm. The amount of Ca^{2+} released is greater in fast fibres than in slow fibres [29,30]. The released Ca^{2+} binds to the regulatory protein troponin C (TnC). TnC, along with troponin I, M and tropomyosin (TnI, TnM and TM respectively), forms a regulatory complex blocking the availability of actin binding sites for the myosin filament. Upon binding calcium, TnC increases its affinity for TnI, releasing its inhibitory effect on actin and displaces TM, exposing the actin binding site, allowing the interaction between the actin and myosin filament.

As well as Ca^{2+} ions, ATP is important for muscle function and force production. The binding of ATP to the myosin detaches it from the actin filament [31]. The ATP molecule is hydrolysed by the ATPase activity of myosin, generating ADP and P_i. Hydrolysis of ATP generates sufficient energy to move the myosin filament into a 'cocked' position, where it is weakly bound to actin. Following the complete exposure of the binding site on the actin filament by the movement of TM, myosin binds strongly to the exposed site. The P_i is released by the myosin filament, producing a power stroke, moving the actin filament inwards and shortening the sarcomere. Myosin then releases the ADP molecule but remains attached to the actin filament in rigor state at the end of the power stroke until another ATP molecule binds. Once another ATP molecule binds, the cycle restarts. SERCA (sarcoplasmic/endoplasmic reticulum calcium ATPase) actively pumps Ca^{2+} ions back into the SR, reducing the Ca^{2+} back to resting levels [27]. The surrounding mitochondrial network provides ATP via the capillary network to the myofilaments. The rate of consumption of ATP is dependent on the myosin heavy chain expressed and in humans, it is lowest in slow fibres and highest in fast fibres [30]. Therefore, skeletal muscle requires a constant supply of ATP for normal function.

1.1.3. Skeletal Muscle Development and Growth

During early embryogenesis, skeletal muscle originates from epithelial structures called somites, which lie parallel to the neural tube [32]. Precursor cells delaminate from the somites and migrate to form the limb buds and the trunk. Somitic cells are pluripotent and their specification to the skeletal muscle lineage is regulated by sonic hedgehog signalling [33]. During limb myogenesis, precursor cells delaminate and migrate from the hypaxial region of the dermomyotome (dorsal region of the somite) into the limb buds. Pax3 is a transcription factor regulating the expression of c-met [34], controlling the delamination and migration of precursor cells [35]. Muscle precursors proliferate until they reach their final destination in the limb, after which they begin to express the myogenic regulatory factor Myf5 [36], followed by MyoD [36], myogenin [37] and MRF4 [38], which regulate myoblast differentiation, fusion and muscle fibre formation [39].

In humans, muscle fibre formation occurs in waves. The first involves myoblast fusion to form primary, embryonic fibres that define the overall structure of the muscle. This occurs during the early stages of development. Secondly, foetal fibres form using the primary fibres as a scaffold to grow longitudinally to the tendons on either side. This occurs during gestational weeks 10-18. Finally, tertiary fibres start to form around weeks 16-17. *De novo* muscle fibre formation subsequently ceases, determining adult muscle fibre number. The importance of embryonic myogenesis can be seen from pregnancies with intrauterine growth restriction (IUGR) due to placental insufficiency. Maternal nutrient restriction during pregnancy has been shown to limit the myoblast cell cycle activity in the foetus, resulting in a reduced number of myonuclei per myofibre and total myofibre number in the offspring [40,41]. This may be partly due to the reduced availability of foetal growth factors that regulate myoblast proliferation such as the insulin-like growth factors (IGFs). Muscle hypertrophy that occurs during late gestation is also affected by IUGR, due to decreased activity of the Akt-mTORC1 pathway [42-44], leading to reduced muscle mass. Low birthweight as a result of IUGR has been shown to be associated with reduced muscle mass and strength in children [45], early adulthood [46] and older adulthood [47,48]. This supports the concept that reduced intrauterine growth and muscle development during pregnancy can persist throughout the life span, influencing muscle health in later life.

After birth, skeletal muscles growth occurs via muscle hypertrophy rather than hyperplasia, increasing in size rather than number. Resident satellite cells proliferate and differentiate in response to stimuli, fusing with adjacent muscle fibres and providing its nucleus to the fibre, contributing to protein synthesis and hypertrophy. Postnatal muscle fibre hypertrophy in humans also involves a net increase in protein synthesis over degradation, regulated by the IGF1/mTOR pathway. As skeletal muscle is heterogeneous, resident progenitor cells in the muscle are capable of differentiating into diverging cell fates [49,50]. Increased intramuscular adipose tissue (IMAT) has

been associated with type 2 diabetes and impaired glucose tolerance in obesity [51], as well as ageing [52] and sarcopenia [53]. However, whether this increased IMAT is a result of the adipogenic differentiation of satellite cells or from preadipocytes is currently unknown.

1.1.4. Exercise and Adaptation to Training

Regular and repeated bouts of exercise in specific intensity and duration can result in significant structural [54,55] and metabolic changes [56] in muscle.

Endurance exercise involves the activation of muscle groups at an intensity that permits sustained activity over a prolonged period, without undue fatigue. This includes walking, running, cycling, swimming, etc., and requires efficient oxygen transport and delivery. Endurance exercise induces changes to the capillary supply to the active muscles [55,57], increasing oxygen delivery and the number and size of mitochondria. After 6 weeks of training, mitochondrial density increases by 50-100% in all fibre types, with a greater increase in type IIa fibres compared with type I and IIx fibres [58,59]. The process of mitochondrial biogenesis is complex and a highly regulated process, including the transcription and translation of nuclear genes, import of protein products into the mitochondria, replication of mitochondrial DNA (mtDNA), transcription of mitochondrial genes, and final assembly of enzyme complexes. Peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC1α) is viewed as the master regulator of mitochondrial biogenesis, acting as a transcriptional coactivator, recruiting and regulating the expression and function of multiple genes important in mitochondrial biogenesis, including nuclear respiratory factor (NRF)-1, NRF-2 and mitochondrial transcription factor A (Tfam) [60,61]. The nuclear abundance of PGC1α increases following endurance exercise [62] suggesting a role for PGC1α in increasing mitochondrial density following exercise. Ectopic PGC1α expression [60] has also been shown to increase mtDNA levels and mitochondrial biogenesis in skeletal muscle cells, as well as increase respiratory capacity and ATP synthesis [60]. It has been implicated in the conversion of glycolytic type II fibres to type I-like fibres, increasing the expression of proteins characteristic of type I fibres (e.g. myoglobin and troponin I (slow)) as well as a resistance to fatigue [63]. Russell et al. [64] showed that endurance exercise increases type I fibres by 18% with an accompanying decrease in type IIa and IIx fibres. As type I fibres have the highest resistance to fatigue, a greater percentage of type I fibres is beneficial for endurance and may improve age-related declines in muscle function. Endurance trained muscles have a more developed SR [65,66], with a decrease in calcium handling proteins such as SERCA [67], suggesting improved muscle fibre efficiency.

Resistance/strength training gives rise to a range of morphological changes in skeletal muscle. A major change is an increased capacity to generate force due to muscle hypertrophy [68]. Muscle hypertrophy occurs at a relatively slow rate, as protein synthesis must exceed protein degradation for an extended length of time before a net increase in contractile proteins is seen. The activation

and fusion of satellite cells to existing fibres provides extra myonuclei, aiding in the hypertrophic process. The major morphological changes are an increase in fibre cross sectional area (CSA), a change in the angle of pennation of individual fibres and an increase in non-contractile tissues such as collagen [58]. The degree of muscle fibre hypertrophy is dependent on $p70^{S6K}$ phosphorylation. $p70^{S6K}$ phosphorylation is dependent on mTOR, integrating metabolic and nutrient stimuli to regulate cell proliferation and growth [69]. Activation of this pathway via resistance training results in increased muscle protein translation, driving hypertrophy. Although satellite cells play a role in the regenerative capacity of muscle following injury, they do not appear to play a major role in contributing to muscle hypertrophy following exercise. In mouse skeletal muscle with 90% of satellite cells depleted, mechanical overload resulted in the same increase in muscle mass and muscle hypertrophy as wild-type muscle, however regeneration was significantly blunted following depletion, suggesting different roles of satellite cells [70].

1.1.5. Skeletal Muscle Atrophy

Skeletal muscle atrophy occurs as a consequence of disuse (e.g. immobilization [71], ageing [72,73]), and is characterized by a decrease in myofibre protein content, fibre size, force production and resistance to fatigue, as reviewed by S. Bodine [74].

There are various triggers and signalling molecules that play a role in skeletal muscle atrophy [75,76]. Myostatin (MSTN) is a key negative regulator of muscle growth [77]. Mouse knockouts of myostatin [78] exhibit enlarged muscles as a result of hypertrophy and hyperplasia. MSTN can also induce atrophy via inhibition of translation and protein synthesis. Myostatin expression is increased in some forms of atrophy. HIV positive men have been shown to have higher serum levels of MSTN [79], implicating MSTN in cachexia-type atrophy. However, mice knockouts of MSTN showing muscle hypertrophy have similar/slightly higher levels of muscle atrophy as wild-type mice in response to unloading [80], suggesting that MSTN may play a role in cachexia-type atrophy rather than disuse atrophy.

It is well established that muscle unloading and disuse results in decreased protein synthesis [76]. 4E-BP1 is a translation initiation factor that acts as a strong inhibitor of the eukaryotic initiation factor (eIF)-4E when unphosphorylated, halting translation and protein synthesis. 4E-BP1 expression is increased in disuse atrophy, together with an increased binding of 4E-BP1 to eIF-4E in unloaded muscle [69,76]. This leads to an inability to efficiently translate mRNA into proteins, resulting in reduced protein synthesis. Another key aspect of muscle atrophy is an increase in muscle protein breakdown. Myofibrillar protein constitutes over half of muscle protein content, and is lost at a fast rate during muscle atrophy [81], via the ubiquitin-proteasome system (UPS) and calpains. The activation of calpains by muscle unloading results in the breakdown of the myofibrillar structure.

Scaffold proteins such as titin, vinculin and nebulin are all known calpain substrates [82]. The release of myofilaments via the calpain system allows breakdown of the filaments via the UPS. The muscle-specific F-box containing ubiquitin protein ligase atrogin-1 and the ring finger-containing ligase MuRF1 exhibit increased expression during atrophy as a consequence of multiple upstream physiological signals including denervation [83], disuse [84], inflammation [85] and cachexia [86], suggesting a common set of pathways leading to enhanced proteolysis. Only a few substrates have been identified for atrogin-1, most of which are involved in growth and survival related processes, including MyoD [87] (a key regulator of muscle growth) and eIF3f [88] (an activator of protein synthesis). MuRF1 has been reported to degrade multiple structural proteins, including troponin I [89], MHCs [90], and myosin light chains [91]. Denervation-induced muscle atrophy is a key mechanism thought to underlie sarcopenia. Denervation-induced muscle atrophy induces the up-regulation of several muscle-specific E3 ubiquitin ligases. The up-regulation of these E3 ligases due to denervation leads to an increased breakdown of muscle structural proteins, leading to loss of muscle mass and subsequent function.

The constant up-regulation of E3 ubiquitin ligases, together with the inhibition of protein synthesis during ageing and as a result of pathological states, results in an inability to maintain muscle mass, leading a the atrophy of skeletal muscle fibres. This leads to the loss of total muscle mass, with muscle atrophy contributing to the reduction in muscle function.

1.1.6. Skeletal Muscle Repair

Muscle repair and regeneration following injury requires re-establishing cellular homeostasis and restoration of normal structure and functional capacity.

Breakdown in the integrity of the sarcolemma is a common feature of muscle injury, resulting in perturbations in calcium homeostasis. The mechanisms contributing to the repair of membrane integrity are not fully understood. Dysferlin, a protein found in the cytoplasm and cell surface of muscle fibres, is increased in damaged fibres [92]. Dysferlin may play a role in membrane sealing by mediating the fusion of cytosolic vesicles with the surface membrane, a process that is positively regulated by cytosolic calcium levels. This is relevant following muscle injury when cytosolic calcium levels are elevated due to dysregulated influx [93]. However, dysferlin-independent muscle resealing mechanisms exist, highlighting the need for further investigation.

Muscle injury also induces the activation of skeletal muscle satellite cells (SCs). SC activation following injury can occur via several mechanisms, as reviewed in [93,94]. Fibroblast growth factor 2 (FGF2) levels decrease in the cytosol following injury and induce SC activation and proliferation. FGF2 can also be found bound to heparin sulphate proteoglycans (HSPs) in the surrounding connective tissue [95]. The release of matrix metalloprotease 2 (MMP2) [96], which is expressed and

stored in skeletal muscle, can cleave HSPs and release FGF2, leading to SC activation. Skeletal muscle expresses neuronal NOS (nNOS) [97], which is activated by calcium. nNOS derived nitric oxide (NO) increases the activation and expression of MMP2, in turn, activating FGF2 and SC activation. MMP2 also causes the release of hepatocyte growth factor (HGF) from HSPs, which is expressed by mesenchymal cells and stored in the extracellular matrix [98]. HGF subsequently binds to c-met which is expressed by SCs, triggering activation and proliferation [98,99].

Skeletal muscle regeneration occurs in 3 stages [94]. Each stage is tightly regulated by muscle-specific transcriptional regulatory factors. The first stage is the proliferation of SCs, following activation, during which the levels of MyoD and Myf5 are elevated [100,101]. Myf5 is expressed in quiescent SCs but together with microRNA-31 which suppresses Myf5 translation, is sequestered in messenger ribonucleoprotein (mRNP) granules, inhibiting it's downstream pathway [102]. As SCs exit the cell cycle and enter the early differentiation stage, MyoD and Myf5 are activated, increasing the expression of muscle-specific genes, including myogenin (MYOG), Myf4 and the myocyte enhancer binding factor-2 (MEF2) family of transcription factors [100,103]. These TFs activate the transcription of genes important for the fusion of myogenic cells, forming multi-nucleated myotubes and transition to terminal differentiation. Muscle-specific genes continue to be highly expressed during terminal differentiation as myotubes grow and mature.

1.2. Sarcopenia and Ageing

Ageing is defined as the gradual, time-dependent decline in functional capabilities that affects most living organisms. In humans, ageing represents the accumulation of the molecular and cellular changes that occur with time that can increase the vulnerability and susceptibility to human disease. Multiple different processes are generally considered to contribute to the ageing process and have been termed 'the hallmarks of aging' [104].

Sarcopenia is defined as the loss of muscle mass and function (strength or performance) with age and is a core component of physical frailty in older people. The etymology is Greek in origin; *sarx* meaning flesh, and *penia,* which can be translated as loss, such that sarcopenia is the 'loss of flesh' and was first coined by Irwin Rosenberg in the late 1980s [105]. Sarcopenia is associated with multiple adverse outcomes such as loss of independence, disability, falls and fractures as well as admissions to nursing homes. As of September 2016, sarcopenia was awarded an ICD-10 code (M62.84) [106], recognising sarcopenia as a disease, as well as the impact it has on the ageing population.

1.2.1. Epidemiology

Sarcopenia affects both men and women. However, the severity and the prevalence between the sexes is variable, depending on developmental and adult influences. In their review, Cruz-Jentoft et

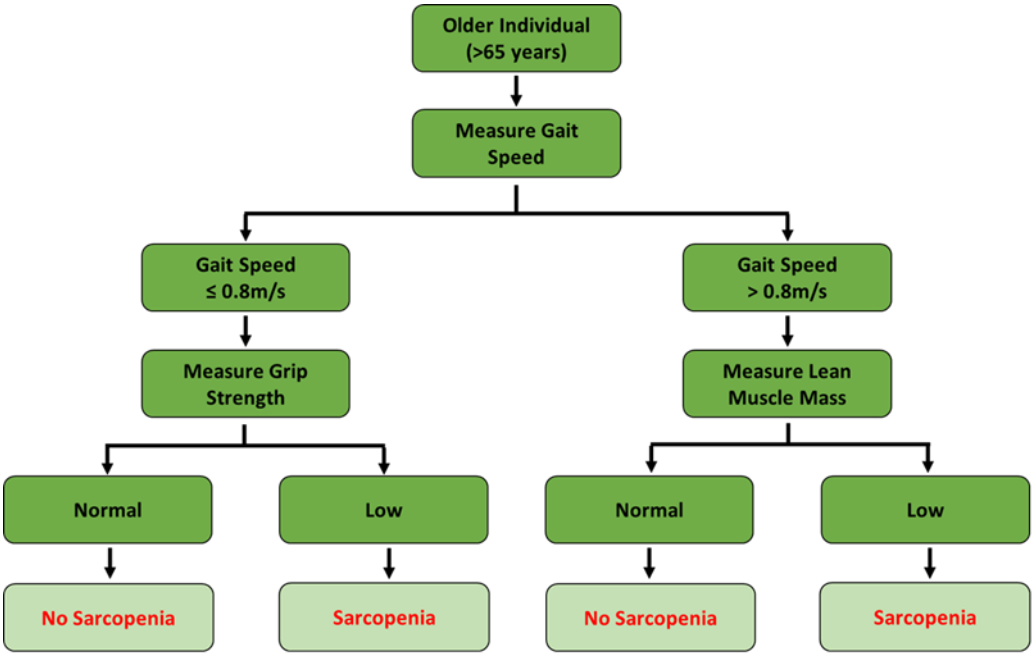


Figure 1.2: EWGSOP characterisation of sarcopenia

Flow diagram showing the basic definition proposed by the EWGSOP for classifying sarcopenia in the older population (>65 years), based firstly on gait speed, and subsequently on grip strength and lean muscle mass.

al. [107] described the prevalence of sarcopenia in community-dwelling populations. The prevalence ranged between 1% and 29% for community-dwelling elderly, and for those living in long-term care institutions, there was a much higher prevalence (14-33%), with a prevalence of up to 68% in men in certain studies. Using a definition of 2 standard deviations below a young reference group for appendicular skeletal mass (ASM)/height^2 but not measures of function, as a definition of sarcopenia, there was a prevalence of 13-24% in 65-70 year olds in a New Mexico cohort, increasing to over 50% in people >80 [108]. A similar study [109] used a cut-off based on skeletal muscle mass determined by dual energy X-ray absorptiometry (DXA), finding a prevalence of 10% and 8% for men and women respectively in people between 60 and 69 years of age, increasing to 40% and 18% for men and women respectively for individuals >80. The majority of studies analysed by Cruz-Jentoft et al. [107] suggest that there is an increased prevalence of sarcopenia with increasing age. Sex may potentially influence sarcopenia in the population, with some studies seeing a higher prevalence in men [110,111], while others in women [108,112]. Most studies however show no difference between the sexes. The prevalence of sarcopenia varies widely in the literature, and is most likely affected by the population studied, the reference population utilized and the operational definition of sarcopenia used.

Due to population and methodological differences, it is difficult to estimate a general prevalence of sarcopenia. However, on average, it is generally accepted by the scientific community that sarcopenia affects 5-13% of the elderly between 60-70 years of age, increasing to 11-50% in

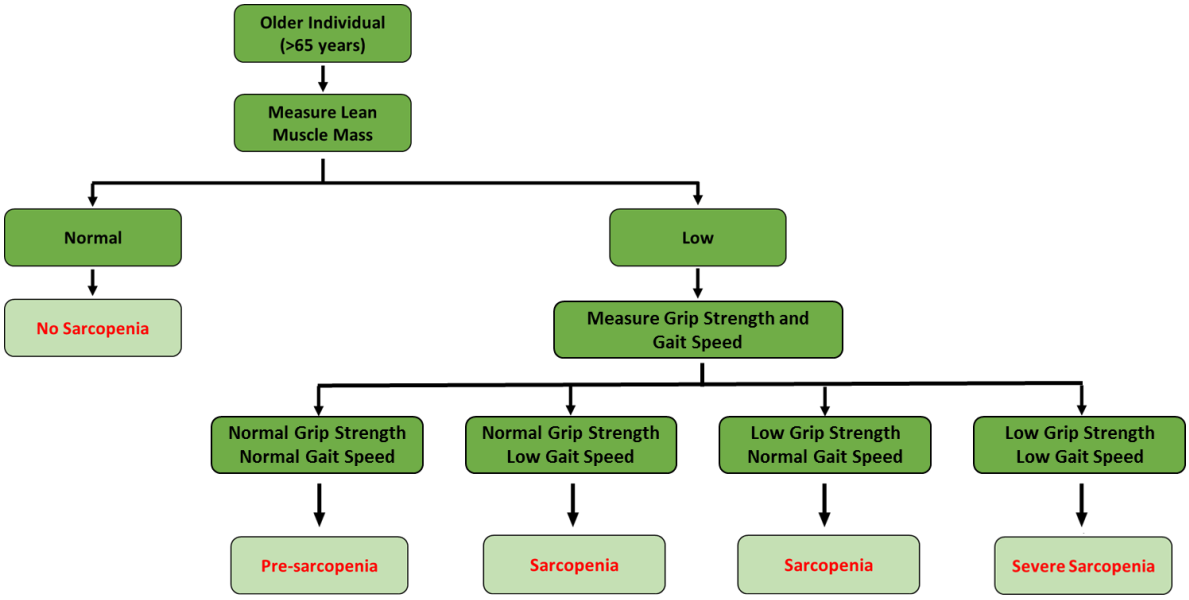


Figure 1.3: EWGSOP extended definition of sarcopenia

Flow diagram showing the extended definition proposed by the EWGSOP for classifying sarcopenia in the older population (>65 years), based firstly on lean muscle mass and subsequently grip strength and gait speed, providing a classification of no sarcopenia, pre-sarcopenia, sarcopenia or severe sarcopenia.

individuals >80 years in community-dwelling populations. The prevalence is considered to be higher in the elderly living in long-term care institutions.

1.2.2. Clinical Parameters for Diagnosis of Sarcopenia

Several working groups have been formed to aid the diagnosis of sarcopenia, due to the variability between populations and the need for ethnic considerations. These include the EWGSOP [113] (European Working Group on Sarcopenia in Older People), the IWGS [114] (International Working Group on Sarcopenia), the AWGS [115] (Asian Working Group on Sarcopenia) and the FNIH (Foundation for the National Institutes of Health) Sarcopenia Project [116]. Each adopt similar principles for sarcopenia case finding in different populations.

The EWGSOP was created to develop operational definitions and diagnostic criteria for use in the clinical setting as well as in research studies into sarcopenia [113]. It contains representatives from other European scientific organisations, including the ESPEN (European Society of Clinical Nutrition and Metabolism), the IANA (International Academy of Nutrition and Ageing), and IAGG-ER (International Association of Gerontology and Geriatrics – European Region). EWGSOP suggested an algorithm for defining sarcopenia in older individuals based on gait speed measurement, and subsequently lean muscle mass and grip strength (figure **1.2**). EWGSOP conceptually divided sarcopenia into sub-categories depending on the severity of the disease and clinical markers; pre-

sarcopenia, with only low lean mass but normal grip strength and gait speed; sarcopenia, with low lean mass and either low grip strength or low gait speed; severe sarcopenia, with low lean mass, low grip strength and low gait speed (figure **1.3**).

A wide range of methods can be used for determining skeletal muscle mass. For clinical practice, EWGSOP suggest the use of either DXA or bioimpedance analysis (BIA) for the measurement of lean muscle mass. As well as determining muscle mass, these techniques also allow the measurement of body fat in patients. Cut-off points provided by EWGSOP depend on one of these techniques being used for measurements. EWGSOP suggest the use of healthy young adults with a similar ethnicity and background as a reference group if possible, with a cut-off of two standard deviations below the mean reference value. This is due to the variability between populations, and as such, more research is needed to provide consensus reference values for different populations. There is no concrete cut-off for low lean mass, with 2 standard deviations below the average of a normative reference group being suggested, while some studies have used an arbitrary 20th percentile as a cut-off, which was shown to be associated with poor health and decreased activity and lower extremity activity [117].

As well as muscle mass, another important aspect of sarcopenia is a decrease in muscle performance, as measured by strength and function. These can be determined by analysing different measures of physical performance. Isometric grip strength is strongly associated with lower extremity power and a good clinical marker of poor mobility [118]. As such, it is a good simple measure of muscle strength, correlating well with leg strength [118]. The EWGSOP cut-off for defining low grip strength is <30kg for men and <20kg for women [113,118]. These can also be stratified based on BMI for males and females [119]. However, due to differences between ethnicities, different populations have different cut-offs for grip strength. For example, the AWGS suggest a cut-off of <26kg for men and <18kg for women [115]. Muscle function can be determined by a wide range of tests. Gait speed has been recognised as having a relationship with leg strength, providing a predictive value for the onset of disability. A commonly used cut-off for gait speed is 0.8 m/s over a 4-meter course [118]. As well as gait speed, several other measures of physical function exist including a timed up-and-go test, where individuals are required to stand up from a chair, walk a short distance, turn around, return and sit back down, which form part of the short physical performance battery (SPBB) [120], evaluating gait, strength and endurance.

Ultimately, for European and Caucasian populations, EWGSOP have developed an algorithm based on gait speed, grip strength and muscle mass measurements for classifying sarcopenia (figure **1.2**). A gait speed >0.8m/s identifies risk of sarcopenia. According to this algorithm, patients then have grip strength and muscle mass measurements to define sarcopenia status.

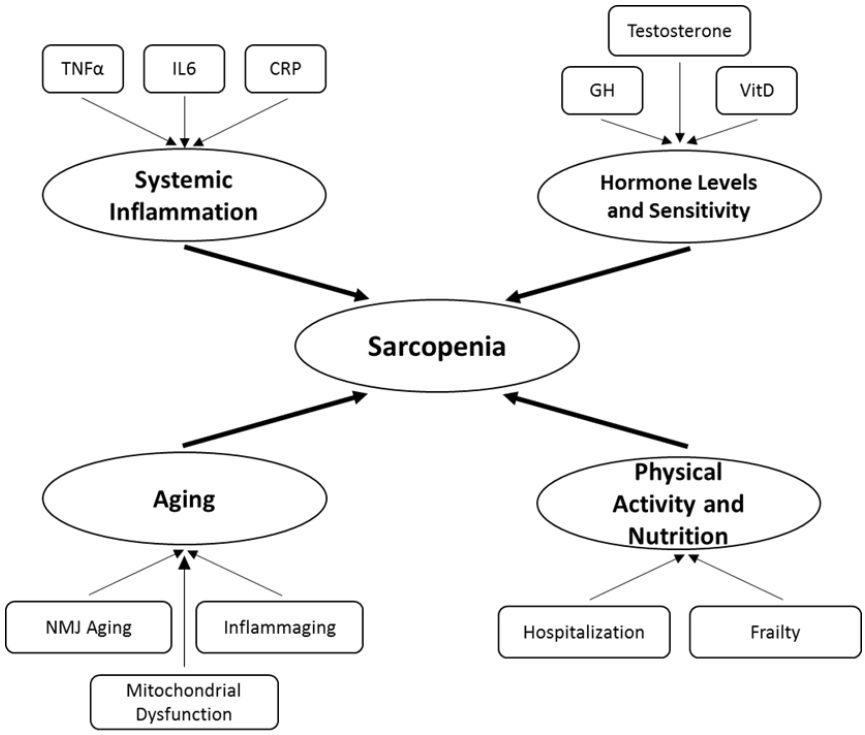


Figure 1.4: Main contributors to the pathogenesis of sarcopenia
Flow diagram of the main mechanisms thought to contribute to the onset and progression of skeletal muscle loss and loss of function associated with sarcopenia.

1.2.3. Pathophysiology of Sarcopenia

Significant changes occur in muscle mass and quality with ageing. Muscle is predominantly composed of two fibre types as described above. During low intensity activity, strength mainly comes from type I fibres, while during high intensity activity, strength comes from both types of fibres. With age, atrophy mainly affects type II fibres [121].

There are several mechanisms that have been proposed to be involved in the onset and progression of sarcopenia, most of which are not fully understood. The key mechanisms thought to contribute to the pathogenesis of sarcopenia are shown in figure **1.4** and described below.

1.2.3.1. Physical Activity

The progression of sarcopenia in the elderly may be accelerated by increased levels of physical inactivity. Hospitalization rates are much higher in the >60 age group compared to younger individuals [122,123], resulting in a higher incidence of bed rest and physical inactivity in the elderly. Extended bed rest poses a serious threat to muscle tissue and as such, limited physical activity during hospitalization has a range of negative outcomes. Loss of lean muscle tissue is inevitable during bed rest, predominantly affecting the muscles of the lower body [124-126]. Loss of lean tissue

during bed rest appears to be greater in the elderly compared to younger individuals experiencing bed rest [127].

Muscle strength can be improved by training and exercise, even in the elderly. An increase in exercise has previously been reported in the elderly to significantly increase muscle cross-sectional area (CSA) [128], as well as an increase in type I and II muscle fibre CSA and muscle thickness following resistance training [129]. These changes in muscle morphology are accompanied by a change in muscle function. There is an increase in stair walking power [129] as well as gait speed and muscle strength [128]. These changes are unrelated to age, sex or functional capability, suggesting a positive role for physical activity in maintaining and even increasing muscle mass and function in the elderly. Regular physical activity has previously been shown to be beneficial in reducing circulating inflammatory markers such as CRP [130] and IL-6 [131], with a reduction in inflammation and reduced levels of protein catabolism in skeletal muscle, maintaining muscle mass and structure.

The decrease in muscle mass seen with prolonged inactivity is primarily due to a reduction in protein synthesis [124,132]. Exercise is known to increase protein synthesis [126], contributing to the maintenance and increase of muscle mass in response to physical activity.

1.2.4. Skeletal Muscle Ageing

Skeletal muscle undergoes profound changes to its structure during ageing. After the age of 30, muscle mass in humans is lost at a rate of 0.5-1% per year, with a dramatic increase in the rate after the age of 65 [133]. This leads to an increased level of frailty as reflected by decreased physical activity and energy expenditure. This ultimately leads to muscle weakness, disability and metabolic disorders. The total decline in muscle mass between 40 and 80 years of age is estimated to be between 30 and 50% [134], with the most prominent decrease seen in the limbs [135,136].

Type II glycolytic fibres tend to be preferentially lost with age and sarcopenia [137,138]. This fibre loss is accompanied by expanded motor units [139], intramuscular infiltration of adipose and connective tissue [53,140] and decreased vascularization [141,142]. Together with a reduced type II fibre number, there is a reduction in type II fibre size, as evident by a decrease in cross-sectional area (CSA) [137]. This reduction in type II fibres and fibre atrophy may contribute to the increased level of fatigue, weakness and inactivity often seen in the elderly, key aspects of the frailty criteria [119], further reinforcing the frailty phenotype. The decrease in CSA results in decreased force production. Physical activity impacts the CSA of the remaining fibres, aiding in slowing down and preventing further atrophy. This however, mainly occurs in type I fibres [134,138].

The endocrine system plays a major role in regulating muscle mass [143]. Production of insulin-like growth factor 1 (IGF1) and many other anabolic molecules is reduced during ageing, resulting in a

decrease in myofibrillar protein synthesis. As a result, aged muscle becomes less responsive to anabolic stimuli, contributing further to muscle loss. With age, there is a decrease in the production of anabolic hormones (e.g. testosterone), which play a key role in regulating protein synthesis and in men, testosterone is a major anabolic stimulant. As well as a reduction in anabolic hormone production, there is increased anabolic resistance seen in the elderly, requiring increased anabolic stimulation in order to generate the same level of protein synthesis and muscle anabolism compared to young people. This combination of reduced anabolic hormone production and reduced sensitivity to anabolic stimulation results in reduced protein synthesis and muscle anabolism in the elderly. Another major pathway contributing to fibre loss is muscle denervation [134], due to the loss of motor units. This denervation results in atrophy of the fibres that do not become re-innervated by axonal sprouting from neighbouring motor units.

Several other explanations have been proposed to explain the age-related loss of muscle mass. These include increased apoptosis [144,145], mitochondrial dysfunction and oxidative stress [146-148], and reduced satellite cell content and function [149] resulting in reduced regenerative potential. However, the mechanisms initiating these processes in aged muscle is complex and remains a topic of debate.

1.2.5. Molecular Mechanisms of Sarcopenia and Ageing

The majority of the genes identified to play a role in aged skeletal muscle are involved in muscle growth. Insulin-like growth factor 1 (IGF1) [150] and ciliary neurotrophic factor receptor α (CNTFRα) [151] are decreased and myostatin (MSTN) [152] is increased in aged muscle. MSTN is an inhibitor of myogenesis and myogenic differentiation, an increase of which results in decreased myogenesis and an inability to repair muscle damage [77,78,153]. This continued inability to repair muscle damage reduces muscle mass. Older muscle is also associated with an increase in genes involved in inflammation and apoptosis, including C1Q-α [154] and FOXO3A [155,156]. FOXO3A is a transcription factor thought to initiate apoptosis by inducing the transcription of genes necessary for cell death [157]. A major pathway that appears in numerous studies comparing skeletal muscle of young vs old humans is mitochondrial structure and dysfunction [143,148,158,159]. Genes associated with these pathways show a consistent decrease in expression in aged muscle in humans, as well as other tissues (e.g. brain and kidney) and other species (e.g. mice and *Drosophila)* [160]. Decreased expression of these genes may result in an increase in reactive oxygen species (ROS) due to inefficient clearance and increased production as a result of mitochondrial dysfunction. ROS cause damage to the mitochondria and other cellular compartments [161-163], resulting in loss of activity and ultimately apoptosis due to irreparable damage.

As well as global gene expression analysis, studies have also looked at micro RNA (miRNA) expression in aged muscle. miRNAs are small non-coding RNAs involved in the post-transcriptional

regulation of gene expression, by binding to the 3'- or 5'-UTR of mRNAs and either inhibiting translation or increasing the rate of mRNA degradation [164]. miR-1 and miR-133 are two muscle-specific miRNAs shown to be involved in myogenesis, playing a vital role in myoblast-myotube transition [165]. miR-1 down-regulates the expression of histone deacetylase 4 (HDAC4), a known negative regulator of myogenesis [166]. miR-1 has been shown to be increased in skeletal muscle following endurance exercise in mice [167], and as such may contribute to the adaptive process of muscle to exercise and muscle regeneration. Similarly, miR-133 increases myoblast differentiation by repressing the transcription of serum response factor (SRF) [166]. miR-133 levels have been shown to be reduced in the muscle of older adults [168], implicating a reduction in miR-133 levels in reduced differentiation and muscle regeneration and repair with age. Several non-muscle-specific miRNAs, including miR-181, miR-26a and miR-424, have also been shown to play a role in skeletal muscle myogenesis [169]. miR-181 promotes differentiation by targeting the homeobox protein Hox-A11, alleviating the repression of the expression of MyoD, and has been shown to be increased after exercise in mice [167], suggesting a role for miR-181 in muscle regeneration and repair following exercise. miR-26a promotes myogenesis and myogenic differentiation by targeting Enhancer of Zeste homolog 2 (EZH2), a member of the polycomb repressive complex, which maintains the proliferative capacity of myoblasts and myogenic progenitor cells. miR-424 has been shown to play a role in regulating muscle differentiation by repressing the expression of CDC25A, the phosphatase responsible for removing the inhibitory phosphorylation marks of CDK2 and allowing cell cycle progression [170]. miR-424 was further shown to reduce rRNA and protein synthesis in muscle cells [171], with rRNA associating with ribosomes, contributing to normal functioning of the ribosome and protein synthesis [172]. miR-424 was shown to be associated with low muscle mass and poorer physical performance in several independent groups of older people, including patients with COPD, community-dwelling elderly with sarcopenia and patients in the intensive care unit [171].

Therefore, there is preliminary evidence to suggest a role for specific miRNAs in sarcopenia and in regulating muscle mass, although further work is required to determine how different miRNAs function together to regulate muscle mass and function in the elderly.

1.2.5.1. Cellular Senescence

Cellular senescence is the state of irreversible cell cycle arrest that can be triggered by a variety of different stimuli, including potentially oncogenic stimuli, and as such can act to supress tumorigenesis. Senescence helps to maintain cellular homeostasis by preventing further growth and proliferation of cells exhibiting DNA damage and genomic instability [173], acting as a protective mechanism. However, it has recently been shown that senescent cells are capable of secreting cytokines and remodelling factors that may have deleterious effects on tissue homeostasis [174,175],

contributing to chronic inflammation and cancer progression. Therefore, cellular senescence plays two important but differing roles.

1.2.5.1.1. Mechanisms of Cellular Senescence

1.2.5.1.1.1. p16-Rb Pathway

The cell cycle is a tightly regulated process [176]. In order for a cell to progress through the cell cycle, it must pass through four major checkpoints: G1-S transition, S phase checkpoint, G2 to M transition and the mitotic spindle checkpoint. The $p16^{INK4a}$-Rb pathway controls the G1-S transition as reviewed by Sharpless et al. [177]. $p16^{INK4a}$ binds to the cyclin dependent kinase (CDK)-4/6, inhibiting the kinase activity and preventing phosphorylation of the retinoblastoma protein (pRb). In its unphosphorylated state, Rb remains bound to the transcription factor E2F1, localizing it to the cytoplasm and preventing its translocation into the nucleus and the transcriptional activation of E2F1 target genes crucial for G1-S transition (e.g. c-Myc). The $p16^{INK4a}$-Rb pathway has also been shown to cooperate with the mitogenic signalling cascade for ROS induction, in turn activating protein kinase C delta (PKCδ) and increasing ROS production [178], resulting in a positive feedback loop and cell cycle arrest. Therefore, $p16^{INK4a}$ plays a major role in initiating and maintaining cellular senescence and the transcriptional regulation of $p16^{INK4a}$ likely plays a major role in initiating senescence pathways.

Expression of $p16^{INK4a}$ is regulated by the hypermethylation of its promoter via components of the polycomb repressive complex (PRC)-1 and PRC2 [179,180]. Bmi1 is a polycomb protein, part of the PRC1 complex, and inhibits senescence by repressing the transcription of $p16^{INK4a}$. Bmi1 catalyses the ubiquitination of histone H2A on lysine 119 (H2AK199ub), leading to chromatin compaction and gene silencing [181]. The catalytic component of the PRC2 complex is EZH2, mediating the di- and trimethylation of lysine 27 on histone H3 (H3K27me2/3) [182], compacting the chromatin. H3K27me3 is recognized by the CBX family member of the PRC1 complex [183], recruiting it to the $p16^{INK4a}$ promoter. As such, the H2AK199ub mark and subsequent gene silencing by PRC1 is dependent on EZH2 activity at the $p16^{INK4a}$ promoter.

To induce cellular senescence, the repressive histone marks at the $p16^{INK4a}$ promoter must be removed [179]. In response to oncogenic stressors, the expression level of the H3K27 demethylase JMJD3 increases [184], removing the H3K27me3 mark from the $p16^{INK4a}$ promoter. Following demethylation, the Jun dimerization protein 2 (JDP2) binds to and sequesters H3K27 from PRC2, maintaining the $p16^{INK4a}$ promoter in an active state. Unmethylated H3K27 is not recognized by the PRC1, resulting in de-compaction of the chromatin around the $p16^{INK4a}$ promoter and subsequent $p16^{INK4a}$ expression.

1.2.5.1.1.2. p53-p21-Rb Pathway

p53 is a tetrameric transcription factor that is highly regulated by posttranslational modifications [185]. p53 is viewed as one of the most powerful tumour suppressive genes owing to its ability to halt cellular progression and its ability to induce apoptosis.

Following activation of the DNA damage response (DDR), the induction of p53 is important for the initiation of cellular senescence [186,187]. DNA damage appears at the telomeres of cells during replicative senescence, with many DNA damage-response proteins found associated with these regions, such as 53BP1 and RAD17 [188]. Senescent cells also express an activated form of Ataxia-telangiectasia mutated (ATM). ATM phosphorylates p53 on serine-15 and serine-20 is phosphorylated by ATM's downstream target CHK2. These modifications prevent the degradation of p53 by the E3 ubiquitin ligase, MDM2 [189]. The DDR also induces the expression and activation of ATM [190], preventing the degradation of p53 and increased expression of its downstream targets.

The p21 gene is a target of p53, and initiates growth arrest and cellular senescence by preventing the phosphorylation of Rb or by binding to and sequestering E2F transcription factors [191]. In senescent cells, increased p53 activity is accompanied by an increase binding of p53 to the p21 gene promoter, activating its transcription. p21 inhibits the activity of CDK1 and CDK2 [192], which are required for the phosphorylation of Rb protein and the subsequent release and activation of E2F-dependent gene expression. p21 can also bind directly to E2F1[193], suppressing its transcriptional activity and repressing the expression of genes required for transition through the G1-S checkpoint.

1.2.5.1.2. Cellular Senescence during Ageing

Many tissues, including skeletal muscle, accommodate a population of cells capable of dividing upon stimulation [9]. Most of these mitotic cells are found in a reversible quiescent state, during which the cell cycle is arrested. However, upon stimulation, these quiescent cells are activated and begin to proliferate. Ageing of tissues is thought to be predominantly due to the accumulation of senescent cells [194,195], which are irreversibly growth arrested. Senescent cells display a different phenotype compared to their quiescent counterparts, showing distinct genetic, morphological and behavioural characteristics.

Replicative senescence is the major mechanism for the accumulation of senescent cells with age. Telomeres are repetitive stretches of DNA found at the end of the chromosomes that act to protect the DNA ends from degradation and recombination during DNA replication. However, due to the inability of the replication machinery to copy the extreme ends of chromosomes, telomeres become shorter with each replication cycle. Eventually, when the telomere length reaches below a critical limit, cells activate the DDR and p53 pathway [188,196], inducing cellular senescence and/or apoptosis. Hayflick and Moorhead [197] first showed that cultured cells *in vitro* had a limited capacity

to divide, after which they stop growing, and was termed replicative senescence. In the 1990s, it was found that telomere shortening is the causal link that triggers replicative senescence [198], as a method of protecting the cell from continuing to divide with DNA damage to the ends of the chromosomes. Evaluation of telomere length in older individuals suggest that those with shorter telomeres have a higher all-cause mortality rate than those individuals with longer telomeres, as well as increased mortality due to diseases/disorders other than cardiovascular disease and cancer.

Telomerase is responsible for maintaining telomere length. However its expression in adult cells is too low to compensate for the level of telomere attrition during DNA replication, resulting in a gradual loss with age [199]. Therefore, the rate of telomere shortening may indicate the rate of ageing [200] of a tissue. Many studies have shown an inverse relationship between telomere length and ageing and age-associated diseases in a variety of tissues [201,202]. A variety of premature ageing diseases such as Werner syndrome and dyskeratosis congenital are characterised by mutations in genes associated with telomere length maintenance and regulation, resulting in faster telomere attrition with age [203]. Ageing is also associated with an accumulation of ROS which has been shown to accelerate telomere shortening [204].

Mice deficient in telomerase have short telomeres and exhibit premature ageing. The first generation of these telomerase-deficient mice show a reduced lifespan, which gets shorter with each subsequent generation combined with an increased incidence of pathologies associated with decreased telomere length and age such as heart failure and decreased tissue regeneration [203]. This is reflected by an increase in p53 expression, implicating the activation of the DDR and cellular senescence.

$p16^{INK4a}$ also plays a major role in mediating cellular senescence. The expression of $p16^{INK4a}$ increases markedly with ageing in the majority of tissues [205,206]. As such, $p16^{INK4a}$ expression may provide a marker of physiological age. The mechanisms for an increase in $p16^{INK4a}$ expression are not fully understood. ROS may play a part in increasing $p16^{INK4a}$ expression via the p38MAPK pathway, acting to reduce the lifespan of tissue-resident stem cells [207]. The levels of Bmi1 have been shown to decrease with age in humans [205], removing its repression on $p16^{INK4a}$ expression. Therefore, $p16^{INK4a}$ may mediate the increase in senescence associated with ageing. $p16^{INK4a}$ is found in the *INK4a/ARF* locus on chromosome 9p21.3 in humans and this locus has consistently been found to be strongly associated with age-related pathologies in many different genome-wide association studies (GWAS), including cardiovascular disease and diabetes [208].

1.2.5.1.3. Cellular Senescence in Skeletal Muscle

Regeneration of skeletal muscle due to injury and damage from normal activity and exercise is dependent on a population of Pax7+ satellite cells (SCs) [209]. These SCs are adult stem cells that exist

in a reversible quiescent state throughout life [9], and are only activated into a proliferative state upon stimulation, to differentiate or to self-renew and maintain the pool of SCs.

In very old sarcopenic mice, there is a significant decrease in the regenerative capacity of skeletal muscle compared to non-sarcopenic old mice [194]. In the old mice, there is increased fibre atrophy, accompanied by a decrease in fibre innervation and central nucleation of fibres, correlating with decreased function. There is a decrease in satellite cell number in old mice compared to adult mice [194]. Older age also produces intrinsic changes in satellite cell regenerative function, which cannot be overcome by a young host environment, as demonstrated by hetero-grafting experiments between old and young mice [194], suggesting that it is not the aged environment, but intrinsic changes producing a reduction in the regenerative capacity of satellite cells.

$p16^{INK4a}$ is a key regulator of cellular senescence and is upregulated in satellite cells from older mice [194]. H2AK199ub is the major repressive mark at the $p16^{INK4a}$ promoter, regulating its expression under normal circumstances [180]. This mark is reduced in the quiescent satellite cells from old mice, suggesting that this de-repression of the *INK4A/ARF* locus, from which $p16^{INK4a}$ is transcribed [177], induces an increase in $p16^{INK4a}$ expression and induction of senescence in quiescent satellite cells. This results in a decrease in the self-renewal capacity of the satellite cells, depleting the SC pool. Knocking-down $p16^{INK4a}$ expression has been shown to restore satellite cells activation and reduce satellite cell senescence in old mice [194]. $p16^{INK4a}$ knock-in in young mice induces early senescence and reduces the activation of satellite cells in response to injury [194]. $p16^{INK4a}$ acts upstream of E2F and pRb [210]. Phosphorylated Rb levels are decreased in old mice, as well as decreased levels of E2F-regulated genes such as cyclin A and Cdc6 [194], which are involved in cell cycle progression. This suggests that the $p16^{INK4a}$-E2F-pRb axis may be a causal factor for the reduced regenerative capacity seen in aged muscle.

1.2.5.1.4. Senescence-Associated Secretory Phenotype

Although in a cell-cycle arrested state, senescent cells remain metabolically active and experience many changes in protein expression and secretion, termed the senescence-associated secretory phenotype (SASP) [174]. The SASP explains how senescent cells can alter the microenvironment of tissues, resulting in multiple different pathologies due to the secreted pro-inflammatory, growth-promoting and remodelling factors. These secreted factors can be divided into 3 main groups: soluble signalling factors (e.g. interleukins, chemokines, growth factors, etc.), proteases and insoluble components of the extracellular matrix (ECM).

IL-6 is the main proinflammatory cytokine secreted by senescent cells. IL-6 acts on nearby cells expressing the IL-6 receptor such as endothelial cells, initiating an inflammatory response, as reviewed by Tanaka et al. [211]. IL-6 expression is associated with DNA damage- and replication-

induced cellular senescence [174]. IL-1β is another proinflammatory cytokine reported to show increased secretion from senescent cells. Upon activation of IL-1R or Toll-like receptor family members, IL-1β induces a cascade of downstream events, resulting in the increased transcription and activation of NF-κB and the JNK and p38 mitogen-activated protein kinase pathways, potently activating inflammatory pathways, as reviewed by Weber et al. [212]. Other soluble factors secreted include chemokines, such as IL-8 (CXCL8), CXCL1, macrophage inflammatory protein 3α and eotaxin. These chemokines are involved in the chemotaxis of leukocytes and other inflammatory cells to the site of secretion [213]. This aggregation of pro-inflammatory cells results in a localized inflammatory response. The secretion of several anti-inflammatory cytokines such as IL-2, IL-10 and IL-12 remains unchanged in senescent cells. This results in an increased ratio of pro-inflammatory cytokines compared to anti-inflammatory cytokines, promoting a proinflammatory phenotype.

As well as inflammatory factors, senescent cells secrete extracellular proteases such as matrix metalloproteases (MMPs) and serine proteases [174,214]. MMP1 and MMP3 show increased secretion by senescent cells. These MMPs cleave and activate some of the soluble inflammatory factors secreted by senescent cells, such as IL-8. Members of the chemokine family can also be cleaved and activated by MMPs. Therefore, secreted proteases favour a proinflammatory phenotype. As well as cleaving and activating inflammatory factors, MMPs cleave most of the components of the ECM. For example, collagen I, II and III are broken down by the collagenases (e.g. MMP1), E-cadherin by MMP3, while laminin, fibronectin and elastin can be degraded by most of the MMPs [215]. Senescent cells also secrete insoluble molecules such as fibronectin into the ECM [216]. Fibronectin can affect cell adhesion, cell survival and cell growth, and its production is increased in premature ageing diseases [217]. This breakdown of the ECM and secretion of fibronectin leads to a breakdown in tissue integrity with altered tissue architecture, leading to loss of function.

The effect of the SASP on normal tissue homeostasis highlights the importance of cellular senescence in ageing, inducing a variety of molecular and phenotypic changes. These changes in skeletal muscle may cause alterations in the normal structure and function of skeletal muscle, leading to reduced muscle function. This results in a vicious cycle ultimately leading to muscle disuse and an inability to efficiently utilize the muscle.

1.2.5.2. Mitochondrial Dysfunction and the Mitochondrial Theory of Ageing

The increased mitochondrial dysfunction seen during ageing results in an increased production of reactive oxygen species (ROS), resulting in further mitochondrial deterioration and cellular damage [104,143]. There is also reduced mitochondrial biogenesis and bioenergetics with age, contributing to the cellular decline associated with ageing [218]. Mutations in mtDNA contribute to mitochondrial dysfunction due to the synthesis of truncated and non-functional proteins of the mitochondrial

respiratory complexes [20,161]. The combination of increased mitochondrial damage and reduced turnover may contribute to the ageing process. The mitochondrial theory of ageing is discussed further below.

Mitochondria are the powerhouses of the cell, responsible for the production of ATP. Every cell contains hundreds of mitochondria constantly producing ATP. ROS are highly reactive molecules generated as by-products of normal cellular oxidative processes. Mitochondria are the main producers of cellular ROS as a by-product of oxidative phosphorylation and aerobic respiration and are therefore highly susceptible to oxidative damage [161]. The majority of mitochondrial ROS are generated at the electron transport chain (ETC), with electrons leaking from the ETC and interacting directly with free oxygen to produce highly reactive free radicals (e.g. superoxide anion (O_2^-)). O_2^- radicals are converted to hydrogen peroxide either spontaneously or via superoxide dismutase (SOD) enzymes [219], which is stable and subsequently gets removed by enzymes in the cytoplasm.

As mitochondria are the main producers of ROS in the cell, this makes mtDNA very prone to oxidative damage due to the close proximity to ROS production and accumulation. The main DNA lesion caused by ROS is 7,8-dihydro-8-oxo-deoxyguanosine (8-oxo-dG), which is highly mutagenic, resulting in a G:C to T:A transversion [220]. It has previously been shown that 8-oxo-dG is much more abundant in mtDNA than nuclear DNA [221], suggesting increased susceptibility of mtDNA to ROS damage. Although cells have evolved defence mechanisms against this (e.g. base excision repair), these mechanisms become less efficient with age [222]. As mtDNA encodes essential components of the ETC, oxidative phosphorylation and mitochondrial protein synthesis, mutations in mtDNA may impair the assembly and function of the respiratory complexes, triggering mitochondrial dysfunction. This leads to a vicious positive feedback cycle, leading to energy depletion and an inability to maintain metabolic demands.

mtDNA is densely packed with genes, containing only one non-coding region, the displacement loop (D-loop), important for replication and transcription of mtDNA [223]. Age-dependent mutations in the D-loop have been reported in multiple tissues, including skeletal muscle [224]. The accumulation of oxidative damage in mtDNA leads to double-stranded breaks. During subsequent repair, intact mtDNA is produced, harbouring a deletion. Deletions of mtDNA are also more frequent in aged tissues in humans. The most common deletion is a 4977bp deletion, found with increased frequency in skeletal muscle with age [158]. This region encodes subunits of the NADH dehydrogenase, cytochrome *c* oxidase and ATP synthase. This deletion therefore impairs the normal functioning of these complexes and the ETC, contributing to the vicious cycle of mitochondrial failure. However, whether such deletions are causative in the development of ageing phenotypes remains to be determined.

Mitochondria play a key role in regulating apoptosis and this process has been implicated in ageing [225]. Oxidative damage in the mitochondria may contribute to an increase in cellular apoptosis [226]. The activation of the permeability transition pore in mitochondria is enhanced in several tissues of aged mice. This pore plays an important role in regulating the initiation of apoptotic pathways [227]. Components of this pore are also susceptible to oxidative damage, as shown in aged flies, suggesting oxidative damage in the mitochondria may modify the regulation of apoptosis initiation [228].

1.2.5.3. Inflammaging

Many aspects of the immune system decline with age, including its reliability and efficiency, increasing an individual's susceptibility to disease [229]. The term 'Inflammaging' has been coined to describe the accumulation of an individual's exposure to antigenic load over their lifetime, caused by both clinical and subclinical infections, and non-infective antigens. The resulting inflammatory response generates ROS and results in tissue damage. This initiates a positive feedback mechanism, in which immune remodelling occurs favouring a chronic proinflammatory phenotype.

During normal ageing, the accumulation of dysfunctional cells leads to the secretion of proinflammatory cytokines and systemic inflammation. In addition to cellular dysfunction, there is an increased skeletal muscle adipose infiltration with age, together with an increased age-related risk of obesity [230]. Adipose tissue is known to release tumour necrosis factor α (TNFα) into the circulation [231]. The age-related increase in adipose tissue may result in an increased secretion of the proinflammatory cytokine TNFα into the circulation, further exacerbating the systemic inflammatory phenotype. Older individuals have repeatedly been shown to have elevated serum levels of interleukin (IL)-6, TNFα and C-reactive protein (CRP) [232]. Increased serum IL-6 [233,234], TNFα [235] and CRP [234] have all been associated with an increased risk of frailty, lower physical performance as measured by the ADL (Activities of Daily Living) score, lower cognitive ability as measured by the MMSE (Mini Mental State Examination) score and increased mortality risk. These factors tend to be secreted by cells and act locally but their presence in the serum suggests that they are being secreted into the circulation, inducing a systemic inflammatory response.

Skeletal muscle is responsive to circulating cytokines. Circulating levels of tumour necrosis factor-alpha (TNFα), interleukin (IL)-6, IL-1 and C-reactive protein (CRP) are increased in older people [235]. IL-1, IL-6 and TNFα exhibit catabolic effects on muscle [236]. Patients with high levels of TNFα have been shown to have a smaller muscle area, low appendicular lean mass (ALM), knee extensor strength and grip strength compared to individuals with lower circulating levels of TNFα [235]. The Longitudinal Ageing Study of Amsterdam showed that increased levels of IL-6 and CRP are associated with an increased decline of muscle mass, although no association with sarcopenia was seen [237]. TNFα and other pro-inflammatory cytokines are potent activators of proteolysis, activating

the UPS (ubiquitin-proteasome system) pathway [238], contributing to muscle breakdown. The Health ABC Study has been used to study the effect of pro-inflammatory cytokines on muscle mass and function [235]. Those with elevated levels of IL-6 and TNFα had decreased muscle mass and muscle strength. This association was found in both well-functioning elderly men and women. This relationship was present after adjusting the data for several inflammation-related disorders, including heart disease, lung disease, osteoarthritis and diabetes. This suggests that there is a direct relationship between the levels of IL-6 and TNFα and muscle mass and strength in the elderly. Moreover, elderly people with elevated levels of IL-6 and CRP were shown to be at increased risk of losing >40% of their muscle strength [237]. However, the underlying molecular mechanisms leading from inflammation to muscle loss and functional decline is not clear.

There are several factors contributing to inflammaging in the older population. One source may be the age-associated accumulation of damaged proteins due to increased damage stimuli and reduced clearance, activating the innate immune response. Several components of damaged cells can activate the innate immune response, including ROS, ATP and ceramides, molecules collectively known as danger/damage signals [239,240]. As the level of damage increases, these signals also increase, resulting in a chronic activation of the innate system with mitochondria playing a key role in inflammaging [239]. The Nlrp3 inflammasome is a multi-protein complex capable of activating pro-caspase 1, initiating the apoptotic response in response to cellular danger. Activators of the Nlrp3 inflammasome also induce ROS production, resulting in oxidative damage. This releases mitochondrial damage-associated molecular patterns (DAMPs, e.g. mtDNA) into the circulation, which are potent activators of the innate immunity and the Nlrp3 inflammasome [240]. Inflammaging can also be caused by cellular senescence. Senescent cells have a secretory phenotype, the senescence-associated secretory phenotype (SASP), which results in the secretion of multiple pro-inflammatory cytokines [174,175]. As the number of senescent cells increases with age, there is an increase in the secretion of pro-inflammatory cytokines, providing a driving factor for inflammaging. Changes to the immune system with age also contribute to inflammaging (immunosenescence) [241]. Immunosenescence refers to the deterioration of the immune response with age, resulting in an impairment of the capacity to respond to infective challenges and the development of long-term immune memory. Hematopoietic stem cells are precursors for cells of the immune system. Their ability to self-renew diminishes, resulting in a reduced pool from which to produce new immune cells. This is accompanied by a reduction in antibody producing B cells, decreasing antibody diversity and affinity. The decrease in adaptive immunity may result in an increase in innate immunity as a compensatory mechanism, increasing the inflammatory phenotype.

With age, there is an increase in the serum levels of cortisol and an associated decrease in the level of DHEAS (dihydroepiandrosterone (DHEA) sulphate) [242]. Cortisol has an immunosuppressive function while DHEAS is immunomodulatory, counteracting the effects of cortisol by antagonism of

glucocorticoid receptors [242]. Therefore, the increase in cortisol levels may be a method in order to counteract the increase inflammation seen during ageing. However, an increase in serum cortisol levels, although acting as an anti-inflammaging mechanism, can have negative consequences, including global immunosuppression as well as frailty due to muscle catabolism.

1.2.5.4. Neuromuscular Ageing

Neurodegeneration is a heterogeneous process that increases with age. Multiple levels of the nervous system are affected, including the motor cortex, spinal cord, peripheral neurones and neuromuscular junctions [243,244].

In the spinal cord, the number of alpha motorneurones (αMNs) declines with age [243]. A motor unit comprises a single αMN and all the muscle fibres connected to it. When an αMN is lost, an adjacent nerve sprouts a projection to innervate the existing denervated muscle fibres creating larger, less efficient motor units. The number of muscle fibres connected to a single motorneurone depends on the function of the muscle. If the muscle is used for precise movements, such as muscles of the eye, a single motorneurone is connected to only a few individual fibres. Muscles requiring less precision and more strength are connected to more fibres. This means that surviving motor units become bigger after denervation, as a single αMN would have to service many more muscle fibres. The motor neurones that re-innervate denervated fast-twitch fibres may come from slow-twitch fibres, which are slower, resulting in less force being generated [24]. This process contributes to decreased efficiency of the motor units that can result in decreased muscular function and muscle fatigue, commonly seen in older people. It has also been suggested that there is a preferential decrease in αMNs supplying fast motor units. Others have also reported a loss of peripheral neurones and changes in their myelin insulation [245]. Age-related changes have also been seen at the neuromuscular junctions (NMJs), with a reduced number synaptic vesicles in the nerve terminal, together with a reduced nerve terminal area [246]. It has also been shown that the compensatory re-innervation of denervated fibres also decreases with age [247]. Taken together, these suggest a neuropathic process resulting in reduced muscle fibre efficiency and force production.

1.2.5.5. Hormone Levels and Sensitivity

As humans age, there are changes in the levels and activities of several hormones important in normal physiology. Ageing is associated with a decreased level of testosterone, an anabolic steroid hormone, known to increase muscle mass and protein synthesis [248,249]. Hence, testosterone has been tested as a treatment for sarcopenia. Increasing the level of testosterone in elderly men to that of levels seen in young men results in an increase in muscle mass, but does not result in functional gain in muscle strength [249]. Administration of testosterone to a level higher than the physiological level has been shown to increase muscle strength as well as muscle mass , however

the risks of such levels may outweigh the benefits [250], with an increased risk of thrombotic complications and the stimulation of previously unrecognised local or metastatic prostate cancer.

With ageing, there is also a decrease in circulating growth hormone (GH) levels. GH is released from the pituitary gland and promotes the secretion of insulin-like growth factor (IGF)-1. IGF1 activates the Akt/mammalian target of rapamycin (mTOR) pathway, inducing muscle hypertrophy by activating protein synthesis [251]. Akt also inhibits the up-regulation of E3 ubiquitin ligases such as the muscle specific RING-finger protein-1 (MuRF1) and atrogin-1 [252], which stimulate the breakdown of muscle protein. A decrease in GH and IGF1 is frequently seen in the elderly [253], accompanied by reduced lean muscle mass and bone mineral density, and increased visceral fat content. Therefore, exogenous GH administration has been trialled as a therapeutic approach. Although, this increases lean muscle mass, it does not lead to functional muscle gains or beneficial metabolic changes. Moreover, side effects were frequently reported and outweigh the benefits of GH treatment [254].

1,25-dihydroxyvitamin D_3 (1,25(OH)2D) is the active form of vitamin D, binding to the vitamin D receptor (VDR) and activating it. VDR is a transcription factor, regulating the expression of multiple genes involved in calcium handling, and muscle cell differentiation and proliferation [255]. Low 25-hydroxyvitamin D (25(OH)D) has been associated with an increased risk of sarcopenia in the elderly, as well as positively associated with skeletal indices [256]. Given the positive effect vitamin D supplementation has on bone function [257], further work is required to determine whether vitamin D supplementation may have a beneficial effect of maintaining muscle mass and preventing further muscle loss.

1.2.5.6. Genomic Instability

Nuclear DNA accumulates somatic mutations over time [104], resulting in inefficient replication and mutated/truncated protein production. Other types of DNA damage found to be associated with ageing include chromosomal aneuploidy and copy number variation. These changes may occur in important genes and pathways (e.g. genes involved in cellular metabolism), leading to dysfunctional cells. This results in increased apoptosis and cellular senescence, becoming detrimental to tissue homeostasis. Premature ageing diseases such as Werner syndrome, are the result of increased levels of DNA damage accumulation [222] [258].

DNA damage can also occur in mitochondrial DNA (mtDNA), as mtDNA is highly prone to ageing-associated mutations due to the highly oxidative environment of the mitochondria. Mice deficient for mtDNA polymerase exhibit premature ageing, with an accompanying impairment in mitochondrial function. This combined effect of mitochondrial dysfunction and mtDNA damage leads to inefficient energy production ultimately affecting tissue function during ageing.

1.2.5.7. Epigenetic Alterations

Histone methylation may play a role in ageing [104]. Deletion of components of the histone methylation complexes (for histone H3, lysine 27 (H3K27)) extends longevity in flies. Inhibition of histone demethylases for H3K27 in worms may also increase longevity. Mammalian paralogs of the *S. cerevisiae* gene *Sir2* can ameliorate aspects of ageing in mice. SIRT6 regulates genomic instability and glucose homeostasis through histone H3K9 deacetlyation, with mice deficient in SIRT6 showing accelerated ageing [259]. The contribution of histone modifications to ageing is complex, and not fully understood.

Ageing is also associated with an increase in transcriptional noise, with aberrant production of many mRNAs [260]. These ageing-associated transcriptional changes also affect non-coding RNAs (ncRNAs) and miRNAs that are associated with the ageing process, targeting components of the longevity network and stem cell activity [166,168,261]. As ncRNAs and miRNAs generally play a role transcriptional regulation, aberrant expression of these RNAs can affect the expression of genes important for maintaining tissue homeostasis, resulting in cellular dysfunction.

1.2.5.8. Loss of Proteostasis

The stress-induced synthesis of cytosolic chaperone proteins is significantly impaired during ageing [104], resulting in decreased chaperone-mediated protein folding and stability. Transgenic flies over-expressing chaperone proteins have an increased lifespan [262], while mice deficient in chaperones show an accelerated ageing phenotype [263]. Activation of HSF1, a transcription factor regulating the heat shock response also increases longevity, suggesting an important role for proteostasis and chaperones during ageing.

With ageing there is also a decrease in the ubiquitin-proteasome system (UPS) and autophagy pathways, the two major pathways for protein quality control [264]. A reduction in these pathways can lead to impaired clearance of misfolded and damaged proteins, causing a build-up of toxic protein aggregates in cells, leading to cellular dysfunction. A decrease in the UPS is seen in the majority of tissues. However, the described effects of ageing on the UPS in skeletal muscle remain inconsistent in the literature, with considerable variability in the observed changes in proteasome function during ageing in skeletal muscle [265,266]. Despite this this, Hepple et al. [267] show an increase in proteasome activity with age, as measured by the chymotrypsin-like activity of the proteasome. Further work is required to confidently say that there is an increase in proteasome activity in skeletal muscle with age, contributing to the reductions in muscle mass.

Alongside the uncertain evidence as to whether or not there is a loss of proteostasis as a result of reduced UPS function in skeletal muscle, there exist other mechanisms of muscle catabolism that

may contribute to the loss of muscle mass seen with ageing. These include the lysosomal system (autophagy), the caspase system and the calpain system. The autophagy pathway has been shown to be activated when the UPS fails, acting as a compensatory mechanism [268]. Increases in these pathways may counter any decrease in the UPS in skeletal muscle, contributing to increase muscle atrophy and loss with age.

1.2.5.9. Deregulated Nutrient Sensing

Genetic analysis has shown that the insulin and IGF1 signalling (IIS) pathway mediates part of the beneficial effects of dietary restriction (DR) on longevity [104]. However, despite a decrease in IIS being beneficial for longevity, IGF1 levels are also decreased during normal ageing. Decreased IIS is common to both physiological and accelerated ageing, while constitutively decreased IIS extends longevity. As such, organisms with constitutively decreased IIS survive longer due to slower rate of cell growth and cell damage. Therefore, decreased IIS in physiologically aged organisms may be an attempt to reduce cell growth and cell damage in order to extend longevity.

The mTOR pathway is another major pathway regulating metabolism. Downregulation of mTORC1 in yeast extends longevity and attenuates any further benefits of DR, suggesting that mTORC1 silencing mimics the effects of DR [269]. mTOR activity increases during ageing in the mouse hypothalamus, contributing to age-related obesity. This suggests that increased stimulation and anabolic activity, signalled via mTOR, contributes to accelerated ageing. Further work is required to determine whether the role of mTOR in ageing is also seen in humans, although strong evidence for the role of mTOR in longevity in mice and yeast suggest it may also play a role in humans.

1.2.5.10. Stem Cell Exhaustion

Tissue regeneration failure is commonly seen in ageing, with a decrease in stem cell number and potential [270]. DNA damage occurs constitutively in all cell types, including stem cells, as a by-product of normal cellular metabolism and replication failures. Replicative and oxidative stress during replication over time leads to a slow attrition of stem cells during adulthood, ultimately degrading tissue homeostasis. Therefore, this loss of stem cell function over time may be sufficient to cause aspects of normal ageing due to the prevention of tissue renewal.

1.2.6. The Skeletal Muscle Transcriptome during Ageing

Recently, the skeletal muscle transcriptome has been studied using microarrays and RNA-seq to define a global baseline transcriptome [271], and to look as the effects of ageing [261,272], exercise [272-274] and multiple disease states, including type II diabetes [273,275] and muscular dystrophy [276].

Lindholm et al. [271] have provided an extensive investigation into the transcriptional profile of human skeletal muscle. It was shown that the gene expression of muscle obtained from different sites and from a similar spatial volume are highly similar in their global transcription profile. Skeletal muscle is highly plastic at all stages of life, changing in response to external stimuli within a relatively short time period [277-279]. Studies have analysed the effect of repeated temporal sampling, looking at the expression of a few select genes [280,281]. Lindholm et al. [271] have however shown that at a global level, if samples are taken within 15 minutes of each other, there is no difference between the transcriptome profiles of the muscle taken from the same or both legs of an individual.

The skeletal muscle transcriptome differs between males and females [271,282], showing a clear sexual dimorphism with respect to muscle mass, size of individual fibres [283], expression of myosin isoforms [283], fatigability [284] and lipid content [285]. This difference is partly driven by testosterone, due to its anabolic effects [286]. In males, the increase in muscle mass occurs during the post-pubertal stage after an increase in testosterone levels. Testosterone is required to maintain the increased muscle mass in males. Testosterone exerts its anabolic effects by influencing gene expression [287]. However, it is not known which genes are controlled by testosterone or which other factors contribute to the sexual dimorphism. Welle et al. [282] and Lindholm et al. [271] used microarrays and RNA-seq respectively to show clear sex differences in the transcriptome of skeletal muscle. Both found >3000 sex-related differences in gene expression. Male profiles are enriched for many glycolytic mRNAs, including lactate dehydrogenase (LDHA), phosphofructokinase muscle (PFKM) and aldolase A (ALDOA), while female profiles were enriched for oxidative mRNAs including peroxisome proliferator-activated receptor gamma coactivator 1 α (PGC1α) and citrate synthase.

1.2.6.1. Skeletal Muscle Transcriptome during Ageing

Ageing is associated with changes in gene expression, and understanding the changes in gene expression that occur with age is vital for understanding sarcopenia and in developing interventions for sarcopenia. The majority of studies have looked at genome-wide transcriptome changes between young and old skeletal muscle to determine age-related changes in expression [73,160,168,261,272,288-291]. However, the high inter-individual variability in skeletal muscle loss in older people has not been examined in detail.

Zahn et al. [160] and Su et al. [272] have utilized microarrays to analyse the genome-wide transcriptome changes that occur with age. Both studies show an increased enrichment of genes differentially expressed in inflammatory pathways, indicating increased inflammation in aged skeletal muscle. Su et al. [272] showed changes in several pathways associated with mitochondrial function and metabolism, with decreased expression of genes associated with the ETC, the TCA cycle and pyruvate metabolism. Decreases in these pathways with age suggest a significant reduction in the

ability to produce energy, with ATP being key for normal muscle function. Mitochondrial dysfunction has repeatedly been implicated in ageing [146,147,159,161]. Zahn et al. [160] suggest a 250-gene signature associated with skeletal muscle ageing, and show that expression of the genes in this signature also correlate with physiological age of skeletal muscle and not just chronological age. This suggests that these genes may explain some of the inter-individual variability in muscle mass and strength observed in the elderly. miRNAs play a key role in regulating gene expression and may underlie some of the age-related changes of gene expression in skeletal muscle. Zacharewicz et al. [261] compared the miRNA profile of skeletal muscle between young and old individuals and identified seven miRNAs that were associated with age (miR-146a-5p, miR-191-5p, miR-320a, miR-483-5p, miR-486-3p, miR-539-5p and miR-628-5p), irrespective of exercise intervention. Four of these seven (miR-191-5p, miR-320a, miR-483-3p and miR-628-5p) were found to be significantly altered between young and old at baseline, before exercise intervention. The pathways predicted to be regulated by these miRNAs included cellular function, development and cell death and survival. Several of these miRNAs also have predicted targets involved in respiration and muscle function. For example, phosphofructokinase is a potential target of miR-320a, resulting in reduced glycolysis in the elderly; and miR-86-3p has several potential targets within the muscle protein synthesis pathway (e.g. RPTOR and TSC-1), and is positively regulated by MRTF-1 and SRF, suggesting attenuated activity of the MRTF-A/SRF pathway in skeletal muscle with age [292].

These studies suggest an interplay between miRNA and mRNA changes during ageing in skeletal muscle may play a key role in regulating skeletal muscle mass and function in the elderly, underlying the pathogenesis of sarcopenia.

1.2.7. Genetic Mechanisms of Sarcopenia

The genetics of muscle mass and strength, the two most commonly recognised risk phenotypes of sarcopenia, is complicated by many confounding factors, including enormous inter-individual variability in muscle mass and strength, the heterogeneity of muscle tissue and the low magnitude of changes generally seen in gene expression studies in muscle. Studies have suggested a 50-80% genetic contribution to lean muscle mass in humans as determined by DXA [293,294], as well as a 30-85% heritability (h^2) for muscle strength.

Specific to older people, Prior et al. [295] have reported that the h^2 of appendicular lean mass (ALM) is 20% in people aged >45 years. A genetic component for quadriceps strength (14%) and knee extension strength (31%) has also been reported, highlighting a genetic component to muscle strength [296]. However, the contribution of environmental factors may also increase with age. Overall, genetic variation may partly explain the inter-individual variation in skeletal muscle

phenotype, although this is complex and makes understanding the molecular aspects of sarcopenia more challenging.

Recently, several genome-wide association studies (GWAS) have identified loci associated with muscle mass and strength. Carola Zillikens et al. [297] carried out a meta-analysis and identified five single-nucleotide polymorphisms (SNPs) associated with lean body mass, including rs2943656, located near the insulin receptor substrate 1 (IRS1) gene and rs9936385, located near the fat and obesity associated (FTO) gene. rs2943656 was also associated with appendicular lean mass. Matteini et al. [298] found further association between rs752045 and hand grip strength. This SNP was found to be located in a DNase I hypersensitivity site within a binding motif for the CCAAT/enhancer-bidning protein beta (CEBPB), which has been implicated in regulating genes associated with inflammation [299], and as such may play a role in muscle repair.

1.2.8. Current Treatments for Sarcopenia

Resistance exercise has been shown to improve muscle mass and strength in the elderly [128], accompanied by an improvement in mobility. Regular aerobic exercise may ameliorate sarcopenic symptoms by increasing skeletal muscle mitochondria and improving insulin sensitivity [300]. However, regular exercise is not always feasible in the elderly.

In addition to physical activity, pharmacological intervention may be considered and is a major area of research. Several drugs have been tested for their effects on muscle mass and strength, but at present, there is little evidence justifying their use in sarcopenia. Testosterone is a steroid hormone with anabolic effects. Some studies have implicated testosterone supplementation in increasing muscle mass [249,301] and grip strength [302]. However, the results in the literature are inconsistent. In young people, testosterone administration results in an increase in muscle mass but not strength. Whereas in older people, testosterone does result in some improvement in muscle strength [249,301,302]. However, testosterone intake has multiple secondary effects including aggressive behaviour, thrombosis, oedema and prostate cancer, contraindicating its use. Growth hormone (GH) supplementation has been shown to improve overall body composition by increasing lean muscle mass and decreasing fat mass [254]. However, there is strong evidence to show that GH does not increase muscle strength or overall functional capacity accompanied by severe side effects (e.g. oedema, diabetes, and joint pain).

Vitamin D levels decrease with age, with extremely low levels seen in the elderly (<30 ng/ml). Low serum vitamin D is associated with sarcopenia and sarcopenic obesity [256]. The effectiveness of vitamin D supplementation for sarcopenia remains controversial, despite several studies showing an improvement in muscle strength following supplementation [303,304]. Vitamin D supplementation

is beneficial for bone function in osteoporosis [257] and therefore may be useful for treating the symptoms of sarcopenia. As a result, further work is needed to determine the effectiveness of vitamin D supplementation in sarcopenia.

Recently, it has been suggested that angiotensin converting enzyme (ACE) inhibitors (ACEi) may be beneficial in the treatment and management of sarcopenia. ACEi work by reducing the activity of the renin-angiotension-aldosterone system by preventing the conversion of angiotensin I to angiotensin II. They have previously been shown to improve endothelial function and angiogenesis while reducing inflammation [305]. This may contribute to improved skeletal muscle blood flow, and as a result improving muscle function. ACEi have also been shown to protect against mitochondrial dysfunction in skeletal muscle [306]. There is some evidence for ACEi use in sarcopenia. In the Health ABC study, participants using ACEi had higher lower extremity muscle mass compared to participants using other antihypertensive drugs [307], while in the Women's Health and Ageing Study, use of ACEi was associated with lower decline of muscle strength after 3 years [308].

1.3. Epigenetics and Regulation of Gene Expression

Epigenetics refers to the heritable modifications in gene expression and gene activity that is not attributable to any changes in the genomic sequence itself. Epigenetics plays a pivotal role in tissues and has been implicated in many human diseases. Human cells all contain an identical genome; however, their phenotypic characteristics and traits differ, allowing for specialized functions to be carried out by a subset of cells. However, as cells all have an identical genome, a mechanism must exist to allow the expression of appropriate genes in the appropriate cells at the appropriate time, to allow cells to carry out their specialized functions.

Epigenetics plays an important role in the interplay between genetics and the environment and contributes to the long lasting effects on the environment on an individual across generations. External stressors in the environment have repeatedly been shown alter an individual's epigenome and DNA methylation pattern. The Dutch Winter Hunger of 1944/45 resulted in a shortage of food, affecting the whole population. Since then, the long-term effect of this on the offspring were investigated, showing changes in DNA methylation associated with the famine, contributing to increased risk of cardiovascular disease and schizophrenia associated with alterations in DNA methylation [309,310]. Stress is another environmental stressor shown to affect DNA methylation. Post-traumatic stress disorder and childhood abuse have been associated with altered DNA methylation [311,312]. Stress during pregnancy is associated with increased risk of psychiatric disorders in the offspring thought to be due to altered regulation of the glucocorticoid receptor gene [313]. These studies and many others have implicated DNA methylation and alterations in the epigenetic regulation of gene expression in many human diseases, with the environment and developmental stressors resulting in an altered epigenome. Changes in DNA methylation have also been shown to

be associated with many age-related diseases, including cancer [314], diabetes [315] and Alzheimer's disease [316], highlighting the importance of understanding changes in DNA methylation.

There are multiple mechanisms of epigenetic control of gene expression. The main mechanisms are DNA methylation, histone modifications and non-coding RNAs/microRNAs.

1.3.1. DNA Methylation

DNA methylation is a chemical modification to the DNA, involving the covalent addition of a methyl ($-CH_3$) moiety to cytosine residues, generating a 5-methylcytosine residue. Methylated cytosine residues are generally found in the dinucleotide 5'-CpG-3', with the *p* being the phosphate bond linking the cytosine and guanine residues, although non-CpG methylation does exist, primarily present in pluripotent stem cells and a select subset of differentiated cell types [317,318]. Most CpGs in the human genome are methylated. Unmethylated CpGs tend to be found in clusters, known as CpG islands, rather than being randomly distributed. CpG islands are generally located in gene promoters and are unmethylated in promoters of ubiquitously expressed genes. In cancer, methylation of CpG islands leads to the repression of tumour suppressor gene expression [319,320], leading to the hypothesis that DNA methylation may play an important role in regulating gene expression.

The majority of DNA methylation marks are created during embryogenesis and development, which are maintained with subsequent cycles of DNA replication [321]. Shortly after fertilisation, the paternal pronucleus undergoes rapid genome-wide DNA demethylation, occurring within 4-8 hours after fertilisation via an active demethylation process. Demethylation of the maternal genome is a much slower process, occurring via passive demethylation. These processes lead to a highly demethylated genome except for at certain loci such as imprinted regions and some transposable elements (TEs). Following this period of demethylation, *de novo* methylation commences. *De novo* methylation is carried out by members of the DNA methyltransferase 3 (DNMT3) family [322]. DNMT3A and DNMT3B are catalytically active, while DNMT3L lacks the catalytic domain and acts as a cofactor for DNMT3A and DNMT3B. The majority of *de novo* DNA methylation occurs during the early developmental stage, with the progressive accumulation of developmental methylation at key promoters driving cell fate during early development. As development is a highly dynamic period, changes in the environment can affect developmental processes and have been shown to impact on DNA methylation in the developing cells [323,324].

Once established during development, DNA methylation patterns must be stably maintained to preserve cell type identity and maintain the epigenome [321]. DNA methylation is maintained by DNA methyltransferase 1 (DNMT1). During DNA replication, DNA becomes hemimethylated, with the

parent strand maintaining methylation marks while the newly synthesised strand is unmethylated. DNMT1 is recruited to the replication foci via interactions with proliferating cell nuclear antigen (PCNA) and other chromatin associated factors (e.g. ubiquitin-like PHD and RING finger domain 1 (UHRF1)) [325]. DNMT1 is expressed during S phase of the cell cycle, when it is required to methylate hemimethylated DNA. DNMT1 knockout studies have shown that without DNMT1, the levels of hemimethylated DNA increases and cells undergo apoptosis, highlighting the importance of DNMT1 for maintaining DNA methylation marks [326,327].

DNA methylation is generally associated with the repression of gene expression. Two models have been described to explain how DNA methylation results in gene silencing. Firstly, methylated CpGs can directly prevent the binding of transcription factors (TFs) to their consensus sequences due to steric clashes between the TF and the methyl group. Secondly, methylated CpGs are recognized by and bind to methyl-CpG binding proteins (MBPs), recruiting co-repressors to silence gene expression [328]. MBPs recognise and bind to methyl-CpGs via methyl-CpG binding domains (MBDs). MBD proteins (MBD1-4) and MeCP2 (methyl-CpG binding protein 2) specifically recognise methylated DNA. These proteins target chromatin-remodelling complexes to regions of methylated DNA. MBD1 recruits the histone H3 lysine 9 (H3K9) methyltransferase SETDB1, resulting in H3K9 methylation, a repressive mark. MeCP2 silences gene expression by recruiting the Sin3a/HDAC co-repressor complex, resulting in the removal of acetyl groups from histones, compacting nucleosomes and silencing gene expression. MeCP2 can also increase the H3K9 methylation levels at promoters [329], further reinforcing the silencing of gene expression. This suggests an important interplay between DNA methylation and other forms of epigenetic control of gene expression.

As well as regulating gene expression, DNA methylation plays an important role in genomic imprinting and X-chromosome inactivation. These processes involve mono-allelic gene expression, which in genomic imprinting is dictated by parental origin, whereas X-inactivation is random. DNA methylation is very important for maintenance of gene silencing on the inactivated X chromosome (Xi). The use of the DNA demethylating agent 5-azacytidine on human Xi-containing cell hybrids results in the reactivation of genes on the X chromosome, with a greater effect seen upon mutation of the DNMT1 gene [330,331]. Gene imprinting is the mechanism by which certain genes are expressed in a parent-of-origin-specific manner. For example, H19 is only expressed from the maternally-inherited allele, and the paternal allele is silenced. Mouse embryos homozygous for a mutation in the DNMT gene die before birth, with a substantially demethylated genome [332]. The expression of H19 was found to be increased in homozygous (DNMT-/-) embryos, compared to wild-type embryos, due to a reactivation of the paternal allele. This suggests that DNA methylation plays a major role in maintaining genomic imprints, although it may not be the only mechanism controlling genomic imprinting.

1.3.2. Histone Modifications

In the nucleus, DNA is organised into nucleosomes, allowing compaction of the genome into the small area of the nucleus. DNA is wrapped around histone protein complexes. Each histone complex is composed of two H3-H4 dimers, forming a tetramer, flanked by two H2A-H2B dimers. H1 linker histone protects the internucleosomal linker DNA. Histones undergo a variety of post-translational modifications, affecting histone function and the ability of DNA to be efficiently transcribed. The two histone modifications that play a key role in regulating gene expression and chromatin structure are histone acetylation and histone methylation [333-335]. Other histone modifications do exist, including histone ubiquitination, sumoylation and phosphorylation.

1.3.2.1. Histone Acetylation

Histones can be acetylated on lysine residues found on the N-terminal tails of histones. The acetylation of lysine is a highly dynamic process, regulated by histone acetyltransferases (HATs) and histone deacetylases (HDACs). HATs transfer an acetyl group from acetyl CoA, neutralizing the positive charge on the lysine, weakening the interaction between the histone and the DNA. As such, histone acetylation results in a loosening of the nucleosome and opening of the chromatin structure. Acetylation has also been shown to alter histone-histone interactions between nearby nucleosomes, as well as interactions with regulatory proteins. All these changes lead to an open chromatin structure, aiding in binding of TFs and transcriptional activation. Multiple HATs exist, including GCN5 and CBP/p300, which are recruited to histones via DNA-bound transcriptional activators [334]. For example, CBP/p300 can be recruited by interaction with DNA-bound TFs such as phosphorylated CREB [336,337].

HDACs antagonise the activity of HATs, removing acetyl groups from lysine residues. This returns the positive charge to the lysine, allowing a tighter interaction between the histone and DNA, creating a closed chromatin structure. Therefore, HDAC activity represses transcription. There are 4 classes of HDACs: class I, II, III and IV. Class III HDACs are known as the SIRT proteins, requiring NAD^+ for its activity. Class IV only contains 1 member, HDAC11 [333]. Class II HDACs have been shown to play an important role in skeletal muscle development. HDAC4 and 5 repress the expression of DACH2, a negative regulator of myogenin [83], regulating denervation-induced skeletal muscle atrophy. HDAC9 [338]/HDAC4 [339] can also interact with MEF2 proteins, repressing their transcriptional activity and inhibiting skeletal muscle development.

1.3.2.2. Histone Methylation

Histones can be methylated on lysine and arginine resides. Lysine can be mono-, di- or tri-methylated, while arginine can be mono- or symmetrically or asymmetrically di-methylated [335].

Multiple histone lysine methyltransferases (HKMTs) have been discovered, all with a SET domain containing the catalytic activity. HKMTs methylate lysine residues on the N-terminal tails of histones . Protein arginine N-methyltransferase (PRMT) proteins methylate histones on arginine residues. Both families of methylases transfer a methyl group from S-adenosyl methionine (SAM) to the appropriate residue. The most common lysine methylation sites are H3K4, H3K9, H3K27, H3K36 and H4K20. The sites of arginine methylation include H3R2, H3R26 and H4R3. These methylation marks can be reversed by histone demethylases: amine oxidases and jumonji C (JmjC)-domain containing, iron-dependent dioxygenases [184]. There are multiple mechanisms by which these methylases are recruited to histones. Long non-coding RNAs (lncRNAs) have been shown to be able to bind to polycomb complex proteins (PRC2), resulting in H3K27 trimethylation (H3K27me3). One such lncRNA is HOTAIR, leading to repression of specific loci [340,341]. The location of the methylated lysine residue and the degree of methylation are associated with differential gene status. For example, H3K4me3 is associated with active transcription, while H3K27me3 is associated with closed chromatin and repressed transcription. H3K4me1 is linked to enhancer function, while H3K4me3 is linked with promoter activity.

Therefore, histone methylation plays a role in many different aspects of transcriptional regulation from chromatin structure to the recruitment of TFs and the regulation of gene expression, as well as interaction with transcriptional initiation and elongation factors and RNA processing.

1.3.3. MicroRNAs

Micro RNAs (miRNAs) are small RNAs, 21-25 nucleotides long, regulating gene expression post-transcriptionally. They are non-coding RNAs, derived from larger precursors that are processed to produce the mature miRNA by two ribonuclease III enzymes. Drosha cleaves primary miRNA (pri-miRNA) into precursor miRNA (pre-miRNA), and Dicer cleaves pre-miRNA into a dsRNA duplex, which gives rise to the mature miRNA [164]. Most miRNAs target the 3'-UTR of mRNA molecules with imperfect complementarity. Mature miRNAs are incorporated into the RNA-induced silencing complex (RISC), which is required for target recognition and subsequent translational repression.

Translational repression is the main mechanism by which miRNAs negatively regulate the expression of their target genes [342]. However, the mechanism by which miRNAs repress translation has not been fully determined. Repression can occur at the initiation stage due to interference with the recognition of the 5'-cap structure on mature mRNAs. This is supported by the fact that IRES-dependent translation is not susceptible to miRNA-mediated repression. Members of the Argonaute (Ago) protein family can bind to mature miRNAs [343]. Ago2 has been shown to bind directly to the cap structure and can compete with eIF4E, inhibiting translational initiation [343]. Ago proteins can also recruit translational repressors, blocking ribosomal scanning and other aspects of

translation. miRNAs can also repress translational elongation, however the mechanisms for this mode of repression are unknown. Lin-4 miRNA represses the translation of lin-14 and lin-28 in *C. elegans* during development, as is evident by reduced protein levels. However, the mRNAs of these genes remain associated with translating polysomes, suggesting elongation repression. As well as translational repression, miRNAs increase mRNA degradation [344]. mRNA degradation by miRNAs is carried out by Ago proteins, primarily Ago2, which interacts with GW182 protein. In doing so, deadenylase complexes are recruited. Following deadenylation, the mRNA is targeted to the 5'-3' mRNA decay pathway, leading to the degradation of target mRNAs and consequently, decreased protein levels.

Several muscle-specific miRNAs (myomiRs) have been discovered that play a key role in skeletal muscle development and function. These myomiRs are miR-1, miR-133 and miR-206. miR-133 plays a role in the early differentiation of myogenic satellite cells into myoblasts and proliferation of myoblasts [166], as well as playing a role in maintaining muscle and muscle regeneration, by repressing the expression of serum response factor (SRF). miR-1 promotes myogenesis by targeting HDAC4 [166], a known negative regulator of myogenesis, via the HDAC4-myogenin pathway. miR-206 is also able to promote satellite cell differentiation and myogenesis by repressing Pax7 expression and the expression of several negative regulators of myogenesis [345]. These roles for the myomiRs during muscle development suggests an important role for miRNAs in skeletal muscle.

1.3.4. The Skeletal Muscle Epigenome

There is limited information on the skeletal muscle epigenome and changes associated with DNA methylation and histone modifications with age in skeletal muscle and during skeletal muscle development.

1.3.4.1. DNA Methylation during Skeletal Muscle Development

Various studies have investigated the role of DNA methylation during skeletal muscle development and the differentiation of myoblasts into myotubes. At a global level, there is an increase in DNA methylation in myogenic progenitor cells (MPCs) compared to embryonic stem cells (ESCs), followed by a gradual hypomethylation during myogenic differentiation. Compared to MPCs and differentiating myotubes, mature muscle tissue shows a loss of the hypermethylated sites. Carrió et al. [346] further showed that although there is little difference between the DNA methylation profiles of myoblasts and myotubes, there is a loss of methylation associated with terminal differentiation and in mature muscle. A dynamic gain of DNA methylation was observed at selected regions in the genome, while there was a significant hypomethylation during muscle differentiation. Many of the demethylated sites were found to be associated with homeobox and Tbox

transcription factor binding sites [347]. The demethylase TET2 (ten-eleven translocase 2) was found to be important in regulating the expression of myogenic genes. TET2 is involved in the demethylation of the myogenin and myomaker gene promoters, two genes that play a key role in the initiation of myogenic differentiation, concomitantly increasing the expression of these genes [348].

In contrast, Miyata et al. [349] found a slight global hypermethylation as cells progressed from myoblasts to myotubes, with a hypermethylation at the promoter of genes associated with muscle contraction and other muscle-related processes. The hypermethylated regions were further enriched for two binding motifs recognized by ID4 [350] and ZNF238 [351], key TFs regulating myoblast differentiation. It was found that there was a hypermethylation of the MYF6 promoter, the expression of which was down-regulated in terminally differentiated myotubes.

Overall, it was suggested that there is a period of *de novo* DNA methylation before the myoblast stage, with a subsequent hypomethylation during myogenic differentiation.

1.3.4.2. DNA Methylation during Skeletal Muscle Ageing

Changes in DNA methylation have been associated with ageing in several different tissues and cell types [352-354]. Few studies have investigated the effects of ageing on DNA methylation in skeletal muscle tissue. Zykovich et al. [355] examined DNA methylation changes in skeletal muscle, and observed a predominant hypermethylation throughout the genome of the aged muscle compared to the young samples. The differentially methylated CpGs (dmCpGs) between the young and old skeletal muscle were over-represented in the intragenic and 3' regions of genes and were under-represented at the 5' regions of genes. The dmCpGs identified were associated with pathways involved in muscle function, including muscle cell differentiation and axon guidance, suggesting that DNA methylation changes may be altering the expression of genes in these pathways in aged muscle. On the other hand, Jin et al. [356] found a global hypomethylation of the skeletal muscle genome, enriched in the gene-body regions, between young and middle-aged pigs as a model for human ageing. Several differentially methylated regions (DMRs) were identified associated with ageing, which were enriched in pathways associated with protein catabolism and proteolysis. There were also DMRs associated with genes that are potentially involved in the human ageing process (according to the Human Ageing Genomic Resources database), including FoxO3 and FGFR1. FoxO3 activates autophagy and proteosomal pathways in muscle while FGFR1 inhibits atrophy of skeletal muscle. FoxO3 was upregulated in aged muscle with lower gene-body methylation, while FGFR1 was downregulated in aged muscle with increased gene-body methylation. Jin et al. [356] showed increased protein catabolism and atrophy associated with skeletal muscle ageing, driven by changes in DNA methylation.

However, caution has to be taken in interpreting DNA methylation changes associated with age in whole muscle tissue samples, as tissue samples contain cell types of non-myogenic origin, which contribute to the methylation profile. As DNA methylation is important in maintain cell identity, it may be important to look at the effect of ageing on isolated muscle cells (e.g. satellite cells, myoblasts, and myofibres) and determine whether these changes are similar to those seen in the whole tissue samples.

1.3.4.3. Histone Modification Changes during Skeletal Muscle Ageing

Histone modifications, in particular histone acetylation and methylation, play a role in regulating gene expression during the activation and differentiation of skeletal muscle satellite cells. Liu et al. [357] have shown that upon activation of quiescent satellite cells, there is an increase in the global levels of H3K27me3, with the genes marked by H3K27me3 showing a reduction in expression in activated satellite cells compared to quiescent satellite cells. The increase in H3K27me3 at the transcription start sites (TSS) of genes suggests a rapid inhibition of quiescent cell-specific genes upon activation, allowing the proliferation and differentiation of satellite cells. This increase in H3K27me3 is correlated with an increase in the H3K27 methyltransferase EZH2, and decrease in the demethylase Jmjd3. With age, there was an increase in the H3K27me3 histone mark in quiescent satellite cells, with more than half of the peaks identified in intergenic regions. There was also an increase in H3K27me3 around TSSs. This increase may reflect a loss of transcriptional potential of quiescent satellite cells with age and may contribute to the functional decline of satellite cells commonly seen with age.

As well as changes in histone methylation, Asp et al. [358] have reported changes in histone acetylation playing a role in myogenic differentiation. Myotubes show a complete loss of H3K9 and H4K12 acetylation compared to myoblasts. These marks are found associated with active promoters, and the loss of these marks in myotubes suggests that genes maintaining myoblasts in a proliferative state are switched off in myotubes and during differentiation. Cao et al. [359] reported an increase in H4 acetylation at regions of DNA sequence to which MyoD binds. This suggests that MyoD is able to alter the epigenome and chromatin structure of specific regions of the genome in skeletal muscle. However, the biological role of this is still unknown.

1.4. Aims

Sarcopenia is becoming a major burden on the healthcare system as the general population lives longer and the elderly population increases. Therefore, as the prevalence of muscle loss increases, an increased incidence of comorbidities is expected, including obesity, frailty, metabolic disorders

and ultimate loss of independence. Sarcopenia has only recently been officially recognised as a disease, highlighting the importance of sarcopenia in the population and the need to understand the pathophysiology of sarcopenia, in order to develop treatment and interventions.

Decreased muscle mass and function is likely to be an inevitable consequence of ageing. However, the degree to which muscle is lost with age varies between individuals. Many suggestions have been given to explain the increased muscle loss seen with age, including mitochondrial dysfunction, reduction in satellite cell number and function, inflammation and neuromuscular dysfunction. Muscle mass is determined by a synergy between protein synthesis and protein breakdown. However, how these pathways contribute to the inter-individual variability remains relatively unknown.

Epigenetics have been previously shown to play a key role in regulating normal muscle function, as well as satellite cell function during myogenesis. However, more investigations are required to understand the role of epigenetics during skeletal muscle ageing, and to determine how epigenetic processes may contribute to decreased muscle mass in the elderly. Epigenetics, in particular DNA methylation, have been used as biomarkers for risk of disease in later life. Therefore, understanding DNA methylation marks and their association with muscle mass in the elderly may provide biomarkers that predict risk of sarcopenia in the elderly.

The aims of this project are to:

- Identify transcriptomic changes in aged skeletal muscle associated with skeletal muscle mass and function.
- Identify individual pathways that may contribute to skeletal muscle ageing, determining how these pathways are regulated and affect skeletal muscle.
- Identify whether myoblasts derived from human muscle biopsies provide a good model to investigate skeletal muscle ageing and sarcopenia.
- Identify differentially methylated CpGs associated with muscle mass and function in the elderly.
- Determine whether transcriptomic and epigenetic changes may partly explain the inter-individual variability in the loss of muscle mass in the elderly.

Chapter 2 –

Materials and Methods

Chapter 2 – Methods

2.1. Materials

Chemical/Reagent	Details
Ethanol, Absolute (200 proof)	Fisher Scientific, # 10644795
DNase and RNase free water, DEPC treated	Fisher Scientific, # BPE561-1
Dulbecco's Modified Eagle Medium (DMEM)	Sigma, # 41966-052
DMEM Powder	Sigma, # D5030-10L
Penicillin-Streptomycin	Gibco, # 15140122
Foetal Bovine Serum	Gibco, # 10500064
Horse Serum	Gibco, # 16050122
Chick Embryo Extract	Stratech, # C3999-USB
Collagenase from Clostridium histolyticum Type IA	Sigma, # C9891-500MG
L-Glutamine	Gibco, # 25030024
Sodium Pyruvate	Sigma, # S8636-100ML
Trypsin-EDTA	Gibco, # 25200-056
Phosphate Buffered Saline (PBS) tablets	Sigma, # P4417
TRI Reagent	Sigma, # T9424
Chloroform, 99.8+%, Certified AR for analysis	Fisher Scientific, #10090120
DNase I Amplification Grade	Sigma, # AMPD1-1KT
Random nonamers	Sigma, # R7647
Deoxynucleotide Triphosphates (dNTPs)	Promega, # U1420
M-MLV Reverse Transcriptase kit	Promega, # M1701
QuantiFast SYBR Green PCR Kit (2000)	Qiagen, # 204056
TaqMan® Gene Expression Master Mix	Applied Biosystems, # 4369016
miScript II RT Kit (50)	Qiagen, # 218161
Hs_miR-675_2 miScript Primer Assay	Qiagen, # MS00032102
Hs_miR-675*_1 miScript Primer Assay	Qiagen, # MS00032109
Hs_miR-1_2 miScript Primer Assay	Qiagen, # MS00008358
Hs_SNORD61_11 miScript Primer Assay	Qiagen, # MS00033705
Hs_SNORD68_11 miScript Primer Assay	Qiagen, # MS00033712
Hs_SNORD96A_11 miScript Primer Assay	Qiagen, # MS00033733
Hs_CYC1_1_SG QuantiTect Primer Assay	Qiagen, # QT00209454

Hs_SMAD1_1_SG QuantiTect Primer Assay	Qiagen, # QT00103824
Hs_SMAD5_1_SG QuantiTect Primer Assay	Qiagen, # QT00031360
Hs_PPIA_4_SG QuantiTect Primer Assay	Qiagen, # QT01866137
Hs_GAPDH_1_SG QuantiTect Primer Assay	Qiagen, # QT00079247
Primer and Double-dye (Taqman-style) assays - H19	Primer Design, # DD-hu-600
CDKN2A - TaqMan® Gene Expression Assay (Hs02902543_mH)	Applied Biosystems, # 4331182
Lipofectamine RNAiMAX Transfection Reagent	Life Technologies, # 13778075
Opti-MEM Reduced Serum Medium	Life Technologies, # 31985062
Silencer® Select Negative Control No. 1 siRNA	Ambion, # 4390843
Silencer® Select GAPDH Positive Control siRNA	Ambion, # 4390849
mirVana™ miRNA Mimic, Negative Control #1	Ambion, # 4464058
mirVana™ miRNA Mimic, miR-1 Positive Control	Ambion, # 4464062
H19 siRNA Silencer Select (5nmol)	Ambion, # n272445
hsa-miR-675-5p mirVana miRNA mimic (5nmol)	Ambion, # 4464066
hsa-miR-675-3p mirVana miRNA mimic (5nmol)	Ambion, # 4464066
Phenylmethanesulfonyl Fluoride (PMSF)	Sigma, # P7626
Leupeptin Hemisulfate Salt	Sigma, # L5793
Sodium Fluoride	Sigma, # S7920
Ethylenediaminotetraacetic Acid Disodium Salt	Sigma, # 03690
Triton X-100	Sigma, # T8787
Tween-20	Sigma, # P1379
40% Acrylamide Solution	Fisher Scientific, # BPE1402-1
Ammonium persulfate	Sigma, # A3678-25G
N,N,N′,N′-Tetramethylethylenediamine	Sigma, # T9281-25ML
Methanol for HPLC, >99.9%	Sigma, # 34860-2.5L-M
Glycine Reagent Plus, >99.9%	Sigma, # G7126-1KG
Sodium Chloride >99.5%	Sigma, # S7653-1KG
Tris Base	Fisher Scientific, # BP152-500
Prism Ultra Protein Ladder (10-245kDa)	Abcam, # ab116028
Anti-beta Actin antibody	Abcam, # ab8224
Anti-GAPDH antibody	Abcam, # ab9485
Rabbit Anti-Mouse IgG H&L (HRP)	Abcam, # ab6728
Goat Anti-Rabbit IgG H&L (HRP)	Abcam, # ab6721
Smad5 Antibody 100ul	Cell Signalling Technologies, # 9517S
SuperSignal West Pico PLUS Checmiluminescent Substrate (200ml)	Thermo Scientific, #34577

Proteinase K	Qiagen, # 19131
RNase A (17,500 U)	Qiagen, # 19101
Trichloro(1H,1H,2H,2H-perfluoro-octyl)silane	Sigma, # 448931
Fluroinet FC-40	Sigma, # F9755
Droplet Generation Oil for EvaGreen	Bio-rad, # 1864006
Ficoll PM-400	GE Healthcare, # 17-0300-10
N-Lauroylsarcosin sodium salt solution	Sigma, # L7414
1H,1H,2H,2H-Perfluoro-1-octanol	Sigma, # 370533
20X SSC Buffer	Fisher Scientific, # BP1325-4
Maxima H- Reverse Transcriptase	Thermo Scientific, # EP0753
RNase Inhibitor	Lucigen, # F83923-1
dNTPs	Clontech, # S1477
Exonuclease I	New England Biolabs, # M0293L
HiFi HotStart ReadyMix	Kapa Biosystems, # KM2602
AgenCourt AMPure XP Beads	Beckman Coulter, # A63881
Nextera XT Library Prep Kit	Illumina, # 15032354
Nextera XT Index Kit v2	Illumina, # 15052163
CpGenome Universal Methylated DNA	Millipore, # S7821
Epitect Control DNA	Qiagen, # 59568
HotStarTaq Plus DNA Polymerase	Qiagen, # 203605
Human genomic DNA	Roche, # 11691112001 Promega, # G3041
Streptavidin Sepharose High Performance	GE Healthcare, # 11565015
PyroMark Gold Q96 Reagents	Qiagen, # 972804
EpiTyper Reagent and SpecroCHIP II Set	Agena Biosciences, # 10249
Oligomycin	New England Biolabs, # 9996L
Carbonyl cyanide-p-trifluoromethoxyphenoylhydrazone (FCCP)	Sigma, # C2920
Rotenone	Sigma, # R8875
Antimycin A from Streptomyces sp.	Sigma, # A8674
45% Glucose Solution	Fisher Scientific, # 10152950
2-deoxyglucose	Fisher Scientific, # 15363581
Bovine Serum Albumin	Sigma, # A9647
Paraformaldehyde	Sigma, # 158127
Goat serum	Sigma, # G9023
Senescence Detection Kit	Abcam, # ab65351

Cell Proliferation ELISA, BrdU (colorimetric)	Sigma, # 000000011647229001
MYH1E Antibody	DSHB, # MF20
Goat anti-rabbit IgG H&L (Alexa Fluor 594)	Abcam, # ab150080
Goat anti-Mouse IgG (H+L), Alexa Fluor 488	Life Technologies, # R37120
Anti-NCAM antibody [EP2567Y]	Abcam, # ab75813
DAPI	Sigma, # D9542

2.2. Methods

2.2.1. Cohort details

The Hertfordshire Sarcopenia Study (HSS) is a retrospective cohort study investigating the lifecourse influences on muscle function, mass and morphology in community dwelling older people. It is part of the Hertfordshire Cohort Study (HCS), comprising of men born in Hertfordshire between 1931 and 1939, previously described by Syddall et al.[360]. Recruitment into the HSS was as described in figure **2.1** and in Patel el al. [361] and consists of only male participants.

The HSSe (HSS extension) cohort is a second subset, which unlike the HSS cohort consists of both male and female participants. Recruitment into the HSSe is described in figure **2.1**. The Hertfordshire Sarcopenia Study was approved by the Hertfordshire Research Ethics Committee (number 07/Q0204/68). All participants gave written informed consent to participate in the study and for their biological samples to be used to investigate the molecular mechanisms in sarcopenia.

Figure **2.2** highlights the difference between the HSS and HSSe cohorts, and what experiments were carried out with each cohort. Despite both HSS and HSSe being part of HCS, the two cohorts remain independent subsets of HCS.

2.2.2. Tissue Collection and Anthropometric Measurements

After a home visit, consent and administration of a health and activity questionnaire, participants were invited to attend the Wellcome Trust Clinical Research Facility at the University Hospital Southampton for further investigation. Briefly, participants were fasted overnight, prior to a muscle biopsy being taken using a Weil Blakesley conchotome with a 6 mm biting tip. The biopsy was taken from either the left or right vastus lateralis muscle over the antero-lateral aspect of the thigh. Other investigations carried out included dual energy X-ray absorptiometry (DXA) to quantify total lean mass, appendicular lean mass, fat mass and bone mineral content/density, anthropometric measurements, isometric grip strength and a validated battery of tests to assess motor function, including chair rise time, a 6 meter timed up-and-go and walking speed. Fasting blood samples were also taken from the anterior cubital fossa for subsequent analysis. Dr Harnish Patel and Dr Alicja Baczynska carried out tissue collection and anthropometric measurements.

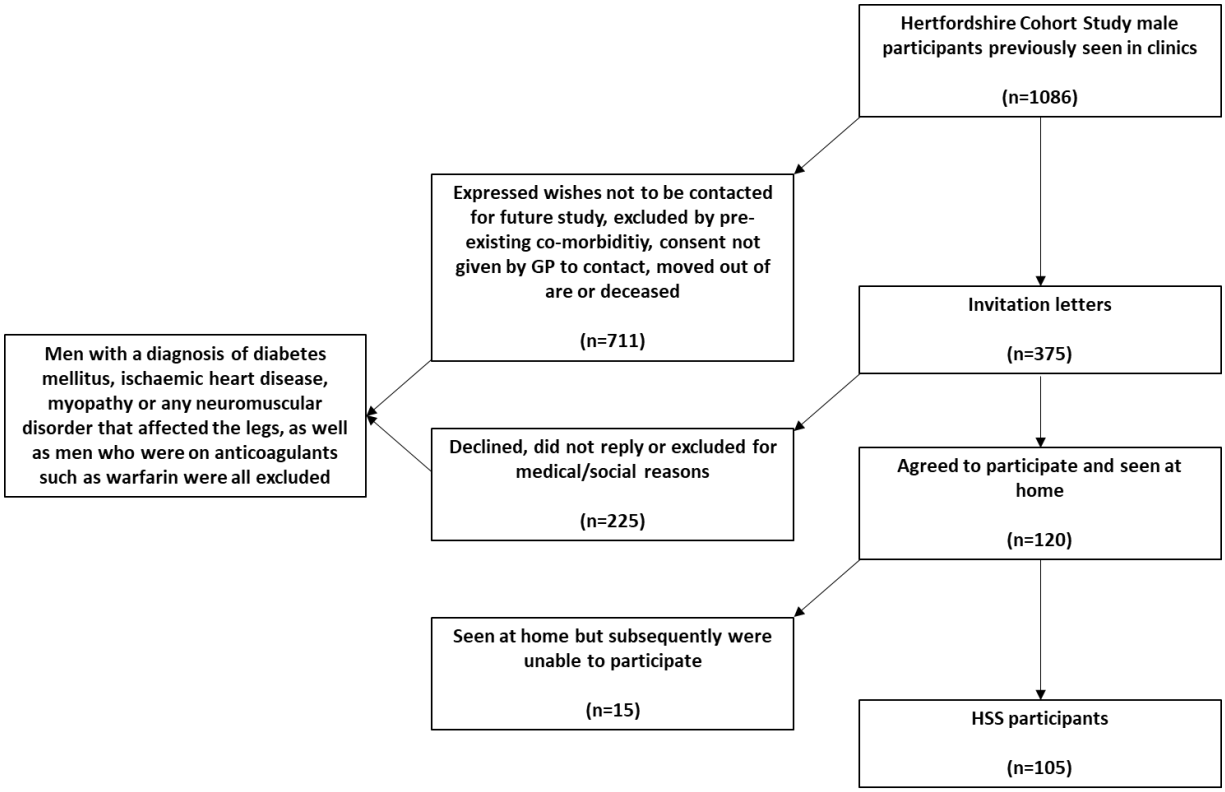


Figure 2.1: Flow chart depicting the recruitment of participants from the Hertfordshire Cohort Study (HCS) to the Hertfordshire Sarcopenia Study (HSS)

Male HCS participants (n=375) were invited to partake in the HSS with permission given from their local GP. 120 participants agreed to take part and were seen at home by the study physician, with details of the study explained to them in detail and written consent obtained. Finally, 105 males agreed to take part in the study and were invited to attend the Wellcome Trust Clinical Research Facility at Southampton General Hospital for investigation, including an overnight stay. Figure adapted from Patel et al. [368].

2.2.3. Myoblast Cell Cultures

Myoblasts cultures were prepared from fresh muscle biopsies from participants of the HSSe cohort. Roughly 40mg of tissue was used per participant. Tissue was minced using a scalpel and subsequently treated with collagenase from *Clostridium histolyticum* type IA for 20 minutes at 37°C. The resulting cell suspension was washed in sterile phosphate buffered saline (PBS) and strained through a 100µm cell strainer to remove large cells and debris, resulting in a cell suspension of satellite cells, myoblasts and fibroblasts. Cell were resuspended in proliferation medium (Dulbecco's Modified Eagle Medium (DMEM) supplemented with 20% Foetal Bovine Serum, 10% Horse Serum, 1% Chick Embryo Extract and 1% Penicillin/Streptomycin solution). Cells were seeded onto 10cm^2 cell culture plates for 3 hours. Fibroblasts adhere to the surface of the plate while myoblasts and satellite cells remain in solution. After 3 hours, cells were transferred to a plate

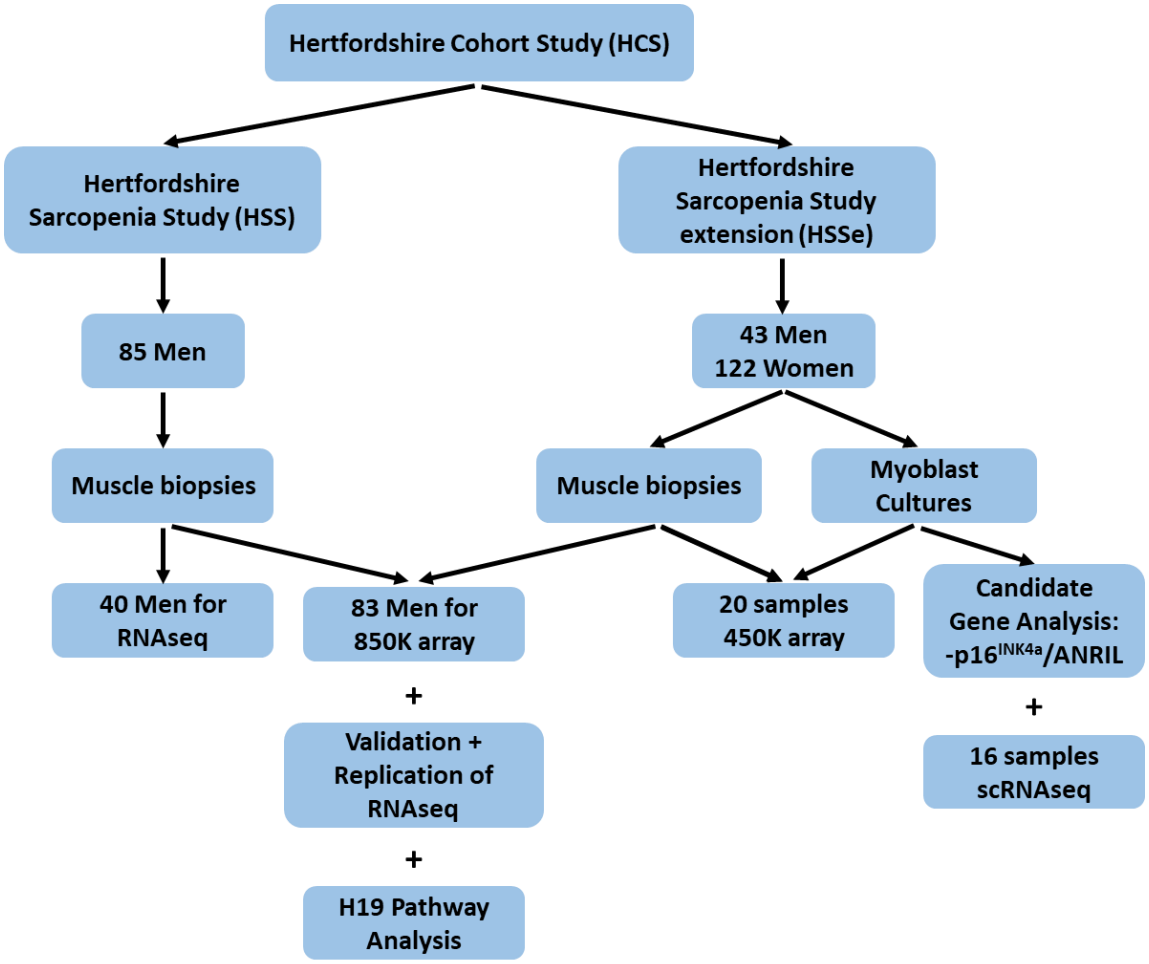


Figure 2.2: Experimental outline of HSS and HSSe Cohorts

Outline of the HSS and HSSe cohorts. Although both cohorts are part of HCS, they form two independent subsets of the HCS, such that the men in HSS are not the same men found in HSSe. Flow diagrams shows which experiments were carried out in which of the two cohorts.

coated with a thin layer of matrigel basement membrane matrix. Cells were grown in a humidified incubator at 37°C in an atmosphere of 5% CO_2. Cells were grown to 80% confluence, passaged to two new matrigel-coated plates, and again grown to 80% confluence. One plate was passaged to five new matrigel-coated plates, and one plate was frozen down in proliferation media supplemented with 5% DMSO and stored in liquid nitrogen vapour phase at the second passage. Final plates were grown to 80% confluence and cells were pelleted for RNA and DNA extractions at the third passage. Cell pellets were stored at -80°C.

2.2.4. Immuno-magnetic Bead Sorting of Cells

Myoblast cultures were sorted using CD56 MicroBeads (Miltenyi Biotech) according to manufacturer's instructions. To enrich the myogenic population, cells were sorted using CD56

MicroBeads prior to experiments, and the number of positive (myogenic) cells calculated by measuring CD56 marker expression by immunocytochemistry (figure **2.3**). All cell populations in the positive fraction were found to be 100% myogenic, showing expression of CD56.

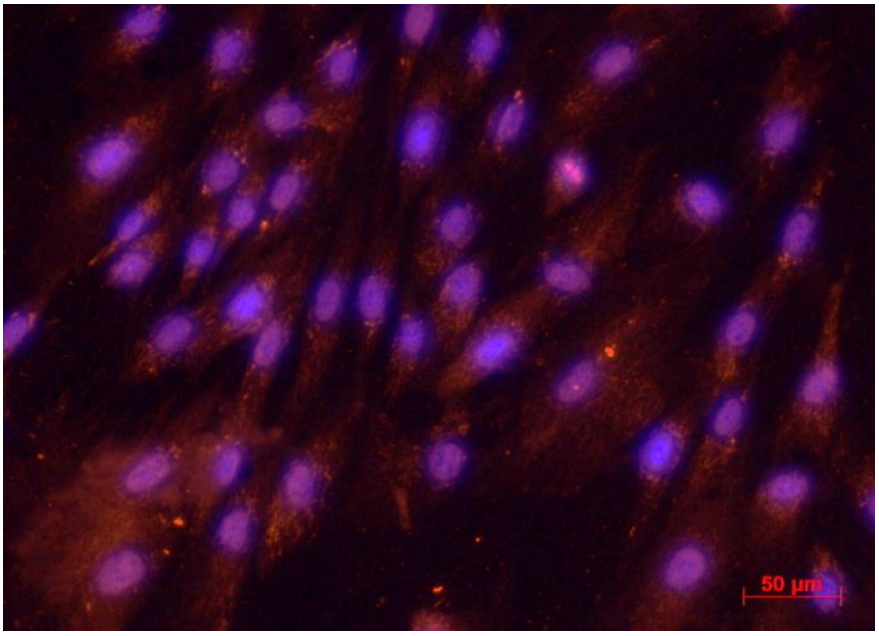


Figure 2.3: CD56 staining of myoblast cultures.
Myoblast cultures were stained with an antibody against CD56 (NCAM), a marker for myoblast cells (red staining). Cells were counterstained with DAPI to visualize cell nuclei (blue staining). As is evident, all the cells express CD56 suggesting a pure culture of myoblast cells.

2.2.5. RNA extraction

RNA from HSS muscle tissue was extracted, quantified and quality checked by Dr Emma Garratt. RNA was extracted from frozen muscle tissue stored in 1ml of culture medium using the *mir*Vana™ miRNA Isolation Kit (Ambion, Life Technologies), following the protocol for total RNA extraction. Protocol was altered to optimise RNA quality and quantity. Briefly, excess frozen culture medium was removed from the tissue without allowing the tissue to thaw, and ~10mg of muscle placed in 600 µl Lysis/Binding buffer. Tissue was homogenised using a Dispomix Homogeniser until all visible clumps were dispersed, before continuing with the protocol according to the manufacturer's instructions for total RNA isolation. RNA from HSSe muscle tissue was subsequently extracted by myself using the same protocol as above. RNA from HSSe myoblast cell pellets was extracted using *mir*Vana ™ miRNA Isolation Kit (Ambion, Life Technologies), following manufacturers protocol for total RNA extraction. Cell pellets were not homogenised in the lysis buffer as with native muscle tissue.

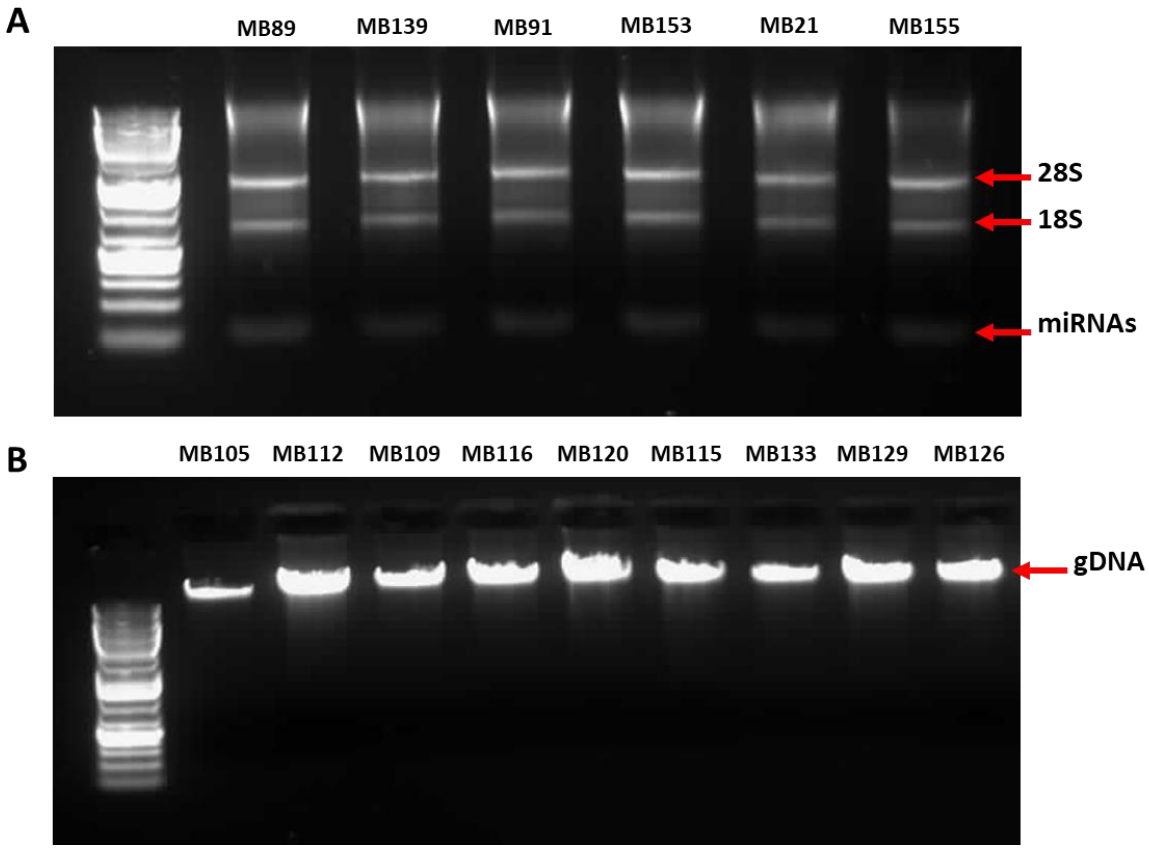


Figure 2.4: RNA and gDNA extractions.

To determine the quality of the RNA and gDNA extracted from the myoblasts and muscle samples, 500ng of RNA (A) and 125ng of DNA (B) were run on a 1% and 0.8% agarose gel respectively. (A) Extracted RNA shows two distinct bands representing 28S and 18S rRNA as well as a faint band at the bottom of the gel representing miRNAs. There is not much smearing of the bands, suggesting good quality of RNA and little degradation. (B) Extracted gDNA shows one sharp band at the top of the gel, with no smearing implying no degradation of the DNA samples.

All RNA was quantified using Qubit 2.0 Fluorometer (Thermo Scientific), 260/280 and 260/230 ratios checked using a NanoDrop 1000 Spectrophotometer (Thermo Scientific) and samples to be sequenced were run on the Agilent Bioanalyzer to check RNA integrity. RNA quality was also checked by agarose gel electrophoresis (figure **2.4a**). All RNA used for sequencing had an RNA Integrity Number (RIN) score > 7. Total RNA was stored at -80°C.

2.2.6. DNA Extraction

Genomic DNA (gDNA) was extracted from HSSe myoblast cell pellets using the GenElute™ Mammalian Genomic DNA Miniprep Kit (Sigma Aldrich) following the manufacturer's protocol. Briefly, the cell pellet was resuspended in 200µl resuspension solution, treated with RNase A

solution and lysed with proteinase K (20mg/ml) and lysis buffer. The solution was passed through a GenElute column, gDNA washed and finally eluted in 3x50µl H_2O. Columns were incubated at room temperature for 5 minutes with H_2O on the column before elution to maximize yields.

gDNA from HSSe muscle tissue was extracted using the high salt method. 500µl of TNES (50mM Tris pH 7.5, 400mM NaCl, 100mM EDTA, 0.5% SDS) and 5µl 20mg/ml proteinase K was added to 20mg of tissue in a gentleMACs™ homogenisation tube (Miltenyi Biotec), and homogenised using a gentleMACs homogeniser, after which they were incubated overnight at 55°C. After complete digestion of the tissue, lysates were transferred to tubes and 500µl of 2.6M NaCl was added and vortexed vigorously. Samples were spun at 13,800g for 10 minutes and supernatant transferred to a new tube. Ice-cold 100% ethanol was added to the supernatant and incubated at -20°C for 30 minutes. Samples were spun at 13,800g for 15 minutes to pellet the DNA and was resuspended in 500µl dH_2O. 5µl 10mg/ml RNase A was added and incubated at 37°C for 1hour. An equal amount of phenol-chloroform was added, samples mixed and spun at 13,800g for 7 minutes. The top aqueous phase was carefully transferred to a new tube. Phenol-chloroform wash was repeated, after which 3M NaAc pH5.2 was added at a 1:10 ratio and 2 volumes of ice-cold 100% ethanol. Samples were stored at -20°C for 3hours, and DNA pelleted by centrifugation at 13,800g for 10 minutes. DNA was washed in 70% ethanol, allowed to air dry and resuspended in dH_2O. Samples were left overnight at 4°C to facilitate resuspension of the DNA.

All genomic DNA was quantified and 260/280 and 260/230 ratios checked using a NanoDrop 1000 Spectrophotometer (Thermo Scientific). DNA quality was checked by agarose gel electrophoresis on a 0.8% agarose gel (figure **2.4b**). All gDNA was stored at -20°C.

2.2.7. Protein Extraction

Muscle tissue was homogenised using a handheld bio-vortexor homogenizer. Protein lysates from muscle tissue were prepared in extraction buffer containing 50mM Tris (pH 7.5), 150mM NaCl, 50mM NaF, 1mM EDTA, 0.5% Triton X-100 with 1mM PMSF and 0.25% leupeptin. Protein lysates from cell cultures were prepared as above without the use of a handheld homogenizer. Total protein lysate concentrations were determined using a Pierce BCA Colorimetric Protein Quantitation Assay (Thermo Scientific, UK). 1 µl of protein lysate added to 9 µl of RNA/DNA free H_2O was added to 200µl of BCA working reagent, mixed and incubated at 37 ˚C for 30 minutes in the dark with gentle agitation. Absorbance at 540 nm measured on the GloMax Plate Reader (Promega).

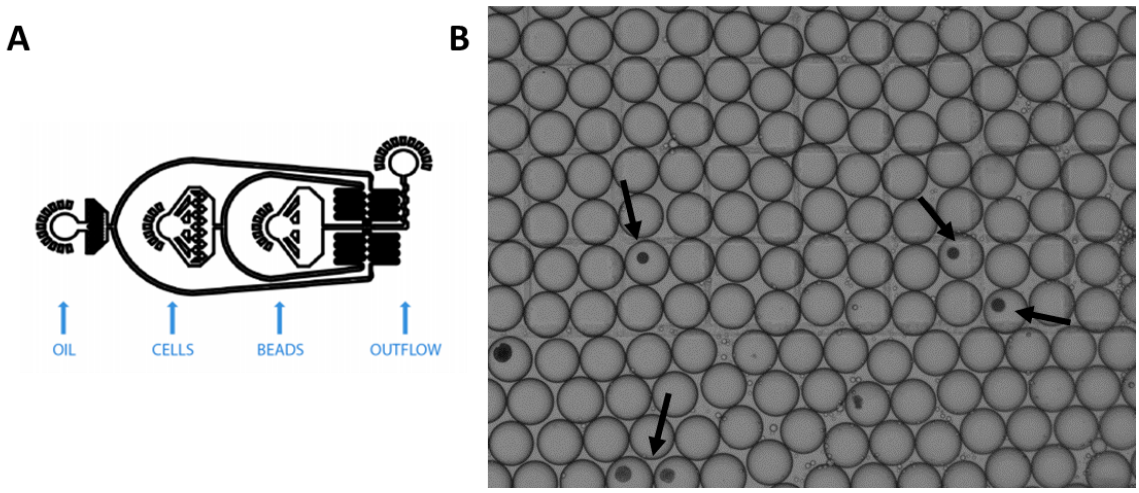


Figure 2.5: DropSeq microfluidics.

(A) Schematic of the microfluidic device used to encapsulate single cells into droplets containing beads with linkers attached to it. Three inlets (oil, cells and beads) and one output (outflow) are present with channels of 125µm diameter. Image taken from Macosko et al. [369]. (B) Visualisation of collected droplets showing a uniform size of droplets (~1nl) with some of the droplets containing a bead, as indicated by an arrow.

2.2.8. Total RNA-Seq Library Prep and Sequencing

1µg total RNA was sent to the Nestle Institute of Science (NIHS) for library prep and sequencing. Briefly, library prep was carried out using TruSeq Stranded Total RNA with Ribo-Zero kit (Illumina) following manufacturer's instructions. Cycling conditions for PCR amplification were optimized by preliminary qPCR to limit PCR duplicates, with 13 cycles performed using Kapa HiFi Hot Start enzyme. Size and quality of libraries were checked by electrophoresis on Caliper GX Microfluidics system and equimolar samples were pooled for multiplexing. Each library was sequenced as paired-end, 2x101bp on a HiSeq 2000. Each library was sequenced on 8 lanes, at 10 million reads per lane.

2.2.9. DropSeq Encapsulation, Library Prep and Sequencing

2.2.9.1. Manufacturing Microfluidic Droplet Devices

Microfluidic devices (figure **2.5a**) were fabricated using polyurethane moulds with a negative of the desired device structure, as developed by Macosko et al. [362]. A polydimethylsiloxane (PDMS) replica made using the sylgard 184 silicone elastomer kit (Dow Corning) was degassed and added to the polyurethane moulds. PDMS was baked at 60°C for 90 minutes and allowed to cool. Inlets and outlets were punched using a 1-mm-diameter biopsy punch. PDMS device was bound to a glass microscope slide using a plasma asher system (Femto, Diener Electronic) followed by baking at 60°C for 15 minutes to release water between the layers. 1% trichloro(1H,1H,2H,2H-perfluoro-

octyl)silane (Sigma) in Fluorinert FC-40 (Sigma) was injected into the outlet and incubated for 5 minutes. Excess fluid was drained by flushing the device with pressurized N_2 to dry the device.

2.2.9.2. Cell Suspension Preparation

Cells were grown to confluence and tryspinzised with 0.25% trypsin-EDTA. Trypsin was quenched with complete proliferation media and cells counted. Cells were centrifuged at 1000g for 5 minutes and resuspended in 1x PBS + 0.01% BSA at a concentration of 100 cells per µl. Cell suspensions were kept on ice until needed.

2.2.9.3. Droplet Encapsulation and Breakage

Beads (144,000) were pelleted by centrifugation at 1000g for 1 minute and TE-TW buffer discarded. Beads were resuspended in 1.2ml lysis buffer (200mM Tris pH 7.5, 6% Ficoll PM-400, 0.2% Sarkosyl, 20mM EDTA, 50mM DTT). Bead suspension was loaded into a 3ml plastic syringe and the cell suspension into a 1ml plastic syringe. Droplet generation oil (Biorad) was loaded into a 5ml syringe. Tubing was connected to the appropriate inlet on the device. Beads were kept in suspension by a magnetic stirrer in the syringe. The solutions were injected into the device at a flow rate of 15,000µl/hr for the oil and 4,000µl/hr for both the beads and the cells. Droplets were examined under a microscope to ensure uniformity and single occupancy of droplets before collection (figure **2.5b**). Droplets were collected in 50ml falcon tubes. Excess oil was removed and 6X SSC buffer and 1ml perfluorooctanol (PFO) was added. Tube was shaken vigorously for 20 seconds by hand. The beads were pelleted by centrifugation at 1000g for 1 minute and supernatant removed leaving 5ml above the oil-aqueous interface. 30ml 6X SSC was added and the aqueous layer transferred to a new tube. Beads were once again pelleted by centrifugation at 1000g for 1 minute, supernatant removed and bead pellet transferred to a non-stick 1.5ml Eppendorf tube. The pellet was washed with 1ml 6X SSC and once with 300µl 5X Maxima H-RT Buffer.

2.2.9.4. Reverse Transcription and Exonuclease I Treatment

200µl RT mix, consisting of 1X Maxima RT buffer, 4% Ficoll PM-400, 1mM dNTPs (Clontech), 1U/µl RNase Inhibitor (Lucigen), 2.5µM Template Switch Oligo (table **2.1**) and 10U/µl Maxima H- RT, was added to the beads and incubated at room temperature for 30 minutes, followed by 90 minutes at 42°C. Beads were pelleted by centrifugation at 1000g for 1 minute, beads were washed once with 1X TE-SDS. Beads were pelleted and washed twice with TE-TW, being pelleted between each by centrifugation at 1000g for 1 minute. Beads were finally washed once with 10mM Tris pH7.5. Beads were pelleted by centrifugation at 1000g for 1 minute and were resuspended in 200µl exonuclease I mix consisting of 1X exonuclease I buffer and 1U/µl exonuclease I (NEB) and incubated for 45 minutes at 37°C. Beads were washed once with 1X TE-SDS, twice with TE-TW and once with H_2O

and were pelleted by centrifugation at 1000g for 1 minute between each wash step. Beads were mixed well and 20µl was added to a Fuchs-Rosenthal haemocytometer chamber to count the beads.

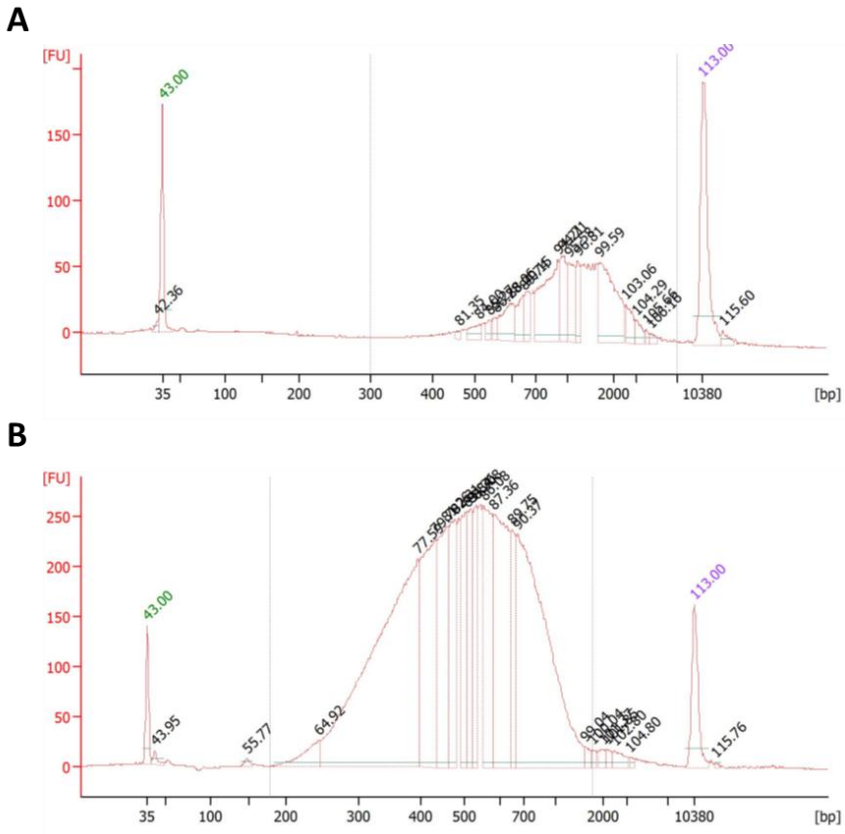


Figure 2.6: Bioanalyzer traces of DropSeq libraries.

(A) Bioanalyzer trace of a representative cDNA library after PCR amplification, showing a good library size after PCR that is not too big or too small, with no contamination after purification. (B) Bioanalyzer trace of a representative cDNA library after tagmentation of the cDNA library, showing good library size and no contamination of the cDNA library. Peaks at the start and the end of the trace represent markers used for quantification.

2000 beads were used per PCR reaction, and the RNA bound to the beads was amplified by PCR in a 50µl mix consisting of 1X Kapa HiFi HotStart Readymix (Kapa Biosystems) and 0.8µM SMART PCR Primer (table **2.1**). Thermocycling conditions were as follows: 95°C for 3 minutes; 4 cycles of: 98°C for 20 seconds, 65°C for 45 seconds and 72°C for 3 minutes; 9 cycles of: 98°C for 20 seconds, 67°C for 20 seconds and 72°C for 3 minutes; finally followed by 72°C for 5 minutes. Libraries were purified with 0.6X Agencourt AMPure XP beads (Beckman Coulter) following the manufacturer's protocol and eluted in 13µl of H_2O. The concentration was quantified on a BioAnalyzer High Sensitivity DNA Chip (Agilent) (figure **2.6a**).

2.2.9.5. Library Prep and Sequencing

For each sample, 500pg purified cDNA was used as input to the Nextera XT tagmentation reaction (Illumina). Tagmentation was carried out to fragment the cDNA followed by tagging of the fragments with the appropriate Nextera Index. Tagmentation was carried out following the manufacturer's protocol, with the exception of 200nM custom primer P5 SMART PCR Hybrid oligo (table **2.1**) was used along with Nextera N70X oligo, where X denotes the Nextera Index used. Samples were amplified as follows: 95°C for 30 secs; then 12 cycles of: 95°C for 10 seconds, 55°C for 30 seconds, 72°C for 30 seconds; followed by a final 5 minutes at 72°C. Tagmented libraries were once again purified with 0.6X Agencourt AMPure XP beads following the manufacturer's protocol, and eluted in 10µl H_2O. Library concentration was quantified on a BioAnalyzer High Sensitivity DNA Chip at least twice (figure **2.6b**).

Libraries were pooled at a final concentration of 2nM. Libraries were sequenced on the Illumina NextSeq 500 using a 2pM concentration of library in 1300µl HT1. 0.3µM Read1CustSeqB primer (table **2.1**) was used for priming read 1. Read 1 was 20bp and read 2 was 50bp.

2.2.10. Methylation Arrays

Illumina Infinium Human Methylation450 BeadChip and Infinium Human MethylationEPIC BeadChip runs were carried out at the Centre for Molecular Medicine and Therapeutics (CCMT). 1µg genomic DNA (gDNA) from 20 HSSe muscle and matched myoblasts was sent for 450K methylation arrays, and 750ng of gDNA from 40 HSS male and 43 HSSe male muscle biopsies was sent for EPIC methylation arrays. gDNA was bisulfite converted, quality checked and run on an Infinium Methylation450K Beadchip or an Infinium MethylationEPIC BeadChip (Illumina). Inter- and intrachip replicates were included for technical validation of the arrays. IDAT files were provided and downloaded for subsequent analysis of differential methylation

2.2.11. Total RNAseq Bioinformatic Analysis

2.2.11.1 QC and pre-processing of reads

Sequencing files were de-multiplexed by the Nestle Institute of Health Sciences (NIHS), resulting in 8 fastq files per sample, for both forward and reverse reads. Fastq files were subsequently sent to the University of Southampton for further analysis. All bioinformatics analysis of the RNAseq data was carried out using the IRIDIS4 computing cluster, part of the High Performance Computing Facility at the University of Southampton. QC metrics at the read level were provided using FastQC. FastQC was performed after every step of processing to ensure consistently high read quality. The 8 individual files per sample were merged into 1 fastq file before processing, using the Linux commands *zcat* and *gzip –c,* to reduce the computational requirement. Adapters were trimmed

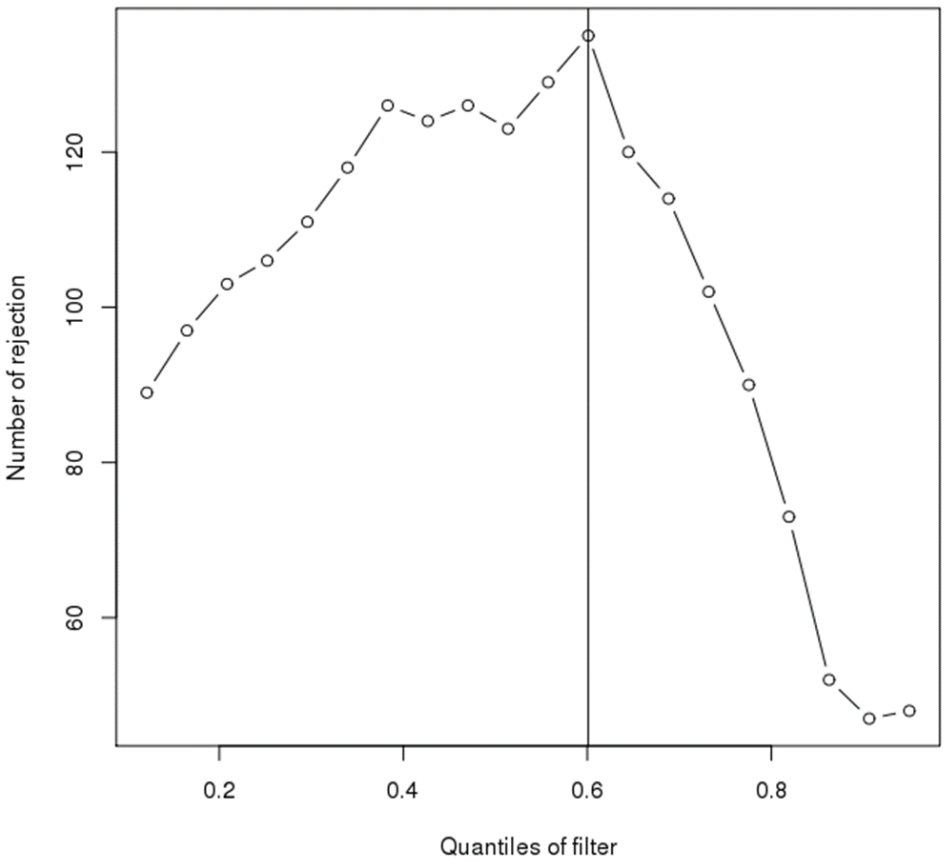


Figure 2.7: Independent filtering

The *results* function of DESeq2 optimises the filter threshold obtained for independent filtering. The threshold in the control vs sarcopenia analysis is shown, with the threshold chosen (vertical line) being the lowest quantile of the filter statistic (mean of the normalized counts) for which the number of rejections is within 1 residual standard deviation to the peak of the curve fit to the number of rejection over all the quantiles.

from reads and subsequently quality trimmed to a minimum base Phred score of 10 using BBDuk from the BBMap suite of tools [363]. The following options were used: *ktrim=rl* (trims Kmers from left and right side of reads), *k=13* (length of kmers to use), *mink=11* (allows smaller Kmers to be trimmed from ends of reads), *hdist=1* (hammering distance, allows 1 mismatch), *rcomp=t* (looks for the reverse complement of the reference sequences as well as the forward sequence), *minlen=25* (minimum length of reads after trimming), *qtrim=rl* (trim bases from the left and right sides of reads based on quality), *trimq=10* (minimum Phred quality of bases to keep), *tpe* (trims both paired reads to the same length) and *tbo* (trims adapters based on pair overlap detection).

2.2.11.2 Read alignment to the genome

Reads were aligned to the current human reference genome (GRCh38) using the TopHat aligner [364] (version 2.0.14). Briefly, TopHat uses a two-step mapping process. Firstly, it utilizes Bowtie [365] (version 2.2.5) to align reads directly to the genome with no gaps. Secondly, any reads which cannot align with Bowtie in the first pass are split into smaller fragments in order to use gapped alignment, and then aligned to the genome using Bowtie. These reads potentially cross exon-exon boundaries. TopHat generates possible splice junctions from two sources. Firstly, it uses a reference GTF file as a source of potential splice junctions. The second source comes from when two fragments from the same read are mapped at a certain distance on the same genomic sequence, suggesting reads span multiple exons, with introns in this case being found ab initio. TopHat uses canonical and non-canonical splice junctions to determine possible locations for gaps in the read alignment. TopHat option *--library-type fr-firststrand* was used.

2.2.11.3. Count table generation

Preliminary analyses of BAM files generated by TopHat were carried out using Samtools [366] (version 1.2.1) and Picard tools [367] (version 1.97) to determine mapping statistics. Files were sorted by read name with Samtools, followed by conversion to SAM files, filtering out reads with mapping quality < 5. Count tables for each sample were generated using the htseq-count function of HTSeq [368] (version 0.6.1) with the following options: *--stranded=reverse* and *--mode=union*. The reference GTF file used with HTSeq was obtained from Ensembl for the human genome release GRCh38.83.

2.2.11.4. Differential gene expression

Differential gene expression analysis was performed in R (version 3.2.1). The DESeq2 [369] software package was used to perform two-class testing for differential gene expression. Two unpaired analyses were carried out, comparing control individuals against pre-sarcopenics and subsequently sarcopenic individuals. Prior to differential expression analysis, DESeq2 carries out independent filtering of the counts, in order to filter out tests from the procedure that have little chance of showing significant evidence in order to optimize and increase the number of significant adjusted p values obtained. DESeq2 carries this out as default, using the default parameters. Although the filter threshold is not the same for each of the analyses carried out, the threshold is optimised for each analysis to achieve a greater number of significantly adjusted p values while decreasing false positives. The filter threshold chosen is calculated automatically by the *results* function of DESeq2, and is the lowest quantile of the mean of the normalized counts, for which the number of rejections is within 1 residual standard deviation to the peak of the curve fit to the number of rejection over all the quantiles (figure **2.7**). A gene with a false discovery rate (FDR) < 0.1 was classed as being

differentially expressed. DESeq2 was subsequently used to analyse differential expression with respect to other muscle phenotypes, including appendicular lean mass index (ALMI), grip strength, gait speed and 6m timed up-and-go (6mTUG). Similar to the categorical analysis, a gene with an FDR < 0.1 was classes as differentially expressed.

2.2.11.5. Gene Set Enrichment Analysis

Gene set enrichment analysis (GSEA) was carried out using GSEA v2.2.2 from the Broad Institute. The following gene set databases were used from MSigDB: hallmark gene sets (H), curated gene sets (C2), motif gene sets (C3) and Gene Ontology (GO) gene sets (C5). The GSEAPreranked tool was used, using a pre-ranked gene list provided to the software. Genes were ranked based on their fold change and nominal p-values. In producing the ranked gene list, nominal p-values were used rather than FDR due to a greater variability in values, resulting in better ranking of genes. Using nominal P values does not change the order of significant genes as determined by FDR. Enrichment statistic used was classic as opposed to the default weighted option. As a pre-ranked list was provided to the software from RNAseq data, the collapse dataset to gene symbols option was changed to *FALSE*. All other options used were kept as default. Gene sets with an FDR < 0.1 were classed as being differentially enriched.

2.2.12. Methylation Array Bioinformatic Analysis

2.2.12.1 QC and pre-processing of file

IDAT files were processed using the minfi package (v.1.24.0) in R using the IlluminaMethylationEPICanno.ilm10b2.hg19 and IlluminaMethylationEPICmanifest datasets. Funnorm normalization was carried to obtain normalized beta values for all samples. PCA plots and were generated and hierarchal clustering based on Euclidean distance was used to determine sample clustering. Madscores (modified Z-scores) were calculated to determine sample outliers, and those with a madscore < -5 were removed from subsequent analysis. Replicates with the lowest madscore were removed from subsequent analysis. Following removal of outliers and replicates, raw data was re-normalized before further anlaysis. CpG probes with a detection p-value > 0.01 or with a beadcount < 3 were set to missing. CpG probes were excluded from subsequent analysis if missing in more than 5% of samples. Probes with a SNP at the CpG interrogation site or at the single nucleotide extension were also removed, as well as probes that have previously been shown to be cross-reactive [370]. Probes with a methylation range <5% were also removed from subsequent analysis. As all the samples in this dataset originated from male donors, the sex chromosomes were not removed.

2.2.12.2. Differential Methylation Analysis

SVA (surrogate variable analysis) was carried out to determine the presence of any unknown confounders in the data which may affect downstream analysis. No surrogate variables were found. ComBat was subsequently used to adjust for batch/chip effect due to interchip replicates clustering apart in the hierarchal clustering analysis. ComBat-adjusted beta values were then used to determine differential methylation.

Differential methylation analysis was carried out using limma (v3.4.2) in R. The models provided to limma were:

Methylation ~ Status + Age
Methylation ~ ALMI + Age
Methylation ~ Grip Strength + Age
Methylation ~ Gait Speed + Age
Methylation ~ Fasting Glucose + Age
Methylation ~ Insulin + Age
Methylation ~ Age

Genomic inflation factors (λ) were calculated for the different models interrogated. Models with λ>1.2 were run through the BACON algorithm to adjust for inflation and bias with inflation and bias adjusted p-values subsequently corrected for multiple testing using the Benjamini-Hochburg method. Differentially methylated regions were determined using DMRcate.

2.2.12.3. Pathway analysis

Pathway analysis was carried out with the significant CpGs in the differential methylation analysis with respect to ALMI (FDR<0.2). CpGs were used if they were associated with an annotated gene as indicated in the Illumina 450K Manifest file. Duplicate genes were removed. Gene networks were generated using GeneMANIA [371] in Cytoscape [372]. Networks generated by GeneMANIA were further clustered using the MCODE algorithm [373] and enriched GO terms determined with BiNGO [374] and the gometh function from the missMethyl package in R.

2.2.13. DropSeq Bioinformatic Analysis

2.2.13.1 QC and pre-processing of files

Raw bcl files for forward and reverse reads were de-multiplexed using bcl2fastq2 (v2.18) to generate fastq files for the forward and reverse reads. FastQC was run on all fastq files to examine Phred quality scores of the reads. All reads had a mean Phred > 30. Fastq files were subsequently passed through the Dropseq alignment pipeline [362]. Briefly: fastq files were converted into unmapped BAM files using Picard's FastqToSam function, followed by SortSam to sort the BAM files

by read name. Reads in the BAM files were tagged with the cell and molecular barcodes, adding the BAM tag XM for molecular barcodes and the XC tag for cell barcodes. Reads with low quality bases in the cell or molecular barcodes were removed and any SMART adapter sequence or trailing polyA tails were trimmed from the reads. BAM files with cell and molecular barcodes extracted were converted back to fastq files. Fastq files were aligned to the hg19 genome using the STAR aligner to generate aligned BAM files. BAM files were subsequently sorted by read name and the aligned and unaligned BAM files were merged using Picard's MergeBamAlignment tool, to combine alignments with their cell and molecular barcodes. Reads are tagged with the BAM tag GE for reads that overlap the exon of a gene as provided by the hg19 refFlat file. Finally, a digital gene expression (DGE) profile is generated for each sample, with the gene expression of each cell per sample.

2.2.13.2. Cell/Gene Level QC

The DGE files were examined to determine the number of cell barcodes that correspond to actual single cells and discard barcodes that correspond to empty beads or noise. Briefly, knee plots of the cumulative fraction of reads were plotted for each DGE. Density plots of the cumulative read fractions were also generated. The number of barcodes that correspond to actual cells was determined as the first inflection point after the initial peak. Cells were discarded if more than 20% of counts were associated with mitochondrial genes. Cells with library sizes and number of features 5 median absolute deviations (MADs) below the median were also discarded.

Genes with low expression or only expressed in a handful of cells are likely to be uninteresting. Therefore, any genes that are expressed in no cells at all were removed from the dataset as well as genes with low expression across all the cells.

2.2.13.3. Data Normalization

To account for the difference sequencing depths between the cells and differences in capture efficiency, reads counts were normalized in order to remove the effects of these confounding factors from downstream quantitative analysis. To do this, it is assumed, that most genes are not differentially expressed between the cells, and therefore and differences between the cells is assumed to be due to bias and is removed by scaling the data. Size factors were calculated for each cell and represent the scaling factor for each cell library. Briefly, counts are pooled from multiple cells to increase the counts size, allowing for accurate size factor estimation. These pooled size factors are subsequently deconvolved into cell factors to allow for cell-specific normalization.

2.2.13.4. Cell Cycle Classification

Cell cycle classification was carried out using a previously published method. Cell cycle genes for each phase of the cell cycle were downloaded from the Reactome database (http://www.reactome.org) and the CellCycleSorting function of the Seurat package (v2.3.4) was used to determine the cell cycle phase of each cell. The algorithm compares the expression of marker genes for the G2/M and S phases and as these marker sets should be anticorrelated, a G2/M and an S score is given to each cell based on the expression patterns of the markers. Cells expressing neither marker sets are most likely not cycling and in G1 phase.

2.2.13.5. Highly Variable Genes and Cell Cluster Identification

The variance in expression for each gene was calculated and then decomposed into biological and technical components. As spike ins were not included in this experiment, the technical variance was calculated using endogenous genes. Briefly, this assumes that most of the genes do not show variable expression, such that any variance in the genes expression is dominated by the technical component and not the biological component. A mean-variance trend is fitted to the endogenous genes, which is then used to estimate the technical component of the variance. The biological component of the variance is subsequently calculated by subtracting the technical component of the variance from the total variance. Highly variable genes (HGVs) were determined as those genes which show a biological component that is significantly greater than zero (FDR < 0.05). In order to remove HVGs that are caused by random noise, HVGs were correlated with each other, and gene pairs with a significantly negative or positive Spearman's rho were taken further to aid in the identification of cell subpopulations and further analysis.

In order to remove the effect of covariates (participant each cell was obtained from and cell cycle phase of the cell), the data was run through the *removeBatchEffect* command from the limma package in R. This generates adjusted expression values for all the cells. The normalized and adjusted expression values for the correlated HVGs were used for cell clustering. Hierarchal clustering based on Euclidean distances between the cells was used to cluster the cells, using Ward's criterion [375,376] to reduce the total variance within each cluster, ensuring each member of the cluster is maximally similar with respect to gene expression pattern. Clusters were defined by applying a dynamic tree cut to the dendogram as described by Langfelder et al. [377]. A Chi-Square Test of Independence was carried out to determine if there was an association between cluster and sarcopenia status of the participant each cell originated.

2.2.13.6. Differential Expression for Marker Gene Identification

Differential expression to determine marker genes between subpopulations was carried out using the edgeR package. Genes were identified that are consistently differentially expressed in one

subpopulation of cells compared to the others. Uninteresting factors were adjusted for in the design formula provided to edgeR (participant and cell cycle phase). Each gene is tested for differential expression between the chosen cluster and every other cluster in the dataset. Marker genes were subsequently identified by taking the top differentially expressed genes from each pairwise comparison between clusters.

2.2.13.7. Gene Pathway Analysis

The top 50 marker genes for each cluster were used to determine GO terms enriched in each cell cluster identified. Enriched GO terms were determined using the topGO package in R. The gene universe provided contained all the genes sequenced after filtering out for genes that do not pass the quality control checks. A Fisher's exact test was carried out using the classic algorithm. P-values were corrected for multiple testing using the Benjamini-Hochburg method.

2.2.14. cDNA Synthesis

Prior to cDNA synthesis, RNA was DNase I treated to remove traces of gDNA. Amplification grade DNase I (Sigma) was added to the RNA for 15 minutes at room temperature, followed by incubation at 70°C for 10 minutes with a stop solution to inactivate the DNase enzyme.

For mRNA analysis, cDNA synthesis was carried out using the M-MLV Reverse Transcriptase (Promega), following the manufacturer's instructions. Briefly, on ice, 500ng DNase-treated RNA was mixed with 10mM dNTP mix (Promega) and 1µM random nonamers (Sigma). RNA was incubated at 70°C for 10 minutes. RNA was subsequently placed on ice and mixed with 5X M-MLV Reverse Transcriptase buffer and M-MLV Reverse Transcriptase. RNA was incubated at room temperature for 10 minutes, followed by 60 minutes at 37°C and 10 minutes at 90°C. For miRNA analysis, cDNA synthesis was carried out using miScript II Reverse Transcriptase Kit (Qiagen). cDNA synthesis was carried out following the manufacturer's instructions. Briefly, on ice, 500ng DNase-treated RNA was mixed with the 5X HiFlex Buffer, 10X Nucleics mix and reverse transcriptase. RNA was incubated at 37°C for 60 minutes, followed by 5 minutes at 95°C to inactivate the enzyme. All cDNA was diluted to 5ng/ul using RNase/DNase free water and stored at -20°C.

2.2.15. Quantitative Real Time PCR (qRT-PCR)

Quantitative real-time polymerase chain reaction (qRT-PCR) was carried out in 384-well plates using a LightCycler 480 system (Roche). qRT-PCR was carried out in duplicate 10µl reactions using either TaqMan Gene expression master mix (Life Technologies) or QuantiFast SYBR Green Master Mix (Qiagen). Primer sequences and where they were obtained can be found in table **2.2**. 1µl of each 10X primer mix was used in each 10µl reaction. 10ng of cDNA was used per qRT-PCR reaction, with

45 cycles carried out for each amplicon analysed. All genes used an annealing temperature of 60°C. β_2-microgloublin (B2M), ribosomal protein L13A (RPL13A) and succinate dehydrogenase A (SDHA) expression was used to normalize the data for age analysis, using the ΔΔCt method. Cyclophilin A (PPIA) and cytochrome c1 (CYC1) expression was used to normalize data for ALMI analysis. For miRNA analysis, 500pg of cDNA was used in each 10µl reaction. RNU6, SNORD68, SNORD96A and SNORD61 were used to normalize the miRNA expression data, using the ΔΔCt method.

2.2.16. gDNA Bisulfite Conversion

500ng of genomic DNA (gDNA) was bisulfite converted using the EZ-96 DNA Methylation-Gold Kit (Zymo Research) following the manufacturers protocol. With each 96-well plate converted, 3x 100% Methylated DNA (Millipore), 3x 0% Methylated DNA (Qiagen) and 3x human genomic DNA (Roche) controls were also converted. Bisulfite converted DNA was subsequently eluted in 30µl RNase/DNase-Free water and stored at -20°C.

2.2.17. Quantitative DNA Methylation Analysis

2.2.17.1. Pyrosequencing

PCR amplification was performed using 2µl of bisulfite converted DNA using the Hotstart Plus DNA Polymerase (Qiagen) in 50µl reactions, using the primers found in table **2.2**. PCR primers specific for bisulfite-converted DNA were designed using the PyroMark Assay Design Software (Qiagen). PCR products were run on a 1.5% agarose gel to ensure amplification of the region of interest (figure **2.8a**). To control for amplification bias and calibrate the assay, the same regions were amplified from bisulfite converted 100% methylated DNA, 0% methylated DNA and human gDNA (Roche). No template controls were included on all plates to ensure no gDNA contamination is present. Briefly, PCR products were immobilised on streptavidin-sepharose beads (GE Healthcare), washed, denatured and released into annealing buffer containing the appropriate sequencing primer. Pyrosequencing was carried out on a PyroMark MD (Qiagen) and methylation percentage calculated using the Pyro Q CpG software (Qiagen).

2.2.17.2. Sequenom

Quantitative DNA methylation analysis was carried out using the Sequenom MassARRAY Compact System (Agena). 1µl (10ng) bisulfite-converted DNA was amplified (Qiagen HotStar Taq Polymerase) using PCR primers specific for bisulfite-converted DNA designed using EpiDesigner (Agena) in a 5µl reaction. PCR products were run on a 1.5% agarose gel to ensure amplification of the region of interest (figure **2.8b**). Each reverse primer contained a T7-promoter tag for *in vitro* transcription (5`-cagtaatacgactcactatagggagaaggct-3`), while the forward primer had a 10mer tag to balance melting temperatures (5`-aggaagagag-3`). Primer sequences available in table **2.2**. PCR products

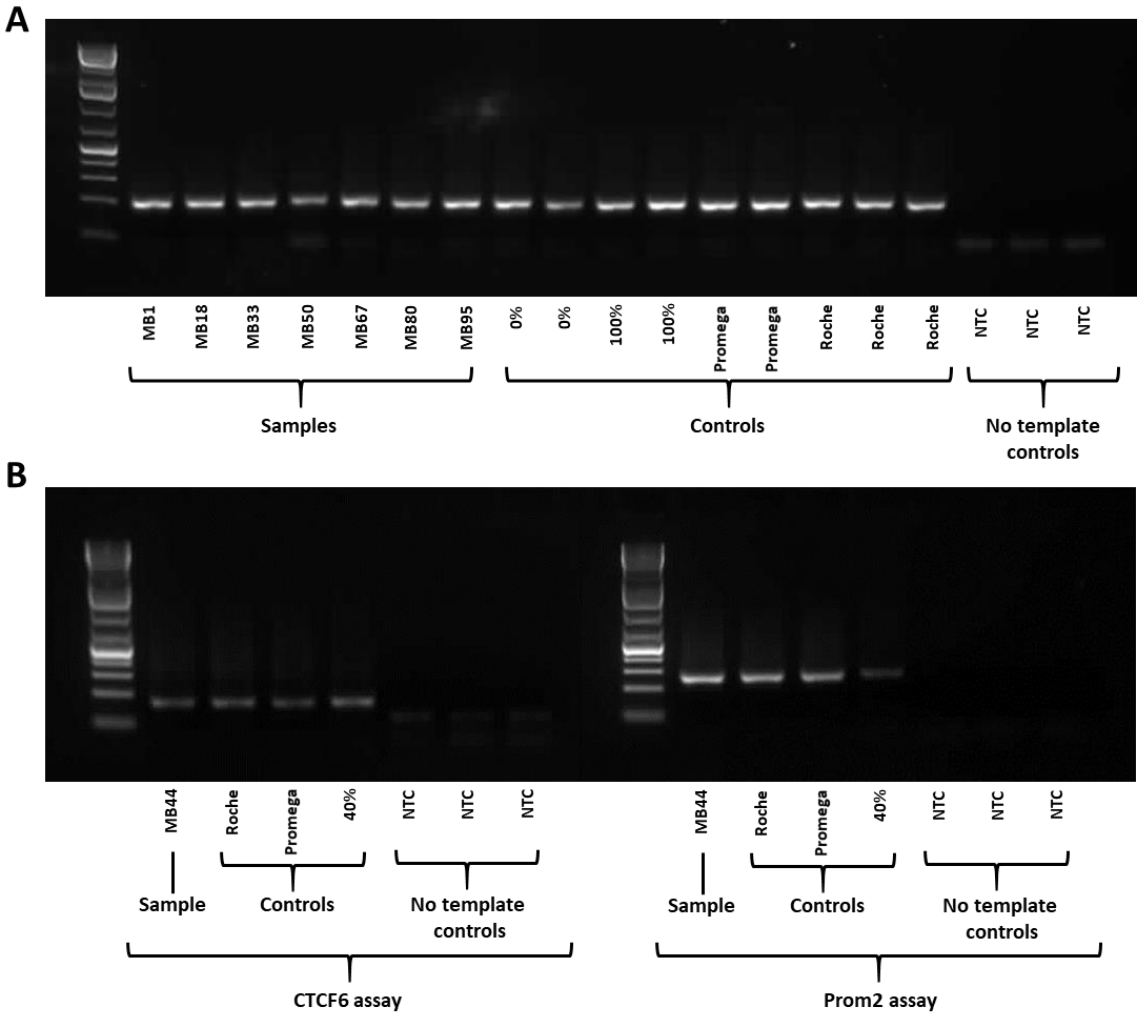


Figure 2.8: Methylation analysis PCR products

PCR products for (A) pyrosequencing and (B) sequenom run on a 1.5% agarose gel to check amplification and contamination. All no template control were negative, with some slight primer dimer. All other samples and controls show a strong band of the correct size with no other bands present.

were Shrimp Alkaline Phosphatase treated (Agena), followed by heat inactivation and simultaneous *in vitro* transcription/uracil-cleavage reactions. Products were desalted and spotted on a 384-pad SpectroCHIP (Agena) using a MassARRAY nanodispender (Samsung). Spectra were acquired using a MassArray MALDI-TOF MS and peak detection and quantitative CpG methylation carried out using the EpiTyper software v1.2 (Agena).

2.2.18. siRNA Knockdown/miRNA Ectopic Expression

Myoblast and myotube cultures from HSSe were transfected with siRNAs or miRNA mimics at passage 5. 10pmol siRNA against H19 (Ambion) and 10pmol miR-675-3p/5p miRNA mimics (Ambion) were transfected using the Lipofectamine RNAiMAX Transfection Reagent (Life

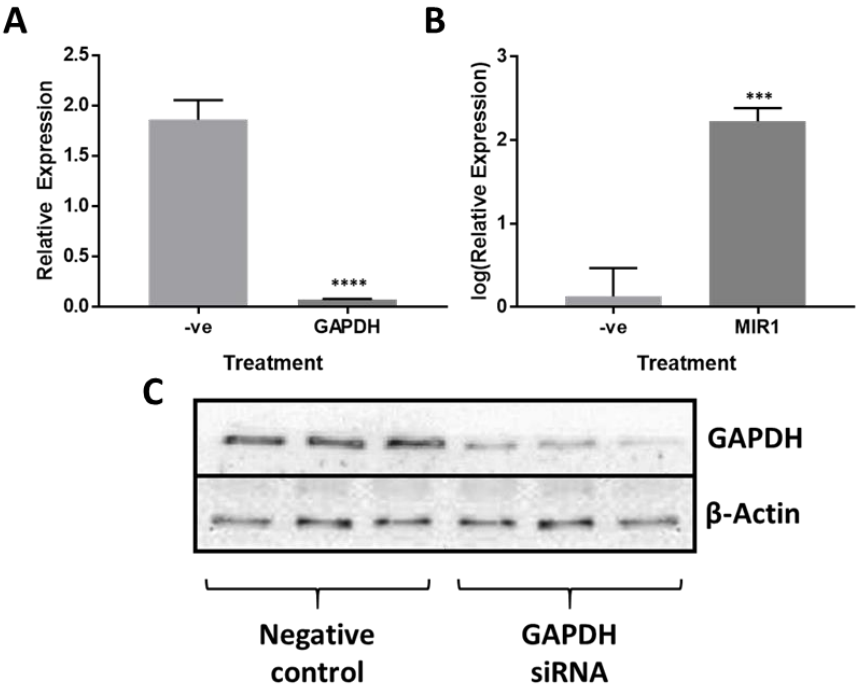


Figure 2.9: Results from the transfection controls

mRNA expression levels of GAPDH (A) and miR-1 (B) after transfection of GAPDH siRNA and miR-1 mimic respectively, to give an indication of transfection efficiency. GAPDH siRNA results in a significant decrease in its mRNA, as well as its protein expression (C), whereas miR-1 mimic induces a significant increase in the levels of miR-1. Paired t-test analysis, ****p<0.0001, ***p<0.001.

Technologies) following the manufacturer's protocol. Briefly, myoblasts were grown to confluence in proliferation media, cells were washed and media changed to differentiation media (day 0). Lipofectamine:siRNA/miRNA complexes were added to the cells at day 1 for 72hrs. After 72hrs, media was changed for fresh differentiation media and cells allowed to differentiate for a further 2 days. At day 5, myotubes were harvested and frozen at -80^{O}C until RNA/protein extraction. For proliferating myoblast experiments, myoblasts were transfected at 60% confluence, and allowed to grow for 3 days before harvesting. GAPDH siRNA and miR-1 miRNA mimic were used as positive controls for the transfections. Results of the GAPDH and miR-1 transfections on myotubes is shown in figure **2.9a-b**. GAPDH siRNA significantly reduced the mRNA and protein level of GAPDH (figure **2.9a+c**) while miR-1 mimic significantly increased miR-1 expression (figure **2.9b**). This shows good transfection efficiency in the myotubes.

2.2.19. Western Blotting

Western blot analysis of total protein levels was performed using a 10% SDS PAGE polyacrylamide gel combined with gel electrophoresis. 20 µg of protein lysate was added to an equal volume of loading buffer (125 mM Tris (pH 6.8), 200 mM DTT, 4% SDS, 20% Glycerol and 0.2% bromophenol

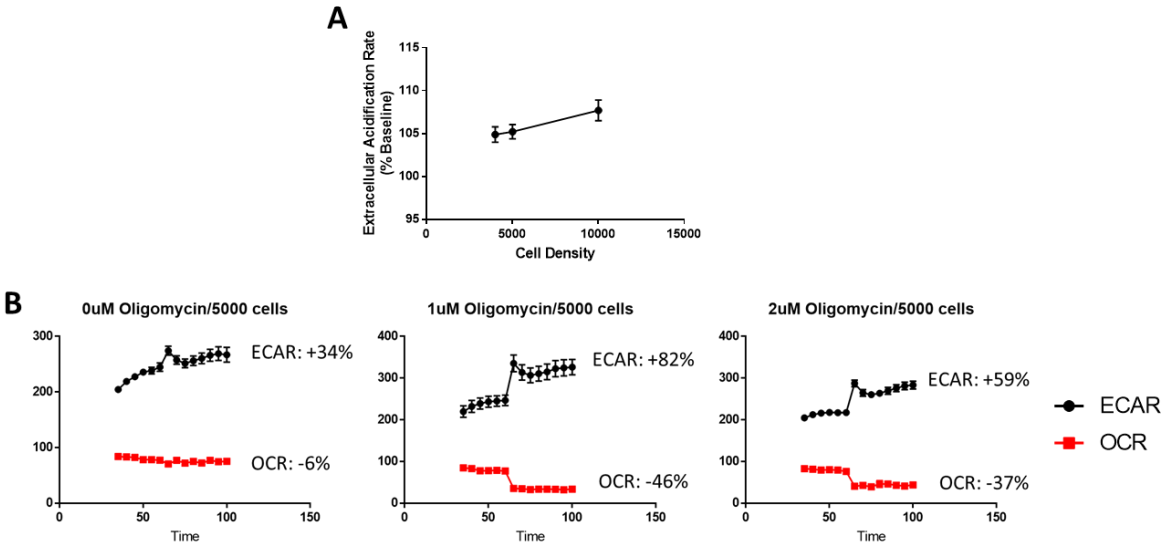


Figure 2.10: Seahorse XF96 assay optimisation results

(A) Optimal seeding density was chosen as a density in the linear range of ECAR increase, with 5000 cells per well chosen. (B) Optimal oligomycin concentration was chosen as the concentration that induced the greatest increase in ECAR as well as the greatest decrease in OCR, with 1µM oligomycin chosen.

blue). Denatured protein lysate was resolved by SDS PAGE and protein transfer was performed in buffer containing 25 mM Tris, 192 mM glycine, 20% (v/v) methanol and 0.1% (w/v) SDS. The membrane was blocked with 5% (w/v) dried skimmed milk powder (Marvel, UK) in 1 x TBS containing 0.05% (v/v) Tween-20. Membrane sections were incubated overnight at 4°C with primary antibody diluted in blocking buffer followed by incubation with horseradish peroxidase-conjugated secondary antibody. Protein bands were detected using SuperSignal West Pico Plus Chemiluminescent Substrate (Thermo Scientific, UK). Protein band intensities were analysed using ImageJ software from raw data images. β-actin was used as a reference protein to allow normalisation of protein loading.

2.2.20. Metabolic Flux Assay

Mitochondrial bioenergetics and glycolysis were measured in myoblasts and myotubes using the Agilent Seahorse XF96 Mito Stress Test and Glycolysis Stress Test. Briefly; myoblasts were plated onto XF96 culture plates and allowed to adhere overnight. Cells were washed in pre-warmed running media (DMEM + 1% pen/strep + 2mM L-glutamine, pH7.4) and incubated in running media for 1hr in a non-CO2 incubator to allow them to pre-equilibrate with the assay media. For the Mito Stress test, pre-warmed oligomycin, FCCP and rotenone & antimycin A were loaded into injector ports A, B and C of the sensor cartridge, respectively. For the Glycolysis Stress test, pre-warmed glucose, oligomycin and 2-deoxyglucose were loaded into injector ports A, B and C respectively. The

final concentrations of the injections were as follows: 1µM oligomycin, 3µM FCCP, 1µM rotenone & antimycin A, 5mM glucose and 50mM 2-deoxyglucose. Cell density and oligomycin concentration were empirically determined by running an initial glycolysis stress test with 3 cell densities and 3 concentrations of oligomycin (figure **2.10**). The optimal cell density is one in the linear range of a graph of cell density vs ECAR (figure **2.10a**). 5000 cells was chosen as being in the middle of the linear range. Optimal oligomycin concentration is one with maximal extracellular acidification rate (ECAR) increase with a corresponding maximal decrease in oxygen consumption rate (OCR). 1µM oligomyin showed the maximum increase in ECAR while decreasing OCR the most (figure **2.10b**).

The cartridge was subsequently calibrated by the XF96 analyser and the assay continued using either the Mito Stress test assay protocol or the Glycolysis Stress test assay protocol. OCR and ECAR were detected under basal conditions followed by the sequential injection of compounds. After completion of the run, cells were lysed and a protein assay carried out to allow for normalization.

2.2.21. Immunocytochemistry

Dr Emma Garratt carried out all immunocytochemistry experiments, with optimisation carried out by both Dr Emma Garratt and Elie Antoun. Myogenic purity of sorted cell cultures was measured using immunocytochemistry with CD56 as a marker. For analysis of differentiation potential, cells were analysed for MYHC expression at differentiation day 2, day 5 and day 10. Briefly, cells were fixed in 3.7% PFA for 10 minutes at room temperature. Following a PBS wash, cells were then permeabilised (0.3% Triton X-100, 1% BSA in PBS) for 7 min at room temperature. Cells were washed in PBST and blocked (5% goat serum, 1% BSA in PBS) for 1 hr at room temperature. Cells were next washed and incubated with the primary antibody overnight at 4^0C with agitation. Cells were washed and incubated with secondary antibodies for 1hr at room temperature, in the dark. To visualise nuclei, cells were incubated in DAPI 1ug/ml.

Cells were imaged using the Axio observer D1 microscope (Zeiss). Five images of randomly selected non-overlapping fields were captured for each sample. Cells on raw unadjusted images were counted using the 'Count' feature of ImageJ. To calculate the fusion index, the number of nuclei incorporated into the myotubes (containing 2+ nuclei) was counted and the ratio of this number to the total number of nuclei was determined.

2.2.22. BrdU Proliferation Assay

Dr Emma Garratt carried out BrdU proliferation assay. Cellular proliferation was determined using the Cell Proliferation ELISA BrdU kit (Sigma) according to manufacturer's instructions. Cells were analysed in triplicate and plated out at 1,000 cells/cm2. The assay was initiated when the cells were actively proliferating and no more than 70% confluent.

2.2.23. β-Galactosidase Senescence Assay

Senescence was determined using a senescence detection kit (Abcam) according to manufacturer's instructions. Cells were plated out at 1,000 cells/cm^2 and analysed when the cells were actively proliferating and no more than 70% confluent. Cells were imaged using the Axio observer D1 microscope (Zeiss). Three randomly selected non-overlapping fields were counted for each sample, including counts for both senescent and non-senescent cells.

2.2.24. Statistics

All bioinformatical analysis was carried out in R (v3.2.1 (total RNAseq) or v3.4.2 (DropSeq and methylation arrays)). All other statistical analysis was carried out in SPSS (v24), unless otherwise specified. Graphs were created using the ggplot2 package in R or GraphPad Prism 7, unless otherwise specified. An FDR<0.1 or a p-value<0.05 was classed as statistically significant.

Table 2.1: *Primers used for DropSeq experiments*

DropSeq Primers	
Template Switch oligo	AAGCAGTGGTATCAACGCAGAGTGAATrGrGrG
SMART PCR primer	AAGCAGTGGTATCAACGCAGAGT
P5 SMART PCR hybrid oligo	AATGATACGGCGACCACCGAGATCTACACGCCTGTCCG CGGAAGCAGTGGTATCAACGCAGAGT* A*C
Custom Read 1 primer	GCCTGTCCGCGGAAGCAGTGGTATCAACGCAG AGTAC

Table 2.2: *List of primers used in RT-PCR and methylation*

RT-PCR Primers

Gene		Sequence	Catalogue number
H19	Forward	GTGACAAGCAGGACATGACAT	
	Reverse	CCGTGGAGGAAGTAAAGAAACA	
	Probe	CCGCAGCCCCACCAGCCTAAGG	
Unspliced Anril	Forward	CAGTGGCTTCCTGTTCATGC	
	Reverse	GGGCTTGACGTCTGATCTGT	
CDKN2A		Quantitect Primer assay	QT01866137
PPIA		Quantitect Primer assay	QT00209454
CYC1		Quantitect Primer assay	QT00103824
SMAD1		Quantitect Primer assay	QT00031360
SMAD5		Quantitect Primer assay	QT00079247
GAPDH		miScript Primer assay	MS00032109
miR-675-5p		miScript Primer assay	MS00032102
miR-675-3p		miScript Primer assay	MS00008358
miR-1		miScript Primer assay	MS00033740
RNU6		miScript Primer assay	MS00033705
SNORD61		miScript Primer assay	MS00033712
SNORD68		miScript Primer assay	MS00033733
SNORD96A		Reference Gene assay	HK-SY-hu-600
B2M		Reference Gene assay	HK-SY-hu-600
RPL13A		Reference Gene assay	HK-SY-hu-600
SDHA		Quantitect Primer assay	QT01866137

Pyrosequencing Primers

Gene		Sequence	Catalogue number
CDKN2A			
CpG3-1	forward	TTTATTTTTTTGGAAGGTGGGAGAGG	
	reverse	[Btn]TTTTTTCTCCCCAACCTCCC	
CpG1	sequencing	GGAGGTTGGGGAGAA	
CpG3-2	sequencing	TGGGAGAGGGTGATT	
CpG9-4	forward	GGTTGTAAAATTTTTTTAAGAAGTAAGTG	
	reverse	[Btn]TTTTTTCTCCCCAACCTCCC	
CpG7-4	sequencing	GATTTTTGTTAGTATTTTAGGA	
CpG8	sequencing	GTAGTAGTAATTGATT	
CpG9	sequencing	AGAAGTAAGTGTTTGTTTTT	

Sequenom Primers

Gene		Sequence	Catalogue number
H19			
Prom1	forward	aggaagagagGGTTTTTTTTAATTGGGGTGGT	
		cagtaatacgactcactatagggagaaggctTAAAACAATA	
	reverse	CTCAACCCTATCTCC	
Prom2	forward	aggaagagagGAGATTTGAGGTGAATTTTAGGGA	
		cagtaatacgactcactatagggagaaggctCAAAACAAAA	
	reverse	TCCCCACAACC	
CTCF3	forward	aggaagagagTTTTTGGTAGGTATAGAAATTGGGG	
		cagtaatacgactcactatagggagaaggctACACCTAACTT	
	reverse	AAATAACCCAAAAC	
CTCF6	forward	aggaagagagGGGTTGTGATGTGTGAGTTTGTATT	
		cagtaatacgactcactatagggagaaggctCACATAAATAT	
	reverse	TTCTAAAAACTTCTCCTTC	

Chapter 3 –

INK4a/ARF Locus Genes Correlate with Ageing and Inflammation in Human Skeletal Muscle

Chapter 3 – INK4a/ARF Locus Genes Correlate With Ageing and Inflammation in Human Skeletal Muscle

3.1. Introduction

The 9p21 locus in humans is the strongest identified susceptibility locus for risk of coronary artery disease in the population [378]. It has subsequently been linked to other pathologies, including cancer [379], type 2 diabetes [380] and frailty in the older people. [381].

This locus encodes multiple genes (figure **3.1**), including three protein coding genes, $p16^{INK4a}$ (CDKN2A), $p15^{INK4b}$ (CDKN2B) and $p14^{ARF}$; and the long non-coding RNA (lncRNA) ANRIL (CDKN2B-AS, **A**ntisense **N**on-coding **R**NA in the **I**NK4a **L**ocus). $p16^{INK4a}$ and $p14^{ARF}$ share the same exon 2 and 3, however have different first exons; exon 1α and 1β in $p16^{INK4a}$ and $p14^{ARF}$ respectively. The two first exons have their own promoters and $p14^{ARF}$ is transcribed in an alternative reading frame (ARF). As such, the two genes are not isoforms of one another and do not share any amino acid homology in their protein products. The third protein coding gene found in this locus is $p15^{INK4b}$, which is located upstream of exon 1β of $p16^{INK4a}$ and consists of 2 exons. The lncRNA ANRIL is transcribed in the antisense direction and consists of 19 exons that span around 120kb of genomic sequence. ANRIL is highly spliced into various linear and circular isoforms, each differentially expressed and exhibiting different functions.

3.1.1. INK4a/ARF Locus and the Cell Cycle

$p16^{INK4a}$ and $p15^{INK4b}$ are members of the INK4 protein family, which also includes $p18^{INK4c}$ and $p19^{INK4d}$ in humans. INK4 proteins bind to the cyclin-dependent kinases CDK4 and CDK6, inhibiting their kinase activity [177]. A key target for CDK4/6 is the retinoblastoma protein (pRb), which is found hypophosphorylated and bound to E2F family members in the cytoplasm when a cell is not actively dividing during G0 phase of the cell cycle. After becoming phosphorylated by CDK4/6, pRb dissociates from E2F, allowing it to translocate to the nucleus and activate the transcription of cell cycle genes required for progression through the G1-S checkpoint. When CDK4/6 are inhibited by INK4 proteins, pRb remains bound to E2F, preventing the transcription of key cell cycle genes and inducing cell growth arrest. The gene $p14^{ARF}$ inhibits the E3 ligase activity of MDM2 [382], preventing its ability to degrade p53 protein and activating the p53-mediated transcriptional network. An important downstream effector molecule of p53 is $p21^{CIP1}$, which when activated by p53 inhibits the activity of CDK2, preventing CDK2-mediated phosphorylation of pRb and progression through the cell cycle. As well as inducing growth arrest, activation of p53 by $p14^{ARF}$ can also induce apoptosis and programmed cell death.

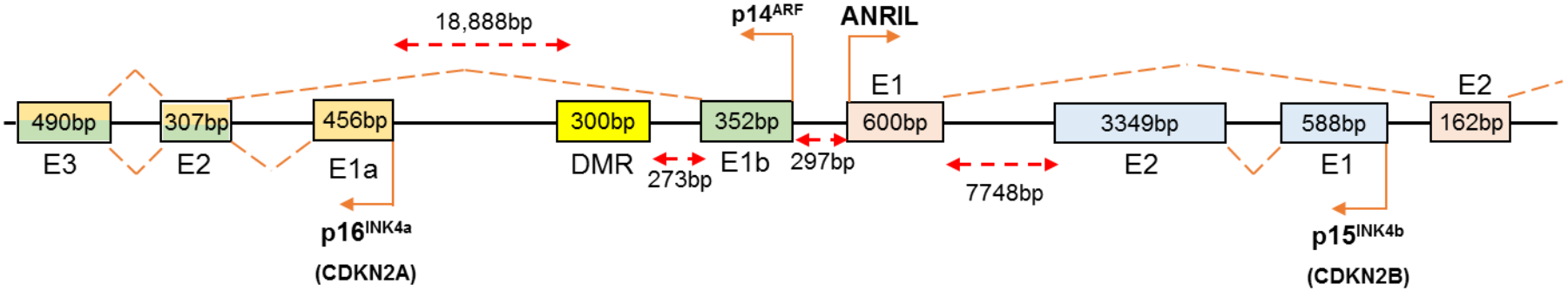


Figure 3.1: Schematic of the 9p21 susceptibility locus

Schematic of the 9p21 susceptibility locus showing the genomic organisation of the INK4A/ARF locus and the position of $p16^{INK4a}$, $p15^{INK4b}$, $p14^{ARF}$ and ANRIL. ANRIL extends further upstream from exon 2, with a total of 19 exons spanning a total of 123.3kb in the genome. ANRIL = Antisense Non-coding RNA in the INK4A/ARF Locus, bp = base pairs, E = Exon, DMR = Differentially Methylated Region, CDKN2A/B = Cyclin-Dependent Kinase Inhibitor 2A/B.

The lncRNA ANRIL also plays a role in regulating the cell cycle. In humans, ANRIL has been shown to promote cell proliferation by regulating the expression of genes from the INK4a/ARF locus [383,384]. ANRIL potently represses the transcription of $p15^{INK4b}$ and $p16^{INK4a}$ to a lesser extent, by recruiting the polycomb repressive complex (PRC)-1 and PRC2. ANRIL specifically associates with members of the PRC1 complex (e.g. BMI1 and CBX7), resulting the repression of the INK4a/ARF locus by induction of H2AK199 ubiquitination and other histone marks associated with polycomb-mediated repression [194]. As well as inhibition by ANRIL-mediated ubiquitination via PRC1, histone methylation by EZH2 also regulates $p16^{INK4a}$ expression, with EZH2 expression decreasing in senescent cells, with a decrease level of H3K27me3 and de-repression of $p16^{INK4a}$ expression [385]. As a result of ANRIL-mediated repression of the transcription of $p15^{INK4b}$ and $p16^{INK4a}$, as well as other epigenetic regulators, there is decreased inhibition of CDK4/6, allowing progression through the G1-S checkpoint and completion of the cell cycle.

3.1.2. INK4A/ARF Locus and Frailty

Frailty, in its complete definition is a syndrome of decreased reserve and resistance to stressors resulting from cumulative declines across multiple physiological systems causing vulnerability to adverse health outcomes including falls, hospitalisation, institutionalization and mortality. There are physical, social and psychological elements that comprise a frailty phenotype. From the molecular point of view, frailty may manifest as a consequence of a decline in the replicative capacity of tissues, which in turn has been suggested to occur due to the activation of tumour suppressor genes such as $p16^{INK4a}$. Frailty is associated with a decrease in physical activity and exercise. Negative correlations have been seen with both the frequency and the amount of exercise and $p16^{INK4a}$ expression levels [386]. $p16^{INK4a}$ expression has been shown to increase with age in various tissues in humans [194,205]. Ageing itself is associated with decreased functional capabilities, with reduced physical activity [387]. This suggests an interplay between ageing, $p16^{INK4a}$, physical activity and functional ability in the elderly, and suggests that molecular ageing is strongly associated with decreased exercise and physical activity.

Interleukin 6 (IL-6) is a proinflammatory cytokine with increased circulating levels in the elderly [388], that has previously been found to be associated with cellular senescence and frailty [389]. Liu et al. [386] have shown that the expression of $p16^{INK4a}$ is associated with serum levels of IL-6 in peripheral blood T-lymphocytes (PBTLs), suggesting that molecular age as predicted by $p16^{INK4a}$ expression may be a better predictor of frailty than chronological age. Further to this, Ferrucci et al. [390] report that increased IL-6 levels result in a higher risk for developing physical disability, with a greater decline in muscle strength and performance after a 3.5 year follow-up in community-dwelling older women. Others have also shown that increased IL-6 levels as a result of chronic systemic

inflammation correlates with decreased muscle strength and power [391], key aspects of the physical frailty syndrome.

ANRIL has previously been shown to be regulate IL6 levels [392]. Experiments in HUVEC cells have shown that ANRIL expression is required for YY1 binding to IL-6 and IL-8 promoter regions, to induce their transcription in response to TNFα stimulation. ANRIL and IL6 were also found to be correlated in PBMC samples from patients with coronary artery disease, a progressive inflammatory process for which ANRIL is the best implicated genetic susceptibility locus. This provides evidence for ANRIL playing a role in mediating inflammation, which is prevalent in the elderly, contributing to frailty and physical disability. Further evidence for ANRIL and the INK4A/ARF locus in frailty is seen from Melzer et al. [381], who found that in multiple independent populations of elderly people (EPIC, InCHIANTI and the Iowa-EPESE studies), the rs2811712 minor allele SNP was associated with reduced physical impairment [381]. This suggests that $p16^{INK4a}$ and the INK4A/ARF locus plays a role in the ageing processes in the general population, and the fact that $p16^{INK4a}$ is associated with increased senescence of skeletal muscle satellite cells [194] implicates the region with skeletal muscle ageing. However, whether ANRIL is associated with skeletal muscle ageing and $p16^{INK4a}$ expression in muscle is currently unknown.

3.1.3. Aims

In the present study, we sought to look at the relationship between $p16^{INK4a}$ and ANRIL with age in aged myoblast and muscle tissue samples from elderly participants of the Hertfordshire Sarcopenia Study (HSS) cohort. We also looked at whether the expression levels of $p16^{INK4a}$ and ANRIL, as determined by RT-PCR also correlate with the expression levels of the cytokine interleukin 6 (IL6), previously shown to be associated with frailty and ageing. We found that $p16^{INK4a}$ and ANRIL show a positive and negative correlation with age in myoblast samples respectively, whereas in muscle tissue samples, both show a positive correlation with age, suggestive of different roles for ANRIL in isolated myoblasts and muscle tissue. The methylation of CpG sites in the promoter region of the ANRIL genes was analysed using pyrosequencing, of which the methylation at 2 CpG sites was found to be associated with the expression of ANRIL in the myoblasts. This study provides evidence for a role for $p16^{INK4a}$ and ANRIL in mediating skeletal muscle ageing and ageing phenotypes, including inflammation and frailty.

3.2. Results

3.2.1. Cohort Characteristics

A total of 166 participants were enrolled into the HSS cohort. Of these, 43 (26%) were males and the rest were females. The mean age among the male and female participants was 77.53 ± 2.70

(mean ± standard deviation) and 77.96 ± 2.55 years respectively. The average age among all 166 participants was 77.85 ± 2.59 years, with no significant difference between the genders. Varying numbers of participants were used in each portion of the study, dependent on tissue availability and failures during RT-PCR and pyrosequencing.

3.2.2. Expression of the Senescence Marker, $p16^{INK4a}$, and ANRIL is Associated with Age

The expression levels of $p16^{INK4a}$ and ANRIL were explored in association with age. Expression was analysed with real-time PCR using Taqman assays against these 2 genes, and the data was analysed using the ΔΔCt method [393], being normalized to the geometric mean of the expression of β_2-microglobulin (B2M), succinate dehydrogenase A (SDHA) and ribosomal protein L13 A (RPL13A). Both $p16^{INK4a}$ and ANRIL were expressed in the aged myoblast samples. Even in the elderly, there is a significant increase in the expression of CDKN2A ($p16^{INK4a}$) between the ages of 73 and 83 years (β=1.628 (95%CI 0.541, 2.743), p = 0.005, r=0.274; figure **3.2a**, table **3.1**) after adjusting for the sex of the participant. These myoblasts also showed a decrease in the expression of unspliced ANRIL with older age (CDKN2B-AS) (β=-2.691 (95%CI -4.510, -0.873), p = 0.004, r=0.276; figure **3.2b**, table **3.1**). As the myoblasts originated from native muscle tissue, we next examined whether the expression levels of $p16^{INK4a}$ and ANRIL were altered in the native muscle tissue as well as the myoblasts. As expected, the levels of $p16^{INK4a}$ increased in the muscle tissue with increasing age (β=1.545 (95%CI 0.271, 2.819), p = 0.018, r=0.254; figure **3.3a**, table **3.1**). However, unlike in the myoblasts, the expression of the unspliced ANRIL isoform increased in the muscle tissue (β=1.529 (95%CI -0.085 3.144), p = 0.063, r=0.226; figure **3.3b**, table **3.1**). The increase in ANRIL was not statistically significant, but suggested a trend ($p < 0.1$) in the expression of ANRIL with age in the native muscle tissue. This is unlike what was seen in the myoblast samples from these same participants, where there was a decrease in ANRIL expression with increasing age.

In conclusion, we have shown that an increase in age is associated with an increased expression of the senescence marker $p16^{INK4a}$ in both myoblasts and skeletal muscle tissue. However, the expression of ANRIL with respect to age differs in the myoblasts and muscle tissue, decreasing in the myoblasts and increasing in muscle tissue. The age-associated difference in $p16^{INK4a}$ and unspliced ANRIL are not significantly associated with gender (appendix **A**).

3.2.3. ANRIL Correlates with the Expression of $p16^{INK4a}$ in Myoblasts but Not in Native Muscle Tissue

It has previously been shown that the different isoforms of ANRIL can inversely regulate the expression of $p16^{INK4a}$ from the INK4/ARF locus, repressing its transcription. We looked at whether

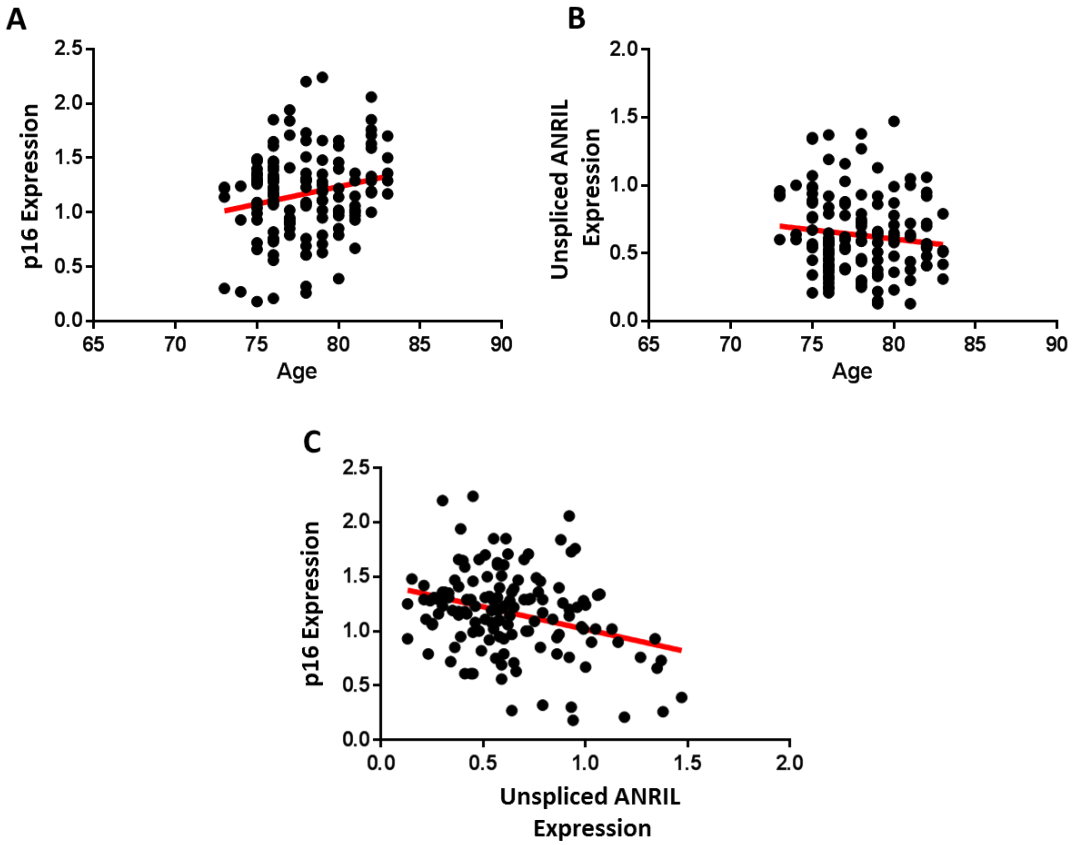


Figure 3.2: Expression of CDKN2A and ANRIL in the myoblasts

Expression of (a) p16 (CDKN2A) and (b) the unspliced ANRIL isoform in myoblast samples with respect to age. (c) p16 expression against ANRIL expression in the myoblast samples. RT-PCR carried out in 2 batches for p16 and anril, therefore results are adjusted for batch, as well as gender. Red line is the linear regression against age (a and b) and unspliced ANRIL (c).

there was a relationship between the expression levels of $p16^{INK4a}$ and unspliced ANRIL. In the myoblasts, as the levels of ANRIL increase, there is an equivalent decrease in the level of expression of $p16^{INK4a}$ ($r = -0.231$, $p = 0.007$; figure **3.2c**, table **3.2**) after adjusting for sex of the participant. This is in agreement with the expression of $p16^{INK4a}$ and ANRIL in the myoblasts with respect to age, where there is an expectation of an inverse relationship between ANRIL and $p16^{INK4a}$ expression due to their opposing directions of change with respect to increasing age.

The relationship between $p16^{INK4a}$ and ANRIL in the native muscle tissue differs to that in the myoblast samples from the same participants. With increasing levels of unspliced ANRIL, there is a comparable increase in the expression level of $p16^{INK4a}$ ($r = 0.261$, $p = 0.008$; figure **3.3c**, table **3.2**). This is in the opposite direction to that seen in the myoblast samples, despite the myoblasts originating from muscle biopsies from the same participant.

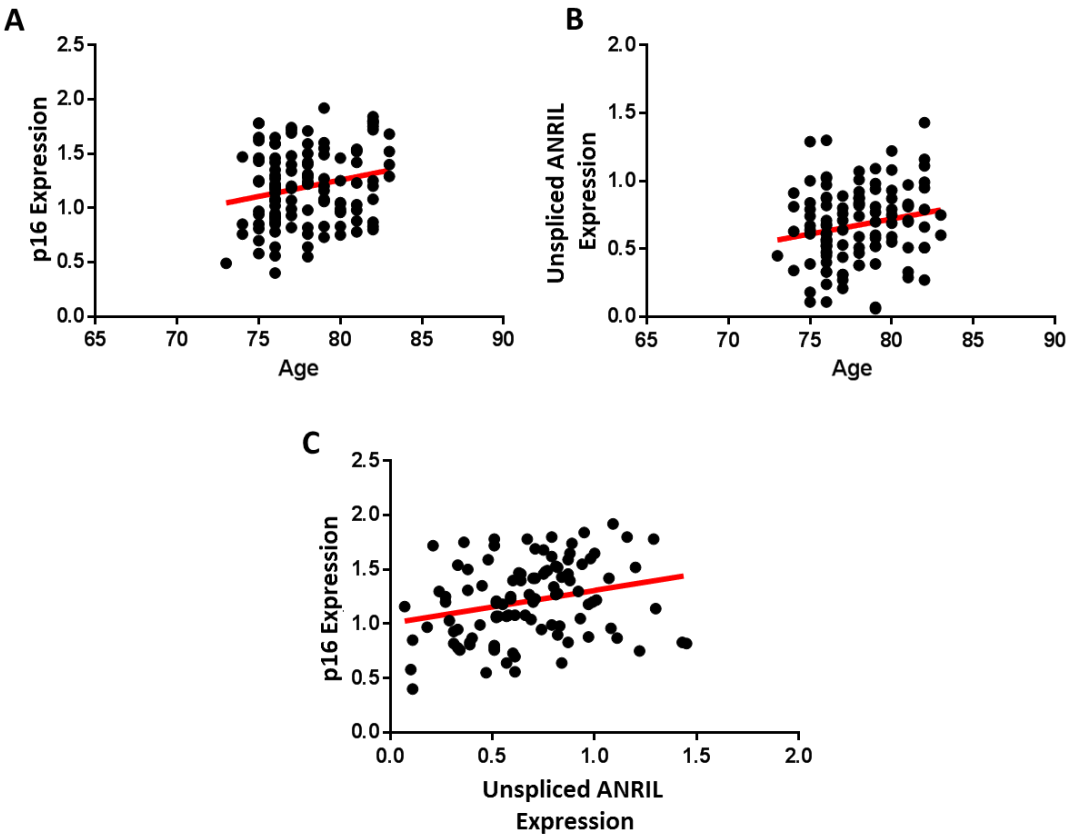


Figure 3.3: Expression of CDKN2A and ANRIL in the muscle tissue

Expression of (a) p16 (CDKN2A) and (b) the unspliced ANRIL isoform in native muscle samples with respect to age. (c) p16 expression against ANRIL expression in the muscle samples. Results are adjusted for gender of the participant. Red line is the linear regression against age (a and b) and unspliced anril (c).

3.2.4. Expression of ANRIL and p16^{INK4a} in Myoblasts Correlates with Expression in Muscle Tissue

We next sought to investigate whether the expression of p16^{INK4a} and ANRIL was correlated in the myoblasts and matched muscle tissue. Muscle tissue samples with a higher expression of p16^{INK4a} also show a high expression of p16^{INK4a} in myoblasts. Therefore, with increased expression of p16^{INK4a} in muscle tissue, there was a similar increase in expression of p16^{INK4a} in the myoblasts ($r = 0.201$, $p = 0.057$; figure **3.4a**, table **3.2**). Although this correlation is not statistically significant, the p value is marginally above significance, suggesting a statistical trend ($p < 0.1$) between the expression of p16^{INK4a} in the native muscle tissue and the myoblasts.

Similar to p16^{INK4a} which sees a positive correlation between expression levels in the muscle tissue and myoblasts, expression of the unspliced ANRIL isoform also shows a positive correlation between its expression in the muscle tissue and myoblasts ($r = 0.184$, $p = 0.062$; figure **3.4b**, table

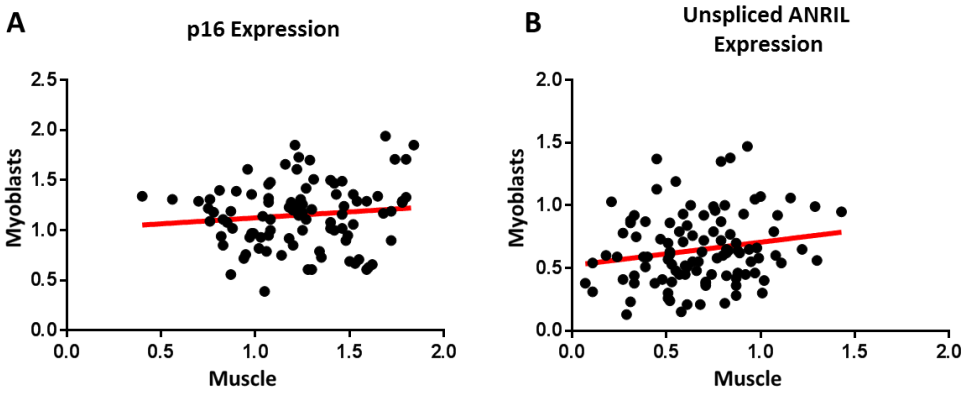


Figure 3.5: Correlation between CDKN2A and ANRIL in the myoblasts and muscle tissue

Expression of (a) p16 (CDKN2A) and (b) the unspliced ANRIL isoform in the myoblasts against the native muscle tissue. Results are adjusted for gender and batch as myoblast RT-PCR was carried out in 2 batches. Red line is the linear regression of the expression in the myoblasts against the expression in the native muscle tissue.

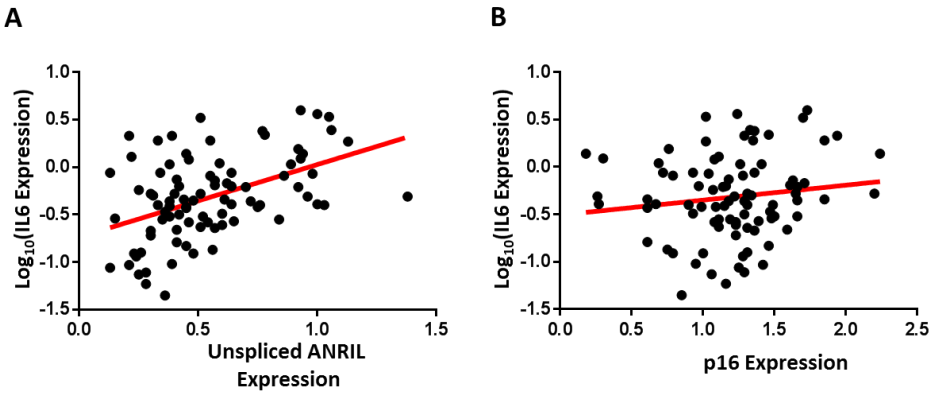


Figure 3.4: Correlation between IL6 and CDKN2A/ANRIL in the myoblasts

Expression of IL6 in the myoblasts against the expression of (a) the unspliced ANRIL isoform and (b) p16 (CDKN2A). IL6 expression values were $\log_{10}$ transformed to satisfy the assumption of normality of the data. Results were adjusted for batch and gender and the red line is the linear regression against unspliced ANRIL (a) and p16 (b) expression.

3.2). Therefore, in samples with high levels of unspliced ANRIL expression in the muscle tissue, there is also a high expression in the myoblasts. Once again, although this correlation is not statistically significant, it is only slightly above significance, suggesting a trend ($p < 0.1$) between the expression of the unspliced ANRIL isoform in the native muscle tissue and the myoblasts. This however is contradictory to the results seen with respect to age, where in the myoblasts we see a decrease of ANRIL expression with age, whereas in the muscle tissue, there is an increase in ANRIL expression with age. Therefore, it would be expected that a negative correlation would be seen between the ANRIL expression in the muscle tissue and the myoblasts.

3.2.5. ANRIL May Regulate the Expression of the Frailty Marker and Pro-Inflammatory Cytokine, IL6

Interleukin 6 (IL6) is a pro-inflammatory cytokine that plays a role in age-associated chronic inflammation and is secreted by multiple cell types including senescent cells. ANRIL has previously been shown to interact with the IL6 promoter, activating its transcription in peripheral mononuclear blood cells (PBMCs) [392]. Therefore, we looked at whether the expression of IL6 in the myoblasts correlated with the expression of ANRIL and $p16^{INK4a}$.

In the myoblasts, with increased ANRIL expression, there was a significant increase in the expression of IL6 (β=1.127 (95% CI 0.304, 1.950), p = 0.008, r=0.3454; figure **3.5a**, table **3.2**). As IL6 is also secreted by senescent cells as part of the senescence-associated secretory phenotype (SASP), we looked at IL6 expression with respect to the levels of $p16^{INK4a}$. As the levels of $p16^{INK4a}$ increased, suggesting an increased level of senescence, there was an increase in the expression of IL6 (β=0.591 (95% CI 0.142, 1.040), p = 0.010, r=0.448; figure **3.5b**, table **3.2**). Although IL6 levels correlate with the expression levels of $p16^{INK4a}$ and unspliced ANRIL in the myoblasts, the expression of IL6 was not associated with age (β=0.028 (95% CI -0.037, 0.093), p = 0.402, r=0.367; table **3.1**). As $p16^{INK4a}$ can be viewed as a marker of molecular age, it appears IL6 better correlates with molecular age as opposed to chronological age in the myoblasts, implicating molecular age as a better marker of frailty than chronological age.

Table 3.1: *Linear regression analysis results*

	N	r	r^2	B	95% CI	P value
Myoblasts						
p16 vs Age	137	0.273	0.075	1.628	0.541,2.743	***0.005*
Unspliced ANRIL vs Age	137	0.275	0.076	-2.691	-4.510, -0.873	***0.004*
IL6 vs Unspliced ANRIL	88	0.694	0.482	0.186	0.084, 0.287	****0.0005*
IL6 vs p16	88	0.377	0.142	0.247	0.055, 0.440	**0.012*
IL6 vs Age	88	0.067	0.005	0.24	-1.084, 1.565	0.719
Muscle Tissue						
p16 vs Age	121	0.254	0.065	1.545	0.271, 2.819	**0.018*
Unspliced ANRIL vs Age	112	0.226	0.051	1.529	-0.085, 3.144	*.0.063*

Table 3.2: *Pearson correlation analysis results*

	N	Pearson correlation coefficient	P value
Myoblasts p16 vs Unspliced ANRIL	137	-0.231	***0.007*
Muscle p16 vs Unspliced ANRIL	105	0.261	***0.008*
p16 Myoblasts vs Muscle	91	0.201	*.0.053*
Unspliced ANRIL Myoblasts vs Muscle	102	0.184	*.0.062*

3.2.6. Methylation at the ANRIL Promoter Correlates with ANRIL Expression in Aged Myoblasts

In order to examine whether the levels of DNA methylation in the promoter region of ANRIL were associated with the expression levels of ANRIL, we analysed the methylation of nine CpG sites located within a 300bp region of the ANRIL promoter, located approximately 900bp upstream of the transcription start site of the ANRIL genes. This site has previously been shown to be an important regulatory region, regulating the expression of ANRIL.

CpG2 shows a significant positive correlation with the expression of ANRIL (β=0.050 (95% CI 0.022, 0.079), p = 0.001, r=0.346; table **3.3**). However, the methylation status at CpG2 for 56% of the samples was 100%. The methylation at CpG9 shows an inverse association with the expression levels of unspliced ANRIL in the myoblasts (β=-0.084 (95% CI -0.162, -0.006), p=0.035, r=0.267; table **3.3**). The methylation at CpG3 also suggests a trend, showing a positive correlation with the expression of ANRIL (p = 0.056, 95% CI = -0.064, 4.860; table **3.3**). The overall methylation across all 9 CpGs is relatively high, ranging from 72% methylated at CpG4 to 100% methylation at CpG2, however, it may only require a small decrease in the methylation status in order to have an impact on the expression levels of ANRIL.

Table 3.3: *DNA methylation analysis of the ANRIL DMR*

	hg38 coordinates	B	95% CI	P value
CpG1	chr9:21,993,722	-0.249	-2.983, 2.485	0.857
CpG2	chr9:21,993,698	1.565	0.109, 3.020	**0.035*
CpG3	chr9:21,993,695	2.398	-0.064, 4.860	*.0.056*
CpG4	chr9:21,993,655	0.687	-2.479, 3.854	0.667
CpG5	chr9:21,993,646	-0.333	-3.442, 2.775	0.832
CpG6	chr9:21,993,639	-0.337	-1.661, 0.984	0.614
CpG7	chr9:21,993,630	0.160	-0.082, 3.401	0.922
CpG8	chr9:21,993,604	0.103	-1.133, 1.338	0.869
CpG9	chr9:21,993,584	-0.655	-1.316, 0.006	*0.052*

3.2.7. Expression of $p16^{INK4a}$ and ANRIL in the myoblasts correlates with measures of muscle mass and function

As cellular senescence is a pathway known to contribute to the loss of muscle mass and function in the elderly, contributing to sarcopenia, we next looked at whether the expression of $p16^{INK4a}$ and ANRIL in the myoblasts showed an association with measures of muscle mass and muscle function (figure **3.6**).

p16^{INK4a} shows a significant negative association with ALMI (β=-0.361 (95%CI -0.711 -0.001), p=0.044, r=0.545, figure **3.6a**), with those showing high p16^{INK4a} expression in the myoblasts exhibiting a low muscle mass. A significant association is also seen between ALMI and ANRIL expression (β=0.886 (95%CI 0.024 1.748), p=0.044, r=0.537, figure **3.6b**), with those with a low muscle mass showing high ANRIL expression in the myoblasts. The expression of p16^{INK4a} also showed a significant negative association with gait speed (β=-0.126 (95%CI -0.219 -0.033), p=0.044, r=0.356, figure **3.6c**), with those exhibiting a high gait speed showing low p16^{INK4a} in the myoblasts. ANRIL expression however does not show a significant association with gait speed (β=-0.327 (95%CI -0.007 0.272), p=0.063, r=0.356, figure **3.6d**), although it does suggest a trend between ANRIL expression in the myoblasts and gait speed in the elderly. p16^{INK4a} and ANRIL expression showed no association with grip strength.

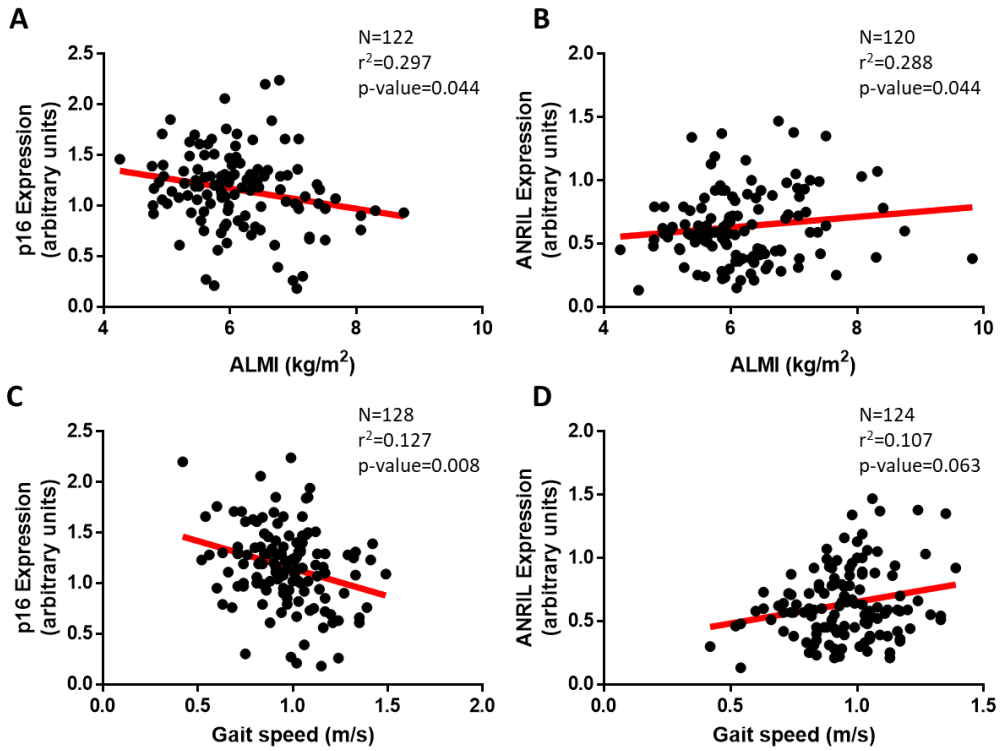


Figure 3.6: Associations between p16INK4a and ANRIL expression and ALMI and gait speed

Association between of p16^{INK4a} expression in the myoblasts and ALMI (a) and gait speed (c), and the association between ANRIL expression in the myoblasts and ALMI (b) and gait speed (d). Results were adjusted for batch and gender and the red line is the linear regression.

3.3. Discussion

Skeletal muscle ageing and the mechanisms contributing to skeletal muscle ageing are multifaceted and require the interplay of multiple different pathways. Here, we show that the genes transcribed from the INK4a/ARF locus are involved in skeletal muscle ageing, contributing to the ageing phenotype. Both $p16^{INK4a}$ and ANRIL expression were found to be differentially associated with ageing in myoblasts and native muscle tissue, suggesting different functions for these genes in the different tissues. To determine what may be regulating the expression of ANRIL in the myoblasts, we looked at the methylation of 9 CpG sites in the promoter region of ANRIL, and we found the expression of ANRIL to be associated with the methylation at 2 of these CpG sites, suggesting a role for DNA methylation in regulating ANRIL expression, which in turn may regulate the expression of $p16^{INK4a}$ and cellular senescence.

3.3.1. Geroconversion from Quiescence to Senescence in Myoblasts and Skeletal Muscle Tissue

Reversible quiescence is a hallmark of mammalian stem cells, playing a key role in the regenerative capacity of tissues during postnatal life. Regulation of quiescence requires the repression of genes associated with senescence as well as tumour suppressor genes. Sousa-Victor et al. [194] showed in mice that the number of satellite cells decreased in old mice compared to adult mice, but it did not drop further in older, sarcopenic mice. This suggests that a loss of satellite cell number cannot fully explain the loss of regenerative capacity seen with age or in sarcopenia. Instead, intrinsic changes are induced in the satellite cells by ageing. It is shown that in geriatric mouse myoblasts, there is a derepression of the INK4a/ARF locus through the loss of H2AK199 ubiquitination, increasing $p16^{INK4a}$ expression [179,194]. In this study, we have shown that in aged human muscle and myoblasts, $p16^{INK4a}$ expression is associated with increased chronological age. The increased expression of $p16^{INK4a}$ in the myoblasts and muscle tissue suggests that there is reduced tissue regenerative capacity in human muscle with age, resulting in an inability to repair muscle damage with age, leading to muscle loss and subsequent loss of muscle function. These differences in expression of $p16^{INK4a}$ suggest an increased level of senescence in skeletal muscle myoblasts with age, which agrees with senescence playing a role in the pathogenesis of sarcopenia.

Myoblasts are the primary progenitor cells in skeletal muscle [9], remaining in a reversible quiescent state. Myoblasts express paired box 7 (Pax7) protein, showing limited gene and protein expression until they are stimulated to proliferate and differentiate into mature myotubes. Without an adequate pool of resident Pax7+ cells, skeletal muscle is unable to repair itself following damage, suggesting that myoblasts and Pax7+ cells are indispensable for skeletal muscle repair following injury [70,394]. Here we show an increase in $p16^{INK4a}$ expression with increasing age in the myoblasts, suggesting an increased transition from reversible quiescence to irreversible senescence in these

cells. Reversible quiescence involves cells not actively dividing, however upon stimulation, are able to re-enter the cell cycle and proliferate and differentiate into muscle progenitor cells, including myoblasts. Upon high levels of $p16^{INK4a}$ expression, cells begin to irreversibly exit the cell cycle and enter a senescent state. As such, this reduces the regenerative capacity of the tissue, resulting in a decreased efficiency for repair and regeneration following stress from trauma (e.g. from exercise or injury). This may provide a possible explanation as to why we see a loss of skeletal muscle mass with age, owing to an inability to efficiently regenerate.

As well as seeing an increase in $p16^{INK4a}$ expression in the myoblasts with age, we also see a similar increase in skeletal muscle tissue. Although some of the expression may be originating from resident myoblasts and satellite cells that have transitioned into an irreversible senescent state, this suggests that the whole tissue is becoming senescent. As skeletal muscle function is highly dependent on the interplay of multiple different cell types that reside in the tissue, this increased $p16^{INK4a}$ expression suggests that many of the cell types in skeletal muscle are becoming senescent, resulting in the reduced communication between the different cells types leading to an inability for skeletal muscle to function as expected. Although there is a reduced communication between the cells in skeletal muscle, there is an increased level in the SASP, a different form of cellular communication resulting in an increased inflammatory response. Unlike normal cellular communication, the SASP results in an increased secretion of pro-inflammatory cytokines, including IL6 and IL8, as well as a an increased secretion of matrix metalloproteases, which play a role in the cleavage and activation of pro-inflammatory cytokines, together with the breakdown of extracellular matrix components [174,175]. This results in the dampening down of normal communication in skeletal muscle and muscle function, and increased inflammation and muscle breakdown.

3.3.2. DNA Methylation May Regulate the Expression of ANRIL in Skeletal Muscle Myoblasts

DNA methylation is a chemical modification to the DNA conferring epigenetic regulation of gene expression. Canonically, DNA methylation induces the repression of gene expression by recruiting methyl-binding proteins and chromatin remodelling complexes (e.g. MECP2-Sin3a-HDAC complex), resulting in the acquisition of repressive histone marks at the locus. We looked at the methylation of 9 CpG sites in the promoter region of ANRIL. Methylation at different CpGs at this site has previously been shown to be associated with different diseases, including obesity and coronary artery disease.

Here we show that methylation at 2 of the CpG sites correlates with the expression of ANRIL in the myoblasts. Although falling just above the threshold for significance (p=0.052), an increase in

methylation at CpG9 was associated with a decrease in ANRIL expression. This is consistent with the canonical view of DNA methylation, where methylation results in the repression of gene expression. The methylation at this DMR has previously been shown to be positively associated with ANRIL expression in umbilical cord tissue from participants of the Growing Up in Singapore Towards healthy Outcomes (GUSTO) cohort [395]. Although this differs to the results seen in the GUSTO cohort, this highlights that methylation at CpG9 may differentially regulate the expression of ANRIL *in vivo* in different cell types and tissues, resulting in altered $p16^{INK4a}$ expression and cellular senescence. Methylation at CpG3, although also falling just above the threshold for significance (p=0.056), also correlates with ANRIL expression. However, increased methylation at CpG3 is associated with an increase in ANRIL expression, which is in concordance with the results in seen in the GUSTO cohort. This however challenges the traditional view that DNA methylation represses gene transcription. Although not consistent with the canonical mechanism of DNA methylation, a positive correlation between DNA methylation and gene expression has recently been shown to play a functional role in human tissues [396]. In the GUSTO cohort, binding of the oestrogen receptor is enhanced by DNA methylation at CpGs 2-3 [395]. The oestrogen receptor is a transcription factor that subsequently binds to the methylated promoter region, inducing transcriptional activation. It may be that DNA methylation at this site allows the binding of transcription factors that preferentially binds to methylated DNA, allowing the initiation of transcription.

Although the overall change in DNA methylation is relatively small (<20% change), these subtle changes in DNA methylation may prove to be functional and contribute to complex diseases [397] such as sarcopenia and muscle loss. The majority of studies report small subtle changes in DNA methylation showing an association with various different outcomes, most notably the effect of the *in utero* environment on later life diseases, as shown from the Dutch Winter Hunger study [309]. However, the biological and clinical relevance of small DNA methylation changes is still unclear. Specific CpGs in an individual cell are either methylated or unmethylated. Therefore, reported DNA methylation levels in tissues is the average DNA methylation of all the individual cells in a sample. Although the total DNA methylation change may be relatively small, if the methylation change is occurring in key cells involved in a particular pathway, the switch from methylated to not methylated, or vice versa, can result in altered gene expression and pathway function. Several studies have shown that relatively small changes in DNA methylation are associated with changes in gene expression [398,399], with Murphy et al. [399] showing that a 1% decrease in methylation at the *IGF2* DMR is associated with a two-fold increase in expression of *IGF2*, which would be theoretically equivalent to the aberrant activation of the maternal allele. Therefore, small changes in DNA methylation can have a profound effect on gene expression and subsequent downstream signalling of pathways.

3.3.3. Older Myoblasts Exhibit the Senescence-Associated Secretory Phenotype

The senescence-associated secretory phenotype (SASP) has provided a new role for senescent cells in tissues. Although existing in a cell-cycle arrested state, senescent cells remain metabolically active, capable of altering the tissue microenvironment by secretion of various molecules [174]. Senescent cells secrete many different factors, including interleukins and other pro-inflammatory markers. The major cytokine of the SASP is interleukin 6 (IL6), a key proinflammatory cytokine. By secreting IL6, senescent cells can affect neighbouring cells due to IL6 binding to the IL6 receptor on the surface of neighbouring cells. IL6 can initiate an immune response by increasing the production of acute phase proteins, inducing the differentiation of activated B cells resulting in immunoglobulin production, the differentiation of T cells and natural killer cell activation. This contributes to the chronic systemic inflammatory phenotype seen in the elderly.

Here we show that in aged myoblasts, the expression of IL6 positively correlates with the expression of $p16^{INK4a}$, such that samples with high expression of $p16^{INK4a}$ also show a high level of expression of IL6. As $p16^{INK4a}$ is a key regulatory protein in inducing a senescent phenotype, this suggests that subjects exhibiting increased levels of senescence, as indicated by the levels of $p16^{INK4a}$ also show an increased expression of IL6. This may result in increased activation of inflammatory pathways in these people, contributing to the chronic inflammation associated with age.

IL6 levels have repeatedly been shown to be positively associated with a decline in measures of function in the elderly, independent of and to a greater degree than individual diseases [233,389]. Increased IL6 levels in the elderly may result in fatigue and decreased energy levels, leading to reduced functionality and overall frailty. Liu et al. [386] have previously shown an association between IL6 and $p16^{INK4a}$ in peripheral blood T lymphocytes. As IL6 has been found to be associated with functional declines typical of the frailty syndrome [388], Liu et al. [386] suggest that due to the correlation between IL6 and $p16^{INK4a}$ expression levels, and the fact that $p16^{INK4a}$ but not IL6 correlates with chronological age, molecular age as determined by $p16^{INK4a}$ may better indicate frailty than chronological age. We have found that in the myoblasts, IL6 expression positively correlates with the expression of $p16^{INK4a}$ but not with chronological age. This suggests that in the myoblasts, the molecular age of the myoblasts, the resident progenitor cells in skeletal muscle, is a better indicator of frailty and reduced muscle functionality than chronological age.

3.3.4. Novel Role for ANRIL in Skeletal Muscle

In endothelial cells (HUVECs and HCAECs), ANRIL expression has been shown to play a role in regulating the expression of IL6 and IL8 in response to stimulation by TNFα. Upon siRNA-mediated silencing of the short and long ANRIL isoforms, IL6 expression is inhibited, resulting in decreased

expression following stimulation by TNFα. It is known that ANRIL binds directly to Yin and Yang 1 (YY1) protein, a transcription factor that has been shown to mediate RNA binding to chromatin regions of target genes. YY1 is also enriched at the promoter region of IL6, whereas siRNA-mediated ANRIL knockdown leads to decreased YY1 binding to the IL6 promoter. This suggests that ANRIL facilitates the binding of YY1 to the IL6 promoter, which results in the induction of IL6 expression in endothelial cells. IL6 and ANRIL expression also show a positive correlation in PBMCs from patients with coronary artery disease (CAD).

We show that in the myoblasts, there is a positive correlation between the expression of ANRIL and the expression of IL6. This is in agreement with what was found by Zhou et al. [392] in patients with CAD. Although the 9p21 locus which contains the ANRIL gene is a strong risk indicator of CAD, here we show that similar relationships as those found in CAD patients are also found in skeletal muscle of elderly patients with varying degrees of loss of muscle mass. This suggests that the ANRIL-IL6 pathway may play an important role in skeletal muscle, with ANRIL increasing the expression of IL6, potentially being recruited to the IL6 promoter by direct binding to YY1. This increase in IL6 expression contributes to the pro-inflammatory phenotype seen in the ageing population, leading to an increased prevalence of frailty and skeletal muscle loss. In the myoblasts, we see a negative association between ANRIL and $p16^{INK4a}$ levels, as well as a positive correlation between ANRIL and IL6 and $p16^{INK4a}$ and IL6. This is contradictory, as both an increase in ANRIL and $p16^{INK4a}$ is associated with an increase in IL6 levels. This suggests that ANRIL is not the only mechanism regulating the expression of $p16^{INK4a}$ in the myoblasts. Therefore, other currently unknown mechanisms may result in an increase in $p16^{INK4a}$ levels, leading to increased senescence. We also only investigated the expression of unspliced ANRIL and its association with IL6 expression. Further investigation into the association of IL6 expression and the individual isoforms of ANRIL may provide a better understanding as to the role of ANRIL in regulating IL6 in skeletal muscle myoblasts. An increase in ANRIL subsequently leads to increased IL6 expression, contributing to the SASP seen in senescent cells. This provides a novel role for ANRIL in skeletal muscle, different to its role in regulating the expression of $p16^{INK4a}$ and $p15^{INK4b}$, tumour suppressor genes involved in regulating the cell cycle and senescence.

We do see a difference in the expression of ANRIL expression with age in the myoblasts and skeletal muscle tissue. In chapter 6, we investigated differences in DNA methylation using the 450K methylation array (Chapter 6). This was carried out to compare the methylation profiles of the myoblasts and matched skeletal muscle tissue (unpublished data). The Horvath ageing clock [353] was not conserved between the myoblasts and skeletal muscle tissue, however this could be a measure of the cellular heterogeneity of the tissue rather than true ageing. Therefore, the differences in ANRIL expression with age in the myoblasts and muscle tissue may be due to the cellular heterogeneity of the muscle tissue as well as the fact that we only analysed the unspliced isoform

of the gene. This suggests that ANRIL expression is being contributed to from the multiple different cell types present in muscle tissue and the different isoforms of ANRIL.

3.3.5. Further Considerations

Although we have shown here that age is associated with an increase in $p16^{INK4a}$ levels in skeletal muscle-derived myoblasts, suggesting an increase in senescence of the myoblasts, and that this may be regulated by the lncRNA ANRIL, there is still unanswered questions. The expression of ANRIL correlates with the methylation at several CpG sites in the promoter region, but it is unknown what causes the change in methylation. Further work is required to determine how the change in methylation results in altered ANRIL expression, and what transcription factors and proteins are involved. Here, we have shown that IL6 expression correlates with the expression of both $p16^{INK4a}$ and ANRIL, but further work is required to elucidate the SASP of the myoblasts, and how it contributes to the loss of muscle function. As IL6 is a key pro-inflammatory cytokine associated with age-associated chronic inflammation, further work is required to determine if/how ANRIL regulates the expression of IL6 in the myoblasts, providing a potentially novel role for ANRIL in the myoblasts. Finally, the number of samples used in this study is relatively small, and further work in a larger number of samples, preferably in an independent cohort of individuals would be required in order to validate the work and to show that the findings are seen in the general population. Several of the correlations and methylation changes observed in this investigation are suggestive of trends, while increasing the number of samples may show stronger significant associations between skeletal muscle and myoblasts and DNA methylation changes.

3.3.6. Conclusion

Muscle growth and maintenance is dependent on multiple different factors, one of the most important being the regenerative capacity of satellite cells and muscle progenitor cells in skeletal muscle. Here we have shown that, *in vitro*, there is an age-associated increase in a key cellular senescence marker in the myoblasts, potentially leading to the decreased regenerative capacity of skeletal muscle in the elderly. $p16^{INK4a}$ is known to be regulated by ANRIL, and here we show that ANRIL expression correlates with the expression of $p16^{INK4a}$. Further solidifying the concept of increased senescence in older skeletal muscle, we see a positive correlation between IL6 levels and $p16^{INK4a}$. IL6 is one of the key cytokines secreted by senescent cells, contributing to a pro-inflammatory phenotype. IL6 expression also correlates with ANRIL expression, suggesting that ANRIL may regulate the transcription of IL6, resulting in its increased secretion by senescent cells. All this contributes to the chronic inflammation seen in the elderly, contributing to skeletal muscle loss and the associated loss of function associated with sarcopenia.

Chapter 4 –

Genome Wide Changes in Skeletal Muscle Associated with the Severity of Sarcopenia

Chapter 4 – Genome Wide Changes in Skeletal Muscle Associated with the Severity of Sarcopenia

4.1 Introduction

The transcriptome is the complete set of transcripts in a cell for a specific developmental stage or physiological condition. The transcriptome encompasses not just coding RNAs (mRNAs), but also includes non-coding RNAs (tRNAs, rRNAs, lncRNAs, miRNAs, etc). Understanding the changes in the transcriptome under different physiological conditions and following external stresses is essential for interpreting the functional elements of the genome, and how the transcriptome contributes to normal development and disease.

4.1.1. The effect of ageing on the muscle transcriptome

No studies to date have looked at the global gene expression changes associated with sarcopenia in the elderly population. The effects of physical activity and other environmental factors have been broadly investigated; however, although studies looking at the contribution of normal ageing to skeletal muscle are available [290,291], studies investigating sarcopenia and the differential diagnosis of sarcopenia are limited.

Most studies to date have focussed on the effect of ageing on skeletal muscle, comparing young muscle with old. Welle et al. [291] carried out a microarray analysis looking for differentially expressed genes between young (21-27 years) and old (67-75 years) men. They found a range of genes that were significantly altered in the aged muscle compared to the young muscle. There was a decrease in genes encoding proteins involved in energy metabolism (e.g. succinate dehydrogenase subunit 3) in the old muscle, together with a decreased expression of genes involved in stress response, DNA repair and inflammation (e.g. IL-6 receptor and GADD45α). There was an increase in genes encoding hnRNPs and other RNA binding/splicing proteins (e.g. U5 snRNP 116kDa and cold induced RNA binding protein) and those encoding proteasome components and ubiquitin associated proteins (e.g. proteasome 26S subunit and ubiquitin specific protease 1), as well as an increase in the expression of genes encoding hormones, proteins involved in transcription and growth factors (e.g. follistatin, HDAC4 and IGF binding protein 6). The decrease in expression of genes associated with energy metabolism may contribute to the decline in mitochondrial function seen in ageing muscle, which is one of the possible contributors to the pathology of sarcopenia. Similarly, an increase in the expression of genes associated with the proteasome may point towards an increase in protein degradation. Since approximately 20% of muscle weight is attributed to protein, an increase in protein degradation may contribute to the loss of muscle mass seen with age and in

sarcopenia. Many of the changes seen in men were also seen in women [290], although there were some sex differences, pointing to possible gender-specific effects of ageing and sarcopenia.

Welle et al. [290] show a clear change in the transcriptome between young and old muscle. Although insight has been provided into the age-related changes in expression in skeletal muscle ageing, these changes tend to be difficult to interpret for several reasons. Many of the fold changes seen in whole genome transcriptome analysis in skeletal muscle tend to be relatively small (<1.5 fold change), and as such it may be the interplay of multiple genes in a common pathway and consistent changes in these genes playing a greater role in the disease in question. Studies often use an arbitrary cut-off for the mean fold change (>1.7 or >2) to define differential expression, which removes these genes. Skeletal muscle is a very heterogeneous tissue, containing many different cell types (e.g. fibroblasts, blood cells, connective tissue and adipose tissue). These different cell types may contribute to the gene expression profile, reflecting non-muscle specific changes in expression. Similar problems are seen with methylation studies, in which tissue heterogeneity may affect differential methylation analysis. As such, with methylation arrays, cell-type correction is often carried out to account for the different cells present in the tissue in question. The physiological age of skeletal muscle may provide a better indicator of muscle function and disease status, compared to chronological age. Therefore, while the genes analysed by Welle et al. **[291]** may provide an explanation as to why muscle loss occurs with age, they potentially only provide some insight the pathology of sarcopenia, not fully explaining why sarcopenia affects the elderly population with differing severities. Gene expression may explain only a proportion of the variability seen in skeletal muscle mass and strength, with other factors contributing significantly (e.g. age, gender, physical activity, nutrition, etc) as well as a proportion of the variance remaining unexplained.

Zahn et al. [160] provided a 250-gene signature for muscle ageing, finding an increased expression of genes involved in pathways regulating cell growth, complement activation and ribosomal and extracellular matrix genes, as well as a decrease in genes involved in mitochondrial oxidative phosphorylation. Zahn et al. [160] look at the type II/type I fibre diameter ratio to give an estimate of the physiological age of skeletal muscle and conclude that the age-related genes are enriched for those that predict physiological age, not just chronological age. Therefore, genes seen by both Welle et al. [290,291] and Zahn et al. [160] may play a role in both physiological and chronological ageing and provide insight into the pathology of sarcopenia.

People age at different rates, especially with respect to skeletal muscle ageing. Some people may remain fit and strong when they are old, whilst others live with increasing levels of frailty and require more support to maintain their independence For example, an 80 year old may be chronologically old; however, their physiological age may be 10-20 years younger based on their engagement in activities of daily living which in part requires a high level of muscle function. Few

studies have investigated the factors contributing to the pathology of sarcopenia in the elderly, or why there is variability in the rates of muscle loss in the elderly.

4.1.2. Aims

The aim of this investigation was to look at how the skeletal muscle transcriptome is altered in elderly participants with sarcopenia compared to healthy aged-matched controls. Most current studies look at the differences in skeletal muscle between young and old participants, with very few comparing sarcopenic and non-sarcopenic people. RNA-seq analysis will be used to map out the gene expression profiles between the two groups of individuals and look at differences to determine which genes and pathways may play a role in sarcopenia and measures of muscle mass and function.

4.2 Results

4.2.1. Cohort Characteristics

A total of 40 male participants from the HSS had RNA extracted from skeletal muscle obtained from biopsy and subjected to total RNAseq. 31 participants did not have any deficits in skeletal muscle mass, grip strength or gait speed (healthy controls), five were classed as pre-sarcopenics and four were classed as sarcopenics. A diagnosis of sarcopenia was made according to the EWGSOP [113] guidelines. Controls had normal appendicular mass index (ALMI), walking speed and grip strength. Pre-sarcopenia was defined as having low ALMI (<7.23kg/m^2) and normal muscle function. Sarcopenia was defined as having a low ALMI (<7.23 kg/m^2) and slow walking speed (≤0.8m/s) or low ALMI and decreased grip strength (<30kg). Participant characteristics can be seen in table **4.1**, with comparisons of the clinical variables in figure **4.1**.

4.2.2. RNAseq and QC Analysis

RNA from forty participants was processed and cDNA libraries were prepared. As cellular RNA is composed predominantly of ribosomal RNA (rRNA), samples were ribo-depleted to remove rRNA, allowing a good signal from other RNA species. All other RNA species were cDNA converted, amplified and purified, followed by sequencing.

Table 4.1: Characteristics of HSS participants that underwent RNAseq analysis

	Controls	Pre-sarcopenics	Sarcopenics
	n=31	n=5	n=4
Age	72.90 ± 2.42	71.84 ± 1.13	74.275 ± 3.59
BMI	27.98 ± 3.74	23.67 ± 0.49 *(p = 0.0123)*	26.40 ± 1.36
Height (m)	1.73 ± 0.06	1.69 ± 0.03	1.66 ± 0.08 *(p =0.0374)*
Weight (kg)	83.68 ± 13.77	67.78 ± 3.40 *(p = 0.0133)*	73.00 ± 8.65
Total Lean Body Mass (kg)	59.57 ± 6.78	49.76 ± 3.07 *(p = 0.00266)*	48.56 ± 4.45 *(p = 0.00266)*
ALM (kg)	24.65 ± 3.03	19.71 ± 1.14 *(p = 0.000821)*	18.63 ± 2.02 *(p = 0.000267)*
AMI (kg/m^2)	8.25 ± 0.75	6.89 ± 0.27 *(p=0.000198)*	6.74 ± 0.25 *(p=0.000188)*
Fat Mass (kg)	22.78 ± 7.76	16.85 ± 2.59 (p = 0.0936)	23.41 ± 6.25
Gait speed (m/s)	1.13 ± 0.16	1.07 ± 0.12	0.91 ± 0.34 *(p = 0.0309)*
Grip Strength (kg)	38.90 ± 6.28	35.00 ± 2.83	25.5 ± 6.61 *(p = 0.000169)*
6mTUG time (s)	10.52 ± 1.42	10.7 ± 1.60	14.36 ± 5.09 *(p = 0.00092)*
Chair rise time (s)	17.40 ± 4.14	17.72 ± 3.31	19.55 ± 5.67 †

Table 4.1: Characteristics of the participants of HSS that underwent RNAseq analysis. BMI = Body Mass Index (kg/m^2), ALM = Appendicular Lean Mass, AMI = Appendicular Mass Index (ALM/height^2), 6mTUG = 6 meter timed up-and-go. P values were calculated using the lm function in R. †Chair rise time was missing for one participant.

FastQC was run on fastq files at multiple stages. FastQC is a Linux program used to analyse the quality of the raw sequencing data. Files were first run through FastQC before concatenating. As is evident in appendix B, both the forward and reverse reads were good quality. All fastq files had similar FastQC outputs as those in appendix B. All samples had the Illumina Universal Adapter present at the end of the reads (appendix B **a** and **k**). None of the reads had an N base at any position (appendix B **b** and **l**). Per base sequence content is shown in appendix B **c** and **m**. In a random library, you would expect there to be no difference between the different bases of a sequence run, so lines should run parallel. However, libraries produced by priming with random hexamers, as is done with Illumina TruSeq library prep kits, have an intrinsic bias in the first 10bp of reads. This is not biased towards an absolute sequence and does not affect downstream analysis. Appendix B **d** and **n**, **e** and **o**, and **f** and **p** show overrepresented Kmers, per sequence GC content and sequence duplication levels respectively, and were consistent across all files. The average quality score distribution over all sequences is shown in appendix B **g** and **q**. Each base in a read generated by RNAseq is given a Phred quality score related to the base-calling error probabilities. A Phred (Q) score of 20 is equivalent to the probability of an incorrect base call 1 in 100 times, such that the base call accuracy

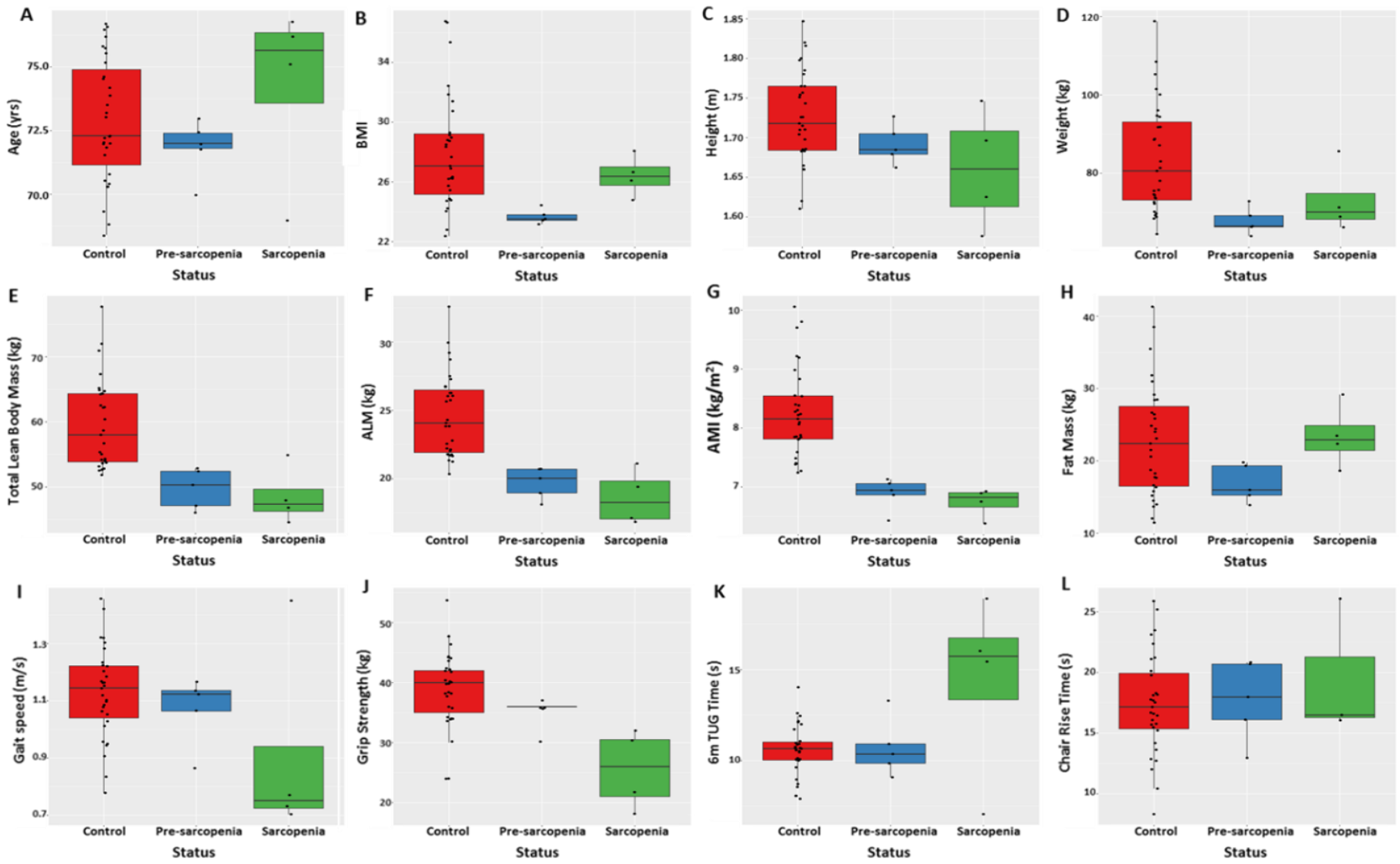


Figure 4.1: Clinical variables as a function of sarcopenia status

A = Age, B = BMI, C = Height, D = Weight, E = Total Lean Body Mass, F = Appendicular lean mass (ALM), G = Appendicular Mass Index (AMI), H = Fat Mass, I = Gait Speed, J = Grip Strength, K = 6m TUG Time, L = Chair Rise Time. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$. Red plots correspond to controls, blue plots correspond to pre-sarcopenics and green plots correspond to sarcopenics. Analysis was carried out in R (v3.2.1) and graphs generated using the ggplot2 package.

is 99%. In general, for both forward and reverse reads, the majority of the reads had a Q > 30, equivalent to > 99.9% accuracy. The reverse reads had more reads with an average Q < 10. Appendix B **h** and **r** show that all the reads were the same length (101bp). The quality of each tile of the flowcell was analysed (appendix B **i** and **s**) to determine if there was a loss of quality in one part of the flowcell. The colours are on a hot to cold scale, with hot colours indicating worse qualities than average. There was a good quality across the whole flowcell for all samples. Appendix B **j** and **t** show box and whisker plots of quality scores across all bases. The majority of the plots were in the green portion of the plot, indicating Q ≥ 28. Towards the end of each read, the general quality begins to fall, with whiskers in the red portion of the plot, indicating Q ≤ 20. However, the mean quality scores remained ≥ 28.

Reads subsequently underwent a combined adapter and quality trimming, trimming bases from the 3'- and 5'- ends with a Q ≤ 10. Forward and reverse reads were trimmed together to ensure the files remained in sync. Files underwent FastQC analysis and the results can be seen in appendix C. N content (appendix C **b** and **l**), per base sequence content (appendix C **c** and **m**), Kmer content (appendix C **d** and **n**), per sequence GC content (appendix C **e** and **o**), and per tile sequence content (appendix C **i** and **s**) remained unaltered. Adapters were removed from the reads (appendix C **a** and **k**), reducing sequence duplication levels (appendix C **f** and **p**). After quality trimming, no reads were present with a Q < 20 (appendix C **g** and **q**). The Phred score for bases in reads were all > 20 after trimming (appendix C **j** and **t**). As such, some of the reads were less than 101bp long (appendix C **h** and **r**) due to removal of adapters and poor quality bases, although the majority of reads remain at 101bp. This does not affect downstream analysis.

The average number of reads for each sample was 80 million (figure **4.3a**). Sample 25 had a reduced number of reads, with 48,108,950 reads, and sample 28 had 62,073,660 reads. This should not affect any differential expression analysis due to downstream normalization with respect to library size.

4.2.3. RNA-seq Alignment Metrics

Raw reads were aligned to the human genome version GRCh38, using TopHat2. Figure **4.3b** shows the number of concordant read pairs that successfully aligned. There was an average of 92.4% concordant pair alignment. Reads left unmapped due to no sequence match in the human genome were discarded.

Aligned files were run through the CollectRnaSeqMetrics script, part of the picard-tools software, in order to determine where the reads align in the genome. Figure **4.2** shows the percentage of bases aligning to rRNA, coding sequences, UTRs, introns and intergenic regions. The total number

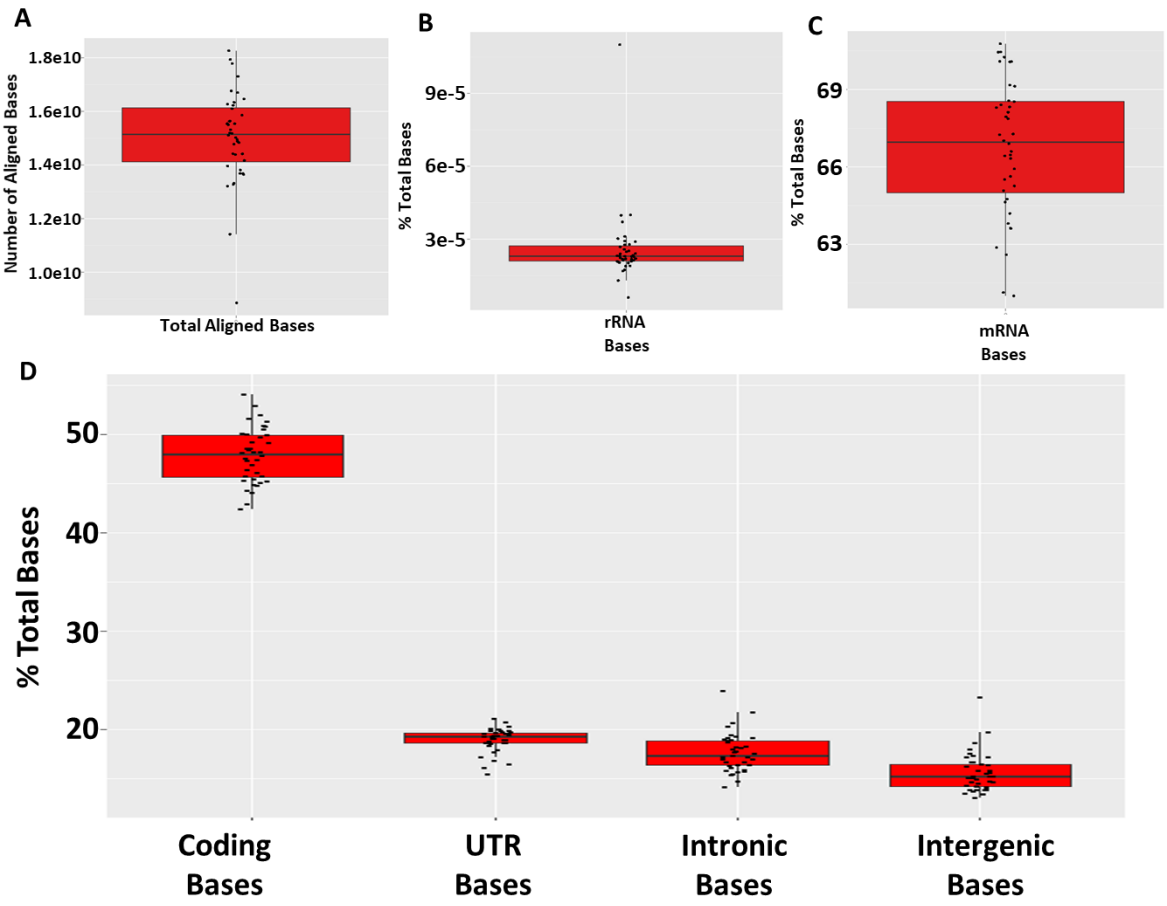


Figure 4.2: RNA-seq Alignment Metrics as determined by Picard tools

(a) Total Number of aligned bases; (b) rRNA bases; (c) mRNA bases; (d) Coding bases; (e) Untranslated region (UTR) bases; (f) Intronic bases; (g) Intergenic bases. Whiskers extend from hinges to either the highest or lowest value that is within 1.5*IQR. Data beyond the end of the whiskers are classed as outliers as specified by Tukey. IQR = inter-quartile range.

of bases aligned to the genome is $1.503 \times 10^{10} \pm 1.742 \times 10^{9}$. Of these bases, 66.76±2.63% were aligned to mRNA sequences. 47.83±2.77% of the bases aligned to coding sequences of mRNAs, while 18.9±1.21% aligned to either the 3'- or 5'-UTR regions. 17.63±1.95% of bases aligned to intronic sequences, most likely a result of pre-mRNA sequences present in the sample or unannotated/novel gene isoforms. 15.60±1.97% of bases aligned to intergenic regions, including non-coding RNAs. Of the total aligned bases, only 0.0026±0.0015% of the bases aligned to rRNA sequences, suggesting successful rRNA depletion. The percent of bases aligned to intronic sequences appeared much higher than expected, as one would not expect a large number of pre-mRNA species to be present in a sample due to rapid mRNA processing in a cell. Therefore, as the large number of bases aligning to intronic sequences are not going to be aligned to mRNA species present in the GTF file, which contain mRNA sequences rather than pre-mRNA species, these will be discarded, and won't affect downstream analysis.

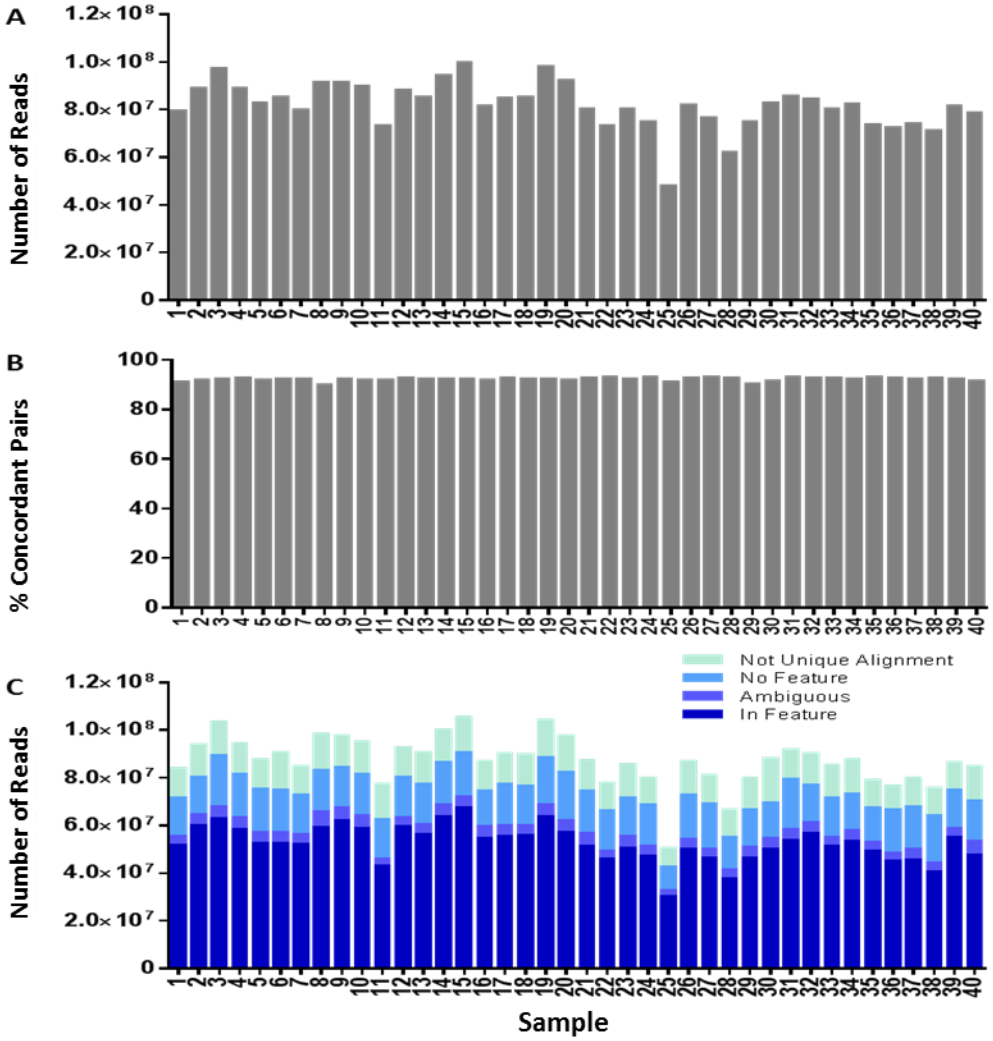


Figure 4.3: Total read metrics

Total read metrics following pre-processing, alignment using TopHat and count table generation with HTSeq. (a) Total number of reads for each sample as determined by FastQC. (b) % concordant pairs aligned to the human genome using TopHat2. (c) HTSeq metrics showing the number of reads associated with annotated features in the human genome, no annotated feature in the genome, ambiguous reads and reads with no unique alignment.

4.2.4. HTSeq Metrics

Count tables were produced using HTSeq, with a reference GTF file for the human genome version GRCh38.83. This results in a table with the number of hits for each gene. Different gene isoforms are grouped together into their parent gene. A summary of the HTSeq output can be seen in figure **4.3c**.

The majority of reads were found in known genomic features (coding genes, miRNAs, lncRNAs, pseudogenes, etc), with the average number of reads in genomic features being 5.27×10^7. 4.42×10^6

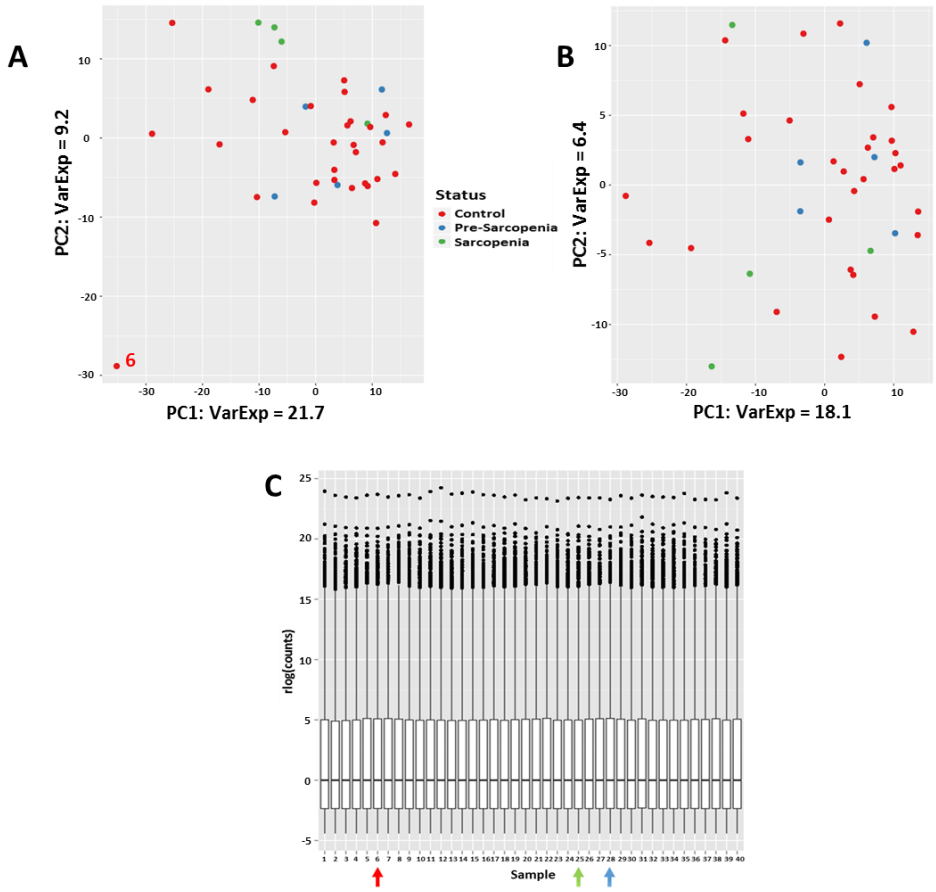


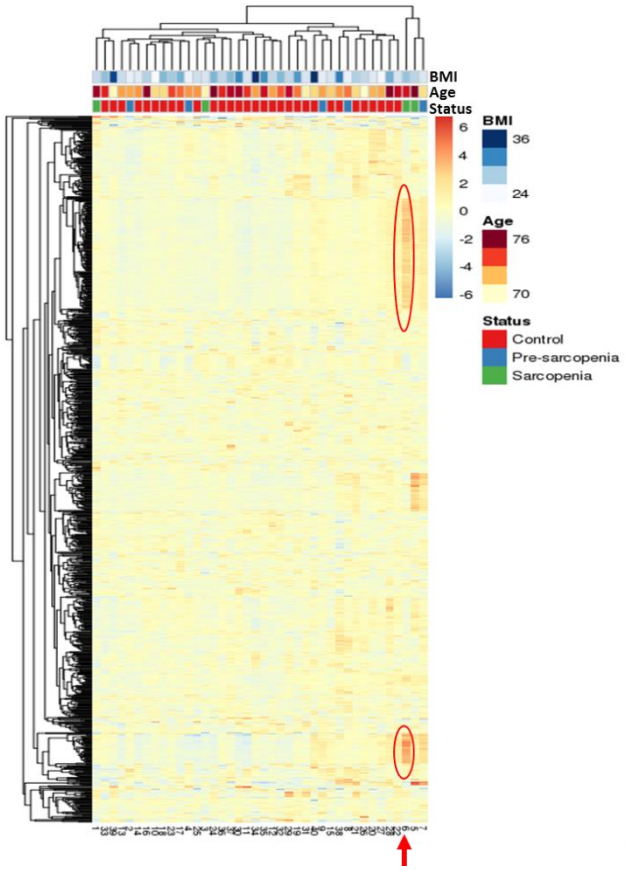
Figure 4.4: Count data visualization

Count data visualization as QC/QA following a variance stabilizing transformation. (a) PCA plot with all 40 samples. Sample 6 however appears to be an outlier. (b) PCA plot of samples removing sample 6 from the analysis (a,b) Red = Control, Blue = Pre-sarcopenia, Green = Sarcopenia. (c) Boxplots of VST transformed count data for all 40 samples. Red arrow = sample6, green arrow = sample 25, blue arrow = sample 28. Plots were produced in R using the ggplot2 package.

reads were classed by HTSeq as being ambiguous due to the read being assigned to more than one overlapping feature. 1.71×10^7 reads were classed as having no feature in the genome. These could come from unannotated genes/gene isoforms/non-coding RNAs. These require further analysis to elucidate where they come from. Finally, 1.31×10^7 reads were classed as having multiple features and no unique alignment, because they aligned to more than one non-overlapping feature in the genome.

4.2.5. Count Table Analysis

Exploratory analyses of the count data were carried out. These analyses, such as clustering and principal-component analysis (PCA), work best with homoscedastic data. RNAseq count data does



Gene	Funtion
CD11a	cytotoxic T-cell mediated killing
CD300E	innate immune response
CD37	T-cell - B-cell interactions
CD41	platelet aggregation and coaggulation
CD48	T cell activation
CD49D	Cytolytic T-cell interactions
CD53	transduces CD2-generated signals in T cells
CSF2RA	controls production, differentiation and function of macrophages
CSF2RB	receptor for IL3, IL5 and CSF
CSF3R	controls production, differentiation and function of granulocytes
CXCL7	activates neutrophils
CXCL8	IL8 attracting neutrophils, basophils and T cells
CXCR1	IL8 receptor neutrophil chemotaxis
CXCR2	IL8 receptor neutrophil chemotaxis
CXCR4	binds LPS
IL1B	mediator of inflammatory response
IL7R	development of lymphocytes
LITAF	activates TNFα expresion upon LPS stimulation
NCF1	host defence
NCF2	host defence
NCF4	host defence
S100A12	pro-inflammatory
S100A8	pro-inflammatory
S100A9	pro-inflammatory
TLR1	activates innate immunity
TLR2	activates innate immunity
TLR6	activates innate immunity
TLR8	activates innate immunity

Figure 4.5: Heatmap of the 1000 most variable genes

Heatmap of the 1000 most variable genes across all samples, showing the difference from the row mean for each sample. Rows and columns are clustered according to their Euclidean distance. Sample 6 is indicated by the red arrow. Circled portions of the heatmap indicate genes that are much more variable in sample 6 compared to the rest of the samples, some of which are listed on the right of the image. Heatmap generated by the pheatmap package in R.

not show homoscedasticity, as the variance increases with the mean. In order to analyse the count data, original data was transformed using the variance stabilizing transformation (VST) function from the DESeq2 package in R. VST transforms the data such that it reduces the dependence of the variance on the mean, resulting in more consistent per-gene standard deviations while keeping the variances unequal for all genes.

Figure **4.4a** shows a principal component analysis (PCA) plot for all 40 samples. The samples were not clustered together into groups and did not separate based on their sarcopenia status. The first two principal components, PC1 and PC2, account for 21.7% and 9.2% of the variation between samples respectively. Although there is no apparent grouping of the samples, sample 6 is grouped away from the rest of the samples, indicating a potential outlier sample. Removing sample 6 from the analysis and repeating the PCA did not affect the grouping, with no evident grouping of the samples (figure **4.4b**). PC1 and PC2 account for 20% and 6.3% of the variation respectively after

removal of sample 6. Boxplots were produced using the transformed count data (figure **4.4c**). The median of the transformed data is zero for each sample, and each sample has outlier genes that have a high level of counts. However, no individual sample showed a deviation from the rest of the samples. Sample 6 (red arrow), which grouped differently in the PCA plot showed a similar boxplot to the remaining samples. Sample 25 (green arrow) and 28 (blue arrow) had reduced read counts (approximately 48million and 62million reads respectively). Both these samples had similar boxplots to the rest of the samples, suggesting that the lower sequencing depth should not affect downstream analysis following data normalization.

The top 1000 genes that show the most variation across the samples were visualised on a heatmap (figure **4.5**). The samples were clustered based on their Euclidean distance from each other, as demonstrated by the dendogram at the top of the heatmap. Samples did not group based on their sarcopenic status. Hierarchal clustering does not split the groups into three separate groups. However, three of the sarcopenic samples (shown in green) cluster together. Several of the control samples formed clusters, while the pre-sarcopenia samples were scattered between the controls. Samples did not cluster based on age or BMI either. Sample 6 (figure **4.5**, red arrow) showed several genes that were overexpressed compared to the other samples. These genes included several pro-inflammatory genes, including several toll-like receptors (TLR1, 2, 6, 8), all of which activate the innate immune system. LITAF (lipopolysaccharide-induced TNF factor) was also up-regulated, which activates the expression of TNFα following stimulation by LPS. Several cytokines (IL1B, IL8) and chemokines (CXCL7, CXCL8) were also up-regulated. These genes may explain sample 6 grouping away from the other samples in the PCA plot (figure **4.4a**).

4.2.6. Differential Gene Expression

Differential gene expression was carried out using DESeq2. Each phenotypic variable was inputted into the DESeq2 design formula individually to determine genes differentially expressed with respect to the different variables. Age was not corrected for, as there was no significant difference between the ages of the participants in each group. Differentially expressed genes are shown in tables **4.2, 4.3** and **4.4** and appendix **D**. Following multiple testing correction using the Benjamani-Hochburg method, a false discovery rate (FDR) was obtained for each gene.

Differential expression analysis was carried out with respect to appendicular lean mass index (ALM/height^2, ALMI), grip strength, gait speed and 6m timed up-and-go (6mTUG). Differentially expressed genes can be seen in tables **4.2**.

Thirteen genes were differentially expressed genes with respect to ALMI, with 54% upregulated. SBK2 (SH3-binding domain kinase family member 2) and PITX1 (paired like homeodomain 1) were

the top two-upregulated genes with respect to ALMI (FDR=$1.34x10^{-8}$ and $1.19x10^{-4}$ respectively). H19 and AC004556.2 were the top two down-regulated genes with respect to ALMI (FDR=0.019 for both). There were 24 genes differentially expressed with respect to gait speed, 21% of the genes were upregulated. The top two genes upregulated with respect to gait speed were CA11 (Carbonic

Table 4.2: *List of differentially expressed genes (FDR<0.1)*

ALMI	Grip Strength	Gait Speed		6mTUG		
Decreased Expression						
AC004556.2	IGKV1D-33	MIR503HG	AC068413.1	OVCH1-AS1		
H19	IGKV1-33	CALCB	XDH	CA11		
ANKRD20A1	GDPD4	HAS2-AS1	BTNL3	OAS3		
CTC-499B15.6	HMOX1	AP001043.1	AL133370.1	MTFR1L		
SLC22A3	AC091045.1	CYP26B1	LINC01451	EIF5		
AL162726.3	HTN1	ONECUT2	CYP3A5	NDRG4		
	ACTRT3	ADGRV1	AC068413.1	RWDD4P1		
		COL25A1	RNU1-28P			
		NELL2	CCL24			
		LIPJ	RNU1-27P			
Increased Expression						
SBK2		CA11		MIR503HG	CALCB	SCHIP1
PITX1		OAS3		BTNL3	IGLV5-45	PFKP
ALAD		IFI6		SAMD11	C12orf77	IGLV4-69
STEAP1B		AC005225.2		AP001043.1	DSC3	EGFEM1P
PCED1B		NDRG4		CERKL	CYP26B1	LOC100652955
AC097662.2				LTK	LINC01451	SAA2
RP11-141M1.4				ADGRV1	LIPJ	AC099489.1
				HAS2-AS1	RNU6-3P	LINC00941
				HTN1	TMEM200A	CPNE8
				SLPI	LY6D	SLC22A2
				ONECUT2	AC079630.1	LGALS7B
				TNFSF11	COL19A1	MTND4P16
				AL133370.1	LINC00629	LINC01038
				DNAH11	FAM216B	CCDC8
				PEX5L	AC067735.1	MIR34AHG
				LRTM1	ACTRT3	SNORA74B
				AC068413.1	GUCY1B2	NME9
				AL162726.3	TRBV4-2	SNORA11
				C1orf140	DISC1	LOC101928100
				AL513303.1	ELAVL2	RN7SL544P
				LRRK2	LOC100418928	AL157788.1
				TNR	IFFO2	AC007221.2
				TRAV26-1	LINC01608	RP11-427J23.1
				COL25A1	NWD1	CCL24
				LINC00475	PROSER2	VN1R11P
				NEUROD1	AL359538.1	LOXL4
				SUSD4	MIR503	AC093690.1.1
				CCDC13	LINC00382	LINC01376
				SRRM1P2	XDH	DLEU2L
				LINC01759	ANGPT2	LOC100420951
						SCHIP1

anhydrase-related protein 11) and OAS3 (2'-5'-Oligoadenylate synthetase 3) with an FDR of 0.00937 and 0.023 respectively. MIR503HG (miR-503 host gene) and CALCB (calcitonin related polypeptide β) were the top 2 down-regulated genes with respect to gait speed (FDR=$1.77x10^{-6}$ and $6.24x10^{-4}$ respectively). There were seven genes differentially expressed with respect to grip strength, these were all down regulated, with IGKV1D-33 (Immunoglobulin Kappa Variable 1D-33) the top gene differentially expressed (FDR=$1.56x10^{-4}$).

Table 4.3: *List of differentially expressed genes in pre-sarcopenia vs healthy controls (FDR<0.1)*

Ensembl Gene ID	Base Mean	Log2(FC)	P Value	FDR	Gene Symbol
ENSG00000196878	202.58	-0.913	3.38E-07	0.0090	LAMB3
ENSG00000276322	6.97	0.945	3.77E-06	0.0505	
ENSG00000185112	207.54	-0.525	6.32E-06	0.0564	FAM43A
ENSG00000042980	23.57	-0.899	1.44E-05	0.0963	ADAM28

We subsequently looked at genes differentially expressed with respect to sarcopenia status. Fewer genes were differentially expressed (FDR<0.1) with respect to pre-sarcopenia compared to sarcopenia (4 and 135 genes respectively, tables **4.3** and **4.4** respectively), reflecting the fact that the pre-sarcopenia definition takes into account only muscle mass, whereas sarcopenia takes into accounts muscle mass and function. The genes altered in pre-sarcopenia include laminin beta 3 (LAMB3, FDR=0.009) and ADAM metallopeptidase domain 28 (ADAM28, FDR=0.0963), both of which show decreased expression in pre-sarcopenia (figures **4.6a, 4.6b**). Of the genes altered in sarcopenia, 64% of the differentially expressed genes showed increased expression in sarcopenia (figures **4.6c, 4.6d).** The top two up-regulated genes were microRNA 503 host gene (MIR503HG, FDR=$6.02x10^{-8}$) and tenascin R (TNR, FDR=0.0014). The top two down-regulated genes were AC005332.9 (FDR=0.0013) and HRAS-like suppressor (HRASLS, FDR=0.0013). Amino-alpha-1,6-glucosidase, 4 alpha-glucanotransferase (AGL) is also down-regulated in sarcopenia (FDR=0.0021), showing decreased expression in those with sarcopenia compared to those with normal muscle mass and function.

LAMB3 is the only gene significantly decreased in both pre-sarcopenia and sarcopenia. There is also some overlap between the genes differentially expressed with respect to sarcopenia and those differentially expressed with respect to the muscle related phenotypic variables (figure **4.7**). AL162726.3, a lncRNA, is the only gene differentially expressed with respect to ALMI and sarcopenia. Nine genes were differentially expressed with respect to gait speed and sarcopenia, and only one with gait speed and sarcopenia. Forty genes were differentially expressed with respect to 6mTUG and sarcopenia. These included MIR503HG, HAS2-AS1 and COL25A1.

75% of the genes differentially expressed with respect to gait speed, are also significantly changed with respect to 6mTUG. These include OAS3 (2'-5'-oligoadenylate synthetase 3) and NDRG4 (N-myc

Table 4.4: *List of top differentially expressed genes in sarcopenia vs healthy controls (FDR<0.1)*

Ensembl Gene ID	Base Mean	Log2(FC)	P Value	FDR	Gene Symbol
ENSG00000223749	31.18	1.445	2.49E-12	6.02E-08	MIR503HG
ENSG00000278730	1195.31	-0.341	1.07E-07	0.0013	AC005332.9
ENSG00000127252	722.19	-0.543	1.65E-07	0.0013	HRASLS
ENSG00000116147	3.82	0.870	3.30E-07	0.0014	TNR
ENSG00000143502	35.96	1.068	3.63E-07	0.0014	SUSD4
ENSG00000170624	8459.63	-0.351	3.97E-07	0.0014	SGCD
ENSG00000188517	43.80	1.078	2.35E-07	0.0014	COL25A1
ENSG00000162688	61939.57	-0.582	7.45E-07	0.0021	AGL
ENSG00000228430	9.29	1.028	7.84E-07	0.0021	AL162726.3
ENSG00000188452	37.76	0.896	1.04E-06	0.0025	CERKL
ENSG00000179813	8.42	0.946	1.44E-06	0.0032	FAM216B
ENSG00000114757	20.97	0.951	2.65E-06	0.0049	PEX5L
ENSG00000227060	4.34	0.922	2.49E-06	0.0049	LINC00629
ENSG00000230943	338.19	-0.563	2.84E-06	0.0049	LINC02541
ENSG00000205622	13.83	0.963	3.62E-06	0.0058	AP001043.1
ENSG00000234754	9.57	0.920	4.49E-06	0.0068	C1orf140
ENSG00000196878	202.58	-0.849	4.84E-06	0.0069	LAMB3
ENSG00000164199	29.68	0.878	5.34E-06	0.0072	ADGRV1
ENSG00000082293	323.11	0.855	9.05E-06	0.0115	COL19A1
ENSG00000279141	4.19	0.811	1.02E-05	0.0123	LINC01451
ENSG00000105877	285.42	0.897	1.17E-05	0.0124	DNAH11
ENSG00000182923	1197.97	-0.307	1.19E-05	0.0124	CEP63
ENSG00000225342	40.38	0.858	1.16E-05	0.0124	AC079630.1
ENSG00000228526	43.55	0.860	1.30E-05	0.0130	MIR34AHG
ENSG00000065882	2321.61	-0.818	1.44E-05	0.0139	TBC1D1
ENSG00000067057	250.34	0.797	1.56E-05	0.0145	PFKP
ENSG00000124107	66.92	0.886	1.75E-05	0.0157	SLPI
ENSG00000118898	442.72	0.781	1.87E-05	0.0161	PPL
ENSG00000266028	301.89	0.515	1.93E-05	0.0161	SRGAP2
ENSG00000181026	170.65	0.570	2.32E-05	0.0165	AEN
ENSG00000188906	1830.60	0.744	2.16E-05	0.0165	LRRK2
ENSG00000222345	4.58	0.898	2.18E-05	0.0165	SNORD19
ENSG00000234546	3.85	0.866	2.29E-05	0.0165	LINC01759
ENSG00000246575	4.73	0.705	2.28E-05	0.0165	AC093162.2
ENSG00000048052	1970.85	-0.578	2.59E-05	0.0178	HDAC9
ENSG00000248690	11.08	0.871	2.82E-05	0.0189	HAS2-AS1
ENSG00000134569	953.55	0.624	3.00E-05	0.0195	LRP4
ENSG00000225733	4868.82	-0.235	3.08E-05	0.0195	FGD5-AS1
ENSG00000149634	74.50	-0.592	3.20E-05	0.0198	SPATA25
ENSG00000058453	649.34	0.613	3.86E-05	0.0233	CROCC
ENSG00000123201	17.53	0.705	3.99E-05	0.0235	GUCY1B2
ENSG00000131080	84.23	0.854	4.18E-05	0.0235	EDA2R
ENSG00000228589	362.35	-0.325	4.13E-05	0.0235	SPCS2P4
ENSG00000081189	35766.76	-0.426	4.84E-05	0.0256	MEF2C
ENSG00000182253	51345.06	-0.365	4.88E-05	0.0256	SYNM
ENSG00000230082	51.99	0.561	4.78E-05	0.0256	PRRT3-AS1
ENSG00000003137	108.67	0.845	5.02E-05	0.0258	CYP26B1
ENSG00000115221	6928.18	-0.410	5.38E-05	0.0271	ITGB6
ENSG00000144771	4.02	0.796	5.77E-05	0.0281	LRTM1
ENSG00000169515	123.17	0.830	5.82E-05	0.0281	CCDC8

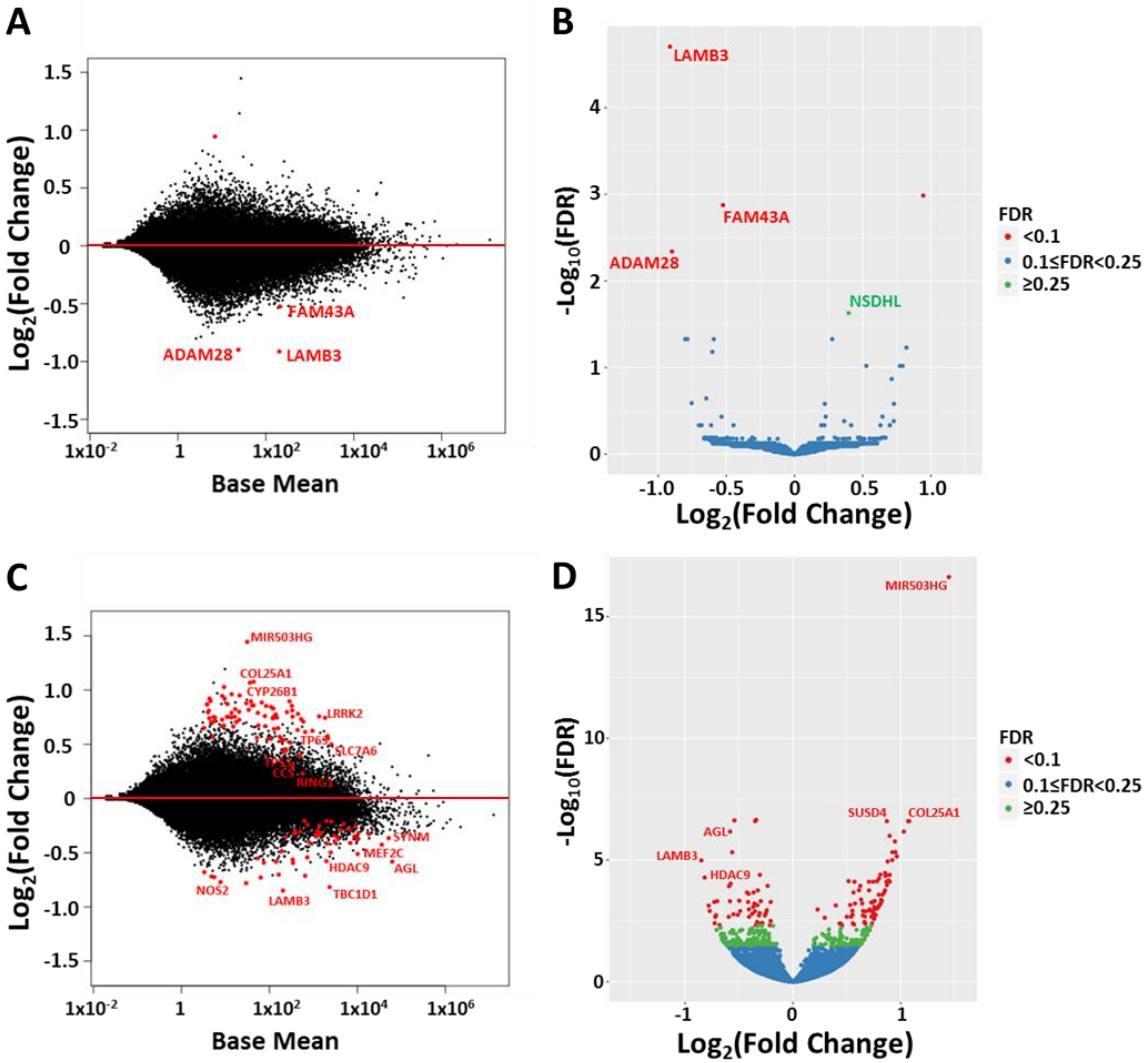


Figure 4.6: MA plots and volcano plots of differentially expressed genes

MA-plots (A + C) and volcano plots (B + D) of differentially expressed genes in pre-sarcopenia (A + B) and sarcopenia (C + D). The MA-plots show the $\log_2$(fold change) against the base mean for each gene as calculated by DESeq2. Genes in red indicate an FDR < 0.1. Volcano plots show the $-\log_{10}$(FDR) against the $\log_2$(fold change) for all genes. Red = FDR < 0.1, Green = 0.25 < FDR ≤ 0.1, Blue = FDR ≥ 0.25.Plots were produced in R using the ggplot2 package.

downstream regulated gene 4), which showed a negative association with poor physical performance, while HAS2-AS1 (hyaluronan synthase 2 antisense RNA 1) showed a positive association with poor physical performance. There was little overlap between the genes differentially expressed with respect to muscle mass and muscle function variables. AL162726.3 was the only gene differentially expressed with respect to ALMI and 6mTUG. The expression of MIR503HG showed a positive association with respect to poor physical performance. MIR503HG shows a negative association with respect to ALMI, however the FDR is greater than 0.1 (FDR=0.130).

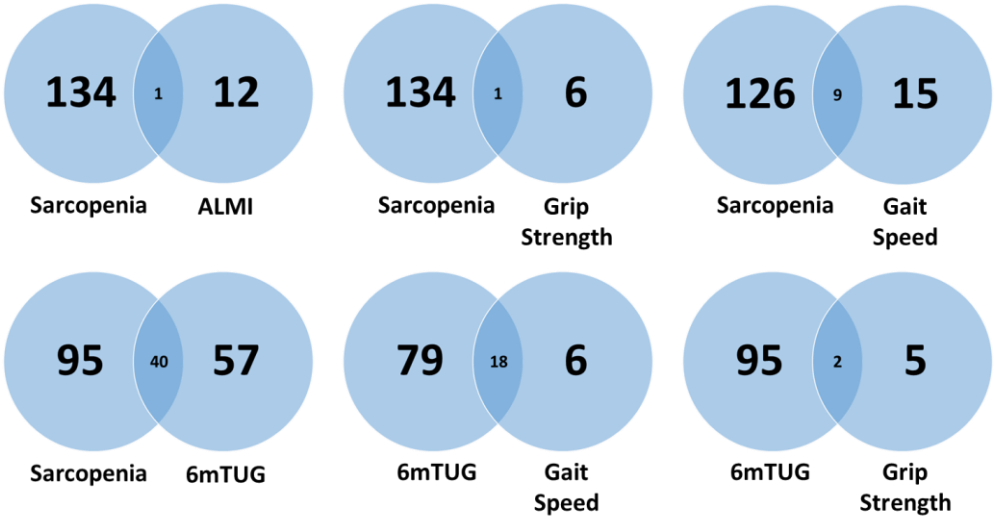


Figure 4.7: Overlap of the differentially expressed genes

Overlap of the differentially expressed genes (FDR < 0.1) between sarcopenia and other phenotypic variables associated with muscle mass and function. Differentially expressed genes in pre-sarcopenia and sarcopenia were combined together for this analysis. ALM = appendicular lean mass, ALMI = Appendicular Lean Mass Index, 6m TUG = 6m timed up-and-go.

4.2.7. Gene Set Enrichment Analysis

Gene set enrichment analysis (GSEA) tests for sets of related genes that are systematically changed, although may individually contribute only subtly to the phenotype. GSEA combines information from a priori sets of genes to increase the signal to noise ratio and improve statistical power. Prior to GSEA analysis, genes were ranked based on their log fold change and p value, generating a list with genes with the biggest positive fold change at the top and those with the biggest negative fold change at the bottom.

4.2.7.1. Phenotypic variables

GSEA was performed with the results from the ALMI, grip strength, gait speed and 6mTUG differential expression analysis (table **4.5**). With increased ALMI, we observed a positive enrichment of genes associated with oxidative phosphorylation (FDR<0.001), citrate cycle/TCA cycle (FDR=0.035) and calcium signalling pathways (FDR=0.079). With respect to ALMI, there was a negative enrichment of genes associated with the p53 pathway (FDR<0.001), reactive oxygen species pathway (FDR=0.055) and apoptosis (FDR=0.052). With respect to gait speed, there was a positive enrichment of genes associated with oxidative phosphorylation (FDR<0.001), mTORC1 signalling (FDR<0.001) and myogenesis (FDR=0.002) as well as a negative enrichment of genes associated with the p53 pathway (FDR<0.001), apoptosis (FDR=0.002) and the ATM pathway (FDR=0.013). With respect to grip strength, there was a positive enrichment of genes associated

with PI3K-Akt-mTOR signalling (FDR<0.001) and myogenesis (FDR=0.001), together with a negative enrichment of genes associated with DNA repair (FDR<0.001) and the p53 pathway (FDR=0.028).

With respect to increased 6mTUG, there is a negative enrichment of genes associated with oxidative phosphorylation, myogenesis and mTORC1 signalling (FDR<0.001), as well as a positive enrichment of genes associated with the inflammatory response, the p53 pathway and apoptosis (FDR<0.001).

Many of the pathways altered with respect to the various muscle-related variables commonly altered. The p53 pathway showed a negative enrichment with respect to increased ALMI, gait speed and grip strength, and decreased 6mTUG, suggesting decreased senescence and decreased DNA damage response, as is further suggested by a negative enrichment of the ATM pathway with respect to gait speed and negative enrichment of DNA repair with respect to grip strength. Increased oxidative phosphorylation and mTOR activity is seen with respect to high ALMI, gait speed, grip strength and low 6mTUG. Decreased apoptosis is also observed with respect to ALMI, gait speed and 6mTUG. Similarly, altered myogenesis and muscle signalling is associated with ALMI, gait speed and grip strength.

Table 4.5: *Top 5 positively and negatively enriched hallmark gene sets*

Gene Set	Gene Set Size	% Genes enriched	Normalized Enrichment score (NES)	NOM P-value	FDR
ALMI					
Positive enrichment					
Interferon alpha response	94	68.1	3.94	<0.001	<0.001
Cell adhesion molecules	131	61.1	3.10	<0.001	<0.001
Oxidative phosphorylation	199	51.8	3.20	<0.001	<0.001
Heme metabolism	195	47.7	2.36	<0.001	0.002
Ribosome	88	76.1	2.30	0.004	0.006
Negative enrichment					
p53 pathway	199	51.8	-3.49	<0.001	<0.001
TNFα signalling via NF-κB	198	32.8	-2.59	<0.001	0.004
Unfolded protein response	111	55.0	-2.47	<0.001	0.001
Hypoxia	197	39.1	-2.48	<0.001	0.004
Estrogen response early	199	48.7	-2.48	<0.001	0.003
Gait Speed					
Positive enrichment					
Oxidative Phosphorylation	199	72.9	9.82	<0.001	<0.001
MTC Targets V1	199	58.8	5.10	<0.001	<0.001
Fatty acid metabolism	158	64.6	4.70	<0.001	<0.001
Adipogenesis	195	42.6	3.92	<0.001	<0.001
MTORC1 Signalling	198	46.5	3.77	<0.001	<0.001
Negative enrichment					
Epithelial mesenchymal transition	197	68.0	-5.88	<0.001	<0.001
TNFα signalling via NF-κB	198	35.9	-3.21	<0.001	<0.001
p53 pathway	199	31.7	-2.92	<0.001	<0.001
Angiogenesis	36	66.7	-2.84	<0.001	<0.001
Apoptosis	159	42.1	-2.30	<0.001	0.002
Grip Strength					
Positive enrichment					
PI3K-Akt-mTOR signalling	105	51.0	3.14	0.002	<0.001
Myogenesis	197	41.1	2.61	<0.001	0.001
Mitotic spindle assembly	193	48.5	2.36	<0.001	0.002
Heme metabolism	195	48.7	2.29	<0.001	0.003
Negative enrichment					
Epithelial mesenchymal transition	195	53.8	-3.08	<0.001	<0.001
E2F targets	199	36.2	-3.01	<0.001	<0.001
DNA repair	143	62.2	-2.67	<0.001	<0.001
Protein secretion	96	40.1	-2.30	0.002	0.004
P53 pathway	198	43.4	-1.94	0.011	0.028

4.2.7.2. Sarcopenia

GSEA was subsequently performed on the results obtained from the differential gene expression analysis in individuals with sarcopenia (table **4.6**). In individuals with sarcopenia, there was a significant increase in the p53 signalling pathways (FDR<0.001), apoptosis (FDR=0.044) and the death pathway (FDR=0.052), suggesting an increase in the atrophy of muscle fibres in sarcopenia, contributing to muscle loss. Consequently, there are multiple gene sets associated with muscle structure and development significantly altered in sarcopenia, including the muscle structure development (FDR<0.001), sarcomere organization (FDR=0.013), myogenesis (FDR<0.001) and muscle tissue development (FDR=0.007) gene sets. This suggests an alteration in the muscle structure in sarcopenia, which may affect muscle function. As expected, there is also a significant decrease in the gene sets associated with mitochondrial biogenesis and function in sarcopenia, including oxidative phosphorylation (FDR<0.001), oxidoreductase complex (FDR<0.001), mitochondrial translation (FDR<0.001) and PGC1α pathway (FDR=0.009) gene sets. We have also shown a decreased enrichment in several gene sets associated with protein synthesis and ribosomal structure and function, including structural constituent of ribosome (FDR<0.001), ribosome biogenesis (FDR=0.002) and translation factor activity RNA binding (FDR<0.001).

Table 4.6: *Gene set enrichment analysis of differentially expressed genes associated with sarcopenia*

Gene Set	Database	Gene Set Size	% Genes Enriched	NES	NOM P-value	FDR
Cell Death and Apoptosis						
p53 Pathway	H	199	39.7	3.86	<0.001	<0.001
p53 Signalling Pathway	KEGG	66	59.09	2.77	<0.001	0.001
E2F Targets	H	199	37.19	2.49	<0.001	0.001
ATM Pathway	BIOCARTA	20	55	2.04	0.004	0.033
Death Receptor Activity	MF	22	68.18	2.13	0.002	0.036
Apoptosis	H	159	25.79	1.70	0.023	0.044
Death Pathway	BIOCARTA	33	57.58	1.90	0.002	0.052
Mitochondria and Metabolism						
Oxidative Phosphorylation	H	199	64.84	-6.31	<0.001	<0.001
Mitochondrial Membrane Part	CC	160	63.13	-4.93	<0.001	<0.001
Mitochondrial Matrix	CC	393	41.22	-4.45	<0.001	<0.001
Respiratory Chain	CC	78	82.05	-4.14	<0.001	<0.001
Electron Transport Chain	BP	91	67.03	-4.03	<0.001	<0.001
Oxidoreductase Complex	CC	92	64.13	-3.91	<0.001	<0.001
mTORC1 Signalling	H	197	55.33	-3.64	<0.001	<0.001
PI3K/Akt/mTOR Signalling	H	104	50.96	-3.37	<0.001	<0.001
Citrate/TCA Cycle	KEGG	30	66.67	-2.44	<0.001	0.003
Response to Oxidative Stress	BP	341	38.71	-2.38	<0.001	0.009
PGC1α Pathway	BIOCARTA	22	81.82	-2.20	0.002	0.009
IGF1 Pathway	BIOCARTA	21	47.62	-1.93	0.006	0.035
Insulin Pathway	BIOCARTA	22	45.45	-1.90	0.014	0.036
Muscle Function and Structure						
Contractile Fibre	CC	207	55.07	-4.53	<0.001	<0.001
Extracellular Matrix	CC	413	38.98	4.15	<0.001	<0.001
Muscle Contraction	BP	229	44.54	-3.75	<0.001	<0.001
I Band	CC	120	54.17	-3.69	<0.001	<0.001
A Band	CC	32	78.13	-3.46	<0.001	<0.001
Muscle Structure Development	BP	418	32.30	-3.35	<0.001	<0.001
Sarcoplasm	CC	65	61.54	-3.21	<0.001	<0.001
Actin Myosin Filament Sliding	BP	38	60.53	-3.19	<0.001	<0.001
Sarcoplasmic Reticulum Membrane	CC	36	66.67	-2.82	<0.001	<0.001
M Band	CC	20	70.00	-2.68	<0.001	<0.001
Myogenesis	H	200	25.00	-2.66	<0.001	<0.001
Structural Constituent of Muscle	MF	41	66.85	-2.73	<0.001	0.001
Myofilament	CC	23	69.57	-2.53	<0.001	0.001
Calcium Activated Potassium Channel Activity	MF	17	82.35	-2.06	<0.001	0.022
GH Pathway	BIOCARTA	28	42.86	-1.90	0.015	0.037
Translation/Protein Synthesis						
Structural Constituent of Ribosome	MF	206	59.71	-4.56	<0.001	<0.001
Ribosome	KEGG	88	75.00	-4.20	<0.001	<0.001
Translation Factor Activity RNA Binding	MF	84	65.48	-3.78	<0.001	<0.001
Ribosome Biogenesis	BP	301	52.42	-2.55	<0.001	0.002
Inflammation						
TNFα Signalling via NF-κB	H	197	39.59	2.29	0.002	0.003
TNFR2 Pathway	BIOCARTA	18	77.78	2.19	0.006	0.021
IL2/STAT5 Signalling	H	196	29.08	1.70	0.021	0.046

4.3 Discussion

Skeletal muscle ageing is a complex process, involving changes in multiple pathways contributing to the decline in muscle mass and function. We have shown several genes that are differentially expressed with respect to ALMI, gait speed, grip strength and 6mTUG. The top down-regulated gene with respect to ALMI was H19, a lncRNA shown to play a role in muscle development and growth. There was little overlap in the genes differentially expressed with respect to muscle mass (ALMI) and muscle function (grip strength, gait speed and 6mTUG). We have also identified genes associated with differential sarcopenia status. There were fewer genes significantly altered in pre-sarcopenia than sarcopenia, possibly reflecting the complex nature of skeletal muscle and the fact that sarcopenia takes into account both muscle mass and function. This potentially suggests pre-sarcopenia is a milder phenotype than sarcopenia. Of the genes differentially expressed in pre-sarcopenia, only LAMB3 is differentially expressed in sarcopenia too. The major pathways associated with sarcopenia and also changing with respect to various muscle phenotypic variables were pathways associated with mitochondrial function and biogenesis (e.g. oxidative phosphorylation, respiration, PGC1α), myogenesis, DNA repair, inflammation and protein synthesis.

4.3.1 Mitochondrial Theory of Ageing and Sarcopenia

Mitochondrial dysfunction and cellular decline due to oxidative stress is one of the leading theories for ageing [104,147,159]. Complexes I and III of the mitochondrial respiratory chain generate free radicals in the form of reactive oxygen species (ROS) during oxidative phosphorylation. ROS produce deleterious effects in tissues, resulting in protein carbonylation and damage to nuclear and mitochondrial DNA (mtDNA) [159,162]. Under normal circumstances, ROS are efficiently cleared from cells by antioxidant enzymes such as SOD1 and catalase. However, it is thought that as we age, there is a reduced efficiency of clearance, resulting in a cellular build-up of ROS and accumulation of stress and damage. This leads to a vicious cycle in which ROS build-up causes mtDNA mutations, producing dysfunctional proteins of the electron transport chain, exacerbating ROS production.

Skeletal muscle is a highly metabolic tissue, requiring the constant production of ATP for normal function. As such, mitochondria play a vital role in the normal physiology of skeletal muscle. Mitochondria are highly susceptible to oxidative stress and ROS damage, affecting ATP production. As protein synthesis relies on ATP production, and as muscle is composed of around 20% protein, a significant reduction in protein synthesis due to mitochondrial dysfunction may lead to a loss of muscle mass.

Although no individual gene associated with mitochondria and mitochondrial function is significantly differentially expressed in sarcopenia, many mitochondrial associated pathways are significantly altered in sarcopenia, as determined by the GSEA. Gene sets such as oxidative phosphorylation, electron transport chain and the respiratory chain pathways are decreased in sarcopenia, implying that the genes associated with these functions show decreased expression. Although these decreases in expression may not be large, the combined changes in multiple genes may significantly impair these pathways, leading to mitochondrial dysfunction. As mitochondria are vital for skeletal muscle function, a slight decrease in many genes associated with the mitochondria may severely affect mitochondrial function, leading to an inability to efficiently produce ATP.

Myoglobin is an iron- and oxygen-binding protein found in skeletal muscle tissue, primarily in type I and IIa fibres. Myoglobin is the primary oxygen-carrying pigment in skeletal muscle, and its main function is the storage of oxygen. Myoglobin is only found in the bloodstream following muscle injury and can be used as a diagnostic tool when found in the blood. Similar to haemoglobin, myoglobin binds oxygen on a heme group. However, myoglobin only contains one heme group, whereas haemoglobin contains four, although the heme groups are identical. Heme is produced via porphyrin synthesis. ALAD (delta-aminolevulinic acid dehydratase) is an enzyme responsible for converting delta-aminolevulinic acid into prophobilinogen, an intermediate step in the heme biosynthesis pathway [400]. ALAD expression shows a positive association with increased muscle mass (ALMI). Although not significant at an FDR of 0.1 in sarcopenia, this is possibly due to the low number of sarcopenic participants in this study. ALAD however is nominally significant in sarcopenia (p value=0.0023). This decrease in ALAD may result in a decrease in the ability to produce heme molecules. This may result in a reduction in the amount of heme available to be incorporated into myoglobin molecules, and as such, results in a reduction in the oxygen-storage capacity of skeletal muscle. As myoglobin allows the use of skeletal muscle during periods of low oxygen, a possible reduction in myoglobin levels due to reduced heme production may increase the speed at which muscle fatigues, possibly limiting its use and resulting in reduced muscle function.

As well as being decreased in sarcopenia, mitochondrial gene sets are increased with respect to increased ALMI and gait speed, and decreased with respect to 6mTUG. This is as expected, because sarcopenia is defined as a decrease in ALMI and function (↓gait speed and ↑6mTUG). As mitochondria play a key role in normal muscle function due to the constant requirement for ATP production, decreased mitochondrial activity can severely affect muscle function. Many of the genes that are decreased significantly at the nominal level (p value < 0.1) in sarcopenia, contributing to the mitochondrial gene set signature, include those associated with the mitochondrial respiratory chain (e.g. NDUFC2 and COX11) and redox status (e.g. PRDX3). This decreased energy output plays a major role in the inability of muscle to function normally, impairing muscle homeostasis and function.

4.3.2. Protein Synthesis and Ribosomal Function in Aged Skeletal Muscle

Protein synthesis is a key processes in regulating skeletal muscle mass, with muscle mass being dependent on the balance between protein synthesis and protein degradation. With age, there is a decrease in anabolic stimulus [143], which together with an increased anabolic resistance in the elderly, results in reduced systemic protein synthesis in the elderly, occurring also in skeletal muscle. Protein synthesis is a tightly regulated process, and the key molecular complex controlling protein synthesis is the ribosome. The ribosome is a multi-protein/RNA complex which converts mRNA sequences into their corresponding protein sequences. Ribosomal biogenesis is tightly regulated with many processes controlling the translation of ribosomal proteins and transcription of ribosomal RNA molecules.

In this study, we have found that the top differentially expressed gene in sarcopenia and also differentially expressed with respect to Gait speed, 6mTUG time and showing nominal significance with ALMI is MIR503HG. MIR503HG is a miRNA host gene, which encodes several mature miRNAs, including miR-424. miR-424 has previously been shown to play a role in skeletal muscle differentiation by repressing the expression of CDC25A, a gene encoding a phosphatase responsible for activating CDK2. Therefore, increased miR-424 expression results in reduced CDC25A and reduced CDK2 activation, promoting cellular quiescence and myoblast differentiation [170]. Despite miR-424 playing a role in promoting cell differentiation, miR-424 has also been shown to reduce rRNA synthesis in skeletal muscle cells [171], contributing to reduced protein synthesis. The normal production of rRNA molecules are important for the normal functioning of the ribosome and protein synthesis. As protein synthesis is the key mechanism regulating muscle mass and function, miR-424 was shown to be associated with low muscle mass and poor physical performance in multiple cohorts of individuals, including the HSS [171]. The increase in MIR503HG associated with sarcopenia and low muscle mass and function seen in this study suggests an increase in miR-424 expression associated with sarcopenia. This increased miR-424 expression may contribute to a reduction in rRNA expression in sarcopenia, resulting in reduced protein synthesis.

This was further confirmed by significant enrichment of genes associated with multiple ribosomal and translation pathways in sarcopenia. Structural constituent of the ribosome and ribosome biogenesis show a significant decreased enrichment in sarcopenia (FDR<0.002). There was also a decrease in translation factor activity RNA binding in sarcopenia (FDR<0.001). Together, this suggests that a reduced biogenesis and activity of the ribosome may contribute to the reduced protein synthesis seen in the ageing population, resulting in reduced muscle mass and function.

miR-424 is co-transcribed with miR-542, with the expression of the two miRNAs showing a very high correlation [171]. Therefore, it may be that the increased MIR503HG associated with poor muscle

mass and function may also contribute to increased miR-542 expression in sarcopenia. miR-542 has previously been shown to positively regulate p53 [401], a key regulating of cellular senescence and cell death [186,402]. Therefore, the increase in MIR503HG may increase cell death in sarcopenia, via increased miR-542 expression. miR-542 has also been shown to contribute to mitochondrial dysfunction in skeletal muscle and reduced muscle function in the elderly [403]. Therefore, an interplay between mitochondrial dysfunction, MIR503HG and cell death may play a role in regulating muscle mass and function in the elderly.

4.3.3. Role of H19 in Skeletal Muscle Ageing

In adults, skeletal muscle is one of the few tissues to express the *H19* gene. We have found that the expression of *H19* shows a negative association with ALMI in aged muscle tissue, suggesting a role for H19 during skeletal muscle development. The *H19* gene produces a product that is 2.4kb long and is maternally imprinted, such that it is only expressed from the maternally inherited allele. Along with its neighbour, the insulin-like growth factor-2 gene (*Igf2)*, *H19* was one of the first imprinted genes to be discovered [404] and plays a role in regulating the imprinted gene network (IGN) [405]. It was subsequently identified in a genetic screen looking at myogenic differentiation in which the key myogenic regulatory factor (MRF) *MyoD* was also identified [406]. The mRNA produced contains no open reading frames sufficiently long enough to have coding potential and no corresponding protein has been found, implying the gene product is a long non-coding RNA (lncRNA). *H19* expression is strongly induced during embryogenesis and is subsequently down-regulated after birth in all tissues, expect for skeletal and cardiac muscle. Skeletal muscle is one of the few tissues that expresses H19 at a relatively high level after birth and may therefore play an important role during myogenesis and myogenic repair and regeneration.

H19 has been shown to act as a molecular sponge for the let-7 family of miRNAs, regulating the downstream expression of let-7 targets genes. Kallen et al. [407] have shown that H19 has four functional binding sites for members of the let-7 family. *Hmga2* is a target for let-7-mediated regulation [408], destabilizing its mRNA in the cytoplasm, facilitating its degradation. With an increase in *H19* expression, it may be possible that there is an increase in let-7 binding to H19, resulting in decreased let-7 availability. Hmga2 reduction is crucial for terminal differentiation of muscle precursor cells. Therefore, an increase in H19 may decrease the degradation of Hmga2, resulting in reduced terminal differentiation of myoblasts and myogenesis [409]. As a result, this can impair muscle repair and the maintenance of muscle mass with age. This suggests a potential role for let-7 signalling downstream of H19 in sarcopenia that requires further investigation.

As well as acting as a lncRNA, the first exon of *H19* encodes two conserved microRNAs (miRNAs): miR-675-3p and miR-675-5p [410]. These two miRNAs promote muscle regeneration and differentiation by negatively regulating two classes of targets: the bone morphogenetic protein

(BMP) pathway transcription factors Smad1 and Smad5, and the DNA replication initiation factor Cdc6 [411]. Dey et al. [411] show decreased myogenin expression upon silencing H19, which is rescued following miR-675-3p/5p expression in C2C12 cells, along with decreased expression of Smad1/5 and Cdc6. Denervation-induced muscle atrophy is regulated by the HDAC4-myogenin axis, with myogenin increasing the expression of the muscle-specific E3 ligases atrogin-1 and MuRF1 [83]. Decreased Smad1/5 activity may therefore result in activation of the HDAC4-myogenin axis, leading to muscle atrophy [412] and loss of muscle mass. However, contradictory to the data in this study, Dey et al. [411] show that H19, via miR-675-3p/5p, is required to promote skeletal muscle regeneration in response to muscle injury. The results seen by Dey et al. [411] occur after muscle injury with CTX (cardiotoxin), which induces an acute injury to the muscle, whereas our results are from muscle tissue from elderly people, which are constantly undergoing stimulation from normal everyday activities and skeletal muscle ageing. The skeletal muscle used in this study is also from aged muscle, while the mice used in the study by Dey et al. [411] were 10-weeks old, considered to be young adults. Therefore it may be the context of the studies that may contribute to these differences.

During normal myoblast differentiation, H19 expression increases, showing high levels in myotubes compared to myoblasts. H19 expression in the myoblasts results in an inhibition of cell proliferation and cell growth [411,413], which is required for muscle differentiation. However, as we have looked at whole muscle homogenates, we are unable to differentiate between the changes in H19 expression in mature myofibres and myoblasts. Therefore, if the increase in H19 expression we are seeing occurs in the myoblasts, this may prevent the proliferation of these myoblasts, promoting the early differentiation of the cells, reducing the pool of reserve cells available for muscle regeneration and repair.

H19 has also been shown to bind the K homology-type splicing regulatory protein (KSRP) [414], a single stranded RBP that is able to integrate different levels of gene expression. H19 and KSRP interact in proliferating, undifferentiated cells, but not after myogenic differentiation. The interaction between H19 and KSRP promotes KSRP's decay-promoting activity of labile mRNAs. KSRP, via H19, interacts with the AU-rich elements of the myogenin 3'-UTR and promotes its degradation in undifferentiated cells [414]. Silencing H19 results in an increased association of KSRP with pri-miR-206 and pri-miR-133b, causing an accumulation of mature miR-206 and miR-133b transcripts, while decreasing the association between KSRP and myogenin. This suggests that with an increase in H19, KSRP binds to the myogenin mRNA, promoting its degradation and preventing KSRP from binding to and promoting the maturation of pri-miR-206 and pri-miR-133b. These two miRNAs are important myomiRs required for myogenic differentiation and myogenesis. Decreased maturation of these myomiRs may impair myogenic differentiation in response to muscle damage, reducing myogenesis and muscle repair.

These differing roles of H19 in the regulation of myogenin expression highlights the complicated nature of myogenic regulation. In undifferentiated cells, H19 may reduce the levels of myogenin by promoting its degradation via KSRP, in turn preventing myogenic differentiation. However, after differentiation, H19, via its host miRNAs, may increase myogenin expression and activation of the HDAC4-myogenin axis, resulting in muscle atrophy. The interplay between the different pathways that H19 may regulate may lead to increased muscle atrophy with a reduced ability for myogenesis, leading to a net result of decreased muscle mass and an impaired muscle function.

4.3.4. AGL in Healthy and Diseased Muscle

Amino-alpha-1,6-glucosidase, 4-alpha-glucanotransferase (AGL) is a glycogen debranching enzyme, involved in glycogen degradation into glucose in muscle and liver tissue, in which glycogen acts as a major source of energy reserve in most organisms. Mutations in the *AGL* gene are associated with the onset of glycogen storage disease type III (GSDIII). GSDIII is an autosomal recessive disorder affecting the liver, skeletal and cardiac muscle, causing liver cirrhosis, skeletal muscle myopathy and cardiomyopathy. AGL is an enzyme with two active sites, for transferase and glucosidase activities, removing side branches from glycogen molecules, allowing a phosphorylase enzyme to continue degrading glycogen into glucose after stalling at branch points [415].

To date, over 50 distinct mutations in the *AGL* gene have been identified [416,417]. Most of these are nonsense mutations, resulting in premature truncation of the protein, while others affect the catalytic activity of AGL, reducing glycogen debranching activity. Symptoms of GSDIII vary widely [418] but include muscle symptoms of varying severities. In children, muscular symptoms manifest as hypotonia and mild weakness, mimicking congenital myopathy. Weakness affects trunk, proximal and distal muscles, causing decreased grip strength and movement, symptoms seen in sarcopenia. In adults, significant muscle weakness is seen. However, there is massive clinical heterogeneity in muscle symptoms, reasons for which are unclear. Muscle glycogen is necessary for normal oxidative metabolism in muscle, providing glucose for energy production and normal muscle function. A limited ability for glycogen debranching activity reduces the accessibility of the available glycogen stores for glucose production, leading to an increase in muscle fatigue and weakness.

In this study, there is a decrease in AGL expression in sarcopenia but no change in pre-sarcopenia. The expression of AGL also shows a nominally significant positive association with ALMI and gait speed. Therefore, a decrease in AGL expression may be contributing to the decrease in muscle mass and function seen in sarcopenia, due to a decreased capacity for glycogen debranching activity and glucose production for energy in skeletal muscle, similar to what is seen with the myopathy symptoms in GSDIII. This can subsequently lead to an indirect decrease in muscle mass as a result of muscle inactivity and muscle disuse atrophy due to increased weakness and fatigue.

4.3.5. DNA Damage and Repair

Levels of DNA damage increase as we age due to the accumulation of somatic mutations over the lifetime, together with other factors, including ROS damage [104]. An inability to efficiently repair DNA damage ultimately affects two main processes: protein synthesis and cell cycle progression. In sarcopenia, we see a positive enrichment in genes associated with the p53 pathway and apoptosis. p53 is a DNA-damage induced transcription factor responsible for activating the transcription of genes involved in DNA repair and ultimately cell death via apoptosis. Genes in the p53 pathway that are significantly increased in sarcopenia include the apoptosis enhancing nuclease (AEN, FDR= 0.0165), growth arrest and DNA-damage-inducible, alpha (GADD45A, FDR=0.0366) and tumour protein p53 (TP53, FDR=0.0916). CDKN1A (cyclin-dependent kinase inhibitor 1A) is a potent inhibitor of cyclin-dependent kinase (CDK)-2 and CDK1, resulting in a halting of cell cycle progression, leading to apoptosis. CDKN1A is increased in sarcopenia (FDR=0.0916) and along with GADD45A suggests an increase in apoptosis possibly via p53-dependent and -independent mechanisms.

As well as being increased in sarcopenia, the p53 and apoptosis pathways are decreased with increased muscle mass (ALMI) and muscle function (↑grip strength, ↑gait speed and ↓6mTUG). This further reinforces the idea that cell death is a contributing factor to a loss of muscle mass and function. However, whether p53 and apoptosis activation leads to the loss of muscle mass and subsequent muscle function or vice versa remains unclear. However, it is clear that these pathways play a key role in the pathogenesis of sarcopenia in the elderly.

It is well known that DNA damage and apoptosis increase in ageing, but we have shown for the first time that in aged skeletal muscle, an increase in p53 pathway and apoptosis is associated with increased muscle loss. The loss of muscle mass is a key phenotype of sarcopenia, and one mechanism by which muscle is lost is via the apoptosis and cell death of muscle cells. This can occur in myofibres, resulting in the loss of muscle integrity and structure, but also in resident satellite cells, contributing to the loss of regenerative capacity seen in ageing and sarcopenia. Apoptosis in satellite cells results in an inability of skeletal muscle repair and regeneration, leading to further muscle loss and a corresponding loss in muscle function. The mechanisms upstream of p53 activation however remain unclear, requiring further investigation. However, genomic instability and DNA damage may be contributing factors.

These pathways suggest an increased level of damage to the genome in sarcopenia. However, the cause of this damage is unclear from this study. One possible explanation is an increase in ROS production and decreased antioxidant activity due to mitochondrial dysfunction. This damage may contribute to an increased cell death and atrophy. Cell death is the ultimate fate of cells that contain

irreparable DNA damage. Therefore, apoptosis of muscle cells and resulting muscle atrophy may contribute to the decrease in muscle mass seen in sarcopenia, due to the decreased ability to efficiently repair DNA damage.

Increased activation of p53 pathways may also be a mechanism to compensate for ribosomal stress and protect the tissue. Increased ribosomal stress due to reduced rRNA synthesis and ribosomal protein synthesis results in an inability to produce sufficient proteins for normal function. As well as resulting in reduced muscle mass, reduced protein synthesis due to ribosomal stress may also impair normal functioning of the cell. Therefore, p53 activation may be a method for protecting against ribosomal stress and to prevent the cell from further proliferating under unfavourable conditions. Further work is required to determine whether there is ribosomal stress associated with sarcopenia, however we have shown that there is increased MIR503HG expression associated with sarcopenia and reduced muscle mass/function. As MIR503HG encodes miR-424/524, two miRNAs previously shown to inhibit rRNA expression [171,403], this ribosomal stress may induce an increased in p53 expression to prevent further proliferation.

4.3.6. Inflammation

Chronic low-grade inflammation is a hallmark of the ageing process and is commonly seen in the elderly suffering from sarcopenia [229]. With age, there is generally a higher level of systemic cytokines and acute-phase proteins, as well as prolonged inflammatory responses to infectious challenge. Numerous age-related diseases associated with sarcopenia also increase inflammation, such as obesity and rheumatological diseases [419]. However, very few studies have looked at the genes and signalling pathways that regulate inflammation in the elderly and how they are changed. We see a decreased enrichment in the gene set 'Regulation of Immune System Process', with an increase in 'TNFα signalling via NF-κB'. This is suggestive of a dysregulation in the immune system, potentially resulting in aberrant inflammation in muscle via NF-κB and TNFα in sarcopenia. Interferon alpha (IFNα) is a type I interferon that is clinically used in the treatment of viral infections in patients with chronic hepatitis B or C [420]. IFNα acts to reduce inflammation and consequent tissue damage in these patients. IFNα therefore plays an anti-inflammatory role, increasing the expression of IL10, resulting in a decreased activity of the NLRP3 inflammasomes and suppressing IL-1β maturation [421]. The decreased IFNα response we see in sarcopenia could suggest that there is increased inflammation in the muscle of those with sarcopenia. This may disrupt muscle function, altering cellular homeostasis, which if sufficiently changed may result in apoptosis and programmed cell death. This may contribute to the sarcopenic phenotype, but what is causing this altered regulation of the immune response is unclear and requires further investigation.

4.3.7. Further Considerations

It is evident from this study that sarcopenia is a complex condition, with multiple different factors and pathways contributing to its underlying pathology. However, the major limitation with this study is the low number of samples studied. Only 40 samples were analysed, and of these, only five were classified as pre-sarcopenic and four as sarcopenic. Therefore, this pilot work requires replication in a larger number of samples to validate the results observed here and whether these changes are seen in the general population. Secondly, the RNAseq analysis was carried out on muscle tissue samples, which is confounded by tissue heterogeneity, which can contribute a signal towards the global transcriptomic profile of the skeletal muscle. Therefore, changes observed may be due to other resident tissues in the muscle, requiring further investigation. Further work is needed to understand fully how these pathways affect normal muscle physiology and contributing to sarcopenia. Mitochondrial dysfunction seems to play a major role in age-related disorder [147,148,159] and is the main pathway affected in sarcopenia. Further investigation is required to elucidate what causes mitochondrial dysfunction in sarcopenia and what effects it is having on skeletal muscle function in sarcopenia. Mutations in AGL are known to cause GSDIII [415], a disease that shows myopathy symptoms. AGL has not been previously associated with or investigated with respect to sarcopenia. However, we see lower expression of AGL in sarcopenia and as it plays a role in glycogen breakdown, AGL may be important in energy production in muscle. The role of AGL in sarcopenia and whether it contributes to the loss of muscle function requires further investigation. H19 is known to play a major role in myogenesis and muscle function. However, the exact pathway downstream of H19 contributing to decreased muscle mass and function remain unclear. It is also unclear what causes higher expression of H19 in sarcopenia, and what transcriptional regulation may be playing a part. It may be interesting to look at the regulation of H19 transcription and the pathway involving miR-675-3p/5p.

4.3.8. Conclusion

Sarcopenia has a multi-factorial aetiology, resulting in a heterogeneous and complex clinical presentation. Although little overlap is seen in the individual genes altered in sarcopenia and with respect to ALMI and muscle function, many of the pathways overlap, with altered mitochondrial function, DNA repair and myogenesis observed with respect to all analysis. H19 also appears to play a major role in regulating muscle mass, showing an association with muscle mass more so than muscle function and sarcopenia status. The little overlap between the different measures of muscle mass and function suggest different genes regulating these processes, culminating in similar pathways being altered. Understanding how these function in healthy muscle and in sarcopenia will provide a better understanding of muscle ageing and what steps can be taken to better ameliorate

the symptoms or slow down the process. The loss of muscle mass and function with age is debilitating and places a considerable demand on health and social care. In order to improve the quality of life of the elderly, we need to better understand the pathways that contributing to the loss of skeletal muscle mass as well as function.

Chapter 5 –

The long non-coding RNA H19 is associated with low muscle mass in community-dwelling older people – findings from the Hertfordshire Sarcopenia Study (HSS)

Chapter 5 – The long non-coding RNA H19 is associated with low muscle mass in community-dwelling older people – findings from the Hertfordshire Sarcopenia Study (HSS)

5.1. Introduction

An interplay of different factors affect skeletal muscle mass, with a close relationship between muscle growth (hypertrophy) and muscle wasting (atrophy). Skeletal muscle is a very plastic and dynamic tissue, with the ability to respond and adapt its function in response to various extrinsic and intrinsic stimuli. Maintenance of adequate skeletal muscle mass as well as function is important for activities of daily living and optimal metabolic function.

5.1.1. Mechanisms controlling muscle mass

Maintenance of skeletal muscle mass requires a fine balance between muscle hypertrophy and muscle atrophy. Dysregulation of this balance can result in decreased muscle mass and function. Muscle hypertrophy in adults involves an increase in protein synthesis and decrease in protein degradation. Several pathways regulate protein synthesis in skeletal muscle; with the IGF1-PI3K-Akt/PKB-mTOR pathway [251,422] playing a key role in positively regulating muscle mass in response to external stimuli, including exercise and hormones (e.g. testosterone [152]). Muscle atrophy involves a decrease in muscle fibre size, due to a net loss of protein. Muscle specific ubiquitin ligases (e.g. MuRF1 and atrogin-1) are upregulated in response to denervation and other stimuli [83] such as inflammation, initiating the breakdown of myogenic regulatory factors and structural proteins.

As well as muscle hypertrophy as a mechanism for increasing muscle mass, satellite cell differentiation and fusion results in muscle hyperplasia. As a consequence of extrinsic cues such as contraction induced muscle damage, activation of intrinsic pathways such as Wnt3a [423] and IGF1 signalling [424] can activate quiescent satellite cells to proliferate, differentiate and contribute to the repair and regeneration of muscle fibres.

To characterise the major signalling pathways perturbed in sarcopenia a previous genome-wide analysis of the skeletal muscle transcriptome in older people identified a significant upregulation of the gene H19 in those with a low appendicular lean mass index (ALMI).

5.1.2. H19 in Skeletal Muscle

H19 is a maternally imprinted, long non-coding RNA 2.4kb in length that has previously been reported to be highly expressed in embryonic tissues but is down regulated shortly after birth in all

tissues except skeletal muscle, cardiac muscle and placental tissue [413,425]. Early work described a strong induction in the expression of H19, initially called *MyoH*, during myoblast differentiation [406], implicating H19 in processes involved in development and myoblast differentiation. Studies have further shown that H19 expression is upregulated during the differentiation of mouse C2C12 cells [411] and bovine myoblast cells [426], as well as during mouse and human skeletal muscle satellite cell differentiation. Moreover, knockdown of H19 expression in C2C12 cells leads to a decrease in myogenesis with a significant reduction in the expression of myogenic markers such as myogenin and MHC [411], with a similar result seen after knockdown of H19 expression in bovine skeletal muscle satellite cells [426]. H19 expression has been studied in human diseases characterised by low lean mass. For example, Lewis et al. [427] recently showed that H19 expression was inversely associated with lean mass in patients with moderate to severe COPD.

H19 acts as a host gene for two miRNAs, with exon 1 of the *H19* gene encoding two conserved microRNAs, miR-675-3p and miR-675-5p. These miRNAs have been shown to be responsible for mediating the effects of H19 on muscle cell differentiation. Dey et al. [411] showed that ectopic expression of either miR-675-3p or -5p rescued the reduction in muscle cell differentiation induced by H19 knockdown in mouse myoblasts. They also found that miR-675-3p and -5p induced differentiation of C2C12 cells even in the presence of growth media, suggesting that they alone are sufficient to induce muscle cell differentiation. miR-675-3p and -5p may influence muscle differentiation by repressing the bone morphogenetic protein (BMP) pathway, targeting the anti-differentiation Smad transcription factors, SMAD1 and SMAD5, which are direct targets of mir-675-3p, and the Cell division control protein 6 (Cdc6), a protein essential for the initiation of DNA replication, a direct target of miR-675-5p. Interestingly co-knockdown of SMAD1, SMAD5 and Cdc6 in C2C12 cells transfected with a H19 siRNA is sufficient to overcome the decrease in differentiation, suggesting that these are important targets for the pro-differentiation function of H19.

Despite SMAD1/5 acting to inhibit differentiation in myoblast cells, the BMP–SMAD1/5/8 axis also acts as a positive regulator of skeletal muscle mass *in vivo*, promoting muscle growth and preventing muscle wasting [412,428]. Over-expression of BMP7 or the BMP receptor ALK3 in skeletal muscle was shown to increase SMAD1/5/8 phosphorylation, increase muscle fibre size (hypertrophy), muscle mass and force generation. BMP signalling via SMAD1/5 also protects muscle fibres from atrophy *in vivo.* Sartori et al. [429] showed that inhibition of BMP signalling, as well as knockdown of SMAD1 and SMAD5 expression in mouse tibialis anterior, significantly reduced myofibre size, induced significant atrophy of the muscle *in vivo,* and strongly exacerbated the effects of denervation and fasting. SMAD1/5, together with the common SMAD, SMAD4, act to negatively regulate the expression of an E3 ubiquitin ligase called MUSA1 (Fbxo30), and as such prevent its action on the breakdown of muscle fibres and muscle atrophy. Winbanks et al. [412] have further shown that inhibition of SMAD1/5 signalling by SMAD6 administration to mouse muscle

increases the expression of MuRF1 and atrogin1, two E3 ubiquitin ligases that have been shown to significantly contribute to atrophy.

5.1.3. Aims

In this study, we sought to validate the association between H19 and skeletal muscle mass identified from the RNAseq analysis, and determine whether H19 acts via the miR-675-3p/5p-SMAD1/5 pathway in humans to regulate skeletal muscle mass. We found that the miR675-3p and 5p, encoded by exon 1 of H19, were negatively associated with ALMI, and positively associated with H19 expression, while SMAD1, a reported target of miR-675-3p was positively associated with ALMI and negatively associated with miR-675-3p expression. Furthermore, we show for the first time that H19 can directly modulate the miR-675/SMAD1 axis in human primary myotubes, suggesting that that increased expression of H19 and its miRNA effectors contributes to the inhibition of SMAD1/5 expression, and may contribute to reduced hypertrophy, increased atrophy and loss of muscle mass in older people.

5.2. Results

5.2.1. Participant characteristics

Summary of anthropometric and physical function characteristics of participants of HSS and HSSe are shown in table **5.1**. In HSS, the mean (SD) age was 72.63 (2.47) years, height 1.74 (0.07)m, weight 82.59 (13.06)kg, BMI 27.33 (3.51)kg/m^2, Total Lean Body Mass 56.02 (6.66)kg, appendicular lean mass 24.06 (3.23)kg, total fat mass 22.32 (7.28)kg, gait speed 1.11 (0.19)m/s and grip strength 37.56

Table 5.1: *Cohort Characteristics*

	HSS (n=57)		HSSe (n=130	
	Mean	**S.D.**	**Mean**	**S.D.**
Age (yrs)	72.63	2.47	78.45	2.55
Height (m)	1.74	0.07	1.64	0.08
Weight (kg)	82.59	13.06	72.36	11.64
BMI (kg/m^2)	27.33	3.51	27.05	3.88
Total Lean Body Mass (kg)	56.02	6.66	39.77	6.98
Appendicular Lean Mass (kg)	24.06	3.23	16.65	3.49
Total Fatmass (kg)	22.32	7.28	29.00	8.24
Gait Speed (m/s)	1.11	0.19	0.96	0.20
Grip Strength (kg)	37.56	7.96	25.17	8.87

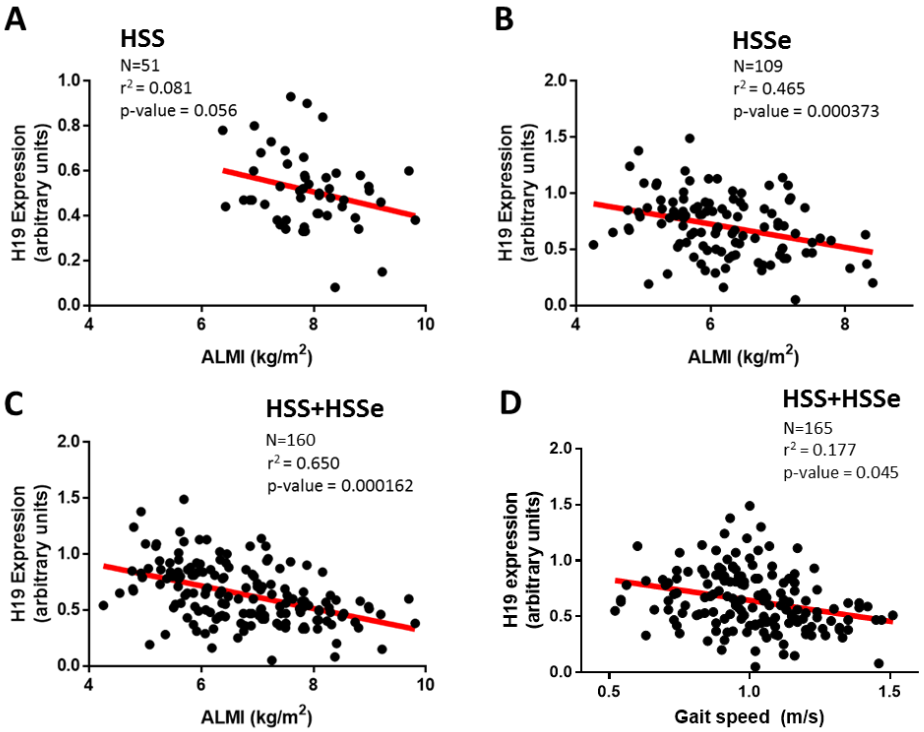


Figure 5.1: Validation and replication of the expression of H19 in HSS and HSSe muscle tissue determined by RT-PCR.

(A) H19 expression in HSS muscle, although only just above significant, shows a similar relationship between H19 expression and ALMI as in the RNAseq. (B) Replication of H19 association with ALMI in HSSe muscle, comprising both men and women. After adjustment for sex and age, there is a significant inverse association between H19 expression and ALMI. (C) Combining HSS and HSSe results improves power, increasing the significance between H19 and ALMI association. (D) H19 expression shows an inverse association between gait speed and H19 expression in the combined HSS and HSSe dataset.

(7.96) kg. In HSSe, 26% of participants were male. The mean age (SD) of whole cohort was 78.45 (2.55) years, height 1.64 (0.08) m, weight 72.36 (11.64) kg, BMI 27.05 (3.88) kg/m^2, Total Lean Body Mass 39.77 (6.98) kg, appendicular lean mass 16.65 (3.49) kg, total fat mass 29.00 (8.24) kg, gait speed 0.96 (0.20) m/s and grip strength 25.17 (8.87) kg.

5.2.2. H19 expression in skeletal muscle

To validate the association between H19 expression and ALMI, qRT-PCR was carried out on biopsy derived skeletal muscle tissue from the HSS participants (n=51). There was a weak inverse association between higher H19 expression and lower ALMI (p=0.056, figure **5.1a**) that did not reach statistical significance. However in HSSe, in a larger cohort (n=109) of men and women, there was a significant inverse association between H19 expression and ALMI (p=0.000373, figure **5.1b**). As HSS and HSSe are two independent

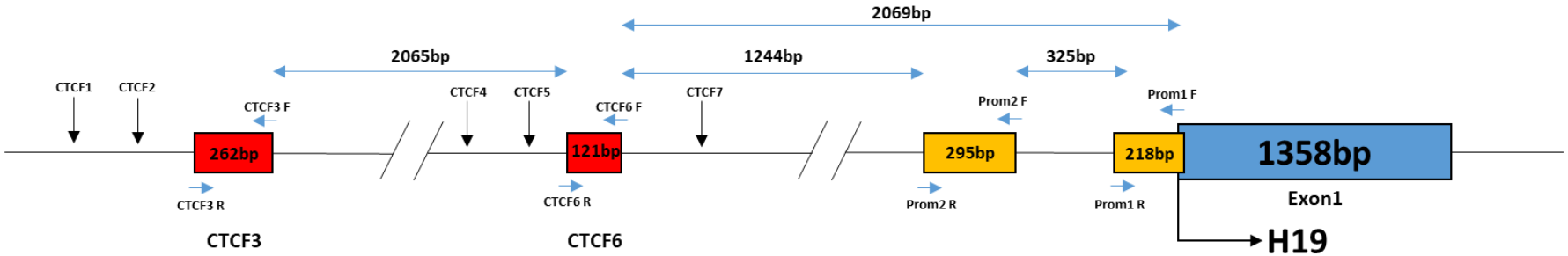


Figure 5.2: Schematic of the H19 promoter region and ICR

Schematic of the H19 promoter region and imprinting control region (ICR) with regions analysed by Sequenom Mass Array for methylation highlighted in red (CTCF3 and CTCF6) and in yellow (promoter region 1 (Prom1) and promoter region 2 (Prom 2)). Exon 1 of the H19 gene is shown in blue. Image not to scale.

groups of participants from the HCS, the data were grouped to increase power; overall, there was a strong inverse association between H19 expression and ALMI in skeletal muscle tissue (p=0.000162, figure **5.1c**) after combining the two groups.

In the RNAseq analysis, H19 expression was not altered with respect to grip strength (FDR>0.1) or gait speed (FDR>0.1). Although, not significant at an FDR < 0.1, H19 expression was increased in those with sarcopenia (nominal p-value = 0.025) and in those with pre-sarcopenia (nominal p-value = 0.0032). However, analysis of the qRT-PCR data showed that there was a significant inverse association with gait speed (p=0.045, figure **5.1d**), but no association with grip strength.

5.2.3. H19 promoter/ICR methylation

DNA methylation plays an important role in maintaining genomic imprints and as H19 is a maternally imprinted gene, we analysed the DNA methylation of several CTCF binding sites in the imprinting control region (ICR) as well as the promoter of H19 (figure **5.2**). CTCF sites 6 and 3 in the ICR are the most widely studied CTCF binding sites in the H19 ICR and have been shown to be key in regulating the H19 locus [430-432]. The methylation of CpGs in CTCF binding site 6 (CTCF6) showed a negative correlation with H19 expression (CpG1: r=-0.231, p=0.043; CpGs 3-4: r=-0.207, p=0.090) and the methylation of CpGs at CTCF binding site 3 (CTCF3) as well as 3 neighbouring CpGs also showed a negative correlation with H19 expression (CpGs 5-8: r=-0.195, p=0.068; CpGs 9-11: r=-0.210, p=0.069). Multiple CpGs in the promoter of H19 (approximately -650bp from the H19 transcription start site) also showed a significant negative correlation with H19 expression in muscle (table **5.2**).

Table 5.2: *Correlation results of DNA methylation at the H19 promoter and ICR and H19 expression*

CpG	Genomic Coordinates (hg19)	N	Pearson r	P-Value	95% CI (lower, upper)
CTCF6					
CpG1	chr11:2021203+	77	-0.231	*0.043**	-0.430, -0.007
CpGs 3-4	chr11:2021208/10+	68	-0.207	0.090.	-0.422, 0.033
CTCF3					
CpGs 5-8	chr11:2023449/52/54/56-	88	-0.195	0.068.	-0.387, 0.015
CpGs 9-11	chr11:2023393/95/99-	87	-0.201	0.069.	-0.393, 0.010
Promoter					
CpG3	chr11:2019850	88	-0.276	*0.009***	-0.457, -0.070
CpG6	chr11:2019800-	90	-0.178	0.093.	-0.370, 0.030
CpGs 7-8	chr11:2019786-	88	-0.236	*0.027**	-0.422, -0.028
CpG16	chr11:2019702-	90	-0.215	*0.042**	-0.402, -0.008
CpGs 17-18	chr11:2019679/81-	91	-0.219	*0.037**	-0.405, -0.014
CpG20	chr11:2019657-	89	-0.244	*0.021**	-0.428, -0.037

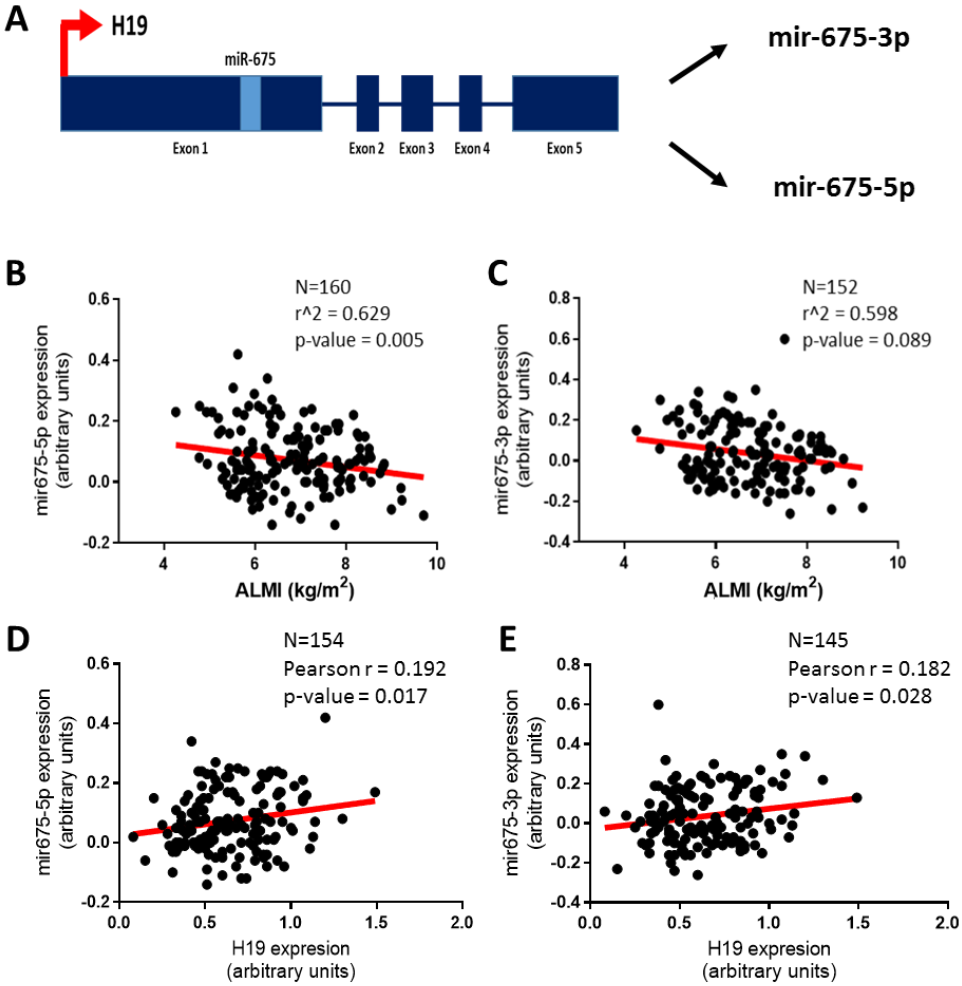


Figure 5.3: miR-675-3p/5p expression in skeletal muscle tissue

(A) Schematic of the H19 gene, showing the localisation of miR-675 in exon 1 of the H19 gene. (B) Association between miR-675-5p expression and ALMI showing a significant negative association. (C) Association between miR-675-3p expression and ALMI, showing a negative trend. (D+E) Correlation between the expression of miR-675-5p (D) and miR-675-3p (E), both showing a significant positive correlation.

5.2.4. Association of miR-675-3p/5p expression with ALMI

Exon 1 of H19 encodes two distinct evolutionary conserved miRNAs, miR-675-3p and miR-675-5p (figure **5.3a**), which have been suggested to act downstream of H19. We sought to investigate whether the expression of these miRNAs also showed a similar association with ALMI, as H19. Using qRT-PCR on RNA extracted from muscle tissue from HSS and HSSe, we found that miR-675-5p showed a significant inverse association with ALMI (p=0.005, figure **5.3b**) while there was a weak association between the expression of miR-675-3p and ALMI (p=0.089, figure **5.3c**). As both miR-675-3p and -5 are encoded for by the H19 mRNA, the association between miR-675-3p/5p and H19

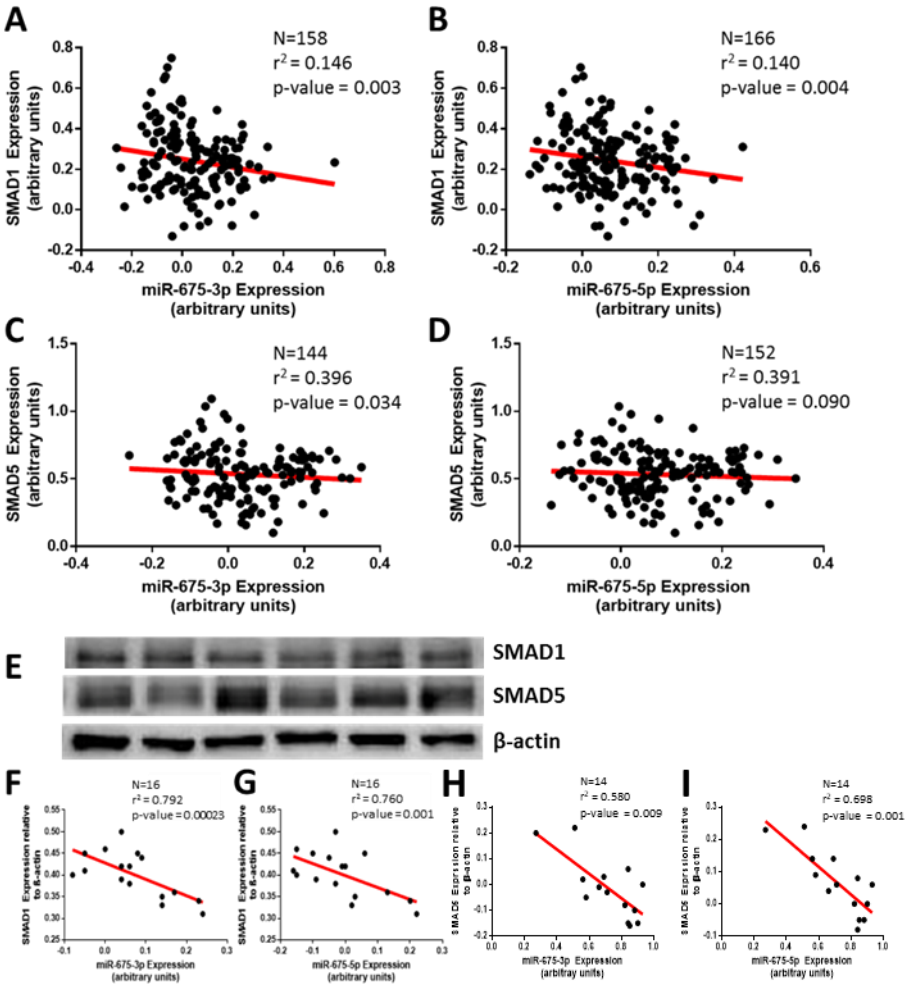


Figure 5.4: SMAD1/5 RNA and protein expression in skeletal muscle and association with miR-675-3p/5p expression

(A+B) Association between SMAD1 mRNA levels and miR-675-3p (A) and miR-675-5p (B) expression. Both showing a significant negative association between SMAD1 and miR-675 expression. (C+D) Association between SMAD5 mRNA and miR-675-3p (C) and miR-675-5p (D) expression. SMAD5 and miR-675-3p show a significant negative association whereas SMAD5 and miR-675-5p shows a negative trend. (E) Western blot for SMAD1 and SMAD5 protein in skeletal muscle tissue, with β-actin as a loading control. (F-I) Association between SMAD1 (F+G) and SMAD5 (H+I) protein levels and miR-675-3p (F+H) and miR-675-5p (G+I) expression. SMAD1 and SMAD5 show a significant negative association with both miR-675-3p and -5p.

expression was investigated. In skeletal muscle, these miRNAs showed a significant positive correlation with H19 expression (p=0.017 and p=0.028 respectively, figures **5.3d-e**). The expression of miR-675-3p and -5p were also significantly correlated (r=0.701, p=1.490x10^{-24}).

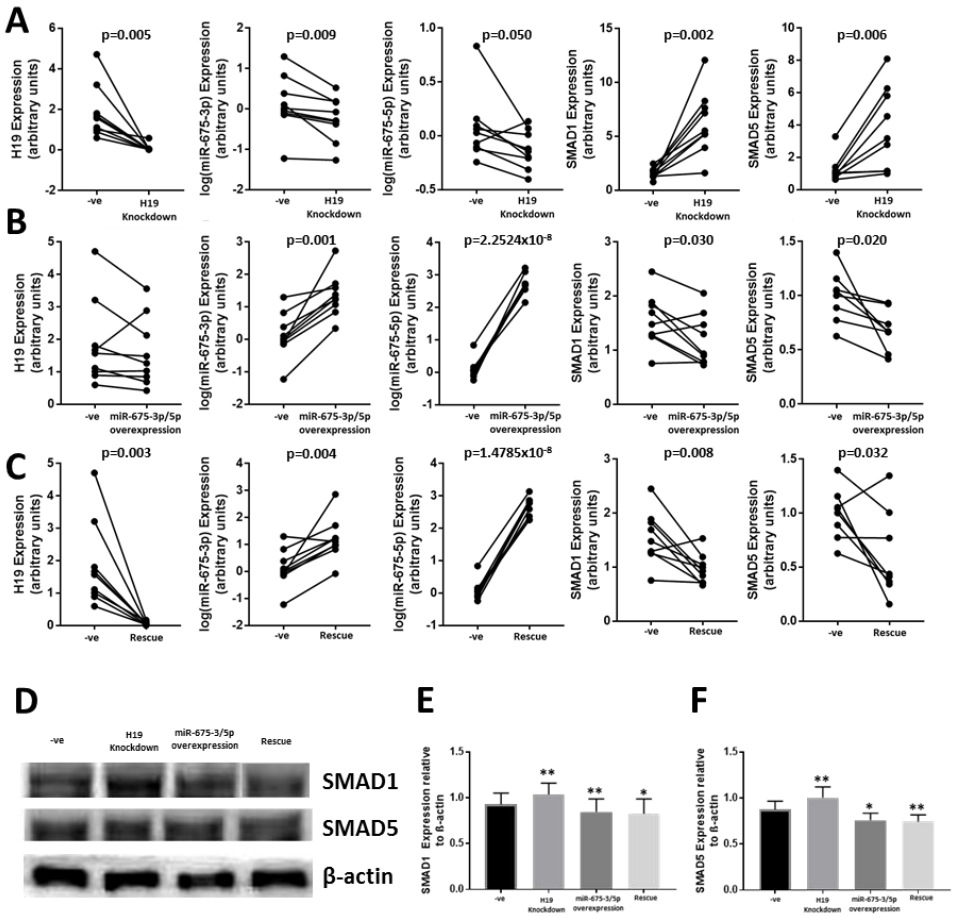


Figure 5.5: Effect of siRNA-mediated H19 knockdown and miR-675-3p/5p ectopic expression on SMAD1/5 levels in myotubes

(A-C) Expression of H19, miR-675-3p, miR-675-5p, SMAD1 and SMAD5 (left to right) following siRNA-mediated H19 knockdown (A), ectopic expression of miR-675-3p/5p (B) and siRNA-mediated H19 knockdown together with ectopic miR-675-3p/5p expression (C). miR-675-3p/5p expression levels were log transformed. (D) Representative western blot of SMAD1 and SMAD5 protein levels following H19 knockdown, ectopic miR-675-3p/5p expression and rescue. B-actin was used as a loading control. (E+F) Quantified protein levels of SMAD1 (E) and SMAD5 (F) following the different treatments. Data are shown as mean ± s.e.m. **p < 0.01, *p < 0.05. All statistical analysis was paired-samples *t* test (two-tailed), n=9 per group.

5.2.5. SMAD1 and SMAD5 are potential targets for miR-675-3p/5p

Dey et al.[411] previously showed that miR-675-3p/5p target SMAD1 and SMAD5 in C2C12 cells and primary mouse myoblasts. miRNAs act by inducing either the degradation of their target mRNA or through inhibiting their translation. Therefore, we investigated whether there was a relationship between the expression of SMAD1/5 and miR-675-3p/5p in human muscle tissue. The expression

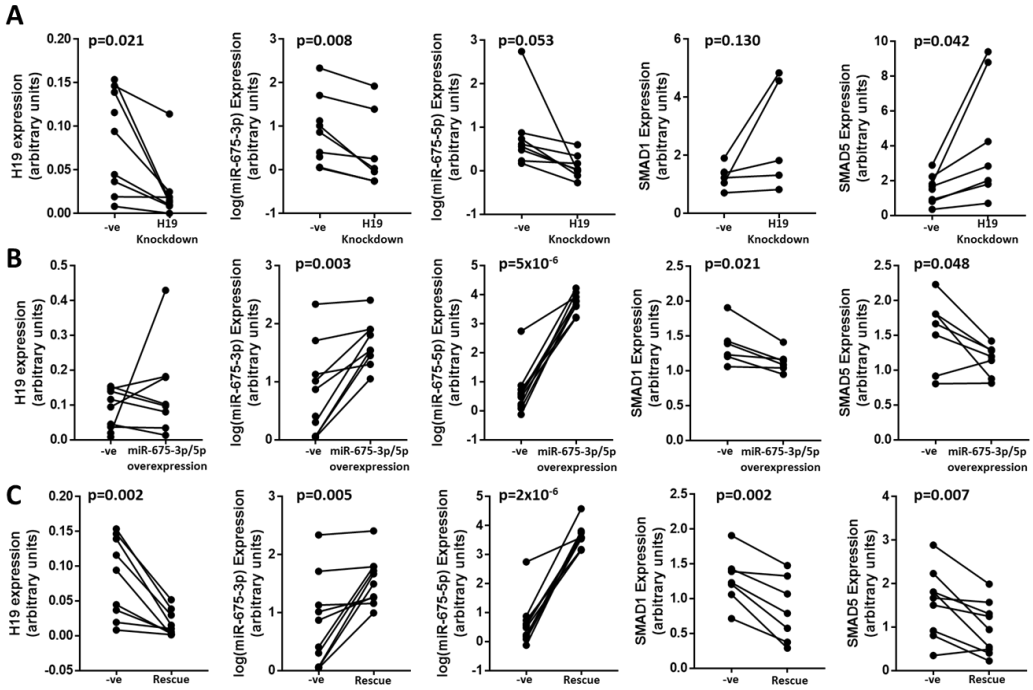


Figure 5.6: Effect of siRNA-mediated H19 knockdown and miR-675-3p/5p ectopic expression on SMAD1/5 levels in myoblasts

(A-C) Expression of H19, miR-675-3p, miR-675-5p, SMAD1 and SMAD5 (left to right) following siRNA-mediated H19 knockdown (A), ectopic expression of miR-675-3p/5p (B) and siRNA-mediated H19 knockdown together with ectopic miR-675-3p/5p expression (C). miR-675-3p/5p expression levels were log transformed. All statistical analysis was paired-samples *t* test (two-tailed), n=6-10 samples per group.

of miR-675-3p showed significant inverse associations with the mRNA expression of SMAD1 (p=0.003, figure **5.4a**) and SMAD5 (p=0.034, figure **5.4c**), and miR675-5p expression was associated with SMAD1 mRNA expression (p=0.004, figure **5.4b**), but only weakly associated with SMAD5 mRNA levels (p=0.090, figure **5.4d**). However, both miR-675-3p and -5p levels were negatively associated with SMAD1 (p=0.00023; p=0.001 respectively) and SMAD5 protein levels (p=0.011; p=0.001 respectively), suggesting that miR-675-3p and -5p may act by both degrading SMAD1/5 mRNA as well as preventing protein translation (figures **5.4e-i**). Higher SMAD1 protein expression was also associated with higher ALMI (p=0.048), but there was no association between SMAD5 protein expression and ALMI.

5.2.6. H19-miR-675-3p/5p directly act to affect the expression of SMAD1/5 in human myotubes

To determine whether alterations in H19 expression mediate the perturbations seen in SMAD1 and 5 expression, we used siRNAs directed against H19 to knock down the expression of H19 in myotube cultures derived from primary human myoblasts isolated from muscle biopsies from HSSe.

Transfection of an siRNA against H19 significantly decreased H19 mRNA expression (p=0.005, figure **5.5a**), with a corresponding decrease in the expression of both miR-675-3p and -5p. This was accompanied by a significant increase in the expression of both SMAD1 and SMAD5 mRNA expression (p=0.002 and 0.006 respectively) and protein expression (SMAD1 p=0.001 and SMAD5 p=0.001, figures **5.5d-f**). Rescuing miR-675-3p/5p expression by ectopic expression of the miRNAs following H19 siRNA knockdown showed a decrease in SMAD1 and SMAD5 mRNA expression (p=0.008 and p=0.032, respectively, figure **5.5c**) and SMAD1 and SMAD5 protein expression (p=0.016 and p=0.003, respectively, figures **5.5d-f**).

5.2.7. H19-miR-675-3p/5p may act via the same pathway in myoblasts

To examine whether H19 acts via a similar pathway in myoblasts as in myotubes, we transfected H19 siRNA and miR-675-3p/5p mimics into primary human myoblasts without allowing them to differentiate in order to examine effects in myoblasts. H19 knockdown with siRNA only (figure **5.6a**) and the rescue experiment (figure **5.6c**) resulted in a significant decrease in H19 expression (p=0.021 and p=0.002 respectively). Knockdown of H19 alone (figure **5.6a**) resulted in a corresponding decrease in miR-675-3p expression (p=0.008), with a decrease in miR-675-5p expression (p=0.053) in myoblasts, both of which are rescued by ectopic expression of miR-675-3p/5p (p=0.005 and $p=2x10^{-6}$ respectively). mRNA level of SMAD5 were significantly increased after knockdown of H19 (p=0.042) with a suggestive increase in SMAD1 mRNA level (p=0.130), which effects on SMAD1 and SMAD5 expression (figure **5.6**) as seen in the myotubes.

5.2.8. SMAD1/5 may act to increase myosin heavy chain content and protect from muscle atrophy

As SMAD1/5 have been previously linked to muscle hypertrophy and decreased atrophy, we next investigated myosin heavy chain content, MHY3 [422], a marker of muscle hypertrophy in the myotubes following siRNA-mediated knockdown of H19 (figure **5.7a**). Knockdown of H19 led to an increase in the expression of MHY3 in the myotubes (p=0.002), which was rescued by ectopic expression of miR-675-3p/5p mimics (p=0.021), implicating SMAD1/5 in increasing protein synthesis or in increasing myoblast differentiation. However, the limitation of this is that we did not measure myotubes diameter directly and so cannot directly link SMAD1/5 with muscle hypertrophy. Therefore further analysis looking at the role of SMAD1/5 on myotubes hypertrophy is required.

We determined whether there was a correlation between SMAD1/5 expression and myofibre size in skeletal muscle. Serial cross sections of muscle tissue (7µm) were stained for both type I (slow

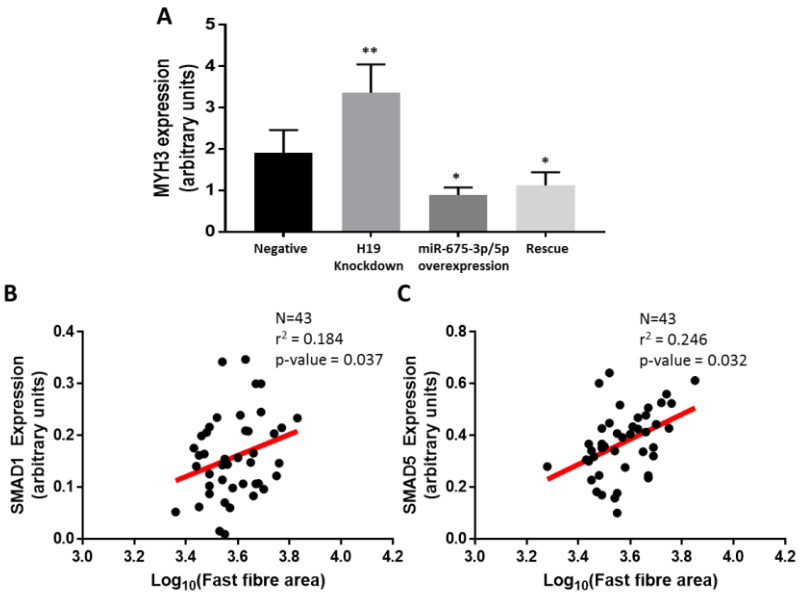


Figure 5.7: Role of H19-miR-675-SMAD1/5 axis in skeletal muscle hypertrophy

(A) Expression of MYH3 following siRNA-mediated H19 knockdown, miR-675-3p/5p ectopic expression and rescue of H19 knockdown by miR-675-3p/5p expression. (B) Association between SMAD1 expression and type II (fast) fibre area (μm^2). (C) Association between SMAD5 expression and type II (fast) fibre area (μm^2). Data are shown as mean ± s.e.m. **p < 0.01, *p < 0.05. Grouped analysis was paired-samples *t* test (two-tailed), while linear regression for continuous analysis.

twitch) and type II (fast twitch) fibres, as previously described [47]. Higher expression of both SMAD1 (figure **5.7b**, p=0.037) and SMAD5 (figure 5**.7c**, p=0.032) were associated with larger type II myofibres. There was a suggestive trend for the positive association between SMAD1 expression and type I fibre area (p=0.114), however there was no association between SMAD5 expression and type I muscle fibre area (p=0.806), suggesting a role for SMAD1/5 in muscle growth.

5.3. Discussion

Here, we showed that higher expression of the long non-coding RNA H19, as identified by genome wide transcriptome analysis, and its miRNA effectors miR-675-3p and -5p were associated with lower ALMI. Moreover, there was a positive association between the expression of SMAD1, a target of miR-675-3p, with ALMI. Furthermore, we showed that H19, via miR-675-3p and -5p acts to regulate the expression (both RNA and protein) of SMAD1 and SMAD5 in skeletal muscle, with implications for muscle mass and atrophy.

5.3.1. Role of H19 in Skeletal Muscle Hypertrophy

We have shown that higher expression levels of both H19 and miR-675-3p and -5p were associated with low muscle mass. There was a positive correlation between H19 expression and miR-675-3p

and -5p expression in muscle tissue, consistent with the operation of this pathway in human muscle. There was a negative association between miR-675-3p and -5p RNA expression with levels of SMAD1 and SMAD5 mRNA. miR-675-3p expression was also inversely associated with SMAD1 and SMAD5 protein expression, with weaker, but still significant, associations observed between miR-675-5p and SMAD1 and SMAD5 protein levels, consistent with miR-675-5p not directly promoting SMAD1/5 mRNA degradation. Interestingly we found that higher SMAD1 protein levels, a target of miR-675-3p, was significantly associated with higher ALMI, suggesting dysregulation of the H19/miR-675/SMAD pathway and that perturbations in this pathway may contribute to the decrease in muscle mass.

H19 can potentially act via several pathways to regulate muscle mass in humans. As well as regulating muscle hypertrophy and atrophy via SMAD1/5, H19 may act via other pathways. H19 has been shown to act as a molecular sponge for the let-7 family of miRNAs [407]. The let-7 family of miRNAs are ubiquitously expressed, with *Hmga2* as one of the targets genes of let-7-mediated regulation. Let-7 promotes the degradation of *Hmga2* mRNA. The degradation of *Hmga2* in myoblasts is essential to allow the formation of multinucleated myofibres and terminal differentiation of myoblasts. The increase in H19 associated with decreased muscle mass suggests a decreased availability of let-7 due to increased binding of let-7 miRNAs to H19. This would result in decreased degradation of *Hmga2,* which may lead to the decreased generation of terminally differentiated myofibres and myogenesis in those with increased expression of H19, contributing to reduced muscle mass. H19 has also been shown to regulate the expression of IGF1R [413,433]. miR-675-3p has been previously shown to directly regulate the expression of IGF1R. IGF1R is the receptor for IGF1, with IGF1 playing a key role in normal growth and hypertrophic responses. These genes aid in muscle repair, and the increased H19 associated with low lean mass suggests a decreased ability to repair muscle tissue in response to injury, leading to decreased muscle mass. Whether or not H19 acts via any of these pathways in skeletal muscle is currently unknown and requires further investigation into the mechanisms of action of H19.

The role of H19 in regulating the miR-675/SMAD1/5 axis has been studied primarily in mice, *in vivo* and *in vitro*, where Dey et al. [411] showed that miR-675-3p and -5p target SMAD1/5 in C2C12 cells, decreasing the expression of both their RNA and protein. However, here we showed that H19 can directly modulate the miR-675/SMAD1 axis in human myotubes. H19 knockdown led to a concomitant decrease in miR-675-3p and -5p expression and an increase in SMAD1/5 mRNA and protein expression. Moreover, transfection of mir-675-5p and 3p mimics directly downregulated SMAD1 and SMAD5 expression suggesting that in human myotubes miR-675 is causally involved in the regulation of the anti-differentiation factors SMAD1/5 and implicating SMAD1/5 signalling in human muscle in the maintenance of muscle mass. Altered expression of the H19/miR-675/SMAD

axis may contribute to loss of muscle mass in a number of ways. Firstly, in mouse models decreased SMAD1 expression decreases muscle hypertrophy and increases muscle atrophy [412,428]. The decrease in SMAD1 expression observed in individuals with low muscle mass suggests that reduced SMAD1 expression may have a similar impact on muscle atrophy in human muscle. This is suggested by the positive association between type II fibre size and SMAD1 expression. Alterations in the H19/miR-675/SMAD pathway may also impact on satellite cell/myoblasts differentiation, as suggested by H19 knockdown and ectopic miR-675-3p and -5p expression causally altering SMAD1/5 expression. Increased H19/mir-675 expression would be expected to reduce the proliferative capacity of the satellite cells and induce precocious differentiation, potentially impairing the regenerative capacity of the muscle tissue and contributing to loss of muscle mass.

There are a number of limitations to this work. Firstly, although we have replicated the association observed between H19 and ALMI from the RNAseq analysis in the HSS cohort and in an independent group of individuals from the HSSe, additional replication is required. Interestingly Lewis et al. [427] have previously reported higher H19 and miR-675 expression was associated with lower fat free mass index (FFMI) and quadriceps strength in COPD patients. However, they did not find any significant association between miR-675 and FFMI in muscle samples from 67 participants from HSS; these 67 participants are also included in this study but expression of H19 was not previously measured in them. The difference in findings between the two studies may be due to the smaller number of samples analysed by Lewis et al.[427] Secondly, differences in the assay design may contribute to the different results seen. Furthermore, we investigated the expression of miR-675-3p and -5 individually, which was not reported by Lewis et al. [427] Nevertheless, the associations seen between H19, its effector miRNAs and the mRNA targets of miR-675 support the robustness of our findings and suggest that this axis is altered with respect to low ALMI. Another limitation with this study is, as shown in chapter 6, the level of differentiation observed at day 5 is relatively low. Therefore, the results we see here suggest a role for H19-miR-675 signalling in differentiating myotubes rather than fully differentiated myotubes, and further work at later stages of differentiation are required to determine whether the results are similar throughout the course of differentiation. Finally, as suggested in chapter 4, the differences observed between our study and the results by Dey et al. [411] may be due to the experimental context.

5.3.2. Imprinting and H19 methylation

The expression of H19 is regulated by imprinting, with the H19 gene being expressed from the maternal allele while expression of the paternally inherited copy is silenced. DNA methylation is the major contributor to genomic imprinting, as well as the regulation of gene expression. The H19 gene has an imprinting control region (ICR) upstream from its transcription start site, between the H19 gene and the IGF2 gene on chromosome 11. Seven CTCF binding sites in the ICR play key roles

in the imprinting of H19 and maintain an allele-specific expression of the gene [430]. These CTCF sites are hypermethylated on the paternal allele, inhibiting the binding of CTCF to the sites. This results in expression of IGF2 due to the removal of the insulator activity caused by CTCF binding and silencing of H19 expression from the paternal allele. Conversely, on the maternal allele, these CTCF sites are hypomethylated, allowing the binding of CTCF to the DNA. CTCF bound to these sites act as an insulator, preventing enhancers downstream of H19 from acting upon IGF2 and initiating its expression. Concomitantly, CTCF binding to these sites initiates H19 transcription from the maternal allele because of chromatin looping.

We found that a significant increase in methylation at two key CTCF sites, CTCF6 and CTCF3, is associated with decreased expression of H19 in skeletal muscle. This supports the role of the ICR in regulating the expression of H19 in skeletal muscle. This is in agreement with Lewis et al. [427] in patients with COPD, who found that methylation of the ICR correlated with the expression of H19 in skeletal muscle. Together with differential methylation at the ICR, we found increased methylation at several sites in the promoter of H19 also associated with decreased H19 expression. DNA methylation at the promoter of genes alters the ability of transcription factors to bind the promoter, altering the promoter's ability to regulate gene expression. It is currently unknown what transcription factors bind this region of the H19 promoter, requiring further work to determine the effect of the differential methylation in this region on H19 expression.

It is currently unknown whether alterations in DNA methylation at the H19 promoter and CTCF sites is causal to the altered expression in H19 or whether it is a consequence of altered H19 expression. It is known that DNA methylation is responsive to changes in the environment and changes during early life. Therefore, it is possible that environmental exposures during early life development (pre- and post-natal development) may induce changes in DNA methylation of the H19 ICR and promoter, resulting in altered H19 expression during later life and predisposing individuals to changes in muscle mass. Further investigation is required to determine whether the relationship between DNA methylation and expression of H19 is a causal relationship, or whether the changes in methylation are a consequence of altered expression during later life.

5.3.3. Conclusion

This study is the first to show that H19 expression in human skeletal muscle is associated with skeletal muscle mass in healthy older people. We have shown that H19 potentially acts through the miR-675/SMAD1/5 pathway to regulate muscle mass and muscle hypertrophy in skeletal muscle. This provides a potential pathway that can inform the development of interventions to help slow the progression of muscle loss and help maintain function and independence in older people.

Chapter 6 –

Cultured Human Myoblasts Provide a Model for the Investigation of Skeletal Muscle Ageing and Interventions

Chapter 6 – Cultured Human Myoblasts Provide a Model for the Investigation of Skeletal Muscle Ageing and Interventions

6.1. Introduction

Sarcopenia affects 5-13% of 60-70 year olds, with the prevalence increasing significantly in those over 80 years of age [107,113]. Sarcopenia is a major cause of frailty and disability in the elderly. The recognition of the significance of the loss of muscle mass and function in the elderly has led sarcopenia to be officially recognized as a disease with an ICD-10 code in 2016 [106]. Therefore, it is important for the development of tools in which the mechanisms of human skeletal muscle ageing can be investigated, as well as tools that can be used to investigate interventions and therapeutics to ameliorate the effects of muscle ageing and retard the ageing of skeletal muscle, aiding in the prevention of the loss of further muscle mass and function.

6.1.1. Skeletal Muscle Ageing

Ageing is associated with gradual changes in total body composition, with increased body fat and decreased muscle mass. After the age of 30, skeletal muscle is lost at a rate of approximately 1% per year, with a dramatic increase in the rate of loss after the age of 65 [134,135]. However, there is significant variability between individuals in the rate of muscle loss. A proportion of the variability between individuals can be explained by fixed genetic factors, however much of the variability remains unexplained. There is some evidence implicating the early life environment on an individual's risk of sarcopenia and muscle loss [48,434-436]. Epigenetics is thought to mediate the interaction between an individual's genotype and the environment, providing a dynamic mechanism in which the environment can affect an individual's risk of disease. Epigenetics has been implicated in affecting an individual's susceptibility to many diseases, including type II diabetes [315] , non-alcoholic fatty liver disease (NAFLD) [437] and cardiovascular disease (CVD) [438,439]. Epigenetic changes at key locations in the genome are thought to be the mechanism by which the environment affects an individual's genome without altering the genetic code [440]. However, the epigenetic changes associated with the loss of muscle mass and function in healthy older people have not been investigated.

Loss of muscle mass is associated with a change in many pathways, including mitochondrial function, cellular proliferation, differentiation and metabolism. Decreased function of muscle

myoblasts (proliferation and differentiation) as well as increased senescence of satellite cells and myoblasts are all thought to contribute to decreased muscle function and contribute to muscle loss. Therefore, it is important to be able to investigate the mechanisms contributing to the loss of myoblast function, as well as investigate interventions to reverse the debilitating effects of ageing on myoblasts. The effects of skeletal muscle ageing have been studied in *Drosophila* and murine models [143,441], but the effects of ageing on skeletal muscle in human models are limited.

6.1.2. Cellular Models of Skeletal Muscle Function

In order to effectively assess the mechanisms underlying skeletal muscle ageing, as well as assess the efficacy of developed therapeutics and interventions, it is necessary to develop *in vitro* models to investigate the cellular and molecular mechanisms that initiate pathophysiological changes and test the efficacy and safety of new therapeutics and interventions in a preclinical setting.

C2C12 cells are mouse myoblast cells that have been cultured and grown from a normal adult C3H mouse leg muscle [442,443]. C2C12 cells are well characterised cells that proliferate and differentiate rapidly upon serum starvation. They have been used in a large number of studies for the investigation of mechanisms involved in muscle function, as well as to investigate several human diseases. Alaxakis et al. [444] have used C2C12 to investigate the effects of the composition of the extracellular matrix on Duchenne muscle dystrophy (DMD). Several groups have also used C2C12 cells as a model to target interventions and drug therapeutics as well as a cell type to develop models for the investigation of disease progression, including Moorwood et al. [445] and Li et al. [446]. Despite the availability of C2C12 cells and mouse myoblasts in which human diseases can be modelled and investigated, complex human disorders like DMD or even ageing may not be appropriate, as mechanisms common to processes like ageing and dystrophies may be different between mice and humans [447-449]. Only a handful of studies have investigated complex diseases using human-derived cells as an *in vitro* model for skeletal muscle [450,451]. This signifies the requirement for human derived myoblasts, in which mechanisms contributing to changes in skeletal muscle can be investigated and interventions can be tested in a more appropriate preclinical setting.

6.1.3. Aims

The aim of this study was to investigate the DNA methylation profiles of paired skeletal muscle tissue and cultured myoblasts with respect to grip strength, gait speed, ALMI, fasting glucose and fasting serum insulin levels. We aimed to compare the muscle and myoblasts at the level of methylation, to determine whether there was a methylation signature associated with muscle

function, and whether this signature is retained between the muscle and myoblast samples. This will help to determine whether cultured myoblasts represent a valuable *in vitro* model system in which human skeletal muscle ageing can be studied. 450K methylation arrays were carried out on matched muscle tissue and myoblast cultures, and DNA methylation profiles compared. Myoblasts retained a similar DNA methylation profile to muscle tissue with respect to muscle-related phenotypes (e.g. ALMI, grip strength, gait speed) but were more variable with respect to metabolic phenotypes (e.g. fasting serum insulin and fasting plasma glucose levels). As myoblasts retained the methylation signatures associated with muscle-related phenotypes, we molecularly characterised the myoblasts, looking at the proliferation rates, levels of cellular senescence, differentiation potential of the myoblasts and mitochondrial and glycolytic respiration, with respect to muscle mass and function. The phenotypic variables of sarcopenia were found to be retained in the myoblast cultures, suggesting that patient derived myoblast cultures may be a useful *in vitro* model to investigate the mechanisms of muscle functions as well as therapeutic interventions.

6.2. Results

6.2.1. Cohort Characteristics

Summary of anthropometric and physical function of participants of HSSe used are shown in table **6.1**. All participants were men. For the samples analysed by 450k methylation array (n=20), the mean age (S.D.) was 78.47 (2.76) years, height 1.72 (0.06) meters, weight 78.93 (12.93) kg, BMI 26.39 (3.89) kg/m^2, Total Lean Body Mass 48.65 (6.63) kg, appendicular lean mass 21.29 (3.51) kg, total fatmass 24.87 (8.42) kg, gait speed 0.97 (0.20) m/s, grip strength 35.32 (8.39) kg, fasting plasma glucose 5.36 (0.43) mmol/l and fasting serum insulin 6.90 (2.81) mU/l. For all the male samples (n=36), the mean age (S.D.) was 78.03 (2.57) years, height 1.73 (0.06) meters, weight 77.86 (11.51) kg, BMI 26.10 (3.29) kg/m^2, Total Lean Body Mass 48.29 (5.71) kg, appendicular lean mass 21.27 (2.93) kg, total fatmass 25.38 (8.79) kg, gait speed 0.98 (0.22) m/s, grip strength 36.11 (6.80) kg, fasting plasma glucose 5.48 (0.65) mmol/l and fasting serum insulin 7.31 (5.27) mU/l.

Table 6.1: *Cohort characteristics for participants of HSSe used in this study*

	HSSe (n=20)		HSSe (n=36)	
	Mean	**S.D.**	**Mean**	**S.D.**
Age (yrs)	78.47	2.76	78.03	2.57
Height (m)	1.72	0.06	1.73	0.06
Weight (kg)	78.93	12.93	77.86	11.51
BMI (kg/m^2)	26.39	3.89	26.10	3.29
Total Lean Body Mass (kg)	48.65	6.63	48.29	5.71
Appendicular Lean Mass (kg)	21.29	3.51	21.27	2.93
Total Fatmass (kg)	24.87	8.42	25.38	8.79
Gait Speed (m/s)	0.97	0.20	0.98	0.22
Grip Strength (kg)	35.32	8.39	36.11	6.80
Fasting Glucose (mmol/l)	5.36	0.43	5.48	0.65
Insulin (mU/l)	6.90	2.81	7.31	5.27

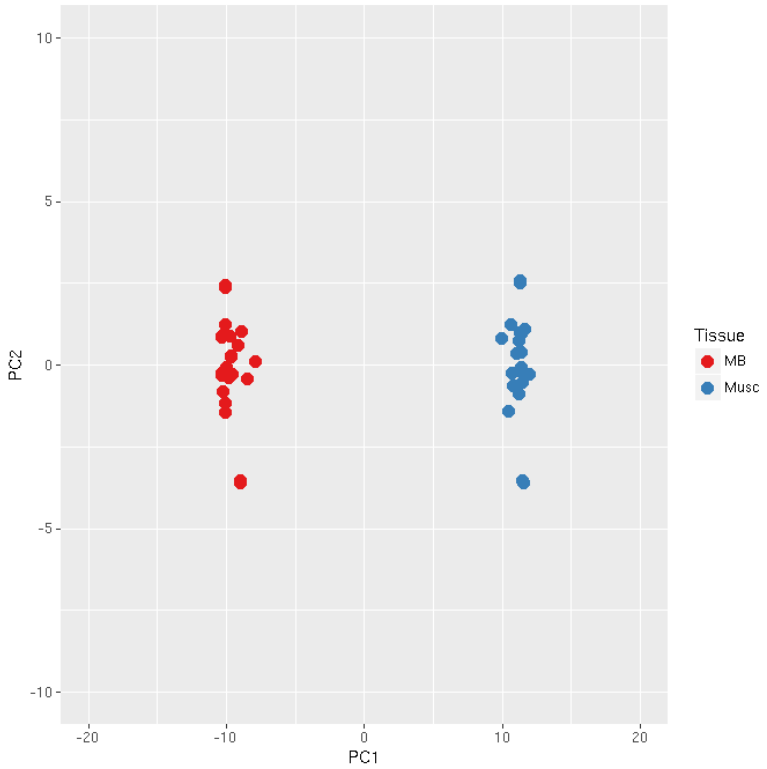


Figure 6.1: PCA plots of samples for the first two principle components Principle component plot of all samples following data normalization. Samples appear to cluster into two very distinct groups, depending on whether they are myoblast samples, or muscle tissue samples. Samples are coloured by the tissue, MB=myoblast, Musc=muscle tissue.

6.2.2. Beta Values Sample QC

DNA methylation analysis was carried out using the Illumina Infinium 450K Beadchip. Beta values for each probe were obtained, with beta values being the ratio between the unmethylated and methylated signals on the array. Several QC steps were carried out at the level of beta values. To obtain beta values, idat files were read into an RGset, a minfi RGChannelSet object which stores the raw intensities in the green and red channels. Before any QC, raw intensities were normalized using the Funnorm (functional normalization) function of minfi to generate beta values for each probes.

Initially, a PCA plot was generated to determine the clustering of samples and identify any outlier samples present in the dataset. A PCA plot of the first two principle components was generated (figure **6.1**), with samples coloured based on their tissue of origin. Samples clearly separate into two distinct clusters, based on whether the DNA originated from myoblast culture or muscle tissue. Methylation is generally cell-specific, with different cell types exhibiting different methylation profiles. Therefore, the separation of methylation profiles on the PCA plot highlights the differences

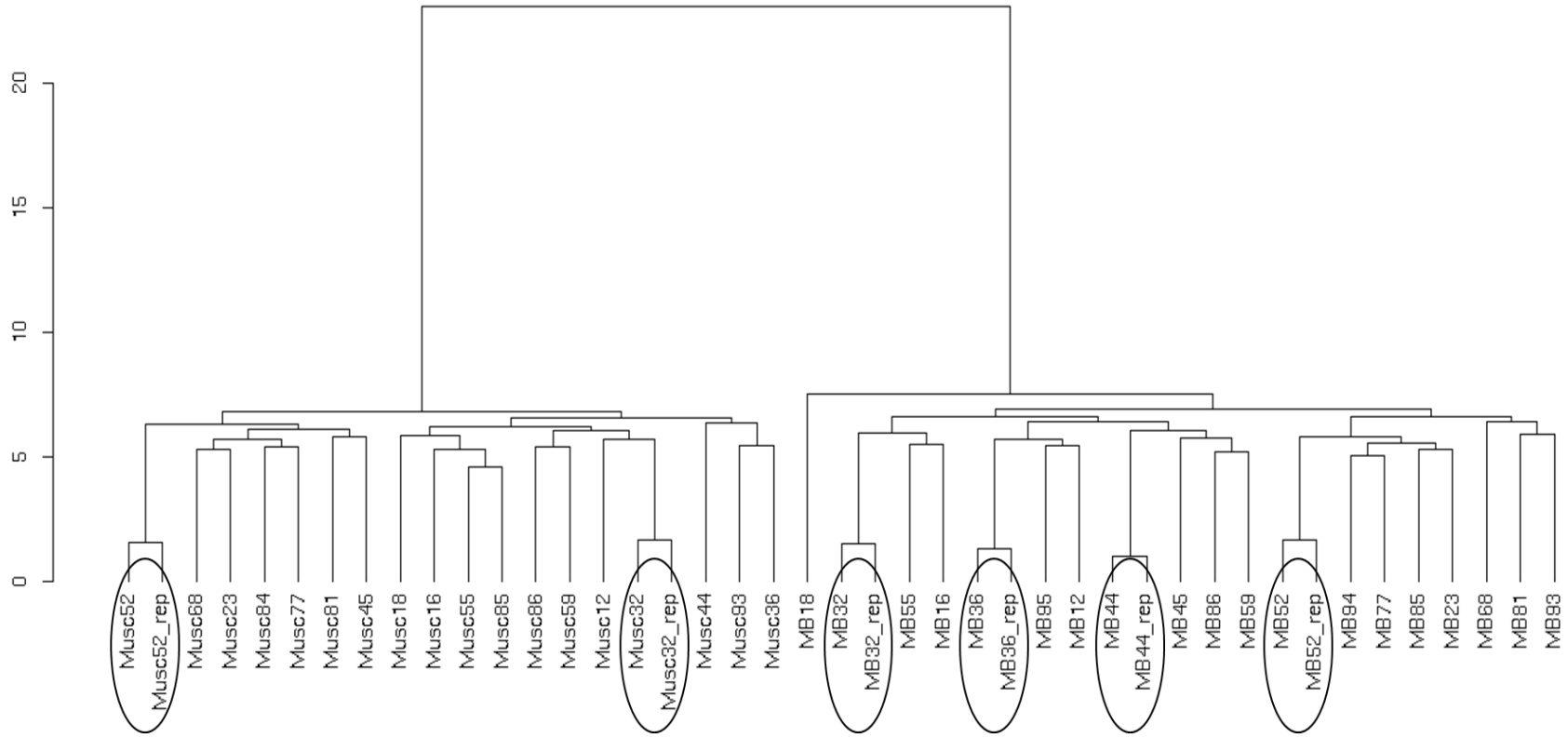


Figure 6.2: Hierarchal clustering of samples based on their Euclidean distances

Hierarchal clustering of normalized beta values based on their Euclidean distance. Two main branches of the dendogram appear, based on tissue of origin. All beadchip replicates clustered together (circled) suggesting good reproducibility across the beadchips.

between myoblasts and native muscle tissue, which is composed of a mainly three different fibre types and a small proportion of myoblasts and other resident cells contributing to the methylation profiles observed. The different fibre types in the muscle may be contributing slightly different methylation profiles to the muscle-associated methylation signature.

Hierarchal clustering was carried out on all samples, based on their Euclidean distances to generate a dendogram in order to look at the clustering of inter- and intra-chip replicates (figure **6.2**). All beadchip replicates clustered together on the same branch of the dendrogram suggesting good reproducibility across the beadchips. All the myoblast samples clustered together on one branch of the dendogram, while all the muscle samples clustered on another branch of the dendogram. This is in agreement with the PCA plot showing two distinct groups of samples based on the tissue of origin of each sample.

Finally, madscores (modified z-score) were calculated for all samples to determine any sample outliers in the dataset (table **6.2**). Samples with a madscore < -5 were classed as outliers. No samples had a madscore < -5. Technical replicates were removed from the dataset based on their madscores.

Table 6.2: *MADscores for all samples plus replicates*

ID	Beadchip	Madscore		ID	Beadchip	Madscore
MB81	1	-0.730		Musc12	3	-0.674
MB12	3	0.627		Musc16	3	1.050
MB16	3	-0.108		Musc18	2	-3.216
MB18	2	1.935		Musc23	2	-1.010
MB23	2	0.492		Musc32	2	-0.373
MB32	2	1.056		Musc32_rep	4	-0.391
MB32_rep	4	1.093		Musc36	3	-0.094
MB36	3	0.976		Musc44	3	1.294
MB36_rep	4	1.137		Musc45	4	-0.024
MB44	3	-2.911		Musc52	1	0.504
MB44_rep	4	-2.652		Musc52_rep	4	0.574
MB45	4	-0.376		Musc55	1	-0.154
MB52	1	1.249		Musc59	3	-0.598
MB52_rep	4	1.222		Musc68	1	-1.019
MB55	1	0.068		Musc77	4	0.000
MB59	3	-0.284		Musc81	1	0.064
MB68	1	0.552		Musc84	2	-3.298
MB77	4	0.682		Musc85	2	-1.441
MB85	2	-2.101		Musc86	2	-0.918
MB86	2	0.202		Musc93	1	0.928
MB93	1	0.169				
MB94	3	-0.076				
MB95	1	0.497				

6.2.3. Beta Values Probe QC

CpG probes with a detection p-value > 0.01 or with a beadcount < 3 were set to missing. CpG probes were excluded from subsequent analysis if missing in more than 5% of samples. 1151 probes were excluded based on their detection p-value and a further 196 probes excluded based on a low beadcount. The detection P value is used to determine whether a probe's signal is above background. A set of negative control probes are used to detect background intensity and a small p value indicates the signal is above background, and likely to be a true signal. 16,945 SNP probes were removed from the dataset, including probes that have a SNP at the CpG site and probes that have a SNP at the single nucleotide extension. Chen et al. [452] have previously shown that there are probes on the 450K array that cross-hybridize to multiple different regions in the genome, resulting in 29,152 probes being removed from the dataset. CpG probes with a methylation range <5% were removed from the dataset. These probes were removed as there is conflicting data in the literature about the biological impact of CpGs with small methylation changes. This allowed the identification of probes with larger methylation differences, removing probes with small variation in methylation, with little chance of being differentially methylated.

6.2.4. Differentially Methylated Probes

Differential methylation was carried out with respect to grip strength, gait speed, ALMI, fasting plasma glucose and fasting serum insulin levels using limma. Limma is an R package used to fit multi-factor models to array data, in order to determine differential methylation/expression. Each phenotypic variable was inputted into the design matrix along with age as a covariate. Following fitting of the model to the beta values, p values were adjusted for multiple testing using the Benjamani-Hochburg method, generating a false discovery rate (FDR) for each probe. For analyses that have a genomic inflation factor (λ) > 1.2, p values were adjusted with the BACON algorithm to reduce inflation and bias, and subsequently adjusted for multiple testing using the Benjamini-Hochburg method. Top 50 differentially methylated probes are shown in appendices **E-N**.

6.2.4.1. Differential Methylation in Muscle Tissue

Differential methylation was carried out with respect to appendicular lean mass index (ALMI), gait speed, grip strength, fasting plasma glucose levels and fasting serum insulin levels. Two CpGs were differentially methylated with respect to grip strength, cg20761395 located in the body of the VPS53 gene (FDR=0.0263) and cg25071674 located in the body of the EIF2C2 gene (FDR=0.0263). There were no dmCpGs with an FDR < 0.2 with respect to gait speed. However, the top two dmCpGs

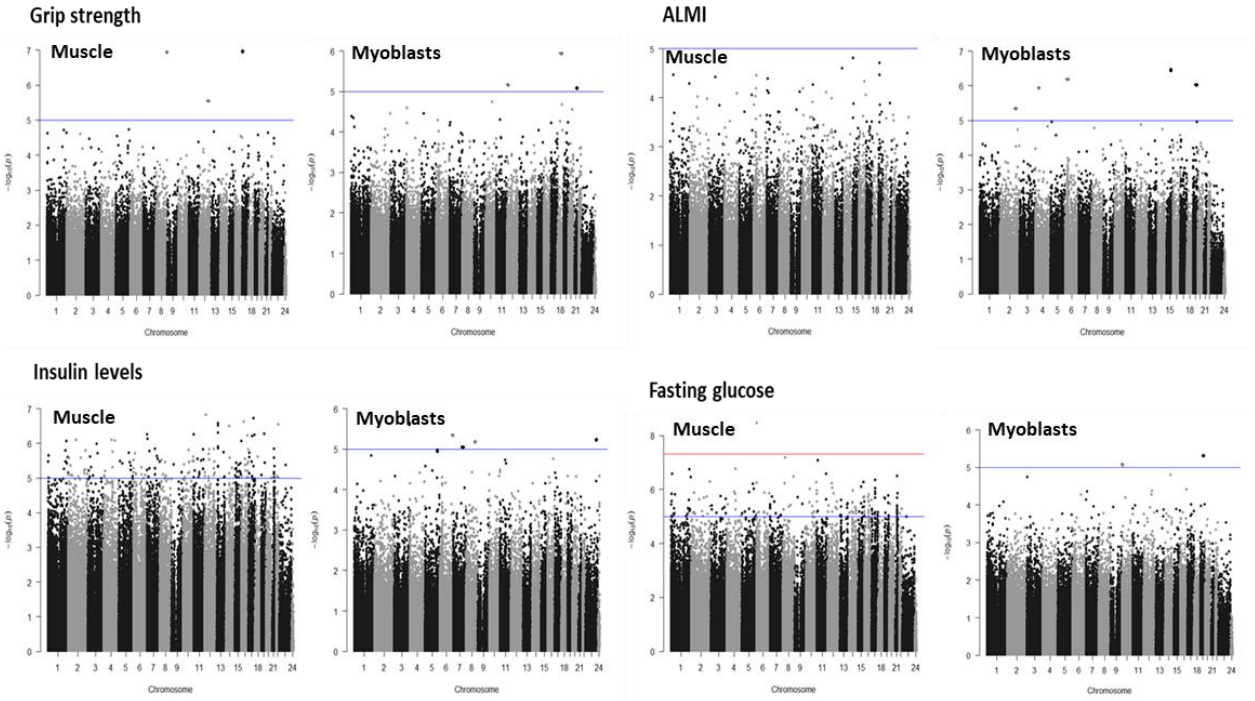


Figure 6.3: Manhattan plots for the differential methylation analysis in muscle and myoblasts The $-\log_{10}$(p values) for individual CpGs across the genome in muscle and myoblasts are plotted, with the chromosome on the x-axis. There are a handle of CpGs differentially methylated with respect to grip strength and ALMI in myoblasts and muscle. There are many dmCpGs with respect to insulin and fasting glucose levels in muscle, but this signal is abolished in the myoblasts. Blue line $-\log_{10}$(p-value) = 5, red line $-\log_{10}$(p-value) = 7.5.

with respect to gait speed were cg18432572 (FDR=0.6848, p=4.80x10^{-6}) in the body of the CYFIP2 gene, and cg16740427 (FDR=0.6848, p=5.02x10^{-6}) within 1500bp of the TSS of the FAM109B gene. There were no dmCpGs (FDR<0.2) with respect to ALMI in the muscle tissue, most likely due to the low number of samples (n=13 with ALMI measurements). The top two dmCpGs with respect to ALMI were cg20052933 (FDR=0.9999, p=6.68x10^{-6}) in the 3'UTR of the FLJ44817 gene, and cg25804146 (FDR=0.9999, p=7.82x10^{-6}) in the body of the VAV1 gene. However, there were many dmCpGs with significant associations with fasting plasma glucose and fasting serum insulin levels (1451 and 2019 dmCpGs with an FDR<0.05 respectively). The top 2 dmCpGs showing an association with respect to fasting plasma glucose levels are cg08932727 located in the body of DUSP22 gene (FDR=0.0015) and cg11064291 located in the intergenic region on chromosome 8 (FDR=0.0122). The top 2 dmCpGs showing significant associations with fasting serum insulin levels were cg04919982 located within 200bp upstream of the TSS of the RERGL gene (FDR=0.0187), and cg06405341 located in the body of the AXIN2 gene (FDR=0.0187). Manhattan plots showing the $-\log_{10}$(p values) for the individual CpGs with respect to grip strength, ALMI, fasting plasma glucose and fasting serum insulin levels can be seen in figure **6.3**.

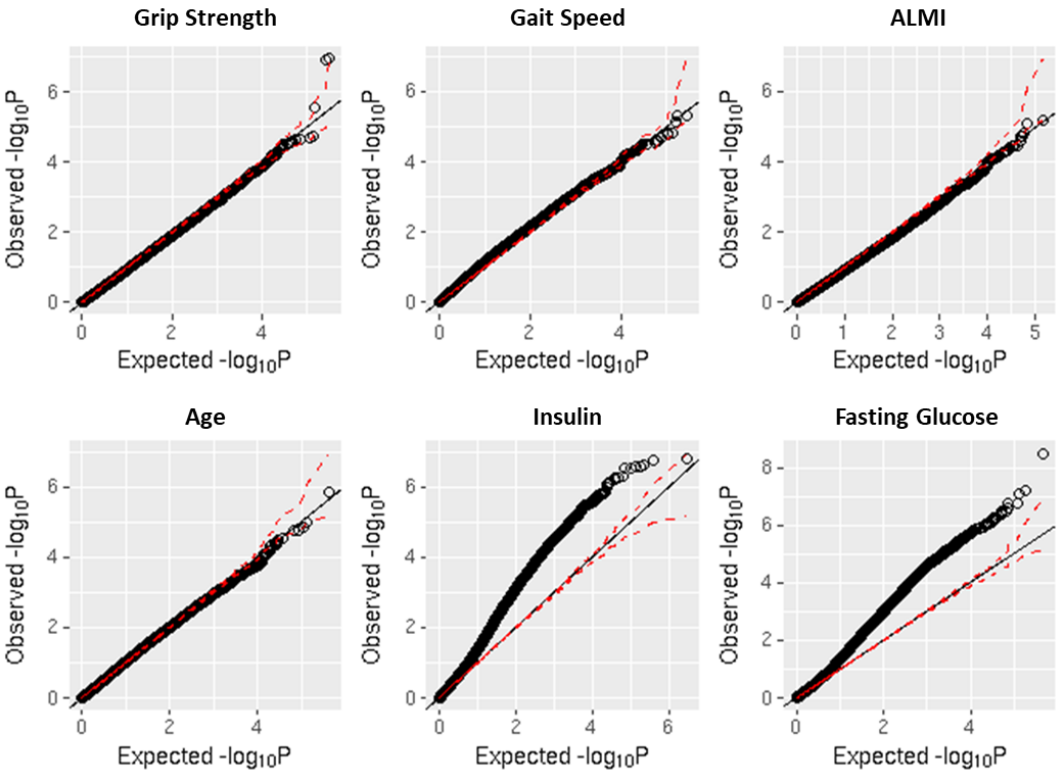


Figure 6.4: Q-Q plots of the observed vs expected p values of the differential methylation analysis in muscle

$-\log_{10}$ of the observed p-values plotted against $-\log_{10}$ of the expected p-values for the analysis with respect to grip strength, gait speed, ALMI, age, insulin and fasting glucose levels. Significant inflation of p-values is observed in the analysis with respect to insulin and fasting glucose levels. Dotted red lines indicate the 95% CI.

Table 6.3: *Genomic inflation factors of the p-values for the differential methylation analysis*

	Genomic Inflation Factor, λ	
	Muscle	**Myoblasts**
Grip Strength	0.868	0.854
ALMI	0.844	0.846
Gait Speed	1.309	0.807
Fasting Glucose	1.436	0.839
Insulin	1.542	0.908
Age	0.990	0.900

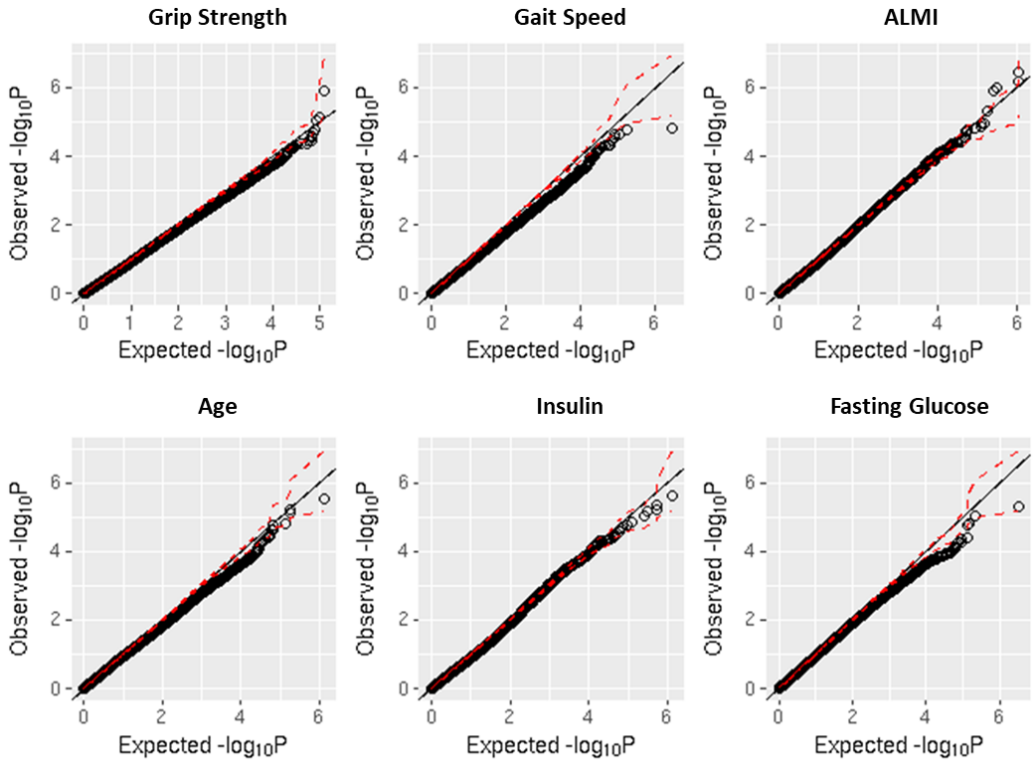


Figure 6.5: Q-Q plots of the observed vs expected p values of the differential methylation analysis in myoblasts

$-\log_{10}$ of the observed p-values plotted against $-\log_{10}$ of the expected p-values for the analysis with respect to grip strength, gait speed, ALMI, age, insulin and fasting glucose levels. No inflation of p-values is observed in any of the analyses. Dotted red lines indicate the 95% CI.

Due to the unexpectedly high numbers of dmCpGs showing significant associations with respect to fasting plasma glucose and fasting serum insulin levels, q-q plots of observed vs expected p-values were examined and genomic inflation factors determined. As evident from figure **6.4**, q-q plots show significant inflation of the observed p values with respect to the expected, and this is further reflected in the genomic inflation factors (λ=1.4365 and 1.5419 for fasting serum insulin and fasting plasma glucose levels respectively)(table **6.3**). This inflation of the observed vs expected p-values was not seen for the muscle-specific phenotypes (e.g. ALMI, grip strength), with all genomic inflation factors < 1.2 (table **6.3**).

6.2.4.2. Differential Methylation in Myoblast Samples

To determine whether the DNA methylation signatures seen in muscle with regards to muscle mass and function were maintained in the cultured myoblasts, DNA from matched myoblasts samples was extracted and interrogated on the Infinium 450K methylation arrays. There were no CpGs found to be differentially methylated with respect to grip strength in the myoblast samples that

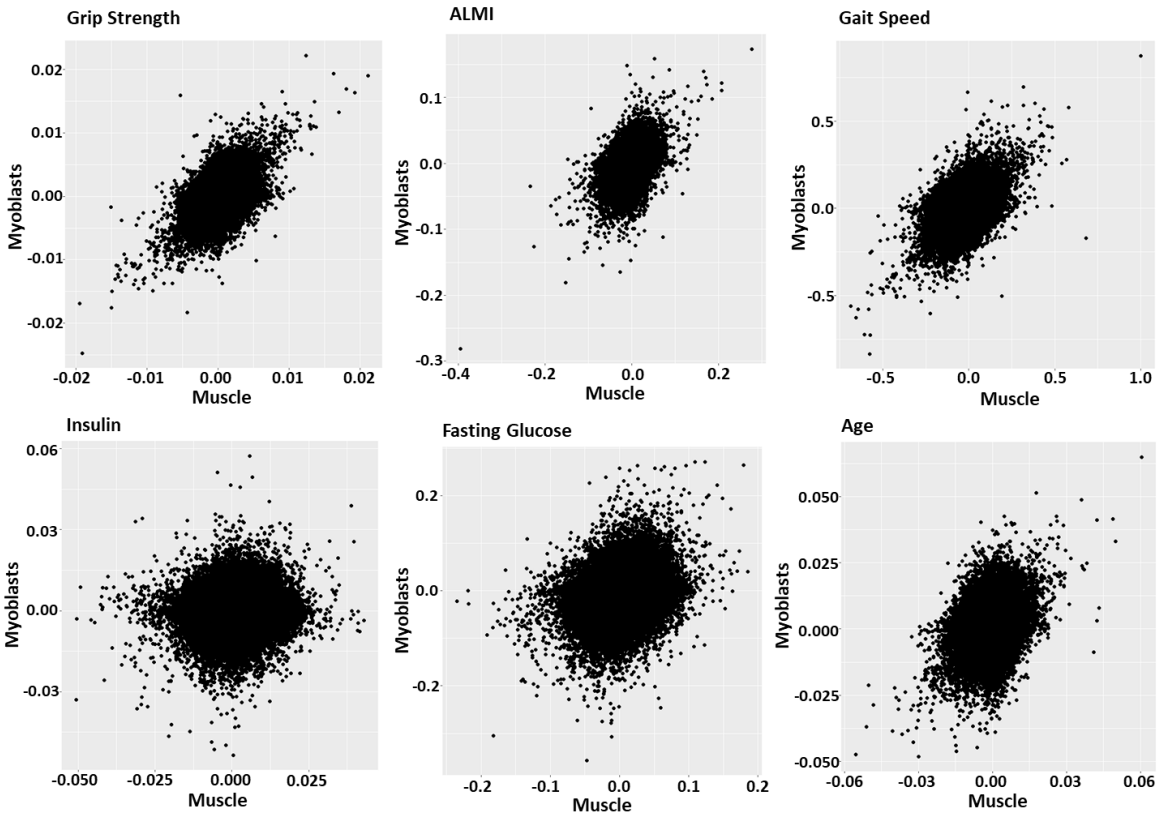


Figure 6.6: Regression coefficients in the myoblasts against those in the muscle samples for all probes.

Regression coefficients are plotted for all the CpGs in the dataset. As is evident, there is a strong correlation between the coefficients for grip strength, ALMI and gait speed. However, these correlations are much weaker for traits associated with metabolic health (insulin and fasting glucose levels).

passed the FDR threshold (FDR<0.05). There were also no CpGs differentially methylated with respect to ALMI at an FDR<0.05, although there were four differentially methylated CpGs with an FDR<0.2 associated with ALMI. These were cg02130555 (FDR=0.1285) located in the 3'UTR region of the NSUN7 gene, cg21505334 (FDR=0.1285) located in the body of the CEACAM5 gene, cg22001110 (FDR=0.1285) located in the 3'UTR region of the C6orf47 gene, and cg23887948 (FDR=0.1285) located in the 5'UTR region of the STRA6 gene. Interestingly, the strong signal seen in the muscle tissue with respect to fasting plasma glucose and fasting serum insulin levels was abolished in the myoblast samples, with no CpG found to be differentially methylated (FDR<0.05).

Due to high level of inflation of observed p values in the muscle samples, q-q plots were examined in the myoblasts for all comparisons. As with the muscle samples, the q-q plots of the observed vs expected p-values for the muscle-specific phenotypes (e.g. ALMI and grip strength) showed no inflation (figure **6.5**), with all genomic inflation factors < 1.2 (table **6.3**). With respect to fasting plasma glucose and fasting serum insulin levels, which showed significant inflation of the p-values

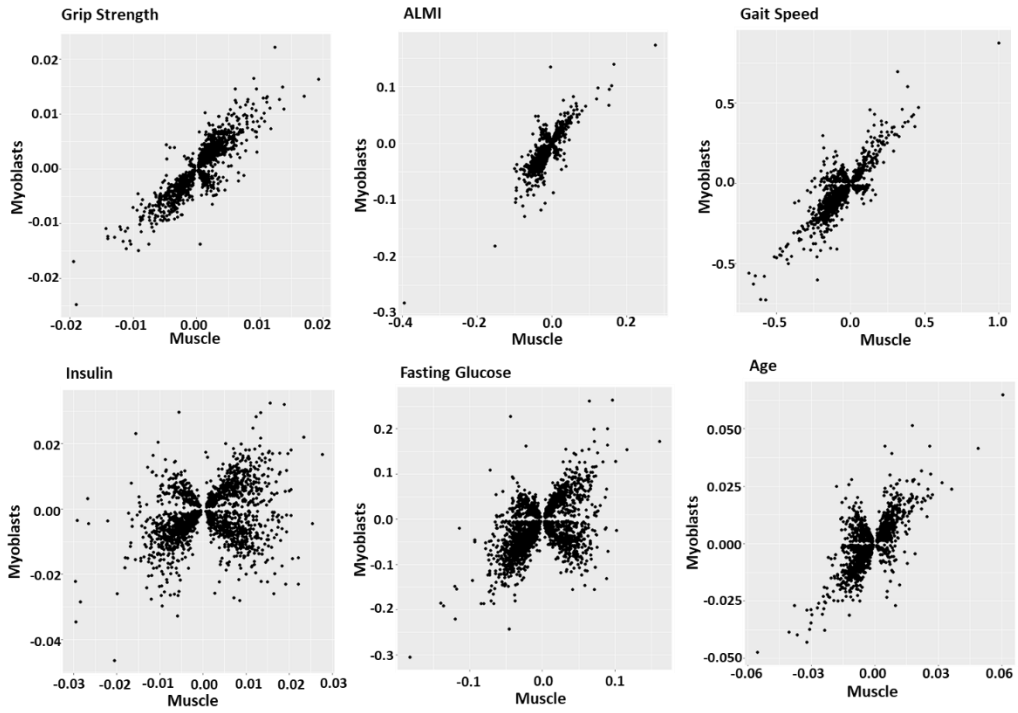


Figure 6.7: Regression coefficients in the myoblasts against those in the muscle samples for probes with p-value <0.05

Regression coefficients are plotted for all the CpGs that have a nominal p-value < 0.05 in each regression analysis. As is evident, there is a very strong correlation between the coefficients for grip strength, ALMI and gait speed. However, these correlations are much weaker for traits associated with metabolic health (insulin and fasting glucose levels).

in the muscle tissue, q-q plots (figure **6.5**) showed that the distributions of the p-values were not inflated in the myoblast samples with genomic inflation factors below 1.2.

6.2.4.3. Myoblasts Retain the Methylation Signature Associated with Muscle Mass and Function

To determine whether the methylation signatures associated with muscle mass and function were retained between the tissues, we looked at the regression coefficients of the regression models and carried out a Spearman correlation of the muscle coefficients vs the myoblast coefficients for all probes (figure **6.6**) and the probes with a nominal p-value < 0.05 (figure **6.7**). Spearman rho coefficients of the correlations between the two are shown in table **6.4**. A strong correlation was seen between the muscle and myoblast samples for the regressions associated with grip strength, gait speed and ALMI. The correlations between the coefficients of the regressions with fasting plasma glucose and fasting serum insulin levels in the myoblasts and muscle tissue was much weaker than those associated with grip strength, gait speed and ALMI. Looking at the probes with a p value < 0.05, similar correlation relationships are seen (figure **6.7**, table **6.4**). This suggests that the inter-individual methylation associations with respect to muscle mass and strength are better conserved between the muscle tissue and myoblast cultures than the methylation associations with traits associated with metabolic health.

Table 6.4: *Spearman rho correlations of regression coefficients between the muscle and myoblasts*

	Spearman rho, ρ	
	All CpGs	**CpGs p-value < 0.05**
Grip Strength	0.175	0.850
ALMI	0.163	0.815
Gait Speed	0.233	0.765
Fasting Glucose	0.056	0.277
Insulin	0.122	0.434
Age	0.101	0.627

6.2.5. Myoblast Functional Characterisation

We have shown epigenetic changes in the myoblasts and muscle samples that are associated with muscle mass and function, and shown correlations between the associations in the two tissues. We next wanted to determine whether there were differences in the functional characteristics and the phenotype of the myoblast samples.

6.2.5.1. Myoblast Proliferation Capacity

To examine the proliferative capacity of the myoblasts and whether any changes in the proliferation were associated with changes in muscle mass and function, the proportion of cells actively dividing was measured in each sample using a BrdU assay, to provide a surrogate measure of the proliferation rate of the myoblasts. There was a positive trend between the BrdU incorporation and ALMI (r^2=0.156, p=0.084), with those showing high ALMI having a higher proportion of myoblasts actively dividing than those with lower ALMI (table **6.5**). There was no association between BrdU incorporation and gait speed or grip strength (tables **6.6** and **6.7**).

6.2.5.2. Senescence

To determine whether changes in muscle mass and function can be attributed to changes in cellular senescence seen in the myoblasts, we examined the levels of cellular senescence in each sample by staining for senescence-associated B-gal (SA-B-gal), a known marker of cellular senescence (figure **6.8**). All samples showed a level of senescence below <10%, with an average of 2.62% (SD=2.16) of cells being senescent, suggesting a low level of cellular senescence. We see a suggestive trend between the levels of cellular senescence and ALMI, with those showing high ALMI showing less

senescent cells than those with low ALMI (table **6.5**). There were no association between the levels of senescent cells and any other phenotypic variable (tables **6.6** and **6.7**).

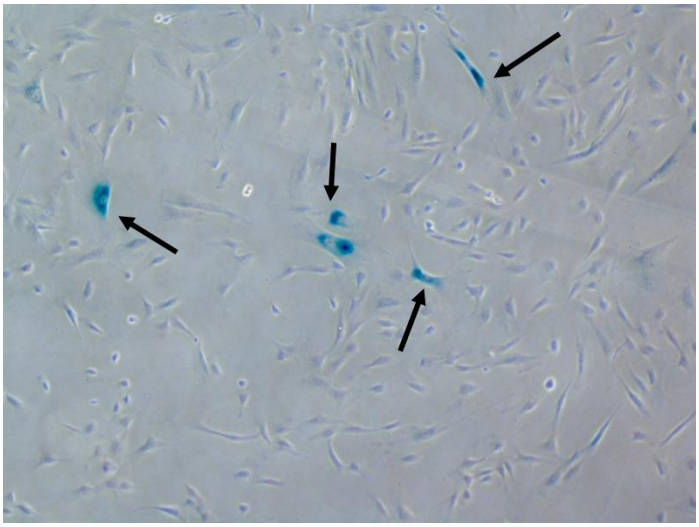


Figure 6.8: Senescence-associated β-gal staining for senescent cells

To examine the level of senescent cells in our myoblast cultures, cells were stained for β-gal activity, a marker of cellular senescence. Arrows indicate senescent cells that are stained blue due to β-gal activity.

6.2.5.3. Differentiation Potential of Myoblasts

As changes in muscle mass and function can be attributed to differences in the differentiation potential of skeletal muscle myoblasts, we differentiated cultured skeletal muscle myoblasts into myotubes and examined the percentage of myoblasts that fused to form multinucleated myotubes. The percentage of multinucleated myotubes was investigated at day 2 (figure **6.9**), day 5 and day 10 of differentiation. The percent of differentiated cells was calculated as the number of nuclei in MHC+ myotubes as a percentage of the total number of DAPI stained nuclei.

At day 2, the average percentage of differentiated MHC+ cells was 5.63% (SD=7.75). There was a significant positive association between the percentage of differentiated cells and gait speed (table **6.6**, r^2=0.203, p=0.018). There was no association between ALMI and grip strength and the differentiation at day 2 (tables **6.5** and **6.7**). At day 5, the average percentage of differentiated MHC+ cells was 19.53% (SD=12.16). There was a higher percentage of differentiated cells at day 5 associated with gait speed (table **6.6**, r^2=0.132, p=0.089). There was no association between the differentiation at day 5 and ALMI (table **6.5**). At day 10, the average percentage of differentiated MHC+ cells was 15.18% (SD=12.93). There was a significant negative association between the percentage of differentiated cells at day 10 and grip strength (table **6.7**, r^2=0.262, p=0.018). There was no association seen between the differentiation at day 10 and ALMI or gait speed (tables **6.5** and **6.6**). A limitation however is that the number of differentiated cells over the 10 days remained relatively low, requiring further culturing of the cells to increase differentiation levels.

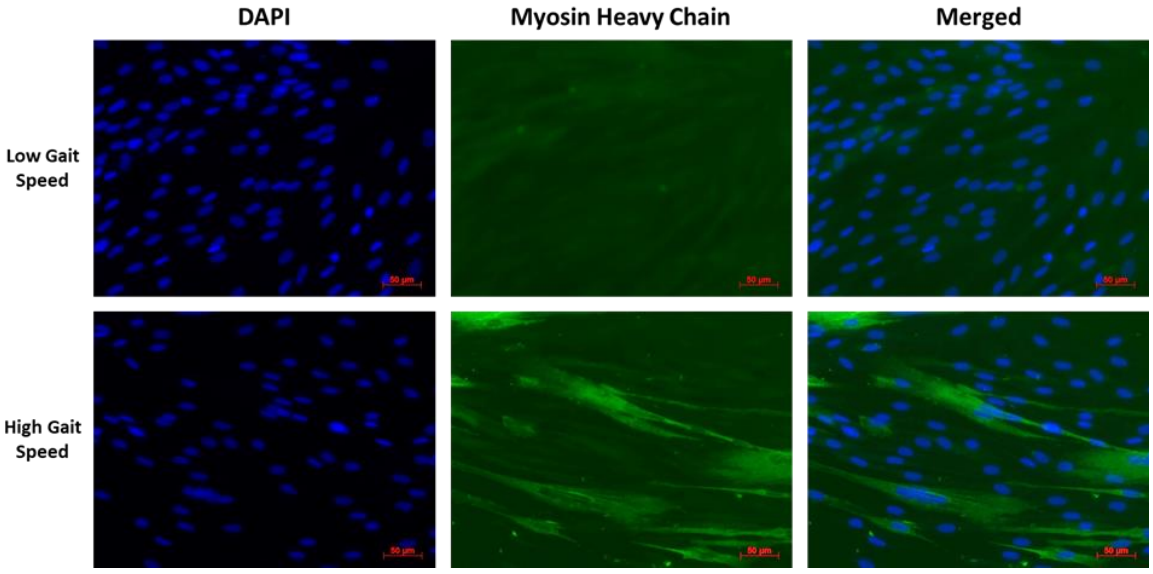


Figure 6.9: Myosin Heavy Chain staining of differentiated myotubes at day 2 of differentiation from participants with low and high gait speed

Top panels are staining from a sample with low gait speed at day 2 and bottom panels are from a sample with high gait speed. DAPI was used to visualize cell nuclei. Scale bars = 50μm.

6.2.5.4. Glycolysis and Mitochondrial Respiration of Myoblasts and Myotubes

As loss of skeletal muscle mass and function has also been attributed to impairments in mitochondrial function, we investigated differences in mitochondrial respiration as well as glycolytic respiration of the cultured myoblasts and differentiated myotubes at terminal differentiation. To determine glycolytic and mitochondrial activity of the myoblasts, we used Seahorse extracellular flux assays (tables **6.5-6.7**).

In the myoblasts, there was a weak positive association between proton leak and grip strength in the myoblasts (r^2=0.179, p=0.082). Maximum respiration showed a weak positive association with gait speed (r^2=0.140, p=0.077), with a greater increase from baseline for those with a faster gait speed. However, both these did not reach statistical significance. There was also a positive association between gait speed and non-glycolytic acidification (r^2=0.263, p=0.050). There was a trend not reaching statistical significance for higher baseline respiration being attributable to ATP production with high ALMI (r^2=0.179, p=0.058), together with a corresponding decrease in non-mitochondrial respiration (r^2=0.274, p=0.009). Myoblasts also showed increased glycolytic capacity (r^2=0278, p=0.033) and decreased non-glycolytic acidification (r^2=0.269, p=0.039) in those with high ALMI.

We also examined the mitochondrial and glycolytic respiration in the cultured myotubes. Grip strength showed a significant positive association with the glycolytic reserve of the myotubes (r^2=0.268, p=0.045). There were no association between the measures of mitochondrial respiration

Table 6.5: *Results of the characterisation of the myoblasts and myotubes with respect to ALMI*

Outcome	Phenotype		N	r^2	p-value
ALMI	Myoblasts	BrdU Incorporation	31	0.156	0.084
		Cellular Senescence	31	0.102	***0.157.***
		ATP Production	29	0.179	***0.058.***
		Maximal Respiration	29	0.06	0.695
		Spare Capacity	29	0.06	0.695
		Proton Leak	29	0.084	0.363
		Non-mitochondrial respiration	29	0.274	***0.009****
		Glycolysis	19	0.054	0.62
		Glycolytic Capacity	18	0.278	***0.033****
		Glycolytic Reserve	19	0.039	0.975
		Non-glycolytic acidification	19	0.269	***0.039****
	Myotubes	Differentiation Potential D2	31	0.051	0.561
		Differentiation Potential D5	31	0.049	0.6
		Differentiation Potential D10	31	0.046	0.658
		ATP Production	30	0.253	***0.012****
		Maximal Respiration	30	0.171	***0.06.***
		Spare Capacity	30	0.171	***0.06.***
		Proton Leak	30	0.055	0.791
		Non-mitochondrial respiration	30	0.205	***0.031****
		Glycolysis	29	0.07	0.515
		Glycolytic Capacity	28	0.049	0.914
		Glycolytic Reserve	28	0.051	0.81
		Non-glycolytic acidification	29	0.142	***0.115.***

and gait speed. However, myotubes show a strong positive association between gait speed and glycolytic capacity (r^2=0.250, p=0.008), glycolytic reserve (r^2=0.310, p=0.002) and non-glycolytic acidification (r^2=0.161, p=0.049). There was also a significant positive association between ATP production at baseline and ALMI (r^2=0.253, p=0.012), together with increased maximal respiration and spare capacity in the myotubes (r^2=0.171, p=0.060) and increased non-mitochondrial respiration (r^2=0.205, p=0.031).

6.2.5.5. Correlations between Myoblast/Myotube Molecular Characteristics

To determine whether the different phenotypes of the myoblasts and myotubes were associated, we examined the correlations between the different molecular characteristics identified in this study. The percent of multinucleated myotubes was calculated as the number of nuclei in MHC+ myotubes with more than 1 nucleus as a percentage of the total number of DAPI stained nuclei. The percentage of multinucleated myotubes was significantly positively correlated between day 2 and day 5 (rho = 0.685, p<0.001), day 2 and day 10 (rho = 0.462, p=0.005) and between day 5 and day 10 (rho = 0.580, p<0.001). BrdU incorporation and the percentage of senescent myoblasts also showed a significant negative correlation (rho = -0.368, p=0.027), while there was a weak positive trend between BrdU incorporation and the percentage of multinucleated myotubes at day 2 (rho = 0.291, p=0.085) and a suggestive trend at day 5 (rho = 0.276, p=0.103), although both failed to reach statistical significance.

Table 6.6: *Results of the characterisation of the myoblasts and myotubes with respect to gait speed*

Outcome	Phenotype		N	r^2	p-value
Gait Speed	Myoblasts	BrdU Incorporation	35	0.285	0.311
		Cellular Senescence	35	0.049	0.962
		ATP Production	33	0.046	0.795
		Maximal Respiration	33	0.14	***0.077.***
		Spare Capacity	33	0.14	***0.077.***
		Proton Leak	33	0.074	0.333
		Non-mitochondrial respiration	33	0.054	0.578
		Glycolysis	21	0.087	0.728
		Glycolytic Capacity	20	0.123	0.391
		Glycolytic Reserve	21	0.135	0.304
		Non-glycolytic acidification	21	0.263	***0.05****
	Myotubes	Differentiation Potential D2	35	0.203	***0.018****
		Differentiation Potential D5	35	0.132	***0.089.***
		Differentiation Potential D10	35	0.049	0.935
		ATP Production	34	0.05	0.648
		Maximal Respiration	34	0.058	0.499
		Spare Capacity	34	0.058	0.499
		Proton Leak	34	0.069	0.37
		Non-mitochondrial respiration	34	0.046	0.783
		Glycolysis	33	0.102	***0.175.***
		Glycolytic Capacity	32	0.25	***0.008****
		Glycolytic Reserve	32	0.31	***0.002****
		Non-glycolytic acidification	33	0.161	***0.049****

Table 6.7: *Results of the characterisation of the myoblasts and myotubes with respect to grip strength*

Outcome	Phenotype		N	r^2	p-value
Grip Strength	Myoblasts	BrdU Incorporation	36	0.147	0.852
		Cellular Senescence	36	0.131	0.58
		ATP Production	34	0.112	0.884
		Maximal Respiration	34	0.126	0.468
		Spare Capacity	34	0.126	0.468
		Proton Leak	34	0.195	**0.082.**
		Non-mitochondrial respiration	34	0.17	**0.149.**
		Glycolysis	22	0.133	0.83
		Glycolytic Capacity	21	0.128	0.907
		Glycolytic Reserve	22	0.133	0.83
		Non-glycolytic acidification	22	0.194	0.238
	Myotubes	Differentiation Potential D2	36	0.126	0.719
		Differentiation Potential D5	36	0.173	**0.168.**
		Differentiation Potential D10	36	0.262	**0.018***
		ATP Production	35	0.165	**0.16.**
		Maximal Respiration	35	0.142	0.295
		Spare Capacity	35	0.142	0.295
		Proton Leak	35	0.138	0.329
		Non-mitochondrial respiration	35	0.116	0.686
		Glycolysis	34	0.111	0.923
		Glycolytic Capacity	33	0.228	**0.116.**
		Glycolytic Reserve	33	0.268	**0.045***
		Non-glycolytic acidification	34	0.114	0.764

As the percentage of basal respiration attributable to ATP production was significantly associated with ALMI in both myoblasts and myotubes, ATP production was correlated with differentiation potential, cellular senescence and BrdU incorporation. Although the overall level of senescence was relatively low in all cultures (average=2.62%), we surprisingly found a significant negative correlation between ATP production in the myoblasts and the percentage of senescent cells (rho = -0.447, p=0.007) and a suggestive trend with BrdU incorporation (rho = 0.272, p=0.113). ATP production in the myotubes showed a significant positive correlation with the percentage of multinucleated myotubes at day 2 (rho = 0.436, p=0.01), a possible trend at day 5 (rho = 0.250, p=0.154) and at day 10 (rho = 0.317, p=0.068), although further replication is required to validate the results at the later 2 time points.

6.3. Discussion

In this study, we have shown methylation changes in muscle tissue associated with the loss of muscle mass and function, as well as with metabolic outcomes. As DNA methylation is one of the key mechanisms that confers tissue identity, we then looked at the methylation in matched myoblasts cultured from the muscle tissue, and found that the methylation changes show a relatively strong correlation between muscle tissue and myoblasts with respect to muscle-related phenotypes (e.g. grip strength and ALMI). However, methylation changes associated with metabolic traits were less correlated between muscle tissue and myoblasts, suggesting different signals from the different fibre types and other resident present in the muscle, which all contribute to the overall methylation signature associated with skeletal muscle. Although, this may be due to the culture conditions of the myoblasts, which were grown in media containing high levels of glucose. To further investigate the suitability of using cultured human myoblasts as an *in vitro* model for the investigation of muscle function, we characterised the myoblasts at a molecular and cellular level. We found that BrdU incorporation, level of senescence, differentiation potential, glycolysis and mitochondrial respiration showed suggestive correlations with measures of ALMI, gait speed and grip strength, although it would be beneficial to repeat in a larger sample group to consolidate the results.

6.3.1. Cultured Human Myoblasts as an *in vitro* model

As humans are living longer, sarcopenia is becoming an ever growing problem, with the muscle weakness and impaired function being key components of the geriatric syndrome of frailty. Therefore, it is key to understand the mechanisms that contribute to normal muscle ageing in the elderly and to understand the variability observed with respect to skeletal muscle loss. Currently, there are limited pharmacological interventions approved for the treatment of sarcopenia [453].

Preclinical studies using *in vitro* models are an important translational link to aid in the interpretation of the results of human trials in terms of understanding the biological mechanisms of the interventions [454].

6.3.1.1. Muscle-Derived Myoblasts Retain an Epigenetic Memory

To determine how suitable cultured myoblasts are for the investigation of therapeutics and in understanding the mechanisms of muscle ageing, we sought to compare the muscle biopsies and myoblasts at an epigenetic level, as well as at a molecular level. There were very few differentially methylated CpGs (dmCpGs) in the muscle tissue with respect muscle related traits (e.g. ALMI and grip strength) and many more dmCpGs with respect to metabolic traits (e.g. fasting plasma glucose and fasting serum insulin levels). As muscle tissue consists of muscle cells, as well as non-muscle cells (e.g. adipose, cartilage and blood vessels) we used myoblast cells as a homogenous cell surrogate for muscle tissue, and found that despite there still being less dmCpGs with respect to ALMI and grip strength, the signal associated with fasting plasma glucose and fasting serum insulin levels was reduced. This suggests that the cellular heterogeneity of muscle tissue is contributing to the methylation signal in muscle tissue, whereas when those cells are removed and only myoblast cells interrogated, the methylation signature associated with the metabolic traits is no longer present. The correlations between the associations of dmCpGs with the muscle-related traits in the myoblasts and muscle tissue were strong. However, these correlations were much weaker for the metabolic traits. This suggests that the myoblasts, although they have been manipulated in culture, retain a memory of the phenotype of the muscle tissue they originated from, such that myoblasts from those that have low ALMI retain a low ALMI phenotype, allowing investigation into the mechanisms contributing to decreased muscle mass in an *in vitro* cell model. This was further reflected in the principle components analysis and q-q plots. PCA showed the myoblasts samples were groups together based on their methylation profiles, whereas the muscle samples were grouped together and separate from the myoblasts. As myoblasts are cells of myogenic origin, the independent grouping of muscle samples suggests that non-myogenic cells are present in muscle and contribute to the methylation profiles. Although it may be the different fibre types present in native muscle tissue that may contribute to the muscle samples grouping away from the myoblasts. The Q-Q plots showed significant inflation in the muscle samples with respect to metabolic traits, inflation that is not seen in the myoblast samples. There was no inflation for the muscle-related phenotypes. Inflation of p-values is a sign of increased false positive results due to increased type I error. This can be due to unknown confounders in the analysis that have not been accounted for. As we have not corrected for cell type composition in the muscle analysis, it is possible that non-myogenic cells in the muscle sample may be contributing to the differential methylation associated with the metabolic traits. It may also be confounded by the different muscle fibre types present,

which may contribute different methylation signals. As the different muscle fibre types exhibit different metabolic and functional characteristics, this may be partly driven by differences in DNA methylation, allowing fibre-specific expression of genes. This fibre-specific methylation may be a potential confounder and may be affecting the overall methylation signature seen in the skeletal muscle. As the different muscle fibre types also exhibit different glucose handling capabilities, this may also be affecting the associations between the muscle methylation and metabolic phenotypes examined in this study.

However, there are some limitations with this analysis. This investigation was carried out in low numbers of samples, which could help explain the low numbers of dmCpGs that pass below the FDR threshold. Repeating the investigation in a larger number of samples will provide a better insight into dmCpGs associated with changes in muscle-related phenotypes. Another limitation with respect to analysing differential methylation in the myoblasts is that the myoblasts were grown in culture medium containing glucose. As such, the cells were responding to the available glucose that could alter the methylation of CpGs, which may suggest why there were few dmCpGs with respect to fasting plasma glucose and fasting serum insulin levels in the myoblasts.

6.3.1.2. Muscle-Derived Myoblasts Retain the Functional Characteristics of Muscle

Functional characterisation of the myoblast cultures have shown that despite the manipulation and *in vitro* culturing of the myoblasts, they retain similar characteristics as to those from the muscle which they originate from. Cellular proliferation has been shown to be reduced in aged skeletal muscle [455], resulting in a reduction in the pool of muscle progenitor cells [456]. Sousa-Victor et al. [194] have also shown that geriatric satellite cells transition from a reversible quiescent state into a senescent state, which inhibits proliferation of the cells, resulting in further reduction of the satellite cell pool. There has also been evidence that a reduction in satellite cell numbers contributes to skeletal muscle fibre atrophy [455], with a loss of nuclei from large fibres, accompanied by a decrease in overall fibre size and an increase in fibre atrophy [457]. Here, we have found that those with low ALMI show a suggestive trend towards high levels of cellular senescence as well as a decreased proliferation rate of the myoblasts. Although it would not have been possible to culture the senescent cells from the muscle biopsies obtained, the association between ALMI and levels of cellular senescence in the cultures myoblasts suggests that skeletal muscle from people with low ALMI contain myoblasts more susceptible to undergoing senescence and are possibly more prone to senesce compared to cells from the skeletal muscle of people with higher ALMI. Although this was seen in myoblasts rather than satellite cells, myoblasts form part of the muscle progenitor cell population, important for regeneration and repair of skeletal muscle. This suggests that the decreased proliferation and increased senescence of the myoblasts may reduce the ability of the

myoblasts to proliferate, to self-renew or to differentiate, contributing to the reduction of muscle mass seen in sarcopenia.

A reduction in the differentiation potential of myoblasts has also been reported in aged muscle [458,459]. When skeletal muscle satellite cells are activated, they transition from quiescence into proliferative myoblasts, which subsequently differentiate into myotubes and fuse into myofibres. A reduced ability of myoblasts to differentiate may affect muscle hypertrophy and the repair of skeletal muscle tissue in response to damage. The proliferation and differentiation of myoblasts into myotubes is important in repairing damage to muscle fibres in response to exercise or injury. A decreased ability of myoblasts to differentiate and fuse into myotubes and form fibres will affect muscle mass and the ability of the muscle to function normally. We have shown that the differentiation potential of myoblasts at several time points during differentiation is impaired in the myoblasts from those with low muscle function (gait speed and grip strength). This suggests that a reduction in the ability of the myoblasts to differentiate as effectively may impair muscle function and strength.

Mitochondrial function is important for normal muscle function, as skeletal muscle is a highly metabolically active tissue. Mitochondrial dysfunction has been repeatedly shown to contribute to the development of sarcopenia [146-148], resulting in a reduction of muscle mass and function. We have previously shown that there are alterations in the skeletal muscle transcriptome associated with altered muscle mass (Chapter 4, unpublished data). We have shown that in those with low ALMI, there were changes in many pathways associated with mitochondrial function and oxidative phosphorylation. Here, we have shown that the myoblasts have altered mitochondrial and glycolytic respiration, with those having a low ALMI showing reduced ATP production and glycolytic capacity. When differentiated into myotubes, there is also reduced ATP production, accompanied by a reduction in the maximal respiration and spare capacity of the myotubes. Those with a low gait speed also show a reduction in the maximal respiration and spare capacity of the myoblasts and a reduction in glycolytic capacity and glycolytic reserve in the myotubes, which is also seen in those with low grip strength. This suggests that a reduction in the normal mitochondrial respiration and glycolysis in the myoblasts and myotubes may be impairing the function of the cells. The impaired respiration in the myoblasts may contribute to the reduced differentiation and proliferation potential of the cells, which affects the renewal of the satellite cell pool and repair of skeletal muscle. The fact that changes in mitochondrial respiration is seen in the myoblasts as well as the myotubes suggests that the muscle may not be able to function as efficiently, resulting in reduced force generation and strength.

6.3.2. Further Considerations

Although here we have shown that human skeletal muscle-derived myoblasts retain an 'epigenetic memory' of the tissue they originate from, as well as the functional characteristics of the tissue, there are some limitations in using primary myoblasts. The purity of cultures needs to be carefully investigated to ensure a pure culture of myoblasts with no infiltrating fibroblasts. Primary cultures also have a limited proliferative capacity, with an increased level of senescence seen with increased culture *in vitro.* Therefore, immortalization of myoblasts may be required, for which further investigation is needed to ensure cells retain an epigenetic and functional 'memory' of the tissue they originated from. Another limitation of this study is that the epigenetic analysis was carried out on a small number of samples, and results would need to be validated and replicated using other technologies and in other cohort samples.

6.3.3. Conclusions

Here, we have provided evidence that skeletal muscle-derived myoblasts may provide an *in vitro* model for the investigation into the mechanisms that contribute to the pathogenesis of sarcopenia in the elderly and the variability of sarcopenia symptoms seen in the elderly. These myoblasts also allow the testing of therapeutics and interventions in a preclinical setting to test the efficacy and efficiency of new compounds to reverse the effects of sarcopenia and prevent further loss of muscle mass and function.

Chapter 7 –

Genome-wide methylation analysis human skeletal muscle obtained from older people

Chapter 7 – Genome-wide methylation analysis human skeletal muscle obtained from older people

7.1. Introduction

DNA methylation is a modification to DNA, involving the covalent addition of a methyl group to the fifth position on the cytosine, usually in a CpG dinucleotide. DNA methylation is among the most studied epigenetic modification, one that is essential for normal mammalian development. Changes in DNA methylation have been shown to be associated with many age-related diseases, including cancer [314], diabetes [315] and Alzheimer's disease [316]; highlighting the importance of understanding DNA methylation changes in ageing skeletal muscle tissue.

7.1.1. Molecular Function of DNA Methylation

Three enzymes catalyse DNA methylation: DNMT1, which adds methyl groups to replicating DNA, and DNMT3a and DNMT3b, which add new methylation marks, primarily during development [321]. Once catalysed by these enzymes, DNA methylation marks are generally considered to be relatively stable, and are transmitted from cell to cell. However, there is increasing evidence that the methylome is especially susceptible to environmental factors in early life although some epigenetic plasticity has been reported in adulthood.

Once generated, DNA methylation patterns regulate the activity of a gene. In a given organism (e.g. humans), all cells contain an identical genome, although gene editing in immune cells (B cells) may contain different rearrangements of the immunoglobulin gene loci. However, cells express different genes allowing them to have distinct functions and phenotypes. DNA methylation is one of the key mechanisms regulating gene expression to allow temporal and spatial differences in gene expression. DNA methylation is highly abundant at gene promoter regions, and is usually associated with the repression of gene expression. DNA methylation inhibits gene expression through two mechanisms: Firstly, DNA methylation represses gene expression by interfering with the binding of transcription factors to the promoter of genes, inhibiting the transcription of the gene [460]. The binding of transcription factors (TFs) to gene promoters is a key process to initiate transcription of genes and therefore inhibiting this process affects downstream gene transcriptional activation. Secondly, DNA methylation impairs gene expression by the recruitment of methyl-CpG binding proteins (e.g. MeCP2 and MBD1), which preferentially bind to methylated-CpGs [461]. These proteins are associated with repressor complexes (e.g. MeCP2 with mSin3a and MBD1 with SETDB1), which

are recruited to the methylated DNA region. The repressor complexes subsequently alter chromatin structure by acting on the histone proteins, leading to a compacted chromatin structure at the region of DNA methylation, repressing the expression of genes in that region.

7.1.2. DNA Methylation in Skeletal Muscle

Very few studies have examined methylation changes across the genome in skeletal muscle with respect to ageing and these have not looked at those methylation changes with respect to muscle function. There have been many studies investigating the role of epigenetics and peripheral blood DNA methylation during ageing. For example, Bjornsson et al. [462] showed a >20% change DNA methylation in peripheral blood associated with ageing, supporting the hypothesis that the age-related loss of normal epigenetic marks may contribute to the onset of disease in later life.

A handful of studies have examined methylation changes in skeletal muscle, and how different environmental factors affect the skeletal muscle methylome. Barrès et al. [463] reported that a bout of acute exercise was sufficient to induce a decrease in global DNA methylation in skeletal muscle tissue, together with reduced DNA methylation at the promoters of genes including PGC1α, TFAM and MEF2A. This was accompanied by an increase in their expression. Jin et al. [356] interrogated genome-wide methylation changes in the skeletal muscle of young and middle-aged pigs as a model for ageing in human studies. A global reduction in methylation was observed in the older skeletal muscle, including methylation changes in age-related genes including FOXO3 and FGFR1. They concluded that changes in DNA methylation may contribute to increased proteolysis and protein catabolism with ageing, suggesting a link between DNA methylation and age-related skeletal muscle atrophy. Zykovich et al. [355] looked at age-related changes in the skeletal muscle methylome. They found that with increased age, there was a predominant hypermethylation phenotype with differentially methylated CpGs (dmCpGs) found primarily in intragenic regions rather than at gene promoters. There was also an enrichment for differentially methylated CpGs in genes related to pathways associated with muscle innervation (e.g. axon guidance) and cytoskeletal organization, a key component of the contractility of skeletal muscle and normal muscle function.

However, the majority of studies to date have investigated changes in DNA methylation between young and old skeletal muscle, with no studies investigating DNA methylation changes between sarcopenic and healthy skeletal muscle, providing evidence to explain the variability in muscle loss seen in older people.

7.1.3. Aims

The aim of this study was to investigate how the skeletal muscle epigenome is altered in older people with varying degrees of skeletal muscle loss. Infinium MethylationEPIC arrays were used to determine the methylation level at up to 850,000 CpGs in the human genome, to determine whether there were any associations between the methylation profiles and different muscle-associated phenotypes, including skeletal muscle mass, grip strength and fasting glucose levels. Differentially methylated CpGs were mapped to genes in order to determine whether there are common pathways perturbed in relation to muscle mass and function in the elderly.

7.2. Results

7.2.1. Cohort Characteristics

Summary of anthropometric and physical function of participants of HSS (n=40) and HSSe (n=43) used for methylation analysis are shown in table **7.1**. All participants from both cohorts were men. In HSS, the mean age (S.D.) was 72.91 (2.44) years, height 1.72 (0.06) m, weight 80.63 (13.69) kg, BMI 27.29 (3.62) kg/m^2, total lean body mass 57.24 (7.55) kg, appendicular lean mass 23.43 (3.58) kg, total fat mass 22.10 (7.25) kg, gait speed 1.10 (0.19) m/s and grip strength 37.08 (7.17) kg. In HSSe, the mean age (S.D.) was 78.11 (2.63) years, height 1.73 (0.06) m, weight 78.06 (10.96) kg, BMI 26.13 (3.24), total lean body mass 48.52 kg (5.63) kg/m^2, appendicular lean mass 21.25 (2.80) kg, total fat mass 25.42 (8.13) kg, gait speed 1.00 (0.22) m/s, grip strength 35.74 (6.80) kg, fasting plasma glucose 7.84 (5.14) mmol/l and fasting serum insulin 5.48 (0.62) mU/l. For the combined analysis, the mean age (S.D.) was 75.60 (3.64) years, height 1.72 (0.06) m, weight 79.30 (12.34) kg, BMI 26.69 (3.45) kg/m^2, total lean body mass 53.05 (7.96) kg, appendicular lean mass 22.37 (3.39) kg, total fat mass 23.7 (7.81) kg, gait speed 1.05 (0.21) m/s and grip strength 36.39 (6.97) kg. Fasting plasma glucose and serum insulin measures were only available for the HSSe participants.

Table 7.1: *Cohort characteristics for participants of HSS, HSSe and combined datasets used in the methylation analysis*

	HSS (n=40)		HSSe (n=43)		Combined (n=83)	
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Age (yrs)	72.91	2.44	78.11	2.63	75.60	3.64
Height (m)	1.72	0.06	1.73	0.06	1.72	0.06
Weight (kg)	80.63	13.69	78.06	10.96	79.30	12.34
BMI (kg/m^2)	27.29	3.62	26.13	3.24	26.69	3.45
Total Lean Body Mass (kg)	57.24	7.55	48.52	5.62	53.05	7.96
Appendicular Lean Mass (kg)	23.43	3.58	21.25	2.80	22.37	3.39
Total Fatmass (kg)	22.10	7.25	25.42	8.13	23.70	7.81
Gait Speed (m/s)	1.10	0.19	1.00	0.22	1.05	0.21
Grip Strength (kg)	37.08	7.17	35.74	6.80	36.39	6.97
Fasting Glucose (mmol/l)	NA	NA	7.84	5.14	7.84*	5.14*
Insulin (mU/l)	NA	NA	5.48	0.62	5.48*	0.62*

7.2.2. Beta Values QC

Several quality control steps were carried out at the level of beta values. Beta values represent the ratio of the methylated and unmethylated signals (green and red channels respectively), generating methylation levels in the range of 0-1. To obtain beta values, idat files were read into an RGset, a minfi RGChannelSet object (a specialized R data structure) that stores the raw intensities in the green and red channels. One of the samples (HSS19) contained no red or green intensities so was discarded from further analysis. Before any QC, raw intensities were normalized using the Funnorm (functional normalization) function of minfi to generate beta values for each probe.

Initially, hierarchal clustering was carried out on all samples, based on their Euclidean distances to generate a dendogram for all 87 samples in order to look at the clustering of inter- and intra-chip replicates (figure **7.1**). MB88 and MB99 had inter- and intra-chip replicates while MB89 only had intra-chip replicates. All intra-chip replicates clustered together on the same branch of the dendrogram suggesting good results across the whole beadchip. MB88 inter-chip replicates also clustered together suggesting good results across different beadchips. However, MB99 inter-chip replicate (blue arrow, figure **7.1**) clustered away from the other two replicates on the dendogram, which were placed on a different beadchip, suggesting slight differences between the different beadchips. As such, the beadchip each sample was run on must be accounted for in downstream analysis.

Secondly, a principal components analysis (PCA) plot was produced to determine the clustering of samples and whether any outlier samples were present in the dataset. A PCA plot of the first two principle components was generated (figure **7.2**). Samples did not cluster based on their sarcopenia status (figure **7.2a**) or which beadchip the samples were run on (figure **7.2b**). Despite no clustering by PCA based on their beadchip, hierarchal clustering suggested that beadchip may affect downstream analysis.

Finally, median absolute deviation scores (MADscores, modified z-scores) were calculated for all samples to determine any sample outliers in the dataset (table **7.2**). Samples with a madscore < -5 were classed as outliers. Six samples had a madscore < -5 and so were excluded from the dataset and from further downstream analysis. All the samples with a madscore < -5 were on different beadchips. Technical replicates were removed from the dataset based on their madscores, leaving a total dataset with 77 samples.

7.2.3. Beta Values Probe QC

CpG probes with a detection p-value > 0.01 or with a beadcount < 3 were set to missing. CpG probes were excluded from subsequent analysis if missing in more than 5% of samples. 1763 probes were

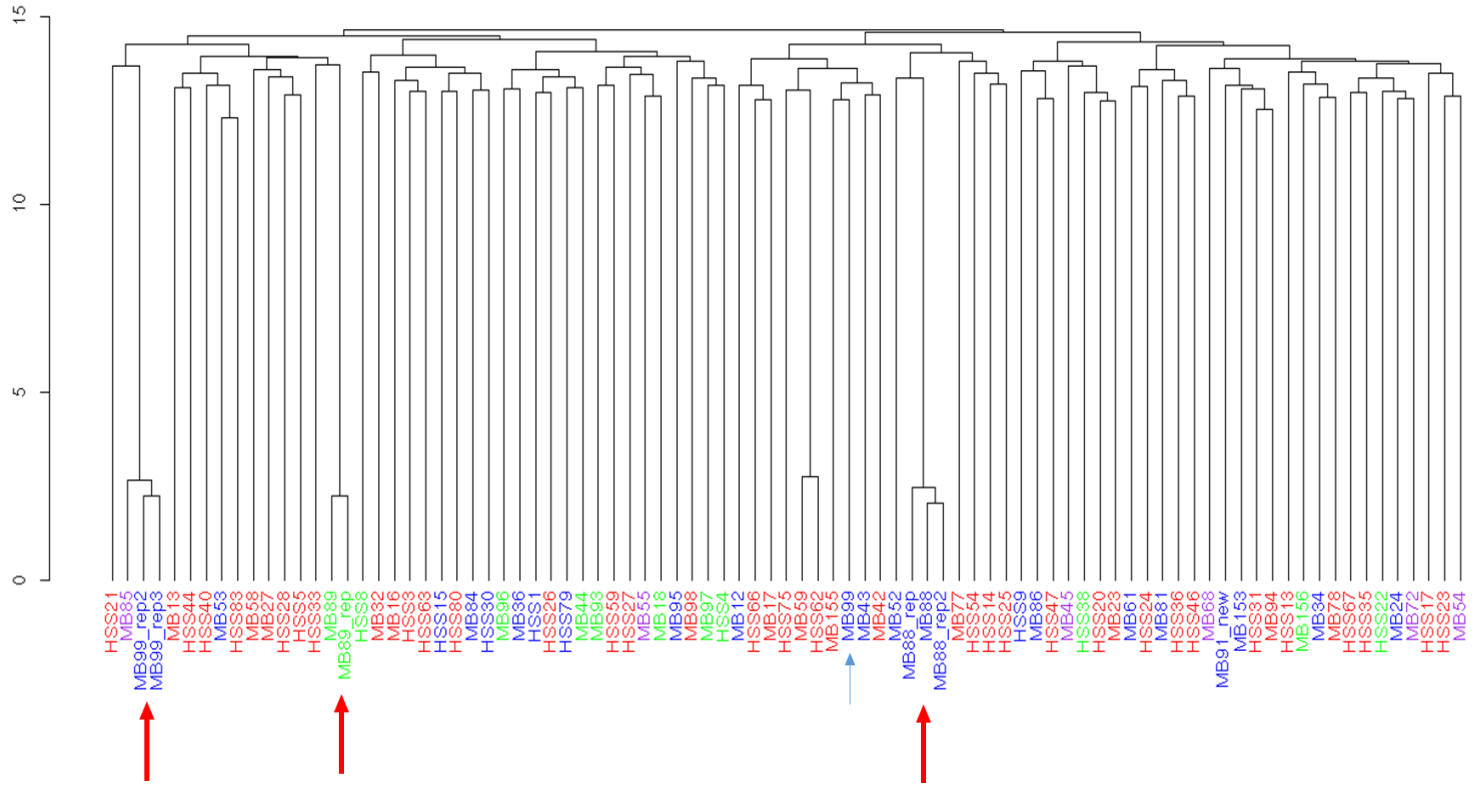


Figure 7.1: Hierarchal clustering of samples based on their Euclidean distances

Hierarchal clustering of normalized beta values based on their Euclidean distance. There is no clear clustering of the samples evident from the dendogram. Intra-chip replicates all clustered together (red arrows) suggesting good quality across the beadchips. MB88 inter-chip replicates clustered together, however MB99 replicates did not cluster together (blue arrow) suggesting possible batch effects between the different beadchips. Samples are coloured based on their sarcopenia status: red=control, blue=pre-sarcopenia, green=sarcopenia, purple=unknown status.

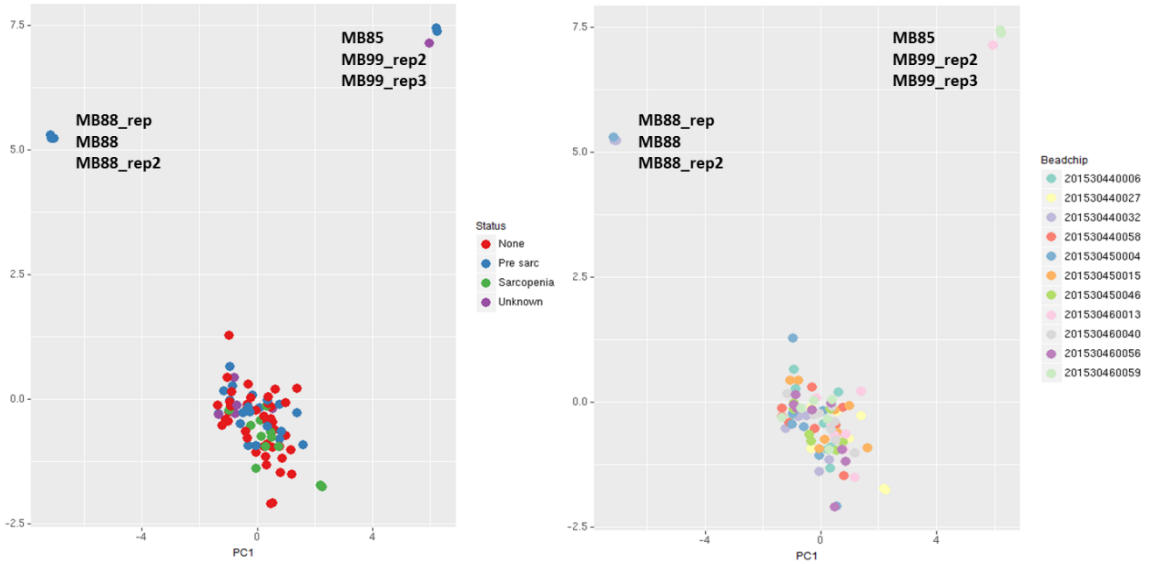


Figure 7.2: PCA plots of samples for the first two principle components

Principle component plots of all samples following data normalization. There is no clear clustering of the samples by their first two principle components, based on their sarcopenia status (A) or the beadchip each sample was run on (B). There do appear to be six outlier samples that cluster away from the rest, requiring further investigation. Samples are coloured by their sarcopenia status in A and by their beadchip in B.

excluded based on their detection p-value and a further 1533 probes excluded based on a low bead count. The detection P value was used to determine whether a probe's signal was above background. A set of negative control probes was used to detect background intensity, with a small p value indicating a signal above background, and likely to be a true signal. 29,707 SNP probes were removed from the dataset, including probes that had a SNP at the CpG site and probes that had a SNP at the single nucleotide extension. Pidsley et al. [370] have previously shown that there are probes on the 850K array that cross-hybridize to multiple different regions in the genome, resulting in 40,846 probes being removed from the dataset. CpG probes with a methylation range <5% were also removed from the dataset. These probes were removed, as there is conflicting data in the literature about the biological impact of CpGs with small methylation changes. Removal of CpGs with a small methylation range also provides a better ability to validate the methylation array using bisulfite pyrosequencing, as small methylation differences (<5%) prove difficult to validate due to the technical capabilities of pyrosequencing. Due to technical replicates not clustering together in the hierarchal clustering, beta values were adjusted for batch (beadchip each sample was run on) using ComBat to obtain batch-adjusted beta values that can be investigated further.

Table 7.2: *Table of madscores (modified z-scores) for all samples*

ID	Chip	z-score	ID	Chip	z-score	ID	Chip	z-score
MB61	1	0.152	MB59	5	-3.902	HSS80	9	-0.077
MB27	1	0.205	HSS22	5	-2.784	MB34	9	-0.612
HSS31	1	-5.936	MB23	5	0.064	HSS75	9	-3.747
MB58	1	0.243	MB88_rep2	5	-0.486	HSS9	9	0.325
HSS8	1	0.679	HSS35	5	1.413	HSS25	9	0.927
MB52	1	0.677	MB12	5	-0.771	HSS38	9	0.966
HSS27	1	-4.95	HSS66	5	-0.826	MB13	9	-2.141
HSS3	1	-3.794	MB81	5	-0.072	MB99	9	-3.472
HSS54	2	-0.032	MB68	6	-11.196	HSS17	10	0
MB24	2	0.786	MB17	6	0.724	HSS79	10	0.232
MB94	2	-0.843	MB43	6	0.52	HSS62	10	0.226
MB89	2	0.467	HSS24	6	0.922	MB32	10	0.263
MB89_rep	2	0.499	MB93	6	-0.387	MB44	10	-2.843
HSS33	2	0.84	HSS83	6	-1.408	HSS44	10	0.76
MB53	2	0.971	HSS1	6	-7.668	MB36	10	-1.328
MB55	2	0.289	MB155	6	-2.735	HSS46	10	0.67
MB72	3	0.275	MB42	7	0.547	MB78	11	-2.515
MB88_rep	3	-0.926	MB97	7	-1.34	HSS40	11	1.028
HSS13	3	0.263	HSS26	7	0.892	MB99_rep3	11	0.259
MB88	3	-0.35	HSS36	7	0.515	MB54	11	0.441
HSS67	3	-0.286	MB91	7	-0.222	HSS4	11	-0.652
MB18	3	0.674	HSS47	7	-10.096	MB45	11	-0.835
MB16	3	0.175	HSS15	7	-0.035	MB99_rep2	11	0.242
MB84	3	-5.793	HSS5	8	-0.347	MB98	11	0.316
MB86	4	0.085	HSS14	8	-13.185			
HSS23	4	0.25	MB96	8	-2.929			
MB95	4	0.864	HSS59	8	-0.721			
HSS63	4	-0.277	HSS30	8	0.775			
HSS28	4	-0.322	HSS21	8	0.318			
MB77	4	0.165	MB153	8	-0.149			
MB156	4	0.927	MB85	8	-4.732			
HSS20	4	-0.151						

7.2.4. Differentially Methylated Probes

Differential methylation was carried out using the R package limma. Each phenotypic variable was inputted into the design matrix along with age as a covariate. Following fitting of the model to the beta values, p values were adjusted for multiple testing using the Benjamini-Hochberg method, generating a false discovery rate (FDR) for each probe. For analyses that had a genomic inflation factor (λ) > 1.2, p values were adjusted with the BACON algorithm [464] to reduce inflation and bias, and subsequently adjusted for multiple testing using the Benjamini-Hochberg method. Differentially methylated probes (FDR < 0.2) are shown in appendices **O-Q**.

7.2.4.1. Continuous Analysis

Differential methylation was carried out with respect to the different phenotypic variables, including appendicular lean mass index (ALMI), grip strength, gait speed, fasting plasma glucose status, fasting serum insulin levels, age, and categorical analysis comparing control, pre-sarcopenia and sarcopenia. 834 CpGs were found to be differentially methylated with respect to ALMI. Of these CpGs, 210 showed a negative association with ALMI. The CpG showing greatest differential methylation with respect to ALMI was cg07259683 (FDR=0.0237) (figure **7.3a**), a CpG in the body of the Meis Homeobox 2 (MEIS2) gene, which showed a positive association with ALMI. cg13877502 showed the strongest negative association with ALMI (FDR=0.0237), located in the 5'UTR region of the KN Motif And Ankyrin Repeat Domains 1 (KANK1) gene. cg08444004 located in the body of the HDAC4 gene and cg13066693 located in the 5'UTR of the ESRRB gene showed a negative association with ALMI (FDR<0.2) while cg26071755 located upstream from the TSS of the HTR2C gene showed a positive association with ALMI (FDR<0.2).

194 CpGs were differentially methylated with respect to age, of which 44 showed a negative association with age. cg09476997 was the CpG showing the strongest positive association with age (FDR=0.0007) (figure **7.3b**), located in the body of the Solute Carrier Family 9, Subfamily A (NHE3, cation Proton Antiporter 3), Member 3 Regulator 2 (SLC9A3R2) gene. cg15552843 showed the strongest negative association with age (FDR=0.0021), located within 200bp of the TSS of the C22orf45 gene. Thirty-nine CpGs were differentially methylated with respect to fasting serum insulin levels, with 26 CpGs showing a negative association with fasting serum insulin levels. cg12046168 showed the strongest positive association with fasting serum insulin levels (FDR=0.0206), located in the body of the Methionine Sulfoxide Reductase B2 (MSRB2) gene. cg09253462 showed the strongest negative association with fasting serum insulin levels (FDR=0.03387), located in the 3'URE of the Zinc Finger Protein 737 (ZNF737) gene.

There were no dmCpGs (FDR < 0.2) with respect to fasting plasma glucose levels, gait speed and grip strength or between control and sarcopenic muscle.

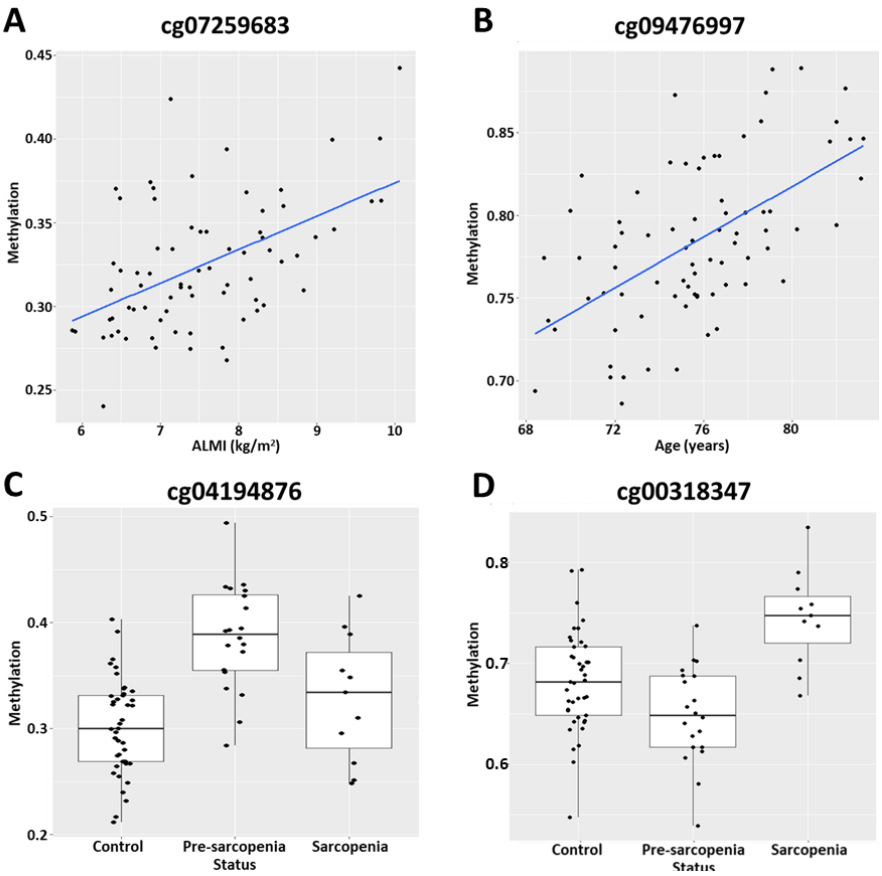


Figure 7.3: The most significantly differentially methylated

(A) cg07259683 showed the strongest positive association with ALMI (FDR=0.0237). (B) cg09476997 showed the strongest positive association with age (FDR=0.0007). (C) cg04194876 was the only differentially methylated CpG in pre-sarcopenia compared with controls (FDR=0.0307). (D) cg00318347 was the most significantly differentially methylated CpG in the sarcopenic muscle compared to pre-sarcopenic muscle (FDR=0.0706).

7.2.4.2. Categorical Analysis

cg04194876 was the only CpG that was significantly differentially methylated (FDR=0.0307, figure **7.3c**) between controls and pre-sarcopenics, showing increased methylation in pre-sarcopenic muscle. This CpG was not differentially methylated in sarcopenic tissue (FDR>0.2) or between pre-sarcopenic and sarcopenic muscle (FDR>0.2); it is not associated with a gene, but is located in an inter-genomic region on chromosome 14. cg19021310 is located in the body of the MEGF11 gene on chromosome 15 and is the gene-associated CpG with the smallest FDR value (FDR=0.2236) that was differentially methylated between pre-sarcopenics and controls.

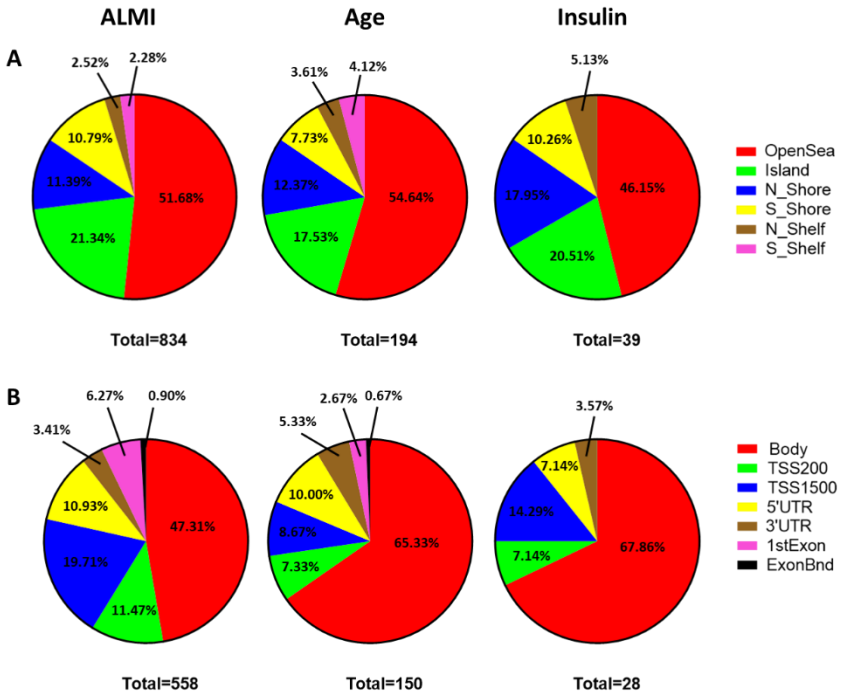


Figure 7.4: Distribution of significant CpGs within different genomic regions (A) The percentage of differentially methylated CpGs located in each of the different CpG locations relative to the CpG Island. Total number is the total number of significant CpGs identified. (B) Percentage of CpGs in each of the different gene region feature categories. Total number is the total number of significant CpGs associated with a gene. All the data was obtained from the Illumina hg19 EPIC annotation file.

Four CpGs were differentially methylated between pre-sarcopenics and sarcopenics. Of the four CpGs, three of them showed increased methylation in the sarcopenic muscle compared to the pre-sarcopenics. cg00318347 was the most significantly differentially methylated CpG (FDR=0.0706, figure **7.3d**), located in the body of the EBF2 gene on chromosome 8.

7.2.4.3. Genomic Locations of dmCpGs

The effect of differential methylation of CpGs differs depending on the location of the CpG with respect to genes and regulatory factors in the DNA. The distribution of the differentially methylated CpGs in the genome is shown in figure **7.4a**. Most of the CpGs differentially methylated with respect to ALMI, age and fasting serum insulin levels were located in open seas, which are isolated CpGs in the genome. Roughly 20% of the dmCpGs were located in CpG islands, CpG-rich regions typically in the promoter regions of genes. Approximately 20% of CpG differentially methylated with respect to ALMI and age and 30% of CpGs with respect to fasting serum insulin were located in CpG shores, within 2kb of a CpG island. CpG shelves, located between 2 and 4kb from a CpG island, contained the fewest dmCpGs.

Examining the functional genome distribution of these dmCpGs (figure **7.4b**), the majority of the CpGs (ALMI, 47%; age, 65%; fasting serum insulin 67%) were located in the gene body. The gene body encompasses everything between the ATG start codon and the stop codon, including introns and exons. The 5' UTR and regions within 200bp of the transcription start site (TSS) contained similar numbers of dmCpGs (approximately 10% with respect to ALMI and age, and 7% with respect to fasting serum insulin levels). Between 200bp and 1500bp from the TSS (TSS1500), there were roughly 19%, 8% and 14% of the dmCpGs with respect to ALMI, age and fasting serum insulin levels, respectively. The 3'UTR had between 3 and 5% of the CpGs, while the 1st exon contained 6% and 2% of the CpGs with respect to ALMI and age, respectively. With ALMI and age, there were a few CpGs located at exon boundaries (0.9% and 0.6%, respectively).

Table 7.3: *List of enriched TF binding sites*

Factor	Total Genes with Factor	Significant Genes with Factor	FDR
Whole Gene Body			
EZH2	719	37	1.64×10^{-5}
CTCF	5817	172	1.41×10^{-3}
5'TSS (500bp upstream, 1000bp downstream)			
EZH2	869	39	2.13×10^{-4}

7.2.4.4. Transcription Factor Enrichment at dmCpG-Associated Genes

To determine whether there was an enrichment of the differentially methylated sites at transcriptional regulatory binding sites, the ENCODE ChIP-Seq Significance Tool was used to look for enriched transcription factors binding sites associated with the genes that had a dmCpG associated with ALMI (table **7.3**). EZH2 was the only transcription factor found to be enriched at the at the 5' region from the TSS (500bp upstream, 100bp downstream) of the genes with significant dmCpGs (FDR=2.13×10^{-4}). As the majority of dmCpGs were found in the gene bodies, we examined for TFs enriched in the whole body of genes. EZH2 (FDR=1.64×10^{-5}) and CTCF (FDR=1.41×10^{-3}) were found to be enriched at the genes in our dataset.

7.2.5. Differentially Methylated Regions

Neighbouring CpGs tend to work together to regulate gene expression. Therefore, we examined for differentially methylated regions with respect to ALMI, age and fasting serum insulin levels, as well

as comparing between control and pre-sarcopenics and pre-sarcopenics and sarcopenics. These comparisons had individual CpGs showing significant differential methylation (FDR < 0.2).

Table 7.4: *List of DMRs identified*

	Chr	Start	End	Width	No. of CpGs	Min FDR	Stouffer	Overlapping Promoters
ALMI	chr6	33244752	33245638	887	23	3.17E-31	0.0015	B3GALT4/ RPS18
	chr6	32063501	32064258	758	21	8.30E-30	0.0063	-
	chr22	51016604	51017019	416	10	1.37E-24	0.0097	CPT1B
	chr9	706836	707410	575	4	2.90E-26	0.0395	KANK1
Age	chr16	2088960	2088997	38	3	5.34E-20	0.0013	NTHL1
	chr16	2087377	2087932	556	3	5.70E-19	0.0090	SLC9A3R2
	chr2	219736167	219736549	383	3	2.19E-11	0.0495	-
	chr8	101348359	101349044	686	10	6.49E-22	0.0555	RNF19A
	chr6	31148332	31148748	417	15	6.97E-23	0.1968	POU5F1
	chr11	315102	315751	650	4	1.44E-12	0.2085	IFITM1
	chr10	77871618	77871958	341	5	2.80E-11	0.2492	-
	chr10	71176501	71176853	353	6	7.36E-13	0.2812	TACR2
	chr20	44746392	44747006	615	8	3.14E-13	0.2931	CD40
	chr1	161067886	161068252	367	7	1.11E-11	0.3131	KLHDC9
	chr13	112985463	112986154	692	6	6.67E-12	0.4086	LINC01044
	chr14	24640895	24641706	812	13	2.76E-17	0.5209	REC8
	chr7	27208454	27209828	1375	18	3.71E-23	0.7369	MIR196B
	chr6	32117725	32118811	1087	15	3.07E-13	0.8956	PRRT1
	chr22	24890045	24891452	1408	18	1.17E-31	0.9679	UPB1
	chr6	32144978	32146779	1802	33	3.86E-21	0.9967	AGPAT1
Insulin	chr10	100993553	100993965	413	9	4.37E-40	0.9981	HPSE2
Con vs Pre	chr2	198649933	198650215	283	6	1.48E-13	0.9239	BOLL
Pre vs sarc	chr5	102898223	102898545	323	5	5.30E-27	0.7090	NUDT12

Using DMRcate, DMRs were identified with respect to ALMI, age, fasting serum insulin, and in the categorical analysis between controls and pre-sarcopenics and pre-sarcopenics and sarcopenics (table **7.4**). DMRcate calculates region-wise p-values by combining p-values for the individual CpGs in the region using Stouffer's method [465]. Stouffer's method is similar to Fisher's method for combining p-values for meta-analysis, but uses test statistics rather than p-values, generating a region-wise p-value. There were 4 DMRs with respect to ALMI, with 3 of the DMRs associated with an annotated gene, overlapping the promoters of the Beta-1,3-Galactosyltransferase 4/Ribosomal protein S18 (B3GALT4/RPS18)(figure **7.5**), carnitine palmitoyltransferase 1B (CPT1B) and KN Motif And Ankyrin Repeat Domains 1 (KANK1) genes. All 4 DMRs with respect to ALMI had a Stouffer value < 0.2, suggesting significance of all the p-values of the CpGs in the DMR. There were 16 DMRs with respect to age, of which 14 overlapped promoters of genes. The top 2 DMRs with respect to age overlapped the Nth Like DNA Glycosylase 1 (NTHL1) and SLC9A3 Regulator 2 (SLC9A3R2) genes

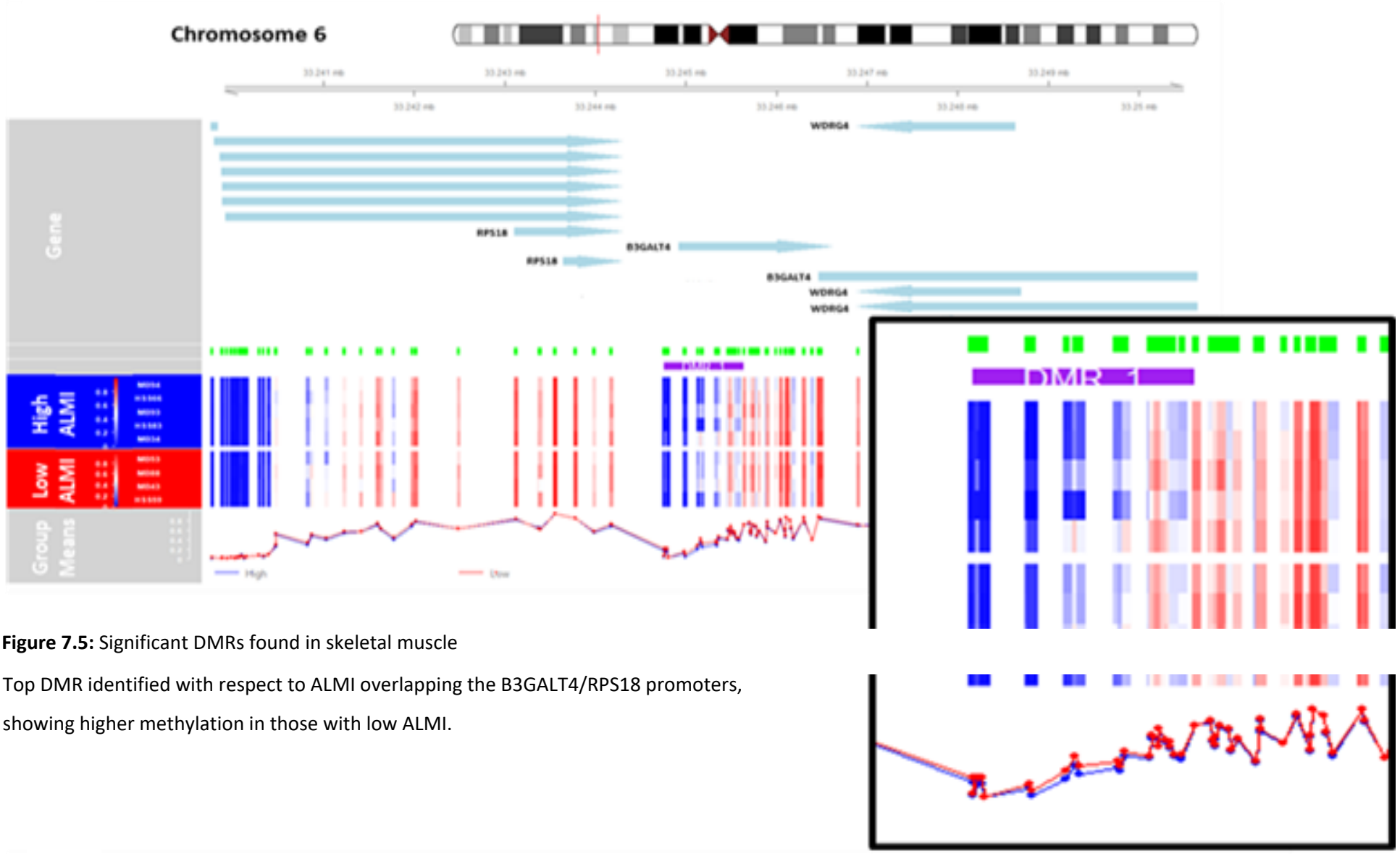


Figure 7.5: Significant DMRs found in skeletal muscle

Top DMR identified with respect to ALMI overlapping the B3GALT4/RPS18 promoters, showing higher methylation in those with low ALMI.

Five DMRs with respect to age had Stouffer values < 0.2, while the remaining were all > 0.2. There was one DMR with respect to fasting serum insulin, overlapping the Inactive Heparanase-2 (HPSE2) gene. There was also only one DMR with the comparison between controls and pre-sarcopenics, overlapping the promoter of the Boule Homolog, RNA Binding Protein (BOLL) gene. There was one DMR with the comparison between pre-sarcopenics and sarcopenics, overlapping the promoter of the Nudix Hydrolase 12 (NUDT12) gene on chromosome 2. The DMRs with respect to fasting serum insulin and between controls and pre-sarcopenics, and pre-sarcopenics and sarcopenics all had Stouffer values > 0.2.

7.2.6. Pathway Analysis

Gene ontology (GO) analysis was carried out on the significant CpGs with respect to ALMI. GO analysis was not carried out on CpGs significantly differentially methylated between controls and pre-sarcopenics, and pre-sarcopenics and sarcopenics due to the low number of dmCpGs. GO terms with an FDR < 0.1 were classed as significantly altered with respect to ALMI or age.

Multiple GO terms associated with skeletal muscle development were associated with ALMI, including regulation of myotube differentiation (GO:0010830, FDR=0.005106), skeletal muscle tissue development (GO:007519, FDR=0.005921) and regulation of skeletal muscle fibre development (GO:0048742, FDR=0.018668). There was also a significant association between ALMI and GO terms associated with apoptosis and cell death, including the apoptotic process (GO:0006915, FDR=0.0191), regulation of apoptotic process (GO:0042981, FDR=0.0366) and negative regulation of programmed cell death (GO:0043069, FDR=0.0637). GO terms associated with muscle structure and cellular organization were also altered with respect to ALMI, including muscle structure development (GO:006106, FDR=00033), cellular component assembly (GO:0022607) and extracellular matrix/structure organization (GO:0030198 and GO:0043062, FDR=0.1980 and 0.0199, respectively).

Cytoscape is a bioinformatic tool used to visualize molecular interactions and determine molecular networks. The MCODE plug-in in Cytoscape allowed us to achieve a better understanding of the function of the proteins at the molecular level, by module analysis of the PPI network. Module analysis of the genes containing a significant dmCpG with respect to ALMI showed multiple different gene clusters with genes showing interactions and co-expression patterns (figures **7.6a** and **7.7a**). 14 significant clusters of genes were identified (table **7.5**). These clusters comprise genes that show interactions at the protein level, as well as co-expression in the same cells and form part of similar pathways, therefore identifying groups of genes with a significant dmCpG that may work together to alter muscle function. The genes in clusters 3 and 13 were investigated further as these clusters contained genes that have previously been reported to play a role in skeletal muscle

development and function. The genes in cluster 3 included EHD2 (EH Domain containing 2), involved in sarcolemmal repair [466]; HTR2C (5-Hydroxytryptamine Receptor 2C), involved in the mobilization of Ca^{2+} ions [467,468]; NID2 (nidogen 2), playing a role at the synapse of the neuromuscular junction [469]i; MYF6 (myogenic factor 6), which plays a role in skeletal muscle differentiation [348] and PAX3 (Paired-box protein 3), which plays a role in maintaining the stem cell-like properties of satellite cells [32,34]. The genes in cluster 13 included HDAC4 (Histone deacetylase 4), which plays a role in denervation-induced muscle atrophy [83]; ESRRB (oestrogen-related receptor b), which may play a role in energy production and metabolism [470,471]; KANK1 (KN motif and ankyrin repeat domains 1), which has a role in actin organisation and cell growth, and localises to muscle-tendon attachment sites in *Drosophila* [472]; PRR5L (Proline rich 5 like), which is involved in mTORC2 signalling [473]; TGFBI (Transforming growth factor beta induced), which has been shown to promote myofibril bundling and fibre growth [474] and TNS1 (Tensin 1), which has been shown to play a role in skeletal muscle regeneration [475].

Of the genes with CpGs showing increased methylation, some of the GO terms showing significant enrichment within gene clusters included extracellular matrix assembly, elastic fibre assembly, key for muscle structure, as well as axon ensheathment, synaptic transmission and ensheathment of neurons. This may suggest changes at the neuromuscular junction and denervation of skeletal muscle, one of the mechanisms contributing to muscle loss and sarcopenia. There was also an enrichment of GO terms associated with muscle function, including regulation of sequestering of calcium ions and regulation of the myoblast differentiation (figure **7.6b**). Of the genes showing decreased methylation, some of the GO terms enriched included positive regulation of striated muscle contraction, striated muscle hypertrophy, regulation of muscle adaptation and regulation of skeletal muscle tissue development (figure **7.7b**), implicating changes in muscle structure and development.

7.2.7. Methylation Clock

The DNA methylation age of each sample was calculated following the method described by Horvath et al.[353]. The DNA methylation age of the muscle tissue did not correlate well with chronological age (pearson r = 0.134, p=0.246, figure **7.8a**), and all the methylation ages were younger than the chronological age. The age acceleration of the muscle was calculated for each sample. Age acceleration is the difference between DNA methylation age and chronological age, with a positive/negative value of epigenetic age acceleration suggesting faster/slower ageing of tissues than expected. Age acceleration showed a significant positive association with ALMI (p=0.000052, figure **7.8b**), with those with high ALMI showing faster ageing of skeletal muscle (smaller absolute age acceleration) than those with low ALMI. Age acceleration showed a non-

significant positive trend with grip strength (p=0.190), but showed no association with gait speed.

To determine whether age acceleration showed any relationship with measures of metabolic

Table 7.5: *Genes in each individual cluster identified by Cytoscape*

Cluster	Genes in cluster	Cluster	Genes in cluster
Cluster 1	SCN3B, CHRM3, TBC1D4, SLC6A4, ASPA, HIVEP3, DOK6, GRIK4, TNXB, GATAD1, SHANK2, UBE4A, NALCN, PRKCZ, GRIA1, PRDM8, CLDN11, ZFPM1, AP3B2, CNTN2, NELL1, OPCML, KCNJ9, MOG, SOX1, RAB37, GRIA2, KCNB2, PHACTR1, MLLT1, PCDH7, PLP1, DZIP1, B3GALNT1	**Cluster 8**	FAH, SNAP23, DDX17, SLC51A, MFSD6L, COL4A2, RCAN2, OGG1, ITGA2B, CSF3R, SRC, HOXC6, RAP1GAP, MAU2, CELF6, SVOPL, OLFML2B, NR2F2, RUNX3, HSPB8, TESC, CTF1, AIRE, DEPDC5
Cluster 2	GAS2L1, CIB3, CACNG6, FBLN2, KLHL41, RASEF, AGAP2, LIMS2, TSPOAP1, DLGAP1, SRCIN1, BDNF, ADAMTS2, CA3, IRGC, CACNG3, CACNA1E, HOXC4, BCAS3, GPR6, ELMO3, NAV2, MYOD1, PDE4B, CDH19, USP18, PDP1, HAGHL, ISL1, STAG3, LTBP4, SORCS2, RGS22, ARHGAP21, XPNPEP1, FOXE3, MATN2, SMOC2, NCOA7, MYADM, HDC, ICAM5, CCDC28B, STEAP2, LMBRD1, KLRG2, DBP, GRIN2D, MAMDC2, ATP12A, DCHS2, IL1RAPL2, MAPK4, ABHD12B	**Cluster 9**	RHBG, PLA2G3, PCYT2
Cluster 3	KIAA0040, RPH3AL, PAX3, LINGO1, GPR45, IRX6, IRF2BPL, ECE2, KCNN4, ATG14, GALNT9, KLK6, SPOCK3, AGXT2, NID2, PANK4, BTBD11, ALK, RAP1GAP2, PSD, SLC24A3, ABR, CELF1, HTR2C, LIG4, ATP6V1C2, C1QTNF2, EHD2, FBXL5, MEIS2, PTPN5, CUX2, MX1, BEND4, ZFHX3, CSGALNACT1, HSPB7, BCL7B, EPHB2, MYF6, GPM6B, GPC2, VSX1, HSPA9, MXRA7, FAM60A, SHISA2, ADRA1B, COL4A3, DMBX1, CCNH, UMOD, EMX1	**Cluster 10**	ATG16L1, DLC1, PTPRT, KIRREL3, KIF21A, ABCC8, HECW1, AJAP1, ARNT2, PLEKHG3, UVRAG, TBL1XR1, KLF5, CAMTA1,
Cluster 4	CNTNAP2, DNAJC6, SLC12A5, SLC17A6, KCNS1, FMN2, CAMK2A, RNF220, GMDS, IQSEC1, AJAP1, SNAP25	**Cluster 11**	TFDP1, AMBRA1, ADRA1A, FMNL2, PRKCH, FBXW7, RBM34, IPO5, CHODL,
Cluster 5	DDR1, LRP6, NUDT5, AKAP14, PRLR, ISG15, CRB3, SBNO2, MDK, AMER3, CCDC105, PLPP7, LYN, KCNK10, HTR4, EPCAM, MAPKBP1, VPS16, EBF2, SEPT8, FOXO3, RNF144B, NFKBIA, NFIX, PIKFYVE, GRK5, HOXC5, AP2A1, CLDN9, HS3ST3A1, GPX3, CFAP77, FADS6, PRKAR1B, CSRNP2, ALDH1A3, SCGN, DPYS, NTHL1, FFAR4, ESM1, CDC25B, SLC13A3, NSMCE1, SLC16A4, TRIM29, DAGLB	**Cluster 12**	RUNX2, RNASET2, AP1B1, SIL1, YWHAH, ALK, YWHAG, AMOTL2, FSCB, CLCF1, EML1, SLC25A37, CCNL1, WDR45B, BUD13, ELANE, SPG11, SRGAP3, DST, RGS17, ANKH, CRLF1, EPHB1, FAM91A1, PPM1D, WWOX, FARP1,
Cluster 6	HNF4A, CHODL, TRIM34, CGA, LPGAT1, PTPRN2, TTLL7, SLIT3, CELF4, TTLL11, PTPRS, KDM8, CDC123, SLC9A9, SYNDIG1, RABGAP1L, TPM1, BARHL1, GAS7	**Cluster 13**	LIPC, TMEM266, RPS3, TGFBI, NTNG1, PGM1, IL18RAP, PRR5L, CEACAM4, KANK1, ESRRB, SCD5, CLDN14, HGD, HDAC4, NRG2, TNS1, PTN,
Cluster 7	GRIA3, PTPRT, EXTL1, CDH22, SGIP1, CA10, CAMK2N1	**Cluster 14**	BCL2L11, KCTD5, TRIP13, MID1, PLEKHA6,

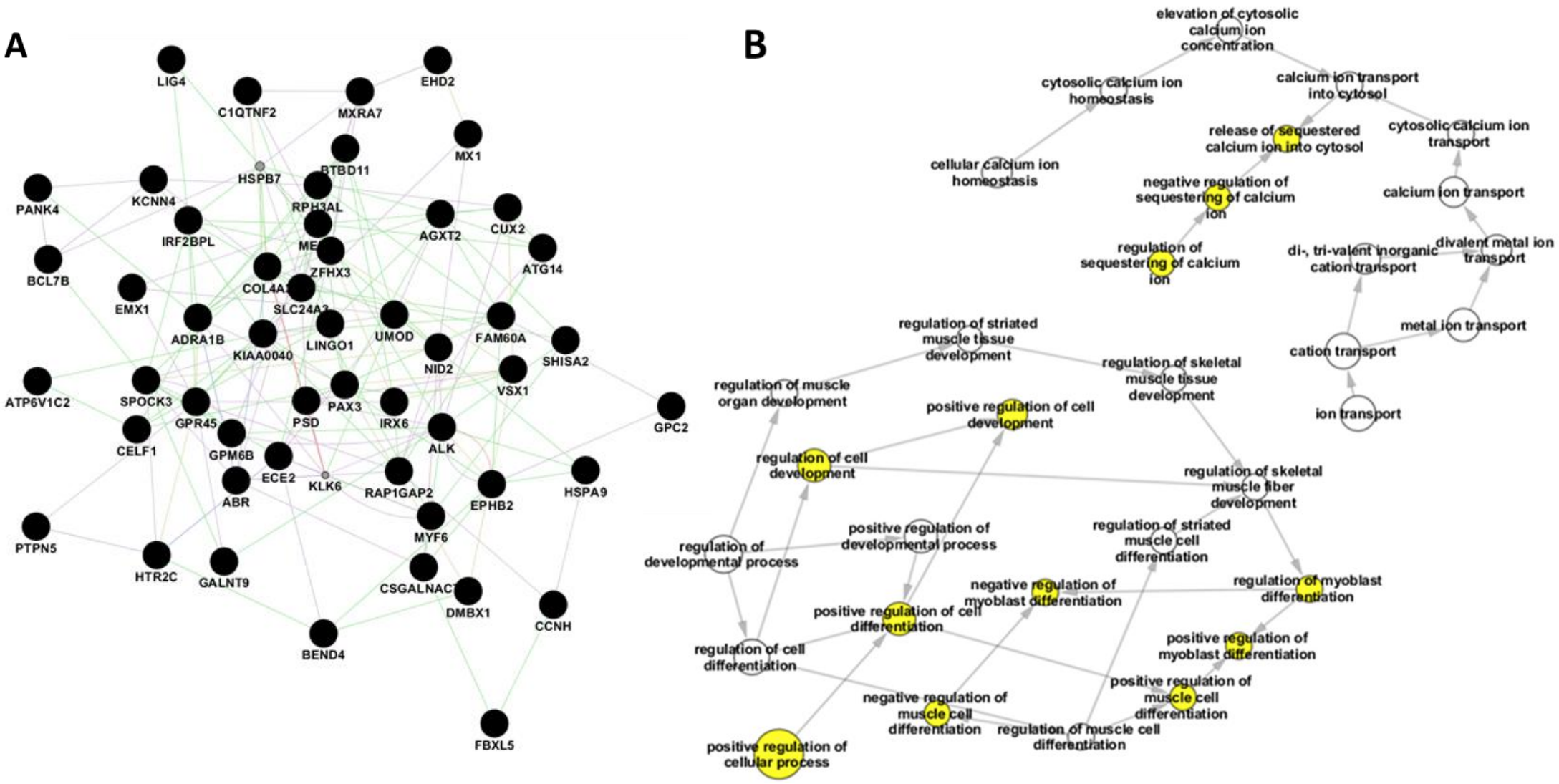


Figure 7.6: Gene clusters showing significant gene ontology enrichment with genes showing increased methylation

(A) Gene cluster 3 showing physical interactions (thick red lines), genetic interactions (green lines), co-expression (purple lines) and co-localizations (blue lines) of genes/proteins. (B) Gene ontology terms significantly enriched within each cluster (yellow nodes) showing enrichment for structural components of muscle. Black nodes are genes with a dmCpG (FDR < 0.2).

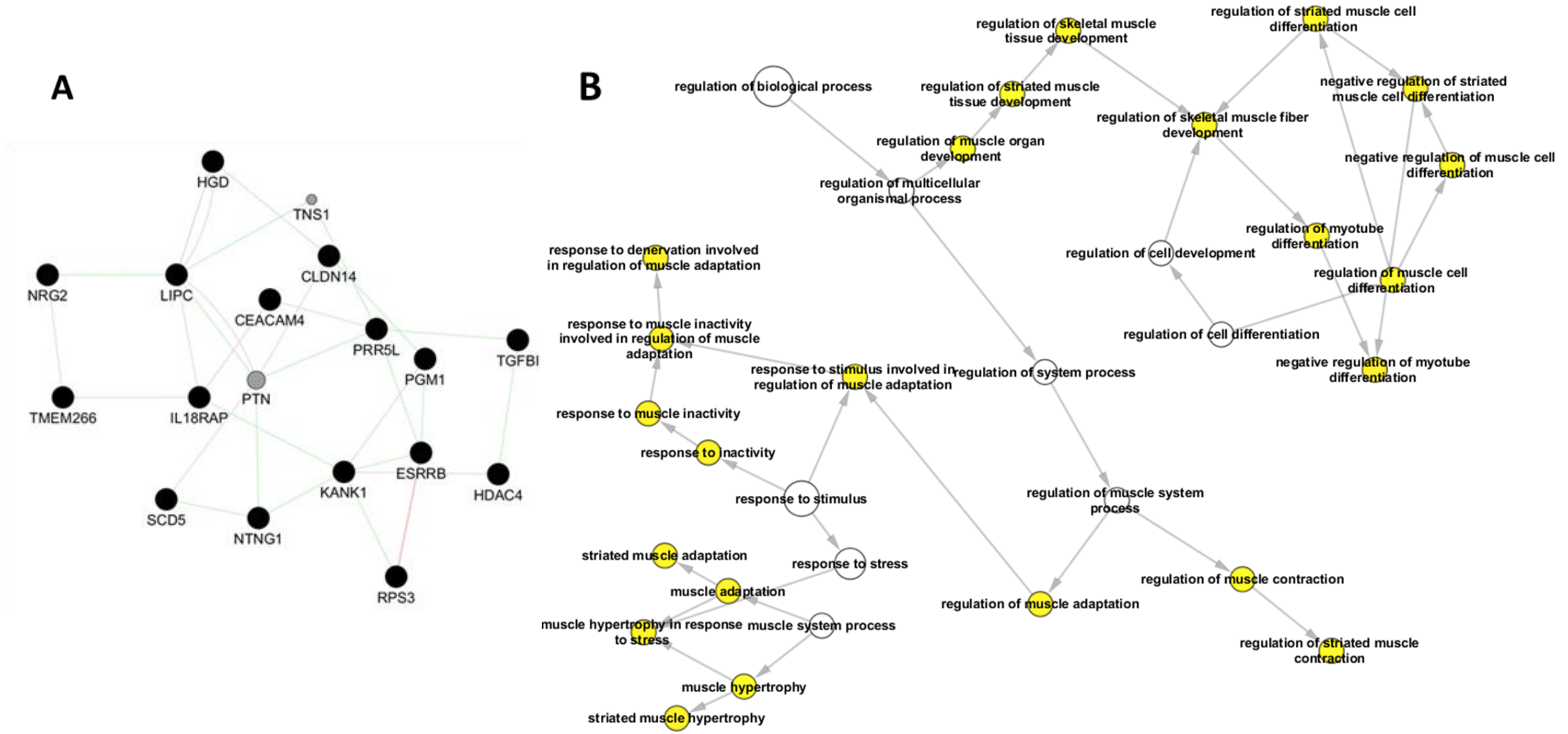


Figure 7.7: Gene clusters showing significant gene ontology enrichment with genes showing decreased methylation

(A) Gene cluster 13 showing physical interactions (thick red lines), genetic interactions (green lines), co-expression (purple lines) and co-localizations (blue lines) of genes/proteins. (B) Gene ontology terms significantly enriched within each cluster (yellow nodes) showing enrichment muscle structure and development. Black nodes are genes with a dmCpG (FDR < 0.2).

health, we examined age acceleration with respect to fasting plasma glucose levels and fasting serum insulin levels. Age acceleration showed no association with fasting plasma glucose levels or fasting serum insulin levels (p=0.393 and p=0.725 respectively). Age acceleration, however, showed a significant decrease in pre-sarcopenic muscle compared to control muscle tissue (p=0.025) but no significant difference between sarcopenic and control muscle tissue (figure **7.8c**).

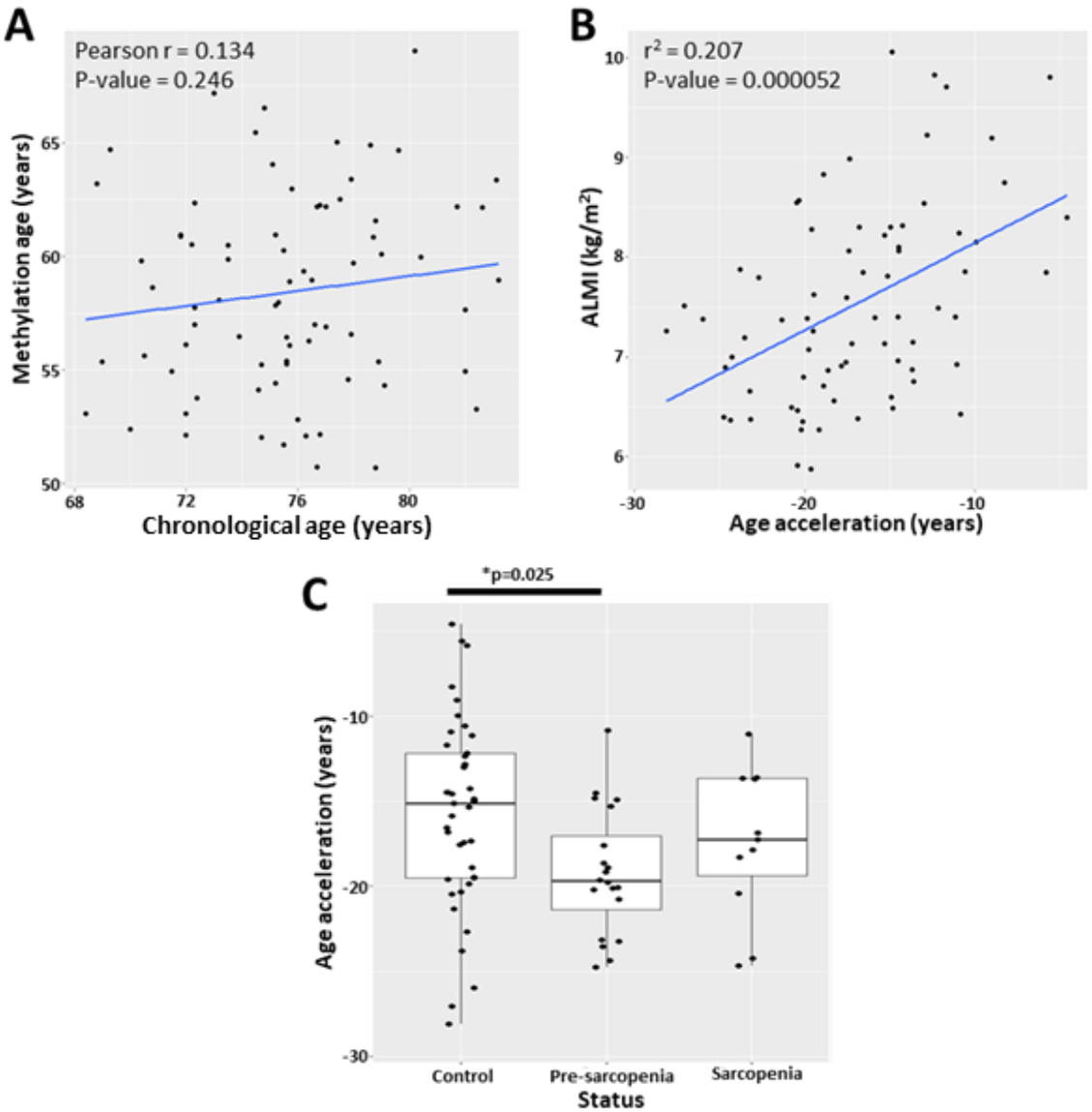


Figure 7.8: DNA methylation age comparison of skeletal muscle tissue

(A) Correlation between DNA methylation age and chronological age. There is a positive trend between the two, although not significant. (B) A significant positive association is seen between ALMI and age acceleration of the skeletal muscle tissue. (C) Pre-sarcopenic skeletal muscle tissue shows slower age acceleration than control muscle tissue, with no difference between control and sarcopenic or pre-sarcopenic and sarcopenic muscle tissue.

7.3. Discussion

In this study, we have shown for the first time methylation changes in skeletal muscle tissue associated with changes in skeletal muscle mass and with respect to sarcopenia status. There was a greater number of dmCpGs found to be associated with skeletal muscle mass than for any other aspect of skeletal muscle phenotype. The majority of the dmCpGs associated with skeletal muscle mass were located in the gene body, with a tendency for hypermethylation of CpGs with higher muscle mass. Several DMRs found to be associated with muscle mass were also located in the promoters of genes that play a role in normal muscle function. The genes associated with the dmCpGs and DMRs showed enrichment in multiple pathways associated with normal muscle function, including myoblast differentiation, muscle signalling and muscle development and adaptation. Changes in these pathways may alter muscle function, resulting in a loss of muscle mass if these pathways are perturbed.

7.3.1. DNA Methylation and Skeletal Muscle Mass and Function

Sarcopenia is a disease with many factors contributing to its initiation and progression. The main mechanisms regulating skeletal muscle mass in later life are muscle hypertrophy (growth) and muscle atrophy (loss). A balance between these two processes is key in maintaining an optimal level of muscle mass in an individual. Several different molecular mechanisms and pathways are critical to these two processes, of which different factors play a key role in regulating, including DNA methylation [356,476]. Muscle hypertrophy and atrophy require *de novo* gene expression, and therefore DNA methylation and transcriptional regulation play a key role in these processes. Kanzleiter et al. [477] studied the effect of exercise in mice on DNA methylation, and found that exercise resulted in altered methylation in many genes associated with muscle growth and differentiation-related genes in murine skeletal muscle. Exercise altered the methylation of CpGs in the promoter region of genes such as IGFBP4 (IGF-binding protein 4) and PLXNA2 (Plexin A2), which act upstream of MyoD and myogenin [478,479], regulating muscle growth and differentiation, implicating DNA methylation in the regulation of muscle hypertrophy. Jin et al. [356] further showed an age-related change in DNA methylation in genes associated with protein catabolism and proteolysis in porcine muscle, key aspects of muscle atrophy.

In this study, we have shown that in aged human skeletal muscle, there was altered methylation associated with skeletal muscle mass in genes that are part of muscle atrophy and muscle growth pathways. HDAC4 showed altered methylation in its gene body. HDAC4 plays a major role in regulating muscle atrophy via the HDAC4-myogenin pathway [83,412]. In the report by Moresi et al. [83], forced expression of HDAC4 resulted in the activation of this pathway, resulting in increased

denervation-induced muscle atrophy, the expression of which may be regulated by the altered methylation in the HDAC4 gene. As well as altered muscle atrophy and hypertrophy pathways, we found changes in methylation in genes associated with the regulation of striated muscle differentiation and myotube differentiation Myoblasts are recruited to an area of muscle development and undergo myogenic differentiation into myotubes and mature muscle fibres. Changes in DNA methylation of genes associated with myogenic differentiation may alter the expression of these genes, and as such affect the downstream process of differentiation and muscle growth. We also found methylation changes in pathways associated with muscle function, including genes involved in the regulation of striated muscle contraction and the regulation of the sequestration of calcium ions (e.g. increased methylation within 1500bp of the HTR2C TSS associated with increased ALMI). Activation of the 5-hydroxytryptamine receptor 2C (HTR2C) has previously been reported to play a role in the mobilization of calcium ions (Ca^{2+}) [467,468]. As Ca^{2+} ions play a vital role in controlling the regulation of muscle contraction in response to stimulus [480], changes in Ca^{2+} signalling pathways may affect muscle function by reducing the ability of skeletal muscle to efficiently respond to stimuli and function normally.

Changes in pathways associated with muscle hypertrophy and muscle growth as well as muscle atrophy may lead to an imbalance between gain and loss in muscle mass and as such, may contribute to the loss of muscle mass seen in the elderly. This loss of muscle mass, together with changes in pathways associated with muscle function, may contribute to a reduction in function, activity and independence of older people. Whether these changes in DNA methylation directly affect muscle mass by altering gene expression, or whether we see changes in DNA methylation due to reduced muscle mass is unknown, and the causal relationship requires further investigation. It is also still unclear whether these changes in methylation affect gene expression, requiring further functional analysis *in vitro* to determine the role of these DNA methylation changes.

One of the contributing factors underlying sarcopenia is muscle denervation [244]. Skeletal muscle denervation involves the loss of nerves at the neuromuscular junction (NMJ), which results in decreased muscle efficiency and gradual atrophy of the denervated muscle. Denervated muscle is not as efficient at force production, as many more muscle fibres become a part of the same motor unit, resulting in a less efficient transfer of nervous stimuli and overall muscle function. It has also been reported previously that there are changes in the myelination of neurones with age, affecting normal neuronal function [245]. Here, we found methylation changes at CpGs associated with axon ensheathment, myelination and regulation of action potential in the neuron. Although we did not see any direct changes in neuronal signalling, previous evidence has suggested that these changes in neuronal signalling and the NMJ are associated in the loss of muscle mass [134]. Myelination of neurons and their axons enhances the propagation of action potentials down the axon and towards

the synapse at the NMJ. We also found alterations in genes associated with cell-cell signalling and synaptic transmission pathways. These changes in neuronal and axonal structure, together with synaptic signalling may suggest a decreased efficiency of signal transmission to the NMJ and to the muscle, resulting in muscle denervation. This implicates DNA methylation in the regulation of muscle denervation, leading to changes in muscle mass and function.

ESRRB had a differentially methylated CpG, showing an inverse association with ALMI. ESRRB (Oestrogen-related receptor b) is part of a family of orphan nuclear receptors that plays a role in regulating energy metabolism [470]. As skeletal muscle function relies on efficient energy production and metabolism, changes in the regulation of transcription of genes that play a role in energy metabolism may impair normal muscle function. ESRRB has also been shown to play a key role in maintaining the self-renewal potential of embryonic stem cells, and maintain pluripotency [470,471]. Therefore, ESRRB may play a role in regulating the self-renewal potential of myoblasts and muscle progenitor cells (MPCs). A reduction in the self-renewal capacity of MPCs has long been thought to contribute to the pathogenesis of muscle loss. Therefore, changes in ESRRB expression due to altered regulation by DNA methylation may contribute to the altered self-renewal capacity of MPCs. Here we saw a decrease in methylation with increased ALMI, suggesting increased ESRRB expression with increased ALMI. Therefore, this suggests that increased ESRRB may contribute to the maintenance of the self-renewal capacity of MPCs, and the ability of skeletal muscle to regenerate, although further investigation into the role of altered methylation on ESRRB expression and the role of ESRRB in skeletal muscle is required. However, only a small percentage of muscle is comprised of MPCs, therefore the majority of the methylation signal observed originates from differentiated muscle. Therefore, ESRRB may be playing another, currently unknown, role in skeletal muscle, requiring further investigation.

7.3.2. DNA Methylation and Other Epigenetic Changes

The genes associated with the ALMI dmCpGs were enriched for EZH2 binding sites. EZH2 is a polycomb group protein, forming a part of the polycomb-repressive complex 2 (PRC2). The EZH2 subunit of PRC2 tri-methylates histone H3 on lysine 27 (H3K27me3), establishing a repressed chromatin structure. DNA methylation and histone methylation on lysine residues appear to be highly interrelated in regulating chromatin structure and gene expression. Both DNA methylation and H3K27me3 are repressive epigenetic marks, silencing the expression of genes in the region. It has been suggested that H3K27me3 is a transient mark, primarily occurring at CpG islands in embryonic stem cells, predisposing the region to obtain more static DNA methylation marks later during development and cell differentiation [481]. EZH2 directly recruits DNA methyltransferases (DNMTs) to the DNA, and is required for DNA methylation of EZH2-target promoters [482]. Therefore,

EZH2 primes the DNA for more stable DNA methylation marks via DNMTs, resulting in further repression of EZH2-target genes. EZH2 is highly expressed in mouse satellite cells, in cells co-expressing Pax7 [483]. EZH2 was found to be necessary for maintaining the pool of satellite cells in postnatal muscle tissue, with the ablation of EZH2 expression resulting in a 40% reduction in $Pax7^+$ cells. It was suggested that EZH2 is required to maintain the proliferative capacity of satellite cells, with EZH2 activity resulting in reduced expression of structural genes such as titin, MYH7 and sarcolipin, suggesting EZH2 prevents the progression and differentiation of satellite cells into myotubes and mature myofibres [483,484]. Therefore, EZH2 contributes to the regulation of the balance between stem cell state of satellite cells and differentiation.

The enrichment of EZH2 binding sites in the list of genes with ALMI dmCpGs suggests that DNA methylation and histone modifications may be acting in conjunction to alter muscle mass in the elderly, with EZH2 activity affecting the differentiation capacity of satellite cells, as well as reducing the satellite cell population in mature muscle. The enrichment of genes with EZH2 binding sites suggests that either the DNA methylation results in the recruitment of EZH2 to the region, altering histone methylation, integrating DNA and histone methylation-mediated regulation of gene expression, EZH2 is recruited to the DNA via another mechanisms, resulting in the recruitment of DNA methyltransferases to the region and induces the changes in DNA methylation that we see [482]. As such, EZH2 may contribute to the DNA methylation of these genes in the elderly, resulting in perturbed satellite cell function. A reduction in quiescent satellite cells in elderly muscle has always been thought to be a key factor in the aetiology of sarcopenia [149], which may be linked to changes in DNA methylation and histone modifications.

7.3.3. Skeletal Muscle DNA Methylation Age

DNA methylation is known to be altered with age, and several studies have shown that with age, there is a global hypomethylation, including in skeletal muscle. However, some studies have shown a predominant hypermethylation phenotype with age in skeletal muscle. S. Horvath [353] have previously shown that DNA methylation can be used to accurately predict chronological age across multiple species and different tissue types. It has been suggested that the DNA methylation age measures the cumulative effect of an epigenetic maintenance system, linking the innate molecular processes such as DNA methylation to the decline in tissue function.

Here, we looked at the DNA methylation (DNAm) age of the skeletal muscle tissue and found that there was a poor correlation between DNAm age and chronological age. Although disappointing, this is not surprising, as S. Horvath [353] similarly found a poor correlation between chronological age and DNAm age in skeletal muscle tissue. It was suggested that this may be due to the presence of satellite cells in skeletal muscle, which may act to rejuvenate skeletal muscle at different rates in

different people, depending on many external factors including activity, muscle injury, etc. Therefore, further work would be required to investigate the skeletal muscle DNAm age, using other established algorithms [354] or other methods.

Despite the poor correlation between DNAm age and chronological age, we found that there was a positive correlation between ALMI and age acceleration, suggesting that those with low ALMI have a slower rate of muscle ageing than expected. This is surprising as one would expect those with low ALMI to have a faster rate of muscle ageing than expected. However, it is possible that in those with low ALMI, the activation and differentiation of satellite cells may occur at a greater rate than those with high ALMI, and as such the DNAm age appears younger, leading to slower age acceleration of the skeletal muscle. It may also be possible that, as previously shown (chapter 6), those with low ALMI have myoblasts that are more likely to show senescence and show a slow proliferation rate. Therefore, these cells may be at a 'younger' state at the methylation levels, leading to reduced age acceleration and younger muscle in those with low ALMI. A similar result is seen in those with pre-sarcopenia compared to control individuals, where those with pre-sarcopenia show a slower rate of ageing of skeletal muscle compared to controls (figure **7.8c**). Further investigation is required to examine the effect of the slower age acceleration in those with low ALMI and in order to understand why the observed result is opposite to what would be expected.

7.3.4. Further Considerations

Although we found methylation changes associated with muscle mass in skeletal muscle tissue from older people, the relatively low number of samples requires the results to be further validated in a larger population to ensure biological replication of the results. Technical validation of the results via pyrosequencing is also required to validate the results of the microarray. Furthermore, we found that EZH2 binding sites were enriched in the list of genes with significant dmCpGs. However, further work using EZH2-specific inhibitors or activators is required to determine whether changes in EZH2 activity in elderly muscle may be playing a direct role in regulating the expression of genes important for maintaining muscle mass. Although we found changes in DNA methylation associated with variations in muscle mass and function, further investigation is required to determine whether these methylation changes are directly involved in altered muscle mass and function, or whether these changes in DNA methylation arise to due altered muscle homeostasis and from changes in environmental factors.

7.3.5. Conclusions

Changes in DNA methylation have been associated with many different age-related diseases and disorders, including cancer and diabetes. Here, we found that DNA methylation was also associated

with changes in skeletal muscle mass in older people, with DNA methylation changes enriched in pathways associated with muscle hypertrophy/atrophy and differentiation as well as muscle innervation and structure. Understanding the mechanisms regulating the expression of genes in key pathways, such as DNA methylation, provides a better understanding of the molecular pathways underpinning sarcopenia, as well as understanding whether DNA methylation is causally affecting muscle mass, informing the evaluation of effective interventions and therapeutics to combat the loss of skeletal muscle with age.

Chapter 8 –

Single Cell RNAseq Shows Cellular Heterogeneity of Skeletal Muscle Myoblasts from Older Individuals

Chapter 8 – Single Cell RNAseq Shows Cellular Heterogeneity of Skeletal Muscle Myoblasts from Older Individuals

8.1. Introduction

Skeletal muscle satellite cells are resident stem cells located between the basal lamina and the sarcolemma of muscle tissue, and provide support for skeletal muscle growth and repair, as well as muscle hypertrophy and maintenance in postnatal muscle tissue. The loss of satellite cells and functional impairment of these cells occurs during ageing, sarcopenia and muscular dystrophy. Satellite cells remain in a quiescent state until activated, when they proliferate as myoblasts, before differentiating and fusing into myotubes to subsequently form mature myofibres. During differentiation, a subpopulation of satellite cells also undergo self-renewal, repopulating the pool of resident satellite cells.

8.1.1. Satellite Cell Heterogeneity

Several studies have reported that satellite cells exist as a relatively heterogeneous cell population, with differing abilities to divide and proliferate, activate and differentiate. Satellite cells are Pax7+ cells. However, these Pax7+ cells express varying levels of other myogenic regulatory factors. Satellite cells can also express Myf5, with Pax7+/Myf5+ cells representing satellite cells that have undergone commitment to the myogenic lineage, whereas Pax7+/Myf5- cells remain as stem cells and contribute to the renewal of the satellite cell pool [485,486]. A subpopulation of satellite cells exist in a low metabolic state. During the division of this subpopulation of cells, strand segregation of the replicated DNA can be asymmetric. All template strands can segregate to the daughter cell, resulting in a cell with stem cell properties, whereas daughter cells that acquire the nascent DNA strands undergo myogenic differentiation [485]. A separate side-population of satellite cells has also been identified [487], expressing the satellite cell markers Pax7, Syndecan-4 and Syndecan-3, as well as the side population marker ABCG2. These cells have been shown to be distinct from myogenic progenitor cells. They form a self-renewing pool of cells capable of dividing into satellite cells and forming multinucleated myotubes; they are thought to aid in the repair of injured muscle fibres. Beauchamp et al. [488] have further identified a distinct population of satellite cells that are CD34- and Myf5-, involved in maintaining the lineage-committed majority, differing from CD34+/Myf5+/M-cadherin+ cells that are quiescent and committed down the myogenic lineage.

Despite the identification of these different markers of subpopulations of satellite cells present in adult skeletal muscle, little is actually known about the functional significance of the different populations, or how much more transcriptional diversity is present in the satellite cell population.

8.1.2. Single-Cell Analysis in Skeletal Muscle

Cho et al. [489] examined the transcriptome of single murine muscle satellite cells and showed that satellite cells *in vivo* existed along an activation continuum, rather than as two distinct states of being. Metabolic heterogeneity was also identified in these satellite cells, with some satellite cell subpopulations that are metabolically poised to respond to certain stimuli and activate down a myogenic commitment. Skeletal muscle needs to respond to many different stimuli, with satellite cells activation requiring precise response to stimuli. The presence of diverse satellite cell populations with cells poised with different metabolic states may allow a rapid and efficient response to a wide range of different stimuli. As well as heterogeneity in the total transcriptome of mouse satellite cells, Chapman et al. [490] showed that there is significant diversity in the cell surface markers of satellite cells, with this heterogeneity being functionally important with regards to the myogenic differentiation capacity of the progenitor cells. Different myofibres were associated with satellite cells expressing different surface markers, suggesting that individual fibres may have different efficiencies in the repair mechanisms by resident satellite cells.

These studies have provided exploratory analysis of satellite cell heterogeneity in murine skeletal muscle in a relatively small number of cells. However, there has been no previous investigation into the transcriptional heterogeneity of human satellite cells.

8.1.3. Aims

The aim of this study was to investigate the transcriptome of myoblast cells at the single cell level, and investigate the heterogeneity of myoblasts from human skeletal muscle tissue. This investigation was designed as a discovery study, using 8 healthy controls, 3 pre-sarcopenic and 5 sarcopenic individuals. DropSeq was used to isolate and encapsulate the transcriptomes of single cells for downstream Poly-A transcriptome sequencing. The heterogeneity of the myoblasts was examined as a whole as well as differences between the myoblasts from control and sarcopenic participants of HSSe. The different subpopulations of myoblasts were examined for genes expressed primarily in the subpopulation, distinguishing the different cell populations from each other. Finally, we examined the pathways enriched in each individual cell population, to determine any functional differences between the cell clusters.

8.2. Results

8.2.1. Cohort Characteristics

16 myoblast samples from participants of HSSe were sequenced (8 controls, 3 pre-sarcopenics and 5 sarcopenics). All myoblast samples were cultured from male participants (table **8.1**). The mean age (S.D.) was 77.5 (2.36) years, height 1.73 (0.06) m, weight 78.13 (11.11) kg, BMI 26.03 (3.05) kg/m^2, total lean body mass 48.59 (6.17) kg, appendicular lean mass 21.62 (3.09) kg, total fat mass 26.68 (9.50) kg, gait speed 0.92 (0.24) m/s and grip strength 33.25 (7.49) kg.

Table 8.1: *Cohort characteristics for participants of HSSe whose myoblast cultures were used for the scRNAseq analysis*

	Myoblasts (n=16)	
	Mean	**S.D.**
Age (yrs)	77.50	2.36
Height (m)	1.73	0.06
Weight (kg)	78.13	11.11
BMI (kg/m^2)	26.03	3.05
Total Lean Body Mass (kg)	48.59	6.17
Appendicular Lean Mass (kg)	21.62	3.09
Total Fat mass (kg)	26.68	9.50
Gait Speed (m/s)	0.92	0.24
Grip Strength (kg)	33.25	7.49

8.2.2. Cell/Gene Level QC Metrics

Digital Gene Expression (DGE) files were examined to determine the number of barcodes corresponding to actual cells, and not corresponding to empty beads or noise. Knee plots of the cumulative fraction of reads and density plots (figure **8.1**) were generated for each DGE file. To determine the number of barcodes that corresponded to actual cells, the first inflection point after the initial peak of the density plot was taken, corresponding to the bend in the knee plot of the cumulative fraction of reads before the plot began to plateau. The average number of cells per sample was 253.56 ± 96.34 (mean ± S.D.). The average total number of reads per sample was 30 million (30,285,289 ± 8,955,094; mean ± S.D.), resulting in an average of 187,193 ± 85,221 reads per cell (mean ± S.D.).

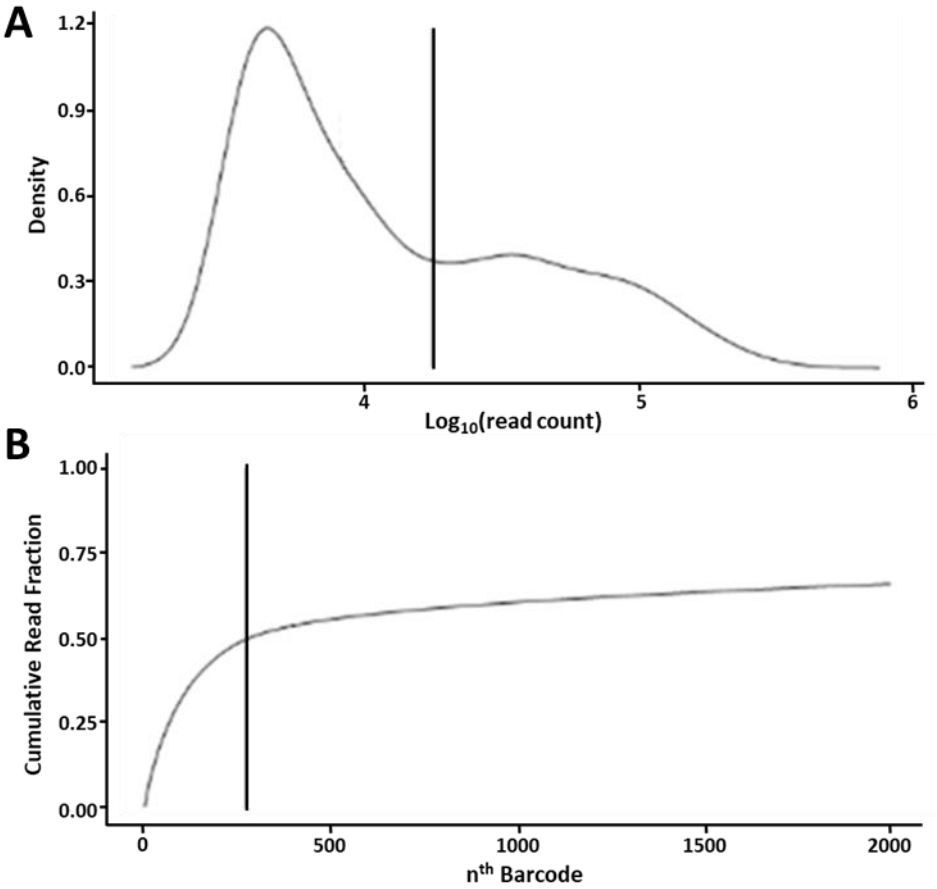


Figure 8.1: Representative density plot and cumulative read fraction (knee) plot to determine actual number of cells sequenced

(A) Density plot of the read counts for all the cell barcodes. Left of the vertical line represents the majority of the density of reads in the sample. The vertical line is plotted at the first inflection point after the initial density peak. (B) Cumulative read fraction (knee) plot. Vertical line represents the same vertical line as in (A), with that barcode being the total number of cells sequenced and all other cell barcodes representing noise and empty barcodes.

A histogram of the number of transcripts per cell (figure **8.2a**) suggested that several barcodes were exposed to multiple cells (doublets), which were removed from downstream analysis due to uncertainty as to which cell the transcripts originated from. The number of transcripts per cell and the number of reads per cell showed a linear relationship (figure **8.2b**) suggesting good sequencing data.

The percentage of reads mapping to mitochondrial genes is a sign of cellular integrity. When a cell lyses, cytoplasmic RNA leaves the cell. However, mitochondrial RNA remain in the mitochondria, as the mitochondrial membrane remains intact. Therefore, a high percentage of reads mapping to mitochondrial genes is a sign of cellular lysis. Cells with more than 20% of reads mapping to

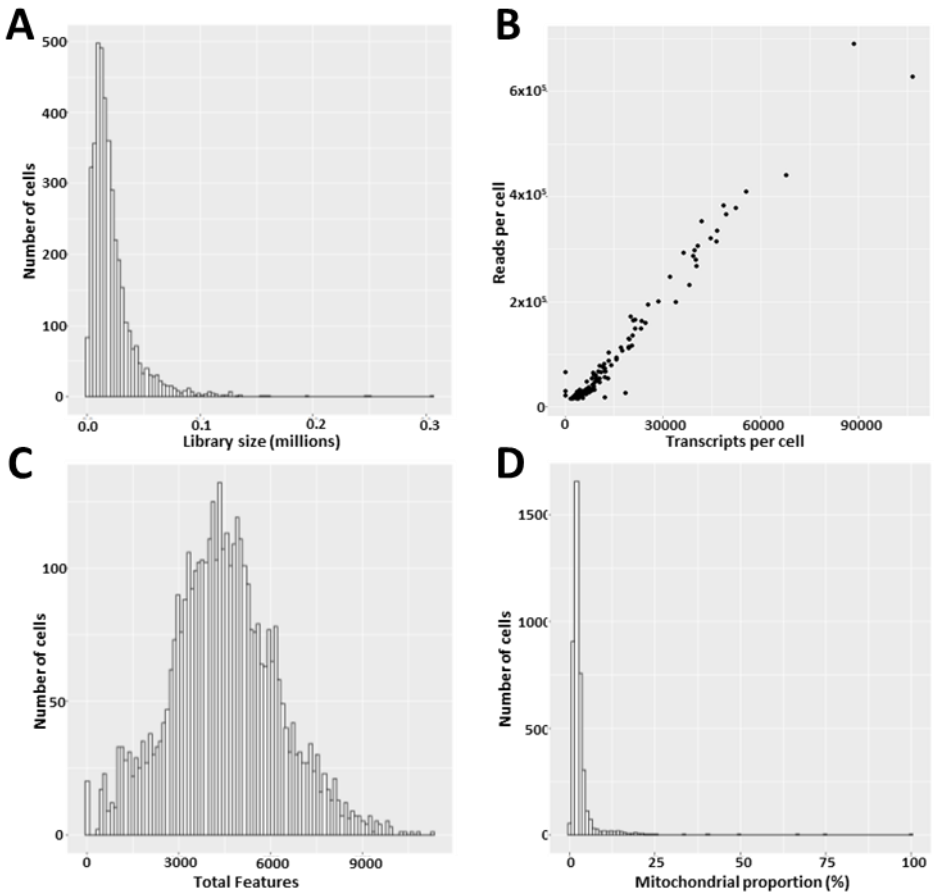


Figure 8.2: Cell QC metric plots

(A) Histogram of the library sizes of all the cells, showing a positively skewed distribution. Most of the cells had similar library sizes; however there were some cells that had much larger library sizes (e.g. 2x and 3x average). These suggested the presence of doublets or triplets in the dataset. (B) Scatter plot of reads per cell against transcripts per cell, showing a strong linear relationship. (C) Histogram of the total features in each cell of the dataset, showing a normal distribution of features, with cell outliers showing very large numbers of features compared to the rest of the cells. (D) Histogram of the percent of reads mapping to mitochondrial genes. As expected, most of the cells had a low proportion of reads mapping to mitochondrial genes, although there were some cells with a high proportion.

mitochondrial genes were discarded. Twenty-six cells were discarded because of this (figure **8.2d**). Sixty-seven cells were discarded due to the presence of a low number of total genes mapped (figure **8.2c**), indicating cells with a low complexity most likely due to incomplete capture of RNA during the encapsulation process. Removal of all the low quality cells resulted in 4009 cells with good quality sequencing data for downstream analysis.

Genes with low expression or only expressed in a handful of cells are unlikely to differ between different subpopulations of cells. Therefore, any genes that were expressed in no cells at all were

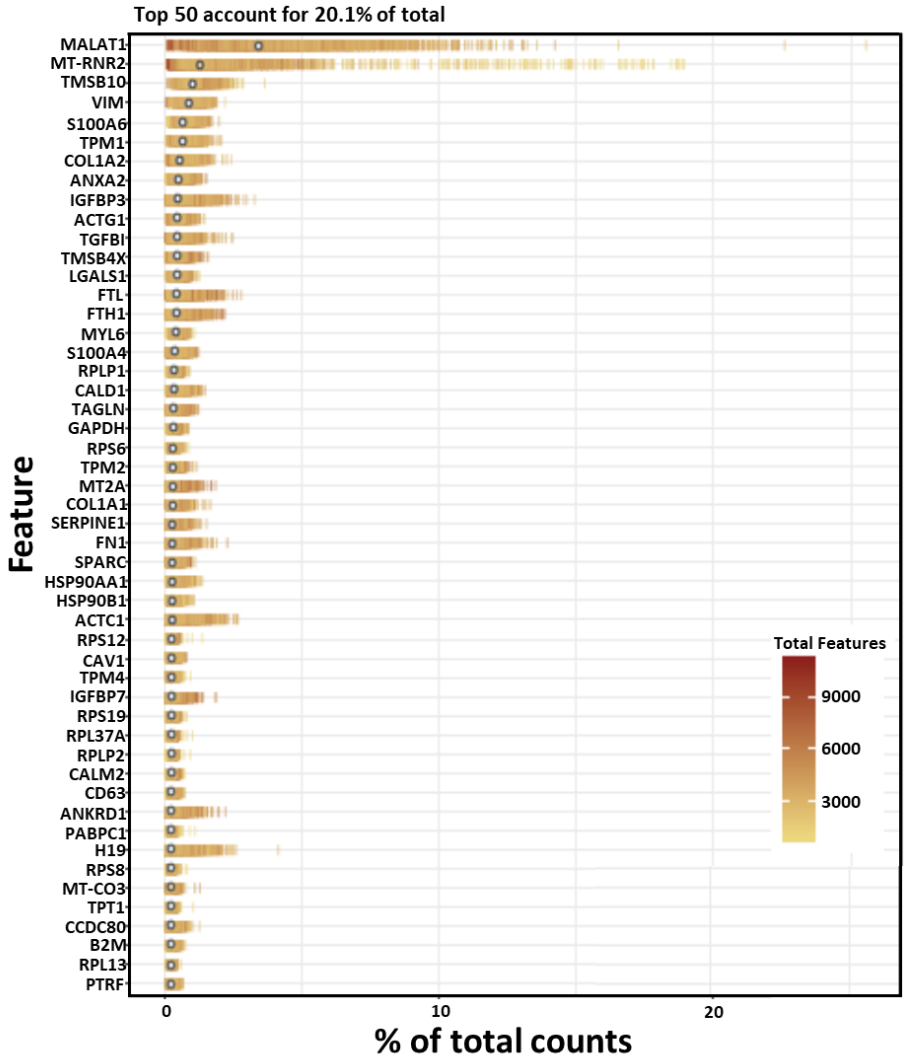


Figure 8.3: Expression plot of the 50 genes with the highest expression across the dataset

The top 50 genes are plotted with the percentage of the total reads that were associated with the gene, with each cell coloured according to the total number of features in the cell. As expected, the majority of these genes were constitutively expressed genes, including mitochondrial genes, structural genes, heat shock proteins and ubiquitously expressed genes. These 50 genes accounted for 20.1% of the total reads in the dataset.

removed from the dataset. Genes with an average log expression < -1 across all cells were also removed from the dataset. 18,877 genes with low expression across the cells were removed from further analysis. The 50 most abundant genes are shown in figure **8.3**. As expected, the majority of the top 50 most abundant genes in the myoblasts were constitutively expressed genes, including mitochondrial genes (e.g. MT-RNR2 and MT-CO3), as well as muscle structural genes (e.g. MYL6 and COL1A2) and ubiquitously expressed genes (e.g. GAPDH and B2M). This suggests that library prep and pre-processing of data files was ideal, because the genes one would expect to show highest expression in these cells are the ones showing highest expression.

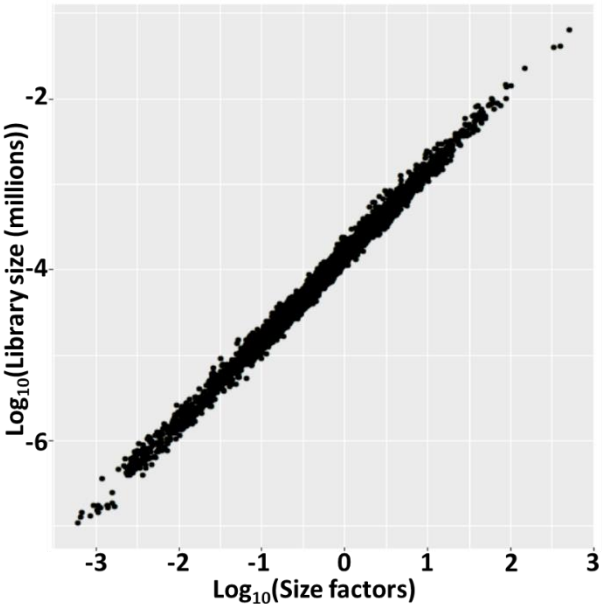


Figure 8.4: Scatter plot of library sizes against size factors
A strong linear relationship between the size factors and total library size suggested the size factors are suitable for normalization of the data to remove bias due to sequencing depth and capture efficiency.

8.2.3. Data Normalization

To normalize the data to account for different sequencing depths and capture efficiency, size factors were computed for each cell using a deconvolution method. As is evident from figure **8.4**, the size factors were tightly correlated with total library size for each cell, suggesting that any cell differences were due to differences in sequencing depth and capture efficiency. Substantial differential expression between cells would result in increased scatter around the trend and disrupt the linear relationship between the size factors and library size, which does not occur in this dataset, suggesting a likely homogenous cell population.

8.2.4. Cell Cycle Classification

Cell cycle classification was carried out using a previously published method. Cell cycle genes for each phase of the cell cycle were downloaded from the Reactome database (http://www.reactome.org) and the CellCycleSorting function of the Seurat package (v2.3.4) was used to determine the cell cycle phase of each cell. The algorithm compares the expression of marker genes for the G2/M and S phases and as these marker sets should be anticorrelated, a G2/M and an S score is given to each cell based on the expression patterns of the markers. Cells expressing neither marker sets are most likely not cycling and in G1 phase. G1 cell cycle scores and G2/M cell

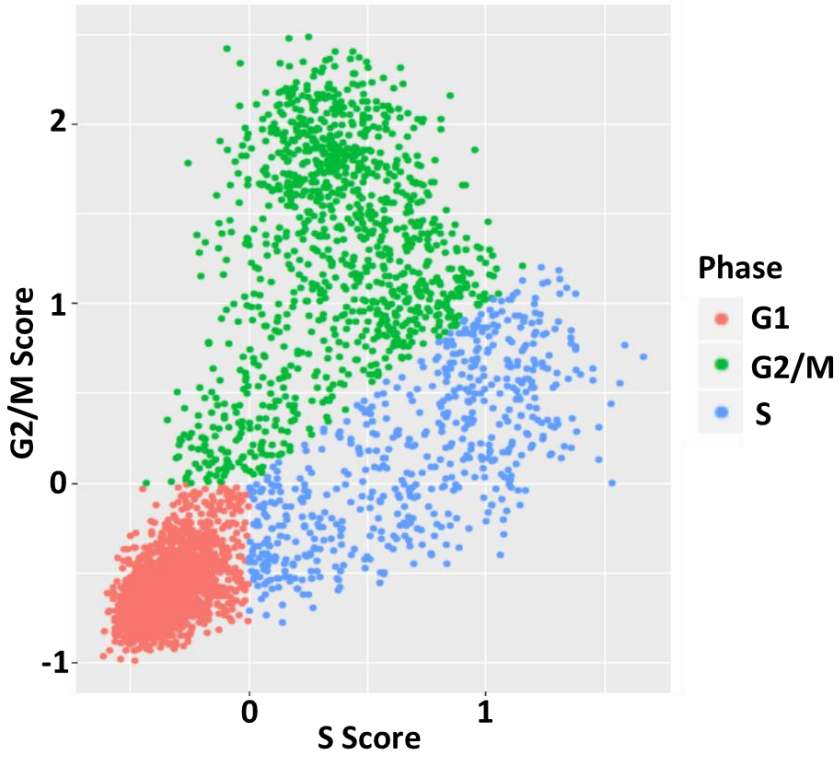


Figure 8.5: Cell cycle classification of the myoblasts

Cell cycle classification based on the expression of G2M and S phase marker genes.

cycle scores were plotted to generate a cell cycle plot (figure **8.5**). 2344 cells were classified as being in G1 phase, 1080 cells were in G2/M phase and 585 cells were classified as being in S phase. This suggests that the cells in the dataset are not synced, and in different cell cycle phases. Therefore, to eliminate confounding effects from cell cycle-induced differences in gene expression, the assigned phase of each cell was adjusted for in downstream analysis, to prevent the discarding of information by only analysing cells in G1 phase. Despite a cell cycle stage being assigned to each cell, this remains an estimation and experimental validation using fluorescence-assisted cell sorting (FACS) to ensure cells sequenced are all in the same cell cycle stage would have been preferable.

8.2.5. Identification of Highly Variable Genes (HVGs)

Highly variables genes (HVGs) are genes that may be involved in driving cellular heterogeneity in the myoblasts. 761 genes were classed as being HVGs in the myoblasts, with an FDR < 0.05 and a biological variance > 0.1. A list of the top 50 HVGs is in table **8.2** and violin plots of the top 10 HVGs (figure **8.6**) showed that they were not driven by any outlier cells. The correlations of the HVGs with each other were calculated. Tightly correlated HVGs distinguishes between HVGs that arise from random noise and those involved in driving systematic differences between subpopulations. From the set of 761 HVGs, 177,518 significant correlations (FDR < 0.05) were identified. The top 50

correlations are shown in table **8.3**. This high number of significant correlations suggests that there is some substructure in the dataset, with subpopulations of cells showing systematic differences in their expression profiles.

Table 8.2: *List of the top 50 highly variable genes ordered based on their biological variance*

	Mean Expression	Total Variance	Biological Variance	Technical Variance	P Value	FDR
ACTC1	3.755	6.041	5.085	0.956	<2.23E-308	<2.23E-308
H19	3.881	5.388	4.449	0.940	<2.23E-308	<2.23E-308
NES	2.816	3.665	2.593	1.072	<2.23E-308	<2.23E-308
PTX3	2.649	3.614	2.525	1.088	<2.23E-308	<2.23E-308
ANKRD1	4.347	3.377	2.498	0.879	<2.23E-308	<2.23E-308
CTGF	3.252	3.113	2.092	1.021	<2.23E-308	<2.23E-308
POSTN	1.224	3.039	2.055	0.984	<2.23E-308	<2.23E-308
ACTA2	2.652	2.934	1.845	1.088	<2.23E-308	<2.23E-308
TNNT2	0.851	2.612	1.811	0.801	<2.23E-308	<2.23E-308
ELN	2.372	2.866	1.757	1.109	<2.23E-308	<2.23E-308
OXTR	3.119	2.655	1.617	1.038	<2.23E-308	<2.23E-308
LRRC17	1.951	2.710	1.593	1.116	<2.23E-308	<2.23E-308
THBS1	4.190	2.450	1.550	0.899	<2.23E-308	<2.23E-308
HIST1H4C	2.993	2.506	1.454	1.053	<2.23E-308	<2.23E-308
SCUBE3	2.410	2.527	1.421	1.106	<2.23E-308	<2.23E-308
HAS2	2.821	2.481	1.410	1.072	<2.23E-308	<2.23E-308
DES	1.331	2.380	1.360	1.019	<2.23E-308	<2.23E-308
DKK1	2.799	2.427	1.353	1.074	<2.23E-308	<2.23E-308
ADAMTS1	2.841	2.412	1.343	1.069	<2.23E-308	<2.23E-308
CHRNA1	1.268	2.326	1.326	0.999	<2.23E-308	<2.23E-308
FST	1.591	2.341	1.260	1.081	<2.23E-308	<2.23E-308
FTH1	5.477	2.005	1.258	0.747	<2.23E-308	<2.23E-308
RGS4	1.706	2.343	1.244	1.098	<2.23E-308	<2.23E-308
RARRES2	0.862	2.043	1.235	0.808	<2.23E-308	<2.23E-308
INHBA	3.725	2.168	1.208	0.960	<2.23E-308	<2.23E-308
HMOX1	1.747	2.293	1.190	1.103	1.84E-303	4.20E-301
SERPINE2	2.849	2.197	1.128	1.069	3.25E-293	7.25E-291
ATP2B1	4.220	2.011	1.116	0.896	<2.23E-308	<2.23E-308
PLAU	1.868	2.227	1.115	1.113	2.24E-269	4.35E-267
SDPR	1.542	2.186	1.114	1.072	9.86E-286	2.07E-283
SERPINE1	5.240	1.885	1.112	0.773	<2.23E-308	<2.23E-308
IGFBP3	5.890	1.810	1.106	0.704	<2.23E-308	<2.23E-308
ALDH1A1	2.273	2.214	1.100	1.113	1.15E-263	2.15E-261
CRYAB	3.485	2.067	1.075	0.992	4.24E-306	9.87E-304
TNNT1	0.596	1.679	1.056	0.622	<2.23E-308	<2.23E-308
SULF1	1.757	2.125	1.021	1.104	1.67E-236	2.69E-234
MYF5	0.979	1.879	1.005	0.875	<2.23E-308	<2.23E-308
PLCB4	2.158	2.102	0.985	1.117	3.69E-219	5.53E-217
MT2A	5.125	1.765	0.980	0.785	<2.23E-308	<2.23E-308
STC2	3.236	1.978	0.954	1.023	8.43E-240	1.38E-237
COL11A1	1.522	1.995	0.927	1.068	2.05E-213	2.90E-211

DSP	1.965	2.040	0.924	1.117	4.39E-197	5.76E-195
IGFBP7	4.725	1.756	0.923	0.832	3.86E-317	9.42E-315
IGF2	0.608	1.535	0.903	0.632	<2.23E-308	<2.23E-308
COL8A1	2.507	1.995	0.895	1.100	5.22E-192	6.76E-190
SLC7A11	1.820	2.003	0.894	1.110	1.06E-188	1.34E-186
SRGN	2.147	2.006	0.889	1.117	8.74E-185	1.07E-182
RPS18	3.877	1.827	0.887	0.940	2.40E-244	4.06E-242
TNFRSF11B	1.184	1.839	0.871	0.968	5.42E-226	8.23E-224
HAPLN1	0.879	1.673	0.855	0.818	2.02E-288	4.40E-286

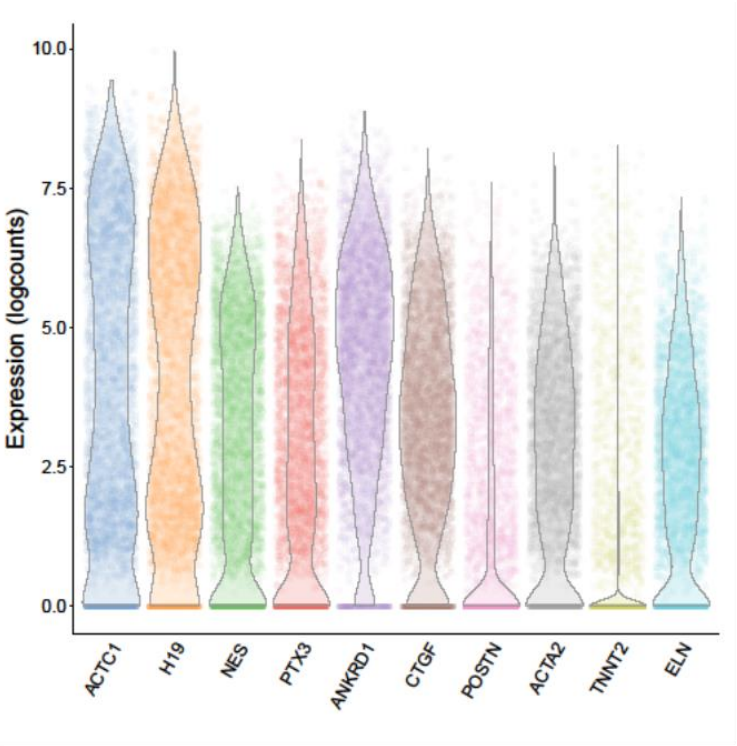


Figure 8.6: Violin plots of the top 10 HVGs

Violin plots for these 10 genes showed random spread in the distribution of the expression of these genes across the cells, suggesting that the variability of the HVGs is not driven by outlier cells.

8.2.6. Cell Clustering

To determine whether there were subpopulations present in the myoblast dataset, the normalized and covariate-adjusted expression values for the correlated HVGs underwent hierarchal clustering on the Euclidean distance between cells. Ten different subpopulations of cells with distinct gene expression patterns were identified (figure **8.7**). There was significant overlap between the different clusters, which might be expected as the cells have been enriched to generate a pure myoblast population during the generation of the myoblast cultures from the muscle biopsies.

Nevertheless, the different clusters may represent myoblasts at different stages of development and differentiation.

Table 8.3: *Top 50 correlations between the HVGs*

Gene 1	Gene 2	Spearman rho	P Value	FDR
FTH1	FTL	0.762	2.00E-06	6.52E-06
RPS18	TMSB4X	0.748	2.00E-06	6.52E-06
ACTC1	H19	0.733	2.00E-06	6.52E-06
H19	NES	0.719	2.00E-06	6.52E-06
RPS18	RPS13	0.709	2.00E-06	6.52E-06
RPS18	RPS23	0.696	2.00E-06	6.52E-06
ACTC1	NES	0.689	2.00E-06	6.52E-06
RPS18	RPS2	0.686	2.00E-06	6.52E-06
FTH1	TXN	0.682	2.00E-06	6.52E-06
TMSB4X	RPS13	0.680	2.00E-06	6.52E-06
FTH1	RPS18	0.677	2.00E-06	6.52E-06
RPS18	RPS15A	0.675	2.00E-06	6.52E-06
TMSB4X	RPS23	0.673	2.00E-06	6.52E-06
RPS18	RPS27A	0.670	2.00E-06	6.52E-06
RPS13	RPS23	0.666	2.00E-06	6.52E-06
FTH1	TMSB4X	0.664	2.00E-06	6.52E-06
RPS18	RPL26	0.661	2.00E-06	6.52E-06
TMSB4X	RPS2	0.658	2.00E-06	6.52E-06
TMSB4X	NDUFA4	0.658	2.00E-06	6.52E-06
RPS23	RPS2	0.651	2.00E-06	6.52E-06
RPS18	NDUFA4	0.651	2.00E-06	6.52E-06
TMSB4X	RPS27A	0.649	2.00E-06	6.52E-06
RPS23	RPS27A	0.646	2.00E-06	6.52E-06
RPS18	RPS15	0.644	2.00E-06	6.52E-06
FTL	TXN	0.644	2.00E-06	6.52E-06
RPS13	RPS27A	0.642	2.00E-06	6.52E-06
NES	ATP2B1	0.641	2.00E-06	6.52E-06
RPS15A	RPS23	0.640	2.00E-06	6.52E-06
TMSB4X	RPS15A	0.637	2.00E-06	6.52E-06
FTH1	RPS13	0.634	2.00E-06	6.52E-06
RPS13	RPS2	0.634	2.00E-06	6.52E-06
RPL26	RPS23	0.633	2.00E-06	6.52E-06
H19	CHRNA1	0.631	2.00E-06	6.52E-06
TMSB4X	RPS15	0.631	2.00E-06	6.52E-06
H19	ATP2B1	0.627	2.00E-06	6.52E-06
RPS15A	RPS13	0.627	2.00E-06	6.52E-06
ACTC1	CHRNA1	0.627	2.00E-06	6.52E-06
RPS2	RPS27A	0.625	2.00E-06	6.52E-06
TMSB4X	OST4	0.624	2.00E-06	6.52E-06
TMSB4X	S100A11	0.620	2.00E-06	6.52E-06
RPL26	RPS13	0.620	2.00E-06	6.52E-06
TMSB4X	RPL26	0.617	2.00E-06	6.52E-06
TMSB4X	TXN	0.616	2.00E-06	6.52E-06

FTH1	RPS23	0.614	2.00E-06	6.52E-06
FTL	NQO1	0.614	2.00E-06	6.52E-06
H19	DES	0.614	2.00E-06	6.52E-06
RPL26	RPS2	0.613	2.00E-06	6.52E-06
RPS18	RPL14	0.604	2.00E-06	6.52E-06
NDUFA4	RPS23	0.603	2.00E-06	6.52E-06
RPS15A	RPS2	0.602	2.00E-06	6.52E-06

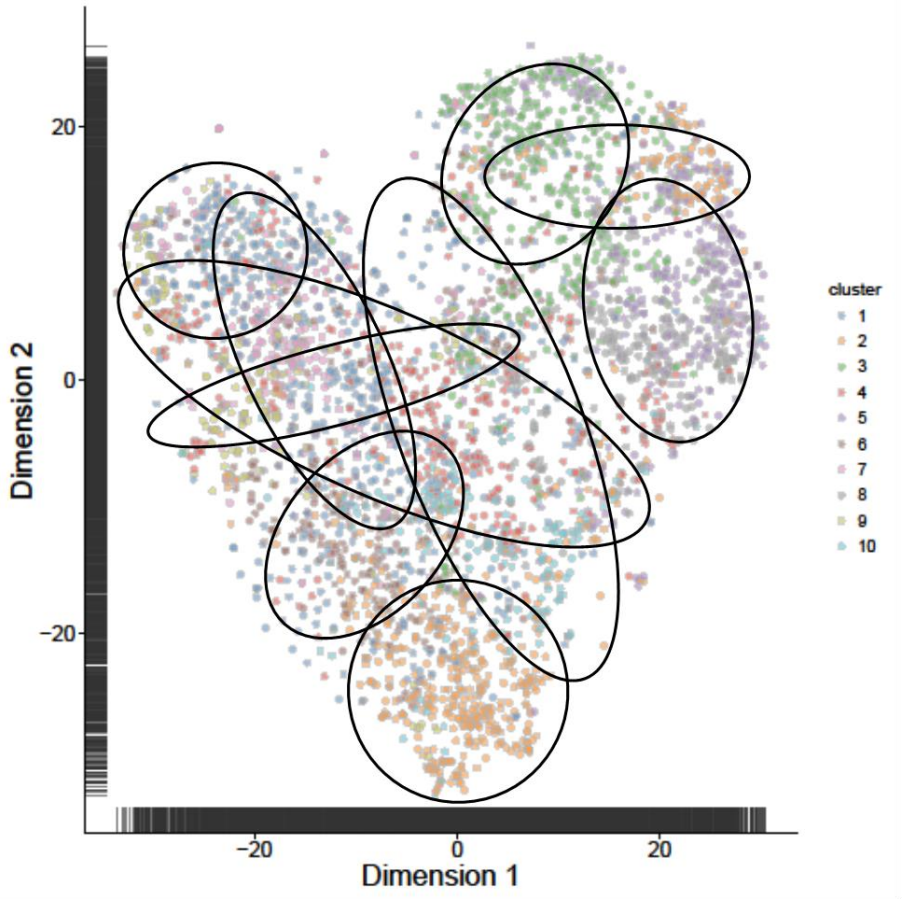


Figure 8.7: t-SNE plot of the cell subpopulations

Hierarchal clustering based on Euclidean distances of the cells identified 11 separate cell subpopulations based on their gene expression. The cells were coloured based on which cell cluster they were part of. There was overlap between the different cell clusters, likely due to subtle differences between the subpopulations as the cells were enriched for myoblasts. Each point represents an individual cell, coloured based on the cluster they belong to.

To determine if there was an association between cluster occupancy between myoblasts from control, pre-sarcopenic and sarcopenic individuals, a Chi-Square Test of Independence was used. This showed a highly significant association between the clustering and sarcopenia status of the cells ($X^2(18) = 150.711$, $p=5.4057x10^{-23}$). Several independent clusters of cells were predominantly

Table 8.4: Number of cells in each cluster, divided by sarcopenia status

			Status				
			Control	Pre-sarc	Sarcopenia		Total
Cluster	1	Count	238	199	326		763
		Expected Count	336.1	175.5	251.4		763.0
		% within Cluster	31.2%	26.1%	42.7%		100.0%
		% within Status	13.5%	21.6%	24.7%		19.0%
	2	Count	273	156	173		602
		Expected Count	265.2	138.4	198.4		602.0
		% within Cluster	45.3%	25.9%	28.7%		100.0%
		% within Status	15.5%	16.9%	13.1%		15.0%
	3	Count	225	111	180		516
		Expected Count	227.3	118.7	170.0		516.0
		% within Cluster	43.6%	21.5%	34.9%		100.0%
		% within Status	12.7%	12.0%	13.6%		12.9%
	4	Count	214	81	147		442
		Expected Count	194.7	101.7	145.6		442.0
		% within Cluster	48.4%	18.3%	33.3%		100.0%
		% within Status	12.1%	8.8%	11.1%		11.0%
	5	Count	192	72	119		383
		Expected Count	168.7	88.1	126.2		383.0
		% within Cluster	50.1%	18.8%	31.1%		100.0%
		% within Status	10.9%	7.8%	9.0%		9.6%
	6	Count	161	74	112		347
		Expected Count	152.9	79.8	114.3		347.0
		% within Cluster	46.4%	21.3%	32.3%		100.0%
		% within Status	9.1%	8.0%	8.5%		8.7%
	7	Count	157	66	58		281
		Expected Count	123.8	64.6	92.6		281.0
		% within Cluster	55.9%	23.5%	20.6%		100.0%
		% within Status	8.9%	7.2%	4.4%		7.0%
	8	Count	113	92	54		259
		Expected Count	114.1	59.6	85.3		259.0
		% within Cluster	43.6%	35.5%	20.8%		100.0%
		% within Status	6.4%	10.0%	4.1%		6.5%
	9	Count	83	30	96		209
		Expected Count	92.1	48.1	68.9		209.0
		% within Cluster	39.7%	14.4%	45.9%		100.0%
		% within Status	4.7%	3.3%	7.3%		5.2%
	10	Count	110	41	56		207
		Expected Count	91.2	47.6	68.2		207.0
		% within Cluster	53.1%	19.8%	27.1%		100.0%
		% within Status	6.2%	4.4%	4.2%		5.2%
Total		Count	1766	922	1321		4009
		Expected Count	1766.0	922.0	1321.0		4009.0
		% within Cluster	44.1%	23.0%	33.0%		100.0%
		% within Status	100.0%	100.0%	100.0%		100.0%

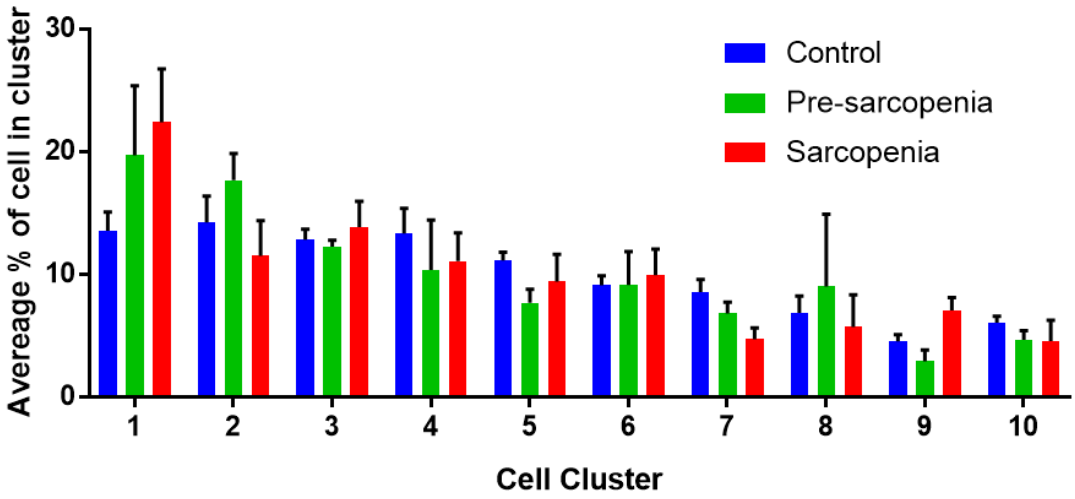


Figure 8.8: Average percentage of cells in each cluster

There were clear differences between some of the clusters, with some of them being predominantly comprised of cells with one sarcopenia status over the other two. Bars are coloured based on sarcopenia status, mean percentage of cells across individuals ± s.e.m.

formed of cells that originated from either control, pre-sarcopenic or sarcopenic individuals (table **8.4**, figure **8.8**). The error bars in figure **8.8** suggest that it is not an equal number of cells from each individual contributing to the occupancy of each cell cluster and there may be some individuals that contribute more cells to the individual clusters. However, the differences between clusters remain, with some cluster more likely to contain cells from a sarcopenic individual compared to another cluster which may be more likely to contain cells from a pre-sarcopenic individual.

Cluster 1 was associated with sarcopenia and pre-sarcopenia, with 24.7% of the cells from sarcopenic individuals found in this cluster and 21.6% of cells from pre-sarcopenic individuals. However, only 13.5% of cells from control individuals were present in this cluster. This was much more than the expected number of cells in cluster 1 from the sarcopenic muscle, with approximately 29.7% more cells from sarcopenic individuals in this cluster than expected, as well as a 13.4% increase in the expected number of cells from pre-sarcopenic individuals. This was associated with a 29.2% decrease in the number of cells from control individuals in this cluster. Cluster 8 was found to be potentially associated with pre-sarcopenic, enriched with cells from pre-sarcopenic individuals. 10.0% of the cells from pre-sarcopenic individuals can be found in cluster 10, with only 6.4% and 4.1% of cells from control and sarcopenic individuals respectively. This represented a 54.4% higher occupancy of cluster 8 than one would expect. Concomitantly, there was a 36.7% reduction in the expected number of cell from sarcopenic individuals in cluster 8.

No one cell cluster was found to be comprised predominantly of cells from control individuals. However, cluster 5 had the highest number of control cells compared to pre-sarcopenic and

sarcopenic cells, with 50.1% of the cells in cluster 5 being control cells, accounting for 10.9% of all control cells. This was a greater proportion of cells contributing to the composition of cluster 5 compared to pre-sarcopenic and sarcopenic cells, which account for 18.8% and 31.1% of all cells in cluster 5 respectively. This was a 13.8% increase in the number of control cells in cluster 5 than would be expected and 18.3% fewer cells from pre-sarcopenic individuals than would be expected. All other clusters appeared to contain approximately equal numbers of control, pre-sarcopenic and sarcopenic cells.

8.2.7. Marker Genes between Cell Subpopulations

We next investigated whether there were genes that can distinguish between the different cell populations. Table **8.5** shows the top 20 marker genes for each cell cluster. There was substantial overlap between the genes which distinguish the different cell clusters. However, the pattern of expression of these marker genes in each cluster was distinct, with some of the genes showing a high expression in one cluster, but decreased or non-existent expression in the other clusters. Many of these marker genes play a key role in normal muscle function including: CHRNA1, the acetylcholine receptor subunit alpha, part of the muscle acetylcholine receptor; DES, desmin, a structural component of skeletal muscle and marker for skeletal muscle myoblasts; H19, a lncRNA that plays a role in skeletal muscle differentiation and proliferation of myoblasts; and MYF5, myogenic factor 5, which plays a key role in regulating skeletal muscle differentiation and myogenesis. In the top 50 marker genes, several other muscle specific genes distinguished the different cell populations, including, MYOD1, TNNT1 and TTN, all of which play key roles in the structure and function of normal muscle. Within the list of marker genes, there were also many genes associated with cellular proliferation and cell division, including PCNA and MKI67. MKI67 is a known marker for cellular proliferation, and showed varied expression across all the cells, suggesting they are at different stages of proliferation potential, and may have started the differentiation program of gene expression.

The different patterns of expression of the marker genes distinguished the individual subpopulations of cells from the rest of the cells, despite an overlap of marker genes between the clusters. Cluster 1 showed the low expression of the following genes: CENPF, BIRC5, STMN1, HMGB1, CDH15, MYF5, NES, CHRNA1 and DES; while there was high expression of the following genes: H19, PTX3, THBS1, FGF2, OST4 and FTH1. The low and high expression of these genes, in combination with the pattern of expression of the remainder of the marker genes for cluster 1 distinguished it from the other subpopulations of cells. Cluster 8 showed the highest expression of PPIB, NDUFA4, RPS15A, RPS18 and RPL19, together with high expression of ACTC1, H19, NES, ATP2B1, NCL and ANP32B. The expression of these genes in combination with the pattern of

Table 8.5: *List of the top 20 marker genes for each cell cluster*

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10
ATP2B1	IGFBP7	H19	ANXA2	FTH1	ANXA2	ATP2B1	FTH1	FTH1	ATP2B1
DST	NAP1L1	NAP1L1	ATP2B1	GREM1	LRRC17	DST	H19	IGFBP7	FTH1
FTH1	RPS18	NES	FTH1	IGFBP7	NAP1L1	IGFBP7	TMSB4X	NES	H19
LRRC17	TMSB4X	STMN1	GREM1	NEAT1	NES	NAP1L1	ACTC1	PTRF	NEAT1
NAP1L1	CHRNA1	TMSB4X	NAP1L1	NES	TMSB4X	NES	ANXA2	STMN1	RPS18
STMN1	FN1	ACTC1	TMSB4X	RPS18	ACTC1	PTRF	RPS18	TIMP1	TMSB4X
TIMP1	H19	CENPF	ASPH	TMSB4X	ATP2B1	STMN1	ATP2B1	CCBE1	ACTC1
TMSB4X	NES	FTH1	FN1	ANXA2	CCBE1	TMSB4X	FTL	IL7R	CDH15
ANXA2	RPL31	RPS18	H19	ASPH	RPS18	ANXA2	GREM1	NAP1L1	H2AFZ
CDH15	RPS23	RPS2	RPL19	ATP2B1	SERPINE1	IL7R	RPS19	RPL19	NES
CENPF	SERPINE1	SPARC	RPS2	H19	COL4A1	SPARC	CHRNA1	TMSB4X	RPS23
CHRNA1	ATP2B1	TUBA1B	CHRNA1	H2AFZ	FTH1	ACTB	RPL31	TUBA1B	TXN
IGFBP7	COL4A1	CHRNA1	FTL	RPL31	H19	H19	SERF2	TXN	ANXA2
PTRF	NDUFA4	DEK	OST4	SPARC	MYF5	OST4	NCL	ACTB	DES
H19	RPS19	DES	RPS15	FTL	RGMB	RPS18	NDUFA4	DEK	FTL
HMGB1	ALCAM	HMGB1	RPS18	NDUFA4	DES	CENPF	RPS27A	FN1	NAP1L1
MYF5	RPLP1	IGFBP7	ALCAM	SMC4	RPS2	CHRNA1	TXN	FTL	RPS19
NES	TAGLN	RPS15	NDUFA4	CENPF	SERPINE2	FTH1	DES	OST4	SMC4
RPS18	TXNRD1	ATP2B1	RPS14	COL1A2	TAGLN	LRRC17	HSP90AA1	RGMB	CHRNA1
SPARC	CD59	CEP55	TXNRD1	FN1	TXN	RPS14	RPS13	RPS18	COL1A2

expression of the rest of the marker genes for cluster 8 distinguished it from other subpopulations of cells. Cluster 5 showed high expression of ACTC1, NES, SMC4, CENPF, ATP2B1 and H19, together with low expression ANXA2, OST4, FTH1, SPARC and TGFBI. This expression pattern together the rest of the cluster 10 marker genes distinguished cluster 10 from the other subpopulations of cells.

8.2.8. Marker Genes Pathway Analysis

To determine which pathways are altered within each subpopulation of cells, the 100 top marker genes for each cluster were inputted into a gene ontology (GO) enrichment analysis. This allowed us to determine the pathways that marker genes were enriched for and the pathways that differed between particular clusters of cells, reflecting a functional difference between the cell clusters. As clusters 1, 5 and 8 were comprised of predominantly sarcopenic, control and pre-sarcopenic cells respectively, we examined differences between these three clusters.

Using the 'Biological process' ontology, there were 234 GO terms enriched for the cluster 1 gene markers (FDR < 0.05). Of these enriched GO terms, several terms were associated with skeletal muscle function and processes involved with the catabolism of cellular components key to function. The enriched terms included: muscle filament sliding (GO:0030049, FDR=0.0408), myoblast differentiation (GO:0045445, FDR=0.0115), muscle structure development (GO:0061061, FDR=0.0208) and positive regulation of striated muscle tissue development (GO: 0045844, FDR=0.0292). Several GO terms associated with the extracellular matrix and cell adhesion were also enriched, including extracellular matrix organisation (GO:0030198, FDR=0.0002), extracellular structure organisation (GO: 0043062, FDR=0.0002) and cell adhesion (GO:0007155, FDR=0.0002). These enriched terms suggest that this cluster, formed predominantly of sarcopenic cells, has changes in key pathways of muscle contraction, together with changes in the extracellular structure of the muscle tissue and subsequent muscle development.

There were 199 GO terms enriched for the cluster 5 gene markers (FDR < 0.05). Many of the significant GO terms were associated with protein translation, including: protein targeting to ER (GO:0045047, FDR=1.12×10^{-14}), translational initiation (GO:0006413, FDR=1.17×10^{-11}), ribosome biogenesis (GO:0042254, FDR=4.74×10^{-7}), rRNA processing (GO:0006364, FDR=1.57×10^{-8}) and RNA catabolic process (GO:0006401, FDR=1.5×10^{-6}). There was also significant enrichment in GO terms associated with cellular proliferation and structural organisation, including extracellular structure organisation (GO:0043062, FDR=1.66×10^{-6}), collagen fibril organisation (GO:0030199, FDR=5.97×10^{-5}), supramolecular fibril organisation (GO:0097435, FDR=0.0006), cell proliferation (GO:0008283, FDR=0.0001) and regulation of cell proliferation (GO:0042127, FDR=0.0007). As cluster 2 is predominantly comprised of healthy control cells, the enrichment of marker genes associated with

cell division, transcription and translation suggests altered proliferative capacity in myoblasts from pre-sarcopenic and sarcopenic individuals.

There were 167 GO terms enriched for the cluster 8 gene markers (FDR < 0.05). Approximately 20% of the GO terms significantly enriched for in cluster 8 were associated with the ribosome, ribosomal biogenesis and translation. These included rRNA processing (GO:0006364, FDR=5.36x10^{-15}), mRNA catabolic processing (GO:0006402, FDR=7.22x10^{-14}), ribosome biogenesis (GO:0042254, FDR=6.14x10^{-13}), translation (GO:0006412, FDR=1.92x10^{-8}), ribosomal large subunit biogenesis (GO:0042273, FDR=0.0224) and ribosomal small subunit biogenesis (GO:0000028, FDR=0.0244). Changes in these pathways suggest that pre-sarcopenia may result due to altered protein synthesis, ultimately potentially leading to sarcopenia and reduced muscle function.

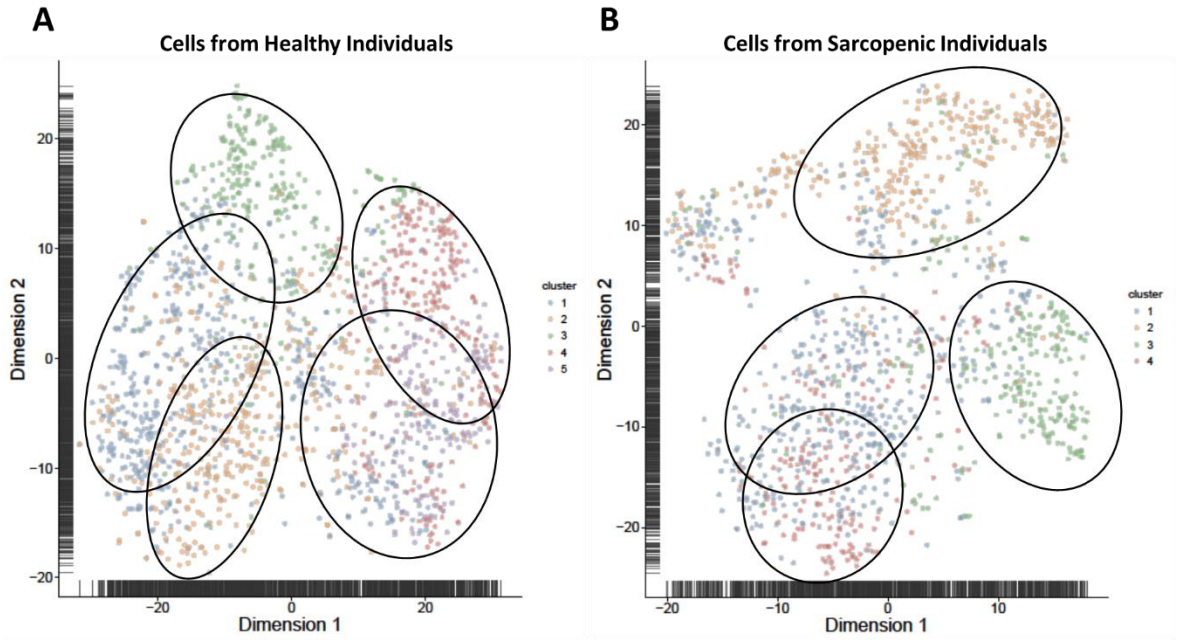


Figure 8.9: t-SNE plots for the cluster analysis of healthy cells and sarcopenic cells individually

(A) t-SNE plot for the cluster analysis of the healthy cells, showing clustering into 11 clusters based on their transcriptomes. There was still some level of overlap between the clusters, suggesting that although distinct clusters, there are some similarities between them, reinforcing the idea of myoblasts along a continuum of proliferative/differentiation capacity. (B) t-SNE plot for the cluster analysis of the sarcopenic cells, showing clustering into 8 clusters based on their transcriptomes. Similar to the healthy cells, there was some level of overlap between the individual clusters. Each point represents an individual cell, coloured based on the cluster they belonged to.

8.2.9. Cell Clustering of Myoblasts from Healthy and Sarcopenic Individuals

Having shown that myoblasts from healthy, pre-sarcopenics and sarcopenic individuals separate into 10 different clusters based on their transcriptomes, we next investigated the clustering of myoblasts from healthy individuals, originating from participants who have no decrease in muscle mass or strength, to determine a baseline clustering pattern of elderly myoblasts. We also examined the clustering of myoblasts from sarcopenic individuals to determine the heterogeneity of myoblasts from sarcopenic muscle.

Table 8.6: *List of the top 20 marker genes for each cell cluster in the analysis of cells from healthy individuals*

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
HSP90AA1	ATP2B1	NAP1L1	ACTC1	ACTC1
NAP1L1	HMGB2	CHRNA1	ATP2B1	HMGB2
RPL31	NAP1L1	MYOD1	HSP90AA1	NAP1L1
TMSB4X	RPL31	TTN	NAP1L1	TMSB4X
FTH1	CHRNA1	H19	FGF2	ATP2B1
IGFBP7	FGF2	MT2A	IGFBP7	CHRNA1
MYOD1	KIF20B	TNNT1	TTN	FTH1
RPS18	RPS18	ACTC1	FTH1	KIF20B
MT2A	GREM1	FGF2	GREM1	H19
RPLP1	MYOD1	IGFBP7	NES	NES
CENPF	RPLP1	RPL31	TNNT1	RPL31
ANXA2	TOP2A	VASH2	CDH15	TOP2A
RPS2	CDH15	KLHL41	CENPF	NUF2
VASH2	FTH1	NES	H19	RPS18
CCDC88A	NUF2	RPL24	MT2A	ANXA2
NDUFA4	ACTC1	PTRF	CHRNA1	MYOD1
RPS23	H19	RP11-47I22.3	RPL31	NDC80
TTN	NDC80	CD44	TMSB4X	CCNA2
RPS13	RPS2	LRRC17	CCDC88A	LRRC17
RPS15A	CCNA2	TXNRD1	LRRC17	NDUFA4

8.2.9.1. Myoblasts from Healthy Individuals

Analysing only the cells that originated from control healthy participants, 5 subpopulations of cells were identified (figure **8.9a**). The existence of multiple subpopulations of healthy control cells is similar to previous analyses which included both controls, pre-sarcopenics and sarcopenics, where the control cells were found to be spread across all the clusters, with varying numbers of cells in each cluster. The top 20 marker genes for each cluster of the healthy cells can be seen in table **8.6**. To determine whether there were any functional differences between the different clusters, we examined the pathways that the marker genes were enriched for. Clusters 1 and 2 had genes

enriched in pathways associated with ribosomal biogenesis, ribosomal function and protein translation (e.g. GO:0042254 FDR=7.06x10^{-14}, GO:0006364 FDR=1.04x10^{-10} and GO:0006412 FDR=4.00x10^{-5}). Several of the clusters had marker genes enriched in pathways associated with mitosis (e.g. GO:0007093 and GO:0140014) and cell proliferation (e.g. GO:0008283). Clusters 3 and 4 had marker genes enriched in pathways associated with skeletal muscle development and functions, including muscle organ development (GO:0007517, FDR=3.44x10^{-6}), striated muscle tissue development (GO:0014706, FDR=1.19x10^{-5}) and actin-myosin filament sliding (GO:0033275, FDR=6.33x10^{-7}). Several clusters also consisted of marker genes associated with protein targeting to the correct cellular localization and cellular organization (e.g. GO:0045047, GO:0006612 and GO:0033365). The enrichment of marker genes of these clusters suggest that myoblasts from healthy muscle exist along a continuum, with a subset of myoblasts actively proliferating, while other myoblasts may have started along a myogenic commitment, and started to express genes associated with myotube differentiation and function.

Table 8.7: *List of the top 20 marker genes for each cell cluster in the analysis of cells from sarcopenic individuals*

Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**
IGFBP7	IGFBP7	NAP1L1	IGFBP7
NES	NAP1L1	NES	NES
PPIB	PPIB	IGFBP7	MFAP5
FTH1	FTH1	VAMP5	SPARC
MFAP5	SPARC	H19	H19
NAP1L1	VAMP5	ACTC1	ITGBL1
H19	H19	ATP2B1	PPIB
ITGBL1	NUPR1	CHRNA1	ATP2B1
NUPR1	ACTC1	P4HA2	CTGF
ACTC1	PTMA	HMGA1	PTMA
CTGF	FTL	DES	NAP1L1
SPARC	P4HA2	LRRC17	SERF2
CHRNA1	SERF2	LOX	TIMP1
FTL	HMGA1	NDUFA4	CHRNA1
TIMP1	HSP90AA1	TXNRD1	OST4
ATP2B1	OST4	C4orf3	ACTC1
HSP90AA1	LRRC17	PDLIM3	DST
DES	SELM	PTTG1IP	SELM
DST	IGFBP3	COL4A1	COL8A1
COL8A1	NDUFA4	DAB2	LOX

8.2.9.2. Myoblasts from Sarcopenic Individuals

We next investigated the level of cellular heterogeneity in myoblasts from sarcopenic individuals alone. Four different subpopulations of cells were identified based on their transcriptomes (figure

8.9b). The top 20 marker genes for each cluster are shown in table **8.7**. Three of the four different clusters had marker genes enriched in pathways associated with response to glucocorticoid and corticosteroid hormones (e.g. GO:0031960, GO:0051384 and GO:0071384). This suggests a stress response may be activated in the myoblasts from sarcopenic muscle, mediated by the glucocorticoids and corticosteroids. All four clusters had marker genes associated with pathways involved in skeletal muscle development and function, including supramolecular fibre organisation (GO:0097435, FDR=0.0247), muscle structure development (GO:0061061, FDR=7.19×10^{-5}), striated muscle development (GO:0014706, FDR=0.0053) and regulation of myoblast differentiation (GO:0045661, FDR=0.0038). This suggests that these cells have a reduced capacity to differentiate into myotubes and participate in the repair and regeneration of sarcopenic muscle in response to injury. Three of the four clusters also had marker genes associated with cell proliferation and cell death pathways, including regulation of cell proliferation (GO:0042127, FDR=0.0038), regulation of myoblast proliferation (GO:2000291, FDR=0.0135), positive regulation of cell death (GO:0010942, FDR=0.0432) and apoptotic process (GO:0006915, FDR=0.0383). This suggests a possible change in the proliferative capacity of myoblasts in sarcopenia. This data shows that myoblasts in skeletal muscle from those with sarcopenia are not functionally homogenous, with a heterogeneity resulting in a subset of cells with a potential reduced proliferative capacity, along with another subset with reduced differentiation capacity.

8.3. Discussion

In this study, we found significant cellular heterogeneity in human skeletal muscle myoblasts. Of all the cells in the dataset, irrespective of sarcopenia status, we found that the cells separated into 11 separate clusters, each with a distinct transcriptional profile able to distinguish the cell subpopulation from others. Several of these cell clusters appear enriched with cells that originated from subjects that classed as control, pre-sarcopenic and sarcopenic according to the EWGSOP definition of sarcopenia [113], with evident differences in the pathways that were changed in each cluster, suggesting functional differences between the clusters.

8.3.1. Cellular Heterogeneity in Human Skeletal Muscle Myoblasts

It is well reported in the literature that there is cellular heterogeneity of myoblasts from mouse skeletal muscle. This heterogeneity confers different functional characteristics to the myoblasts, with differing differentiation potential and proliferation capabilities. Rocheteau et al. [485] have shown a population of satellite cells that express Pax7 at a high level, with a low metabolic state and which are less primed for myogenic commitment. A decrease in this subpopulation of cells in skeletal muscle may suggest a reduced regenerative capacity of the tissue, due to a reduction in

the pool of cells capable of proliferating and committing to a myogenic lineage. This is similar to the results found by Pavlath et al. [491], who found that skeletal muscle myoblasts have varying regenerative capabilities, resulting in impaired repair to skeletal muscle following exogenous trauma. Several other studies have also used mouse models to show that satellite cells in muscle consist of different subsets of cells, some of which are committed progenitors prone to myogenic differentiation, as well as a distinct population capable of self-renewal to maintain the pool of satellite cells [486,492]. These intrinsic differences between the myoblasts suggests that a loss of myoblasts with high regenerative capacities in the elderly may result in impaired muscle repair due to the predominant presence of myoblasts with a low regenerative capacity. Therefore, it may be of interest to separate cells with high and low regenerative capacities, and determine the intrinsic differences between the two populations of cells, and how they contribute to skeletal muscle ageing.

In this study, we found that myoblasts from healthy, pre-sarcopenic and sarcopenic muscle tissue were heterogeneous and showed differences in their global transcriptomic profiles. The fact that the myoblasts from those with sarcopenia differed from myoblasts from healthy individuals suggests that intrinsic changes in the myoblasts may contribute to changes in muscle mass and function associated with sarcopenia and ageing. The marker genes for the subpopulation of cells that is comprised predominantly of sarcopenic cells are enriched in transcriptomic pathway associated with normal muscle function as well as pathways associated with DNA and RNA breakdown. The level of DNA damage is greater in the elderly, and DNA damage can induce cells to transition into a senescent state. This suggests that myoblasts from sarcopenic individuals may show increased senescence compared to myoblasts from healthy people. These differences in the myoblasts between healthy and sarcopenic participants suggest an impaired capacity for muscle regeneration in response to external damage. A similar pattern of gene expression and pathway enrichment is seen in the myoblasts from the subpopulation of cells that is comprised predominantly of pre-sarcopenic cells. When looking at the subpopulation of cells comprised mainly of control healthy cells, we saw that many of the marker genes were enriched in pathways associated with cell division and cellular proliferation. This suggests that myoblasts from healthy individuals show increased levels of cellular proliferation compared to myoblasts from pre-sarcopenic and sarcopenics, which is key for maintenance of the pool of proliferative cells in skeletal muscle.

8.3.2. Consequences of Myoblast Heterogeneity in Sarcopenia

Cellular heterogeneity of cells can confer different functional capabilities to the tissue of origin of the cells. Heterogeneity of cells has been shown to have different functional capabilities in several

tissues. Two populations of stem cells in the skin with distinct molecular and functional properties play a role in maintaining homeostasis within different epidermal regions [493,494]. In cancer, substantial intra-tumour heterogeneity is observed, with different subpopulations of tumour cells with differing tumourigenicity, anti-tumour immunity and capabilities for the production of secreted factors and metastases [495]. Therefore, the heterogeneity of cells within the same tissue can affect the function of the tissue, depending on the functional differences between the subpopulations of cells.

Here, we found that myoblasts from sarcopenic individuals showed extensive heterogeneity based on their global transcriptomes. Although the myoblasts from sarcopenic individuals showed a similar level of heterogeneity compared to healthy control myoblasts, the marker genes associated with each sarcopenic subpopulation of cells suggest that these subpopulations play different roles compared to the myoblasts from healthy individuals, with altered function. In the total analysis of all the cells, although there was one cluster formed predominantly of sarcopenic cells, the rest of the cells were divided among the clusters with cells from pre-sarcopenics and healthy controls. To examine this further, we looked at just the sarcopenic cells, and found that they indeed did separate into eight distinct subpopulations, with marker genes enriched in different pathways. The most common enriched pathway was the response to glucocorticoids (GCs), including cortisol (CORT). GCs are stress hormones released by the hypothalamic-pituitary-adrenal (HPA) axis in response to different stressors. GCs, including cortisol, have been shown to negatively affect muscle function. GCs induce skeletal muscle atrophy [496] by increasing the rate of protein catabolism [497] and increasing the expression of the muscle-specific ubiquitin ligases Atrogin-1 [157] and MuRF-1 [498]. CORT has been shown to affect muscle mitochondrial function and activity [499]. Therefore, the enrichment of genes in pathways associated with response to GCs suggest that sarcopenic cells are under cellular stress and may be prone to increased protein catabolism and atrophy as well as showing decreased mitochondrial function. Mitochondrial dysfunction is a key mechanism thought to contribute to the pathogenesis of sarcopenia [146], and mitochondrial dysfunction in the myoblasts may affect their differentiation and the regeneration of muscle. Some of the sarcopenic clusters also showed reduced muscle function and genes in pathways associated with muscle contraction, suggesting that there may be a population of myoblasts in sarcopenic muscle that may be further along a myogenic differentiation lineage, potentially in response to muscle injury with the myoblasts required to aid in repair and regeneration. Some of the sarcopenic cell clusters also showed reduced cell proliferation, suggesting that these cells may not be able to self-renew to replenish the pool of myoblasts in muscle, resulting in a depletion of these cells, another key mechanism thought to contribute to the pathogenesis of sarcopenia [149].

The myoblasts from sarcopenic individuals appeared functionally different from myoblasts from healthy individuals, which although we found they separated into eleven subpopulations based on their transcriptomes, all showed marker genes associated with different aspects of cell proliferation, differentiation and muscle function. This is suggestive of myoblasts from healthy individuals being along a continuum, with some myoblasts showing a higher proliferative capacity and low differentiation potential, another subset of myoblasts having a high differentiation potential and low proliferation capacity, while some myoblasts may have begun differentiating and are expressing genes vital for muscle structure and muscle contraction. All these suggest that myoblasts from healthy individuals play a role in normal muscle structure and function, regardless of where along the continuum they may exist.

8.3.3. Further Considerations

Although here we found that different populations of myoblasts exist in human skeletal muscle, further work is required to determine how functionally different the subpopulations of cells are and how the different cell populations may affect muscle function. It may be interesting to determine whether there are differences in protein levels for some of the marker genes of the different populations of cells, to see whether the transcriptomic differences between subsets of myoblasts also correspond to differences in protein levels and as an extension, cellular function. It may also be of interest to determine whether the proportion of cells in the different functional clusters of cells is associated with measures of muscle function and may explain the different in muscle strength and function.

8.3.4. Conclusions

In this study, we have shown for the first time on a large scale that human myoblasts show intrinsic differences in their global transcriptome profiles. These differences led to a number of subpopulations of myoblasts emerging, which may have distinct functional characteristics. We found that myoblasts from healthy control, pre-sarcopenics and sarcopenics showed differences at the single cell level, with sarcopenic and pre-sarcopenic cells showing changes in pathways key to normal muscle function, whereas healthy cells were enriched for genes key to cellular proliferation and differentiation. We also found that myoblasts from sarcopenic individuals were heterogeneous on their own. Investigating the different populations of myoblasts in sarcopenia may provide a better understanding to how satellite cells and myoblasts contribute to the pathogenesis of sarcopenia.

Chapter 9 –

Discussion

Chapter 9 – Discussion

Sarcopenia is a multi-factorial disease affecting the older population. As life expectancy increases, age-related conditions such as sarcopenia and its co-morbidities become more prevalent, affecting a greater number of people. As sarcopenia is associated with a loss of muscle mass and function, the ability to carry out normal everyday activities is reduced, resulting in decreased independence and increased reliance on others for help. Sarcopenia places a substantial burden on public health care systems, with the only published study showing direct costs of sarcopenia costing $18.5 billion in 2000, representing 1.5% of total health care expenditure in the United States, associated with hospitalization, nursing home care and home healthcare [500]. This however is likely to be much higher today, with the increased population of >65 year olds. To date, however the underlying causes of sarcopenia are largely unknown. The understanding of the molecular and epigenetic mechanisms contributing to the pathogenesis of sarcopenia and the factors that contribute to the variability in sarcopenia diagnosis amongst the elderly would allow for the identification of targeted treatment and management strategies that help older people maintain their independence. In this study, we have examined the epigenetic and transcriptional changes associated with sarcopenia and measures of muscle mass and function, and characterised an *in vitro* model to further investigate muscle ageing and in which interventions for the treatment of sarcopenia can be tested for efficacy.

9.1. Decreased Regenerative Capacity May Contribute to Decreased Skeletal Muscle Repair

Skeletal muscle is a very dynamic tissue, constantly undergoing changes in response to external stimuli as a result of normal muscle function and muscle loading. Normal muscle growth and repair requires muscle progenitor satellite cells, which reside between the sarcolemma and the basal lamina of skeletal muscle. When activated, these satellite cells can proliferate asymmetrically into myoblasts and another satellite cell to renew the resident pool, differentiating and fusing together to form myotubes, aiding in muscle repair and growth. These satellite cells normally exist in a quiescent state, becoming proliferative and differentiating upon stimulation. Quiescence is a reversible state of cell cycle arrest. On the other hand, cellular senescence is an irreversible state of cell cycle arrest, the levels of which increase with age, due to multiple factors.

9.1.1. Cellular Senescence

A key regulator of senescence is CDKN2A ($p16^{INK4a}$) [501]. $p16^{INK4a}$ expression has been shown to increase with age in multiple tissues, including skeletal muscle, and is associated with decreased regenerative capacities of tissues. In this study, we have found that with increasing chronological age, there is a corresponding increase in expression of $p16^{INK4a}$ levels in both muscle tissue, and isolated myoblasts. There is also a trend for increased level of cellular senescence in the myoblasts as seen by senescence-associated β-gal staining, although further replication of this is required. This suggests that the cultured myoblasts from muscle tissue with increased $p16^{INK4a}$ expression have increased levels of senescence, suggesting that the myoblasts retain the phenotype of the tissue from which they were cultured. We do also see a correlation between $p16^{INK4a}$ expression and senescence-associated β-gal staining in the myoblasts. The increase in $p16^{INK4a}$ levels may be a key driving force inducing the senescence of myoblasts in aged skeletal muscle. As these myoblasts are critical for skeletal muscle growth and regeneration, increased senescence of the myoblasts suggests a decreased regenerative capacity of the muscle. A decrease in the regenerative capacity of skeletal muscle has been reported in sarcopenic muscle in both humans and mice [194,195]. The decrease in p16 expression was accompanied by an increase in the expression of ANRIL with age, a lncRNA which negatively regulates the expression of $p16^{INK4a}$, suggesting that ANRIL may mediate this age-induced expression of $p16^{INK4a}$. Interestingly this was only observed in myoblasts, suggesting that ANRIL-mediated regulation of $p16^{INK4a}$ expression occurs primarily in the myoblasts, resulting in increased senescence and reduced regenerative capacity of the muscle tissue. This is not observed in the muscle tissue itself, possibly due to the heterogeneous nature of muscle, with other resident cell types contributing to the expression patterns of $p16^{INK4a}$and ANRIL in skeletal muscle.

Senescent cells, as well as having an inability to proliferate and differentiate, exhibit an increased metabolic activity, with an increased secretion of soluble factors including pro-inflammatory cytokines, such as interleukin 6 (IL6) [174]. In this study, the expression of $p16^{INK4a}$ was positively correlated with the expression of IL6 in the myoblasts, suggesting that the myoblasts, which have transitioned from quiescence to senescence, have increased production and secretion of IL6. IL6 is a potent pro-inflammatory cytokine, inducing an inflammatory phenotype, a phenotype that is regularly increased in elderly individuals. This phenomenon may contribute to the chronic low-grade systemic inflammation seen in the aging population. IL6 has previously been shown to induce mitochondrial dysfunction [502,503], while mitochondrial dysfunction has been shown to induce cellular senescence [175]. Interestingly in the myoblasts, mitochondrial dysfunction is also observed, suggesting that the increased senescence associated with age, leads to increased IL6 expression, which in turn may contribute to mitochondrial dysfunction. This suggests an interplay between

inflammation, cellular senescence and mitochondrial dysfunction, resulting in a positive feedback loop causing an increase in skeletal muscle loss and loss of function.

9.1.2. Mitochondrial Dysfunction and ROS

To date many pathways have been implicated in the pathogenesis of sarcopenia, however most of these studies have compared muscle from young versus old participants. Here, we carried out total RNA-seq on 40 age-matched participants, comparing the transcriptomes of sarcopenic and healthy age-matched control muscle. A number of pathways were found to be differentially regulated in sarcopenia, compared to healthy age-matched participants. RNAseq analysis showed that oxidative phosphorylation, mitochondrial biogenesis, mTOR pathways and glycolysis are consistently altered with respect to measures of muscle mass and function. These pathways altered in aged skeletal muscle converge to a common pathway, resulting in mitochondrial dysfunction. Mitochondrial dysfunction is a key process that increases during aging [161]. We have found that changes in the mitochondria and mitochondrial-associated pathways are consistently associated with decreased muscle mass and muscle function. Although individual genes were not significantly changed with respect to the different measures of muscle mass and function, small changes that do not pass the significance threshold in multiple genes contributing to a single pathway can significantly decrease the activity of the associated pathway. Mitochondrial dysfunction can lead to multiple aberrant events, including a build-up of reactive oxygen species (ROS) and a decrease in ATP production [161]. As mitochondria have a highly oxidative environment, they are highly susceptible to ROS, leading to damage to proteins of the respiratory complexes and damage to mitochondrial DNA (mtDNA) [221], leading to the production of mutated mitochondrial proteins. This leads to further mitochondrial dysfunction, resulting in an increasing deterioration of mitochondrial function. As skeletal muscle is highly dependent on energy production, a decrease in mitochondrial function can severely affect normal muscle function, leading to muscle disuse, which itself can lead to disuse atrophy and further muscle loss.

Although changes in mitochondrial pathways were associated with gene expression changes in the muscle tissue, mitochondrial pathways did not associate with the changes in methylation or with the single cell RNAseq results due to different gene sets being used for the pathway analysis. As this is a limitation and as such not allowing comparison of mitochondrial pathway changes between total RNAseq, single cell RNAseq and methylation analysis, replication of the pathway analysis and gene set enrichment using a common gene set database between the different datasets may provide a better insight into common pathway changes across the different ‘omics’. However, we did see the presence of a differentially methylated region (DMR) in a region that has been identified as a promoter-flanking region of the CPT1B gene. The methylation of the CpGs in this region show a higher level of methylation in those with high ALMI. Differential methylation is known to affect

the regulation of gene expression. Despite being contradictory with the canonical view of DNA methylation repressing gene transcription, Maples et al. [504] have previously shown that in skeletal muscle, an increased methylation at CpGs in the promoter of the CPT1B gene was correlated with increased expression of CPT1B, due to inhibition of USF binding, which acts as a repressor for CPT1B expression in skeletal muscle. Therefore, the increased methylation that we find in the DMR at the CPT1B promoter may correspond to an increased expression of CPT1B. As CPT1B catalyses the rate limiting step of mitochondrial FAO, the decreased expression of CPT1B due to decreased DMR methylation in those with low ALMI, may promote a reduction of muscle mass due to inefficient FAO and energy production. Skeletal muscle is a very metabolically active tissue. Mitochondrial fatty acid oxidation (FAO) provides substrates for gluconeogenesis and electrons for the electron transport chain (ETC) for energy production [505]. Therefore, a decrease in gluconeogenesis or energy production via the ETC may affect muscle function and lead to muscle atrophy.

As CPT1B affects electron generation for the ETC, a change in methylation of key regulatory region of the CPT1B gene, and subsequent changes in expression may affect mitochondrial function and respiration. We also found functional changes in cultured myoblasts and myotubes, associated with mitochondrial function. In myotubes from those with low ALMI, we found a decreased production of ATP, reduced maximal respiration and spare capacity, and an increased level of non-mitochondrial respiration. In the myoblasts, there was impaired ATP production, with an increase in non-mitochondrial respiration. This suggests that mitochondrial function is impaired, resulting in decreased ATP production, and energy generation. As skeletal muscle relies on ATP production to function, an impairment in ATP production and subsequent mitochondrial dysfunction can affect muscle mass. A large body of evidence points towards mitochondrial dysfunction playing a role in muscle atrophy and the loss of muscle mass. Muscle mitochondrial play a key role in regulating apoptosis in skeletal muscle. Therefore, mitochondrial dysfunction can leads to an increased apoptosis of cells in skeletal muscle, contributing to the loss of mass.

9.1.3. Increased DNA Damage and Cell Death

One of the pathways changed with respect to ALMI, gait speed, grip strength and sarcopenia was the p53 pathway and DNA damage response pathways, which can induce cellular senescence. Both mtDNA and nuclear DNA are susceptible to DNA damage and increased genomic instability with age.

Under normal circumstances, the cell employs several mechanisms to repair the damage done to the DNA, in order to reduce the effect of DNA mutations and oxidative damage (e.g. base excision repair), a potent inducer of DNA damage. The levels of ROS increase with mitochondrial dysfunction. In sarcopenia and in those with low muscle mass, there is growing evidence of an

increase in mitochondrial dysfunction and a build-up of ROS. In this study, sarcopenia was associated with an increase in p53 expression, as well as increased enrichment in genes associated with p53-mediated pathways and apoptosis. The p53 protein is a key effector of DNA damage [185,186]. DNA damage is sensed by several different proteins (e.g. ATM, CHK2, ATR, etc), resulting in post-translational changes to the p53 protein. These modifications increase the half-life of p53 as well as enhances its ability to bind to DNA promoter sequences and induce gene transcription. p53 induces the transcription of the p21 gene, which in turn inhibits CDK2, resulting in growth arrest and cellular senescence. Single cell transcriptomics of healthy and sarcopenic skeletal muscle myoblasts showed sarcopenic cells are associated with altered DNA ligation, suggestive of increased DNA damage and the requirement for DNA repair. This increase in DNA damage and DNA repair pathways associated with sarcopenia seen in both whole muscle tissue, a heterogeneous tissue, and muscle progenitor cells (MPCs, myoblasts) suggests an increase in DNA damage-mediated cell death, resulting in reduced muscle regenerative capacity, increased muscle atrophy and muscle loss, a pathway that appears to be intrinsic to MPCs. An altered glucocorticoid/cortisol response seen at the transcriptome level of the sarcopenic myoblasts suggests that sarcopenic cells are under stress, not seen in healthy myoblasts. This increased stress may further contribute to increased cell death and reduced regenerative capacity seen in sarcopenia

Therefore, there is preliminary evidence that in sarcopenia, a subset of myoblasts show altered expression of genes associated with DNA damage and DNA repair. An increase in damage to mtDNA and nuclear DNA as a result of mitochondrial dysfunction and other cellular mechanisms may be inducing senescence, resulting in a reduced regenerative capacity of skeletal muscle in sarcopenia. These changes appear to be intrinsic to myogenic cells rather than other cell types present in skeletal muscle tissue.

As well as inducing cellular senescence, p53 also induces apoptosis [402]. Several p53-regulated genes can enhance the secretion of cytochrome *c* into the cytoplasm from the mitochondria. Cytochrome *c* subsequently interacts with APAF1, a p53-regulated gene, to initiate the apoptosis pathway by activating caspase-3. In the muscle tissue RNAseq analysis, we show that with increased muscle mass, there is a decrease in apoptosis pathways. This suggests that apoptosis is one mechanism by which skeletal muscle is lost with age, resulting in decreased muscle function. Multiple different upstream factors may contribute to the increase in apoptosis; however, our results suggest an increase in mitochondrial dysfunction and increase DNA damage as potential inducers of apoptosis in elderly and sarcopenic muscle.

9.1.4. Altered Myoblast Proliferation and Differentiation in Aged Skeletal Muscle

Single cell sequencing of myoblasts identified intrinsic changes in the transcriptomes of sarcopenic myoblasts that may underpin the reduced proliferative capacity. A change in the proliferative capacity and differentiation of myoblasts and satellite cells in skeletal muscle can negatively impact on the regenerative capacity of muscle tissue [194,456]. A change in the proliferation of myoblasts may result in a decreased pool of resident MPCs. Decreased proliferation may be due to an increased level of senescence. However, intrinsic changes in the myoblasts may also contribute to a reduced proliferative capacity of myoblasts in those with a low muscle mass. Genes that were differentially expressed between myoblasts were enriched in pathways associated with cellular proliferation and the regulation of the cell cycle. This suggests that changes in the global transcriptome of sarcopenic muscle may drive the reduced proliferation of the myoblasts. This suggests that in sarcopenia, myoblasts are less able to proliferate and renew the pool of progenitor cells in the skeletal muscle.

The ability of myoblasts to differentiate into myotubes and fuse to form mature myofibres is a key aspect of skeletal muscle regeneration. A reduction in the differentiation potential of myoblasts may influence downstream muscle regeneration. We have shown that in those with low muscle function, there is a delay in the initiation of differentiation and fusion of myoblasts into multinucleated myotubes. This suggests a delay in the regeneration of skeletal muscle in response to injury and damage imposed upon the muscle as a result of normal function. This data agrees with the idea that impaired regeneration may be a driving force of muscle ageing and the pathogenesis of sarcopenia in the elderly.

9.2. Decreased Hypertrophy and Increased Atrophy in Aged Skeletal Muscle

We have found that in skeletal muscle tissue, those with low ALMI have increased expression of H19. H19, via miR-675-3p/5p, can regulate skeletal muscle hypertrophy by regulating the post-transcriptional expression of SMAD1/5, members of the bone morphogenetic protein (BMP) family, shown to induce muscle hypertrophy and protect from atrophy *in vivo*. Muscle hypertrophy is the main mechanism of muscle growth in adult skeletal muscle. Muscle hypertrophy requires a balance between protein synthesis and protein catabolism. When protein catabolism overtakes protein synthesis, skeletal muscle begins to atrophy, resulting in a loss of muscle mass. Several mechanisms contribute to muscle hypertrophy and atrophy, all of which are tightly regulated.

9.2.1. Role of H19 Signalling in Regulating Muscle Hypertrophy

Muscle denervation increases with age, and has previously been implicated in the pathogenesis of sarcopenia. As well as decreasing the efficiency of muscle force production, denervation induces

muscle fibre atrophy. Skeletal muscle is one of the few tissues that express the long non-coding RNA (lncRNA) H19 during adulthood. H19 is a maternally imprinted lncRNA, found to play a role in skeletal muscle development. Taking a genome-wide approach to investigate the transcriptome of aged skeletal muscle, we found that the expression of H19 showed a negative association with ALMI, suggesting a role for H19 in regulating muscle mass. As well as being a lncRNA, H19 acts as a miRNA host gene, encoding two miRNAs: miR-675-3p and -5p. Dey et al. [411] have shown that SMAD1/5 genes are targets for miR-675-3p, and in agreement, we found that miR-675-3p and -5p targets and reduces the expression of SMAD1/5 in skeletal muscle. SMAD1/5 are members of the bone morphogenic protein (BMP) family, which have been shown to protect against muscle atrophy, and promote hypertrophy in post-natal muscle tissue [412,428]. SMAD1/5 induce muscle hypertrophy by activating mTOR-dependent anabolic pathways. SMAD1/5 phosphorylation and activation results in the phosphorylation of the S6 ribosomal protein (S6RP), a downstream target of the Akt/mTOR pathway. The Akt/mTOR pathway is a key regulator of muscle hypertrophy, resulting in the activation of $p70^{S6K}$ and the phosphorylation of S6RP and downstream activation of 4E-BP1. This results in increased translation and protein synthesis.

As well as regulating muscle hypertrophy and atrophy via SMAD1/5, H19 may act via other pathways. H19 acts as a molecular sponge for the let-7 family of miRNAs [407]. One of the targets genes of let-7 mediated regulation is *Hmga2*, promoting its degradation [408]. The degradation of *Hmga2* is required to allow the formation of multinucleated myofibres and terminal differentiation of myoblasts. The increase in H19 associated with decreased muscle mass suggests a decreased availability of let-7 and decreased *Hmga2* degradation. As such, this results in the decreased generation of terminally differentiated myofibres and myogenesis, contributing to reduced muscle mass. miR-675-3p has also been shown to regulate the expression IGF1R in cells [413]. We have previously shown that H19, via miR-675-3p, may regulate the expression of IGF1R (unpublished data), suggesting a role for H19 in the regulation of IGF signalling and muscle growth. These genes aid in muscle repair, and the increased H19 associated with low lean mass suggests a decreased ability to repair muscle tissue in response to injury, leading to decreased muscle mass.

9.2.2. Role of HDAC4-Myogenin on Muscle Atrophy

Skeletal muscle atrophy is achieved via several different mechanisms, one of which is the HDAC4-myogenin pathway. The HDAC4-myogenin pathway induces muscle atrophy in response to denervation of muscle fibres [83]. HDAC4 induces the expression of myogenin in denervated muscle, and myogenin subsequently activates the expression of the muscle-specific E3 ubiquitin ligases MuRF1 and atrogin-1. Therefore, HDAC4 is a key protein in regulating denervation-induced skeletal muscle atrophy. SMAD1/5 activity inhibits HDAC4 resulting in reduced myogenin expression and

denveration-induced atrophy. However, the mechanisms by which SMAD1/5 regulate the expression and activity of HDAC4 is unknown. One potential mechanism of the regulation of HDAC4 expression is DNA methylation. We found differential methylation at a CpG found in the intronic region of the HDAC4 gene. Those with increased ALMI were found to show a reduced methylation at this region. Although the canonical view is such that DNA methylation represses gene expression, many others have found that DNA methylation at intronic regions of genes regulates intronic enhancers and alternative promoters [320,506], activating gene expression. This suggests that the decreased methylation in the intron of HDAC4 in those with high ALMI may result in decreased expression of HDAC4. This has further downstream consequences of reduced myogenin expression and expression of the muscle-specific E3 ubiquitin ligases. Therefore, one method by which protection from atrophy in skeletal muscle of those with high ALMI may involve DNA methylation-mediated regulation of HDAC4 expression. This however does require experimental validation to determine the effect of intronic methylation of the HDAC4 gene.

As fibre denervation increases with age and is potentially a contributing factor to sarcopenia, an increase in HDAC4-myogenin signalling, together with decreased SMAD1/5 signalling in sarcopenia and decreased muscle mass could result in further fibre atrophy and cell death.

The combined effect of increase cellular apoptosis and fibre atrophy, together with a decrease in myogenesis may result in a significant decrease in muscle mass and the subsequent loss of function associated with sarcopenia.

9.3. Implications for Sarcopenia in the Population

The ever-growing burden of sarcopenia on the general population has resulted in sarcopenia being assigned an ICD-10 code [106], recognising sarcopenia as a disease and highlights the importance of investigation into the mechanisms contributing to and treatments for sarcopenia. There is substantial variation in the loss of muscle mass and function affecting the elderly, with not everyone over the age of 65 being affected by sarcopenia. We have identified potential pathways altered in the myoblasts and muscle tissue from those with sarcopenia and low muscle mass.

H19/miR-675-3p/5p and its downstream pathway was found to be associated with reduced muscle mass in the elderly. This provides a potential biomarker for those at risk of increased loss of muscle mass and function, and developing sarcopenia. Discovery of this pathway in relation to sarcopenia provides a potential route for intervention, with the potential to develop therapeutics that can target aspects of this pathway. Targeting H19 and altering this pathway, preventing H19 signalling via miR-675-3p/5p, may result in de-inhibition of SMAD1/5 expression and muscle hypertrophy during ageing, reducing the loss of muscle mass seen in the elderly. We have also provided a model in which developed molecules against the H19 pathway can be tested to determine whether they

have the potential to reverse the sarcopenic phenotype and restore the proliferative capacity and differentiation potential of sarcopenic myoblasts.

Methylation changes at specific CpG have been used as predicting risk of disease for many diseases, including cancer [507,508], type 2 diabetes [509] and obesity [510,511]. After examining the methylome of elderly skeletal muscle, we have identified CpG marks and DMRs that are associated with reduced muscle mass and may be useful in predicting elderly individuals who are more susceptible to increased loss of muscle mass. Currently, it is unknown as to why there is a great deal of variability between individuals with regards to the amount of muscle loss seen with age. If it were possible to predict who was at a greater risk of developing sarcopenia and increased muscle loss by looking at pre-defined risk markers, it may be possible to intervene to aid in the halting of the loss of muscle mass and prevent further decreases. Although there are currently limited treatment options available for sarcopenia, some interventions have shown evidence for aiding in the reduction of the loss of muscle mass. Exercise in the elderly has shown some evidence to promote muscle hypertrophy in the elderly and halt the progression of sarcopenia, and could be an option for those who are identified as being at increased risk of muscle loss and of developing sarcopenia.

As well as providing potential biomarkers and risk markers for sarcopenia in order to identify those at an increased risk of loss of muscle mass, we have also identified several pathways that are altered in aged skeletal muscle that may contribute to the pathogenesis of sarcopenia. Mitochondrial dysfunction appears to be the major mechanism altered in sarcopenic muscle, altered at both the transcriptional level, with the potential involvement of methylation changes associated with mitochondrial function. The balance between muscle hypertrophy and atrophy also plays a key role in aged skeletal muscle, with altered H19 expression and changes in cell death and apoptosis contributing to the sarcopenic phenotype. Understanding the pathways that are changed in the skeletal muscle of elderly people will allow the development of newer, more efficient therapeutics. It is already known that there is increased mitochondrial dysfunction with age in many different tissues, including skeletal muscle. However, we have shown that between healthy and sarcopenic elderly people, there is a greater increase in mitochondrial dysfunction, at both the transcriptome and functional levels. This may provide a potential explanation for the inter-individual variability of muscle loss, and why sarcopenia affects part of the elderly population and not the whole population. Many different DNA methylation changes have been found to be associated with altered expression of key genes in skeletal muscle (e.g. $p16^{INK4a}$ and H19) and altered muscle mass (e.g. CPT1B and HDAC4). If it were possible to pharmacologically target these DNA methylation marks and reverse them, it may be possible to reverse some of the phenotypic outcomes associated with sarcopenia. H19 may be of particular interest to target, as methylation defects at the imprinting control region of the H19 gene is associated with Beckwith-Wiedemann syndrome [512]

and Silver-Russel syndrome [513], both of which manifest clinically with changes in their musculature among other symptoms. Therefore, understanding the pathways and mechanisms that contribute to muscle loss and sarcopenia provides a route to targeting these pathway pharmacologically in the hope of preventing further muscle loss and the reversal of sarcopenia symptoms.

9.4. Future Directions

Here, we have shown changes in several pathways, as well as transcriptomic, epigenetic and functional changes all associated with regulating muscle mass and in the pathogenesis of sarcopenia in the elderly. However, there are some limitations to the work carried out here.

Firstly, several investigations have utilized small numbers of samples. Further replication and validation of results in larger, independent cohorts would be beneficial to examine whether these effects are seen in the general population. We have begun some replication of results using HSSe to replicate results obtained from HSS, however further work is still required. Secondly, with regards to some of the methylation changes seen in our studies, some of the changes appear to contradict the canonical view of DNA methylation regulation of gene expression. Although it has been shown that increased methylation can activate gene expression in certain circumstances, further investigation into the methylation changes seen here is required to determine how and if the increased DNA methylation activates the expression of the genes of interest. Finally, work carried out in some of the studies here has used samples from only male participants of HSS and HSSe, although some have of the studies here have used samples from both men and women. Although we have shown that some results replicate in both men and women (e.g. H19 association with muscle mass), replication of other results in samples obtained from females would be beneficial to investigate whether the changes associated with reduced muscle mass and sarcopenia hold true for both men and women.

In this study, we have investigated epigenetic, transcriptomic and functional changes in skeletal muscle myoblasts and myotubes and their roles in regulating muscle mass and function. It would be interesting to see how the different 'omics' integrate. This can help determine whether there are epigenetic changes that regulate changes seen at the transcriptomic level, and subsequently whether these changes contribute to the reduced function of the myoblasts and myotubes observed in this study. It is widely reported that changes at the epigenetic level regulate the transcriptome, with DNA methylation changes being important to regulate gene transcription. Changes in the transcription of genes results in altered protein expression of genes, which alters cellular function. Therefore, it would be interesting to see how the different factors in skeletal muscle integrate and work together to regulate skeletal muscle mass and function. It is currently unknown whether the changes in DNA methylation seen in this study are causal, resulting in altered

gene expression and functional consequences, or whether DNA methylation acts as a bystander, providing a biomarker for changes seen at the molecular level. As we found dmCpGs enriched in the promoter regions of EZH2 target genes, it may be possible that changes in EZH2 activity may mediate differential methylation in skeletal muscle, however further work is required to understand the functional significance of this. Understanding the role of DNA methylation in skeletal muscle further will allow for better targeted therapeutics and provide novel biomarkers to identify those with increased risk of muscle loss.

Appendices

Appendix A: Expression of p16 (CDKN2A) and the unspliced ANRIL isoform in the myoblasts and muscle tissue separated by gender

Appendix B: Representative images of FastQC results, before concatenating the 8 fastq files

Appendix C: FastQC results for sample 1, following concatenation, quality and adapter trimming

Appendix D: DEGs associated with Sarcopenia

Appendix E: Top 50 dmCpGs associated with ALMI in Muscle (450K array)

Appendix F: Top 50 dmCpGs associated with Gait in Muscle (450K array)

Appendix G: Top 50 dmCpGs associated with Fasting Glucose in Muscle (450K array)

Appendix: H: Top 50 dmCpGs associated with Grip in Muscle (450K array)

Appendix I: Top 50 dmCpGs associated with Insulin in Muscle (450K array)

Appendix J: Top 50 dmCpGs associated with ALMI in Myoblasts (450K array)

Appendix K: Top 50 dmCpGs associated with Gait in Myoblasts (450K array)

Appendix L: Top 50 dmCpGs associated with Fasting Glucose in Myoblasts (450K array)

Appendix M: Top 50 dmCpGs associated with Grip in Myoblasts (450K array)

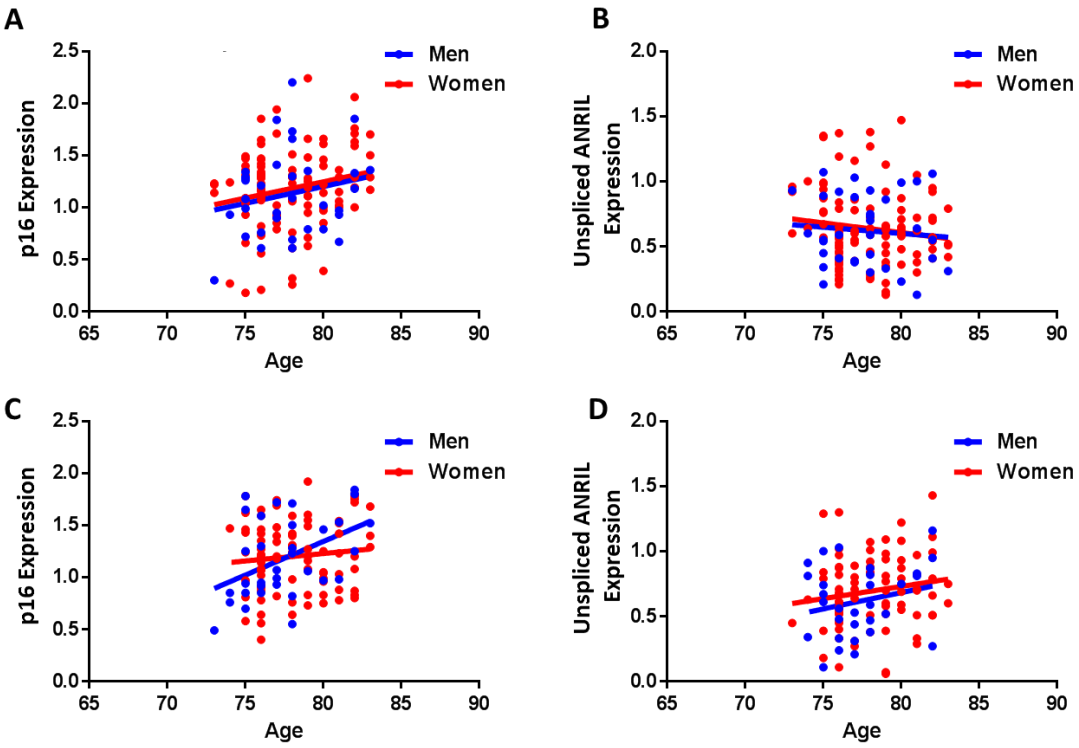
Appendix N: Top 50 dmCpGs associated with Insulin in Myoblasts (450K array)

Appendix O: dmCpGs associated with ALMI (850K array)

Appendix P: dmCpGs associated with Age (850K array)

Appendix Q: dmCpGs associated with fasting serum insulin levels (850K array)

Appendix A

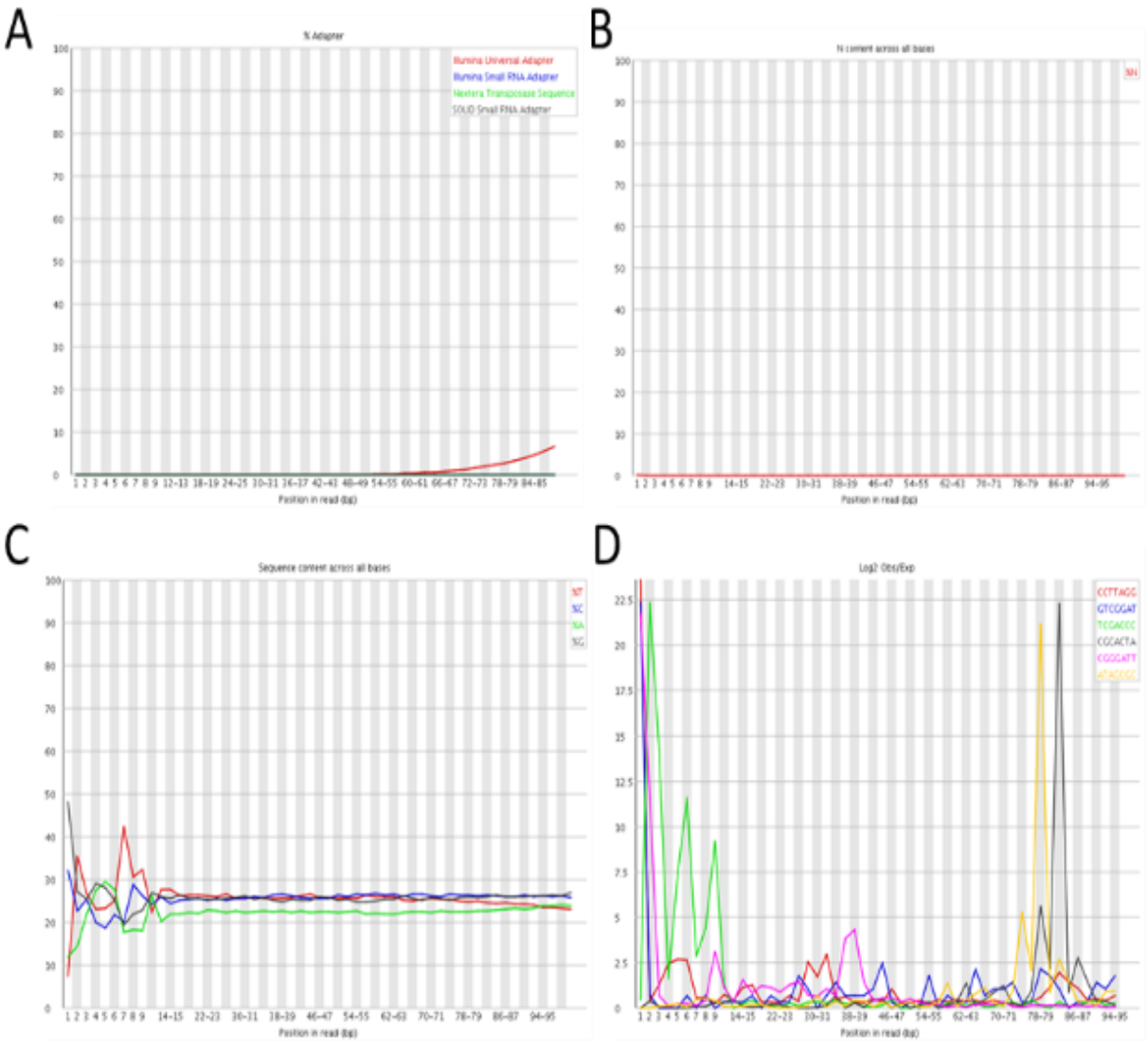


Expression of p16 (CDKN2A) and the unspliced ANRIL isoform in the myoblasts and muscle tissue separated by gender. Expression of p16 (a and c) and ANRIL (b and d) was not associated with gender in either the myoblasts (a and b) or muscle tissue (c and d) (slopes were all significantly different, $p > 0.05$). p16 and ANRIL was significantly associated with age in both men and women and in both the myoblast and muscle samples ($P < 0.05$), except for p16 expression in women in the muscle samples ($p = 0.334$).

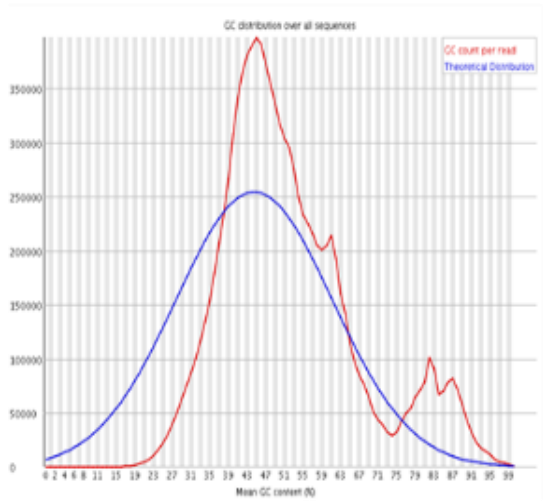
Appendix B

Representative images of FastQC results, before concatenating the 8 fastq files

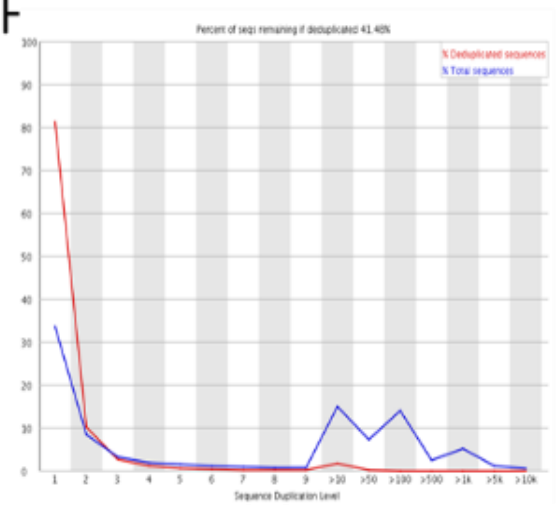
A-J correspond to the forward reads. A = adapter content across the read (red = Illumina Universal Adapter); B = N content across the read; C = per base sequence content (red = %G, blue = %A, green = %T, black = %C); D = overrepresented Kmers; E = per sequence GC content (red = GC count across the read, blue = theoretical distribution of GC); F = Duplicate sequence level (red = % Duplicated sequence, blue = % Total sequences); G = per sequence quality scores; H = Sequence length distribution (midpoint of x-axis = 101bp); I = Per tile sequence quality; J = per base sequence quality, box and whisker plot across the read length (green = very good quality, score > 28, orange = reasonable quality, 20 < score ≤ 28, red = poor quality, score ≤ 20).



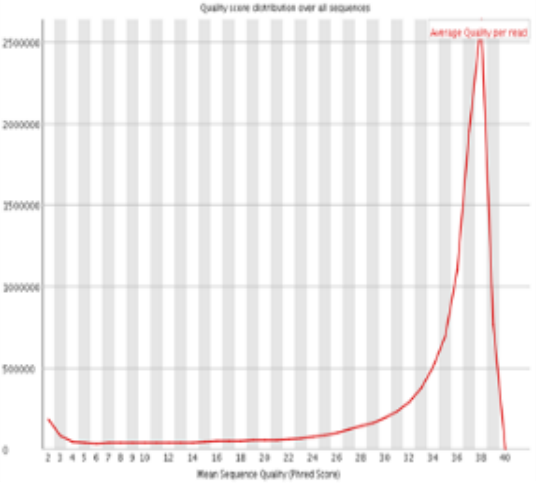
E


GC distribution over all sequences
GC count per read
Theoretical Distribution
Mean GC content (%)

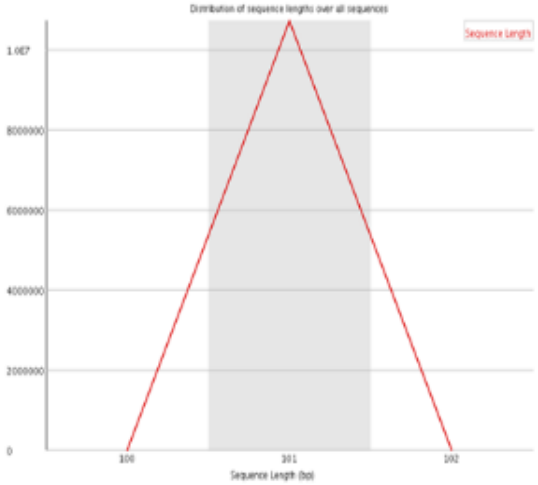
F


Percent of seqs remaining if deduplicated 41.46%
% Deduplicated sequences
% Total sequences
Sequence Duplication Level

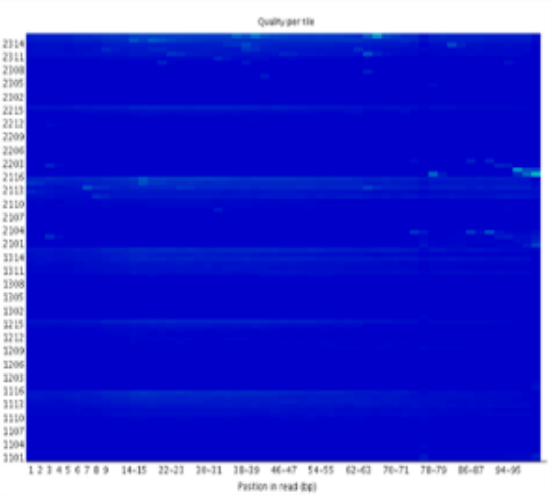
G


Quality score distribution over all sequences
Average Quality per read
Mean Sequence Quality (Phred Score)

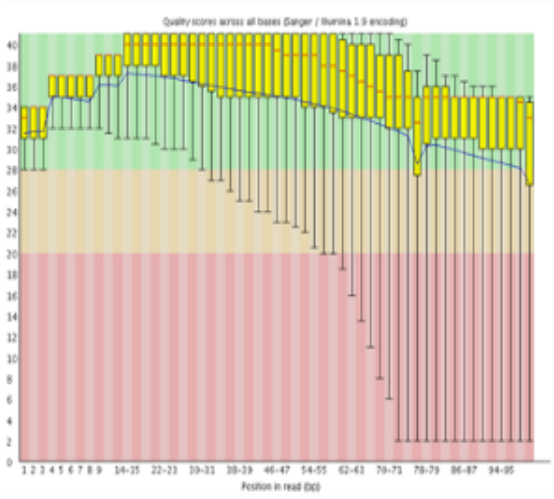
H


Distribution of sequence lengths over all sequences
Sequence Length
Sequence Length (bp)

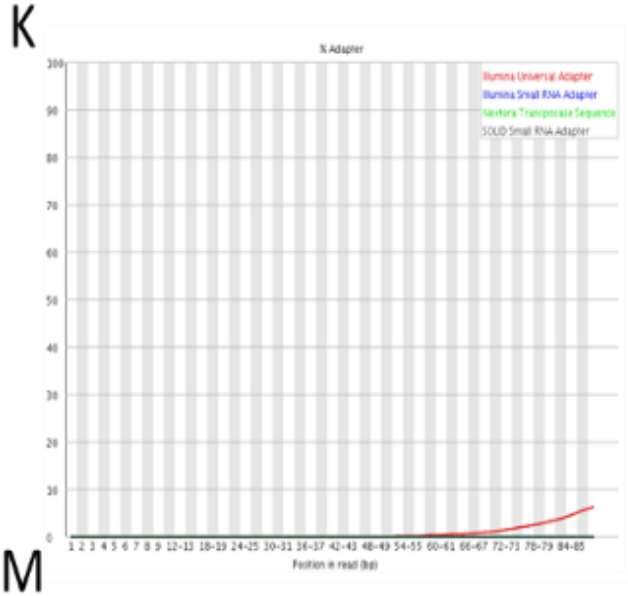
I


Quality per tile
Position in read (bp)

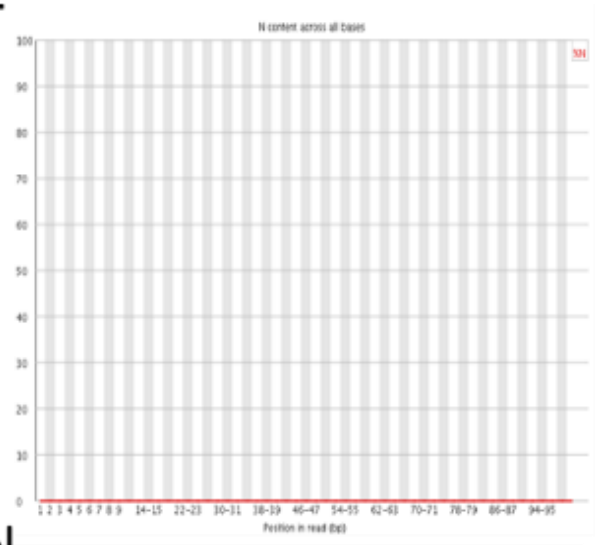
J


Quality scores across all bases (Sanger / Illumina 1.9 encoding)
Position in read (bp)

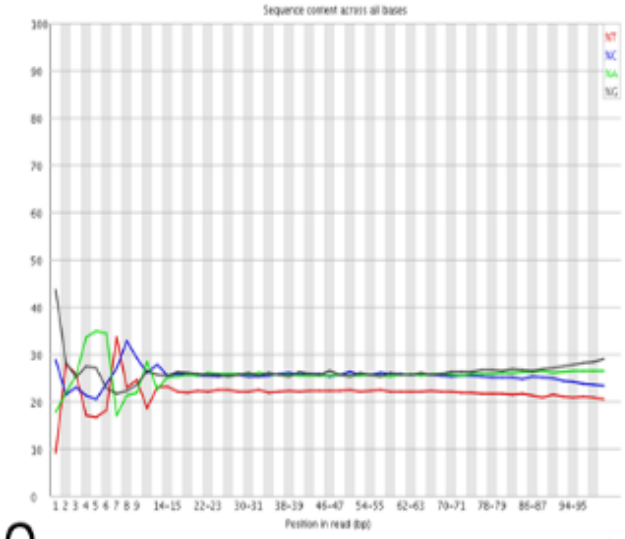
K


% Adapter
Illumina Universal Adapter
Illumina Small RNA Adapter
Nextera Transposase Sequence
SOLID Small RNA Adapter
Position in read (bp)

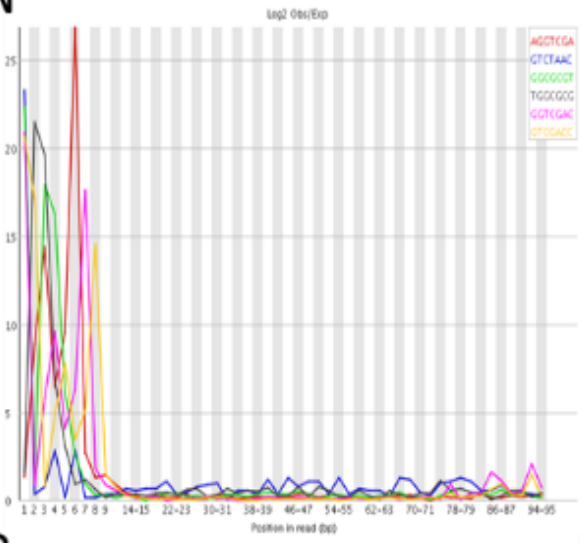
L


N content across all bases
Position in read (bp)

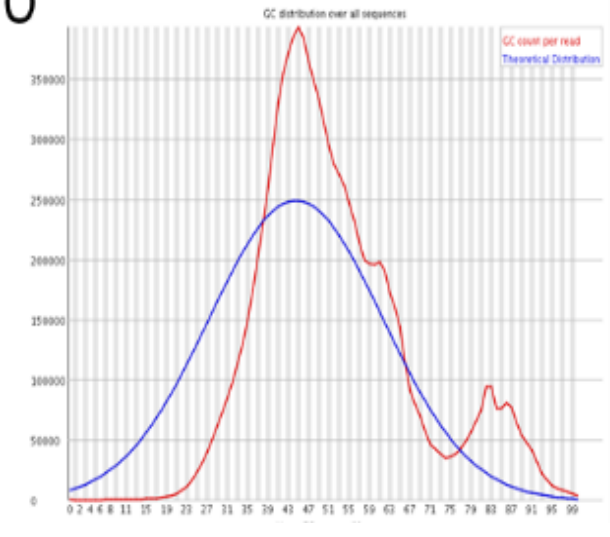
M


Sequence content across all bases
Position in read (bp)

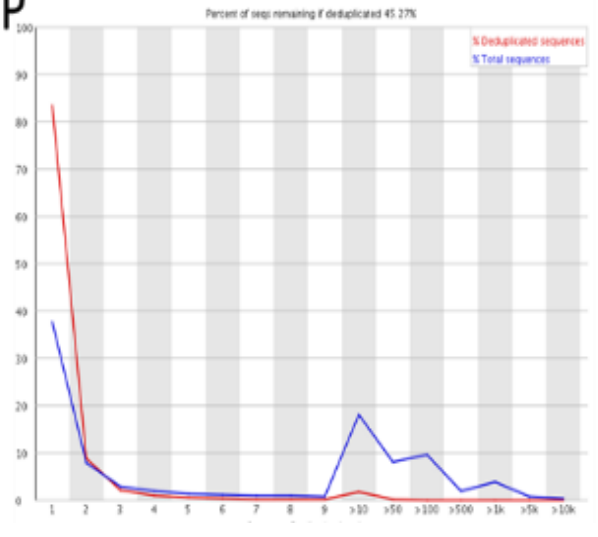
N


Log2 Obs/Exp
Position in read (bp)

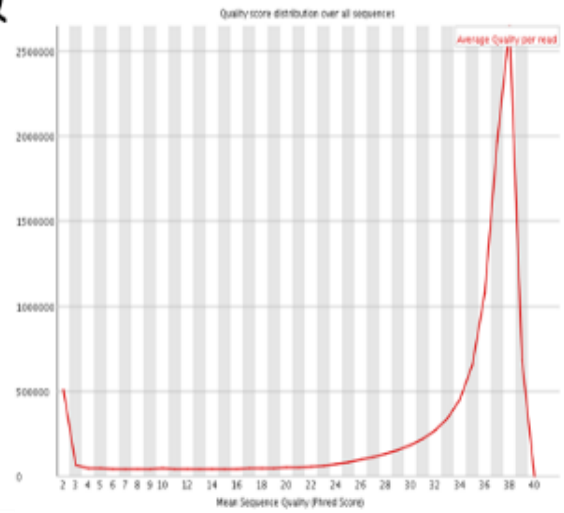
O


GC distribution over all sequences
GC count per read
Theoretical Distribution

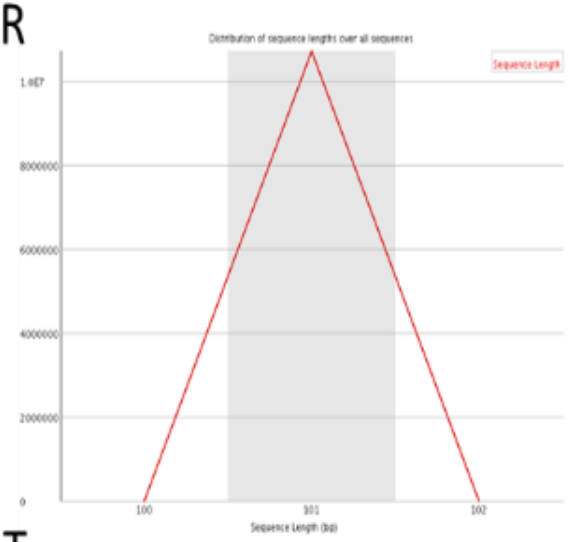
P


Percent of seqs remaining if deduplicated 45.27%
% Deduplicated sequences
% Total sequences

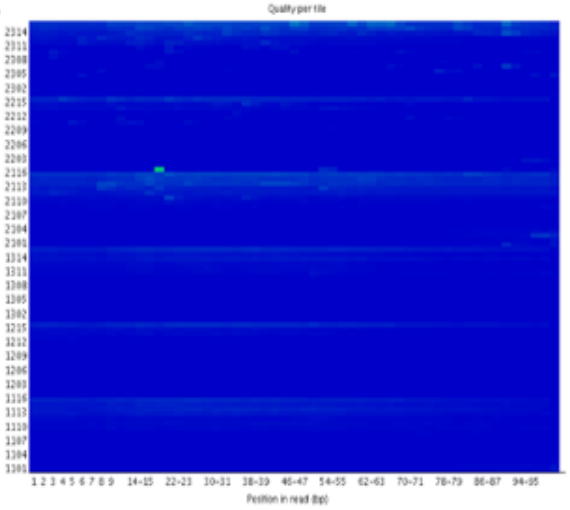
Q


Quality score distribution over all sequences
Average Quality per read
2500000
2000000
1500000
1000000
500000
0
2 3 4 5 6 7 8 9 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Mean Sequence Quality (Phred Score)

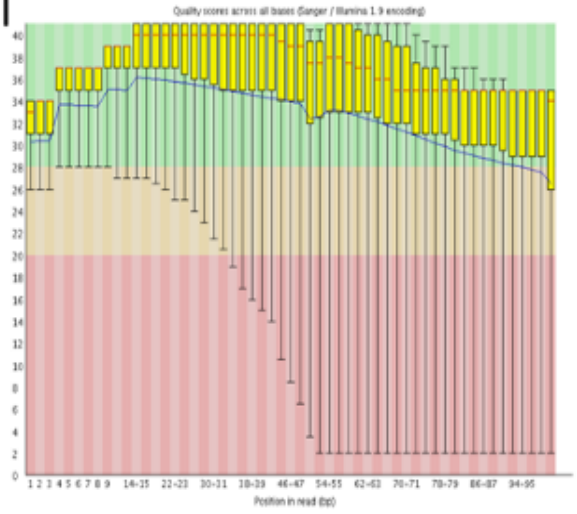
R


Distribution of sequence lengths over all sequences
Sequence Length
1.0E7
8000000
6000000
4000000
2000000
0
100
101
102
Sequence Length (bp)

S


Quality per tile
2314
2311
2308
2305
2302
2215
2212
2209
2206
2203
2116
2113
2110
2107
2104
2101
1314
1311
1308
1305
1302
1215
1212
1209
1206
1203
1116
1113
1110
1107
1104
1101
1 2 3 4 5 6 7 8 9 14-15 22-23 30-31 38-39 46-47 54-55 62-63 70-71 78-79 86-87 94-95
Position in read (bp)

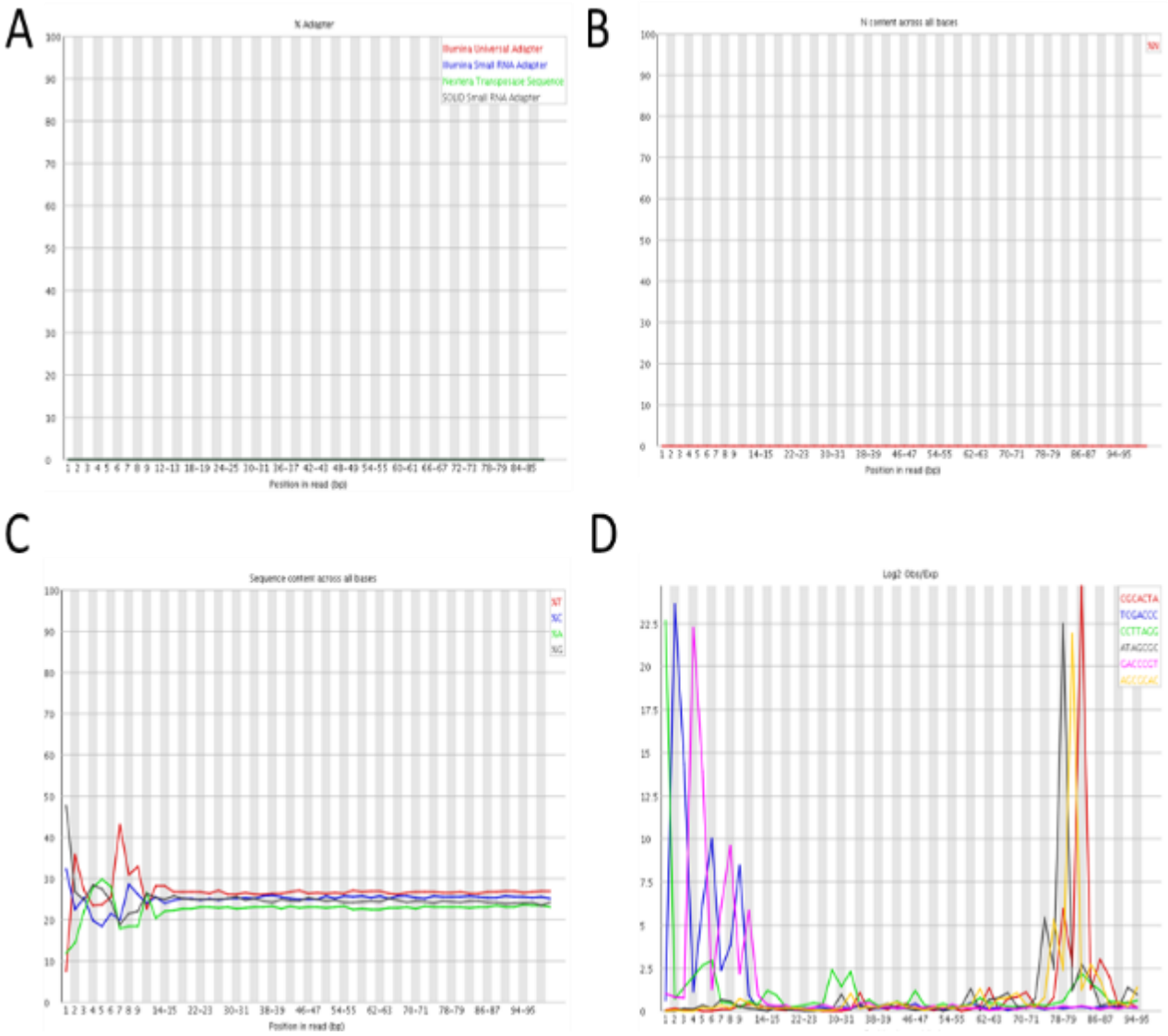
T


Quality scores across all bases (Sanger / Illumina 1.9 encoding)
40
38
36
34
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
1 2 3 4 5 6 7 8 9 14-15 22-23 30-31 38-39 46-47 54-55 62-63 70-71 78-79 86-87 94-95
Position in read (bp)

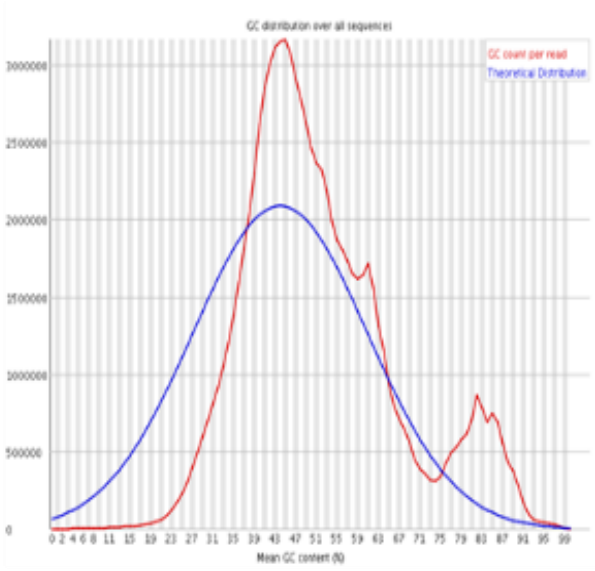
Appendix C

FastQC results for sample 1, following concatenation, quality and adapter trimming

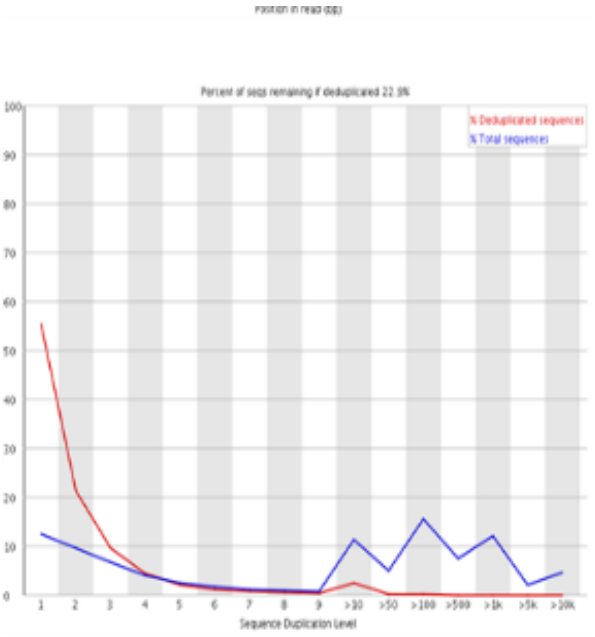
A-J correspond to the forward reads, K-T correspond to the reverse reads. A+K = adapter content across the read (red = Illumina Universal Adapter); B+L = N content across the read; C+M = per base sequence content (red = %G, blue = %A, green = %T, black = %C); D+N = overrepresented Kmers; E+O = per sequence GC content (red = GC count across the read, blue = theoretical distribution of GC); F+P = Duplicate sequence level (red = % Duplicated sequence, blue = % Total sequences); G+Q = per sequence quality scores; H+R = Sequence length distribution (midpoint of x-axis = 101bp); I+S = Per tile sequence quality; J+T = per base sequence quality, box and whisker plot across the read length (green = very good quality, score > 28, orange = reasonable quality, 20 < score ≤ 28, red = poor quality, score ≤ 20).



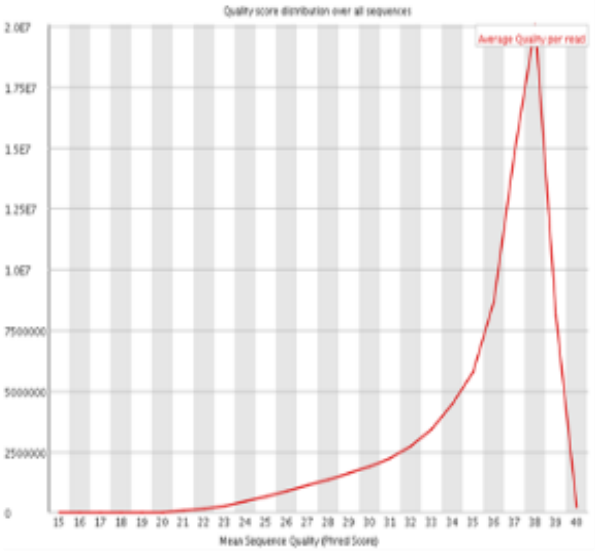
E


GC distribution over all sequences
GC count per read
Theoretical Distribution
Mean GC content (%)

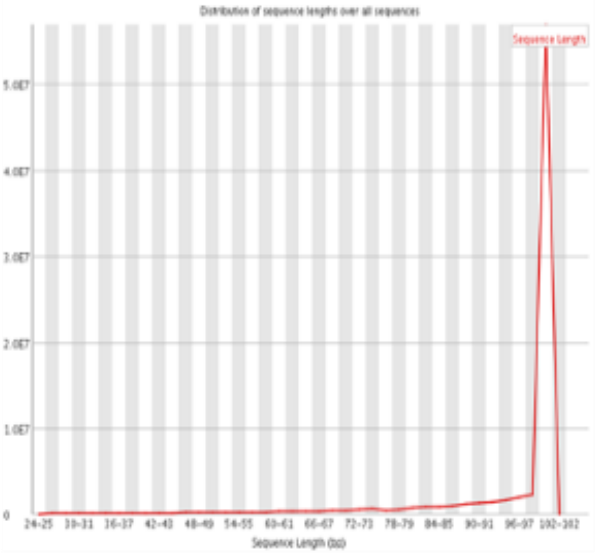
F


Percent of seqs remaining if deduplicated 22.8%
% Deduplicated sequences
% Total sequences
Sequence Duplication Level

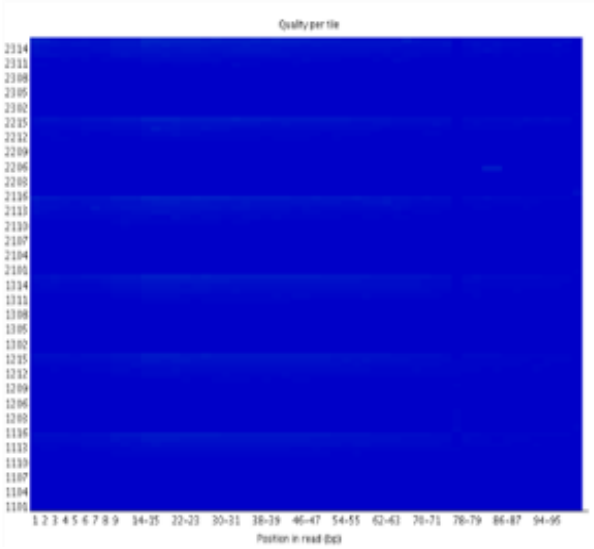
G


Quality score distribution over all sequences
Average Quality per read
Mean Sequence Quality (Phred Score)

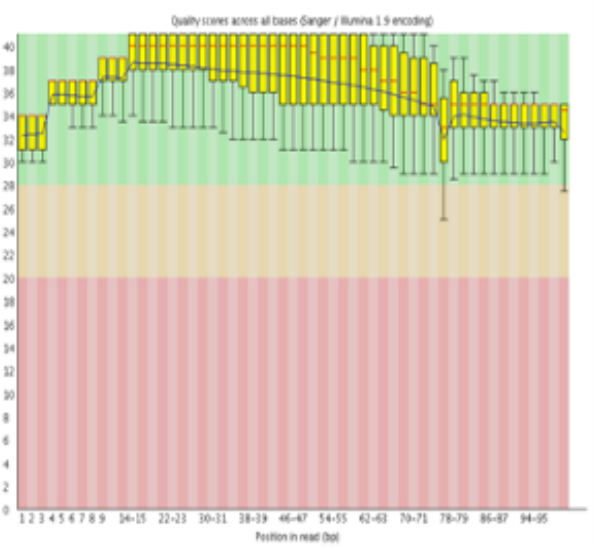
H


Distribution of sequence lengths over all sequences
Sequence Length
Sequence Length (bp)

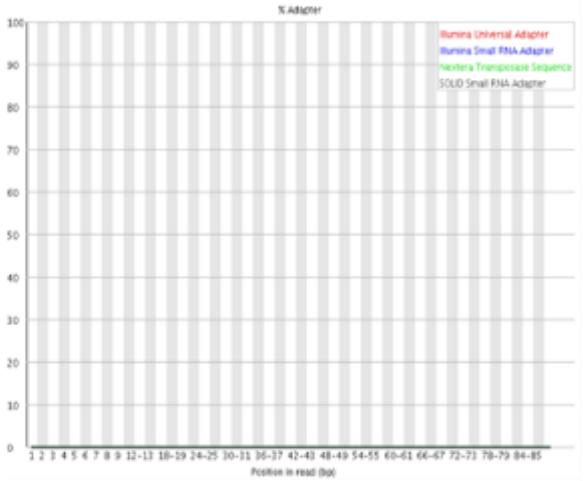
I


Quality per tile
Position in read (bp)

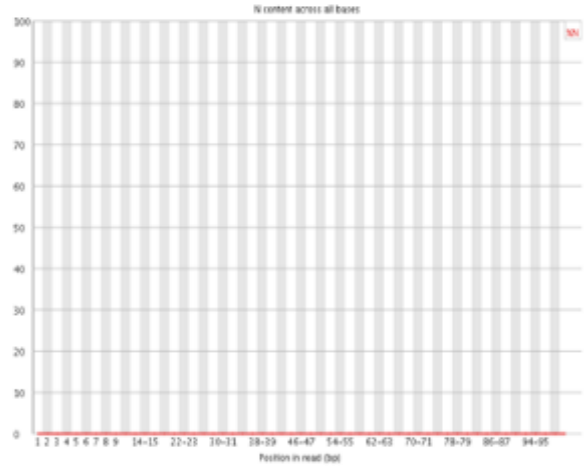
J


Quality scores across all bases (Sanger / Illumina 1.9 encoding)
Position in read (bp)

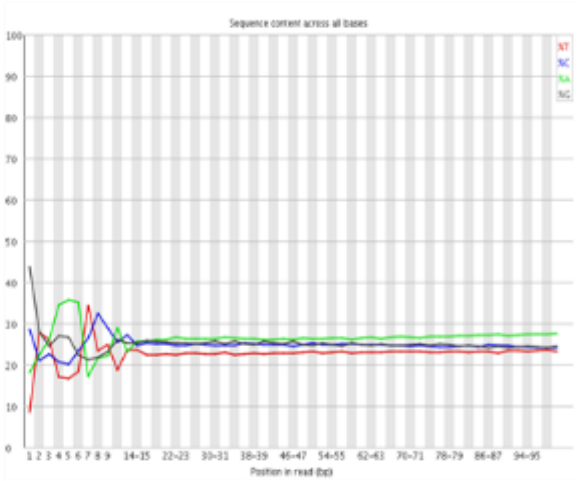
K


% Adapter
Illumina Universal Adapter
Illumina Small RNA Adapter
Nextera Transposase Sequence
SOLID Small RNA Adapter
Position in read (bp)

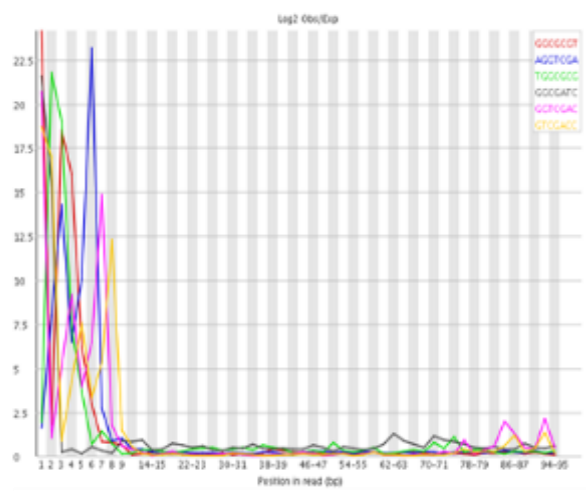
L


N content across all bases
%N
Position in read (bp)

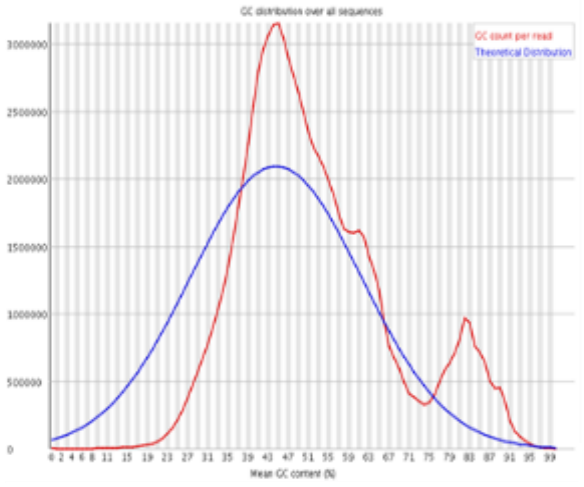
M


Sequence content across all bases
%T
%C
%A
%G
Position in read (bp)

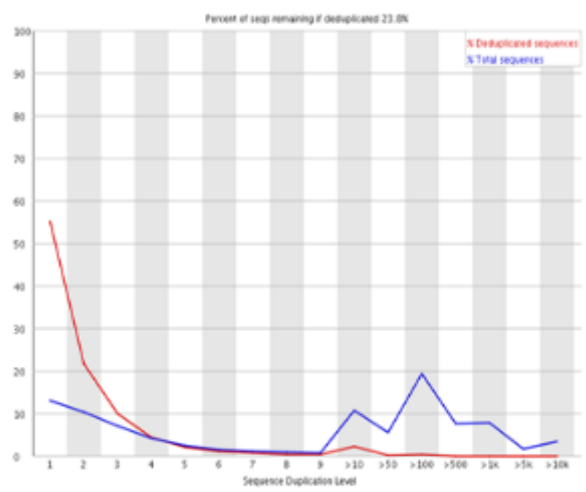
N


Log2 Obs/Exp
Position in read (bp)

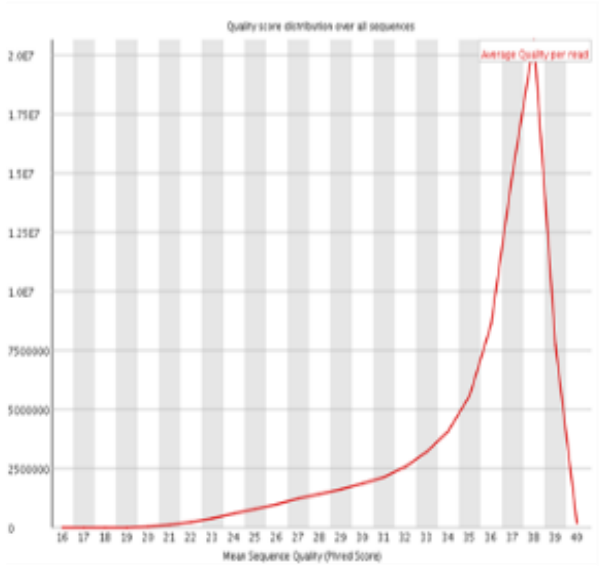
O


GC distribution over all sequences
GC count per read
Theoretical Distribution
Mean GC content (%)

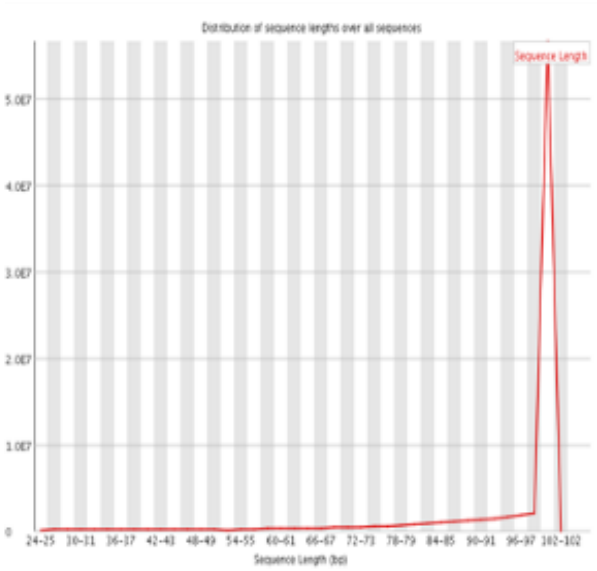
P


Percent of seqs remaining if deduplicated 23.8%
% Deduplicated sequences
% Total sequences
Sequence Duplication Level

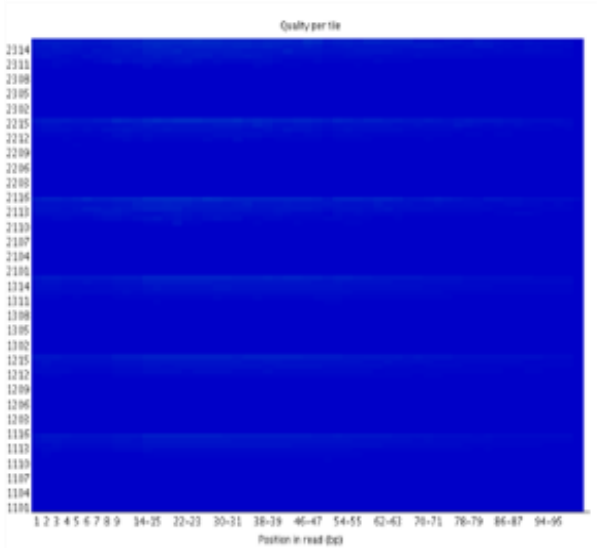
Q


Quality score distribution over all sequences
Average Quality per read
Mean Sequence Quality (Phred Score)

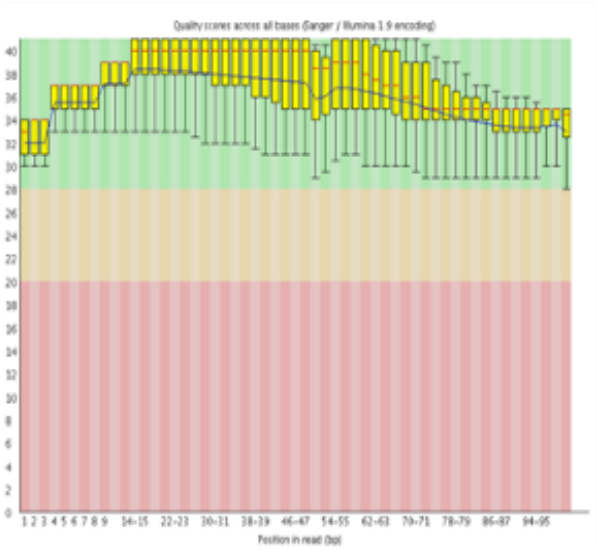
R


Distribution of sequence lengths over all sequences
Sequence Length
Sequence Length (bp)

S


Quality per tile
Position in read (bp)

T


Quality scores across all bases (Sanger / Illumina 1.9 encoding)
Position in read (bp)

Appendix D – DEGs associated with Sarcopenia

Ensembl Gene ID	Base Mean	Log2(FC)	P Value	FDR	Gene Symbol
ENSG00000223749	31.18	1.445	2.49E-12	6.02E-08	MIR503HG
ENSG00000278730	1195.31	-0.341	1.07E-07	0.0013	AC005332.9
ENSG00000127252	722.19	-0.543	1.65E-07	0.0013	HRASLS
ENSG00000116147	3.82	0.870	3.30E-07	0.0014	TNR
ENSG00000143502	35.96	1.068	3.63E-07	0.0014	SUSD4
ENSG00000170624	8459.63	-0.351	3.97E-07	0.0014	SGCD
ENSG00000188517	43.80	1.078	2.35E-07	0.0014	COL25A1
ENSG00000162688	61939.57	-0.582	7.45E-07	0.0021	AGL
ENSG00000228430	9.29	1.028	7.84E-07	0.0021	AL162726.3
ENSG00000188452	37.76	0.896	1.04E-06	0.0025	CERKL
ENSG00000179813	8.42	0.946	1.44E-06	0.0032	FAM216B
ENSG00000114757	20.97	0.951	2.65E-06	0.0049	PEX5L
ENSG00000227060	4.34	0.922	2.49E-06	0.0049	LINC00629
ENSG00000230943	338.19	-0.563	2.84E-06	0.0049	LINC02541
ENSG00000205622	13.83	0.963	3.62E-06	0.0058	AP001043.1
ENSG00000234754	9.57	0.920	4.49E-06	0.0068	C1orf140
ENSG00000196878	202.58	-0.849	4.84E-06	0.0069	LAMB3
ENSG00000164199	29.68	0.878	5.34E-06	0.0072	ADGRV1
ENSG00000082293	323.11	0.855	9.05E-06	0.0115	COL19A1
ENSG00000279141	4.19	0.811	1.02E-05	0.0123	LINC01451
ENSG00000105877	285.42	0.897	1.17E-05	0.0124	DNAH11
ENSG00000182923	1197.97	-0.307	1.19E-05	0.0124	CEP63
ENSG00000225342	40.38	0.858	1.16E-05	0.0124	AC079630.1
ENSG00000228526	43.55	0.860	1.30E-05	0.0130	MIR34AHG
ENSG00000065882	2321.61	-0.818	1.44E-05	0.0139	TBC1D1
ENSG00000067057	250.34	0.797	1.56E-05	0.0145	PFKP
ENSG00000124107	66.92	0.886	1.75E-05	0.0157	SLPI
ENSG00000118898	442.72	0.781	1.87E-05	0.0161	PPL
ENSG00000266028	301.89	0.515	1.93E-05	0.0161	SRGAP2
ENSG00000181026	170.65	0.570	2.32E-05	0.0165	AEN
ENSG00000188906	1830.60	0.744	2.16E-05	0.0165	LRRK2
ENSG00000222345	4.58	0.898	2.18E-05	0.0165	SNORD19
ENSG00000234546	3.85	0.866	2.29E-05	0.0165	LINC01759
ENSG00000246575	4.73	0.705	2.28E-05	0.0165	AC093162.2
ENSG00000048052	1970.85	-0.578	2.59E-05	0.0178	HDAC9
ENSG00000248690	11.08	0.871	2.82E-05	0.0189	HAS2-AS1
ENSG00000134569	953.55	0.624	3.00E-05	0.0195	LRP4
ENSG00000225733	4868.82	-0.235	3.08E-05	0.0195	FGD5-AS1
ENSG00000149634	74.50	-0.592	3.20E-05	0.0198	SPATA25
ENSG00000058453	649.34	0.613	3.86E-05	0.0233	CROCC
ENSG00000123201	17.53	0.705	3.99E-05	0.0235	GUCY1B2
ENSG00000131080	84.23	0.854	4.18E-05	0.0235	EDA2R
ENSG00000228589	362.35	-0.325	4.13E-05	0.0235	SPCS2P4
ENSG00000081189	35766.76	-0.426	4.84E-05	0.0256	MEF2C
ENSG00000182253	51345.06	-0.365	4.88E-05	0.0256	SYNM
ENSG00000230082	51.99	0.561	4.78E-05	0.0256	PRRT3-AS1
ENSG00000003137	108.67	0.845	5.02E-05	0.0258	CYP26B1
ENSG00000115221	6928.18	-0.410	5.38E-05	0.0271	ITGB6
ENSG00000144771	4.02	0.796	5.77E-05	0.0281	LRTM1
ENSG00000169515	123.17	0.830	5.82E-05	0.0281	CCDC8
ENSG00000198453	435.62	0.632	7.23E-05	0.0342	ZNF568

ENSG00000100344	33.77	0.818	7.47E-05	0.0346	PNPLA3
ENSG00000104848	2466.99	-0.497	7.91E-05	0.0360	KCNA7
ENSG00000225511	16.70	0.824	8.09E-05	0.0362	LINC00475
ENSG00000116717	335.79	0.755	8.64E-05	0.0366	GADD45A
ENSG00000257542	162.51	-0.702	8.56E-05	0.0366	OR7E47P
ENSG00000260971	141.77	-0.584	8.55E-05	0.0366	AC119674.1
ENSG00000126749	383.97	-0.298	8.95E-05	0.0366	EMG1
ENSG00000250978	1346.70	0.757	8.89E-05	0.0366	AC079467.1
ENSG00000066629	2056.91	-0.415	9.73E-05	0.0379	EML1
ENSG00000136273	621.90	-0.206	9.61E-05	0.0379	HUS1
ENSG00000225083	63.06	-0.729	9.70E-05	0.0379	GRTP1-AS1
ENSG00000161513	118.95	0.740	1.00E-04	0.0384	FDXR
ENSG00000134339	20.55	0.726	1.02E-04	0.0385	SAA2
ENSG00000139117	126.62	0.642	1.14E-04	0.0404	CPNE8
ENSG00000163171	2081.29	0.576	1.12E-04	0.0404	CDC42EP3
ENSG00000166741	331.41	0.812	1.11E-04	0.0404	NNMT
ENSG00000237854	238.70	-0.357	1.12E-04	0.0404	LINC00674
ENSG00000164070	184.55	0.530	1.22E-04	0.0428	HSPA4L
ENSG00000134318	10161.95	-0.329	1.25E-04	0.0433	ROCK2
ENSG00000188039	20.46	0.800	1.30E-04	0.0436	NWD1
ENSG00000197562	501.97	0.397	1.29E-04	0.0436	RAB40C
ENSG00000007171	29.51	-0.781	1.33E-04	0.0439	NOS2
ENSG00000062524	10.80	0.753	1.44E-04	0.0463	LTK
ENSG00000088179	3528.82	-0.368	1.43E-04	0.0463	PTPN4
ENSG00000260337	54.43	0.790	1.47E-04	0.0467	LOC105370978
ENSG00000245648	13.78	0.791	1.51E-04	0.0472	LOC101928100
ENSG00000222939	78.56	0.761	1.59E-04	0.0493	AC079467.1
ENSG00000116652	139.48	0.776	1.64E-04	0.0500	DLEU2L
ENSG00000164040	817.10	-0.408	1.70E-04	0.0512	PGRMC2
ENSG00000138193	590.21	0.709	1.74E-04	0.0516	PLCE1
ENSG00000204227	563.39	0.229	1.75E-04	0.0516	RING1
ENSG00000120658	120.70	0.748	1.82E-04	0.0528	ENOX1
ENSG00000158526	734.56	-0.243	1.85E-04	0.0532	TSR2
ENSG00000235884	5.60	0.712	1.94E-04	0.0542	LINC00941
ENSG00000267296	108.65	0.551	1.96E-04	0.0542	CEBPA-AS1
ENSG00000275149	4.28	0.741	1.92E-04	0.0542	RP11-427J23.1
ENSG00000198780	7.71	-0.771	2.01E-04	0.0548	FAM169A
ENSG00000229175	3.26	0.651	2.02E-04	0.0548	LINC00382
ENSG00000241316	185.91	-0.489	2.04E-04	0.0548	SUCLG2-AS1
ENSG00000178184	56.01	0.674	2.10E-04	0.0556	PARD6G
ENSG00000136404	1151.88	-0.353	2.13E-04	0.0558	TM6SF1
ENSG00000167842	505.91	-0.308	2.32E-04	0.0596	MIS12
ENSG00000198176	1350.49	-0.250	2.31E-04	0.0596	TFDP1
ENSG00000169495	5.80	0.745	2.36E-04	0.0599	HTRA4
ENSG00000128272	10024.38	-0.377	2.45E-04	0.0615	ATF4
ENSG00000253426	53.90	-0.552	2.52E-04	0.0628	AC022784.3
ENSG00000134363	134.43	0.767	2.61E-04	0.0640	FST
ENSG00000168268	98.33	0.552	2.63E-04	0.0640	NT5DC2
ENSG00000071282	10093.45	-0.511	2.66E-04	0.0643	LMCD1
ENSG00000030304	146.59	0.745	2.75E-04	0.0652	MUSK
ENSG00000158008	646.92	-0.713	2.76E-04	0.0652	EXTL1
ENSG00000142864	9096.17	-0.261	2.81E-04	0.0659	SERBP1
ENSG00000106258	13.49	0.748	2.85E-04	0.0659	CYP3A5
ENSG00000168118	1429.00	-0.315	2.87E-04	0.0659	RAB4A

ENSG00000105327	157.71	0.649	2.90E-04	0.0660	BBC3
ENSG00000163071	79.57	0.667	3.04E-04	0.0686	SPATA18
ENSG00000146005	16.29	0.754	3.07E-04	0.0687	PSD2
ENSG00000103064	2556.63	0.499	3.29E-04	0.0696	SLC7A6
ENSG00000113924	4.83	-0.719	3.26E-04	0.0696	HGD
ENSG00000119865	211.89	0.540	3.23E-04	0.0696	CNRIP1
ENSG00000137571	3074.83	-0.407	3.28E-04	0.0696	SLCO5A1
ENSG00000159110	343.10	-0.595	3.27E-04	0.0696	IFNAR2
ENSG00000235920	8.41	0.727	3.20E-04	0.0696	AC073109.2
ENSG00000184378	11.57	0.752	3.44E-04	0.0721	ACTRT3
ENSG00000173992	341.84	0.295	3.48E-04	0.0725	CCS
ENSG00000278627	5.36	0.752	3.71E-04	0.0765	AC005962.2
ENSG00000064419	1983.79	-0.207	4.04E-04	0.0826	TNPO3
ENSG00000146926	1478.99	-0.332	4.23E-04	0.0856	ASB10
ENSG00000223305	15.04	0.663	4.26E-04	0.0856	RN7SKP30
ENSG00000154122	4264.16	-0.290	4.51E-04	0.0900	ANKH
ENSG00000134769	6581.30	-0.283	4.56E-04	0.0902	DTNA
ENSG00000177551	8.36	0.666	4.69E-04	0.0906	NHLH2
ENSG00000204711	5.58	-0.726	4.67E-04	0.0906	C9orf135
ENSG00000090447	239.37	0.443	4.81E-04	0.0916	TFAP4
ENSG00000124762	523.40	0.732	4.82E-04	0.0916	CDKN1A
ENSG00000141510	207.27	0.431	4.93E-04	0.0916	TP53
ENSG00000153956	18231.03	-0.328	4.89E-04	0.0916	CACNA2D1
ENSG00000253619	9.23	0.672	4.93E-04	0.0916	AC068413.1
ENSG00000120659	5.17	0.567	5.18E-04	0.0955	TNFSF11
ENSG00000117640	3006.63	-0.343	5.36E-04	0.0978	MTFR1L
ENSG00000267072	3.32	-0.679	5.39E-04	0.0978	NAGPA-AS1
ENSG00000136381	2536.88	-0.213	5.47E-04	0.0985	IREB2
ENSG00000120256	236.21	0.455	5.57E-04	0.0996	LRP11

Appendix E – Top 50 dmCpGs associated with ALMI in Muscle

Probe	Average Methylation	P Value	hg19 coordinates	Probe Location	Gene Symbol
cg00000029	0.340	6.68E-06	chr14:69995178-	3'UTR	FLJ44817
cg00000108	0.915	7.82E-06	chr19:6822454+	Body	VAV1
cg00000109	0.792	1.54E-05	chr15:25756044+		
cg00000165	0.149	1.91E-05	chr6:29521783+		
cg00000236	0.445	1.93E-05	chr19:13875014-	TSS1500	MRI1
cg00000289	0.650	2.52E-05	chr13:112237440-		
cg00000292	0.343	3.39E-05	chr19:13875950+	Body	MRI1
cg00000321	0.231	3.39E-05	chr1:38606161+		
cg00000363	0.470	3.48E-05	chr6:35733824+		
cg00000622	0.017	3.71E-05	chr3:84259272+		
cg00000658	0.834	4.02E-05	chr7:14192484+	Body	DGKB
cg00000714	0.112	4.66E-05	chr16:68401882-	Body	SMPD3
cg00000721	0.909	4.79E-05	chr8:19243209+	Body	SH2D4A
cg00000734	0.053	5.16E-05	chr1:246770778+	Body	CNST
cg00000769	0.031	5.40E-05	chr11:4566853-	1stExon	OR52M1
cg00000884	0.875	6.43E-05	chr10:13425688+		
cg00000905	0.046	6.45E-05	chr6:28973483-	TSS1500	ZNF311
cg00000924	0.497	6.64E-05	chr19:11071239+	TSS1500	SMARCA4
cg00000948	0.920	6.92E-05	chr17:72199373+	TSS1500	RPL38
cg00000957	0.859	7.13E-05	chr7:4854126+	Body	RADIL
cg00001099	0.730	7.42E-05	chr9:112914149+	Body	AKAP2
cg00001245	0.024	7.64E-05	chr12:38550112-		
cg00001249	0.702	7.72E-05	chr19:18672850-	5'UTR	C19orf50
cg00001261	0.462	7.97E-05	chr11:69706527+		
cg00001269	0.848	7.97E-05	chr7:25898688-		
cg00001349	0.737	8.16E-05	chr4:160118923-		
cg00001364	0.542	8.37E-05	chr11:72504909+	TSS200	STARD10
cg00001446	0.806	8.63E-05	chr5:160154401+	5'UTR	ATP10B
cg00001534	0.903	8.79E-05	chr16:85122369+	3'UTR	KIAA0513
cg00001582	0.044	9.31E-05	chr6:28583971+		
cg00001583	0.025	9.53E-05	chr2:100466754+	Body	AFF3
cg00001594	0.031	9.75E-05	chr14:106003576+		
cg00001687	0.969	1.02E-04	chr7:150554915-	Body	ABP1
cg00001747	0.018	1.03E-04	chr12:111751588-	Body	CUX2
cg00001791	0.808	1.04E-04	chr2:219762664+		
cg00001793	0.854	1.21E-04	chr12:7281343-	1stExon	RBP5
cg00001809	0.786	1.23E-04	chr10:134361124+	Body	INPP5A
cg00001854	0.660	1.29E-04	chr21:26934682-	Body	MIR155HG
cg00001874	0.638	1.31E-04	chr15:35262084-	TSS200	AQR
cg00001930	0.688	1.38E-04	chr19:17759636+	Body	UNC13A
cg00002028	0.030	1.38E-04	chr6:10494999+		
cg00002033	0.207	1.42E-04	chr3:137505160-		
cg00002080	0.835	1.45E-04	chr19:835499+		
cg00002116	0.015	1.54E-04	chr2:85482129-	Body	TCF7L1
cg00002145	0.758	1.57E-04	chr7:25898758-		
cg00002190	0.822	1.74E-04	chr6:32056637+	Body	TNXB
cg00002224	0.141	1.76E-04	chr9:74061323+		
cg00002236	0.034	1.78E-04	chr11:85438818+	Body	SYTL2
cg00002406	0.022	1.92E-04	chr7:156309329+		
cg00002426	0.434	1.92E-04	chr7:157461368-	Body	PTPRN2

Appendix F – Top 50 dmCpGs associated with Gait in Muscle

Probe	Average Methylation	P Value	hg19 coordinates	Probe Location	Gene Symbol
cg18432572	0.626	4.80E-06	chr5:156737087+	Body	CYFIP2
cg16740427	0.277	5.02E-06	chr22:42469976-	TSS1500	FAM109B
cg22236485	0.792	7.30E-06	chr8:128742664-		
cg02761778	0.655	1.48E-05	chr16:69643871+	5'UTR	NFAT5
cg15128638	0.410	1.55E-05	chr3:124493484+	Body	ITGB5
cg12995372	0.653	1.68E-05	chr7:11066024+	Body	PHF14
cg11594299	0.246	2.00E-05	chr7:4924002-	TSS1500	RADIL
cg14424836	0.935	2.15E-05	chr9:131515105-	Body	ZER1
cg11500788	0.760	2.53E-05	chr8:28355368+	5'UTR	FZD3
cg21163392	0.914	2.64E-05	chr12:112681365+	Body	C12orf51
cg27190839	0.716	2.65E-05	chr11:33097035+	TSS1500	LOC283267
cg26580095	0.319	2.83E-05	chr1:92545650+	TSS1500	BTBD8
cg26296941	0.841	3.06E-05	chr15:93114062-	Body	LOC100144604
cg21294945	0.882	3.07E-05	chr1:151996246+		
cg05144884	0.872	3.25E-05	chr16:2770429+	1stExon	PRSS27
cg06096677	0.426	3.27E-05	chr17:6338330-	1stExon	AIPL1
cg09233204	0.887	3.64E-05	chr10:94819559-	TSS1500	CYP26C1
cg02756176	0.133	3.98E-05	chr4:187491362-		
cg04064241	0.879	4.22E-05	chr4:121966889-	Body	C4orf31
cg00447469	0.946	4.39E-05	chr13:114251707+	Body	TFDP1
cg07393322	0.080	4.63E-05	chr22:43117318-	TSS1500	A4GALT
cg14846292	0.628	4.86E-05	chr12:49726500+	3'UTR	C1QL4
cg02751327	0.886	4.96E-05	chr12:64788110+		
cg08948149	0.909	5.00E-05	chr2:45739148-	Body	SRBD1
cg18033052	0.610	5.08E-05	chr22:18893141+	TSS1500	DGCR6
cg20333027	0.319	5.21E-05	chr2:231728977-	TSS1500	ITM2C
cg09950716	0.031	6.14E-05	chr8:124408885-	TSS200	ATAD2
cg11424665	0.158	6.16E-05	chr3:129158811+	TSS200	IFT122
cg07573936	0.572	6.16E-05	chr15:55805566+		
cg13916874	0.738	6.35E-05	chr6:4286300-		
cg07799277	0.056	6.47E-05	chr5:98368443+		
cg03786808	0.815	6.48E-05	chr17:41994384-	TSS200	C17orf88
cg17931890	0.044	6.66E-05	chr2:176962423-		
cg14308867	0.873	7.32E-05	chr6:31603476+	Body	BAT2
cg12837905	0.469	7.64E-05	chr10:131811621-		
cg04234556	0.922	7.69E-05	chr1:29059141-		
cg25671376	0.822	8.61E-05	chr6:89901084-	Body	GABRR1
cg03843170	0.718	8.79E-05	chr6:41253757-	Body	TREM1
cg05754905	0.904	8.88E-05	chr1:160992352-	TSS1500	F11R
cg11309456	0.623	8.91E-05	chr12:110942599+	Body	RAD9B
cg20808206	0.797	9.22E-05	chr15:81130947-	5'UTR	KIAA1199
cg20367788	0.647	9.73E-05	chr15:67534499-	Body	AAGAB
cg08219459	0.846	1.04E-04	chr16:68734735+		
cg06864151	0.899	1.04E-04	chr7:1913918-	Body	MAD1L1
cg13407274	0.902	1.08E-04	chr9:139367397+	Body	SEC16A
cg08862181	0.692	1.15E-04	chr17:10277045-	TSS1500	MYH13
cg20076301	0.763	1.16E-04	chr8:54568339+		
cg04843801	0.042	1.24E-04	chr3:42642061+	TSS200	NKTR
cg01937808	0.812	1.28E-04	chr16:4025588-	Body	ADCY9
cg13543683	0.773	1.29E-04	chr9:101175220-	Body	GABBR2

Appendix G – Top 50 dmCpGs associated with Fasting Glucose in Muscle

Probe	Average Methylation	P Value	hg19 coordinates	Probe Location	Gene Symbol
cg08932727	0.240	3.36E-09	chr6:328182+	Body	DUSP22
cg11064291	0.863	6.42E-08	chr8:26980219-		
cg17542333	0.817	8.36E-08	chr11:10673844-	5'UTR	MRVI1
cg06126494	0.765	1.69E-07	chr4:111715333-		
cg24834199	0.823	1.82E-07	chr1:225613299+	5'UTR	LBR
cg08306303	0.923	2.56E-07	chr11:119567761-	Body	PVRL1
cg09015905	0.753	2.58E-07	chr1:10692435+		
cg11088471	0.733	3.20E-07	chr21:38085281-	Body	SIM2
cg01598421	0.873	3.33E-07	chr8:134132043-	Body	TG
cg25878744	0.902	3.52E-07	chr1:244303505+		
cg12910797	0.795	4.36E-07	chr17:46651722+	5'UTR	HOXB3
cg17896879	0.071	5.11E-07	chr2:233414901+	5'UTR	TIGD1
cg13420004	0.887	5.23E-07	chr15:89421300-	Body	HAPLN3
cg09967440	0.831	5.36E-07	chr10:127444969+	Body	C10orf137
cg15203214	0.610	5.52E-07	chr6:379489+		
cg08325885	0.876	5.83E-07	chr8:22297108+	TSS1500	PPP3CC
cg19120580	0.818	5.93E-07	chr12:49173414+	Body	ADCY6
cg25551358	0.925	6.18E-07	chr5:139079288+		
cg04723679	0.873	7.64E-07	chr15:62548919+		
cg02734505	0.694	8.52E-07	chr12:54763081-	3'UTR	ZNF385A
cg24963001	0.631	9.13E-07	chr17:46651952+	TSS200	HOXB3
cg17144149	0.709	9.14E-07	chr17:46656572+	TSS1500	HOXB4
cg18055067	0.846	9.28E-07	chr5:148803404-	Body	LOC728264
cg18776472	0.884	9.40E-07	chr10:50732819-	Body	ERCC6
cg27026695	0.858	1.07E-06	chr14:65006016+	TSS1500	HSPA2
cg05168453	0.794	1.10E-06	chr10:134437948+	Body	INPP5A
cg26612230	0.923	1.20E-06	chr4:95508526+	Body	PDLIM5
cg26916621	0.728	1.21E-06	chr17:46657346-	TSS200	MIR10A
cg05977002	0.156	1.24E-06	chr5:92908051-	Body	FLJ42709
cg14578879	0.735	1.31E-06	chr6:100904055-	Body	SIM1
cg21520549	0.881	1.33E-06	chr2:36589883-	Body	CRIM1
cg21880624	0.934	1.35E-06	chr21:40194678-	Body	ETS2
cg12218099	0.829	1.39E-06	chr6:159188159-	Body	EZR
cg22699448	0.058	1.41E-06	chr12:75724083-	TSS1500	CAPS2
cg07911171	0.623	1.41E-06	chr1:7340656+	Body	CAMTA1
cg13657511	0.908	1.42E-06	chr10:134459241-	Body	INPP5A
cg15649236	0.711	1.44E-06	chr17:46657504-	TSS200	MIR10A
cg27190654	0.387	1.44E-06	chr4:109387521-		
cg03426023	0.233	1.48E-06	chr16:54972078+		
cg10758369	0.811	1.48E-06	chr1:48459236-		
cg06454084	0.849	1.67E-06	chr19:47158242+	Body	DACT3
cg23944298	0.767	1.74E-06	chr3:127248864-		
cg00078759	0.899	1.79E-06	chr16:1600969-	Body	IFT140
cg06179765	0.030	1.81E-06	chr7:23287352+	Body	GPNMB
cg17118478	0.907	1.85E-06	chr2:191543902+	Body	NAB1
cg00317355	0.077	1.88E-06	chr16:49308483-		
cg11304734	0.114	1.90E-06	chr9:37485592+	TSS1500	POLR1E
cg08096367	0.917	1.95E-06	chr15:95384471+		
cg20693334	0.499	2.02E-06	chr17:46654330-	Body	HOXB4
cg02730156	0.456	2.02E-06	chr21:38069950+		

Appendix H – Top 50 dmCpGs associated with Grip in Muscle

Probe	Average Methylation	P Value	hg19 coordinates	Probe Location	Gene Symbol
cg20761395	0.858	1.16E-07	chr17:511517-	Body	VPS53
cg25071674	0.860	1.20E-07	chr8:141595345-	Body	EIF2C2
cg06257038	0.888	2.90E-06	chr12:127268720+		
cg16656232	0.918	1.83E-05	chr5:167543888+	Body	ODZ2
cg05126581	0.825	1.93E-05	chr1:216144819-	Body	USH2A
cg09380555	0.849	2.13E-05	chr13:53277872-	Body	LECT1
cg20814813	0.736	2.29E-05	chr1:247506185-		
cg24795748	0.661	2.30E-05	chr21:47404165+	Body	COL6A1
cg06354193	0.540	2.40E-05	chr1:3183072-	Body	PRDM16
cg22900476	0.425	2.46E-05	chr2:173647945-	Body	RAPGEF4
cg00471142	0.911	2.67E-05	chr19:39806765-	TSS1500	LRFN1
cg06742628	0.330	2.80E-05	chr5:16886424+	Body	MYO10
cg06071083	0.038	2.85E-05	chr16:68002365+	1stExon	SLC12A4
cg16555341	0.045	3.19E-05	chr16:89160798-	Body	ACSF3
cg25271146	0.818	3.25E-05	chrX:24382955+	1stExon	FAM48B1
cg12622520	0.766	3.40E-05	chr3:49192768-	TSS1500	LAMB2L
cg00008544	0.795	3.94E-05	chr5:179778527+	Body	GFPT2
cg15767191	0.734	4.16E-05	chr11:123900083-	TSS1500	OR10G8
cg26499561	0.887	4.73E-05	chrX:24952147+	Body	POLA1
cg12865252	0.886	5.19E-05	chr5:33457705-	Body	TARS
cg01547344	0.833	5.38E-05	chr5:153282014+		
cg02967812	0.552	5.46E-05	chr17:80066314-	Body	CCDC57
cg16993043	0.179	5.81E-05	chr1:200008026+	Body	NR5A2
cg16363238	0.871	5.85E-05	chr6:159618265+	Body	FNDC1
cg02848093	0.914	6.03E-05	chr3:74659531+		
cg21771679	0.620	6.16E-05	chr7:131206773+	Body	PODXL
cg08734527	0.492	6.25E-05	chr1:18092835-	5'UTR	ACTL8
cg02603251	0.522	6.43E-05	chr5:156798048+	Body	CYFIP2
cg02152271	0.825	6.55E-05	chr15:25199270-	5'UTR	SNRPN
cg00584089	0.754	6.55E-05	chr2:202776922+		
cg11471401	0.375	6.58E-05	chr12:52888480-	TSS1500	KRT6A
cg16635196	0.700	6.68E-05	chr3:169758644-	5'UTR	GPR160
cg04145522	0.965	7.00E-05	chr16:87738833+	Body	LOC100129637
cg03796321	0.033	7.90E-05	chr15:86162442+	Body	AKAP13
cg00146756	0.477	8.05E-05	chr4:6729110-		
cg26689077	0.406	8.29E-05	chr12:53599806-	5'UTR	ITGB7
cg20152891	0.098	8.46E-05	chr19:49944506-	Body	SLC17A7
cg20541456	0.422	8.72E-05	chr5:156696520+	1stExon	CYFIP2
cg09319125	0.879	9.27E-05	chr17:77537803+		
cg14172137	0.807	9.41E-05	chr6:114725242+		
cg01436487	0.767	9.46E-05	chr7:1890148-	Body	MAD1L1
cg08778123	0.911	1.03E-04	chr5:51813703+		
cg23576723	0.874	1.04E-04	chr10:105206636-	3'UTR	CALHM2
cg26858540	0.893	1.08E-04	chr19:56603322+	Body	ZNF787
cg09894136	0.715	1.09E-04	chr22:43833175+	Body	MPPED1
cg22708738	0.072	1.11E-04	chr19:50316212+	Body	FUZ
cg10595753	0.897	1.15E-04	chr17:80898281-	Body	TBCD
cg14234083	0.599	1.17E-04	chr17:71731988-		
cg13324103	0.271	1.27E-04	chr10:29948428-	5'UTR	SVIL
cg17152266	0.668	1.34E-04	chr19:48694091+		

Appendix I – Top 50 dmCpGs associated with Insulin in Muscle

Probe	Average Methylation	P Value	hg19 coordinates	Probe Location	Gene Symbol
cg04919982	0.851	1.51E-07	chr12:18243179-	TSS200	RERGL
cg06405341	0.872	1.84E-07	chr17:63535223+	Body	AXIN2
cg01311051	0.166	2.35E-07	chr16:30064897+	5'UTR	ALDOA
cg00729885	0.223	2.56E-07	chr13:36046056-	Body	NBEA
cg26799416	0.890	2.76E-07	chr22:42611379+	1stExon	TCF20
cg14774585	0.247	3.03E-07	chr13:36046577+	Body	NBEA
cg14542554	0.270	3.12E-07	chr14:61122327-		
cg01952313	0.093	4.90E-07	chr13:36050881-	Body	MIR548F5
cg01047748	0.687	5.19E-07	chr19:35849024+	TSS1500	FFAR3
cg20737909	0.711	5.39E-07	chr11:1507020+	Body	HCCA2
cg03184424	0.851	5.45E-07	chr7:811786+	Body	HEATR2
cg18238808	0.883	5.72E-07	chr17:71280700+	3'UTR	CDC42EP4
cg00024494	0.788	5.79E-07	chr16:73096842-		
cg12910797	0.795	6.15E-07	chr17:46651722+	5'UTR	HOXB3
cg23058673	0.782	7.26E-07	chr7:5422234-	Body	TNRC18
cg24368912	0.836	7.97E-07	chr4:100243455+	TSS1500	ADH1B
cg01939872	0.891	7.98E-07	chr2:99948976-	Body	TXNDC9
cg05369942	0.075	8.18E-07	chr4:142228617-		
cg24834199	0.823	8.48E-07	chr1:225613299+	5'UTR	LBR
cg08293531	0.951	8.53E-07	chr15:101601180+	Body	LRRK1
cg16396948	0.808	1.02E-06	chr3:111603394-	Body	PHLDB2
cg26406074	0.123	1.23E-06	chr13:36044768-	Body	NBEA
cg13446906	0.168	1.37E-06	chr13:36051073+	Body	MIR548F5
cg24610349	0.849	1.43E-06	chr7:98696198-	Body	SMURF1
cg06768361	0.935	1.44E-06	chr12:117480333+	Body	TESC
cg03612413	0.449	1.45E-06	chr21:38066918-		
cg08787837	0.114	1.47E-06	chr16:86029554+		
cg09676376	0.794	1.50E-06	chr12:54779077-	Body	ZNF385A
cg19324627	0.805	1.51E-06	chr11:116708329+	1stExon	APOA1
cg04141648	0.882	1.54E-06	chr10:69456667-	TSS1500	CTNNA3
cg20881175	0.832	1.57E-06	chr2:218741790-	Body	TNS1
cg11753771	0.965	1.59E-06	chr1:214638070-	Body	PTPN14
cg18690368	0.872	1.65E-06	chr7:29924746+	Body	WIPF3
cg10682299	0.659	1.73E-06	chr4:111550154-	Body	PITX2
cg01783816	0.126	1.81E-06	chr13:103523190-	Body	ERCC5
cg23833461	0.767	1.83E-06	chr17:9929528-	1stExon	GAS7
cg07854457	0.085	1.85E-06	chr5:126898200-		
cg21593835	0.791	1.88E-06	chr3:123354036-	Body	MYLK
cg04747539	0.879	1.93E-06	chr6:22887383-		
cg25285794	0.095	1.97E-06	chr12:54402699-	TSS200	HOXC8
cg25767112	0.886	2.19E-06	chr1:8394362-	Body	SLC45A1
cg22335876	0.055	2.30E-06	chr7:121940581+		
cg27016272	0.888	2.37E-06	chr3:32344486-	Body	CMTM8
cg12361352	0.043	2.37E-06	chr18:31158540+	5'UTR	ASXL3
cg09967440	0.831	2.41E-06	chr10:127444969+	Body	C10orf137
cg25878744	0.902	2.51E-06	chr1:244303505+		
cg25930644	0.853	2.53E-06	chr17:8531915-	5'UTR	MYH10
cg15978565	0.075	2.53E-06	chr2:11894464-	5'UTR	LPIN1
cg16564033	0.713	2.54E-06	chr21:38067741+		
cg11809342	0.901	2.55E-06	chr2:160007430-	Body	TANC1

Appendix J – Top 50 dmCpGs associated with ALMI in Myoblasts

Probe	Average Methylation	P Value	hg19 coordinates	Probe Location	Gene Symbol
cg23887948	0.903	3.57E-07	chr15:74494854+	5'UTR	STRA6
cg22001110	0.717	6.63E-07	chr6:31626313+	3'UTR	C6orf47
cg21505334	0.912	9.69E-07	chr19:42213020-	Body	CEACAM5
cg02130555	0.916	1.17E-06	chr4:40811505+	3'UTR	NSUN7
cg00982548	0.948	4.65E-06	chr2:198649783-	5'UTR	BOLL
cg06583549	0.781	1.10E-05	chr19:46387962-	1stExon	IRF2BP1
cg20752472	0.896	1.10E-05	chr5:14710734-	3'UTR	ANKH
cg05822022	0.873	1.30E-05	chr12:56397822+	Body	SUOX
cg22008026	0.098	1.46E-05	chr4:155413392+	TSS1500	DCHS2
cg08513527	0.878	1.64E-05	chr8:33333687-		
cg01284448	0.903	1.77E-05	chr14:76114486+	3'UTR	FLVCR2
cg23212579	0.924	1.86E-05	chr2:219256018-	Body	SLC11A1
cg25938042	0.044	2.69E-05	chr16:87637283+	Body	JPH3
cg13744194	0.087	2.69E-05	chr5:68788087-	TSS200	OCLN
cg26135975	0.859	3.56E-05	chr17:10635580+		
cg20884780	0.926	3.67E-05	chr15:90192525+	Body	KIF7
cg08949428	0.807	3.81E-05	chr13:98451284+		
cg10314139	0.626	3.81E-05	chr10:71560231-	TSS1500	COL13A1
cg14788994	0.721	3.94E-05	chr17:8274570+	Body	KRBA2
cg27268835	0.699	4.29E-05	chr16:51475605+		
cg15228451	0.824	4.46E-05	chr20:34649125-		
cg15210307	0.791	4.77E-05	chr1:32498089-	Body	KHDRBS1
cg14184780	0.582	5.00E-05	chr9:95475751+	Body	BICD2
cg10324810	0.931	5.03E-05	chr16:2762695-	Body	PRSS27
cg15518491	0.106	5.28E-05	chr1:70820790-	Body	HHLA3
cg23411240	0.879	5.42E-05	chr2:238285512-	Body	COL6A3
cg00361562	0.930	5.49E-05	chr2:198649771-	5'UTR	BOLL
cg13459498	0.024	5.58E-05	chr20:42544792-	5'UTR	TOX2
cg05487269	0.809	5.59E-05	chr16:3000729-	3'UTR	FLYWCH1
cg13983566	0.803	5.77E-05	chr15:64336065-	5'UTR	DAPK2
cg08572875	0.911	5.87E-05	chr17:296110+	TSS1500	FAM101B
cg00588262	0.885	5.89E-05	chr19:42348588-	TSS200	LYPD4
cg03422651	0.878	5.99E-05	chr16:2542324+	5'UTR	TBC1D24
cg18584550	0.886	6.17E-05	chr19:1636987-	Body	TCF3
cg12917258	0.054	6.20E-05	chr10:99258018+	TSS1500	UBTD1
cg07281948	0.696	6.51E-05	chr12:57601935+	Body	LRP1
cg04794887	0.080	6.54E-05	chr4:1858231+	TSS1500	LETM1
cg17820989	0.062	6.67E-05	chr8:145734432-	TSS200	MFSD3
cg25858160	0.390	6.68E-05	chr20:55499436+		
cg16712828	0.841	7.11E-05	chr20:3193986+	Body	ITPA
cg27333274	0.051	7.15E-05	chr16:57279026-	TSS200	ARL2BP
cg12785082	0.777	7.71E-05	chr7:101662332-	Body	CUX1
cg07439880	0.053	8.03E-05	chr7:2395427+	Body	EIF3B
cg23537415	0.168	8.30E-05	chr19:56169445-	Body	U2AF2
cg01332054	0.777	8.31E-05	chr15:74494861+	5'UTR	STRA6
cg07703190	0.927	8.33E-05	chr21:47165787-		
cg17156011	0.686	8.37E-05	chr1:224357281-		
cg09124518	0.838	8.41E-05	chr10:134119682+	Body	STK32C
cg18396459	0.825	9.17E-05	chr20:2384367+	Body	TGM6
cg07434260	0.091	9.26E-05	chr19:10698036-	TSS200	AP1M2

Appendix K – Top 50 dmCpGs associated with Gait in Myoblasts

Probe	Average Methylation	P Value	hg19 coordinates	Probe Location	Gene Symbol
cg06578851	0.411	1.50E-05	chrX:101315561-		
cg06680826	0.907	1.65E-05	chr15:95022460+	3'UTR	MCTP2
cg21280014	0.108	2.21E-05	chr7:157644665+	Body	PTPRN2
cg09866409	0.858	2.32E-05	chr3:46608802-	5'UTR	LRRC2
cg13475977	0.933	2.76E-05	chr13:21047397-	Body	CRYL1
cg26143141	0.909	3.52E-05	chr4:7740821+	Body	SORCS2
cg22934516	0.076	4.84E-05	chr11:35413951-	Body	SLC1A2
cg24517295	0.164	4.88E-05	chr3:14641410+		
cg04308167	0.653	4.97E-05	chr17:19440767-	Body	SLC47A1
cg18784855	0.836	5.42E-05	chr13:61791681+		
cg24155871	0.092	5.86E-05	chr1:75139347+	1stExon	C1orf173
cg13854930	0.825	6.30E-05	chr12:55051924+		
cg21742463	0.867	6.58E-05	chr12:92822923+	TSS1500	CLLU1OS
cg00446536	0.903	6.92E-05	chr2:6636481-		
cg12591207	0.850	6.94E-05	chrX:19008006-	3'UTR	GPR64
cg12715458	0.044	7.61E-05	chr11:8892757-	5'UTR	ST5
cg02644729	0.783	7.62E-05	chr2:97361045-	Body	FER1L5
cg08817693	0.406	7.73E-05	chr12:107348944+	TSS1500	C12orf23
cg10988946	0.880	8.38E-05	chr14:19828288-		
cg02505588	0.863	9.05E-05	chr5:105753028-		
cg05596468	0.949	9.22E-05	chr8:1885915-	Body	ARHGEF10
cg13398192	0.114	9.40E-05	chr13:108518152-	1stExon	FAM155A
cg24371906	0.857	1.04E-04	chr1:212719743+		
cg23161794	0.055	1.09E-04	chr6:30068942-		
cg08476956	0.851	1.10E-04	chr16:3645613+	Body	BTBD12
cg24469784	0.907	1.13E-04	chr2:241723513+	Body	KIF1A
cg15850513	0.800	1.14E-04	chrX:47091513-	TSS1500	USP11
cg07530759	0.854	1.42E-04	chr7:141541627+	TSS1500	PRSS37
cg13180232	0.843	1.46E-04	chr10:35425499-	5'UTR	CREM
cg22106284	0.838	1.58E-04	chr7:75053027-	Body	POM121C
cg01454528	0.059	1.61E-04	chr17:49243444-	TSS200	NME2
cg04904531	0.828	1.86E-04	chr1:118405119-		
cg07916058	0.870	1.90E-04	chr1:180644805+	Body	XPR1
cg09042759	0.904	1.91E-04	chr6:32134176+	Body	EGFL8
cg01666652	0.895	1.92E-04	chr14:22987689+		
cg07274618	0.033	2.10E-04	chr17:74070698-	TSS200	GALR2
cg27393126	0.058	2.21E-04	chr1:75139364+	1stExon	C1orf173
cg06100421	0.430	2.36E-04	chr7:130131136-	TSS200	MESTIT1
cg13906954	0.840	2.38E-04	chr3:56713650-	Body	C3orf63
cg08116651	0.887	2.38E-04	chr16:66835939-	TSS1500	CCDC79
cg07509167	0.931	2.39E-04	chr8:1822568+	Body	ARHGEF10
cg01530032	0.485	2.46E-04	chr17:19435805+	TSS1500	SLC47A1
cg05500015	0.046	2.47E-04	chr11:89867819+	5'UTR	NAALAD2
cg21227784	0.918	2.50E-04	chr20:60511759+	Body	CDH4
cg18477362	0.894	2.54E-04	chr3:166452764-		
cg25635591	0.874	2.55E-04	chr17:57642417-	TSS1500	DHX40
cg14871472	0.915	2.57E-04	chr6:151330408-	Body	MTHFD1L
cg18304305	0.863	2.60E-04	chr5:42720521-	3'UTR	GHR
cg22579849	0.025	2.62E-04	chr9:34620091-	Body	DCTN3
cg27656573	0.910	2.63E-04	chr22:30127118-	3'UTR	ZMAT5

Appendix L – Top 50 dmCpGs associated with Fasting Glucose in Myoblasts

Probe	Average Methylation	P Value	hg19 coordinates	Probe Location	Gene Symbol
cg06879822	0.073	4.85E-06	chr19:42719482+	Body	DEDD2
cg27519996	0.053	8.82E-06	chr10:13141822+	TSS1500	OPTN
cg25120290	0.044	1.59E-05	chr14:85996873-	1stExon	FLRT2
cg02761554	0.068	1.79E-05	chr3:9885425-	Body	RPUSD3
cg01364935	0.835	3.86E-05	chr16:73185687-		
cg07737560	0.030	4.17E-05	chr12:101470827-	Body	ANO4
cg11825059	0.047	4.37E-05	chr7:14876371+	Body	DGKB
cg25734928	0.236	5.26E-05	chr6:105398579-		
cg21682902	0.907	5.26E-05	chr12:96390960-	TSS1500	HAL
cg06988995	0.043	6.35E-05	chr10:43633847-	TSS200	CSGALNACT2
cg24580076	0.917	7.26E-05	chr7:915073+	TSS1500	C7orf20
cg19384775	0.060	8.26E-05	chr1:205012139+	TSS200	CNTN2
cg06436663	0.905	9.60E-05	chr17:41718597-	3'UTR	MEOX1
cg04518808	0.124	9.68E-05	chr19:54384822+	TSS1500	PRKCG
cg13669089	0.047	1.04E-04	chr11:31391497+	5'UTR	DNAJC24
cg07139440	0.039	1.05E-04	chr1:151030981+	5'UTR	CDC42SE1
cg21201657	0.157	1.06E-04	chr19:5624574+	Body	SAFB
cg23381232	0.778	1.10E-04	chr7:95696973-	Body	DYNC1I1
cg04280590	0.856	1.12E-04	chr9:33510387-	Body	SUGT1P1
cg09188978	0.937	1.18E-04	chr12:78227745-	Body	NAV3
cg10673351	0.445	1.22E-04	chr11:2471097-	Body	KCNQ1
cg19674178	0.253	1.29E-04	chr15:49752768+	Body	FGF7
cg25808809	0.217	1.34E-04	chr13:114917199-		
cg04329136	0.925	1.42E-04	chr15:94443775-		
cg02613601	0.817	1.44E-04	chr13:32183063-		
cg04837616	0.589	1.47E-04	chr17:75880542+	TSS1500	FLJ45079
cg02666105	0.093	1.54E-04	chr16:31724574+	5'UTR	ZNF720
cg13382100	0.139	1.62E-04	chr12:132414849-	Body	PUS1
cg00448707	0.857	1.65E-04	chr1:68512928+	Body	DIRAS3
cg23350294	0.187	1.65E-04	chr8:57359992-	TSS1500	PENK
cg12737833	0.528	1.67E-04	chr2:101500543+	5'UTR	NPAS2
cg11172423	0.593	1.71E-04	chr1:43205623-	1stExon	CLDN19
cg08109266	0.833	1.72E-04	chr14:37773865+	Body	MIPOL1
cg25141757	0.882	1.72E-04	chr5:57759521-		
cg25022915	0.149	1.76E-04	chr17:49243257-	TSS1500	NME2
cg08277036	0.913	1.76E-04	chr10:134451258-	Body	INPP5A
cg01074767	0.753	1.77E-04	chr12:7262137+	TSS1500	C1RL
cg25259479	0.571	1.82E-04	chr7:157408632+	Body	PTPRN2
cg06056415	0.097	1.83E-04	chr5:174162510-		
cg16359901	0.167	1.90E-04	chr1:6269635+	Body	RNF207
cg19003337	0.887	1.91E-04	chr17:39780836+	5'UTR	KRT17
cg22851880	0.423	1.93E-04	chr1:36038701-	TSS1500	TFAP2E
cg06739711	0.568	1.97E-04	chr14:62331920+		
cg22302985	0.031	1.99E-04	chr13:108520945-	TSS1500	FAM155A
cg18059699	0.888	2.08E-04	chr3:45467287-	Body	LARS2
cg17835614	0.600	2.16E-04	chr7:72409374+	Body	POM121
cg02567144	0.043	2.18E-04	chr6:52285149-	1stExon	EFHC1
cg18181923	0.629	2.20E-04	chr14:99682269+	Body	BCL11B
cg17011453	0.435	2.22E-04	chr1:81537650-		

Appendix M – Top 50 dmCpGs associated with Grip in Myoblasts

Probe	Average Methylation	P Value	hg19 coordinates	Gene Location	Gene Symbol
cg26994526	0.161	1.19E-06	chr18:47719735+	Body	MYO5B
cg12127290	0.768	7.10E-06	chr12:21531643+	3'UTR	IAPP
cg03254566	0.702	8.66E-06	chr21:44006596-		
cg03353654	0.398	1.77E-05	chr10:80677715-		
cg14218844	0.356	2.07E-05	chr18:56709536-	Body	LOC390858
cg00146756	0.462	2.54E-05	chr4:6729110-		
cg12033822	0.894	2.74E-05	chr20:44990591-	5'UTR	SLC35C2
cg22054885	0.687	3.42E-05	chr5:23507450-	TSS1500	PRDM9
cg24393552	0.930	3.48E-05	chr2:239014989+	Body	ESPNL
cg10989261	0.046	4.11E-05	chr1:6844145+	TSS1500	CAMTA1
cg04643444	0.726	4.41E-05	chr1:31653173-	3'UTR	NKAIN1
cg23034788	0.934	4.91E-05	chr11:47364745+	Body	MYBPC3
cg14442841	0.141	5.08E-05	chr6:20127257+	Body	MBOAT1
cg07412079	0.104	5.42E-05	chr18:33766607+	TSS1500	MOCOS
cg05896377	0.875	5.74E-05	chr7:3920297-	Body	SDK1
cg16742703	0.657	5.86E-05	chr19:51357857-	TSS1500	KLK3
cg14858014	0.495	5.93E-05	chr4:6728936+		
cg06009010	0.770	5.99E-05	chr17:650835+	Body	GEMIN4
cg05367893	0.676	6.80E-05	chr7:2186465-	Body	MAD1L1
cg12746947	0.555	7.55E-05	chr1:38227094+	Body	EPHA10
cg21473651	0.858	8.42E-05	chr2:178052747-		
cg00220492	0.040	8.95E-05	chr15:36871852-	1stExon	C15orf41
cg20837735	0.869	8.96E-05	chr18:61144177-	1stExon	SERPINB5
cg05200373	0.736	9.73E-05	chr6:33170101-	Body	SLC39A7
cg20545896	0.912	1.05E-04	chr7:141408315+	5'UTR	WEE2
cg24968336	0.101	1.05E-04	chr19:38810212+	TSS1500	KCNK6
cg00137839	0.813	1.08E-04	chr17:80916178+	Body	B3GNTL1
cg27602828	0.749	1.11E-04	chr18:3845348-	TSS200	DLGAP1
cg03071124	0.552	1.15E-04	chr7:157447178-	Body	PTPRN2
cg11238518	0.918	1.20E-04	chr2:208543957-		
cg03593908	0.398	1.31E-04	chr13:60842299-		
cg18658231	0.898	1.34E-04	chr2:170283027+		
cg25801742	0.316	1.36E-04	chr11:58736104-		
cg20705938	0.189	1.40E-04	chr4:111558988-	TSS1500	PITX2
cg02160738	0.728	1.41E-04	chr11:60961816+		
cg04967200	0.852	1.45E-04	chr10:134115376+	Body	STK32C
cg10120657	0.203	1.50E-04	chr17:77786354+		
cg13595143	0.409	1.65E-04	chr9:472966+		
cg25941299	0.171	1.70E-04	chr7:113727916-		
cg22499001	0.901	1.71E-04	chr17:62050280-	TSS200	SCN4A
cg19130189	0.774	1.72E-04	chr8:89241632+	Body	MMP16
cg15836427	0.870	1.72E-04	chr7:155640618-		
cg19837938	0.650	1.77E-04	chr5:23507458-	TSS1500	PRDM9
cg03144357	0.823	1.77E-04	chr8:145741925-	Body	RECQL4
cg17356750	0.395	1.78E-04	chr7:114329849+	Body	FOXP2
cg07344583	0.066	1.78E-04	chr15:89456631+	1stExon	MFGE8
cg25000623	0.225	1.79E-04	chr17:72848918+	Body	GRIN2C
cg16525281	0.143	1.81E-04	chr19:38810227+	TSS1500	KCNK6
cg12766106	0.907	1.87E-04	chr9:139803690+	Body	TRAF2
cg13862848	0.622	1.89E-04	chr13:44272416-	5'UTR	ENOX1

Appendix N – Top 50 dmCpGs associated with Insulin in Myoblasts

Probe	Average Methylation	P Value	hg19 coordinates	Probe Location	Gene Symbol
cg27290207	0.708	2.44E-06	chr3:190572892+	3'UTR	LOC647309
cg24019399	0.940	4.50E-06	chr6:168677763+		
cg11244095	0.869	6.11E-06	chrX:135382675+	TSS1500	GPR112
ch.8.2197061F	0.082	6.81E-06	chr8:109278439+		
cg17339147	0.059	9.06E-06	chr7:116963500+	TSS200	WNT2
cg11153969	0.902	1.44E-05	chr1:207277031+	TSS1500	C4BPA
cg15201635	0.064	1.74E-05	chr16:68482637-	TSS1500	SMPD3
cg13793810	0.916	1.85E-05	chr11:63435340-	Body	ATL3
cg00223767	0.727	2.24E-05	chr11:75831715+	Body	UVRAG
cg22203890	0.383	2.61E-05	chr5:980708+		
cg24282401	0.851	3.27E-05	chr10:133747754+	TSS1500	PPP2R2D
cg02150354	0.880	3.36E-05	chr5:92914202+	Body	FLJ42709
cg05134476	0.869	3.64E-05	chr8:75227222-	Body	JPH1
cg18138206	0.184	3.68E-05	chr10:133110731-	TSS1500	TCERG1L
cg17547898	0.867	3.83E-05	chr4:123769295+	Body	FGF2
cg18009704	0.813	3.86E-05	chr17:3117981+	TSS1500	OR1A1
cg03979024	0.486	4.13E-05	chr10:419189-	Body	DIP2C
cg13845521	0.073	4.55E-05	chrY:21238886+	Body	TTTY14
cg02757561	0.926	4.60E-05	chr10:124642139+	Body	LOC399815
cg20378921	0.786	4.64E-05	chr3:14251555+		
cg07912144	0.976	4.67E-05	chr18:77247106-	Body	NFATC1
cg05136650	0.701	4.71E-05	chr20:44523309-	Body	CTSA
cg10783870	0.901	4.88E-05	chr6:33255751+	Body	WDR46
cg07662121	0.039	4.98E-05	chr16:15489794+	1stExon	MPV17L
cg10598168	0.838	5.10E-05	chr5:147647955+	TSS1500	SPINK5L3
cg03643423	0.826	5.16E-05	chr17:11929694-	Body	MAP2K4
cg23319085	0.824	6.08E-05	chr2:74700469-	3'UTR	CCDC142
cg21647584	0.711	6.15E-05	chrX:134847724-	5'UTR	CT45A1
cg14470792	0.927	6.42E-05	chr4:183168158-		
cg21064080	0.068	6.72E-05	chr6:70576821+	5'UTR	COL19A1
cg05845592	0.039	6.79E-05	chr16:28634766-	5'UTR	SULT1A1
cg20806676	0.925	6.89E-05	chr6:29579306-	Body	GABBR1
cg14224789	0.891	6.89E-05	chr4:1990073+	Body	WHSC2
cg04616566	0.911	6.95E-05	chr15:72020560-	Body	THSD4
cg13054563	0.044	7.04E-05	chr14:23236577+	Body	OXA1L
cg17801256	0.092	7.19E-05	chr1:35734334-	TSS1500	ZMYM4
cg18465149	0.842	7.60E-05	chr6:31859141+	Body	EHMT2
cg25172682	0.047	8.67E-05	chr17:34122758+	TSS200	MMP28
cg24435704	0.738	9.00E-05	chr7:139763773+	TSS1500	PARP12
cg09689404	0.055	1.05E-04	chr10:3724497-		
cg02605601	0.043	1.05E-04	chr10:3724557-		
cg17297328	0.078	1.06E-04	chr1:108743728-	TSS1500	SLC25A24
cg06934038	0.240	1.08E-04	chr12:94468759-		
cg02104967	0.931	1.12E-04	chr17:1386368-	Body	MYO1C
cg08273362	0.913	1.19E-04	chr19:1005172+	Body	GRIN3B
cg20588982	0.038	1.23E-04	chr20:32308170-	TSS200	PXMP4
cg24011501	0.936	1.23E-04	chr17:74928863-	Body	MGAT5B
cg23900696	0.919	1.24E-04	chr12:13325724+		
cg25993608	0.081	1.30E-04	chr18:76621396+		
cg20272962	0.080	1.32E-04	chr17:56326765+	Body	LPO

Appendix O – dmCpGs associated with ALMI

Probe	Average Methylation	P Value	FDR	hg19 coordinates	Gene Location	Gene Symbol
cg07259683	0.325	1.76E-07	0.0237	chr15:37196490x-	Body	MEIS2
cg13649056	0.118	1.43E-07	0.0237	chr9:136474626-		
cg13877502	0.145	1.29E-07	0.0237	chr9:707387+	5'UTR	KANK1
cg16026813	0.088	1.95E-07	0.0237	chr10:103245720+	Body	BTRC
cg19648320	0.271	1.11E-07	0.0237	chr16:77387723-	Body	ADAMTS18
cg02574198	0.100	2.75E-07	0.0279	chr14:101544201-		
cg15730313	0.058	3.42E-07	0.0297	chr5:191628-	1stExon	LRRC14B
cg05043823	0.396	4.20E-07	0.0320	chr15:77951054+	TSS1500	LINGO1-AS2
cg04673462	0.171	1.05E-06	0.0639	chr1:38461896-		
cg04695372	0.101	1.04E-06	0.0639	chr17:77639509+		
cg14183712	0.311	1.40E-06	0.0708	chr7:157811918+	Body	PTPRN2
cg22813735	0.533	1.39E-06	0.0708	chr21:19528053+	5'UTR	CHODL
cg11147442	0.345	1.88E-06	0.0882	chr12:57210954-		
cg05835113	0.647	2.75E-06	0.1044	chr7:48294858+	Body	ABCA13
cg14431361	0.155	2.62E-06	0.1044	chr17:70611386+		
cg26842596	0.393	2.64E-06	0.1044	chr19:15121297-	TSS1500	CCDC105
cg10779909	0.063	3.37E-06	0.1078	chr9:706773+	TSS200	KANK1
cg12831439	0.271	3.21E-06	0.1078	chr8:55377905+		
cg13067553	0.166	3.17E-06	0.1078	chr1:147782452+		
cg12635602	0.170	3.75E-06	0.1110	chr11:58731357+		
cg14996491	0.088	4.02E-06	0.1110	chr1:42384284-	1stExon	HIVEP3
cg17870569	0.623	3.89E-06	0.1110	chr2:84513930+		
cg19341106	0.415	4.32E-06	0.1141	chr7:157811843+	Body	PTPRN2
cg00039147	0.394	4.95E-06	0.1204	chr17:35194628-		
cg05334416	0.253	4.87E-06	0.1204	chr15:37338491+	Body	MEIS2
cg00303982	0.046	8.76E-06	0.1213	chr8:145556952-	Body	SCRT1
cg01953068	0.154	1.01E-05	0.1213	chr1:18733763-		
cg02228185	0.137	1.18E-05	0.1213	chr17:3379567+	1stExon	ASPA
cg02487452	0.681	1.01E-05	0.1213	chr19:855899-	Body	ELANE
cg03358592	0.109	1.05E-05	0.1213	chr9:707410+	5'UTR	KANK1
cg04037640	0.358	1.03E-05	0.1213	chr6:32065034+	Body	TNXB
cg04371212	0.574	9.58E-06	0.1213	chr10:118563276-		
cg04482825	0.382	1.08E-05	0.1213	chr15:37190548-	Body	MEIS2
cg04693537	0.848	1.17E-05	0.1213	chr6:80580652+		
cg04743412	0.051	1.02E-05	0.1213	chr14:63641459-		
cg05570491	0.046	1.02E-05	0.1213	chr1:42384437-	TSS200	HIVEP3
cg06020989	0.400	7.77E-06	0.1213	chr9:21709125-		
cg06396682	0.476	9.77E-06	0.1213	chr17:72538243+	Body	CD300C
cg07064066	0.583	6.62E-06	0.1213	chr21:39493375+	TSS200	DSCR8
cg07166084	0.565	1.15E-05	0.1213	chr1:242217440-		
cg11715228	0.190	5.21E-06	0.1213	chr19:6464013+	TSS1500	CRB3
cg11726060	0.781	1.01E-05	0.1213	chr20:41614446+	Body	PTPRT
cg12729160	0.182	1.01E-05	0.1213	chr1:37875985+		
cg13066693	0.774	1.09E-05	0.1213	chr14:76878042-	5'UTR	ESRRB
cg13336045	0.787	6.75E-06	0.1213	chr11:124048399-		
cg13594863	0.870	8.50E-06	0.1213	chr11:17496264+	Body	ABCC8
cg14599998	0.071	1.11E-05	0.1213	chr2:65086822-		
cg15542777	0.258	5.70E-06	0.1213	chr20:62288970+	TSS200	RTEL1
cg16558362	0.811	7.29E-06	0.1213	chr1:204318846-	5'UTR	PLEKHA6
cg17201651	0.146	8.44E-06	0.1213	chr10:111683631+	TSS1500	XPNPEP1

cg18954541	0.049	1.14E-05	0.1213	chr20:45279985+	1stExon	SLC13A3
cg19446499	0.618	7.25E-06	0.1213	chr18:57671509+		
cg20710266	0.264	8.27E-06	0.1213	chr15:96877855+	Body	NR2F2
cg21493505	0.201	1.04E-05	0.1213	chr19:56879418-	TSS1500	ZNF542
cg23378923	0.728	8.06E-06	0.1213	chr8:80455644-		
cg23707097	0.071	7.60E-06	0.1213	chr10:104179659-	5'UTR	PSD
cg25923240	0.207	1.14E-05	0.1213	chr8:55372569+	3'UTR	SOX17
cg26949694	0.103	7.82E-06	0.1213	chr11:27742060-	Body	BDNF
cg27413462	0.035	1.02E-05	0.1213	chr11:14993866+	TSS200	CALCA
cg04988514	0.084	1.29E-05	0.1223	chr19:48947560+	3'UTR	GRIN2D
cg12688576	0.064	1.22E-05	0.1223	chr6:18387407+	TSS200	RNF144B
cg14308648	0.888	1.27E-05	0.1223	chr10:3568949-		
cg14696334	0.081	1.28E-05	0.1223	chr15:91500193+	Body	RCCD1
cg19163067	0.427	1.25E-05	0.1223	chr1:36042438-	Body	TFAP2E
cg06315909	0.859	1.40E-05	0.1280	chr17:76227508-	5'UTR	LOC283999
cg18533870	0.527	1.41E-05	0.1280	chr11:62650385-	ExonBnd	SLC3A2
cg22226398	0.663	1.40E-05	0.1280	chr1:235295459-	Body	RBM34
cg02198701	0.420	1.46E-05	0.1305	chr16:55364614-	3'UTR	IRX6
cg26704870	0.089	1.50E-05	0.1319	chr2:73144279-	TSS1500	EMX1
cg01164054	0.228	1.66E-05	0.1323	chr22:22959261+		
cg01472075	0.302	1.61E-05	0.1323	chr16:69984927+	5'UTR	CLEC18A
cg03636968	0.122	1.63E-05	0.1323	chr11:123526191+	TSS1500	SCN3B
cg10133284	0.372	1.68E-05	0.1323	chr16:57518223+	5'UTR	DOK4
cg15013171	0.337	1.55E-05	0.1323	chr22:50157867+		
cg17478763	0.053	1.60E-05	0.1323	chr19:13133351-	TSS1500	NFIX
cg17818196	0.075	1.64E-05	0.1323	chr19:6464013-	Body	TPM1
cg24928110	0.169	1.59E-05	0.1323	chr20:41614446+		
cg02483353	0.445	1.81E-05	0.1338	chr1:37875985+	5'UTR	STEAP2
cg02964324	0.710	1.83E-05	0.1338	chr14:76878042-	Body	HECW1
cg12263124	0.083	1.78E-05	0.1338	chr11:124048399-		
cg17672639	0.650	1.82E-05	0.1338	chr11:17496264+	Body	SLC25A37
cg24994249	0.173	1.79E-05	0.1338	chr2:65086822-	TSS1500	UBE4A
cg26293019	0.127	1.77E-05	0.1338	chr20:62288970+	Body	ISL1
cg04611801	0.085	1.95E-05	0.1350	chr1:204318846+	Body	BZRAP1
cg05192049	0.518	1.91E-05	0.1350	chr10:111683631-		
cg19863411	0.140	1.95E-05	0.1350	chr20:45279985+	TSS1500	PCYT2
cg24163575	0.128	1.92E-05	0.1350	chr18:57671509+		
cg24938286	0.222	1.93E-05	0.1350	chr15:96877855-	Body	MOG
cg06387622	0.337	2.00E-05	0.1352	chr19:56879418-	Body	LOC285768
cg12243007	0.098	1.99E-05	0.1352	chr8:80455644-	5'UTR	RPH3AL
cg00289249	0.851	2.14E-05	0.1368	chr10:104179659-	Body	MAU2
cg03490289	0.399	2.06E-05	0.1368	chr8:55372569+		
cg06299447	0.223	2.13E-05	0.1368	chr11:27742060+	TSS200	FBXO41
cg07547549	0.162	2.10E-05	0.1368	chr11:14993866-	Body	SLC12A5
cg17341345	0.232	2.07E-05	0.1368	chr19:48947560-	TSS1500	CCNH
cg00387200	0.438	2.18E-05	0.1370	chr6:18387407-		
cg02063488	0.244	2.24E-05	0.1370	chr10:3568949+	Body	FBXO41
cg20930114	0.076	2.28E-05	0.1370	chr15:91500193+	TSS1500	SEPT10
cg23115387	0.221	2.20E-05	0.1370	chr1:36042438-	1stExon	FOXE3
cg23717174	0.221	2.26E-05	0.1370	chr17:76227508+	Body	FAH
cg25318967	0.868	2.23E-05	0.1370	chr11:62650385-		
cg22890775	0.672	2.45E-05	0.1463	chr1:235295459-	5'UTR	RGS17
cg02216061	0.765	2.54E-05	0.1469	chr16:55364614-	TSS1500	HIRIP3
cg04214993	0.756	2.52E-05	0.1469	chr2:73144279+	TSS1500	LINC01609
cg12804006	0.493	2.52E-05	0.1469	chr22:22959261-	ExonBnd	ATP12A

cg02333645	0.165	2.62E-05	0.1471	chr16:69984927-	5'UTR	NCOA7
cg13576552	0.030	2.59E-05	0.1471	chr11:123526191-		
cg13684632	0.358	2.64E-05	0.1471	chr22:50157867-		
cg16416322	0.194	2.66E-05	0.1471	chr19:13133351-		
cg22378341	0.430	2.61E-05	0.1471	chr15:63336009-		
cg04828580	0.206	2.71E-05	0.1473	chr8:96085448+		
cg08240539	0.276	2.70E-05	0.1473	chr7:89844832-	TSS1500	FTSJ2
cg05748163	0.219	2.82E-05	0.1476	chr7:43291318+	Body	EBF2
cg07115820	0.193	2.88E-05	0.1476	chr14:101544258+	1stExon	EPX
cg07541744	0.290	2.91E-05	0.1476	chr8:23399008+	Body	IRX6
cg10166994	0.849	2.89E-05	0.1476	chr11:118230048+	Body	SBK2
cg14148981	0.444	2.81E-05	0.1476	chr5:50688913-		
cg17622024	0.830	2.81E-05	0.1476	chr17:56402029-		
cg17855264	0.031	2.90E-05	0.1476	chr13:112567087+		
cg21337909	0.416	2.88E-05	0.1476	chr17:79869801+	Body	TNXB
cg03242265	0.120	3.02E-05	0.1492	chr15:30261379+		
cg12427162	0.853	3.05E-05	0.1492	chr6:29631295-	Body	SFT2D2
cg12810297	0.232	3.07E-05	0.1492	chr6:1100823-		
cg18633600	0.106	2.99E-05	0.1492	chr17:184018+	Body	LRTM2
cg24405700	0.709	3.04E-05	0.1492	chr19:19457892+	Body	B4GALNT3
cg15740106	0.414	3.15E-05	0.1504	chr8:49729253+		
cg21975349	0.309	3.15E-05	0.1504	chr2:73496907-		
cg26531634	0.218	3.17E-05	0.1504	chr20:44658225+	TSS1500	CANT1
cg06087453	0.335	3.34E-05	0.1530	chr5:86709409-	Body	LINC01331
cg17859258	0.855	3.35E-05	0.1530	chr10:123442326-	Body	KIF21A
cg18675097	0.100	3.26E-05	0.1530	chr2:73496909+	5'UTR	NKAPL
cg19824242	0.381	3.30E-05	0.1530	chr2:110372285+		
cg20912272	0.395	3.34E-05	0.1530	chr1:47882936-		
cg06260527	0.069	3.48E-05	0.1567	chr15:80466823+	TSS1500	CSRNP2
cg15937073	0.062	3.48E-05	0.1567	chr10:72886986-	TSS200	HIVEP3
cg00429758	0.065	4.03E-05	0.1580	chr6:153439690-	TSS1500	ABCC10
cg01378878	0.083	3.85E-05	0.1580	chr16:30008249-	TSS200	GPX3
cg01479210	0.856	3.83E-05	0.1580	chr8:112249126-		
cg02307184	0.335	3.85E-05	0.1580	chr13:25262660-		
cg04047221	0.053	4.01E-05	0.1580	chr6:126179986+	1stExon	SOX1
cg04823169	0.082	3.97E-05	0.1580	chr9:100149720+	Body	TRIM14
cg05361406	0.127	3.72E-05	0.1580	chr1:118364092-	5'UTR	HSPB8
cg05368740	0.058	3.85E-05	0.1580	chr9:37603181+	Body	HSPB6
cg06339338	0.795	3.83E-05	0.1580	chr3:40621906-	Body	IFNAR2
cg07097417	0.100	3.87E-05	0.1580	chr10:93639778-	5'UTR	LPGAT1
cg11531339	0.378	3.98E-05	0.1580	chr7:2282600+		
cg11645413	0.528	3.55E-05	0.1580	chr8:25897763+	Body	FAM69A
cg14761454	0.514	3.99E-05	0.1580	chr17:56270249-	TSS1500	ATM
cg15173586	0.714	3.65E-05	0.1580	chr16:55363277+	Body	FAM91A1
cg15930643	0.336	3.70E-05	0.1580	chr19:56046689-	3'UTR	FLJ37543
cg16851139	0.357	3.59E-05	0.1580	chr14:52708188+	Body	SRC
cg17590951	0.055	3.90E-05	0.1580	chr4:53546720-		
cg18699242	0.066	3.57E-05	0.1580	chr15:37180636-	TSS1500	USP18
cg18753680	0.270	3.93E-05	0.1580	chr6:32063459+	TSS1500	JPH4
cg22115076	0.114	3.90E-05	0.1580	chr1:147782472-		
cg06498897	0.411	4.10E-05	0.1589	chr1:168198488+		
cg13155023	0.262	4.09E-05	0.1589	chr1:199717229+	Body	DISC1FP1
cg00329728	0.865	4.16E-05	0.1602	chr12:1940452+	5'UTR	GABRA5
cg00977526	0.291	4.43E-05	0.1605	chr12:655829-		

cg03546977	0.152	4.63E-05	0.1605	chr5:153885282-	TSS200	CMYA5
cg04478991	0.390	4.35E-05	0.1605	chr5:171916934+	TSS200	OSTalpha
cg06785646	0.317	4.67E-05	0.1605	chr17:77006381-		
cg06843147	0.841	4.30E-05	0.1605	chr5:73760703+		
cg06942685	0.152	4.51E-05	0.1605	chr12:39800627-	TSS1500	ZNF542
cg08717672	0.804	4.41E-05	0.1605	chr6:28227127+	Body	EML1
cg09546902	0.099	4.59E-05	0.1605	chr2:105273432+	Body	FOXO3
cg09674170	0.069	4.67E-05	0.1605	chr3:128765374-	TSS200	CLDN9
cg10299976	0.219	4.25E-05	0.1605	chr12:51477693-		
cg14205822	0.244	4.63E-05	0.1605	chr1:42384564+	ExonBnd	TBX18
cg14964181	0.198	4.34E-05	0.1605	chr6:43394948-	5'UTR	NUDT1
cg17155270	0.833	4.53E-05	0.1605	chr5:150399802+	5'UTR	MX1
cg17203647	0.102	4.38E-05	0.1605	chr10:91109674-	Body	EBF2
cg19550524	0.210	4.23E-05	0.1605	chr5:23821294+		
cg20699780	0.321	4.64E-05	0.1605	chr13:112723477+	Body	ABR
cg23445341	0.075	4.65E-05	0.1605	chr9:100850334+	Body	SBNO2
cg25320780	0.200	4.44E-05	0.1605	chr12:119616913+	5'UTR	SLC16A4
cg25415932	0.081	4.36E-05	0.1605	chr19:36246882+	TSS1500	OGG1
cg08114317	0.122	4.71E-05	0.1607	chr21:34620140-	TSS1500	KIAA1751
cg12501546	0.063	4.74E-05	0.1612	chr1:212002770-		
cg03140788	0.125	4.79E-05	0.1618	chr2:20069134-	Body	UMOD
cg01323542	0.247	4.82E-05	0.1620	chr1:93395042-	TSS200	AIRE
cg19641841	0.326	4.90E-05	0.1636	chr11:108092087+		
cg05631332	0.447	5.04E-05	0.1665	chr8:124784802+		
cg12159166	0.848	5.06E-05	0.1665	chr5:61001539-	Body	UVRAG
cg22481770	0.076	5.02E-05	0.1665	chr20:36024647-		
cg03295517	0.235	5.10E-05	0.1669	chr16:17112593-		
cg04194876	0.330	5.43E-05	0.1677	chr22:18632429-		
cg04983687	0.963	5.32E-05	0.1677	chr14:24048528-	Body	ZFPM1
cg08532057	0.121	5.44E-05	0.1677	chr5:50673588+	TSS1500	NUPL1
cg08915171	0.594	5.20E-05	0.1677	chr16:55241597+		
cg11551577	0.574	5.31E-05	0.1677	chr11:90372594+	TSS200	DSCR8
cg14399863	0.070	5.48E-05	0.1677	chr15:27113913-	1stExon	KANK1
cg15212525	0.371	5.49E-05	0.1677	chr2:107202311-		
cg15637159	0.813	5.47E-05	0.1677	chr5:78985489-		
cg17152389	0.168	5.22E-05	0.1677	chr3:195943198-	TSS1500	DDX17
cg19816125	0.240	5.49E-05	0.1677	chr15:36787793+	Body	PRKAR1B
cg20963059	0.809	5.17E-05	0.1677	chr2:169975296-		
cg24317217	0.258	5.42E-05	0.1677	chr19:56879432+		
cg26002422	0.111	5.41E-05	0.1677	chr14:100346149+		
cg01574513	0.660	5.60E-05	0.1691	chr6:108977419-		
cg11253122	0.698	5.57E-05	0.1691	chr16:3062382+		
cg17077180	0.308	5.63E-05	0.1691	chr10:94455543-		
cg20467929	0.077	5.65E-05	0.1691	chr6:85472244-	TSS1500	KIF19
cg04000684	0.894	5.73E-05	0.1708	chr7:2282529-	TSS1500	LOC101929439
cg03256036	0.795	5.80E-05	0.1711	chr21:42802197-	Body	SRGAP3
cg12711706	0.266	5.78E-05	0.1711	chr8:25897201-	TSS1500	LOC101927668
cg22471156	0.795	5.83E-05	0.1713	chr11:124043067-	Body	DAGLB
cg21733150	0.071	5.97E-05	0.1737	chr17:935017-	Body	SGIP1
cg27099280	0.061	5.95E-05	0.1737	chr19:1113357+	1stExon	CELF6
cg08189993	0.199	6.10E-05	0.1749	chr1:110933345-		
cg20252015	0.845	6.10E-05	0.1749	chr3:9791051+	Body	PGM1
cg23997421	0.867	6.09E-05	0.1749	chr1:1935535-	TSS1500	GTSF1L
cg03275306	0.673	6.18E-05	0.1761	chr19:13734932-	1stExon	BCL2A1
cg16179589	0.471	6.20E-05	0.1761	chr16:20359971-		

cg08635011	0.803	6.31E-05	0.1779	chr21:45705543-	Body	CUX2
cg13068215	0.316	6.32E-05	0.1779	chr14:65068527-	Body	FMN2
cg19544707	0.777	6.38E-05	0.1789	chr2:222826960+		
cg12640402	0.302	6.42E-05	0.1791	chr11:75649913+	1stExon	IL18RAP
cg05278780	0.183	6.47E-05	0.1795	chr3:156324118-		
cg06912623	0.749	6.52E-05	0.1795	chr5:173599404+		
cg20161471	0.198	6.52E-05	0.1795	chr14:64226259+	Body	TRPM8
cg20182111	0.111	6.59E-05	0.1805	chr16:88558223-	Body	ATP13A4
cg14945587	0.509	6.70E-05	0.1826	chr13:25875436-	Body	SYNDIG1
cg00017557	0.770	7.76E-05	0.1830	chr2:105273376-	5'UTR	C6orf52
cg00130808	0.410	9.94E-05	0.1830	chr21:39493391-	TSS200	LOC145845
cg00501467	0.185	1.03E-04	0.1830	chr9:707166+	Body	PHACTR1
cg00540570	0.385	7.90E-05	0.1830	chr12:70861459+	Body	SRC
cg00613752	0.105	1.10E-04	0.1830	chr1:93770564-	1stExon	GPR6
cg00689225	0.134	9.70E-05	0.1830	chr22:38903025-	Body	NFKBIA
cg00712342	0.740	9.70E-05	0.1830	chr7:612781+	TSS1500	TPM3
cg00727590	0.100	7.83E-05	0.1830	chr2:241118175-	TSS200	PLA2G3
cg00762681	0.402	8.69E-05	0.1830	chr3:70231495-	TSS1500	CCDC160
cg00888313	0.203	8.09E-05	0.1830	chr19:19488846+		
cg00919323	0.587	1.03E-04	0.1830	chr16:85981720-	Body	ATP6V1C2
cg01335180	0.537	1.09E-04	0.1830	chr1:20314957+	Body	HSPA7
cg01409343	0.789	7.81E-05	0.1830	chr1:38461687-	Body	TMEM49
cg01456368	0.044	7.22E-05	0.1830	chr17:72321958+	5'UTR	DNAH1
cg01708284	0.175	1.09E-04	0.1830	chr15:76474132+	Body	DMRT1
cg01862457	0.269	9.30E-05	0.1830	chr3:9154819-	Body	BCAS3
cg02023345	0.135	8.41E-05	0.1830	chr7:20181444+	Body	TMEM26
cg02070906	0.366	9.73E-05	0.1830	chr7:6473557-		
cg02166450	0.061	1.06E-04	0.1830	chr1:67029321+		
cg02394908	0.893	8.60E-05	0.1830	chr15:72612204-	3'UTR	ORAI2
cg02600660	0.763	1.03E-04	0.1830	chr22:39323967-	Body	SLC9A9
cg02638448	0.863	8.97E-05	0.1830	chr1:64100814+		
cg02860797	0.120	9.99E-05	0.1830	chr20:42356299+	TSS1500	OGG1
cg02880176	0.055	1.05E-04	0.1830	chr15:80263218+	TSS200	PLD6
cg02934525	0.173	9.54E-05	0.1830	chr17:36107292+		
cg02988811	0.280	7.84E-05	0.1830	chr12:111515293-		
cg03405983	0.295	1.09E-04	0.1830	chr1:240286481-	5'UTR	LYNX1
cg03425379	0.591	9.60E-05	0.1830	chr14:103746266+		
cg03614796	0.204	1.08E-04	0.1830	chr2:103035314-	Body	SMOC2
cg03652298	0.636	7.76E-05	0.1830	chr5:154496708+	Body	KCNK10
cg03668598	0.734	8.90E-05	0.1830	chr1:224781135+		
cg03731131	0.227	8.28E-05	0.1830	chr2:234891119-	Body	KCNN4
cg03879823	0.230	7.21E-05	0.1830	chr3:193217253+	Body	PHACTR1
cg03884592	0.038	7.14E-05	0.1830	chr20:24613506-	1stExon	HIVEP3
cg04149015	0.279	9.53E-05	0.1830	chr6:10687886-	Body	PRSS36
cg04169260	0.081	1.05E-04	0.1830	chr15:37178909-	TSS200	CTF1
cg04950461	0.655	1.02E-04	0.1830	chr6:13015202+		
cg05025860	0.441	7.61E-05	0.1830	chr20:36024669-		
cg05041720	0.789	7.49E-05	0.1830	chr6:110300308-	5'UTR	MID1
cg05358404	0.393	9.30E-05	0.1830	chr14:35871954+	TSS1500	RTEL1
cg05611362	0.273	9.76E-05	0.1830	chr1:154151224+	3'UTR	SOGA3
cg05815196	0.148	8.34E-05	0.1830	chr22:31536549-	TSS200	PLA2G3
cg06292784	0.487	1.02E-04	0.1830	chrX:133369578+	Body	PTPRT
cg06643679	0.469	9.79E-05	0.1830	chr19:5344050+		
cg07150255	0.752	7.62E-05	0.1830	chr2:10902658-	TSS1500	OR5V1

cg07212541	0.135	9.76E-05	0.1830	chr1:161576580+	Body	ALDH1A3
cg07467336	0.446	8.58E-05	0.1830	chr17:57915740-		
cg07573020	0.821	7.49E-05	0.1830	chr3:52351859+	3'UTR	ANKH
cg07805089	0.477	1.07E-04	0.1830	chr9:911252+	TSS1500	GATAD1
cg07844369	0.889	8.62E-05	0.1830	chr17:59199909+		
cg08023265	0.193	6.78E-05	0.1830	chr10:63212226-		
cg08242668	0.823	1.03E-04	0.1830	chr5:158869634-		
cg08615111	0.157	7.47E-05	0.1830	chr13:112712795-	Body	PLCH2
cg08721974	0.196	9.28E-05	0.1830	chr7:102087746+	TSS1500	CELF1
cg08727311	0.220	7.74E-05	0.1830	chr3:143081579-		
cg08852332	0.499	7.81E-05	0.1830	chr4:187764597-	Body	GRIK4
cg08940169	0.939	8.48E-05	0.1830	chr3:9791144+	Body	ZFPM1
cg08997362	0.222	9.33E-05	0.1830	chr17:17109817-	TSS200	ANKK1
cg09575222	0.692	1.03E-04	0.1830	chr4:117847521-	Body	LOC341056
cg09605287	0.203	1.04E-04	0.1830	chr1:205253947-	TSS200	HOXC-AS2
cg09881857	0.739	7.73E-05	0.1830	chr8:143858548-	Body	MIR548W
cg09929374	0.351	1.07E-04	0.1830	chr7:38734128+	TSS1500	DPYS
cg09988118	0.508	9.50E-05	0.1830	chr6:168853834-	Body	TBX18
cg09998546	0.823	7.75E-05	0.1830	chr14:88712112+		
cg10055566	0.252	7.02E-05	0.1830	chr8:61816898+	1stExon	NKAPL
cg10156217	0.118	9.52E-05	0.1830	chr19:44278274+	3'UTR	ARF6
cg10282491	0.198	9.62E-05	0.1830	chr6:13015221-	TSS200	CLDN9
cg10443049	0.052	7.96E-05	0.1830	chr1:42384474-	Body	NXPH1
cg10467747	0.359	9.07E-05	0.1830	chr16:31159854+		
cg10694470	0.219	9.76E-05	0.1830	chr16:30907890+	TSS200	PRDM8
cg11212329	0.357	1.08E-04	0.1830	chr2:25450598+		
cg11717194	0.054	1.09E-04	0.1830	chr4:115256327-		
cg11929163	0.594	1.02E-04	0.1830	chrX:10723977-	3'UTR	RPS3
cg12413881	0.824	9.94E-05	0.1830	chr20:62288950-		
cg12695707	0.413	1.04E-04	0.1830	chr6:127785848+	TSS1500	CCDC105
cg13334727	0.087	1.08E-04	0.1830	chr22:31536540+	TSS200	SEPT8
cg13480617	0.142	7.53E-05	0.1830	chr20:41595708-	TSS200	PRLR
cg13519586	0.876	9.01E-05	0.1830	chr11:45362186+		
cg13568213	0.417	7.96E-05	0.1830	chr6:29324414+	Body	TMEM8C
cg13688599	0.417	1.09E-04	0.1830	chr15:101442710+		
cg13738195	0.053	7.32E-05	0.1830	chr8:56753149-	TSS1500	DCHS2
cg13827984	0.176	1.06E-04	0.1830	chr5:14707001-		
cg13835436	0.522	7.46E-05	0.1830	chr7:92075926-		
cg14155350	0.160	6.93E-05	0.1830	chr16:31813373+		
cg14376041	0.832	9.49E-05	0.1830	chr1:31158138-		
cg14574489	0.123	7.27E-05	0.1830	chr6:64232389-	Body	SLC9A9
cg14847009	0.120	7.36E-05	0.1830	chr1:2435555+	TSS1500	KIAA0040
cg15273270	0.881	8.04E-05	0.1830	chr11:47575259+	Body	DTHD1
cg15609098	0.088	9.29E-05	0.1830	chr1:24542899-		
cg16146322	0.058	9.09E-05	0.1830	chr11:120765651+	TSS1500	FAM123C
cg16443148	0.055	7.78E-05	0.1830	chr16:88540241-	TSS200	CCDC78
cg16566518	0.032	8.22E-05	0.1830	chr11:113258469+	Body	FAM84B
cg17367518	0.212	1.08E-04	0.1830	chr11:122889971+	TSS1500	TMEM120A
cg17515932	0.693	6.88E-05	0.1830	chr12:54390571+		
cg17721068	0.102	8.71E-05	0.1830	chr17:60899289+	TSS200	CCDC17
cg17860198	0.123	7.39E-05	0.1830	chr8:105479674-		
cg17862113	0.076	1.05E-04	0.1830	chr6:85472307+	3'UTR	MYOD1
cg17984804	0.305	7.44E-05	0.1830	chr11:22543031+		
cg18087266	0.273	1.02E-04	0.1830	chr6:28227220+	TSS1500	RUNX3
cg18278486	0.097	6.91E-05	0.1830	chr14:50362354-	Body	LOC100270746

cg18357676	0.152	7.20E-05	0.1830	chr16:3062368-		
cg19014419	0.096	1.07E-04	0.1830	chr7:8481537-	TSS200	ZNF300
cg19322065	0.120	9.80E-05	0.1830	chr10:5418635+		
cg19698273	0.335	7.93E-05	0.1830	chr4:81118600+	Body	ATCAY
cg19796584	0.089	7.30E-05	0.1830	chr5:136002882+		
cg20010264	0.838	8.79E-05	0.1830	chr1:11990078-	5'UTR	SIL1
cg20546985	0.051	9.27E-05	0.1830	chr11:75129011-	TSS1500	MXRA7
cg20788020	0.765	1.07E-04	0.1830	chr6:170547477-		
cg20921659	0.483	9.88E-05	0.1830	chr19:15121333+	Body	ALK
cg20998200	0.255	7.24E-05	0.1830	chr5:132113601+	Body	SLC17A6
cg21306988	0.237	7.29E-05	0.1830	chr5:35230725-	5'UTR	LIG4
cg21960251	0.675	1.09E-04	0.1830	chr3:101753750+		
cg22253931	0.194	8.93E-05	0.1830	chr9:136387235-	5'UTR	MAPK4
cg22328837	0.858	8.52E-05	0.1830	chr20:2670413+	Body	RUNX2
cg22345063	0.178	8.35E-05	0.1830	chr4:155413469-		
cg23034420	0.455	8.09E-05	0.1830	chr7:64699983+		
cg23249562	0.701	1.01E-04	0.1830	chr11:69809807-	Body	OCLN
cg23419907	0.320	1.06E-04	0.1830	chr18:66084877+		
cg23440585	0.890	1.05E-04	0.1830	chr10:125736702-	Body	ELMO2
cg23527621	0.209	9.19E-05	0.1830	chr3:143021680-	Body	ECE2
cg23782083	0.843	8.44E-05	0.1830	chr1:175162515-	3'UTR	ASH2L
cg23856257	0.270	9.85E-05	0.1830	chr4:36324571+	Body	EPS8L2
cg23983371	0.225	9.80E-05	0.1830	chr7:6970760+	Body	OPCML
cg24125901	0.499	8.77E-05	0.1830	chr2:131513144-	Body	CRLF1
cg24229798	0.513	8.99E-05	0.1830	chr16:776667+	TSS200	GPR20
cg24374161	0.101	1.09E-04	0.1830	chr8:127569022-	5'UTR	AMBRA1
cg24615110	0.047	7.96E-05	0.1830	chr7:75624479-		
cg24850474	0.471	9.14E-05	0.1830	chr7:156190074-	Body	ATP13A1
cg24926361	0.136	9.59E-05	0.1830	chr1:46089850+	Body	GSC2
cg25503513	0.102	9.56E-05	0.1830	chr3:5138170-	TSS1500	DNAJC6
cg25538183	0.183	8.18E-05	0.1830	chr11:17743282+	TSS1500	SLC9A8
cg25645748	0.321	7.67E-05	0.1830	chr8:126913178+	TSS200	AKAP14
cg25783189	0.810	8.27E-05	0.1830	chr1:25257629+	TSS1500	FGR
cg25824543	0.199	7.60E-05	0.1830	chr6:26987575-	TSS200	ZNF175
cg26040401	0.773	8.14E-05	0.1830	chr14:105012309+		
cg26239297	0.405	1.06E-04	0.1830	chr5:150284504+	Body	ABHD12B
cg26791400	0.544	9.28E-05	0.1830	chr1:37701025-	Body	LINC01116
cg26866348	0.112	8.13E-05	0.1830	chr19:3908397-	TSS200	LRRC14B
cg27363750	0.299	9.60E-05	0.1830	chr17:80246233+		
cg27364650	0.166	1.00E-04	0.1830	chr5:138504070+		
cg27441048	0.050	9.91E-05	0.1830	chr17:74707485-	TSS1500	BCL7B
cg15158067	0.264	1.10E-04	0.1839	chr8:105677501-	TSS1500	C1orf233
cg00864424	0.318	1.11E-04	0.1840	chr2:29698784-	TSS1500	SCPEP1
cg08576208	0.078	1.13E-04	0.1846	chr11:22370099-		
cg12743978	0.094	1.12E-04	0.1846	chr13:108867134+	TSS200	CHODL
cg13197754	0.462	1.12E-04	0.1846	chr4:160081865+	Body	TESC
cg16286964	0.687	1.13E-04	0.1846	chr18:48181217+		
cg22255037	0.154	1.12E-04	0.1846	chr6:45422120-	TSS1500	FAM19A1
cg00976619	0.485	1.14E-04	0.1846	chr1:48191702-	Body	CCNL1
cg01156834	0.134	1.14E-04	0.1846	chr21:43434317+	TSS1500	SRCIN1
cg03160107	0.519	1.15E-04	0.1846	chr5:68826794-	TSS1500	SVOPL
cg04932082	0.085	1.13E-04	0.1846	chr2:220154319+		
cg13347820	0.819	1.14E-04	0.1846	chr20:45005588+		
cg22659765	0.730	1.14E-04	0.1846	chr3:183977549+		

cg26071755	0.083	1.15E-04	0.1846	chr8:37997055+	TSS1500	HTR2C
cg04930848	0.163	1.17E-04	0.1850	chr11:719828+		
cg09826217	0.686	1.17E-04	0.1850	chr11:133372412+	Body	CDH4
cg13160449	0.741	1.16E-04	0.1850	chr19:18716439+	Body	ARRB1
cg20213269	0.109	1.16E-04	0.1850	chr8:142377541-	TSS200	SAC3D1
cg25148589	0.098	1.15E-04	0.1850	chr11:46582058+	1stExon	GRIA2
cg01117274	0.642	1.17E-04	0.1856	chr5:74245169+	Body	SGCZ
cg19935326	0.078	1.18E-04	0.1859	chr19:19772686+		
cg00343514	0.217	1.22E-04	0.1859	chr22:19136922+		
cg02722204	0.857	1.22E-04	0.1859	chr1:65774850+	Body	LINC00908
cg04517429	0.159	1.22E-04	0.1859	chr20:48428907-	TSS1500	DCHS2
cg04753078	0.473	1.20E-04	0.1859	chrX:119029933+	Body	TNXB
cg06525317	0.181	1.23E-04	0.1859	chr1:27953220+		
cg06689785	0.268	1.22E-04	0.1859	chr19:52074410+	Body	DOK6
cg07847809	0.376	1.20E-04	0.1859	chr18:29156663-	Body	LOC101926942
cg08247465	0.822	1.21E-04	0.1859	chr14:51347240-	Body	ALK
cg09098195	0.117	1.24E-04	0.1859	chr2:177502087+	TSS1500	BEND4
cg11981204	0.622	1.23E-04	0.1859	chr5:191547+	Body	SORCS2
cg14441584	0.363	1.19E-04	0.1859	chr8:101872688+	TSS200	LOC101927911
cg14688652	0.081	1.23E-04	0.1859	chr3:43899898+	5'UTR	IL1RAPL2
cg17029694	0.284	1.23E-04	0.1859	chr7:72972316-	Body	PTK6
cg17698886	0.141	1.19E-04	0.1859	chr1:1536396-	TSS1500	PTPRS
cg20558369	0.718	1.23E-04	0.1859	chr17:55054810+	Body	ALOXE3
cg22850886	0.075	1.19E-04	0.1859	chr4:4858059+		
cg23226006	0.387	1.23E-04	0.1859	chr21:19617145+	5'UTR	PDP1
cg26314722	0.738	1.23E-04	0.1859	chr12:117529361-		
cg26766694	0.242	1.23E-04	0.1859	chr16:84971417+	Body	IQSEC1
cg01483824	0.750	1.29E-04	0.1869	chr3:68039681+	Body	GRIN2D
cg02400458	0.838	1.28E-04	0.1869	chr3:156876725-		
cg04600297	0.088	1.31E-04	0.1869	chr17:36763024+		
cg05991401	0.129	1.25E-04	0.1869	chr7:138350033+	Body	RAB37
cg06061408	0.778	1.30E-04	0.1869	chr6:91113506-	1stExon	GPR45
cg07605285	0.145	1.28E-04	0.1869	chr4:178612243-		
cg08061745	0.269	1.27E-04	0.1869	chr3:103991976-		
cg10365886	0.463	1.28E-04	0.1869	chrX:113817830+	Body	TNXB
cg11813441	0.256	1.29E-04	0.1869	chr2:53314927+		
cg12480390	0.628	1.31E-04	0.1869	chr20:60089107+	Body	KIRREL3
cg13095627	0.326	1.30E-04	0.1869	chr11:75036905+	TSS1500	IGFL4
cg13422751	0.348	1.29E-04	0.1869	chr11:64808357+	Body	CDH22
cg13698057	0.874	1.30E-04	0.1869	chr4:158141936+	TSS1500	FNDC5
cg13929657	0.163	1.26E-04	0.1869	chr8:15094114+	Body	LOC101927630
cg15697919	0.463	1.27E-04	0.1869	chr15:99123953+		
cg16653138	0.119	1.28E-04	0.1869	chr1:26730264+	1stExon	PRKCZ
cg18187198	0.282	1.29E-04	0.1869	chr18:74266122-		
cg19509675	0.079	1.26E-04	0.1869	chr4:155413618+		
cg19666541	0.208	1.29E-04	0.1869	chr6:32063619-		
cg20950812	0.202	1.29E-04	0.1869	chr18:20674968-	Body	IRX6
cg24302133	0.443	1.30E-04	0.1869	chr18:67485018-	3'UTR	CACNA1E
cg02079455	0.536	1.33E-04	0.1877	chr10:92197387+	Body	FBLN2
cg12148129	0.409	1.32E-04	0.1877	chr2:29532803-		
cg12784089	0.873	1.33E-04	0.1877	chr4:42155260+	Body	ARNT2
cg15597945	0.280	1.32E-04	0.1877	chr4:7423056+	Body	LRP6
cg26965687	0.258	1.32E-04	0.1877	chr17:2869230-	5'UTR	SMYD1
cg13001433	0.373	1.34E-04	0.1884	chrX:103812042-		
cg02208176	0.154	1.41E-04	0.1891	chr20:62166255+	TSS200	KANK1

cg02433458	0.516	1.41E-04	0.1891	chr19:5341265+	Body	LOC101928651
cg02533819	0.840	1.41E-04	0.1891	chr17:8012617-	Body	ATF7
cg03131724	0.162	1.35E-04	0.1891	chr1:26901676+		
cg03407679	0.764	1.38E-04	0.1891	chr8:94932053-	Body	HTR7
cg06117341	0.162	1.39E-04	0.1891	chr1:234867300-	1stExon	CLDN9
cg06493994	0.130	1.42E-04	0.1891	chr3:13060974-	1stExon	SCGN
cg07398767	0.126	1.39E-04	0.1891	chr19:48945886+	TSS1500	CA10
cg07978099	0.068	1.38E-04	0.1891	chr15:80624605+	5'UTR	OSGIN1
cg08189694	0.106	1.36E-04	0.1891	chr19:42750575+	Body	LOC100507424
cg08565703	0.575	1.42E-04	0.1891	chr17:72741152-		
cg09281405	0.081	1.36E-04	0.1891	chr2:105858511+	5'UTR	CLDN11
cg10121508	0.083	1.40E-04	0.1891	chr10:26680729-	TSS1500	GAS2L1
cg10224345	0.762	1.40E-04	0.1891	chr19:51231557-		
cg11574457	0.755	1.39E-04	0.1891	chr6:32063874-		
cg12573892	0.110	1.37E-04	0.1891	chr5:176543790+		
cg12744311	0.129	1.40E-04	0.1891	chr11:126796374-		
cg15231363	0.137	1.39E-04	0.1891	chr19:46544662+	Body	VSX1
cg18706316	0.186	1.37E-04	0.1891	chr20:44811586-	Body	NALCN
cg21559386	0.592	1.40E-04	0.1891	chr1:33339332+	TSS1500	HNF4A
cg21801870	0.822	1.42E-04	0.1891	chr7:17596110+	Body	AJAP1
cg22359581	0.802	1.35E-04	0.1891	chr16:86060428+	Body	EHD2
cg23048358	0.432	1.36E-04	0.1891	chr1:2005180-	TSS1500	GAS7
cg24130365	0.133	1.39E-04	0.1891	chr6:27570442-	Body	PKD1L2
cg26251865	0.280	1.38E-04	0.1891	chr14:57603592-	1stExon	IRGC
cg00051979	0.053	1.42E-04	0.1893	chr15:31516111-	5'UTR	C10orf4
cg00660167	0.093	1.43E-04	0.1896	chr16:55359615+		
cg13302217	0.376	1.43E-04	0.1899	chr1:181770173+	3'UTR	ZFPM1
cg20666917	0.201	1.44E-04	0.1901	chr3:13641006+	Body	C2orf55
cg01028226	0.059	1.45E-04	0.1904	chr2:177502274-		
cg11787646	0.225	1.45E-04	0.1904	chr15:80884783-	Body	EPHB2
cg22101814	0.104	1.45E-04	0.1904	chr12:12340008+	TSS1500	TTC39A
cg00982389	0.145	1.46E-04	0.1910	chr2:88367360+	ExonBnd	LOC654841
cg09116560	0.251	1.46E-04	0.1912	chr21:37326485-	TSS1500	PTPN5
cg04299200	0.206	1.48E-04	0.1919	chr9:706712+		
cg09250897	0.658	1.48E-04	0.1919	chr5:60960751+	Body	GMDS
cg16149607	0.674	1.48E-04	0.1919	chr12:53966263+		
cg24786671	0.201	1.48E-04	0.1919	chr11:58731433-	TSS1500	KRTAP23-1
cg27413739	0.882	1.48E-04	0.1919	chr10:92610352-	Body	FARP1
cg01464052	0.079	1.50E-04	0.1924	chr16:3062653+	TSS200	C1orf51
cg03015304	0.205	1.50E-04	0.1924	chr6:25652602+	Body	SHANK2
cg04295144	0.266	1.50E-04	0.1924	chr17:50237205-	Body	ICAM5
cg00177101	0.115	1.67E-04	0.1926	chr16:83986941+		
cg00453605	0.765	1.61E-04	0.1926	chr12:2950364+		
cg00472710	0.169	1.64E-04	0.1926	chr20:24853492+		
cg01206911	0.061	1.72E-04	0.1926	chr3:170136775+	TSS1500	PUS7
cg01660716	0.400	1.65E-04	0.1926	chr22:29702026+		
cg01788396	0.117	1.71E-04	0.1926	chr2:82461714-	TSS1500	CCDC28B
cg02586483	0.573	1.57E-04	0.1926	chr1:17051321-	TSS1500	MYADM
cg02806012	0.110	1.58E-04	0.1926	chr18:20674506+		
cg03002271	0.632	1.64E-04	0.1926	chr19:23386650+	Body	MYO3B
cg03075534	0.058	1.67E-04	0.1926	chr20:25058189+		
cg03556458	0.882	1.69E-04	0.1926	chr13:101978015+	Body	TGFBI
cg03734783	0.473	1.55E-04	0.1926	chr20:43029652+	1stExon	ACCN3
cg04617948	0.060	1.68E-04	0.1926	chr1:4792941-	TSS1500	FOXI2

cg04917197	0.066	1.58E-04	0.1926	chr19:48229109+		
cg05800416	0.222	1.68E-04	0.1926	chr17:9940652-	TSS200	CSGALNACT1
cg05950212	0.250	1.58E-04	0.1926	chr16:81222820-	TSS1500	HIST2H2AB
cg06157219	0.079	1.60E-04	0.1926	chr19:44220214+	5'UTR	PRKCZ
cg06290690	0.088	1.60E-04	0.1926	chr10:95462161+	Body	TRIM14
cg06501840	0.243	1.56E-04	0.1926	chr17:80200974+	Body	LINC01490
cg06558502	0.088	1.64E-04	0.1926	chr16:88601529-		
cg06603096	0.253	1.70E-04	0.1926	chr2:99439848-	Body	FADS6
cg06639875	0.647	1.67E-04	0.1926	chr1:16502584-	Body	GRK5
cg07124969	0.843	1.66E-04	0.1926	chr1:23055560-		
cg07158903	0.865	1.59E-04	0.1926	chr1:51811072-	Body	VPS13A
cg07236061	0.831	1.70E-04	0.1926	chr2:228093337-	Body	LINC01331
cg07405182	0.087	1.71E-04	0.1926	chr11:18814421+		
cg07527259	0.683	1.65E-04	0.1926	chr5:534332+	Body	CCDC168
cg08742290	0.745	1.72E-04	0.1926	chr6:2043676-	TSS1500	RNASET2
cg09799585	0.327	1.66E-04	0.1926	chr9:110918931+	Body	OLFML2B
cg10196453	0.326	1.51E-04	0.1926	chr21:31722001-	Body	KLRG2
cg10268345	0.478	1.62E-04	0.1926	chr13:98980912+	Body	PKD1
cg10452531	0.052	1.72E-04	0.1926	chr1:150255137+	TSS200	LTBR
cg10770023	0.273	1.67E-04	0.1926	chr11:70516824-	5'UTR	CPT1B
cg11623225	0.048	1.62E-04	0.1926	chr19:10407184-		
cg11942594	0.830	1.57E-04	0.1926	chr9:120545390+		
cg12058478	0.157	1.51E-04	0.1926	chr11:26330272-	Body	AJAP1
cg12323063	0.167	1.71E-04	0.1926	chr3:151178727+	Body	MAP2K6
cg12648537	0.028	1.52E-04	0.1926	chr7:105163001-	TSS1500	RGS22
cg13990958	0.194	1.53E-04	0.1926	chr14:105667315-	TSS1500	DAGLB
cg14046043	0.097	1.54E-04	0.1926	chr1:32665584+	TSS1500	GIGYF2
cg14221766	0.054	1.62E-04	0.1926	chr19:54367947+		
cg14390189	0.900	1.63E-04	0.1926	chr17:2649139-	Body	HEATR2
cg14564793	0.046	1.62E-04	0.1926	chr2:171312565+		
cg14972362	0.361	1.59E-04	0.1926	chr1:240161447-		
cg15897840	0.416	1.63E-04	0.1926	chr5:135387559-		
cg16076328	0.247	1.64E-04	0.1926	chr7:150745709-	Body	EPCAM
cg16204989	0.364	1.69E-04	0.1926	chr10:129535138-	Body	NALCN
cg16797300	0.055	1.72E-04	0.1926	chr10:112290331+		
cg17562163	0.739	1.57E-04	0.1926	chr8:19460097-	Body	LOC101928694
cg17576960	0.840	1.69E-04	0.1926	chr1:149860711+	Body	FMNL2
cg17710091	0.248	1.57E-04	0.1926	chr1:2063875-		
cg17792806	0.174	1.55E-04	0.1926	chr9:100850337+	TSS200	TCEAL7
cg17830985	0.027	1.67E-04	0.1926	chr12:81160899-	TSS200	SHISA2
cg18236734	0.153	1.58E-04	0.1926	chr19:57610699+	TSS200	HTR3E
cg18388270	0.340	1.57E-04	0.1926	chr17:72879149+		
cg18418570	0.243	1.59E-04	0.1926	chr10:121104212-		
cg18676350	0.852	1.52E-04	0.1926	chr16:88356689-	ExonBnd	TSC2
cg19897537	0.622	1.58E-04	0.1926	chr9:80008320-	TSS1500	MGAT4D
cg19961251	0.629	1.64E-04	0.1926	chr5:73720202+		
cg20735682	0.429	1.66E-04	0.1926	chr1:40598455-		
cg22084098	0.056	1.69E-04	0.1926	chr13:103389305+	TSS1500	MAPKBP1
cg22149516	0.058	1.69E-04	0.1926	chr6:167371091+	TSS200	SLCO4C1
cg22280475	0.110	1.51E-04	0.1926	chr1:161956573+	Body	EBF2
cg22979422	0.040	1.64E-04	0.1926	chr7:139138741-	TSS1500	CS
cg23762517	0.072	1.54E-04	0.1926	chr16:2140404+	1stExon	HIVEP3
cg24019564	0.201	1.50E-04	0.1926	chr12:6493278-	TSS1500	RUNX3
cg25077287	0.295	1.66E-04	0.1926	chr22:51016644-	Body	TRIM29
cg26005098	0.276	1.59E-04	0.1926	chr6:33216697+	Body	TSGA10IP

cg26192612	0.139	1.63E-04	0.1926	chr3:167659189-	Body	C6orf155
cg27506210	0.130	1.69E-04	0.1926	chr1:4718532-	5'UTR	DEPDC5
cg03781234	0.803	1.73E-04	0.1928	chr17:67497879+	Body	KIRREL3
cg06436488	0.234	1.73E-04	0.1928	chr8:101118463+		
cg22289837	0.064	1.73E-04	0.1931	chr7:6487999-	TSS1500	CA3
cg22749310	0.062	1.74E-04	0.1935	chr2:233561679+	Body	GPR120
cg00048759	0.113	1.81E-04	0.1939	chr9:36818572+	TSS200	STAG3
cg00276548	0.803	1.77E-04	0.1939	chr7:801377-		
cg01774131	0.703	1.79E-04	0.1939	chr9:98982105+		
cg02073162	0.188	1.79E-04	0.1939	chr3:153346517+		
cg02114946	0.067	1.76E-04	0.1939	chr5:135157736-	5'UTR	SLC7A4
cg02970679	0.844	1.81E-04	0.1939	chr2:47597118+	TSS1500	EPX
cg03607973	0.570	1.82E-04	0.1939	chr13:101830996+		
cg04339360	0.741	1.82E-04	0.1939	chr15:89942816+	Body	KLF5
cg07377519	0.052	1.82E-04	0.1939	chr15:58777264-	TSS1500	AP2A1
cg07397572	0.787	1.79E-04	0.1939	chr2:153296747+	Body	WWOX
cg07962581	0.748	1.82E-04	0.1939	chr1:48154068-	Body	CELF4
cg08564546	0.467	1.80E-04	0.1939	chrX:102585107-	Body	MAMDC2
cg09261636	0.620	1.78E-04	0.1939	chr13:26625359+	Body	NTNG1
cg09571713	0.112	1.78E-04	0.1939	chr3:183817931-	Body	SPOCK3
cg11527279	0.060	1.78E-04	0.1939	chr1:149212246-	Body	TTLL11
cg12253175	0.181	1.80E-04	0.1939	chr17:36176065+	TSS200	AGAP2
cg13282929	0.411	1.77E-04	0.1939	chr16:2104303+	5'UTR	SLC23A2
cg13387826	0.382	1.78E-04	0.1939	chr4:141420609+	1stExon	PRKCZ
cg17504042	0.677	1.75E-04	0.1939	chr6:12468608-		
cg20204986	0.154	1.80E-04	0.1939	chr6:168665386+	Body	WT1
cg20410290	0.917	1.82E-04	0.1939	chr15:42066199-		
cg20820767	0.289	1.79E-04	0.1939	chr5:101632314+	Body	RNF220
cg22904711	0.081	1.82E-04	0.1939	chr8:25897534-	Body	KCNN4
cg26004771	0.234	1.81E-04	0.1939	chr12:56694490-		
cg27250253	0.340	1.75E-04	0.1939	chr1:42384310-		
cg06792975	0.565	1.84E-04	0.1951	chr1:25257566-	5'UTR	MUC7
cg02958263	0.097	1.87E-04	0.1951	chr11:119991619-	Body	RABGAP1L
cg03001484	0.082	1.86E-04	0.1951	chr11:65714726+	TSS200	LMBRD1
cg06955856	0.130	1.87E-04	0.1951	chr6:72129509+	5'UTR	TCP11L2
cg08882472	0.594	1.87E-04	0.1951	chr22:32150725-	Body	DST
cg12685397	0.292	1.85E-04	0.1951	chr11:126506851-	5'UTR	OSBPL6
cg13692134	0.064	1.85E-04	0.1951	chr16:84560887+		
cg13980232	0.418	1.87E-04	0.1951	chr8:86350278-		
cg14249564	0.516	1.87E-04	0.1951	chr10:95327107+	Body	MICALCL
cg20000688	0.415	1.87E-04	0.1951	chr7:99775422+	Body	LINC01182
cg22633280	0.032	1.84E-04	0.1951	chr15:95216999+		
cg22733910	0.059	1.86E-04	0.1951	chr8:8578086-	TSS1500	PDE4B
cg06186150	0.082	1.89E-04	0.1955	chr8:86322830+	TSS1500	PANK4
cg09106510	0.265	1.89E-04	0.1955	chr22:21386494-		
cg16815476	0.215	1.89E-04	0.1955	chr17:56269818+		
cg25776467	0.351	1.89E-04	0.1955	chr16:81471588-	Body	MUC13
cg25984601	0.503	1.88E-04	0.1955	chr13:73635568-	Body	SLIT3
cg08147890	0.050	1.90E-04	0.1957	chr19:50269083-	Body	KBTBD2
cg13332816	0.050	1.91E-04	0.1965	chr16:79124403+	TSS1500	HSPA9
cg05039475	0.079	1.93E-04	0.1972	chr18:35103983-		
cg05444854	0.482	1.92E-04	0.1972	chr9:72762622-	Body	DMBX1
cg08093959	0.769	1.94E-04	0.1972	chr1:108017535-		
cg11235799	0.376	1.93E-04	0.1972	chr4:168154775-		

cg11862846	0.148	1.93E-04	0.1972	chr9:124854846+		
cg16297171	0.309	1.92E-04	0.1972	chr12:58132093+	Body	MATN2
cg00024377	0.089	2.34E-04	0.1974	chr20:4954404-	Body	DLGAP1
cg00062245	0.095	2.23E-04	0.1974	chr1:2005260-	Body	ITGA2B
cg00206463	0.294	2.14E-04	0.1974	chr16:54209282-	5'UTR	TRIM26
cg00560831	0.342	2.10E-04	0.1974	chr11:32448067+		
cg00840960	0.058	2.07E-04	0.1974	chr17:75786172-	TSS1500	HTR4
cg00917437	0.093	2.24E-04	0.1974	chr1:45082840-	TSS1500	KIAA0831
cg00963674	0.859	2.03E-04	0.1974	chr19:44278628+	Body	KDM5A
cg00983948	0.288	1.98E-04	0.1974	chr1:178456093+		
cg01020752	0.314	2.32E-04	0.1974	chr13:50182026-		
cg01365639	0.360	2.18E-04	0.1974	chr4:71337873+		
cg01420492	0.123	2.27E-04	0.1974	chr1:174769548+		
cg01424901	0.220	2.22E-04	0.1974	chr6:70507230+	Body	RAP1GAP
cg01679262	0.663	2.00E-04	0.1974	chr12:106698101+	5'UTR	IPO5
cg01695672	0.889	2.25E-04	0.1974	chr6:56607099+	Body	ZNF311
cg02870741	0.529	2.02E-04	0.1974	chr2:179142589+	Body	CIB3
cg02939404	0.288	2.11E-04	0.1974	chr5:74350132+	TSS1500	MYF6
cg03537215	0.493	1.95E-04	0.1974	chr6:40243983+	TSS1500	CNTN2
cg03788567	0.828	2.34E-04	0.1974	chr11:12340382+	TSS1500	RPL6
cg03982897	0.162	2.13E-04	0.1974	chr4:13923646+		
cg04032578	0.149	2.12E-04	0.1974	chr10:110671907+	TSS200	ZNF300
cg04058122	0.134	2.05E-04	0.1974	chr1:66257553-	TSS1500	MFN2
cg04211115	0.092	2.11E-04	0.1974	chr1:2458298-	1stExon	HS3ST3A1
cg04302433	0.108	1.98E-04	0.1974	chr11:71050395-		
cg04420569	0.649	2.42E-04	0.1974	chr7:13064237-		
cg04706201	0.698	2.02E-04	0.1974	chr3:124647658-	Body	C5orf56
cg04711474	0.292	2.42E-04	0.1974	chr5:168271898-	TSS200	CACNG6
cg05036656	0.107	2.43E-04	0.1974	chr7:32919201-		
cg05079049	0.419	2.05E-04	0.1974	chr5:137911760+		
cg05166530	0.885	1.98E-04	0.1974	chr3:35680791+		
cg05209461	0.782	2.31E-04	0.1974	chr1:46977292-		
cg05341384	0.140	2.19E-04	0.1974	chr2:106959853-		
cg05467676	0.457	2.02E-04	0.1974	chr17:74371760+		
cg05514713	0.417	2.36E-04	0.1974	chr7:2527816-	Body	PHACTR1
cg05567646	0.089	2.41E-04	0.1974	chr8:99017690-	TSS200	COQ3
cg06307913	0.235	2.07E-04	0.1974	chr18:3666162+	5'UTR	PRDM8
cg06837426	0.184	2.06E-04	0.1974	chr17:42462260-	1stExon	ESM1
cg06917617	0.064	1.96E-04	0.1974	chr6:30179435-	Body	CR1L
cg06978661	0.166	2.43E-04	0.1974	chrX:38084656-	Body	LINC00689
cg07068406	0.797	1.98E-04	0.1974	chr5:148034030+	Body	EPHB1
cg07318247	0.600	2.25E-04	0.1974	chr14:55879046+		
cg07370221	0.160	2.10E-04	0.1974	chr12:402444-	Body	GALNT9
cg07538447	0.592	2.10E-04	0.1974	chr6:158937485-	Body	DLC1
cg07702888	0.344	2.44E-04	0.1974	chr6:20303425-	TSS1500	RHBG
cg07746514	0.408	2.32E-04	0.1974	chr10:50975940-	3'UTR	MLEC
cg07796220	0.317	2.35E-04	0.1974	chr20:37349817+	1stExon	MFSD6L
cg07813134	0.704	2.43E-04	0.1974	chr1:21948680-		
cg07816439	0.176	2.28E-04	0.1974	chr13:98606196+	TSS200	C1QTNF2
cg07992196	0.158	2.42E-04	0.1974	chr6:28963009-	Body	LOC285419
cg08215400	0.299	2.46E-04	0.1974	chr19:16272369-	Body	CD163L1
cg08255598	0.518	2.34E-04	0.1974	chr12:81100762+	TSS1500	PCGF1
cg08314989	0.129	2.45E-04	0.1974	chr1:205011262-	5'UTR	B3GALNT1
cg08384240	0.275	2.11E-04	0.1974	chr12:112847769+	Body	LYN
cg08444004	0.874	2.42E-04	0.1974	chr11:1990108-	Body	HDAC4

cg08583701	0.170	2.25E-04	0.1974	chr5:150284619+		
cg08603868	0.176	2.09E-04	0.1974	chr1:12039793-		
cg08618604	0.157	2.47E-04	0.1974	chr17:13504008-	5'UTR	PPAPDC3
cg08623193	0.078	2.23E-04	0.1974	chr10:134330970-	Body	ARHGAP21
cg09012411	0.826	2.27E-04	0.1974	chr19:51774264-		
cg09363375	0.624	2.41E-04	0.1974	chr5:131773770-		
cg09521623	0.324	2.40E-04	0.1974	chr19:54495468-	1stExon	GPM6B
cg09847153	0.179	2.29E-04	0.1974	chr4:41875470+		
cg09864050	0.693	2.35E-04	0.1974	chr16:55364843+		
cg09891302	0.910	2.00E-04	0.1974	chr1:32427280-		
cg10148231	0.354	2.39E-04	0.1974	chr3:49389917-		
cg10219192	0.202	2.28E-04	0.1974	chr16:86331661+		
cg10452534	0.291	1.98E-04	0.1974	chr3:72226965+	1stExon	ELMO3
cg10549944	0.123	2.20E-04	0.1974	chr6:13166210-	Body	LOC101929268
cg10664316	0.321	2.46E-04	0.1974	chr6:99842244+		
cg11362622	0.113	2.24E-04	0.1974	chr4:81118794-	Body	BTBD11
cg11410305	0.207	2.42E-04	0.1974	chr5:54281271-	5'UTR	C10orf71
cg11519672	0.277	2.14E-04	0.1974	chr1:207818818-	Body	SPG11
cg11735242	0.818	2.32E-04	0.1974	chr7:158808509+		
cg11769885	0.043	2.25E-04	0.1974	chr3:134647778-	TSS200	HIVEP3
cg11804805	0.207	2.30E-04	0.1974	chr15:98703755-	Body	UQCRB
cg12043145	0.510	2.09E-04	0.1974	chr12:132904796-		
cg12077142	0.848	2.44E-04	0.1974	chr8:13115701-	Body	FBXO16
cg12078738	0.171	2.42E-04	0.1974	chr1:156338793-	TSS200	CGA
cg12091498	0.120	2.44E-04	0.1974	chr12:121139042-		
cg12308216	0.177	2.44E-04	0.1974	chr17:8702369+	5'UTR	DDR1
cg12560128	0.431	2.09E-04	0.1974	chr12:55329199-	TSS1500	OAS2
cg12613839	0.267	1.99E-04	0.1974	chr5:159797845-	Body	ADAMTS2
cg12616923	0.802	2.24E-04	0.1974	chr4:124671519-		
cg12800962	0.359	2.01E-04	0.1974	chr12:7592407-		
cg12816198	0.395	2.09E-04	0.1974	chr2:74735875-	TSS1500	IRF5
cg12969644	0.086	2.03E-04	0.1974	chr3:160822444-	TSS200	RASEF
cg13555278	0.278	2.29E-04	0.1974	chr8:56854528+	1stExon	EXTL1
cg13676220	0.169	2.22E-04	0.1974	chr2:240086178-	Body	C9orf171
cg13734043	0.245	2.40E-04	0.1974	chr6:101744310+	1stExon	EPX
cg13833632	0.327	2.34E-04	0.1974	chr11:107590917-		
cg13866093	0.359	2.40E-04	0.1974	chr9:134165109-	3'UTR	GPX5
cg14040722	0.633	2.35E-04	0.1974	chr10:25009527-	TSS1500	C20orf95
cg14358492	0.725	2.40E-04	0.1974	chr1:95167783-		
cg14397918	0.804	1.99E-04	0.1974	chr6:1247581-	Body	APBA1
cg14540387	0.201	2.47E-04	0.1974	chrX:13835295+	Body	BCL2L10
cg14594362	0.166	2.14E-04	0.1974	chr1:38461540-	3'UTR	GRIN2D
cg14732015	0.309	2.03E-04	0.1974	chr7:19150174-		
cg14782841	0.698	2.38E-04	0.1974	chr6:44305562+	Body	CACNG8
cg15166311	0.045	2.45E-04	0.1974	chr7:849153+		
cg15175162	0.205	2.39E-04	0.1974	chr5:180542623+	TSS1500	FBXL5
cg15199678	0.087	2.32E-04	0.1974	chr16:67233247-	Body	DLX6AS
cg15303560	0.779	2.30E-04	0.1974	chr8:49494374+	TSS200	RASGRP4
cg15406486	0.067	2.09E-04	0.1974	chr19:31117613+		
cg15578140	0.193	2.09E-04	0.1974	chr12:107836508-	Body	MIR548F3
cg15729869	0.073	2.27E-04	0.1974	chr10:50508713-	TSS200	BARHL1
cg15999590	0.383	2.41E-04	0.1974	chr15:44870890+	TSS1500	MLX
cg16196928	0.231	2.08E-04	0.1974	chr1:4644030-		
cg16274899	0.696	2.15E-04	0.1974	chr1:42384491+	5'UTR	PLEKHG3

cg16518861	0.063	2.39E-04	0.1974	chr8:97246246-	Body	NFKBIA
cg16717727	0.505	2.36E-04	0.1974	chr14:97205952+	Body	CDH19
cg16970761	0.845	2.36E-04	0.1974	chr8:28300399-	TSS200	PLD4
cg17034781	0.821	1.99E-04	0.1974	chr6:87804962+	Body	UST
cg17044200	0.841	2.13E-04	0.1974	chr12:47226255+	Body	MATN2
cg17154586	0.087	2.41E-04	0.1974	chr6:30853255-		
cg17284119	0.701	2.29E-04	0.1974	chr12:113416061-	Body	C6orf186
cg17410930	0.549	2.03E-04	0.1974	chr5:178741205+		
cg17518550	0.797	2.19E-04	0.1974	chr8:128228107-	TSS200	C1orf150
cg17886106	0.244	2.24E-04	0.1974	chr20:48626437+		
cg18115040	0.033	2.11E-04	0.1974	chr7:128577593-	TSS200	HOXD10
cg18352516	0.046	2.33E-04	0.1974	chr9:85678242+		
cg19112186	0.198	2.07E-04	0.1974	chr1:26348484-	5'UTR	CPT1B
cg19187564	0.321	2.22E-04	0.1974	chr9:135333078-	TSS200	CEACAM4
cg19251564	0.114	2.27E-04	0.1974	chr17:56270198-		
cg19425603	0.107	1.95E-04	0.1974	chr10:64038330+	Body	PCDH7
cg19702785	0.081	2.29E-04	0.1974	chr6:28502727-	Body	KCNS1
cg19843271	0.699	2.12E-04	0.1974	chr20:37229509+	TSS1500	BUD13
cg20005074	0.149	2.17E-04	0.1974	chr9:91454488-		
cg20329510	0.112	2.47E-04	0.1974	chr9:72078829-	TSS200	LIMS2
cg20369111	0.405	2.46E-04	0.1974	chr15:52404232-	Body	NTHL1
cg20447047	0.155	2.12E-04	0.1974	chr19:48947631-	Body	AGXT2
cg20584905	0.067	2.24E-04	0.1974	chr1:244397867-	TSS200	PAX3
cg20887417	0.875	1.98E-04	0.1974	chr19:54475129+		
cg20962117	0.067	2.38E-04	0.1974	chr1:178007283-		
cg21244516	0.470	2.39E-04	0.1974	chr4:15657657+	Body	NAV2
cg21350984	0.085	2.34E-04	0.1974	chr7:96642510+		
cg21564495	0.242	1.99E-04	0.1974	chr19:38917025+	Body	FAM19A5
cg21627412	0.047	2.44E-04	0.1974	chr8:144162594+	TSS1500	DZIP1
cg21725716	0.120	2.14E-04	0.1974	chr7:147718109-		
cg21820954	0.741	1.99E-04	0.1974	chr9:135457839+		
cg21880843	0.902	2.01E-04	0.1974	chr17:40718810-	Body	GREB1
cg21955446	0.642	2.47E-04	0.1974	chr2:177799828+	Body	SCD5
cg21961149	0.382	2.14E-04	0.1974	chr14:65175225-		
cg22152931	0.835	2.47E-04	0.1974	chr14:35871715-	Body	PRKRIP1
cg22169327	0.764	2.36E-04	0.1974	chr18:64217671-		
cg22287711	0.835	2.06E-04	0.1974	chr14:105391100-	Body	NRG2
cg22307471	0.522	2.24E-04	0.1974	chr6:149079585+	Body	NELL1
cg22382805	0.255	2.07E-04	0.1974	chr8:98903197-	Body	COL4A2
cg22656404	0.196	2.04E-04	0.1974	chr17:37753185-		
cg22694975	0.365	2.01E-04	0.1974	chr6:110644424+		
cg23867647	0.260	1.97E-04	0.1974	chr12:83021797+	Body	HSPB6
cg23933759	0.741	2.16E-04	0.1974	chr1:247712383-	5'UTR	CLDN14
cg24147489	0.752	2.27E-04	0.1974	chr9:92097197-	Body	HGD
cg24247132	0.308	2.16E-04	0.1974	chr2:176981328+		
cg24277926	0.837	2.39E-04	0.1974	chr3:177315226-		
cg24642169	0.119	2.44E-04	0.1974	chr22:51016638-		
cg25092052	0.140	2.16E-04	0.1974	chr19:42133492+		
cg25945004	0.900	2.35E-04	0.1974	chr6:166418922-	Body	TFDP1
cg26005232	0.079	2.44E-04	0.1974	chr4:30838902-	TSS1500	NUDC
cg26243278	0.083	2.22E-04	0.1974	chr20:43727089-		
cg26353287	0.568	2.06E-04	0.1974	chr11:116644707+		
cg26413672	0.253	2.40E-04	0.1974	chr7:35794235+		
cg26451824	0.834	2.29E-04	0.1974	chr2:128422210+	TSS1500	AMOTL2
cg26930408	0.359	2.29E-04	0.1974	chr16:2090610-	Body	FAM170B

cg27325694	0.682	2.35E-04	0.1974	chr5:35032783+		
cg27577859	0.161	2.25E-04	0.1974	chr2:223163784-	5'UTR	SETD3
cg27648238	0.277	2.11E-04	0.1974	chr12:12182089+		
cg16811201	0.734	2.48E-04	0.1978	chr9:17907485-	Body	HSD17B3
cg04769341	0.177	2.49E-04	0.1979	chr11:19515089+	1stExon	TCEAL7
cg20859918	0.816	2.49E-04	0.1980	chr12:95779645-		
cg03568842	0.144	2.50E-04	0.1981	chr22:49023277-	Body	NEURL1B
cg12527286	0.501	2.50E-04	0.1981	chr13:96297281-		
cg27520536	0.281	2.50E-04	0.1981	chr19:1725535-		
cg00047694	0.082	2.53E-04	0.1986	chr16:84295650-	Body	LOC732275
cg00567703	0.099	2.53E-04	0.1986	chr2:11744869+	5'UTR	HOXC4
cg03671627	0.104	2.54E-04	0.1986	chr4:83698365-	Body	PIKFYVE
cg05156901	0.242	2.52E-04	0.1986	chr15:37396505-	5'UTR	CPT1B
cg12380429	0.874	2.54E-04	0.1986	chr7:102065448+	Body	PC
cg13375188	0.369	2.53E-04	0.1986	chr3:86283923+		
cg14541541	0.807	2.52E-04	0.1986	chr5:139282860+	5'UTR	PRR5L
cg17343120	0.680	2.54E-04	0.1986	chr11:21166228+	Body	ATG16L1
cg17722888	0.695	2.52E-04	0.1986	chr13:110961594-	Body	MLLT1
cg16875863	0.106	2.54E-04	0.1987	chr8:119878888+	3'UTR	GRIN2D
cg00530093	0.351	2.65E-04	0.1990	chr1:172925965+		
cg00616135	0.211	2.58E-04	0.1990	chr19:36246816+	TSS1500	LACTB
cg00902464	0.889	2.62E-04	0.1990	chr21:37836671+		
cg00968616	0.815	2.62E-04	0.1990	chr3:120361491-	3'UTR	CUEDC1
cg01895008	0.214	2.65E-04	0.1990	chr7:64458642-	Body	KLHL41
cg02042600	0.096	2.66E-04	0.1990	chr12:125713569-	5'UTR	SNAP23
cg02544836	0.598	2.64E-04	0.1990	chr1:212731669+	Body	KCTD5
cg02630694	0.064	2.67E-04	0.1990	chr10:74014582-	TSS1500	CDC123
cg04518426	0.806	2.66E-04	0.1990	chr13:114265942+	Body	HDC
cg05309081	0.764	2.58E-04	0.1990	chr1:27247829-	Body	NCOR1
cg06726390	0.065	2.65E-04	0.1990	chr12:125628211+	TSS1500	VPS16
cg07290269	0.182	2.68E-04	0.1990	chr5:42954090-	TSS1500	FAM60A
cg07388347	0.113	2.69E-04	0.1990	chr15:102084204-		
cg07782669	0.827	2.68E-04	0.1990	chr3:134094802+	TSS1500	LOC101926908
cg09172260	0.327	2.66E-04	0.1990	chr10:50340100-		
cg09433509	0.129	2.56E-04	0.1990	chr5:172940317-		
cg09444685	0.090	2.67E-04	0.1990	chr14:99945149+	Body	ADRA1B
cg10195561	0.534	2.66E-04	0.1990	chr8:120685601+	TSS1500	TRIM34
cg11894504	0.158	2.60E-04	0.1990	chr9:99035862-	Body	MKNK2
cg12119693	0.717	2.69E-04	0.1990	chrX:102585202-	Body	WDR45L
cg12400702	0.316	2.58E-04	0.1990	chr5:127174575+	Body	CAND2
cg13493526	0.172	2.67E-04	0.1990	chr5:172086305+	1stExon	C14orf4
cg14230851	0.644	2.65E-04	0.1990	chr7:53466279-		
cg14305907	0.161	2.62E-04	0.1990	chr16:54228323-	5'UTR	KIF12
cg14653977	0.232	2.62E-04	0.1990	chr16:86377053+	5'UTR	GBGT1
cg14692377	0.035	2.69E-04	0.1990	chr12:54413101-	1stExon	SLC6A4
cg15264354	0.139	2.67E-04	0.1990	chr2:209185667+		
cg15440524	0.138	2.62E-04	0.1990	chr22:51016646+	TSS200	KCCAT211
cg16066272	0.071	2.58E-04	0.1990	chr11:66638774-	TSS200	NID2
cg16333198	0.351	2.62E-04	0.1990	chr10:102650389+		
cg16563255	0.257	2.61E-04	0.1990	chr11:36382508-	5'UTR	ZFHX3
cg17966227	0.312	2.68E-04	0.1990	chr2:234176772+		
cg19650019	0.581	2.60E-04	0.1990	chr19:6261277-		
cg20005350	0.127	2.57E-04	0.1990	chr19:48947572+		
cg20173993	0.309	2.69E-04	0.1990	chr10:131820725-	Body	CSF3R

cg20389954	0.380	2.65E-04	0.1990	chr15:63413614+		
cg21529539	0.188	2.59E-04	0.1990	chr1:101094117+	Body	SLC24A3
cg21721681	0.090	2.64E-04	0.1990	chr17:55942898-	3'UTR	DBP
cg24288916	0.184	2.64E-04	0.1990	chr2:170370674-		
cg24394120	0.375	2.62E-04	0.1990	chr15:42800833+		
cg24892659	0.128	2.55E-04	0.1990	chr16:2737220+	Body	NSMCE1
cg25236749	0.155	2.69E-04	0.1990	chr10:12236833+		
cg25885280	0.212	2.55E-04	0.1990	chr15:50541121-	Body	SHANK2
cg26188212	0.727	2.58E-04	0.1990	chr17:15990620+		
cg27342941	0.762	2.59E-04	0.1990	chr20:2820927+	Body	IFNAR2
cg00005297	0.108	2.70E-04	0.1991	chr12:31479853-	Body	DOK7
cg02797577	0.140	2.71E-04	0.1992	chr12:133481466-	TSS200	GPR78
cg24479859	0.732	2.71E-04	0.1992	chr8:73643832+		
cg25132772	0.906	2.71E-04	0.1992	chr1:1536101+	Body	CAMTA1
cg17010394	0.063	2.72E-04	0.1996	chr13:100611085-	5'UTR	TTLL7
cg25123670	0.173	2.72E-04	0.1996	chr5:159355481+		
cg06269342	0.235	2.73E-04	0.1998	chr11:5640272-	Body	MLLT1
cg12062133	0.345	2.73E-04	0.1998	chr19:2042391-		
cg27239280	0.151	2.73E-04	0.1998	chr17:80581925+	TSS1500	CTF1
cg00983520	0.081	2.74E-04	0.2000	chr3:12838781-	1stExon	CPT1B
cg10207408	0.647	2.74E-04	0.2000	chr14:77492788-	1stExon	FSCB

Appendix P – dmCpGs associated with Age

Probe	Average Methylation	P Value	FDR	hg19 coordinates	Gene Location	Gene Symbol
cg09476997	0.784	1.11E-10	0.0001	chr16:2087932+	Body	SLC9A3R2
cg15552843	0.296	4.68E-09	0.0014	chr22:24890809+	TSS200	C22orf45
cg12492885	0.377	3.97E-08	0.0081	chr22:24890814+	TSS200	C22orf45
cg17826389	0.793	8.73E-08	0.0133	chr16:2088995+	3'UTR	SLC9A3R2
cg05194102	0.864	1.37E-07	0.0167	chr16:2088960+	3'UTR	SLC9A3R2
cg02469909	0.346	2.81E-07	0.0244	chr22:24890831+	TSS200	C22orf45
cg01346158	0.783	2.78E-07	0.0244	chr16:2088997+	3'UTR	SLC9A3R2
cg03716592	0.909	3.21E-07	0.0244	chr19:55606078-	Body	PPP1R12C
cg26577790	0.780	4.10E-07	0.0278	chr22:40103865-		
cg14818279	0.839	4.65E-07	0.0283	chr5:148442554+	1stExon	SH3TC2
cg19337857	0.723	6.45E-07	0.0358	chr8:129830804+		
cg08462078	0.774	7.11E-07	0.0361	chr12:109871143+	Body	MYO1H
cg27431877	0.680	1.25E-06	0.0586	chr12:124911924+	Body	NCOR2
cg18200219	0.705	1.44E-06	0.0626	chr8:70360215-	Body	LINC01603
cg02102832	0.800	1.55E-06	0.0630	chr16:1581834-	Body	IFT140
cg13592980	0.105	2.33E-06	0.0808	chr19:2032733+		
cg03364108	0.225	3.45E-06	0.0808	chr22:24890794+	TSS200	C22orf45
cg06003986	0.133	4.10E-06	0.0808	chr18:35104868+	Body	BRUNOL4
cg22051740	0.295	4.39E-06	0.0808	chr16:58535358+	Body	NDRG4
cg02966574	0.161	4.42E-06	0.0808	chr15:68872082+	Body	CORO2B
cg17112695	0.784	2.56E-06	0.0808	chr12:58128450-	Body	AGAP2
cg11324379	0.851	2.57E-06	0.0808	chr19:31676224-		
cg26430984	0.842	2.93E-06	0.0808	chr5:177708457+	Body	COL23A1
cg07484727	0.689	3.26E-06	0.0808	chr16:1587309+	Body	IFT140
cg10195146	0.850	3.39E-06	0.0808	chr4:3473737-	Body	DOK7
cg20638896	0.853	3.72E-06	0.0808	chr8:2004621+	Body	MYOM2
cg09378940	0.052	3.76E-06	0.0808	chr8:71581306-	1stExon	LACTB2
cg23077884	0.750	3.78E-06	0.0808	chr2:157395835+	Body	GPD2
cg04147875	0.837	3.81E-06	0.0808	chr16:1587610-	Body	IFT140
cg11145643	0.872	3.81E-06	0.0808	chr14:100839818-	5'UTR	WARS
cg00520876	0.747	3.91E-06	0.0808	chr16:1207981+	Body	CACNA1H
cg07805940	0.902	4.08E-06	0.0808	chr12:130909013-	Body	RIMBP2
cg11755933	0.849	4.20E-06	0.0808	chr11:67266128-	Body	PITPNM1
cg20890260	0.788	4.57E-06	0.0808	chr8:20113038+	TSS1500	LZTS1
cg09479956	0.848	4.64E-06	0.0808	chr2:236649951+	Body	AGAP1
cg00008254	0.151	5.50E-06	0.0884	chr1:12676745+	Body	DHRS3
cg16988810	0.148	5.70E-06	0.0884	chr6:31762650+	Body	VARS
cg07232063	0.538	6.38E-06	0.0884	chr1:86953050-	Body	CLCA1
cg21549434	0.905	5.72E-06	0.0884	chr5:177708304+	Body	COL23A1
cg09972661	0.745	5.75E-06	0.0884	chr12:21919124+	Body	KCNJ8
cg03948287	0.582	5.89E-06	0.0884	chr4:62406882+	Body	ADGRL3
cg23543912	0.822	6.08E-06	0.0884	chr18:74667616+	Body	ZNF236
cg05276829	0.831	6.26E-06	0.0884	chr2:54861832-	Body	SPTBN1
cg04878869	0.697	6.27E-06	0.0884	chr12:131917579+		
cg03503516	0.393	6.76E-06	0.0903	chr22:24890811+	TSS200	C22orf45
cg10187583	0.632	6.82E-06	0.0903	chr11:117815479+		
cg09440559	0.189	7.22E-06	0.0936	chr1:16576546-	3'UTR	FBXO42
cg04239375	0.099	7.56E-06	0.0960	chr9:107526778+	Body	NIPSNAP3B
cg03078593	0.814	7.99E-06	0.0994	chr12:26789311-	Body	ITPR2

cg03119780	0.877	8.32E-06	0.1014	chr10:22886226+	Body	PIP4K2A
cg21810935	0.868	8.58E-06	0.1026	chr15:62149916+	Body	VPS13C
cg23312375	0.838	8.86E-06	0.1038	chr13:114814024+	Body	RASA3
cg13391812	0.713	9.09E-06	0.1045	chr3:120402723+	TSS1500	HGD
cg01451760	0.188	9.41E-06	0.1051	chr1:59362552+	Body	LINC01135
cg10174643	0.741	9.59E-06	0.1051	chr2:218680226-	Body	TNS1
cg01058902	0.817	9.66E-06	0.1051	chr1:201707751-	TSS1500	NAV1
cg19040750	0.764	1.15E-05	0.1209	chr1:205483697+	5'UTR	CDK18
cg02157636	0.634	1.15E-05	0.1209	chr11:68476835+	Body	MTL5
cg07637085	0.279	1.18E-05	0.1221	chr20:30148787+	Body	HM13
cg10706843	0.816	1.25E-05	0.1265	chr2:128335454+	Body	MYO7B
cg11826475	0.159	1.40E-05	0.1284	chr12:120704006-	TSS1500	PXN
cg05032848	0.110	1.45E-05	0.1284	chr17:33390854+	TSS200	RFFL
cg05128414	0.176	1.59E-05	0.1284	chr1:197882836+	Body	LHX9
cg13907146	0.328	1.60E-05	0.1284	chr3:50243565+	5'UTR	SLC38A3
cg22721472	0.658	1.64E-05	0.1284	chr19:31765880+	3'UTR	TSHZ3
cg12061176	0.142	1.67E-05	0.1284	chr13:43891253+	Body	ENOX1
cg14061378	0.239	1.69E-05	0.1284	chr9:139539092-		
cg22572217	0.851	1.29E-05	0.1284	chr12:6234790+	TSS1500	VWF
cg05595588	0.859	1.36E-05	0.1284	chr16:58149339-	Body	CFAP20
cg08440987	0.782	1.37E-05	0.1284	chr11:128479784+		
cg18140596	0.674	1.42E-05	0.1284	chr2:217720268+		
cg15383827	0.861	1.50E-05	0.1284	chr22:38423466-	Body	POLR2F
cg10710472	0.922	1.55E-05	0.1284	chr1:8394347-	Body	SLC45A1
cg05006473	0.842	1.56E-05	0.1284	chr4:1504481+		
cg02429905	0.495	1.60E-05	0.1284	chr6:32119944-	TSS1500	PRRT1
cg16060379	0.860	1.61E-05	0.1284	chr7:47366289-	Body	TNS3
cg06768939	0.830	1.63E-05	0.1284	chr17:79024447+	Body	BAIAP2
cg23570810	0.549	1.64E-05	0.1284	chr11:315102+	Body	IFITM1
cg05723552	0.775	1.66E-05	0.1284	chr16:57026029-	5'UTR	NLRC5
cg00087546	0.746	1.68E-05	0.1284	chr1:3308286+	Body	PRDM16
cg27423261	0.054	1.71E-05	0.1284	chr18:2572145+	5'UTR	NDC80
cg25105276	0.734	1.75E-05	0.1300	chr8:128509516+		
cg03334072	0.637	1.82E-05	0.1333	chr7:150435779-	5'UTR	GIMAP5
cg23647410	0.769	1.84E-05	0.1335	chr1:221051066-		
cg17004353	0.623	1.93E-05	0.1382	chr11:3073441+	Body	CARS
cg24813176	0.224	1.99E-05	0.1396	chr2:219736167+	Body	WNT6
cg05408686	0.860	1.99E-05	0.1396	chr16:57416050+	Body	CX3CL1
cg16444411	0.872	2.04E-05	0.1396	chr4:152065724-	Body	SH3D19
cg16762283	0.786	2.07E-05	0.1396	chr15:88737177-	Body	NTRK3
cg09315878	0.832	2.08E-05	0.1396	chr1:1152580+	3'UTR	SDF4
cg11286341	0.886	2.08E-05	0.1396	chr8:26448913+	Body	DPYSL2
cg04936838	0.794	2.25E-05	0.1410	chr6:43993067-	Body	LOC101929705
cg11476317	0.798	2.25E-05	0.1410	chr13:24476792-		
cg15909006	0.715	2.26E-05	0.1410	chr16:2086421-	Body	SLC9A3R2
cg00221794	0.730	2.27E-05	0.1410	chr2:20776620+		
cg26631178	0.898	2.28E-05	0.1410	chr17:45773680+	1stExon	TBKBP1
cg14236577	0.829	2.30E-05	0.1410	chr11:126606270+	Body	KIRREL3
cg19605950	0.867	2.30E-05	0.1410	chr9:124542057+	Body	DAB2IP
cg04210113	0.857	2.32E-05	0.1410	chr1:186676183+		
cg26606147	0.809	2.34E-05	0.1410	chr6:36297851-	ExonBnd	C6orf222
cg08732561	0.657	2.35E-05	0.1410	chr10:121027241+	Body	GRK5
cg08481464	0.535	2.36E-05	0.1410	chr10:77871618-	Body	C10orf11
cg26825643	0.818	2.40E-05	0.1419	chr17:47046294-	TSS1500	GIP
cg23251468	0.905	2.51E-05	0.1461	chr18:56639990-	Body	ZNF532

cg14453864	0.852	2.52E-05	0.1461	chr5:36648779-	Body	SLC1A3
cg12642651	0.762	2.56E-05	0.1473	chr17:9861712-	Body	GAS7
cg11585670	0.830	2.63E-05	0.1490	chr1:177136968-		
cg23779331	0.054	2.66E-05	0.1490	chr17:4614312-	Body	ARRB2
cg12277278	0.145	2.67E-05	0.1490	chr20:61636767-		
cg10665789	0.863	2.69E-05	0.1490	chr1:226921868+	Body	ITPKB
cg09012411	0.827	2.75E-05	0.1512	chr1:95167783-		
cg02960654	0.755	2.79E-05	0.1518	chr2:148421642-		
cg24232662	0.841	2.83E-05	0.1524	chr12:100583529+	TSS200	MIR1827
cg08929398	0.240	2.89E-05	0.1544	chr15:68497730-	5'UTR	CALML4
cg11926430	0.617	2.92E-05	0.1548	chr1:209883607+	Body	HSD11B1
cg09915396	0.105	3.03E-05	0.1575	chr17:2907895+	Body	RAP1GAP2
cg26924177	0.860	3.05E-05	0.1575	chrX:68117048+		
cg19048532	0.732	3.07E-05	0.1575	chr7:27148002+	Body	HOXA3
cg07885326	0.852	3.08E-05	0.1575	chr22:49609918-		
cg14290576	0.346	3.13E-05	0.1575	chr9:94181612-	5'UTR	NFIL3
cg25833890	0.793	3.13E-05	0.1575	chr20:24449136+	TSS1500	SYNDIG1
cg00537196	0.626	3.17E-05	0.1585	chr14:52688271-		
cg01992765	0.780	3.24E-05	0.1603	chr21:45622493-		
cg21686213	0.731	3.27E-05	0.1603	chr11:315118+	3'UTR	IFITM1
cg11656163	0.826	3.29E-05	0.1603	chr4:139482559+		
cg13713253	0.874	3.35E-05	0.1621	chrX:71128586+		
cg00796466	0.211	3.44E-05	0.1632	chr2:43386412+		
cg09485430	0.782	3.45E-05	0.1632	chr3:64039543+		
cg13701803	0.745	3.45E-05	0.1632	chr5:142120421-		
cg16960876	0.860	3.48E-05	0.1632	chr20:30076034+		
cg06814493	0.185	3.54E-05	0.1639	chr1:3585499+	5'UTR	TP73
cg10802700	0.919	3.56E-05	0.1639	chr17:80790467-	Body	TBCD
cg00709728	0.858	3.58E-05	0.1639	chr3:124778476-		
cg17339776	0.768	3.68E-05	0.1673	chr4:8202708+	5'UTR	SH3TC1
cg17975128	0.917	3.78E-05	0.1705	chr4:141678884-	TSS1500	TBC1D9
cg03840467	0.897	3.87E-05	0.1722	chr17:75278162+	Body	Sep-09
cg24008238	0.751	3.88E-05	0.1722	chr6:168366528+	Body	MLLT4
cg06805323	0.825	3.91E-05	0.1722	chr6:31634821+	Body	CSNK2B
cg27390669	0.640	3.94E-05	0.1722	chr20:24404156-		
cg02341271	0.717	3.97E-05	0.1722	chr18:21960369-	5'UTR	OSBPL1A
cg00888336	0.722	4.00E-05	0.1722	chr16:80654763-	Body	CDYL2
cg13549428	0.822	4.01E-05	0.1722	chr2:38366942+	Body	CYP1B1-AS1
cg03029371	0.739	4.20E-05	0.1789	chr20:61287281-	5'UTR	SLCO4A1
cg22005304	0.218	4.24E-05	0.1789	chr1:17944807+	Body	ARHGEF10L
cg01335210	0.874	4.26E-05	0.1789	chr10:130422176+		
cg03692239	0.727	4.29E-05	0.1791	chr1:3240386+	Body	PRDM16
cg09395844	0.897	4.34E-05	0.1792	chr11:36020643+	Body	LDLRAD3
cg27052709	0.764	4.35E-05	0.1792	chr17:76495017+	Body	DNAH17
cg25580383	0.886	4.39E-05	0.1797	chr1:218726598-	Body	MIR548F3
cg18390528	0.859	4.44E-05	0.1804	chr11:126294509-	Body	KIRREL3
cg20429150	0.755	4.48E-05	0.1809	chr14:103623601-		
cg04742719	0.280	4.60E-05	0.1824	chr7:100463759-	Body	SLC12A9
cg21116900	0.176	4.63E-05	0.1824	chr12:100750760+	TSS200	SLC17A8
cg13772158	0.824	4.60E-05	0.1824	chr2:127878592-		
cg09496391	0.639	4.64E-05	0.1824	chr8:701366+		
cg22390040	0.341	4.80E-05	0.1828	chr8:17220621-	Body	MTMR7
cg21477285	0.165	4.88E-05	0.1828	chr1:3087072-	Body	PRDM16
cg16389705	0.202	4.89E-05	0.1828	chr5:139051131-	5'UTR	CXXC5

cg00146659	0.911	4.78E-05	0.1828	chr4:7532516-	Body	SORCS2
cg24711094	0.885	4.80E-05	0.1828	chr11:72941203+	5'UTR	P2RY2
cg25481491	0.812	4.81E-05	0.1828	chr12:1904523+	Body	CACNA2D4
cg26550337	0.773	4.82E-05	0.1828	chr18:77203542+	Body	NFATC1
cg08066631	0.863	4.87E-05	0.1828	chr11:2850353-	Body	KCNQ1
cg23963984	0.082	5.03E-05	0.1834	chr10:79659038+	Body	DLG5
cg21511972	0.751	5.00E-05	0.1834	chr2:119234738+		
cg17746707	0.841	5.02E-05	0.1834	chr1:9048966+		
cg14698025	0.668	5.06E-05	0.1834	chr10:1245418-	Body	ADARB2
cg13104874	0.755	5.08E-05	0.1834	chr16:2079201+	Body	SLC9A3R2
cg06464885	0.752	5.09E-05	0.1834	chr5:111553674+	Body	EPB41L4A
cg05481201	0.795	5.19E-05	0.1855	chr7:1585452+	3'UTR	TMEM184A
cg00304423	0.565	5.20E-05	0.1855	chr22:22875030+	TSS1500	ZNF280A
cg22344703	0.219	5.26E-05	0.1857	chr2:219736312-	Body	WNT6
cg01179469	0.690	5.28E-05	0.1857	chr11:126709252+	Body	KIRREL3
cg11813198	0.788	5.32E-05	0.1857	chr12:57572413+	Body	LRP1
cg25651984	0.804	5.36E-05	0.1857	chrX:140991592+	TSS200	MAGEC1
cg03337918	0.861	5.38E-05	0.1857	chr8:96148290-	5'UTR	PLEKHF2
cg21545013	0.835	5.39E-05	0.1857	chr1:8407891-		
cg04084354	0.187	5.50E-05	0.1873	chr1:56721795+		
cg26143874	0.767	5.53E-05	0.1873	chr18:76363295+		
cg16411445	0.899	5.56E-05	0.1873	chr21:40145923-	TSS1500	NCRNA00114
cg02216727	0.741	5.56E-05	0.1873	chr17:38520653-	1stExon	GJD3
cg21924650	0.274	5.68E-05	0.1903	chr16:766396+	Body	METRN
cg06585307	0.660	5.79E-05	0.1927	chr16:1581015+	Body	IFT140
cg03785432	0.280	5.87E-05	0.1931	chr8:41656940-	Body	ANK1
cg17987384	0.601	5.88E-05	0.1931	chr17:26634469+	TSS200	FLJ40504
cg21714849	0.876	5.91E-05	0.1931	chr6:126413224+		
cg25051510	0.764	5.92E-05	0.1931	chr12:124672621+		
cg08092331	0.068	5.97E-05	0.1936	chr12:121019201+	TSS200	POP5
cg03505776	0.201	6.07E-05	0.1946	chr2:85034469-	Body	DNAH6
cg05579613	0.231	6.06E-05	0.1946	chr10:44144736+	TSS1500	ZNF32
cg06892281	0.830	6.17E-05	0.1960	chr20:30282950-	Body	BCL2L1
cg24271593	0.623	6.18E-05	0.1960	chr1:92395225-		
cg22670501	0.707	6.22E-05	0.1965	chr14:97367749+		
cg22351187	0.269	6.30E-05	0.1979	chr12:52586089+	TSS1500	KRT80

Appendix Q – dmCpGs associated with fasting serum insulin levels

Probe	Average Methylation	P Value	FDR	hg19 coordinates	Gene Location	Gene Symbol
cg12046168	0.026	3.53E-08	0.0206	chr10:23384686+	Body	MSRB2
cg19643314	0.032	1.12E-07	0.0326	chr1:205600634+	5'UTR	ELK4
cg09253462	0.812	2.30E-07	0.0339	chr19:20727168+	3'UTR	ZNF737
cg06339716	0.876	2.36E-07	0.0339	chr1:65045551+	Body	MIR4794
cg12518898	0.024	2.89E-07	0.0339	chr10:23384664+	Body	MSRB2
cg22713339	0.862	5.17E-07	0.0504	chr2:32465387-	Body	NLRC4
cg26502984	0.746	9.12E-07	0.0762	chrX:134306770-	TSS1500	CXorf48
cg16158641	0.023	1.13E-06	0.0829	chr11:122526738+	5'UTR	UBASH3B
cg12733590	0.841	2.09E-06	0.1161	chr1:248782815+		
cg07341489	0.896	2.37E-06	0.1161	chr8:141768559-	Body	PTK2
cg22888140	0.904	2.55E-06	0.1161	chr6:76070431+	Body	FILIP1
cg17863743	0.179	2.78E-06	0.1161	chr10:100993587-	Body	HPSE2
cg05127295	0.911	3.41E-06	0.1161	chr5:54965611-	Body	SLC38A9
cg15520688	0.874	3.55E-06	0.1161	chr14:106912909+		
cg16825290	0.078	3.55E-06	0.1161	chr10:100993826+	Body	HPSE2
cg09511126	0.146	3.67E-06	0.1161	chr10:100993597+	Body	HPSE2
cg00270448	0.883	3.73E-06	0.1161	chr17:41408386-		
cg02650323	0.868	3.95E-06	0.1161	chr8:37593312+	TSS1500	ERLIN2
cg19459056	0.787	4.13E-06	0.1161	chr22:36014303-	TSS1500	MB
cg01655006	0.877	4.30E-06	0.1161	chr18:22126203+		
cg10784720	0.872	4.36E-06	0.1161	chr14:106916597+		
cg04600798	0.695	4.37E-06	0.1161	chr8:145640322-	Body	SLC39A4
cg24849133	0.793	4.81E-06	0.1224	chrX:24708803-		
cg00760938	0.750	5.05E-06	0.1233	chr2:217357159+		
cg17510003	0.067	5.49E-06	0.1272	chr11:131737914+	Body	NTM
cg03619790	0.789	5.83E-06	0.1272	chr2:234941312-		
cg13737332	0.173	5.87E-06	0.1272	chr10:100993583-	Body	HPSE2
cg20829683	0.916	6.67E-06	0.1394	chr7:139416404-	Body	HIPK2
cg07697887	0.881	7.82E-06	0.1578	chr11:134607652-		
cg19428722	0.160	9.89E-06	0.1929	chr10:100994478-	Body	HPSE2
cg01649456	0.912	1.08E-05	0.1989	chr14:105519407-	Body	GPR132
cg08927006	0.237	1.11E-05	0.1989	chr10:100993556-	Body	HPSE2
cg10444806	0.131	1.14E-05	0.1989	chr17:38084428+	TSS1500	ORMDL3
cg11494358	0.908	1.17E-05	0.1989	chr1:248757080-	TSS200	OR2T10
cg04836492	0.829	1.24E-05	0.1989	chr1:248790459-	TSS200	OR2T11
cg25638978	0.938	1.28E-05	0.1989	chr11:134605124-		
cg03031183	0.893	1.28E-05	0.1989	chr14:106916953-		
cg24926711	0.322	1.32E-05	0.1989	chr10:100993662-	Body	HPSE2
cg11456854	0.883	1.33E-05	0.1989	chr2:43584626-	Body	THADA

References

1 Millward, D. J., Garlick, P. J., Stewart, R. J., Nnanyelugo, D. O. & Waterlow, J. C. *1975* Skeletal-muscle growth and protein turnover. *Biochem J* **150**(2), 235-243.

2 DeFronzo, R. A. & Tripathy, D. *2009* Skeletal muscle insulin resistance is the primary defect in type 2 diabetes. *Diabetes Care* **32 Suppl 2**S157-163.

3 Thiebaud, D., Jacot, E., DeFronzo, R. A., Maeder, E., Jequier, E. & Felber, J. P. *1982* The effect of graded doses of insulin on total glucose uptake, glucose oxidation, and glucose storage in man. *Diabetes* **31**(11), 957-963.

4 Dominguez, R. *2004* Actin-binding proteins--a unifying hypothesis. *Trends Biochem Sci* **29**(11), 572-578.

5 Behrmann, E., Muller, M., Penczek, P. A., Mannherz, H. G., Manstein, D. J. & Raunser, S. *2012* Structure of the rigor actin-tropomyosin-myosin complex. *Cell* **150**(2), 327-338.

6 Frontera, W. R. & Ochala, J. *2015* Skeletal muscle: a brief review of structure and function. *Calcified tissue international* **96**(3), 183-195.

7 Pavlath, G. K., Rich, K., Webster, S. G. & Blau, H. M. *1989* Localization of muscle gene products in nuclear domains. *Nature* **337**(6207), 570-573.

8 Rotundo, R. L. *1990* Nucleus-specific translation and assembly of acetylcholinesterase in multinucleated muscle cells. *The Journal of cell biology* **110**(3), 715-719.

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

10 Relaix, F. & Zammit, P. S. *2012* Satellite cells are essential for skeletal muscle regeneration: the cell on the edge returns centre stage. *Development* **139**(16), 2845-2856.

11 Hoppeler, H., Luthi, P., Claassen, H., Weibel, E. R. & Howald, H. *1973* The ultrastructure of the normal human skeletal muscle. A morphometric analysis on untrained men, women and well-trained orienteers. *Pflugers Arch* **344**(3), 217-232.

12 Bayol, S. A., Bruce, C. R. & Wadley, G. D. *2014* Growing healthy muscles to optimise metabolic health into adult life. *J Dev Orig Health Dis* **5**(6), 420-434.

13 Schiaffino, S. & Reggiani, C. *2011* Fiber types in mammalian skeletal muscles. *Physiol Rev* **91**(4), 1447-1531.

14 Gomes, A. V., Potter, J. D. & Szczesna-Cordary, D. *2002* The role of troponins in muscle contraction. *IUBMB Life* **54**(6), 323-333.

15 Tonino, P., Kiss, B., Strom, J., Methawasin, M., Smith, J. E., 3rd, Kolb, J., Labeit, S. & Granzier, H. *2017* The giant protein titin regulates the length of the striated muscle thick filament. *Nat Commun* **8**(1), 1041.

16 Ottenheijm, C. A., Granzier, H. & Labeit, S. *2012* The sarcomeric protein nebulin: another multifunctional giant in charge of muscle strength optimization. *Front Physiol* **3**37.

17 Costa, M. L., Escaleira, R., Cataldo, A., Oliveira, F. & Mermelstein, C. S. *2004* Desmin: molecular interactions and putative functions of the muscle intermediate filament protein. *Braz J Med Biol Res* **37**(12), 1819-1830.

18 Rossi, A. E. & Dirksen, R. T. *2006* Sarcoplasmic reticulum: the dynamic calcium governor of muscle. *Muscle & nerve* **33**(6), 715-731.

19 Block, B. A., Imagawa, T., Campbell, K. P. & Franzini-Armstrong, C. *1988* Structural evidence for direct interaction between the molecular components of the transverse tubule/sarcoplasmic reticulum junction in skeletal muscle. *The Journal of cell biology* **107**(6 Pt 2), 2587-2600.

20 Short, K. R., Bigelow, M. L., Kahl, J., Singh, R., Coenen-Schimke, J., Raghavakaimal, S. & Nair, K. S. *2005* Decline in skeletal muscle mitochondrial function with aging in humans. *Proceedings of the National Academy of Sciences of the United States of America* **102**(15), 5618-5623.

21 Porter, M. M., Vandervoort, A. A. & Lexell, J. *1995* Aging of human muscle: structure, function and adaptability. *Scandinavian journal of medicine & science in sports* **5**(3), 129-142.

22 Weisleder, N., Brotto, M., Komazaki, S., Pan, Z., Zhao, X., Nosek, T., Parness, J., Takeshima, H. & Ma, J. *2006* Muscle aging is associated with compromised Ca2+ spark signaling and segregated intracellular Ca2+ release. *The Journal of cell biology* **174**(5), 639-645.
23 Galpin, A. J., Raue, U., Jemiolo, B., Trappe, T. A., Harber, M. P., Minchev, K. & Trappe, S. *2012* Human skeletal muscle fiber type specific protein content. *Anal Biochem* **425**(2), 175-182.
24 Stienen, G. J., Kiers, J. L., Bottinelli, R. & Reggiani, C. *1996* Myofibrillar ATPase activity in skinned human skeletal muscle fibres: fibre type and temperature dependence. *The Journal of physiology* **493 (Pt 2)**299-307.
25 Gollnick, P. D., Sjodin, B., Karlsson, J., Jansson, E. & Saltin, B. *1974* Human soleus muscle: a comparison of fiber composition and enzyme activities with other leg muscles. *Pflugers Arch* **348**(3), 247-255.
26 Johnson, M. A., Polgar, J., Weightman, D. & Appleton, D. *1973* Data on the distribution of fibre types in thirty-six human muscles. An autopsy study. *Journal of the neurological sciences* **18**(1), 111-129.
27 Calderón, J. C., Bolaños, P. & Caputo, C. . *2014* The excitation–contraction coupling mechanism in skeletal muscle. *Biophysical Reviews* **6**(1), 133-160.
28 Rebbeck, R. T., Karunasekara, Y., Board, P. G., Beard, N. A., Casarotto, M. G. & Dulhunty, A. F. *2014* Skeletal muscle excitation-contraction coupling: who are the dancing partners? *Int J Biochem Cell Biol* **48**28-38.
29 Baylor, S. M. & Hollingworth, S. *2012* Intracellular calcium movements during excitation-contraction coupling in mammalian slow-twitch and fast-twitch muscle fibers. *J Gen Physiol* **139**(4), 261-272.
30 Lee, Y. S., Ondrias, K., Duhl, A. J., Ehrlich, B. E. & Kim, D. H. *1991* Comparison of calcium release from sarcoplasmic reticulum of slow and fast twitch muscles. *J Membr Biol* **122**(2), 155-163.
31 Geeves, M. A. *1991* The dynamics of actin and myosin association and the crossbridge model of muscle contraction. *Biochem J* **274 (Pt 1)**1-14.
32 Buckingham, M. *2001* Skeletal muscle formation in vertebrates. *Curr Opin Genet Dev* **11**(4), 440-448.
33 Borycki, A. G., Brunk, B., Tajbakhsh, S., Buckingham, M., Chiang, C. & Emerson, C. P., Jr. *1999* Sonic hedgehog controls epaxial muscle determination through Myf5 activation. *Development* **126**(18), 4053-4063.
34 Epstein, J. A., Shapiro, D. N., Cheng, J., Lam, P. Y. & Maas, R. L. *1996* Pax3 modulates expression of the c-Met receptor during limb muscle development. *Proceedings of the National Academy of Sciences of the United States of America* **93**(9), 4213-4218.
35 Bladt, F., Riethmacher, D., Isenmann, S., Aguzzi, A. & Birchmeier, C. *1995* Essential role for the c-met receptor in the migration of myogenic precursor cells into the limb bud. *Nature* **376**(6543), 768-771.
36 Kablar, B., Krastel, K., Ying, C., Asakura, A., Tapscott, S. J. & Rudnicki, M. A. *1997* MyoD and Myf-5 differentially regulate the development of limb versus trunk skeletal muscle. *Development* **124**(23), 4729-4738.
37 Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esumi, E., Li, S., Nonaka, I. & Nabeshima, Y. *1993* Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature* **364**(6437), 532-535.
38 Olson, E. N., Arnold, H. H., Rigby, P. W. & Wold, B. J. *1996* Know your neighbors: three phenotypes in null mutants of the myogenic bHLH gene MRF4. *Cell* **85**(1), 1-4.
39 Valdez, M. R., Richardson, J. A., Klein, W. H. & Olson, E. N. *2000* Failure of Myf5 to support myogenic differentiation without myogenin, MyoD, and MRF4. *Dev Biol* **219**(2), 287-298.
40 Bayol, S., Jones, D., Goldspink, G. & Stickland, N. C. *2004* The influence of undernutrition during gestation on skeletal muscle cellularity and on the expression of genes that control muscle growth. *The British journal of nutrition* **91**(3), 331-339.
41 Costello, P. M., Rowlerson, A., Astaman, N. A., Anthony, F. E., Sayer, A. A., Cooper, C., Hanson, M. A. & Green, L. R. *2008* Peri-implantation and late gestation maternal

undernutrition differentially affect fetal sheep skeletal muscle development. *The Journal of physiology* **586**(9), 2371-2379.

42 Jensen, C. B., Martin-Gronert, M. S., Storgaard, H., Madsbad, S., Vaag, A. & Ozanne, S. E. *2008* Altered PI3-kinase/Akt signalling in skeletal muscle of young men with low birth weight. *PloS one* **3**(11), e3738.

43 Shen, W., Boyle, D. W. & Liechty, E. A. *2005* Changes in 4E-BP1 and p70S6K phosphorylation in skeletal muscle of the ovine fetus after prolonged maternal fasting: effects of insulin and IGF-I. *Pediatr Res* **58**(5), 833-839.

44 Suryawan, A., Jeyapalan, A. S., Orellana, R. A., Wilson, F. A., Nguyen, H. V. & Davis, T. A. *2008* Leucine stimulates protein synthesis in skeletal muscle of neonatal pigs by enhancing mTORC1 activation. *American journal of physiology. Endocrinology and metabolism* **295**(4), E868-875.

45 Moura-Dos-Santos, M., Wellington-Barros, J., Brito-Almeida, M., Manhaes-de-Castro, R., Maia, J. & Gois Leandro, C. *2013* Permanent deficits in handgrip strength and running speed performance in low birth weight children. *Am J Hum Biol* **25**(1), 58-62.

46 Loos, R. J., Beunen, G., Fagard, R., Derom, C. & Vlietinck, R. *2001* Birth weight and body composition in young adult men--a prospective twin study. *Int J Obes Relat Metab Disord* **25**(10), 1537-1545.

47 Patel, H. P., White, M. C., Westbury, L., Syddall, H. E., Stephens, P. J., Clough, G. F., Cooper, C. & Sayer, A. A. *2015* Skeletal muscle morphology in sarcopenia defined using the EWGSOP criteria: findings from the Hertfordshire Sarcopenia Study (HSS). *BMC geriatrics* **15**171.

48 Sayer, A. A., Syddall, H. E., Dennison, E. M., Gilbody, H. J., Duggleby, S. L., Cooper, C., Barker, D. J. & Phillips, D. I. *2004* Birth weight, weight at 1 y of age, and body composition in older men: findings from the Hertfordshire Cohort Study. *The American journal of clinical nutrition* **80**(1), 199-203.

49 Asakura, A., Komaki, M. & Rudnicki, M. *2001* Muscle satellite cells are multipotential stem cells that exhibit myogenic, osteogenic, and adipogenic differentiation. *Differentiation* **68**(4-5), 245-253.

50 Yada, E., Yamanouchi, K. & Nishihara, M. *2006* Adipogenic potential of satellite cells from distinct skeletal muscle origins in the rat. *J Vet Med Sci* **68**(5), 479-486.

51 Goodpaster, B. H., Krishnaswami, S., Resnick, H., Kelley, D. E., Haggerty, C., Harris, T. B., Schwartz, A. V., Kritchevsky, S. & Newman, A. B. *2003* Association between regional adipose tissue distribution and both type 2 diabetes and impaired glucose tolerance in elderly men and women. *Diabetes Care* **26**(2), 372-379.

52 Delmonico, M. J., Harris, T. B., Visser, M., Park, S. W., Conroy, M. B., Velasquez-Mieyer, P., Boudreau, R., Manini, T. M., Nevitt, M., Newman, A. B., Goodpaster, B. H., Health, A. & Body. *2009* Longitudinal study of muscle strength, quality, and adipose tissue infiltration. *The American journal of clinical nutrition* **90**(6), 1579-1585.

53 Marcus, R. L., Addison, O., Dibble, L. E., Foreman, K. B., Morrell, G. & Lastayo, P. *2012* Intramuscular adipose tissue, sarcopenia, and mobility function in older individuals. *J Aging Res* **2012**629637.

54 Chinsomboon, J., Ruas, J., Gupta, R. K., Thom, R., Shoag, J., Rowe, G. C., Sawada, N., Raghuram, S. & Arany, Z. *2009* The transcriptional coactivator PGC-1alpha mediates exercise-induced angiogenesis in skeletal muscle. *Proceedings of the National Academy of Sciences of the United States of America* **106**(50), 21401-21406.

55 Denis, C., Chatard, J. C., Dormois, D., Linossier, M. T., Geyssant, A. & Lacour, J. R. *1986* Effects of endurance training on capillary supply of human skeletal muscle on two age groups (20 and 60 years). *J Physiol (Paris)* **81**(5), 379-383.

56 Cartee, G. D. *2015* Mechanisms for greater insulin-stimulated glucose uptake in normal and insulin-resistant skeletal muscle after acute exercise. *American journal of physiology. Endocrinology and metabolism* **309**(12), E949-959.

57 Adolfsson, J., Ljungqvist, A., Tornling, G. & Unge, G. *1981* Capillary increase in the skeletal muscle of trained young and adult rats. *The Journal of physiology* **310**529-532.

58 Egan, B. & Zierath, J. R. *2013* Exercise metabolism and the molecular regulation of skeletal muscle adaptation. *Cell Metab* **17**(2), 162-184.
59 Howald, H., Hoppeler, H., Claassen, H., Mathieu, O. & Straub, R. *1985* Influences of endurance training on the ultrastructural composition of the different muscle fiber types in humans. *Pflugers Arch* **403**(4), 369-376.
60 Wu, Z., Puigserver, P., Andersson, U., Zhang, C., Adelmant, G., Mootha, V., Troy, A., Cinti, S., Lowell, B., Scarpulla, R. C. & Spiegelman, B. M. *1999* Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**(1), 115-124.
61 Baar, K., Wende, A. R., Jones, T. E., Marison, M., Nolte, L. A., Chen, M., Kelly, D. P. & Holloszy, J. O. *2002* Adaptations of skeletal muscle to exercise: rapid increase in the transcriptional coactivator PGC-1. *FASEB J* **16**(14), 1879-1886.
62 Little, J. P., Safdar, A., Cermak, N., Tarnopolsky, M. A. & Gibala, M. J. *2010* Acute endurance exercise increases the nuclear abundance of PGC-1alpha in trained human skeletal muscle. *Am J Physiol Regul Integr Comp Physiol* **298**(4), R912-917.
63 Lin, J., Wu, H., Tarr, P. T., Zhang, C. Y., Wu, Z., Boss, O., Michael, L. F., Puigserver, P., Isotani, E., Olson, E. N., Lowell, B. B., Bassel-Duby, R. & Spiegelman, B. M. *2002* Transcriptional co-activator PGC-1 alpha drives the formation of slow-twitch muscle fibres. *Nature* **418**(6899), 797-801.
64 Russell, A. P., Feilchenfeldt, J., Schreiber, S., Praz, M., Crettenand, A., Gobelet, C., Meier, C. A., Bell, D. R., Kralli, A., Giacobino, J. P. & Deriaz, O. *2003* Endurance training in humans leads to fiber type-specific increases in levels of peroxisome proliferator-activated receptor-gamma coactivator-1 and peroxisome proliferator-activated receptor-alpha in skeletal muscle. *Diabetes* **52**(12), 2874-2881.
65 Seene, T., Kaasik, P. & Umnova, M. *2009* Structural rearrangements in contractile apparatus and resulting skeletal muscle remodelling: effect of exercise training. *J Sports Med Phys Fitness* **49**(4), 410-423.
66 Inashima, S., Matsunaga, S., Yasuda, T. & Wada, M. *2003* Effect of endurance training and acute exercise on sarcoplasmic reticulum function in rat fast- and slow-twitch skeletal muscles. *European journal of applied physiology* **89**(2), 142-149.
67 Kinnunen, S. & Manttari, S. *2012* Specific effects of endurance and sprint training on protein expression of calsequestrin and SERCA in mouse skeletal muscle. *J Muscle Res Cell Motil* **33**(2), 123-130.
68 Schoenfeld, B. J., Ogborn, D. & Krieger, J. W. *2016* Effects of Resistance Training Frequency on Measures of Muscle Hypertrophy: A Systematic Review and Meta-Analysis. *Sports Med* **46**(11), 1689-1697.
69 Bodine, S. C., Stitt, T. N., Gonzalez, M., Kline, W. O., Stover, G. L., Bauerlein, R., Zlotchenko, E., Scrimgeour, A., Lawrence, J. C., Glass, D. J. & Yancopoulos, G. D. *2001* Akt/mTOR pathway is a crucial regulator of skeletal muscle hypertrophy and can prevent muscle atrophy in vivo. *Nat Cell Biol* **3**(11), 1014-1019.
70 McCarthy, J. J., Mula, J., Miyazaki, M., Erfani, R., Garrison, K., Farooqui, A. B., Srikuea, R., Lawson, B. A., Grimes, B., Keller, C., Van Zant, G., Campbell, K. S., Esser, K. A., Dupont-Versteegden, E. E. & Peterson, C. A. *2011* Effective fiber hypertrophy in satellite cell-depleted skeletal muscle. *Development* **138**(17), 3657-3666.
71 Psatha, M., Wu, Z., Gammie, F. M., Ratkevicius, A., Wackerhage, H., Lee, J. H., Redpath, T. W., Gilbert, F. J., Ashcroft, G. P., Meakin, J. R. & Aspden, R. M. *2012* A longitudinal MRI study of muscle atrophy during lower leg immobilization following ankle fracture. *J Magn Reson Imaging* **35**(3), 686-695.
72 Drummond, M. J., Dickinson, J. M., Fry, C. S., Walker, D. K., Gundermann, D. M., Reidy, P. T., Timmerman, K. L., Markofski, M. M., Paddon-Jones, D., Rasmussen, B. B. & Volpi, E. *2012* Bed rest impairs skeletal muscle amino acid transporter expression, mTORC1 signaling, and protein synthesis in response to essential amino acids in older adults. *American journal of physiology. Endocrinology and metabolism* **302**(9), E1113-1122.
73 Suetta, C., Frandsen, U., Jensen, L., Jensen, M. M., Jespersen, J. G., Hvid, L. G., Bayer, M., Petersson, S. J., Schroder, H. D., Andersen, J. L., Heinemeier, K. M., Aagaard, P., Schjerling,

P. & Kjaer, M. *2012* Aging affects the transcriptional regulation of human skeletal muscle disuse atrophy. *PloS one* **7**(12), e51238.
74 Bodine, S. C. *2013* Disuse-induced muscle wasting. *The international journal of biochemistry & cell biology* **45**(10), 2200-2208.
75 Bonaldo, P. & Sandri, M. *2013* Cellular and molecular mechanisms of muscle atrophy. *Dis Model Mech* **6**(1), 25-39.
76 Jackman, R. W. & Kandarian, S. C. *2004* The molecular basis of skeletal muscle atrophy. *Am J Physiol Cell Physiol* **287**(4), C834-843.
77 Sharma, M., Langley, B., Bass, J. & Kambadur, R. *2001* Myostatin in muscle growth and repair. *Exerc Sport Sci Rev* **29**(4), 155-158.
78 McPherron, A. C., Lawler, A. M. & Lee, S. J. *1997* Regulation of skeletal muscle mass in mice by a new TGF-beta superfamily member. *Nature* **387**(6628), 83-90.
79 Gonzalez-Cadavid, N. F., Taylor, W. E., Yarasheski, K., Sinha-Hikim, I., Ma, K., Ezzat, S., Shen, R., Lalani, R., Asa, S., Mamita, M., Nair, G., Arver, S. & Bhasin, S. *1998* Organization of the human myostatin gene and expression in healthy men and HIV-infected men with muscle wasting. *Proc Natl Acad Sci U S A* **95**(25), 14938-14943.
80 McMahon, C. D., Popovic, L., Oldham, J. M., Jeanplong, F., Smith, H. K., Kambadur, R., Sharma, M., Maxwell, L. & Bass, J. J. *2003* Myostatin-deficient mice lose more skeletal muscle mass than wild-type controls during hindlimb suspension. *American journal of physiology. Endocrinology and metabolism* **285**(1), E82-87.
81 Munoz, K. A., Satarug, S. & Tischler, M. E. *1993* Time course of the response of myofibrillar and sarcoplasmic protein metabolism to unweighting of the soleus muscle. *Metabolism* **42**(8), 1006-1012.
82 Huang, J. & Forsberg, N. E. *1998* Role of calpain in skeletal-muscle protein degradation. *Proc Natl Acad Sci U S A* **95**(21), 12100-12105.
83 Moresi, V., Williams, A. H., Meadows, E., Flynn, J. M., Potthoff, M. J., McAnally, J., Shelton, J. M., Backs, J., Klein, W. H., Richardson, J. A., Bassel-Duby, R. & Olson, E. N. *2010* Myogenin and class II HDACs control neurogenic muscle atrophy by inducing E3 ubiquitin ligases. *Cell* **143**(1), 35-45.
84 Gustafsson, T., Osterlund, T., Flanagan, J. N., von Walden, F., Trappe, T. A., Linnehan, R. M. & Tesch, P. A. *2010* Effects of 3 days unloading on molecular regulators of muscle size in humans. *Journal of applied physiology* **109**(3), 721-727.
85 Adams, V., Mangner, N., Gasch, A., Krohne, C., Gielen, S., Hirner, S., Thierse, H. J., Witt, C. C., Linke, A., Schuler, G. & Labeit, S. *2008* Induction of MuRF1 is essential for TNF-alpha-induced loss of muscle function in mice. *Journal of molecular biology* **384**(1), 48-59.
86 Asp, M. L., Tian, M., Wendel, A. A. & Belury, M. A. *2010* Evidence for the contribution of insulin resistance to the development of cachexia in tumor-bearing mice. *Int J Cancer* **126**(3), 756-763.
87 Tintignac, L. A., Lagirand, J., Batonnet, S., Sirri, V., Leibovitch, M. P. & Leibovitch, S. A. *2005* Degradation of MyoD mediated by the SCF (MAFbx) ubiquitin ligase. *J Biol Chem* **280**(4), 2847-2856.
88 Csibi, A., Cornille, K., Leibovitch, M. P., Poupon, A., Tintignac, L. A., Sanchez, A. M. & Leibovitch, S. A. *2010* The translation regulatory subunit eIF3f controls the kinase-dependent mTOR signaling required for muscle differentiation and hypertrophy in mouse. *PLoS One* **5**(2), e8994.
89 Kedar, V., McDonough, H., Arya, R., Li, H. H., Rockman, H. A. & Patterson, C. *2004* Muscle-specific RING finger 1 is a bona fide ubiquitin ligase that degrades cardiac troponin I. *Proc Natl Acad Sci U S A* **101**(52), 18135-18140.
90 Clarke, B. A., Drujan, D., Willis, M. S., Murphy, L. O., Corpina, R. A., Burova, E., Rakhilin, S. V., Stitt, T. N., Patterson, C., Latres, E. & Glass, D. J. *2007* The E3 Ligase MuRF1 degrades myosin heavy chain protein in dexamethasone-treated skeletal muscle. *Cell Metab* **6**(5), 376-385.

91 Cohen, S., Brault, J. J., Gygi, S. P., Glass, D. J., Valenzuela, D. M., Gartner, C., Latres, E. & Goldberg, A. L. *2009* During muscle atrophy, thick, but not thin, filament components are degraded by MuRF1-dependent ubiquitylation. *J Cell Biol* **185**(6), 1083-1095.
92 Bansal, D., Miyake, K., Vogel, S. S., Groh, S., Chen, C. C., Williamson, R., McNeil, P. L. & Campbell, K. P. *2003* Defective membrane repair in dysferlin-deficient muscular dystrophy. *Nature* **423**(6936), 168-172.
93 Tidball, J. G. *2011* Mechanisms of muscle injury, repair, and regeneration. *Compr Physiol* **1**(4), 2029-2062.
94 Tedesco, F. S., Dellavalle, A., Diaz-Manera, J., Messina, G. & Cossu, G. *2010* Repairing skeletal muscle: regenerative potential of skeletal muscle stem cells. *J Clin Invest* **120**(1), 11-19.
95 DiMario, J., Buffinger, N., Yamada, S. & Strohman, R. C. *1989* Fibroblast growth factor in the extracellular matrix of dystrophic (mdx) mouse muscle. *Science* **244**(4905), 688-690.
96 Kherif, S., Lafuma, C., Dehaupas, M., Lachkar, S., Fournier, J. G., Verdiere-Sahuque, M., Fardeau, M. & Alameddine, H. S. *1999* Expression of matrix metalloproteinases 2 and 9 in regenerating skeletal muscle: a study in experimentally injured and mdx muscles. *Dev Biol* **205**(1), 158-170.
97 Nakane, M., Schmidt, H. H., Pollock, J. S., Forstermann, U. & Murad, F. *1993* Cloned human brain nitric oxide synthase is highly expressed in skeletal muscle. *FEBS Lett* **316**(2), 175-180.
98 Yamada, M., Sankoda, Y., Tatsumi, R., Mizunoya, W., Ikeuchi, Y., Sunagawa, K. & Allen, R. E. *2008* Matrix metalloproteinase-2 mediates stretch-induced activation of skeletal muscle satellite cells in a nitric oxide-dependent manner. *The international journal of biochemistry & cell biology* **40**(10), 2183-2191.
99 Wozniak, A. C., Pilipowicz, O., Yablonka-Reuveni, Z., Greenway, S., Craven, S., Scott, E. & Anderson, J. E. *2003* C-met expression and mechanical activation of satellite cells on cultured muscle fibers. *Journal of Histochemistry & Cytochemistry* **51**(11), 1437-1445.
100 Cornelison, D. D. W. & Wold, B. J. *1997* Single-cell analysis of regulatory gene expression in quiescent and activated mouse skeletal muscle satellite cells. *Developmental Biology* **191**(2), 270-283.
101 Yablonkareuveni, Z. & Rivera, A. J. *1994* Temporal Expression of Regulatory and Structural Muscle Proteins during Myogenesis of Satellite Cells on Isolated Adult-Rat Fibers. *Developmental Biology* **164**(2), 588-603.
102 Crist, C. G., Montarras, D. & Buckingham, M. *2012* Muscle satellite cells are primed for myogenesis but maintain quiescence with sequestration of Myf5 mRNA targeted by microRNA-31 in mRNP granules. *Cell Stem Cell* **11**(1), 118-126.
103 Molkentin, J. D., Black, B. L., Martin, J. F. & Olson, E. N. *1995* Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* **83**(7), 1125-1136.
104 Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. *2013* The hallmarks of aging. *Cell* **153**(6), 1194-1217.
105 Rosenberg, I. H. *1997* Sarcopenia: origins and clinical relevance. *J Nutr* **127**(5 Suppl), 990S-991S.
106 Anker, S. D., Morley, J. E. & von Haehling, S. *2016* Welcome to the ICD-10 code for sarcopenia. *J Cachexia Sarcopenia Muscle* **7**(5), 512-514.
107 Cruz-Jentoft, A. J., Landi, F., Schneider, S. M., Zuniga, C., Arai, H., Boirie, Y., Chen, L. K., Fielding, R. A., Martin, F. C., Michel, J. P., Sieber, C., Stout, J. R., Studenski, S. A., Vellas, B., Woo, J., Zamboni, M. & Cederholm, T. *2014* Prevalence of and interventions for sarcopenia in ageing adults: a systematic review. Report of the International Sarcopenia Initiative (EWGSOP and IWGS). *Age and ageing* **43**(6), 748-759.
108 Baumgartner, R. N., Koehler, K. M., Gallagher, D., Romero, L., Heymsfield, S. B., Ross, R. R., Garry, P. J. & Lindeman, R. D. *1998* Epidemiology of sarcopenia among the elderly in New Mexico. *Am J Epidemiol* **147**(8), 755-763.
109 Melton, L. J., 3rd, Khosla, S., Crowson, C. S., O'Connor, M. K., O'Fallon, W. M. & Riggs, B. L. *2000* Epidemiology of sarcopenia. *J Am Geriatr Soc* **48**(6), 625-630.

110 Gallagher, D., Visser, M., De Meersman, R. E., Sepulveda, D., Baumgartner, R. N., Pierson, R. N., Harris, T. & Heymsfield, S. B. *1997* Appendicular skeletal muscle mass: effects of age, gender, and ethnicity. *Journal of applied physiology* **83**(1), 229-239.
111 Pongchaiyakul, C., Limpawattana, P., Kotruchin, P. & Rajatanavin, R. *2013* Prevalence of sarcopenia and associated factors among Thai population. *J Bone Miner Metab* **31**(3), 346-350.
112 Kirchengast, S. & Huber, J. *2009* Gender and age differences in lean soft tissue mass and sarcopenia among healthy elderly. *Anthropol Anz* **67**(2), 139-151.
113 Cruz-Jentoft, A. J., Baeyens, J. P., Bauer, J. M., Boirie, Y., Cederholm, T., Landi, F., Martin, F. C., Michel, J. P., Rolland, Y., Schneider, S. M., Topinkova, E., Vandewoude, M., Zamboni, M. & European Working Group on Sarcopenia in Older, P. *2010* Sarcopenia: European consensus on definition and diagnosis: Report of the European Working Group on Sarcopenia in Older People. *Age and ageing* **39**(4), 412-423.
114 Fielding, R. A., Vellas, B., Evans, W. J., Bhasin, S., Morley, J. E., Newman, A. B., Abellan van Kan, G., Andrieu, S., Bauer, J., Breuille, D., Cederholm, T., Chandler, J., De Meynard, C., Donini, L., Harris, T., Kannt, A., Keime Guibert, F., Onder, G., Papanicolaou, D., Rolland, Y., Rooks, D., Sieber, C., Souhami, E., Verlaan, S. & Zamboni, M. *2011* Sarcopenia: an undiagnosed condition in older adults. Current consensus definition: prevalence, etiology, and consequences. International working group on sarcopenia. *J Am Med Dir Assoc* **12**(4), 249-256.
115 Chen, L. K., Liu, L. K., Woo, J., Assantachai, P., Auyeung, T. W., Bahyah, K. S., Chou, M. Y., Chen, L. Y., Hsu, P. S., Krairit, O., Lee, J. S., Lee, W. J., Lee, Y., Liang, C. K., Limpawattana, P., Lin, C. S., Peng, L. N., Satake, S., Suzuki, T., Won, C. W., Wu, C. H., Wu, S. N., Zhang, T., Zeng, P., Akishita, M. & Arai, H. *2014* Sarcopenia in Asia: consensus report of the Asian Working Group for Sarcopenia. *J Am Med Dir Assoc* **15**(2), 95-101.
116 Studenski, S. A., Peters, K. W., Alley, D. E., Cawthon, P. M., McLean, R. R., Harris, T. B., Ferrucci, L., Guralnik, J. M., Fragala, M. S., Kenny, A. M., Kiel, D. P., Kritchevsky, S. B., Shardell, M. D., Dam, T. T. & Vassileva, M. T. *2014* The FNIH sarcopenia project: rationale, study description, conference recommendations, and final estimates. *J Gerontol A Biol Sci Med Sci* **69**(5), 547-558.
117 Newman, A. B., Kupelian, V., Visser, M., Simonsick, E., Goodpaster, B., Nevitt, M., Kritchevsky, S. B., Tylavsky, F. A., Rubin, S. M., Harris, T. B. & Health, A. B. C. S. I. *2003* Sarcopenia: alternative definitions and associations with lower extremity function. *J Am Geriatr Soc* **51**(11), 1602-1609.
118 Lauretani, F., Russo, C. R., Bandinelli, S., Bartali, B., Cavazzini, C., Di Iorio, A., Corsi, A. M., Rantanen, T., Guralnik, J. M. & Ferrucci, L. *2003* Age-associated changes in skeletal muscles and their effect on mobility: an operational diagnosis of sarcopenia. *Journal of applied physiology* **95**(5), 1851-1860.
119 Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W. J., Burke, G., McBurnie, M. A. & Cardiovascular Health Study Collaborative Research, G. *2001* Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* **56**(3), M146-156.
120 Guralnik, J. M., Ferrucci, L., Pieper, C. F., Leveille, S. G., Markides, K. S., Ostir, G. V., Studenski, S., Berkman, L. F. & Wallace, R. B. *2000* Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *J Gerontol A Biol Sci Med Sci* **55**(4), M221-231.
121 Evans, W. J. & Campbell, W. W. *1993* Sarcopenia and age-related changes in body composition and functional capacity. *J Nutr* **123**(2 Suppl), 465-468.
122 Reeves, G. K., Balkwill, A., Cairns, B. J., Green, J., Beral, V. & Million Women Study, C. *2014* Hospital admissions in relation to body mass index in UK women: a prospective cohort study. *BMC Med* **12**45.
123 Vallgarda, S. *1999* Is old age necessarily connected with high hospital admission rates? *Scand J Public Health* **27**(2), 137-142.

124 Ferrando, A. A., Lane, H. W., Stuart, C. A., Davis-Street, J. & Wolfe, R. R. *1996* Prolonged bed rest decreases skeletal muscle and whole body protein synthesis. *Am J Physiol* **270**(4 Pt 1), E627-633.
125 English, K. L. & Paddon-Jones, D. *2010* Protecting muscle mass and function in older adults during bed rest. *Curr Opin Clin Nutr Metab Care* **13**(1), 34-39.
126 Ferrando, A. A., Tipton, K. D., Bamman, M. M. & Wolfe, R. R. *1997* Resistance exercise maintains skeletal muscle protein synthesis during bed rest. *J Appl Physiol (1985)* **82**(3), 807-810.
127 Tanner, R. E., Brunker, L. B., Agergaard, J., Barrows, K. M., Briggs, R. A., Kwon, O. S., Young, L. M., Hopkins, P. N., Volpi, E., Marcus, R. L., LaStayo, P. C. & Drummond, M. J. *2015* Age-related differences in lean mass, protein synthesis and skeletal muscle markers of proteolysis after bed rest and exercise rehabilitation. *The Journal of physiology* **593**(18), 4259-4273.
128 Fiatarone, M. A., O'Neill, E. F., Ryan, N. D., Clements, K. M., Solares, G. R., Nelson, M. E., Roberts, S. B., Kehayias, J. J., Lipsitz, L. A. & Evans, W. J. *1994* Exercise training and nutritional supplementation for physical frailty in very elderly people. *N Engl J Med* **330**(25), 1769-1775.
129 Suetta, C., Andersen, J. L., Dalgas, U., Berget, J., Koskinen, S., Aagaard, P., Magnusson, S. P. & Kjaer, M. *2008* Resistance training induces qualitative changes in muscle morphology, muscle architecture, and muscle function in elderly postoperative patients. *J Appl Physiol (1985)* **105**(1), 180-186.
130 Lakka, T. A., Lakka, H. M., Rankinen, T., Leon, A. S., Rao, D. C., Skinner, J. S., Wilmore, J. H. & Bouchard, C. *2005* Effect of exercise training on plasma levels of C-reactive protein in healthy adults: the HERITAGE Family Study. *Eur Heart J* **26**(19), 2018-2025.
131 Forti, L. N., Njemini, R., Beyer, I., Eelbode, E., Meeusen, R., Mets, T. & Bautmans, I. *2014* Strength training reduces circulating interleukin-6 but not brain-derived neurotrophic factor in community-dwelling elderly individuals. *Age* **36**(5), 9704.
132 Paddon-Jones, D., Sheffield-Moore, M., Cree, M. G., Hewlings, S. J., Aarsland, A., Wolfe, R. R. & Ferrando, A. A. *2006* Atrophy and impaired muscle protein synthesis during prolonged inactivity and stress. *The Journal of clinical endocrinology and metabolism* **91**(12), 4836-4841.
133 Jackson, A. S., Janssen, I., Sui, X., Church, T. S. & Blair, S. N. *2012* Longitudinal changes in body composition associated with healthy ageing: men, aged 20-96 years. *The British journal of nutrition* **107**(7), 1085-1091.
134 Faulkner, J. A., Larkin, L. M., Claflin, D. R. & Brooks, S. V. *2007* Age-related changes in the structure and function of skeletal muscles. *Clinical and experimental pharmacology & physiology* **34**(11), 1091-1096.
135 Sugawara, J., Miyachi, M., Moreau, K. L., Dinenno, F. A., DeSouza, C. A. & Tanaka, H. *2002* Age-related reductions in appendicular skeletal muscle mass: association with habitual aerobic exercise status. *Clinical physiology and functional imaging* **22**(3), 169-172.
136 Janssen, I., Heymsfield, S. B., Wang, Z. M. & Ross, R. *2000* Skeletal muscle mass and distribution in 468 men and women aged 18-88 yr. *Journal of applied physiology* **89**(1), 81-88.
137 Lexell, J., Henriksson-Larsen, K., Winblad, B. & Sjostrom, M. *1983* Distribution of different fiber types in human skeletal muscles: effects of aging studied in whole muscle cross sections. *Muscle & nerve* **6**(8), 588-595.
138 Nilwik, R., Snijders, T., Leenders, M., Groen, B. B., van Kranenburg, J., Verdijk, L. B. & van Loon, L. J. *2013* The decline in skeletal muscle mass with aging is mainly attributed to a reduction in type II muscle fiber size. *Experimental gerontology* **48**(5), 492-498.
139 Doherty, T. J. & Brown, W. F. *1993* The estimated numbers and relative sizes of thenar motor units as selected by multiple point stimulation in young and older adults. *Muscle & nerve* **16**(4), 355-366.
140 Kent-Braun, J. A., Ng, A. V. & Young, K. *2000* Skeletal muscle contractile and noncontractile components in young and older women and men. *Journal of applied physiology* **88**(2), 662-668.

141 Rivard, A., Fabre, J. E., Silver, M., Chen, D., Murohara, T., Kearney, M., Magner, M., Asahara, T. & Isner, J. M. *1999* Age-dependent impairment of angiogenesis. *Circulation* **99**(1), 111-120.
142 Ryan, N. A., Zwetsloot, K. A., Westerkamp, L. M., Hickner, R. C., Pofahl, W. E. & Gavin, T. P. *2006* Lower skeletal muscle capillarization and VEGF expression in aged vs. young men. *Journal of applied physiology* **100**(1), 178-185.
143 Demontis, F., Piccirillo, R., Goldberg, A. L. & Perrimon, N. *2013* Mechanisms of skeletal muscle aging: insights from Drosophila and mammalian models. *Dis Model Mech* **6**(6), 1339-1352.
144 Degens, H. & Alway, S. E. *2006* Control of muscle size during disuse, disease, and aging. *International journal of sports medicine* **27**(2), 94-99.
145 Marzetti, E., Calvani, R., Bernabei, R. & Leeuwenburgh, C. *2012* Apoptosis in skeletal myocytes: a potential target for interventions against sarcopenia and physical frailty - a mini-review. *Gerontology* **58**(2), 99-106.
146 Johnston, A. P., De Lisio, M. & Parise, G. *2008* Resistance training, sarcopenia, and the mitochondrial theory of aging. *Applied physiology, nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme* **33**(1), 191-199.
147 Lanza, I. R. & Nair, K. S. *2009* Muscle mitochondrial changes with aging and exercise. *The American journal of clinical nutrition* **89**(1), 467S-471S.
148 Meng, S. J. & Yu, L. J. *2010* Oxidative stress, molecular inflammation and sarcopenia. *International journal of molecular sciences* **11**(4), 1509-1526.
149 Alway, S. E., Myers, M. J. & Mohamed, J. S. *2014* Regulation of satellite cell function in sarcopenia. *Frontiers in aging neuroscience* **6**246.
150 Leger, B., Derave, W., De Bock, K., Hespel, P. & Russell, A. P. *2008* Human sarcopenia reveals an increase in SOCS-3 and myostatin and a reduced efficiency of Akt phosphorylation. *Rejuvenation research* **11**(1), 163-175B.
151 Guillet, C., Auguste, P., Mayo, W., Kreher, P. & Gascan, H. *1999* Ciliary neurotrophic factor is a regulator of muscular strength in aging. *J Neurosci* **19**(4), 1257-1262.
152 Kovacheva, E. L., Hikim, A. P. S., Shen, R. Q., Sinha, I. & Sinha-Hikim, I. *2010* Testosterone Supplementation Reverses Sarcopenia in Aging through Regulation of Myostatin, c-Jun NH2-Terminal Kinase, Notch, and Akt Signaling Pathways. *Endocrinology* **151**(2), 628-638.
153 Yang, W., Zhang, Y., Li, Y., Wu, Z. & Zhu, D. *2007* Myostatin induces cyclin D1 degradation to cause cell cycle arrest through a phosphatidylinositol 3-kinase/AKT/GSK-3 beta pathway and is antagonized by insulin-like growth factor 1. *The Journal of biological chemistry* **282**(6), 3799-3808.
154 Naito, A. T., Sumida, T., Nomura, S., Liu, M. L., Higo, T., Nakagawa, A., Okada, K., Sakai, T., Hashimoto, A., Hara, Y., Shimizu, I., Zhu, W. D., Toko, H., Katada, A., Akazawa, H., Oka, T., Lee, J. K., Minamino, T., Nagai, T., Walsh, K., Kikuchi, A., Matsumoto, M., Botto, M., Shiojima, I. & Komuro, I. *2012* Complement C1q Activates Canonical Wnt Signaling and Promotes Aging-Related Phenotypes (vol 149, pg 1298, 2012). *Cell* **150**(3), 659-660.
155 Furuyama, T., Yamashita, H., Kitayama, K., Higami, Y., Shimokawa, I. & Mori, N. *2002* Effects of aging and caloric restriction on the gene expression of Foxo1, 3, and 4 (FKHR, FKHRL1, and AFX) in the rat skeletal muscles. *Microsc Res Techniq* **59**(4), 331-334.
156 Williamson, D. L., Raue, U., Slivka, D. R. & Trappe, S. *2010* Resistance Exercise, Skeletal Muscle FOXO3A, and 85-Year-Old Women. *J Gerontol a-Biol* **65**(4), 335-343.
157 Sandri, M., Sandri, C., Gilbert, A., Skurk, C., Calabria, E., Picard, A., Walsh, K., Schiaffino, S., Lecker, S. H. & Goldberg, A. L. *2004* Foxo transcription factors induce the atrophy-related ubiquitin ligase atrogin-1 and cause skeletal muscle atrophy. *Cell* **117**(3), 399-412.
158 Arnheim, N. & Cortopassi, G. *1992* Deleterious mitochondrial DNA mutations accumulate in aging human tissues. *Mutat Res* **275**(3-6), 157-167.
159 Marzetti, E., Calvani, R., Cesari, M., Buford, T. W., Lorenzi, M., Behnke, B. J. & Leeuwenburgh, C. *2013* Mitochondrial dysfunction and sarcopenia of aging: from signaling pathways to clinical trials. *The international journal of biochemistry & cell biology* **45**(10), 2288-2301.

160 Zahn, J. M., Sonu, R., Vogel, H., Crane, E., Mazan-Mamczarz, K., Rabkin, R., Davis, R. W., Becker, K. G., Owen, A. B. & Kim, S. K. *2006* Transcriptional profiling of aging in human muscle reveals a common aging signature. *PLoS genetics* **2**(7), e115.
161 Cui, H., Kong, Y. & Zhang, H. *2012* Oxidative stress, mitochondrial dysfunction, and aging. *J Signal Transduct* **2012**646354.
162 Liochev, S. I. *2013* Reactive oxygen species and the free radical theory of aging. *Free radical biology & medicine* **60**1-4.
163 Moylan, J. S. & Reid, M. B. *2007* Oxidative stress, chronic disease, and muscle wasting. *Muscle & nerve* **35**(4), 411-429.
164 Carthew, R. W. & Sontheimer, E. J. *2009* Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**(4), 642-655.
165 Rao, P. K., Kumar, R. M., Farkhondeh, M., Baskerville, S. & Lodish, H. F. *2006* Myogenic factors that regulate expression of muscle-specific microRNAs. *Proceedings of the National Academy of Sciences of the United States of America* **103**(23), 8721-8726.
166 Chen, J. F., Mandel, E. M., Thomson, J. M., Wu, Q., Callis, T. E., Hammond, S. M., Conlon, F. L. & Wang, D. Z. *2006* The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nature genetics* **38**(2), 228-233.
167 Safdar, A., Abadi, A., Akhtar, M., Hettinga, B. P. & Tarnopolsky, M. A. *2009* miRNA in the regulation of skeletal muscle adaptation to acute endurance exercise in C57Bl/6J male mice. *PloS one* **4**(5), e5610.
168 Drummond, M. J., McCarthy, J. J., Sinha, M., Spratt, H. M., Volpi, E., Esser, K. A. & Rasmussen, B. B. *2011* Aging and microRNA expression in human skeletal muscle: a microarray and bioinformatics analysis. *Physiological genomics* **43**(10), 595-603.
169 Tan, L. J., Liu, S. L., Lei, S. F., Papasian, C. J. & Deng, H. W. *2012* Molecular genetic studies of gene identification for sarcopenia. *Human genetics* **131**(1), 1-31.
170 Sarkar, S., Dey, B. K. & Dutta, A. *2010* MiR-322/424 and -503 are induced during muscle differentiation and promote cell cycle quiescence and differentiation by down-regulation of Cdc25A. *Mol Biol Cell* **21**(13), 2138-2149.
171 Connolly, M., Paul, R., Farre-Garros, R., Natanek, S. A., Bloch, S., Lee, J., Lorenzo, J. P., Patel, H., Cooper, C., Sayer, A. A., Wort, S. J., Griffiths, M., Polkey, M. I. & Kemp, P. R. *2018* miR-424-5p reduces ribosomal RNA and protein synthesis in muscle wasting. *J Cachexia Sarcopenia Muscle* **9**(2), 400-416.
172 Dahlberg, A. E. *1989* The functional role of ribosomal RNA in protein synthesis. *Cell* **57**(4), 525-529.
173 Lowe, S. W. *1999* Hot papers - Cancer - Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a by M. Serrano, A.W. Lin, M.E. McCurrach, D. Beach, S.W. Lowe - Comments. *Scientist* **13**(22), 15-15.
174 Coppe, J. P., Desprez, P. Y., Krtolica, A. & Campisi, J. *2010* The senescence-associated secretory phenotype: the dark side of tumor suppression. *Annu Rev Pathol* **5**99-118.
175 Wiley, C. D., Velarde, M. C., Lecot, P., Liu, S., Sarnoski, E. A., Freund, A., Shirakawa, K., Lim, H. W., Davis, S. S., Ramanathan, A., Gerencser, A. A., Verdin, E. & Campisi, J. *2016* Mitochondrial Dysfunction Induces Senescence with a Distinct Secretory Phenotype. *Cell metabolism* **23**(2), 303-314.
176 Golias, C. H., Charalabopoulos, A. & Charalabopoulos, K. *2004* Cell proliferation and cell cycle control: a mini review. *Int J Clin Pract* **58**(12), 1134-1141.
177 Sharpless, N. E. & DePinho, R. A. *1999* The INK4A/ARF locus and its two gene products. *Curr Opin Genet Dev* **9**(1), 22-30.
178 Talior, I., Tennenbaum, T., Kuroki, T. & Eldar-Finkelman, H. *2005* PKC-delta-dependent activation of oxidative stress in adipocytes of obese and insulin-resistant mice: role for NADPH oxidase. *Am J Physiol Endocrinol Metab* **288**(2), E405-411.
179 LaPak, K. M. & Burd, C. E. *2014* The molecular balancing act of p16(INK4a) in cancer and aging. *Mol Cancer Res* **12**(2), 167-183.
180 Rayess, H., Wang, M. B. & Srivatsan, E. S. *2012* Cellular senescence and tumor suppressor gene p16. *Int J Cancer* **130**(8), 1715-1725.

181 Cao, R., Tsukada, Y. & Zhang, Y. *2005* Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Molecular cell* **20**(6), 845-854.
182 Cao, R., Wang, L. J., Wang, H. B., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S. & Zhang, Y. *2002* Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science* **298**(5595), 1039-1043.
183 Kaustov, L., Hui, O. Y., Amaya, M., Lemak, A., Nady, N., Duan, S. L., Wasney, G. A., Li, Z. H., Vedadi, M., Schapira, M., Min, J. R. & Arrowsmith, C. H. *2011* Recognition and Specificity Determinants of the Human Cbx Chromodomains. *Journal of Biological Chemistry* **286**(1), 521-529.
184 Agger, K., Cloos, P. A. C., Rudkjaer, L., Williams, K., Andersen, G., Christensen, J. & Helin, K. *2009* The H3K27me3 demethylase JMJD3 contributes to the activation of the INK4A-ARF locus in response to oncogene- and stress-induced senescence. *Genes & development* **23**(10), 1171-1176.
185 Toledo, F. & Wahl, G. M. *2006* Regulating the p53 pathway: in vitro hypotheses, in vivo veritas. *Nat Rev Cancer* **6**(12), 909-923.
186 Qian, Y. & Chen, X. *2013* Senescence regulation by the p53 protein family. *Methods in molecular biology* **965**37-61.
187 Qian, Y. & Chen, X. *2010* Tumor suppression by p53: making cells senescent. *Histol Histopathol* **25**(4), 515-526.
188 Takai, H., Smogorzewska, A. & de Lange, T. *2003* DNA damage foci at dysfunctional telomeres. *Curr Biol* **13**(17), 1549-1556.
189 Herbig, U., Jobling, W. A., Chen, B. P. C., Chen, D. J. & Sedivy, J. M. *2004* Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21(CIP1), but not p16(INK4a). *Molecular cell* **14**(4), 501-513.
190 Shiloh, Y. & Ziv, Y. *2013* The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat Rev Mol Cell Biol* **14**(4), 197-210.
191 Abbas, T. & Dutta, A. *2009* p21 in cancer: intricate networks and multiple activities. *Nat Rev Cancer* **9**(6), 400-414.
192 Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K. & Elledge, S. J. *1993* The P21 Cdk-Interacting Protein Cip1 Is a Potent Inhibitor of G1 Cyclin-Dependent Kinases. *Cell* **75**(4), 805-816.
193 Dimri, G. P., Nakanishi, M., Desprez, P. Y., Smith, J. R. & Campisi, J. *1996* Inhibition of E2F activity by the cyclin-dependent protein kinase inhibitor p21 in cells expressing or lacking a functional retinoblastoma protein. *Molecular and Cellular Biology* **16**(6), 2987-2997.
194 Sousa-Victor, P., Gutarra, S., Garcia-Prat, L., Rodriguez-Ubreva, J., Ortet, L., Ruiz-Bonilla, V., Jardi, M., Ballestar, E., Gonzalez, S., Serrano, A. L., Perdiguero, E. & Munoz-Canoves, P. *2014* Geriatric muscle stem cells switch reversible quiescence into senescence. *Nature* **506**(7488), 316-+.
195 Sousa-Victor, P., Perdiguero, E. & Munoz-Canoves, P. *2014* Geroconversion of aged muscle stem cells under regenerative pressure. *Cell Cycle* **13**(20), 3183-3190.
196 Karlseder, J., Broccoli, D., Dai, Y. M., Hardy, S. & de Lange, T. *1999* p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science* **283**(5406), 1321-1325.
197 Hayflick, L. & Moorhead, P. S. *1961* The serial cultivation of human diploid cell strains. *Experimental cell research* **25**585-621.
198 Bodnar, A. G., Ouellette, M., Frolkis, M., Holt, S. E., Chiu, C. P., Morin, G. B., Harley, C. B., Shay, J. W., Lichtsteiner, S. & Wright, W. E. *1998* Extension of life-span by introduction of telomerase into normal human cells. *Science* **279**(5349), 349-352.
199 Collins, K. & Mitchell, J. R. *2002* Telomerase in the human organism. *Oncogene* **21**(4), 564-579.
200 Shammas, M. A. *2011* Telomeres, lifestyle, cancer, and aging. *Curr Opin Clin Nutr Metab Care* **14**(1), 28-34.
201 Collado, M., Blasco, M. A. & Serrano, M. *2007* Cellular senescence in cancer and aging. *Cell* **130**(2), 223-233.

202 Cawthon, R. M., Smith, K. R., O'Brien, E., Sivatchenko, A. & Kerber, R. A. *2003* Association between telomere length in blood and mortality in people aged 60 years or older. *Lancet* **361**(9355), 393-395.
203 Blasco, M. A. *2005* Telomeres and human disease: ageing, cancer and beyond. *Nat Rev Genet* **6**(8), 611-622.
204 Passos, J. F., Saretzki, G. & von Zglinicki, T. *2007* DNA damage in telomeres and mitochondria during cellular senescence: is there a connection? *Nucleic Acids Res* **35**(22), 7505-7513.
205 Ressler, S., Bartkova, J., Niederegger, H., Bartek, J., Scharffetter-Kochanek, K., Jansen-Durr, P. & Wlaschek, M. *2006* p16INK4A is a robust in vivo biomarker of cellular aging in human skin. *Aging cell* **5**(5), 379-389.
206 Krishnamurthy, J., Torrice, C., Ramsey, M. R., Kovalev, G. I., Al-Regaiey, K., Su, L. & Sharpless, N. E. *2004* Ink4a/Arf expression is a biomarker of aging. *The Journal of clinical investigation* **114**(9), 1299-1307.
207 Ito, K., Hirao, A., Arai, F., Takubo, K., Matsuoka, S., Miyamoto, K., Ohmura, M., Naka, K., Hosokawa, K., Ikeda, Y. & Suda, T. *2006* Reactive oxygen species act through p38 MAPK to limit the lifespan of hematopoietic stem cells. *Nat Med* **12**(4), 446-451.
208 Jeck, W. R., Siebold, A. P. & Sharpless, N. E. *2012* Review: a meta-analysis of GWAS and age-associated diseases. *Aging Cell* **11**(5), 727-731.
209 Hill, M., Wernig, A. & Goldspink, G. *2003* Muscle satellite (stem) cell activation during local tissue injury and repair. *Journal of anatomy* **203**(1), 89-99.
210 Chicas, A., Wang, X. W., Zhang, C. L., McCurrach, M., Zhao, Z., Mert, O., Dickins, R. A., Narita, M., Zhang, M. & Lowe, S. W. *2010* Dissecting the Unique Role of the Retinoblastoma Tumor Suppressor during Cellular Senescence. *Cancer cell* **17**(4), 376-387.
211 Tanaka, T., Narazaki, M. & Kishimoto, T. *2014* IL-6 in inflammation, immunity, and disease. *Cold Spring Harb Perspect Biol* **6**(10), a016295.
212 Weber, A., Wasiliew, P. & Kracht, M. *2010* Interleukin-1 (IL-1) pathway. *Sci Signal* **3**(105), cm1.
213 Proost, P., Wuyts, A. & van Damme, J. *1996* The role of chemokines in inflammation. *Int J Clin Lab Res* **26**(4), 211-223.
214 West, M. D., Pereira-Smith, O. M. & Smith, J. R. *1989* Replicative senescence of human skin fibroblasts correlates with a loss of regulation and overexpression of collagenase activity. *Exp Cell Res* **184**(1), 138-147.
215 Rodriguez, D., Morrison, C. J. & Overall, C. M. *2010* Matrix metalloproteinases: what do they not do? New substrates and biological roles identified by murine models and proteomics. *Biochim Biophys Acta* **1803**(1), 39-54.
216 Kumazaki, T., Kobayashi, M. & Mitsui, Y. *1993* Enhanced expression of fibronectin during in vivo cellular aging of human vascular endothelial cells and skin fibroblasts. *Exp Cell Res* **205**(2), 396-402.
217 Rasoamanantena, P., Thweatt, R., Labat-Robert, J. & Goldstein, S. *1994* Altered regulation of fibronectin gene expression in Werner syndrome fibroblasts. *Exp Cell Res* **213**(1), 121-127.
218 Austin, S. & St-Pierre, J. *2012* PGC1alpha and mitochondrial metabolism--emerging concepts and relevance in ageing and neurodegenerative disorders. *Journal of cell science* **125**(Pt 21), 4963-4971.
219 Fukui, M. & Zhu, B. T. *2010* Mitochondrial superoxide dismutase SOD2, but not cytosolic SOD1, plays a critical role in protection against glutamate-induced oxidative stress and cell death in HT22 neuronal cells. *Free Radic Biol Med* **48**(6), 821-830.
220 Dizdaroglu, M., Jaruga, P., Birincioglu, M. & Rodriguez, H. *2002* Free radical-induced damage to DNA: mechanisms and measurement. *Free Radic Biol Med* **32**(11), 1102-1115.
221 Mecocci, P., MacGarvey, U., Kaufman, A. E., Koontz, D., Shoffner, J. M., Wallace, D. C. & Beal, M. F. *1993* Oxidative damage to mitochondrial DNA shows marked age-dependent increases in human brain. *Ann Neurol* **34**(4), 609-616.

222 Rossi, D. J., Bryder, D., Seita, J., Nussenzweig, A., Hoeijmakers, J. & Weissman, I. L. *2007* Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature* **447**(7145), 725-729.
223 Taanman, J. W. *1999* The mitochondrial genome: structure, transcription, translation and replication. *Biochim Biophys Acta* **1410**(2), 103-123.
224 Del Bo, R., Bordoni, A., Martinelli Boneschi, F., Crimi, M., Sciacco, M., Bresolin, N., Scarlato, G. & Comi, G. P. *2002* Evidence and age-related distribution of mtDNA D-loop point mutations in skeletal muscle from healthy subjects and mitochondrial patients. *J Neurol Sci* **202**(1-2), 85-91.
225 Wang, C. & Youle, R. J. *2009* The role of mitochondria in apoptosis*. *Annual review of genetics* **43**95-118.
226 Zheng, J., Edelman, S. W., Tharmarajah, G., Walker, D. W., Pletcher, S. D. & Seroude, L. *2005* Differential patterns of apoptosis in response to aging in Drosophila. *Proc Natl Acad Sci U S A* **102**(34), 12083-12088.
227 Mather, M. & Rottenberg, H. *2000* Aging enhances the activation of the permeability transition pore in mitochondria. *Biochem Biophys Res Commun* **273**(2), 603-608.
228 Yan, L. J. & Sohal, R. S. *1998* Mitochondrial adenine nucleotide translocase is modified oxidatively during aging. *Proc Natl Acad Sci U S A* **95**(22), 12896-12901.
229 Franceschi, C., Capri, M., Monti, D., Giunta, S., Olivieri, F., Sevini, F., Panourgia, M. P., Invidia, L., Celani, L., Scurti, M., Cevenini, E., Castellani, G. C. & Salvioli, S. *2007* Inflammaging and anti-inflammaging: a systemic perspective on aging and longevity emerged from studies in humans. *Mech Ageing Dev* **128**(1), 92-105.
230 Michalakis, K., Goulis, D. G., Vazaiou, A., Mintziori, G., Polymeris, A. & Abrahamian-Michalakis, A. *2013* Obesity in the ageing man. *Metabolism: clinical and experimental* **62**(10), 1341-1349.
231 Makki, K., Froguel, P. & Wolowczuk, I. *2013* Adipose tissue in obesity-related inflammation and insulin resistance: cells, cytokines, and chemokines. *ISRN Inflamm* **2013**139239.
232 Michaud, M., Balardy, L., Moulis, G., Gaudin, C., Peyrot, C., Vellas, B., Cesari, M. & Nourhashemi, F. *2013* Proinflammatory cytokines, aging, and age-related diseases. *J Am Med Dir Assoc* **14**(12), 877-882.
233 Leng, S. X., Xue, Q. L., Tian, J., Walston, J. D. & Fried, L. P. *2007* Inflammation and frailty in older women. *Journal of the American Geriatrics Society* **55**(6), 864-871.
234 Puzianowska-Kuznicka, M., Owczarz, M., Wieczorowska-Tobis, K., Nadrowski, P., Chudek, J., Slusarczyk, P., Skalska, A., Jonas, M., Franek, E. & Mossakowska, M. *2016* Interleukin-6 and C-reactive protein, successful aging, and mortality: the PolSenior study. *Immun Ageing* **13**21.
235 Visser, M., Pahor, M., Taaffe, D. R., Goodpaster, B. H., Simonsick, E. M., Newman, A. B., Nevitt, M. & Harris, T. B. *2002* Relationship of interleukin-6 and tumor necrosis factor-alpha with muscle mass and muscle strength in elderly men and women: the Health ABC Study. *J Gerontol A Biol Sci Med Sci* **57**(5), M326-332.
236 Zoico, E. & Roubenoff, R. *2002* The role of cytokines in regulating protein metabolism and muscle function. *Nutr Rev* **60**(2), 39-51.
237 Schaap, L. A., Pluijm, S. M., Deeg, D. J. & Visser, M. *2006* Inflammatory markers and loss of muscle mass (sarcopenia) and strength. *The American journal of medicine* **119**(6), 526 e529-517.
238 Mitch, W. E. & Goldberg, A. L. *1996* Mechanisms of muscle wasting. The role of the ubiquitin-proteasome pathway. *The New England journal of medicine* **335**(25), 1897-1905.
239 Franceschi, C. & Campisi, J. *2014* Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. *J Gerontol A Biol Sci Med Sci* **69 Suppl 1**S4-9.
240 Zhang, Q., Raoof, M., Chen, Y., Sumi, Y., Sursal, T., Junger, W., Brohi, K., Itagaki, K. & Hauser, C. J. *2010* Circulating mitochondrial DAMPs cause inflammatory responses to injury. *Nature* **464**(7285), 104-107.

241 Gruver, A. L., Hudson, L. L. & Sempowski, G. D. *2007* Immunosenescence of ageing. *J Pathol* **211**(2), 144-156.
242 Guazzo, E. P., Kirkpatrick, P. J., Goodyer, I. M., Shiers, H. M. & Herbert, J. *1996* Cortisol, dehydroepiandrosterone (DHEA), and DHEA sulfate in the cerebrospinal fluid of man: relation to blood levels and the effects of age. *The Journal of clinical endocrinology and metabolism* **81**(11), 3951-3960.
243 Drey, M. *2011* Sarcopenia - pathophysiology and clinical relevance. *Wien Med Wochenschr* **161**(17-18), 402-408.
244 Kim, T. N. & Choi, K. M. *2013* Sarcopenia: definition, epidemiology, and pathophysiology. *J Bone Metab* **20**(1), 1-10.
245 Peters, A. *2002* The effects of normal aging on myelin and nerve fibers: a review. *J Neurocytol* **31**(8-9), 581-593.
246 Fahim, M. A. & Robbins, N. *1982* Ultrastructural studies of young and old mouse neuromuscular junctions. *J Neurocytol* **11**(4), 641-656.
247 Balice-Gordon, R. J. *1997* Age-related changes in neuromuscular innervation. *Muscle Nerve Suppl* **5**S83-87.
248 Griggs, R. C., Kingston, W., Jozefowicz, R. F., Herr, B. E., Forbes, G. & Halliday, D. *1989* Effect of testosterone on muscle mass and muscle protein synthesis. *J Appl Physiol (1985)* **66**(1), 498-503.
249 Snyder, P. J., Peachey, H., Hannoush, P., Berlin, J. A., Loh, L., Lenrow, D. A., Holmes, J. H., Dlewati, A., Santanna, J., Rosen, C. J. & Strom, B. L. *1999* Effect of testosterone treatment on body composition and muscle strength in men over 65 years of age. *J Clin Endocrinol Metab* **84**(8), 2647-2653.
250 Mudali, S. & Dobs, A. S. *2004* Effects of testosterone on body composition of the aging male. *Mech Ageing Dev* **125**(4), 297-304.
251 Schiaffino, S. & Mammucari, C. *2011* Regulation of skeletal muscle growth by the IGF1-Akt/PKB pathway: insights from genetic models. *Skeletal muscle* **1**(1), 4.
252 Sacheck, J. M., Ohtsuka, A., McLary, S. C. & Goldberg, A. L. *2004* IGF-I stimulates muscle growth by suppressing protein breakdown and expression of atrophy-related ubiquitin ligases, atrogin-1 and MuRF1. *American journal of physiology. Endocrinology and metabolism* **287**(4), E591-601.
253 Zadik, Z., Chalew, S. A., McCarter, R. J., Jr., Meistas, M. & Kowarski, A. A. *1985* The influence of age on the 24-hour integrated concentration of growth hormone in normal individuals. *J Clin Endocrinol Metab* **60**(3), 513-516.
254 Papadakis, M. A., Grady, D., Black, D., Tierney, M. J., Gooding, G. A., Schambelan, M. & Grunfeld, C. *1996* Growth hormone replacement in healthy older men improves body composition but not functional ability. *Ann Intern Med* **124**(8), 708-716.
255 Ceglia, L. & Harris, S. S. *2013* Vitamin D and its role in skeletal muscle. *Calcif Tissue Int* **92**(2), 151-162.
256 Kim, T. N., Park, M. S., Lim, K. I., Choi, H. Y., Yang, S. J., Yoo, H. J., Kang, H. J., Song, W., Choi, H., Baik, S. H., Choi, D. S. & Choi, K. M. *2013* Relationships between sarcopenic obesity and insulin resistance, inflammation, and vitamin D status: the Korean Sarcopenic Obesity Study. *Clin Endocrinol (Oxf)* **78**(4), 525-532.
257 Recker, R., Lips, P., Felsenberg, D., Lippuner, K., Benhamou, L., Hawkins, F., Delmas, P. D., Rosen, C., Emkey, R., Salzmann, G., He, W. & Santora, A. C. *2006* Alendronate with and without cholecalciferol for osteoporosis: results of a 15-week randomized controlled trial. *Curr Med Res Opin* **22**(9), 1745-1755.
258 Chen, L., Huang, S., Lee, L., Davalos, A., Schiestl, R. H., Campisi, J. & Oshima, J. *2003* WRN, the protein deficient in Werner syndrome, plays a critical structural role in optimizing DNA repair. *Aging cell* **2**(4), 191-199.
259 Mostoslavsky, R., Chua, K. F., Lombard, D. B., Pang, W. W., Fischer, M. R., Gellon, L., Liu, P., Mostoslavsky, G., Franco, S., Murphy, M. M., Mills, K. D., Patel, P., Hsu, J. T., Hong, A. L., Ford, E., Cheng, H. L., Kennedy, C., Nunez, N., Bronson, R., Frendewey, D., Auerbach, W., Valenzuela, D., Karow, M., Hottiger, M. O., Hursting, S., Barrett, J. C., Guarente, L.,

Mulligan, R., Demple, B., Yancopoulos, G. D. & Alt, F. W. *2006* Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. *Cell* **124**(2), 315-329.

260 Harries, L. W., Hernandez, D., Henley, W., Wood, A. R., Holly, A. C., Bradley-Smith, R. M., Yaghootkar, H., Dutta, A., Murray, A., Frayling, T. M., Guralnik, J. M., Bandinelli, S., Singleton, A., Ferrucci, L. & Melzer, D. *2011* Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. *Aging Cell* **10**(5), 868-878.

261 Zacharewicz, E., Della Gatta, P., Reynolds, J., Garnham, A., Crowley, T., Russell, A. P. & Lamon, S. *2014* Identification of microRNAs linked to regulators of muscle protein synthesis and regeneration in young and old skeletal muscle. *PloS one* **9**(12), e114009.

262 Morrow, G., Samson, M., Michaud, S. & Tanguay, R. M. *2004* Overexpression of the small mitochondrial Hsp22 extends Drosophila life span and increases resistance to oxidative stress. *FASEB J* **18**(3), 598-599.

263 Min, J. N., Whaley, R. A., Sharpless, N. E., Lockyer, P., Portbury, A. L. & Patterson, C. *2008* CHIP deficiency decreases longevity, with accelerated aging phenotypes accompanied by altered protein quality control. *Mol Cell Biol* **28**(12), 4018-4025.

264 Koga, H., Kaushik, S. & Cuervo, A. M. *2011* Protein homeostasis and aging: The importance of exquisite quality control. *Ageing Res Rev* **10**(2), 205-215.

265 Attaix, D., Ventadour, S., Codran, A., Bechet, D., Taillandier, D. & Combaret, L. *2005* The ubiquitin-proteasome system and skeletal muscle wasting. *Essays Biochem* **41**173-186.

266 Combaret, L., Dardevet, D., Bechet, D., Taillandier, D., Mosoni, L. & Attaix, D. *2009* Skeletal muscle proteolysis in aging. *Curr Opin Clin Nutr Metab Care* **12**(1), 37-41.

267 Hepple, R. T., Qin, M., Nakamoto, H. & Goto, S. *2008* Caloric restriction optimizes the proteasome pathway with aging in rat plantaris muscle: implications for sarcopenia. *American journal of physiology. Regulatory, integrative and comparative physiology* **295**(4), R1231-1237.

268 Pandey, U. B., Nie, Z., Batlevi, Y., McCray, B. A., Ritson, G. P., Nedelsky, N. B., Schwartz, S. L., DiProspero, N. A., Knight, M. A., Schuldiner, O., Padmanabhan, R., Hild, M., Berry, D. L., Garza, D., Hubbert, C. C., Yao, T. P., Baehrecke, E. H. & Taylor, J. P. *2007* HDAC6 rescues neurodegeneration and provides an essential link between autophagy and the UPS. *Nature* **447**(7146), 859-863.

269 Johnson, S. C., Rabinovitch, P. S. & Kaeberlein, M. *2013* mTOR is a key modulator of ageing and age-related disease. *Nature* **493**(7432), 338-345.

270 Ruzankina, Y. & Brown, E. J. *2007* Relationships between stem cell exhaustion, tumour suppression and ageing. *British journal of cancer* **97**(9), 1189-1193.

271 Lindholm, M. E., Huss, M., Solnestam, B. W., Kjellqvist, S., Lundeberg, J. & Sundberg, C. J. *2014* The human skeletal muscle transcriptome: sex differences, alternative splicing, and tissue homogeneity assessed with RNA sequencing. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **28**(10), 4571-4581.

272 Su, J., Ekman, C., Oskolkov, N., Lahti, L., Strom, K., Brazma, A., Groop, L., Rung, J. & Hansson, O. *2015* A novel atlas of gene expression in human skeletal muscle reveals molecular changes associated with aging. *Skeletal muscle* **5**35.

273 Stephens, N. A., Xie, H., Johannsen, N. M., Church, T. S., Smith, S. R. & Sparks, L. M. *2015* A transcriptional signature of "exercise resistance" in skeletal muscle of individuals with type 2 diabetes mellitus. *Metabolism: clinical and experimental* **64**(9), 999-1004.

274 Neubauer, O., Sabapathy, S., Ashton, K. J., Desbrow, B., Peake, J. M., Lazarus, R., Wessner, B., Cameron-Smith, D., Wagner, K. H., Haseler, L. J. & Bulmer, A. C. *2014* 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* **116**(3), 274-287.

275 Sreekumar, R., Halvatsiotis, P., Schimke, J. C. & Nair, K. S. *2002* Gene expression profile in skeletal muscle of type 2 diabetes and the effect of insulin treatment. *Diabetes* **51**(6), 1913-1920.

276 Noguchi, S., Tsukahara, T., Fujita, M., Kurokawa, R., Tachikawa, M., Toda, T., Tsujimoto, A., Arahata, K. & Nishino, I. *2003* cDNA microarray analysis of individual Duchenne muscular dystrophy patients. *Human molecular genetics* **12**(6), 595-600.
277 Fluck, M. *2006* Functional, structural and molecular plasticity of mammalian skeletal muscle in response to exercise stimuli. *The Journal of experimental biology* **209**(Pt 12), 2239-2248.
278 Baldwin, K. M. & Haddad, F. *2002* Skeletal muscle plasticity: cellular and molecular responses to altered physical activity paradigms. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists* **81**(11 Suppl), S40-51.
279 Matsakas, A. & Patel, K. *2009* Skeletal muscle fibre plasticity in response to selected environmental and physiological stimuli. *Histology and histopathology* **24**(5), 611-629.
280 Guerra, B., Gomez-Cabrera, M. C., Ponce-Gonzalez, J. G., Martinez-Bello, V. E., Guadalupe-Grau, A., Santana, A., Sebastia, V., Vina, J. & Calbet, J. A. *2011* Repeated muscle biopsies through a single skin incision do not elicit muscle signaling, but IL-6 mRNA and STAT3 phosphorylation increase in injured muscle. *Journal of applied physiology* **110**(6), 1708-1715.
281 Lundby, C., Nordsborg, N., Kusuhara, K., Kristensen, K. M., Neufer, P. D. & Pilegaard, H. *2005* Gene expression in human skeletal muscle: alternative normalization method and effect of repeated biopsies. *European journal of applied physiology* **95**(4), 351-360.
282 Welle, S., Tawil, R. & Thornton, C. A. *2008* Sex-related differences in gene expression in human skeletal muscle. *PloS one* **3**(1), e1385.
283 Staron, R. S., Hagerman, F. C., Hikida, R. S., Murray, T. F., Hostler, D. P., Crill, M. T., Ragg, K. E. & Toma, K. *2000* Fiber type composition of the vastus lateralis muscle of young men and women. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* **48**(5), 623-629.
284 Kent-Braun, J. A., Ng, A. V., Doyle, J. W. & Towse, T. F. *2002* Human skeletal muscle responses vary with age and gender during fatigue due to incremental isometric exercise. *Journal of applied physiology* **93**(5), 1813-1823.
285 Forsberg, A. M., Nilsson, E., Werneman, J., Bergstrom, J. & Hultman, E. *1991* Muscle composition in relation to age and sex. *Clinical science* **81**(2), 249-256.
286 Brodsky, I. G., Balagopal, P. & Nair, K. S. *1996* Effects of testosterone replacement on muscle mass and muscle protein synthesis in hypogonadal men--a clinical research center study. *The Journal of clinical endocrinology and metabolism* **81**(10), 3469-3475.
287 Ringold, G. M. *1985* Steroid hormone regulation of gene expression. *Annual review of pharmacology and toxicology* **25**529-566.
288 Drummond, M. J., McCarthy, J. J., Fry, C. S., Esser, K. A. & Rasmussen, B. B. *2008* Aging differentially affects human skeletal muscle microRNA expression at rest and after an anabolic stimulus of resistance exercise and essential amino acids. *American journal of physiology. Endocrinology and metabolism* **295**(6), E1333-1340.
289 Kim, J. Y., Park, Y. K., Lee, K. P., Lee, S. M., Kang, T. W., Kim, H. J., Dho, S. H., Kim, S. Y. & Kwon, K. S. *2014* Genome-wide profiling of the microRNA-mRNA regulatory network in skeletal muscle with aging. *Aging* **6**(7), 524-544.
290 Welle, S., Brooks, A. I., Delehanty, J. M., Needler, N., Bhatt, K., Shah, B. & Thornton, C. A. *2004* Skeletal muscle gene expression profiles in 20-29 year old and 65-71 year old women. *Experimental gerontology* **39**(3), 369-377.
291 Welle, S., Brooks, A. I., Delehanty, J. M., Needler, N. & Thornton, C. A. *2003* Gene expression profile of aging in human muscle. *Physiological genomics* **14**(2), 149-159.
292 Lamon, S., Wallace, M. A. & Russell, A. P. *2014* The STARS signaling pathway: a key regulator of skeletal muscle function. *Pflugers Arch* **466**(9), 1659-1671.
293 Deng, H. W., Lai, D. B., Conway, T., Li, J., Xu, F. H., Davies, K. M. & Recker, R. R. *2001* Characterization of genetic and lifestyle factors for determining variation in body mass index, fat mass, percentage of fat mass, and lean mass. *Journal of clinical densitometry : the official journal of the International Society for Clinical Densitometry* **4**(4), 353-361.

294 Hsu, F. C., Lenchik, L., Nicklas, B. J., Lohman, K., Register, T. C., Mychaleckyj, J., Langefeld, C. D., Freedman, B. I., Bowden, D. W. & Carr, J. J. *2005* Heritability of body composition measured by DXA in the diabetes heart study. *Obesity research* **13**(2), 312-319.

295 Prior, S. J., Roth, S. M., Wang, X., Kammerer, C., Miljkovic-Gacic, I., Bunker, C. H., Wheeler, V. W., Patrick, A. L. & Zmuda, J. M. *2007* Genetic and environmental influences on skeletal muscle phenotypes as a function of age and sex in large, multigenerational families of African heritage. *Journal of applied physiology* **103**(4), 1121-1127.

296 Tiainen, K., Sipila, S., Alen, M., Heikkinen, E., Kaprio, J., Koskenvuo, M., Tolvanen, A., Pajala, S. & Rantanen, T. *2004* Heritability of maximal isometric muscle strength in older female twins. *Journal of applied physiology* **96**(1), 173-180.

297 Zillikens, M. C., Demissie, S., Hsu, Y. H., Yerges-Armstrong, L. M., Chou, W. C., Stolk, L., Livshits, G., Broer, L., Johnson, T., Koller, D. L., Kutalik, Z., Luan, J., Malkin, I., Ried, J. S., Smith, A. V., Thorleifsson, G., Vandenput, L., Hua Zhao, J., Zhang, W., Aghdassi, A., Akesson, K., Amin, N., Baier, L. J., Barroso, I., Bennett, D. A., Bertram, L., Biffar, R., Bochud, M., Boehnke, M., Borecki, I. B., Buchman, A. S., Byberg, L., Campbell, H., Campos Obanda, N., Cauley, J. A., Cawthon, P. M., Cederberg, H., Chen, Z., Cho, N. H., Jin Choi, H., Claussnitzer, M., Collins, F., Cummings, S. R., De Jager, P. L., Demuth, I., Dhonukshe-Rutten, R. A. M., Diatchenko, L., Eiriksdottir, G., Enneman, A. W., Erdos, M., Eriksson, J. G., Eriksson, J., Estrada, K., Evans, D. S., Feitosa, M. F., Fu, M., Garcia, M., Gieger, C., Girke, T., Glazer, N. L., Grallert, H., Grewal, J., Han, B. G., Hanson, R. L., Hayward, C., Hofman, A., Hoffman, E. P., Homuth, G., Hsueh, W. C., Hubal, M. J., Hubbard, A., Huffman, K. M., Husted, L. B., Illig, T., Ingelsson, E., Ittermann, T., Jansson, J. O., Jordan, J. M., Jula, A., Karlsson, M., Khaw, K. T., Kilpelainen, T. O., Klopp, N., Kloth, J. S. L., Koistinen, H. A., Kraus, W. E., Kritchevsky, S., Kuulasmaa, T., Kuusisto, J., Laakso, M., Lahti, J., Lang, T., Langdahl, B. L., Launer, L. J., Lee, J. Y., Lerch, M. M., Lewis, J. R., Lind, L., Lindgren, C., Liu, Y., Liu, T., Liu, Y., Ljunggren, O., Lorentzon, M., Luben, R. N., Maixner, W., McGuigan, F. E., Medina-Gomez, C., Meitinger, T., Melhus, H., Mellstrom, D., Melov, S., Michaelsson, K., Mitchell, B. D., Morris, A. P., Mosekilde, L., Newman, A., Nielson, C. M., O'Connell, J. R., Oostra, B. A., Orwoll, E. S., Palotie, A., Parker, S. C. J., Peacock, M., Perola, M., Peters, A., Polasek, O., Prince, R. L., Raikkonen, K., Ralston, S. H., Ripatti, S., Robbins, J. A., Rotter, J. I., Rudan, I., Salomaa, V., Satterfield, S., Schadt, E. E., Schipf, S., Scott, L., Sehmi, J., Shen, J., Soo Shin, C., Sigurdsson, G., Smith, S., Soranzo, N., Stancakova, A., Steinhagen-Thiessen, E., Streeten, E. A., Styrkarsdottir, U., Swart, K. M. A., Tan, S. T., Tarnopolsky, M. A., Thompson, P., Thomson, C. A., Thorsteinsdottir, U., Tikkanen, E., Tranah, G. J., Tuomilehto, J., van Schoor, N. M., Verma, A., Vollenweider, P., Volzke, H., Wactawski-Wende, J., Walker, M., Weedon, M. N., Welch, R., Wichmann, H. E., Widen, E., Williams, F. M. K., Wilson, J. F., Wright, N. C., Xie, W., Yu, L., Zhou, Y., Chambers, J. C., Doring, A., van Duijn, C. M., Econs, M. J., Gudnason, V., Kooner, J. S., Psaty, B. M., Spector, T. D., Stefansson, K., Rivadeneira, F., Uitterlinden, A. G., Wareham, N. J., Ossowski, V., Waterworth, D., Loos, R. J. F., Karasik, D., Harris, T. B., Ohlsson, C. & Kiel, D. P. *2017* Large meta-analysis of genome-wide association studies identifies five loci for lean body mass. *Nat Commun* **8**(1), 80.

298 Matteini, A. M., Tanaka, T., Karasik, D., Atzmon, G., Chou, W. C., Eicher, J. D., Johnson, A. D., Arnold, A. M., Callisaya, M. L., Davies, G., Evans, D. S., Holtfreter, B., Lohman, K., Lunetta, K. L., Mangino, M., Smith, A. V., Smith, J. A., Teumer, A., Yu, L., Arking, D. E., Buchman, A. S., Chibinik, L. B., De Jager, P. L., Evans, D. A., Faul, J. D., Garcia, M. E., Gillham-Nasenya, I., Gudnason, V., Hofman, A., Hsu, Y. H., Ittermann, T., Lahousse, L., Liewald, D. C., Liu, Y., Lopez, L., Rivadeneira, F., Rotter, J. I., Siggeirsdottir, K., Starr, J. M., Thomson, R., Tranah, G. J., Uitterlinden, A. G., Volker, U., Volzke, H., Weir, D. R., Yaffe, K., Zhao, W., Zhuang, W. V., Zmuda, J. M., Bennett, D. A., Cummings, S. R., Deary, I. J., Ferrucci, L., Harris, T. B., Kardia, S. L., Kocher, T., Kritchevsky, S. B., Psaty, B. M., Seshadri, S., Spector, T. D., Srikanth, V. K., Windham, B. G., Zillikens, M. C., Newman, A. B., Walston, J. D., Kiel, D. P. & Murabito, J. M. *2016* GWAS analysis of handgrip and lower body strength in older adults in the CHARGE consortium. *Aging cell* **15**(5), 792-800.

299 Harries, L. W., Pilling, L. C., Hernandez, L. D. G., Bradley-Smith, R., Henley, W., Singleton, A. B., Guralnik, J. M., Bandinelli, S., Ferrucci, L. & Melzer, D. *2012* CCAAT-enhancer-binding protein-beta expression in vivo is associated with muscle strength. *Aging cell* **11**(2), 262-268.

300 Iolascon, G., Di Pietro, G., Gimigliano, F., Mauro, G. L., Moretti, A., Giamattei, M. T., Ortolani, S., Tarantino, U. & Brandi, M. L. *2014* Physical exercise and sarcopenia in older people: position paper of the Italian Society of Orthopaedics and Medicine (OrtoMed). *Clin Cases Miner Bone Metab* **11**(3), 215-221.

301 Ferrando, A. A., Sheffield-Moore, M., Yeckel, C. W., Gilkison, C., Jiang, J., Achacosa, A., Lieberman, S. A., Tipton, K., Wolfe, R. R. & Urban, R. J. *2002* Testosterone administration to older men improves muscle function: molecular and physiological mechanisms. *Am J Physiol Endocrinol Metab* **282**(3), E601-607.

302 Bakhshi, V., Elliott, M., Gentili, A., Godschalk, M. & Mulligan, T. *2000* Testosterone improves rehabilitation outcomes in ill older men. *J Am Geriatr Soc* **48**(5), 550-553.

303 Annweiler, C., Schott, A. M., Berrut, G., Fantino, B. & Beauchet, O. *2009* Vitamin D-related changes in physical performance: a systematic review. *J Nutr Health Aging* **13**(10), 893-898.

304 Morley, J. E., Argiles, J. M., Evans, W. J., Bhasin, S., Cella, D., Deutz, N. E., Doehner, W., Fearon, K. C., Ferrucci, L., Hellerstein, M. K., Kalantar-Zadeh, K., Lochs, H., MacDonald, N., Mulligan, K., Muscaritoli, M., Ponikowski, P., Posthauer, M. E., Rossi Fanelli, F., Schambelan, M., Schols, A. M., Schuster, M. W., Anker, S. D., Society for Sarcopenia, C. & Wasting, D. *2010* Nutritional recommendations for the management of sarcopenia. *J Am Med Dir Assoc* **11**(6), 391-396.

305 Abadir, P. M., Walston, J. D. & Carey, R. M. *2012* Subcellular characteristics of functional intracellular renin-angiotensin systems. *Peptides* **38**(2), 437-445.

306 Zoll, J., Monassier, L., Garnier, A., N'Guessan, B., Mettauer, B., Veksler, V., Piquard, F., Ventura-Clapier, R. & Geny, B. *2006* ACE inhibition prevents myocardial infarction-induced skeletal muscle mitochondrial dysfunction. *Journal of applied physiology* **101**(2), 385-391.

307 Di Bari, M., van de Poll-Franse, L. V., Onder, G., Kritchevsky, S. B., Newman, A., Harris, T. B., Williamson, J. D., Marchionni, N. & Pahor, M. *2004* Antihypertensive medications and differences in muscle mass in older persons: The Health, Aging and Body Composition study. *Journal of the American Geriatrics Society* **52**(6), 961-966.

308 Onder, G., Penninx, B. W. J. H., Balkrishnan, R., Fried, L. P., Chaves, P. H. M., Williamson, J., Carter, C., Di Bari, M., Guralnik, J. M. & Pahor, M. *2002* Relation between use of angiotensin-converting enzyme inhibitors and muscle strength and physical function in older women: an observational study. *Lancet* **359**(9310), 926-930.

309 Heijmans, B. T., Tobi, E. W., Stein, A. D., Putter, H., Blauw, G. J., Susser, E. S., Slagboom, P. E. & Lumey, L. H. *2008* Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences of the United States of America* **105**(44), 17046-17049.

310 Nishioka, M., Bundo, M., Kasai, K. & Iwamoto, K. *2012* DNA methylation in schizophrenia: progress and challenges of epigenetic studies. *Genome Med* **4**(12), 96.

311 McGowan, P. O., Sasaki, A., D'Alessio, A. C., Dymov, S., Labonte, B., Szyf, M., Turecki, G. & Meaney, M. J. *2009* Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nat Neurosci* **12**(3), 342-348.

312 Sadeh, N., Spielberg, J. M., Logue, M. W., Wolf, E. J., Smith, A. K., Lusk, J., Hayes, J. P., Sperbeck, E., Milberg, W. P., McGlinchey, R. E., Salat, D. H., Carter, W. C., Stone, A., Schichman, S. A., Humphries, D. E. & Miller, M. W. *2016* SKA2 methylation is associated with decreased prefrontal cortical thickness and greater PTSD severity among trauma-exposed veterans. *Mol Psychiatr* **21**(3), 357-363.

313 Palma-Gudiel, H., Cordova-Palomera, A., Eixarch, E., Deuschle, M. & Fananas, L. *2015* Maternal psychosocial stress during pregnancy alters the epigenetic signature of the glucocorticoid receptor gene promoter in their offspring: a meta-analysis. *Epigenetics : official journal of the DNA Methylation Society* **10**(10), 893-902.

314 Lopez-Serra, P. & Esteller, M. *2012* DNA methylation-associated silencing of tumor-suppressor microRNAs in cancer. *Oncogene* **31**(13), 1609-1622.
315 Bansal, A. & Pinney, S. E. *2017* DNA methylation and its role in the pathogenesis of diabetes. *Pediatr Diabetes* **18**(3), 167-177.
316 Irier, H. A. & Jin, P. *2012* Dynamics of DNA methylation in aging and Alzheimer's disease. *DNA Cell Biol* **31 Suppl 1**S42-48.
317 Ziller, M. J., Muller, F., Liao, J., Zhang, Y., Gu, H., Bock, C., Boyle, P., Epstein, C. B., Bernstein, B. E., Lengauer, T., Gnirke, A. & Meissner, A. *2011* Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS genetics* **7**(12), e1002389.
318 Patil, V., Ward, R. L. & Hesson, L. B. *2014* The evidence for functional non-CpG methylation in mammalian cells. *Epigenetics : official journal of the DNA Methylation Society* **9**(6), 823-828.
319 Xiong, J. & Epstein, R. J. *2009* Growth inhibition of human cancer cells by 5-aza-2'-deoxycytidine does not correlate with its effects on INK4a/ARF expression or initial promoter methylation status. *Molecular cancer therapeutics* **8**(4), 779-785.
320 Yang, X., Han, H., De Carvalho, D. D., Lay, F. D., Jones, P. A. & Liang, G. *2014* Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer cell* **26**(4), 577-590.
321 Law, J. A. & Jacobsen, S. E. *2010* Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* **11**(3), 204-220.
322 Chedin, F. *2011* The DNMT3 family of mammalian de novo DNA methyltransferases. *Prog Mol Biol Transl Sci* **101**255-285.
323 Yoon, A., Tammen, S. A., Park, S., Han, S. N. & Choi, S. W. *2017* Genome-wide hepatic DNA methylation changes in high-fat diet-induced obese mice. *Nutrition research and practice* **11**(2), 105-113.
324 Joubert, B. R., Felix, J. F., Yousefi, P., Bakulski, K. M., Just, A. C., Breton, C., Reese, S. E., Markunas, C. A., Richmond, R. C., Xu, C. J., Kupers, L. K., Oh, S. S., Hoyo, C., Gruzieva, O., Soderhall, C., Salas, L. A., Baiz, N., Zhang, H., Lepeule, J., Ruiz, C., Ligthart, S., Wang, T., Taylor, J. A., Duijts, L., Sharp, G. C., Jankipersadsing, S. A., Nilsen, R. M., Vaez, A., Fallin, M. D., Hu, D., Litonjua, A. A., Fuemmeler, B. F., Huen, K., Kere, J., Kull, I., Munthe-Kaas, M. C., Gehring, U., Bustamante, M., Saurel-Coubizolles, M. J., Quraishi, B. M., Ren, J., Tost, J., Gonzalez, J. R., Peters, M. J., Haberg, S. E., Xu, Z., van Meurs, J. B., Gaunt, T. R., Kerkhof, M., Corpeleijn, E., Feinberg, A. P., Eng, C., Baccarelli, A. A., Benjamin Neelon, S. E., Bradman, A., Merid, S. K., Bergstrom, A., Herceg, Z., Hernandez-Vargas, H., Brunekreef, B., Pinart, M., Heude, B., Ewart, S., Yao, J., Lemonnier, N., Franco, O. H., Wu, M. C., Hofman, A., McArdle, W., Van der Vlies, P., Falahi, F., Gillman, M. W., Barcellos, L. F., Kumar, A., Wickman, M., Guerra, S., Charles, M. A., Holloway, J., Auffray, C., Tiemeier, H. W., Smith, G. D., Postma, D., Hivert, M. F., Eskenazi, B., Vrijheid, M., Arshad, H., Anto, J. M., Dehghan, A., Karmaus, W., Annesi-Maesano, I., Sunyer, J., Ghantous, A., Pershagen, G., Holland, N., Murphy, S. K., DeMeo, D. L., Burchard, E. G., Ladd-Acosta, C., Snieder, H., Nystad, W., Koppelman, G. H., Relton, C. L., Jaddoe, V. W., Wilcox, A., Melen, E. & London, S. J. *2016* DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am J Hum Genet* **98**(4), 680-696.
325 Probst, A. V., Dunleavy, E. & Almouzni, G. *2009* Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol* **10**(3), 192-206.
326 Chen, T., Hevi, S., Gay, F., Tsujimoto, N., He, T., Zhang, B., Ueda, Y. & Li, E. *2007* Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. *Nat Genet* **39**(3), 391-396.
327 Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E. & Jaenisch, R. *2001* Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat Genet* **27**(1), 31-39.

328 Klose, R. J. & Bird, A. P. *2006* Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* **31**(2), 89-97.
329 Fuks, F., Hurd, P. J., Wolf, D., Nan, X., Bird, A. P. & Kouzarides, T. *2003* The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol Chem* **278**(6), 4035-4040.
330 Csankovszki, G., Nagy, A. & Jaenisch, R. *2001* Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. *J Cell Biol* **153**(4), 773-784.
331 Mohandas, T., Sparkes, R. S. & Shapiro, L. J. *1981* Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science* **211**(4480), 393-396.
332 Li, E., Beard, C. & Jaenisch, R. *1993* Role for DNA methylation in genomic imprinting. *Nature* **366**(6453), 362-365.
333 Bannister, A. J. & Kouzarides, T. *2011* Regulation of chromatin by histone modifications. *Cell Res* **21**(3), 381-395.
334 Eberharter, A. & Becker, P. B. *2002* Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO Rep* **3**(3), 224-229.
335 Greer, E. L. & Shi, Y. *2012* Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet* **13**(5), 343-357.
336 Puigserver, P., Adelmant, G., Wu, Z., Fan, M., Xu, J., O'Malley, B. & Spiegelman, B. M. *1999* Activation of PPARgamma coactivator-1 through transcription factor docking. *Science* **286**(5443), 1368-1371.
337 Wallberg, A. E., Yamamura, S., Malik, S., Spiegelman, B. M. & Roeder, R. G. *2003* Coordination of p300-mediated chromatin remodeling and TRAP/mediator function through coactivator PGC-1alpha. *Molecular cell* **12**(5), 1137-1149.
338 Haberland, M., Arnold, M. A., McAnally, J., Phan, D., Kim, Y. & Olson, E. N. *2007* Regulation of HDAC9 gene expression by MEF2 establishes a negative-feedback loop in the transcriptional circuitry of muscle differentiation. *Mol Cell Biol* **27**(2), 518-525.
339 Cohen, T. J., Barrientos, T., Hartman, Z. C., Garvey, S. M., Cox, G. A. & Yao, T. P. *2009* The deacetylase HDAC4 controls myocyte enhancing factor-2-dependent structural gene expression in response to neural activity. *Faseb Journal* **23**(1), 99-106.
340 Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M. C., Hung, T., Argani, P., Rinn, J. L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R. B., van de Vijver, M. J., Sukumar, S. & Chang, H. Y. *2010* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**(7291), 1071-1076.
341 Wu, L., Murat, P., Matak-Vinkovic, D., Murrell, A. & Balasubramanian, S. *2013* Binding interactions between long noncoding RNA HOTAIR and PRC2 proteins. *Biochemistry* **52**(52), 9519-9527.
342 Fabian, M. R., Sonenberg, N. & Filipowicz, W. *2010* Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* **79**351-379.
343 Ye, Z., Jin, H. & Qian, Q. *2015* Argonaute 2: A Novel Rising Star in Cancer Research. *J Cancer* **6**(9), 877-882.
344 Jonas, S. & Izaurralde, E. *2015* Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* **16**(7), 421-433.
345 Liu, N., Williams, A. H., Maxeiner, J. M., Bezprozvannaya, S., Shelton, J. M., Richardson, J. A., Bassel-Duby, R. & Olson, E. N. *2012* microRNA-206 promotes skeletal muscle regeneration and delays progression of Duchenne muscular dystrophy in mice. *J Clin Invest* **122**(6), 2054-2065.
346 Carrio, E., Diez-Villanueva, A., Lois, S., Mallona, I., Cases, I., Forn, M., Peinado, M. A. & Suelves, M. *2015* Deconstruction of DNA methylation patterns during myogenesis reveals specific epigenetic events in the establishment of the skeletal muscle lineage. *Stem Cells* **33**(6), 2025-2036.
347 Tsumagari, K., Baribault, C., Terragni, J., Varley, K. E., Gertz, J., Pradhan, S., Badoo, M., Crain, C. M., Song, L., Crawford, G. E., Myers, R. M., Lacey, M. & Ehrlich, M. *2013* Early de

novo DNA methylation and prolonged demethylation in the muscle lineage. *Epigenetics : official journal of the DNA Methylation Society* **8**(3), 317-332.

348 Zhong, X., Wang, Q. Q., Li, J. W., Zhang, Y. M., An, X. R. & Hou, J. *2017* Ten-Eleven Translocation-2 (Tet2) Is Involved in Myogenic Differentiation of Skeletal Myoblast Cells in Vitro. *Scientific reports* **7**43539.

349 Miyata, K., Miyata, T., Nakabayashi, K., Okamura, K., Naito, M., Kawai, T., Takada, S., Kato, K., Miyamoto, S., Hata, K. & Asahara, H. *2015* DNA methylation analysis of human myoblasts during in vitro myogenic differentiation: de novo methylation of promoters of muscle-related genes and its involvement in transcriptional down-regulation. *Human molecular genetics* **24**(2), 410-423.

350 Jen, Y., Weintraub, H. & Benezra, R. *1992* Overexpression of Id protein inhibits the muscle differentiation program: in vivo association of Id with E2A proteins. *Genes & development* **6**(8), 1466-1479.

351 Yokoyama, S., Ito, Y., Ueno-Kudoh, H., Shimizu, H., Uchibe, K., Albini, S., Mitsuoka, K., Miyaki, S., Kiso, M., Nagai, A., Hikata, T., Osada, T., Fukuda, N., Yamashita, S., Harada, D., Mezzano, V., Kasai, M., Puri, P. L., Hayashizaki, Y., Okado, H., Hashimoto, M. & Asahara, H. *2009* A systems approach reveals that the myogenesis genome network is regulated by the transcriptional repressor RP58. *Developmental cell* **17**(6), 836-848.

352 Heyn, H., Li, N., Ferreira, H. J., Moran, S., Pisano, D. G., Gomez, A., Diez, J., Sanchez-Mut, J. V., Setien, F., Carmona, F. J., Puca, A. A., Sayols, S., Pujana, M. A., Serra-Musach, J., Iglesias-Platas, I., Formiga, F., Fernandez, A. F., Fraga, M. F., Heath, S. C., Valencia, A., Gut, I. G., Wang, J. & Esteller, M. *2012* Distinct DNA methylomes of newborns and centenarians. *Proceedings of the National Academy of Sciences of the United States of America* **109**(26), 10522-10527.

353 Horvath, S. *2013* DNA methylation age of human tissues and cell types. *Genome biology* **14**(10), R115.

354 Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., Bibikova, M., Fan, J. B., Gao, Y., Deconde, R., Chen, M., Rajapakse, I., Friend, S., Ideker, T. & Zhang, K. *2013* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular cell* **49**(2), 359-367.

355 Zykovich, A., Hubbard, A., Flynn, J. M., Tarnopolsky, M., Fraga, M. F., Kerksick, C., Ogborn, D., MacNeil, L., Mooney, S. D. & Melov, S. *2014* Genome-wide DNA methylation changes with age in disease-free human skeletal muscle. *Aging cell* **13**(2), 360-366.

356 Jin, L., Jiang, Z., Xia, Y., Lou, P., Chen, L., Wang, H., Bai, L., Xie, Y., Liu, Y., Li, W., Zhong, B., Shen, J., Jiang, A., Zhu, L., Wang, J., Li, X. & Li, M. *2014* Genome-wide DNA methylation changes in skeletal muscle between young and middle-aged pigs. *BMC Genomics* **15**653.

357 Liu, L., Cheung, T. H., Charville, G. W., Hurgo, B. M., Leavitt, T., Shih, J., Brunet, A. & Rando, T. A. *2013* Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. *Cell reports* **4**(1), 189-204.

358 Asp, P., Blum, R., Vethantham, V., Parisi, F., Micsinai, M., Cheng, J., Bowman, C., Kluger, Y. & Dynlacht, B. D. *2011* Genome-wide remodeling of the epigenetic landscape during myogenic differentiation. *Proceedings of the National Academy of Sciences of the United States of America* **108**(22), E149-158.

359 Cao, Y., Yao, Z., Sarkar, D., Lawrence, M., Sanchez, G. J., Parker, M. H., MacQuarrie, K. L., Davison, J., Morgan, M. T., Ruzzo, W. L., Gentleman, R. C. & Tapscott, S. J. *2010* Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming. *Developmental cell* **18**(4), 662-674.

360 Syddall, H. E., Aihie Sayer, A., Dennison, E. M., Martin, H. J., Barker, D. J. & Cooper, C. *2005* Cohort profile: the Hertfordshire cohort study. *International journal of epidemiology* **34**(6), 1234-1242.

361 Patel, H. P., Syddall, H. E., Martin, H. J., Stewart, C. E., Cooper, C. & Sayer, A. A. *2010* Hertfordshire sarcopenia study: design and methods. *BMC geriatrics* **10**43.

362 Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A.

K., Regev, A. & McCarroll, S. A. *2015* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**(5), 1202-1214.
363 Bushnell, B. *BBMap*, <sourceforge.net/projects/bbmap/> (
364 Trapnell, C., Pachter, L. & Salzberg, S. L. *2009* TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**(9), 1105-1111.
365 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. *2009* Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**(3), R25.
366 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & Genome Project Data Processing, S. *2009* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**(16), 2078-2079.
367 *Picard tools*, <http://picard.sourceforge.net> (
368 Anders, S., Pyl, P. T. & Huber, W. *2015* HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**(2), 166-169.
369 Love, M. I., Huber, W. & Anders, S. *2014* Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**(12), 550.
370 Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., Van Djik, S., Muhlhausler, B., Stirzaker, C. & Clark, S. J. *2016* Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome biology* **17**(1), 208.
371 Montojo, J., Zuberi, K., Rodriguez, H., Kazi, F., Wright, G., Donaldson, S. L., Morris, Q. & Bader, G. D. *2010* GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* **26**(22), 2927-2928.
372 Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. & Ideker, T. *2003* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**(11), 2498-2504.
373 Bader, G. D. & Hogue, C. W. *2003* An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* **4**2.
374 Maere, S., Heymans, K. & Kuiper, M. *2005* BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**(16), 3448-3449.
375 Ward, J. H. *1963* Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* **58**236-244.
376 Ward, P. A. S. & Taylor, D. J. *2001* A hierarchical cluster algorithm for dynamic, centralized timestamps. *Int Con Distr Comp S*585-593.
377 Langfelder, P., Zhang, B. & Horvath, S. *2008* Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* **24**(5), 719-720.
378 Holdt, L. M. & Teupser, D. *2012* Recent studies of the human chromosome 9p21 locus, which is associated with atherosclerosis in human populations. *Arterioscler Thromb Vasc Biol* **32**(2), 196-206.
379 Congrains, A., Kamide, K., Ohishi, M. & Rakugi, H. *2013* ANRIL: molecular mechanisms and implications in human health. *Int J Mol Sci* **14**(1), 1278-1292.
380 Broadbent, H. M., Peden, J. F., Lorkowski, S., Goel, A., Ongen, H., Green, F., Clarke, R., Collins, R., Franzosi, M. G., Tognoni, G., Seedorf, U., Rust, S., Eriksson, P., Hamsten, A., Farrall, M., Watkins, H. & consortium, P. *2008* Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum Mol Genet* **17**(6), 806-814.
381 Melzer, D., Frayling, T. M., Murray, A., Hurst, A. J., Harries, L. W., Song, H., Khaw, K., Luben, R., Surtees, P. G., Bandinelli, S. S., Corsi, A. M., Ferrucci, L., Guralnik, J. M., Wallace, R. B., Hattersley, A. T. & Pharoah, P. D. *2007* A common variant of the p16(INK4a) genetic region is associated with physical function in older people. *Mechanisms of ageing and development* **128**(5-6), 370-377.
382 Vandenberk, B., Brouwers, B., Hatse, S., Wildiers, H. *2011* P16INK4a: A central player in cellular senescence and a promising aging biomarker in elderly cancer patients. *J Geriatr Oncol* **2**(4), 259-269.

383 Yap, K. L., Li, S., Munoz-Cabello, A. M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M. J. & Zhou, M. M. *2010* Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Molecular cell* **38**(5), 662-674.

384 Kotake, Y., Nakagawa, T., Kitagawa, K., Suzuki, S., Liu, N., Kitagawa, M. & Xiong, Y. *2011* Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene* **30**(16), 1956-1962.

385 Bracken, A. P., Kleine-Kohlbrecher, D., Dietrich, N., Pasini, D., Gargiulo, G., Beekman, C., Theilgaard-Monch, K., Minucci, S., Porse, B. T., Marine, J. C., Hansen, K. H. & Helin, K. *2007* The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells. *Genes & development* **21**(5), 525-530.

386 Liu, Y., Sanoff, H. K., Cho, H., Burd, C. E., Torrice, C., Ibrahim, J. G., Thomas, N. E. & Sharpless, N. E. *2009* Expression of p16(INK4a) in peripheral blood T-cells is a biomarker of human aging. *Aging cell* **8**(4), 439-448.

387 Milanovic, Z., Pantelic, S., Trajkovic, N., Sporis, G., Kostic, R. & James, N. *2013* Age-related decrease in physical activity and functional fitness among elderly men and women. *Clin Interv Aging* **8**549-556.

388 Ershler, W. B. & Keller, E. T. *2000* Age-associated increased interleukin-6 gene expression, late-life diseases, and frailty. *Annu Rev Med* **51**245-270.

389 Cohen, H. J., Pieper, C. F., Harris, T., Rao, K. M. & Currie, M. S. *1997* The association of plasma IL-6 levels with functional disability in community-dwelling elderly. *J Gerontol A Biol Sci Med Sci* **52**(4), M201-208.

390 Ferrucci, L., Penninx, B. W., Volpato, S., Harris, T. B., Bandeen-Roche, K., Balfour, J., Leveille, S. G., Fried, L. P. & Md, J. M. *2002* Change in muscle strength explains accelerated decline of physical function in older women with high interleukin-6 serum levels. *J Am Geriatr Soc* **50**(12), 1947-1954.

391 Reuben, D. B., Cheh, A. I., Harris, T. B., Ferrucci, L., Rowe, J. W., Tracy, R. P. & Seeman, T. E. *2002* Peripheral blood markers of inflammation predict mortality and functional decline in high-functioning community-dwelling older persons. *J Am Geriatr Soc* **50**(4), 638-644.

392 Zhou, X., Han, X., Wittfeldt, A., Sun, J., Liu, C., Wang, X., Gan, L. M., Cao, H. & Liang, Z. *2016* Long non-coding RNA ANRIL regulates inflammatory responses as a novel component of NF-kappaB pathway. *RNA Biol* **13**(1), 98-108.

393 Livak, K. J. & Schmittgen, T. D. *2001* Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**(4), 402-408.

394 Lepper, C., Partridge, T. A. & Fan, C. M. *2011* An absolute requirement for Pax7-positive satellite cells in acute injury-induced skeletal muscle regeneration. *Development* **138**(17), 3639-3646.

395 Lillycrop, K., Murray, R., Cheong, C., Teh, A. L., Clarke-Harris, R., Barton, S., Costello, P., Garratt, E., Cook, E., Titcombe, P., Shunmuganathan, B., Liew, S. J., Chua, Y. C., Lin, X., Wu, Y., Burdge, G. C., Cooper, C., Inskip, H. M., Karnani, N., Hopkins, J. C., Childs, C. E., Chavez, C. P., Calder, P. C., Yap, F., Lee, Y. S., Chong, Y. S., Melton, P. E., Beilin, L., Huang, R. C., Gluckman, P. D., Harvey, N., Hanson, M. A., Holbrook, J. D., EpiGen, C. & Godfrey, K. M. *2017* ANRIL Promoter DNA Methylation: A Perinatal Marker for Later Adiposity. *EBioMedicine* **19**60-72.

396 Wan, J., Oliver, V. F., Wang, G., Zhu, H., Zack, D. J., Merbs, S. L. & Qian, J. *2015* Characterization of tissue-specific differential DNA methylation suggests distinct modes of positive and negative gene expression regulation. *BMC Genomics* **16**49.

397 Leenen, F. A., Muller, C. P. & Turner, J. D. *2016* DNA methylation: conducting the orchestra from exposure to phenotype? *Clin Epigenetics* **8**92.

398 Maccani, J. Z., Koestler, D. C., Lester, B., Houseman, E. A., Armstrong, D. A., Kelsey, K. T. & Marsit, C. J. *2015* Placental DNA Methylation Related to Both Infant Toenail Mercury and Adverse Neurobehavioral Outcomes. *Environmental health perspectives* **123**(7), 723-729.

399 Murphy, S. K., Adigun, A., Huang, Z., Overcash, F., Wang, F., Jirtle, R. L., Schildkraut, J. M., Murtha, A. P., Iversen, E. S. & Hoyo, C. *2012* Gender-specific methylation differences in relation to prenatal exposure to cigarette smoke. *Gene* **494**(1), 36-43.
400 Ebert, P. S., Hess, R. A., Frykholm, B. C. & Tschudy, D. P. *1979* Succinylacetone, a potent inhibitor of heme biosynthesis: effect on cell growth, heme content and delta-aminolevulinic acid dehydratase activity of malignant murine erythroleukemia cells. *Biochemical and biophysical research communications* **88**(4), 1382-1390.
401 Wang, Y. M., Huang, J. W., Castella, M., Huntsman, D. G. & Taniguchi, T. *2014* p53 Is Positively Regulated by miR-542-3p. *Cancer research* **74**(12), 3218-3227.
402 Fridman, J. S. & Lowe, S. W. *2003* Control of apoptosis by p53. *Oncogene* **22**(56), 9030-9040.
403 Garros, R. F., Paul, R., Connolly, M., Lewis, A., Garfield, B. E., Natanek, S. A., Bloch, S., Mouly, V., Griffiths, M. J., Polkey, M. I. & Kemp, P. R. *2017* MicroRNA-542 Promotes Mitochondrial Dysfunction and SMAD Activity and Is Elevated in Intensive Care Unit-acquired Weakness. *Am J Respir Crit Care Med* **196**(11), 1422-1433.
404 Pachnis, V., Belayew, A. & Tilghman, S. M. *1984* Locus unlinked to alpha-fetoprotein under the control of the murine raf and Rif genes. *Proceedings of the National Academy of Sciences of the United States of America* **81**(17), 5523-5527.
405 Martinet, C., Monnier, P., Louault, Y., Benard, M., Gabory, A. & Dandolo, L. *2016* H19 controls reactivation of the imprinted gene network during muscle regeneration. *Development* **143**(6), 962-971.
406 Davis, R. L., Weintraub, H. & Lassar, A. B. *1987* Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**(6), 987-1000.
407 Kallen, A. N., Zhou, X. B., Xu, J., Qiao, C., Ma, J., Yan, L., Lu, L., Liu, C., Yi, J. S., Zhang, H., Min, W., Bennett, A. M., Gregory, R. I., Ding, Y. & Huang, Y. *2013* The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Molecular cell* **52**(1), 101-112.
408 Lee, Y. S. & Dutta, A. *2007* The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes & development* **21**(9), 1025-1030.
409 Li, Z., Gilbert, J. A., Zhang, Y., Zhang, M., Qiu, Q., Ramanujan, K., Shavlakadze, T., Eash, J. K., Scaramozza, A., Goddeeris, M. M., Kirsch, D. G., Campbell, K. P., Brack, A. S. & Glass, D. J. *2012* An HMGA2-IGF2BP2 axis regulates myoblast proliferation and myogenesis. *Developmental cell* **23**(6), 1176-1188.
410 Cai, X. & Cullen, B. R. *2007* The imprinted H19 noncoding RNA is a primary microRNA precursor. *Rna* **13**(3), 313-316.
411 Dey, B. K., Pfeifer, K. & Dutta, A. *2014* The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes & development* **28**(5), 491-501.
412 Winbanks, C. E., Chen, J. L., Qian, H., Liu, Y., Bernardo, B. C., Beyer, C., Watt, K. I., Thomson, R. E., Connor, T., Turner, B. J., McMullen, J. R., Larsson, L., McGee, S. L., Harrison, C. A. & Gregorevic, P. *2013* The bone morphogenetic protein axis is a positive regulator of skeletal muscle mass. *The Journal of cell biology* **203**(2), 345-357.
413 Keniry, A., Oxley, D., Monnier, P., Kyba, M., Dandolo, L., Smits, G. & Reik, W. *2012* The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nat Cell Biol* **14**(7), 659-665.
414 Giovarelli, M., Bucci, G., Ramos, A., Bordo, D., Wilusz, C. J., Chen, C. Y., Puppo, M., Briata, P. & Gherzi, R. *2014* H19 long noncoding RNA controls the mRNA decay promoting function of KSRP. *Proceedings of the National Academy of Sciences of the United States of America* **111**(47), E5023-5028.
415 Cheng, A., Zhang, M., Okubo, M., Omichi, K. & Saltiel, A. R. *2009* Distinct mutations in the glycogen debranching enzyme found in glycogen storage disease type III lead to impairment in diverse cellular functions. *Human molecular genetics* **18**(11), 2045-2052.
416 Lucchiari, S., Santoro, D., Pagliarani, S. & Comi, G. P. *2007* Clinical, biochemical and genetic features of glycogen debranching enzyme deficiency. *Acta myologica : myopathies and cardiomyopathies : official journal of the Mediterranean Society of Myology / edited by the Gaetano Conte Academy for the study of striated muscle diseases* **26**(1), 72-74.

417 Shen, J. J. & Chen, Y. T. *2002* Molecular characterization of glycogen storage disease type III. *Current molecular medicine* **2**(2), 167-175.
418 Kishnani, P. S., Austin, S. L., Arn, P., Bali, D. S., Boney, A., Case, L. E., Chung, W. K., Desai, D. M., El-Gharbawy, A., Haller, R., Smit, G. P., Smith, A. D., Hobson-Webb, L. D., Wechsler, S. B., Weinstein, D. A., Watson, M. S. & Acmg. *2010* Glycogen storage disease type III diagnosis and management guidelines. *Genetics in medicine : official journal of the American College of Medical Genetics* **12**(7), 446-463.
419 Bal, A., Unlu, E., Bahar, G., Aydog, E., Eksioglu, E. & Yorgancioglu, R. *2007* Comparison of serum IL-1 beta, sIL-2R, IL-6, and TNF-alpha levels with disease activity parameters in ankylosing spondylitis. *Clinical rheumatology* **26**(2), 211-215.
420 Carotenuto, P., van Riel, D., Artsen, A., Bruijns, S., Uytdehaag, F. G., Laman, J. D., van Nunen, A. B., Zondervan, P. E., De Man, R. A., Osterhaus, A. D. & Pontesilli, O. *2005* Antiviral treatment with alpha interferon up-regulates CD14 on liver macrophages and its soluble form in patients with chronic hepatitis B. *Antimicrobial agents and chemotherapy* **49**(2), 590-599.
421 Guarda, G., Braun, M., Staehli, F., Tardivel, A., Mattmann, C., Forster, I., Farlik, M., Decker, T., Du Pasquier, R. A., Romero, P. & Tschopp, J. *2011* Type I interferon inhibits interleukin-1 production and inflammasome activation. *Immunity* **34**(2), 213-223.
422 Jacquemin, V., Furling, D., Bigot, A., Butler-Browne, G. S. & Mouly, V. *2004* IGF-1 induces human myotube hypertrophy by increasing cell recruitment. *Experimental cell research* **299**(1), 148-158.
423 Otto, A., Schmidt, C., Luke, G., Allen, S., Valasek, P., Muntoni, F., Lawrence-Watt, D. & Patel, K. *2008* Canonical Wnt signalling induces satellite-cell proliferation during adult skeletal muscle regeneration. *Journal of cell science* **121**(Pt 17), 2939-2950.
424 Zhang, L., Wang, X. H., Wang, H., Du, J. & Mitch, W. E. *2010* Satellite cell dysfunction and impaired IGF-1 signaling cause CKD-induced muscle atrophy. *J Am Soc Nephrol* **21**(3), 419-427.
425 Gabory, A., Jammes, H. & Dandolo, L. *2010* The H19 locus: role of an imprinted non-coding RNA in growth and development. *Bioessays* **32**(6), 473-480.
426 Xu, X., Ji, S., Li, W., Yi, B., Li, H., Zhang, H. & Ma, W. *2017* LncRNA H19 promotes the differentiation of bovine skeletal muscle satellite cells by suppressing Sirt1/FoxO1. *Cell Mol Biol Lett* **22**10.
427 Lewis, A., Lee, J. Y., Donaldson, A. V., Natanek, S. A., Vaidyanathan, S., Man, W. D., Hopkinson, N. S., Sayer, A. A., Patel, H. P., Cooper, C., Syddall, H., Polkey, M. I. & Kemp, P. R. *2016* Increased expression of H19/miR-675 is associated with a low fat-free mass index in patients with COPD. *J Cachexia Sarcopenia Muscle* **7**(3), 330-344.
428 Sartori, R. & Sandri, M. *2015* Bone and morphogenetic protein signalling and muscle mass. *Curr Opin Clin Nutr Metab Care* **18**(3), 215-220.
429 Sartori, R., Schirwis, E., Blaauw, B., Bortolanza, S., Zhao, J., Enzo, E., Stantzou, A., Mouisel, E., Toniolo, L., Ferry, A., Stricker, S., Goldberg, A. L., Dupont, S., Piccolo, S., Amthor, H. & Sandri, M. *2013* BMP signaling controls muscle mass. *Nature genetics* **45**(11), 1309-1318.
430 Ollikainen, M., Smith, K. R., Joo, E. J., Ng, H. K., Andronikos, R., Novakovic, B., Abdul Aziz, N. K., Carlin, J. B., Morley, R., Saffery, R. & Craig, J. M. *2010* DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intrauterine components to variation in the human neonatal epigenome. *Human molecular genetics* **19**(21), 4176-4188.
431 Huang, R. C., Galati, J. C., Burrows, S., Beilin, L. J., Li, X., Pennell, C. E., van Eekelen, J., Mori, T. A., Adams, L. A. & Craig, J. M. *2012* DNA methylation of the IGF2/H19 imprinting control region and adiposity distribution in young adults. *Clin Epigenetics* **4**(1), 21.
432 Pidsley, R., Fernandes, C., Viana, J., Paya-Cano, J. L., Liu, L., Smith, R. G., Schalkwyk, L. C. & Mill, J. *2012* DNA methylation at the Igf2/H19 imprinting control region is associated with cerebellum mass in outbred mice. *Mol Brain* **5**42.

433 Ghazal, S., McKinnon, B., Zhou, J., Mueller, M., Men, Y., Yang, L., Mueller, M., Flannery, C., Huang, Y. & Taylor, H. S. *2015* H19 lncRNA alters stromal cell growth via IGF signaling in the endometrium of women with endometriosis. *EMBO Mol Med* **7**(8), 996-1003.
434 Robinson, S. M., Simmonds, S. J., Jameson, K. A., Syddall, H. E., Dennison, E. M., Cooper, C., Sayer, A. A. & Hertfordshire Cohort Study, G. *2012* Muscle strength in older community-dwelling men is related to type of milk feeding in infancy. *J Gerontol A Biol Sci Med Sci* **67**(9), 990-996.
435 Patel, H. P., Jameson, K. A., Syddall, H. E., Martin, H. J., Stewart, C. E., Cooper, C. & Sayer, A. A. *2012* Developmental influences, muscle morphology, and sarcopenia in community-dwelling older men. *J Gerontol A Biol Sci Med Sci* **67**(1), 82-87.
436 Sayer, A. A., Dennison, E. M., Syddall, H. E., Jameson, K., Martin, H. J. & Cooper, C. *2008* The developmental origins of sarcopenia: using peripheral quantitative computed tomography to assess muscle size in older people. *J Gerontol A Biol Sci Med Sci* **63**(8), 835-840.
437 Wang, N., Chen, Y., Ning, Z., Li, Q., Han, B., Zhu, C., Chen, Y., Xia, F., Jiang, B., Wang, B., Wang, X., Jensen, M. D. & Lu, Y. *2016* Exposure to Famine in Early Life and Nonalcoholic Fatty Liver Disease in Adulthood. *The Journal of clinical endocrinology and metabolism* **101**(5), 2218-2225.
438 Sun, C., Burgner, D. P., Ponsonby, A. L., Saffery, R., Huang, R. C., Vuillermin, P. J., Cheung, M. & Craig, J. M. *2013* Effects of early-life environment and epigenetics on cardiovascular disease risk in children: highlighting the role of twin studies. *Pediatr Res* **73**(4 Pt 2), 523-530.
439 Murray, R., Bryant, J., Titcombe, P., Barton, S. J., Inskip, H., Harvey, N. C., Cooper, C., Lillycrop, K., Hanson, M. & Godfrey, K. M. *2016* DNA methylation at birth within the promoter of ANRIL predicts markers of cardiovascular risk at 9 years. *Clin Epigenetics* **8**90.
440 Godfrey, K. M., Lillycrop, K. A., Burdge, G. C., Gluckman, P. D. & Hanson, M. A. *2007* Epigenetic mechanisms and the mismatch concept of the developmental origins of health and disease. *Pediatr Res* **61**(5 Pt 2), 5R-10R.
441 Romanick, M., Thompson, L. V. & Brown-Borg, H. M. *2013* Murine models of atrophy, cachexia, and sarcopenia in skeletal muscle. *Biochim Biophys Acta* **1832**(9), 1410-1420.
442 Sharples, A. P., Al-Shanti, N. & Stewart, C. E. *2010* C2 and C2C12 murine skeletal myoblast models of atrophic and hypertrophic potential: relevance to disease and ageing? *Journal of cellular physiology* **225**(1), 240-250.
443 Burattini, S., Ferri, P., Battistelli, M., Curci, R., Luchetti, F. & Falcieri, E. *2004* C2C12 murine myoblasts as a model of skeletal muscle development: morpho-functional characterization. *Eur J Histochem* **48**(3), 223-233.
444 Alexakis, C., Partridge, T. & Bou-Gharios, G. *2007* Implication of the satellite cell in dystrophic muscle fibrosis: a self-perpetuating mechanism of collagen overproduction. *Am J Physiol Cell Physiol* **293**(2), C661-669.
445 Moorwood, C., Lozynska, O., Suri, N., Napper, A. D., Diamond, S. L. & Khurana, T. S. *2011* Drug discovery for Duchenne muscular dystrophy via utrophin promoter activation screening. *PloS one* **6**(10), e26169.
446 Li, B., Lin, M., Tang, Y., Wang, B. & Wang, J. H. *2008* A novel functional assessment of the differentiation of micropatterned muscle cells. *J Biomech* **41**(16), 3349-3353.
447 Elsea, S. H. & Lucas, R. E. *2002* The mousetrap: what we can learn when the mouse model does not mimic the human disease. *ILAR J* **43**(2), 66-79.
448 Chandrasekharan, K., Yoon, J. H., Xu, Y., deVries, S., Camboni, M., Janssen, P. M., Varki, A. & Martin, P. T. *2010* A human-specific deletion in mouse Cmah increases disease severity in the mdx model of Duchenne muscular dystrophy. *Sci Transl Med* **2**(42), 42ra54.
449 Demetrius, L. *2006* Aging in mouse and human systems: a comparative study. *Annals of the New York Academy of Sciences* **1067**66-82.
450 Mamchaoui, K., Trollet, C., Bigot, A., Negroni, E., Chaouch, S., Wolff, A., Kandalla, P. K., Marie, S., Di Santo, J., St Guily, J. L., Muntoni, F., Kim, J., Philippi, S., Spuler, S., Levy, N., Blumen, S. C., Voit, T., Wright, W. E., Aamiri, A., Butler-Browne, G. & Mouly, V. *2011*

Immortalized pathological human myoblasts: towards a universal tool for the study of neuromuscular disorders. *Skeletal muscle* **1**34.

451 Nesmith, A. P., Wagner, M. A., Pasqualini, F. S., O'Connor, B. B., Pincus, M. J., August, P. R. & Parker, K. K. *2016* A human in vitro model of Duchenne muscular dystrophy muscle formation and contractility. *The Journal of cell biology* **215**(1), 47-56.

452 Chen, Y. A., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., Gallinger, S., Hudson, T. J. & Weksberg, R. *2013* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics : official journal of the DNA Methylation Society* **8**(2), 203-209.

453 Cesari, M., Fielding, R., Benichou, O., Bernabei, R., Bhasin, S., Guralnik, J. M., Jette, A., Landi, F., Pahor, M., Rodriguez-Manas, L., Rolland, Y., Roubenoff, R., Sinclair, A. J., Studenski, S., Travison, T. & Vellas, B. *2015* Pharmacological Interventions in Frailty and Sarcopenia: Report by the International Conference on Frailty and Sarcopenia Research Task Force. *J Frailty Aging* **4**(3), 114-120.

454 Carter, C. S., Marzetti, E., Leeuwenburgh, C., Manini, T., Foster, T. C., Groban, L., Scarpace, P. J. & Morgan, D. *2012* Usefulness of preclinical models for assessing the efficacy of late-life interventions for sarcopenia. *J Gerontol A Biol Sci Med Sci* **67**(1), 17-27.

455 Bortoli, S., Renault, V., Mariage-Samson, R., Eveno, E., Auffray, C., Butler-Browne, G. & Pietu, G. *2005* Modifications in the myogenic program induced by in vivo and in vitro aging. *Gene* **347**(1), 65-72.

456 Bigot, A., Duddy, W. J., Ouandaogo, Z. G., Negroni, E., Mariot, V., Ghimbovschi, S., Harmon, B., Wielgosik, A., Loiseau, C., Devaney, J., Dumonceaux, J., Butler-Browne, G., Mouly, V. & Duguez, S. *2015* 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* **13**(6), 1172-1182.

457 Brack, A. S., Bildsoe, H. & Hughes, S. M. *2005* Evidence that satellite cell decrement contributes to preferential decline in nuclear number from large fibres during murine age-related muscle atrophy. *Journal of cell science* **118**(Pt 20), 4813-4821.

458 Harthan, L. B., McFarland, D. C. & Velleman, S. G. *2013* The effect of syndecan-4 and glypican-1 expression on age-related changes in myogenic satellite cell proliferation, differentiation, and fibroblast growth factor 2 responsiveness. *Comp Biochem Phys A* **166**(4), 590-602.

459 Barberi, L., Scicchitano, B. M., De Rossi, M., Bigot, A., Duguez, S., Wielgosik, A., Stewart, C., McPhee, J., Conte, M., Narici, M., Franceschi, C., Mouly, V., Butler-Browne, G. & Musaro, A. *2013* Age-dependent alteration in muscle regeneration: the critical role of tissue niche. *Biogerontology* **14**(3), 273-292.

460 Watt, F. & Molloy, P. L. *1988* Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes & development* **2**(9), 1136-1143.

461 Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N. & Bird, A. *1998* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**(6683), 386-389.

462 Bjornsson, H. T., Sigurdsson, M. I., Fallin, M. D., Irizarry, R. A., Aspelund, T., Cui, H., Yu, W., Rongione, M. A., Ekstrom, T. J., Harris, T. B., Launer, L. J., Eiriksdottir, G., Leppert, M. F., Sapienza, C., Gudnason, V. & Feinberg, A. P. *2008* Intra-individual change over time in DNA methylation with familial clustering. *JAMA* **299**(24), 2877-2883.

463 Barres, R., Yan, J., Egan, B., Treebak, J. T., Rasmussen, M., Fritz, T., Caidahl, K., Krook, A., O'Gorman, D. J. & Zierath, J. R. *2012* Acute exercise remodels promoter methylation in human skeletal muscle. *Cell metabolism* **15**(3), 405-411.

464 van Iterson, M., van Zwet, E. W., Consortium, B. & Heijmans, B. T. *2017* Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome biology* **18**(1), 19.

465 Stouffer, S. A. *The American soldier*. (Princeton University Press, 1949).

466 Marg, A., Schoewel, V., Timmel, T., Schulze, A., Shah, C., Daumke, O. & Spuler, S. *2012* Sarcolemmal Repair Is a Slow Process and Includes EHD2. *Traffic* **13**(9), 1286-1294.
467 Akiyoshi, J., Nishizono, A., Yamada, K., Nagayama, H., Mifune, K. & Fujii, I. *1995* Rapid desensitization of serotonin 5-HT2C receptor-stimulated intracellular calcium mobilization in CHO cells transfected with cloned human 5-HT2C receptors. *J Neurochem* **64**(6), 2473-2479.
468 Cussac, D., Boutet-Robinet, E., Ailhaud, M. C., Newman-Tancredi, A., Martel, J. C., Danty, N. & Rauly-Lestienne, I. *2008* Agonist-directed trafficking of signalling at serotonin 5-HT2A, 5-HT2B and 5-HT2C-VSV receptors mediated Gq/11 activation and calcium mobilisation in CHO cells. *Eur J Pharmacol* **594**(1-3), 32-38.
469 Fox, M. A., Ho, M. S. P., Smyth, N. & Sanes, J. R. *2008* A synaptic nidogen: Developmental regulation and role of nidogen-2 at the neuromuscular junction. *Neural Dev* **3**.
470 Festuccia, N., Owens, N. & Navarro, P. *2017* Esrrb, an estrogen-related receptor involved in early development, pluripotency, and reprogramming. *FEBS Lett*.
471 Latos, P. A., Goncalves, A., Oxley, D., Mohammed, H., Turro, E. & Hemberger, M. *2015* Fgf and Esrrb integrate epigenetic and transcriptional networks that regulate self-renewal of trophoblast stem cells. *Nat Commun* **6**7776.
472 Clohisey, S. M. R., Dzhindzhev, N. S. & Ohkura, H. *2014* Kank Is an EB1 Interacting Protein that Localises to Muscle-Tendon Attachment Sites in Drosophila. *PloS one* **9**(9).
473 Gan, X., Wang, J., Wang, C., Sommer, E., Kozasa, T., Srinivasula, S., Alessi, D., Offermanns, S., Simon, M. I. & Wu, D. *2012* PRR5L degradation promotes mTORC2-mediated PKC-delta phosphorylation and cell migration downstream of Galpha12. *Nat Cell Biol* **14**(7), 686-696.
474 Kim, H. R. & Ingham, P. W. *2009* The extracellular matrix protein TGFBI promotes myofibril bundling and muscle fibre growth in the zebrafish embryo. *Dev Dyn* **238**(1), 56-65.
475 Ishii, A. & Lo, S. H. *2001* A role of tensin in skeletal-muscle regeneration. *Biochemical Journal* **356**737-745.
476 Carrio, E. & Suelves, M. *2015* DNA methylation dynamics in muscle development and disease. *Frontiers in aging neuroscience* **7**19.
477 Kanzleiter, T., Jahnert, M., Schulze, G., Selbig, J., Hallahan, N., Schwenk, R. W. & Schurmann, A. *2015* Exercise training alters DNA methylation patterns in genes related to muscle growth and differentiation in mice. *American journal of physiology. Endocrinology and metabolism* **308**(10), E912-920.
478 Suzuki, T., Do, M. K., Sato, Y., Ojima, K., Hara, M., Mizunoya, W., Nakamura, M., Furuse, M., Ikeuchi, Y., Anderson, J. E. & Tatsumi, R. *2013* Comparative analysis of semaphorin 3A in soleus and EDL muscle satellite cells in vitro toward understanding its role in modulating myogenin expression. *The international journal of biochemistry & cell biology* **45**(2), 476-482.
479 Awede, B., Thissen, J., Gailly, P. & Lebacq, J. *1999* Regulation of IGF-I, IGFBP-4 and IGFBP-5 gene expression by loading in mouse skeletal muscle. *FEBS Lett* **461**(3), 263-267.
480 Calderon, J. C., Bolanos, P. & Caputo, C. *2014* The excitation-contraction coupling mechanism in skeletal muscle. *Biophys Rev* **6**(1), 133-160.
481 Ohm, J. E., McGarvey, K. M., Yu, X., Cheng, L., Schuebel, K. E., Cope, L., Mohammad, H. P., Chen, W., Daniel, V. C., Yu, W., Berman, D. M., Jenuwein, T., Pruitt, K., Sharkis, S. J., Watkins, D. N., Herman, J. G. & Baylin, S. B. *2007* A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nature genetics* **39**(2), 237-242.
482 Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J. M., Bollen, M., Esteller, M., Di Croce, L., de Launoit, Y. & Fuks, F. *2006* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**(7078), 871-874.
483 Juan, A. H., Derfoul, A., Feng, X., Ryall, J. G., Dell'Orso, S., Pasut, A., Zare, H., Simone, J. M., Rudnicki, M. A. & Sartorelli, V. *2011* Polycomb EZH2 controls self-renewal and safeguards the transcriptional identity of skeletal muscle stem cells. *Genes & development* **25**(8), 789-794.

484 Caretti, G., Di Padova, M., Micales, B., Lyons, G. E. & Sartorelli, V. *2004* The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes & development* **18**(21), 2627-2638.
485 Rocheteau, P., Gayraud-Morel, B., Siegl-Cachedenier, I., Blasco, M. A. & Tajbakhsh, S. *2012* A subpopulation of adult skeletal muscle stem cells retains all template DNA strands after cell division. *Cell* **148**(1-2), 112-125.
486 Kuang, S., Kuroda, K., Le Grand, F. & Rudnicki, M. A. *2007* Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell* **129**(5), 999-1010.
487 Meeson, A. P., Hawke, T. J., Graham, S., Jiang, N., Elterman, J., Hutcheson, K., Dimaio, J. M., Gallardo, T. D. & Garry, D. J. *2004* Cellular and molecular regulation of skeletal muscle side population cells. *Stem Cells* **22**(7), 1305-1320.
488 Beauchamp, J. R., Heslop, L., Yu, D. S., Tajbakhsh, S., Kelly, R. G., Wernig, A., Buckingham, M. E., Partridge, T. A. & Zammit, P. S. *2000* Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. *The Journal of cell biology* **151**(6), 1221-1234.
489 Cho, D. S. & Doles, J. D. *2017* Single cell transcriptome analysis of muscle satellite cells reveals widespread transcriptional heterogeneity. *Gene* **636**54-63.
490 Chapman, M. R., Balakrishnan, K. R., Li, J., Conboy, M. J., Huang, H., Mohanty, S. K., Jabart, E., Hack, J., Conboy, I. M. & Sohn, L. L. *2013* Sorting single satellite cells from individual myofibers reveals heterogeneity in cell-surface markers and myogenic capacity. *Integr Biol (Camb)* **5**(4), 692-702.
491 Pavlath, G. K., Thaloor, D., Rando, T. A., Cheong, M., English, A. W. & Zheng, B. *1998* Heterogeneity among muscle precursor cells in adult skeletal muscles with differing regenerative capacities. *Dev Dyn* **212**(4), 495-508.
492 Chakkalakal, J. V., Christensen, J., Xiang, W., Tierney, M. T., Boscolo, F. S., Sacco, A. & Brack, A. S. *2014* Early forming label-retaining muscle stem cells require p27kip1 for maintenance of the primitive state. *Development* **141**(8), 1649-1659.
493 Andersen, M. S. & Jensen, K. B. *2016* Stem cell heterogeneity revealed. *Nat Cell Biol* **18**(6), 587-589.
494 Sada, A., Jacob, F., Leung, E., Wang, S., White, B. S., Shalloway, D. & Tumbar, T. *2016* Defining the cellular lineage hierarchy in the interfollicular epidermis of adult skin. *Nat Cell Biol* **18**(6), 619-631.
495 Tellez-Gabriel, M., Ory, B., Lamoureux, F., Heymann, M. F. & Heymann, D. *2016* Tumour Heterogeneity: The Key Advantages of Single-Cell Analysis. *International journal of molecular sciences* **17**(12).
496 Schakman, O., Kalista, S., Barbe, C., Loumaye, A. & Thissen, J. P. *2013* Glucocorticoid-induced skeletal muscle atrophy. *The international journal of biochemistry & cell biology* **45**(10), 2163-2172.
497 Lofberg, E., Gutierrez, A., Wernerman, J., Anderstam, B., Mitch, W. E., Price, S. R., Bergstrom, J. & Alvestrand, A. *2002* Effects of high doses of glucocorticoids on free amino acids, ribosomes and protein turnover in human muscle. *Eur J Clin Invest* **32**(5), 345-353.
498 Waddell, D. S., Baehr, L. M., van den Brandt, J., Johnsen, S. A., Reichardt, H. M., Furlow, J. D. & Bodine, S. C. *2008* The glucocorticoid receptor and FOXO1 synergistically activate the skeletal muscle atrophy-associated MuRF1 gene. *American journal of physiology. Endocrinology and metabolism* **295**(4), E785-797.
499 Duclos, M., Martin, C., Malgat, M., Mazat, J. P., Chaouloff, F., Mormede, P. & Letellier, T. *2001* Relationships between muscle mitochondrial metabolism and stress-induced corticosterone variations in rats. *Pflugers Arch* **443**(2), 218-226.
500 Janssen, I., Shepard, D. S., Katzmarzyk, P. T. & Roubenoff, R. *2004* The healthcare costs of sarcopenia in the United States. *Journal of the American Geriatrics Society* **52**(1), 80-85.
501 Gil, J. & Peters, G. *2006* Regulation of the INK4b-ARF-INK4a tumour suppressor locus: all for one or one for all. *Nat Rev Mol Cell Biol* **7**(9), 667-677.

502 White, J. P., Puppa, M. J., Sato, S., Gao, S., Price, R. L., Baynes, J. W., Kostek, M. C., Matesic, L. E. & Carson, J. A. *2012* IL-6 regulation on skeletal muscle mitochondrial remodeling during cancer cachexia in the ApcMin/+ mouse. *Skeletal muscle* **2**14.
503 Ji, C., Chen, X., Gao, C., Jiao, L., Wang, J., Xu, G., Fu, H., Guo, X. & Zhao, Y. *2011* IL-6 induces lipolysis and mitochondrial dysfunction, but does not affect insulin-mediated glucose transport in 3T3-L1 adipocytes. *J Bioenerg Biomembr* **43**(4), 367-375.
504 Maples, J. M., Brault, J. J., Witczak, C. A., Park, S., Hubal, M. J., Weber, T. M., Houmard, J. A. & Shewchuk, B. M. *2015* Differential epigenetic and transcriptional response of the skeletal muscle carnitine palmitoyltransferase 1B (CPT1B) gene to lipid exposure with obesity. *American journal of physiology. Endocrinology and metabolism* **309**(4), E345-356.
505 Wanders, R. J., Ruiter, J. P., L, I. J., Waterham, H. R. & Houten, S. M. *2010* The enzymology of mitochondrial fatty acid beta-oxidation and its application to follow-up analysis of positive neonatal screening results. *J Inherit Metab Dis* **33**(5), 479-494.
506 Maunakea, A. K., Nagarajan, R. P., Bilenky, M., Ballinger, T. J., D'Souza, C., Fouse, S. D., Johnson, B. E., Hong, C., Nielsen, C., Zhao, Y., Turecki, G., Delaney, A., Varhol, R., Thiessen, N., Shchors, K., Heine, V. M., Rowitch, D. H., Xing, X., Fiore, C., Schillebeeckx, M., Jones, S. J., Haussler, D., Marra, M. A., Hirst, M., Wang, T. & Costello, J. F. *2010* Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**(7303), 253-257.
507 Mahapatra, S., Klee, E. W., Young, C. Y., Sun, Z., Jimenez, R. E., Klee, G. G., Tindall, D. J. & Donkena, K. V. *2012* Global methylation profiling for risk prediction of prostate cancer. *Clin Cancer Res* **18**(10), 2882-2895.
508 Wu, H. C., Yang, H. I., Wang, Q., Chen, C. J. & Santella, R. M. *2017* Plasma DNA methylation marker and hepatocellular carcinoma risk prediction model for the general population. *Carcinogenesis* **38**(10), 1021-1028.
509 Gillberg, L. & Ling, C. *2015* The potential use of DNA methylation biomarkers to identify risk and progression of type 2 diabetes. *Front Endocrinol (Lausanne)* **6**43.
510 Mendelson, M. M., Marioni, R. E., Joehanes, R., Liu, C., Hedman, A. K., Aslibekyan, S., Demerath, E. W., Guan, W., Zhi, D., Yao, C., Huan, T., Willinger, C., Chen, B., Courchesne, P., Multhaup, M., Irvin, M. R., Cohain, A., Schadt, E. E., Grove, M. L., Bressler, J., North, K., Sundstrom, J., Gustafsson, S., Shah, S., McRae, A. F., Harris, S. E., Gibson, J., Redmond, P., Corley, J., Murphy, L., Starr, J. M., Kleinbrink, E., Lipovich, L., Visscher, P. M., Wray, N. R., Krauss, R. M., Fallin, D., Feinberg, A., Absher, D. M., Fornage, M., Pankow, J. S., Lind, L., Fox, C., Ingelsson, E., Arnett, D. K., Boerwinkle, E., Liang, L., Levy, D. & Deary, I. J. *2017* Association of Body Mass Index with DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach. *PLoS Med* **14**(1), e1002215.
511 Huang, R. C., Garratt, E. S., Pan, H., Wu, Y., Davis, E. A., Barton, S. J., Burdge, G. C., Godfrey, K. M., Holbrook, J. D. & Lillycrop, K. A. *2015* Genome-wide methylation analysis identifies differentially methylated CpG loci associated with severe obesity in childhood. *Epigenetics : official journal of the DNA Methylation Society* **10**(11), 995-1005.
512 Poole, R. L., Leith, D. J., Docherty, L. E., Shmela, M. E., Gicquel, C., Splitt, M., Temple, I. K. & Mackay, D. J. *2012* Beckwith-Wiedemann syndrome caused by maternally inherited mutation of an OCT-binding motif in the IGF2/H19-imprinting control region, ICR1. *Eur J Hum Genet* **20**(2), 240-243.
513 Bliek, J., Terhal, P., van den Bogaard, M. J., Maas, S., Hamel, B., Salieb-Beugelaar, G., Simon, M., Letteboer, T., van der Smagt, J., Kroes, H. & Mannens, M. *2006* Hypomethylation of the H19 gene causes not only Silver-Russell syndrome (SRS) but also isolated asymmetry or an SRS-like phenotype. *Am J Hum Genet* **78**(4), 604-614.