Sex-Specific Epigenetic Adaptations to Exercise Training

by

Shanie Landen

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Principal supervisor: Prof Nir Eynon

Co-supervisors: Dr Sarah Voisin and A/Prof Séverine Lamon

Institute for Health and Sports (IHeS)

Victoria University

August 20, 2021

Abstract

Physical activity is the most effective intervention to enhance health and prevent chronic diseases, such as obesity, type 2 diabetes, and cardiovascular disease. Studies and consortiums have aimed to understand the underlying molecular mechanisms that bring about a healthier phenotype with exercise training, and growing evidence suggests that epigenetic changes, which are molecular modifications to the DNA, play a large role in regulating exercise training adaptations.

DNA methylation is the most widely studied epigenetic modification in exercise training studies, as it has been shown that both acute and maintained exercise (i.e. training) induce changes in the DNA methylome and subsequent gene function in human skeletal muscle. However, to date, studies identifying skeletal muscle epigenetic adaptations to exercise training have not investigated whether there is a sex-specific effect, despite skeletal muscle being one of the tissues with the most sex-biased gene expression.

A majority of the animal and human studies that have guided our understanding of the underlying molecular adaptations to exercise training have either included only males or pooled males and females together without considering potential sex differences. However, biological sex has been identified as a confounding variable across many biological disciplines, and sex-specific analysis can be critical to the interpretation, validation, reproducibility and generalizability of research findings [1]. **Thus, the overarching aim of this thesis was to investigate the sex-specific epigenome-wide response to exercise training.** Sixty-five healthy males and females (females n = 20; males n = 45) from the Gene SMART (Skeletal Muscle Adaptive Response to Training) study completed four weeks of high-intensity interval training (HIIT) to assess sex-specific training-induced DNA methylation changes. This thesis involved adding the female cohort to the already existing male cohort, of which most of the

participant data was collected prior to the commencement of this thesis. Participants underwent a four-week control period prior to commencing the training intervention. To determine whether training induced similar changes in physiological fitness in males and females, three measurements were assessed – maximum oxygen consumption (VO_2max), peak power output (PP), and lactate threshold (LT) – at three time points (before control, before and after the HIIT intervention). To assess sex differences in DNA methylation and other molecular measurements (fibre type proportions and gene expression), skeletal muscle biopsies were collected at each time point and analysed with the Illumina HumanMethylation EPIC array.

In Chapter 3, we have shown that there are 56,813 differentially methylated positions (DMPs) in the autosomes of male and female skeletal muscle at rest (false discovery rate [FDR] < 0.005), using a large scale meta-analysis of three independent cohorts (Gene SMART, FUSION, and GSE38291) comprising 369 individuals. These DMPs were mostly hypomethylated in males (94%), and were annotated to 10,240 differentially methylated regions (DMRs) and 8,420 differentially methylated genes (DMGs). Gene set enrichment analysis (GSEA) revealed enrichment of sex-differential methylation among muscle contraction, anatomical structure, and metabolism related pathways. Overlapping DMGs with genes known to have sex-biased skeletal muscle expression (differentially expressed genes [DEGs] from GTex), revealed a significant enrichment of DEGs among DMGs. We confirmed over-representation of DEGs among DMGs with transcriptomic data in an additional cohort (FUSION) which was also included in the DNA methylation meta-analysis. Lastly, using qPCR, we verified gene expression sex differences of three top genes identified from the differential methylation and expression analysis in an additional cohort included in the DNA methylation meta-analysis (Gene SMART).

In Chapter 4, we investigated the underlying biological factors contributing to the observed sex differences in basal skeletal muscle DNA methylation. Using a meta-analysis

approach in the Gene SMART and FUSION cohorts, we have shown that type I muscle fibre proportions were associated with DNA methylation at 16% of sex-biased DNA methylation loci. We found that circulating sex hormone levels (estrogen, testosterone, free testosterone, and sex hormone-binding globulin) in the Gene SMART cohort were not associated with differential methylation at the sex-biased DNA methylation loci. Lastly, we identified that the meta-analysis sex-DMPs were enriched for transcription factor binding sites (TFBSs) of 41 transcription factors (TF) , as previously established by uniform processing of multiple ChIP-seq data sets, including sex hormone-related androgen (*AR*), estrogen (*ESR1*), and glucocorticoid (*NR3C1*) receptors.

In Chapter 5, after elucidating the basal skeletal muscle DNA methylome sex differences and their biological contributors, we investigated whether there are sex differences in exercise training-induced DNA methylation changes. First, we found that both males and females improved the physiological fitness measurements PP and LT, but not VO_2max, in response to the HIIT, with no sex differences in the degree of the responses. We identified 1,261 CpGs whose methylation changed after four weeks of HIIT at a stringent FDR threshold < 0.005. We found no sex-specific DNA methylation changes after four weeks of HIIT (sex-by-training interaction) at a stringent FDR threshold < 0.005. A global examination of all the statistical tests performed genome-wide did not reveal an inflation of near zero p-values, suggesting that males and females do not differ in their epigenetic response to four weeks of HIIT. Given the relatively short training intervention, we then aimed to investigate whether there were sex differences in DNA methylation associated with cardiorespiratory fitness (CRF), an indicator of lifelong physical activity levels. We found 27,987 DMPs associated with CRF (FDR < 0.005), and no sex differences in the association between CRF and DNA methylation.

The experimental design and meta-analysis of this thesis provided large-scale epigenome-wide insight on skeletal muscle epigenetic sex differences, and elucidated the role of DNA methylation in exercise training adaptations in both males and females (**Chapter 5**). It yielded a comprehensive understanding of the profound sex-specific skeletal muscle DNA methylation and transcriptomic profiles (**Chapter 3**) and the underlying biological factors (**Chapter 4**) that distinguish male and female skeletal muscle DNA methylomes. Specifically, muscle fibre type proportions were associated with sites displaying sex differences in DNA methylation; nonetheless, the vast majority of loci that exhibit sex-biased DNA methylation differ regardless of sex differences in fibre type proportions. In addition, although circulating hormones were not associated with sex-differential DNA methylation, the enrichment of hormone-responsive TFBSs suggests that hormones underlie a portion of the DNA methylation sex differences in skeletal muscle. However, the influence of other biological factors, such as the sex chromosomes, on the sex differences observed in the autosomal DNA methylome remains to be determined. Lastly, despite the plethora of sex differences in the skeletal muscle DNA methylome at rest, the DNA methylomes of males and females responded similarly to exercise training as well as lifelong physical activity. These novel findings shed light on the epigenetic response of skeletal muscle to exercise training in healthy males and females. Integrating the DNA methylome with downstream -omics, such as transcriptomics, proteomics, and metabolomics, will further elucidate the pathways and networks involved in the skeletal muscle response to exercise training as well as any sex-specific adaptations. Future studies should include males and females in exercise training studies, take sex and other sex-related factors into consideration in study design and analysis, as well as integrate other OMIC layers to better characterise the skeletal muscle response to exercise training in humans.

Student Declaration

"I, Shanie Landen, declare that the PhD thesis entitled "Sex-specific epigenetic adaptations to exercise training" is no more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work."

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

Signature: Date: 08/03/2021

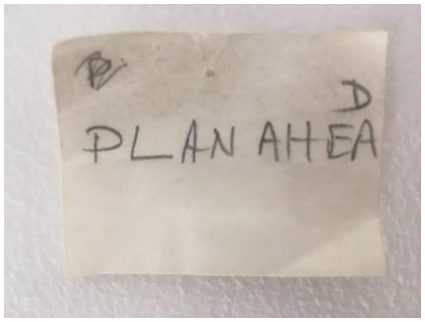
Acknowledgements

Firstly I would like to thank my primary supervisor, Nir, for giving me this opportunity to research in this field which I am so passionate about. He was enormously supportive, looked out for and believed in me, supported me to attend international conferences, and helped me to improve my critical thinking and presentation skills, among many others. He was always pushing me to get work done, which I do really appreciate, despite it causing me to avoid any one-on-one contact towards the end of my PhD. He put together a very supportive, hard-working, yet fun group environment which I do not take for granted. I would like to thank my co-supervisor Sarah for teaching me to code in R (not only basic coding but genome-wide data!!), for putting up with (with the utmost patience) my 11pm threads of messages about my analysis, for guiding me along every phase of my PhD, and providing continuous constructive feedback, making me a better researcher. The two of you are a complimentary team which I am so fortunate to have ended up working with. I would like to thank my third co-supervisor, Sev, who was pivotal in helping me choose my PhD topic, always provided lightning-quick feedback, and helped tremendously with my two review papers included in the literature review. The greatest gift I take with me is this passion and drive for excellence in research that you three have exemplified.

I would like to thank everyone in our Genetics and epigenetics of exercise group. Without them I would have been in the lab from 6am to 9pm almost daily for 1.5 years. Thanks guys for the memorable sunrises and sunsets experienced together in the ex phys lab. Macsue, thank you for exemplifying how to work efficiently. I am grateful that we got to complete our PhD's side by side and help each other tackle all the struggles along the way. Thank you Javier Alvarez for your countless hours helping with GXTs and trainings. Thank you Danielle Hiam for teaching me Aussie slang, always providing a helping hand, and splitting the sex-specific review. Thanks to the interns and rest of the people from our group who helped in data collection. Thank you Andrew Palmer and Kirsten who joined the group recently and are taking this research topic to the next level, your projects are going to bring about exciting results. Thank you to Andrew Garnham who biked across

the city in the cold dark Melbourne winter mornings to perform all of my (as well as every other lab in Melbourne's) biopsies; every biopsy day included an insightful chat with him. Thanks to the lab technicians for your help in organizing the lab space and troubleshooting every issue. Thank you to Alba Moreno for guiding me through the qPCRs as well as many meaningful science discussions. Thank you to the rest of the PhD students and post docs at VU's institute for health and sport who provided a stimulating and fun environment which promotes the development of our great international minds. Thank you to the administrative staff at VU who help orchestrate everything in the background, and to VU for my PhD scholarship over these years. Thank you to Chris Shaw and his lab members for showing me how to perform the immunohistochemistry for fibre typing. Thank you to the participants who worked hard and donated some of their muscle for me to explore my curiosities.

I would like to thank my parents, Tsipi and Udi, who always support my ambitions, and supported my decision to move to the other side of the world despite their desire to keep me close. They were unconditionally prepared to listen and advise me throughout the roller coaster that a PhD can be, but of course only when I felt like talking. I wish COVID had not stopped us from seeing each other and visiting in the last year and a half of my PhD, but that makes the moments we were able to see each other over the last four years all the more special. Thank you to my brother Ben who helped me with all of my grammar and vocabulary questions along the way, as he always does, even when juggling raising two babies who I hope to spend more time with soon. Thank you to my grandparents, Rina and Yosi, who constantly remind me to work hard, be organized, and:



(Yosi Landen, 2012)

Thank you to my husband's mom, Mzia, who congratulated and appreciated the hard work required to achieve every milestone. Thanks Adi and Zack for visiting me in Australia on this journey, and for nurturing my love of science since day one in undergrad classes at UC Davis. Thanks Nitzan and my cousins for instilling confidence in me in moments of weakness. Thanks to the rest of my friends and family who are my support squad and sent me flowers from across the world. Thanks to my running team who copped up listening to my PhD hardships and helped run the stress out of me. Lastly, I have so many thanks to say to my beloved husband, Ilya. Five years ago, he suggested that I pursue a PhD in Australia and that we go on this priceless adventure. He stayed up (more accurately, fell asleep) listening to me rehearse my presentations on repeat until I felt that they were perfect. I would not have been able to complete my PhD without his IT knowledge in using the cloud to do all of the computing for my analysis, and his help in figuring out complicated coding loops which turned screams of frustration into those of joy. He brought me meals to the lab when I was stuck there until late. His support in my goals of completing a PhD while still training as a competitive athlete has helped to bring out the best in me, which there is nothing more I could ask for. I dedicate my thesis to all of the people mentioned above who have taught, helped, and guided me for which I am deeply grateful; it takes quite a lot of people to complete a PhD.

List of Publications

The following work has been accepted for publication at peer-reviewed journals in support of this thesis:

1. **Landen, S**., Voisin, S., Craig, J. M., McGee, S. L., Lamon, S., & Eynon, N. (2019). Genetic and epigenetic sex-specific adaptations to endurance exercise. ***Epigenetics***, *14*(6), 523-535. (part of Chapter 2)
2. **Landen, S**., Jacques, M., Hiam, D., Romero, J. A., Harvey, N. R., Haupt, L. M., ... & Eynon, N. (2021). Skeletal muscle methylome and transcriptome integration reveals profound sex differences related to muscle function and substrate metabolism. ***Clin Epigenetics.*** 2021 Nov 3;13(1):202. doi: 10.1186/s13148-021-01188-1. (Chapters 3 and 4)
3. **Landen, S**.*, Hiam, D.*, Lamon, S., Voisin, S., Eynon, N. Physiological and molecular sex differences in human skeletal muscle in response to exercise training. ***J Physiol***. 2021 Nov 11. doi: 10.1113/JP279499. **shared first authorship* (part of Chapter 2)

The following work is being prepared for publication at peer-reviewed journals in support of this thesis:

1. **Landen, S**., Jacques, M., Hiam, D., Romero, J. A., Harvey, N. R., Haupt, L. M., ... & Eynon, N. (*Being prepared for submission*). Sex comparison in the epigenetic response to exercise training. (Chapter 5)

The following work has been published or submitted in a peer-reviewed journal during my candidature, and is outside the scope of this thesis:

1. Woessner, M.N., Hiam, D., Smith, C., Lin, X., Zarekookandeh, N., Tacey, A., Parker, L., **Landen, S**.,…, Voisin, S., Eynon, N., Duque, G., Levinger, I., Osteoglycin across the adult lifespan (*accepted; **The Journal of Clinical Endocrinology & Metabolism***).
2. Jacques, M., **Landen, S**., Alvarez Romero, J., Yan, X., Garnham, A., Hiam, D., ... & Voisin, S. (2021). Individual physiological and mitochondrial responses during 12 weeks of intensified exercise. ***Physiological Reports***, 9(15), e14962.
3. Voisin, S., Jacques, M., **Landen, S**., …, Thomis, M., Sharples, A. P., Schürmann, A., Roden, M., Horvath, S., and Eynon, N. (2021) Meta-analysis of genome-wide DNA methylation and integrative omics of age in human skeletal muscle. ***Journal of Cachexia, Sarcopenia and Muscle***, 12, 1064– 1078, Hiam, D., Smith, C., Voisin, S., Denham, J., Yan, X., **Landen, S.,** ... & Herrmann, M. (2020). Aerobic capacity and telomere length in human skeletal muscle and leukocytes across the lifespan. ***Aging*** (Albany NY), 12(1), 359.
4. Hiam, D., Voisin, S., Yan, X., **Landen, S**., Jacques, M., Papadimitriou, I. D., ... & Eynon, N. (2019). The association between bone mineral density gene variants and osteocalcin at baseline, and in response to exercise: The Gene SMART study. ***Bone***, 123, 23-27.

5. Yan X, Dvir N, … **Landen S**, … Eynon N. (2018) The ACE I/D gene variant predicts ACE enzyme content in blood but not the ACE, UCP2 and UCP3 protein content in human skeletal muscle in the Gene SMART study. ***J Appl Physiol***. 2018 Jun 21.

The following work has been published as a book chapter:

1. Jacques M*, **Landen S***, Palmer A, Eynon N. (2020) Handbook of Stress: Epigenetic effects of exercise on human skeletal muscle. ***Elsevier***. **shared first authorship*
2. Jacques M, **Landen S**, Voisin S, Eynon N. (2018) Summary Findings on Genetics and Sport Performance. The Routledge Handbook of Sport and Exercise System Genetics. Taylor & Francis.
3. Jacques M, **Landen S**, Voisin S, Lamon S, Eynon N. (2018) Nurture vs Nature: The Genetics and Epigenetics of Exercise. Research Methods in Physical Activity and Health.

Conferences

Conference Poster Presentations (4)

- Cell Metabolism Symposium – *Spain 2019*
- Australian physiological society (AuPS) – *Melbourne 2017 and Sydney 2018*
- Victoria University PhD student conference – *Melbourne 2017*

Conference Oral Presentations (5)

- Canadian society of exercise physiology (CSEP) – *Canada 2018*
- Victoria University high degree research conference – *Melbourne 2018, Melbourne 2019*
- Muscle Network Symposium – *Melbourne 2019, online 2020*
- Australian Physiological Society (AuPS) – *Canberra 2019* (invited speaker)

Funding, Grants, and Awards

2017-2021 Victoria University Central Grant Research Scheme (~$28,000 per year)

2019 Secomb grant [$253], Victoria University

2019 Physiological Society grant [500 pounds]

2019 Victoria University high degree research (HDR) conference people's choice oral presentation [$50]

2020 Three minute thesis (3MT) oral presentation Victoria University runner-up [$250]

2018-2020 Victoria University HDR committee chair

2019 Australian physiological society (AuPS) research training award

Table of Contents

List of Tables

List of Figures

List of Abbreviations

4WP = four week post [training]

ACE = angiotensin converting enzyme

ACTN3 = alpha-actin-3

Akt = protein kinase B

ALDH1A1 = Aldehyde Dehydrogenase 1 Family Member A1

AR = androgen receptor

ARE = androgen response element

ATP = adenosine triphosphate

BP = biological process

CC = cellular component

CHO = carbohydrates

CpG = cytosine phosphate guanine dinucleotide

CSA = cross-sectional area

dbGAP = database of genotypes and phenotypes

DEG = differentially expressed gene

DMG = differentially methylation gene

DMP = differentially methylated position

DMR = differentially methylated region

DNMT = DNA methyltransferase

E1 = estrone

E2 = estradiol

E3 = estriol

eQTL = expression quantitative trait loci

ER = estrogen receptor

ERE = estrogen response element

ERα = estrogen receptor alpha

ERβ = estrogen receptor beta

ETC = electron transport chain

EWAS = epigenome-wide association study

FDR = false discovery rate

FOXO3A = Forkhead Box O3

FSH = follicle stimulation hormone

FT = fibre type

GEM = genome-scale metabolic model

Gene SMART = skeletal muscle adaptive response to training

GEO = gene expression omnibus

GGT7 = Gamma-Glutamyltransferase 7

GLUT4 = glucose transporter type 4

GO = gene ontology

GPER = G-protein coupled estrogen receptor

GSEA = Gene set enrichment analysis

GTEx = genotype-tissue expression

GWAS = genome-wide association study

GXT = graded exercise test

HAT = histone acetyltransferase

HDAC = histone deacetylase

HIIT = high intensity interval training

HMFDR = harmonic mean of the individual component FDR

IHC = immunohistochemistry

KEGG = Kyoto encyclopedia of genes and genomes

Lfsr = local false sign rate

LH = luteinising hormone

LT = lactate threshold

meQTL = methylation quantitative trait loci

MF = molecular function

miR = myomiR

miRNA = micro RNA

mTOR = mammalian target of rapamycin

MU = motor units

NR3C1 = glucocorticoid receptor

OR = odds ratio

OXPHOS = oxidative phosphorylation

PC = principal component

PCA = principal component analysis

PcG = polycomb group

PGC1α = peroxisome proliferator activated receptor gamma coactivator 1 alpha

PP = peak power output

PR = progesterone receptor

PRC1 = polycomb repressive complex 1

PRC2 = polycomb repressive complex 2

qPCR = quantitative polymerase chain reaction

RAS = renin-angiotensin system

RER = respiratory exchange ratio

ROS = reactive oxygen species

SHBG = sex hormone-binding globulin

SNP = single nucleotide polymorphism

SRY = sex-determining region Y

T2D = type two diabetes

TCA = tricarboxylic acid

TET = ten-eleven translocation

TF = transcription factor

TFBS = transcription factor binding site

TssA = active promoter

TssAFlnk = flanking active promoter

VO_2max = velocity of maximum oxygen consumption

Chapter 1 : Introduction

Exercise training induces many physiological adaptations that ultimately promote health. It is becoming increasingly evident that there are sex differences in the physiological (e.g. metabolic) response to exercise training [2], with the underlying molecular mechanisms poorly characterised. Exercise training adaptations are mediated by increased transcription of key regulatory, metabolic, and myogenic genes. Epigenetic modifications, particularly DNA methylation, are emerging as important crucial events for increased transcription [3]. A handful of studies reported changes in DNA methylation patterns after exercise training [4]. Nevertheless, the field of exercise epigenetics is relatively new and there remains much to elucidate regarding downstream phenotypic changes.

Skeletal muscle is the largest tissue in the human body and is functionally involved in exercise, therefore serving as an ideal tissue to study exercise training adaptations. Several studies have identified transcriptome-wide changes in skeletal muscle in response to exercise training [5], with a recent meta-analysis reporting differences between males and females [6]. Despite skeletal muscle displaying transcriptomic sex differences both at baseline [7-11] and in response to exercise training [6], neither sex differences in the skeletal muscle epigenome at baseline nor in response to exercise training have been investigated to date. To examine whether there is a sex-specific DNA methylome response to exercise training, it was important to first understand whether there are sex differences at baseline. Therefore, **the aim of the first study (Chapter 3) was to determine whether there are baseline epigenomic (i.e. DNA methylome) differences between male and female skeletal muscle using a large-scale meta-analysis of three independent cohorts.**

The aim of the second study (Chapter 4) was to explore biological factors underlying the DNA methylation differences between males and females at baseline, by

investigating the effect of fibre type proportions, circulating hormone levels, and transcription factors. Lastly, with the understanding of the baseline DNA methylome sex differences, we could address whether there are training-induced sex differences. **Therefore the aim of the third study (Chapter 5) was to investigate whether there is a sex-specific response to exercise training in the skeletal muscle DNA methylome.**

Commencing with a literature review (**Chapter 2**), this thesis further comprises three experimental chapters:

I. **Chapter 3**: Skeletal muscle methylome and transcriptome integration reveals profound sex differences related to muscle function and metabolism.

II. **Chapter 4:** Biological factors contributing to DNA methylome sex differences in human skeletal muscle.

III. **Chapter 5**: Sex-specific DNA methylation in skeletal muscle in response to exercise training and lifelong physical activity.

The main findings of this thesis are summarised with a general discussion (**Chapter 6**), including the limitations of each study presented and recommendations for future research.

Chapter 2 : Review of Literature

This chapter consists of a combination of two review articles, one published in a peer-reviewed journal [12] and the other under review in a peer-reviewed journal [13].

VICTORIA UNIVERSITY

THE NEW WAY TO DO UNI

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS

This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

Title of Paper/Journal/Book:	Genetic and epigenetic sex-specific adaptations to endurance exercise. Published.
Surname:	Landen
First name:	Shanie
Institute:	Institute for Health and Sport
Candidate's Contribution (%):	72

Status:

Accepted and in press:		Date:	
Published:	x	Date:	2019

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

Signature	Date
Shanie Landen Digitally signed by Shanie Date: 2021.05.08 16:26:55 +10'00'	2021

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:

1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83776954731
CRICOS Provider No. 00124K (Melbourne),
02475D (Sydney), RTO 3113

3. There are no other authors of the publication according to these criteria;
4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and
5. The original data will be held for at least five years from the date indicated below and is stored at the following **location(s)**:

Name(s) of Co-Author(s)	Contribution (%)	Nature of Contribution	Signature	Date
Séverine Lamon	5%	Writing and editing of manuscript	[redacted]	08/08/2021
Jeffrey M. Craig	2%	Editing of manuscript		08/08/2021
Sean L. McGee	5%	Writing and editing of manuscript		08/08/2021
Nir Eynon	8%	Writing and editing of manuscript		08/08/2021
Sarah Voisin	8%	Writing and editing of manuscript		08/08/2021

Updated: September 2019

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

VICTORIA UNIVERSITY

THE NEW WAY TO DO UNI

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS

This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

Title of Paper/Journal/Book:	Physiological and molecular sex differences in human skeletal muscle in response to exercise and training. Published.
Surname:	Landen
First name:	Shanie
Institute:	Institute for Health and Sport
Candidate's Contribution (%):	37

Status:

Accepted and in press:		Date:	
Published:	x	Date:	2019

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

Shanie Landen Digitally signed by Shanie Date: 2021.05.08 16:26:55 +10'00'	2021
Signature	**Date**

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:

1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83776954731
CRICOS Provider No. 00124K (Melbourne),
02475D (Sydney), RTO 3113

3. There are no other authors of the publication according to these criteria;
4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and
5. The original data will be held for at least five years from the date indicated below and is stored at the following **location(s)**:

Name(s) of Co-Author(s)	Contribution (%)	Nature of Contribution	Signature	Date
Danielle Hiam	37	Writing and editing of manuscript	[redacted]	08/08/2021
Macsue Jacques	3%	Editing of manuscript		08/08/2021
Séverine Lamon	10%	Writing and editing of manuscript		08/08/2021
Nir Eynon	8%	Writing and editing of manuscript		08/08/2021
Sarah Voisin	5%	Editing of manuscript		08/08/2021

Updated: September 2019

2.1 Abstract

In recent years, the interest in personalised interventions such as medicine, nutrition, and exercise is rapidly rising to maximise health outcomes and ensure the most appropriate treatments. Exercising regularly is recommended for both healthy and diseased populations to improve health. However, there are sex-specific adaptations to exercise that often are not taken into consideration. While endurance exercise training alters the human skeletal muscle epigenome and subsequent gene expression, it is still unknown whether it does so differently in men and women, potentially leading to sex-specific physiological adaptations. Elucidating sex differences in genetics, epigenetics, gene regulation and expression in response to exercise training will have great health implications, as it may drive discovery and deepen the current understanding of the health benefits of physical activity in humans.

2.2 Introduction

As hunter-gatherers, males and females had different contributions to subsistence in society. While there are divergent views of their respective roles among different societies, anthropologists agree that there has always been a division of labour between males and females. For example, while the indigenous males and females of Paraguay both engaged in a high level of activity by traveling great distances and carrying heavy items, males were often hunting while women were moving the household and involved in childcare [14, 15]. The roles of males and females in society throughout history are correlated with their physiological strengths and weaknesses. Sex differences are defined as significant differences in the means of a phenotype between biological males and females, as opposed to sexual dimorphisms, which are defined as two distinct forms of a trait that differentiate members of the same species by their sex (such as ovaries versus testes) [16, 17].

As a result of females being significantly under-represented in exercise research [18-20] our current understanding of exercise physiology is overwhelmingly inferred from cohorts solely containing males, or mixed cohorts where potential sex differences have not been considered. Yet, it is well-known that males and females display distinct musculoskeletal, cardiovascular, molecular, and metabolic features [2]. Compared with females, males tend to be taller, heavier, have greater lean body mass and lower fat mass, and have higher proportion of fast-twitch (type II) muscle fibres in some muscle groups [21, 22]. During exercise, males rely more on carbohydrates (CHO) and proteins, as opposed to females, who rely more on fats as a fuel source [23, 24]. There are numerous contributors to these physiological differences between the sexes, including sex hormones, genetics, and gene-by-environment interactions (e.g., epigenetics) [25]. This narrative literature review will provide comprehensive physiological and molecular insights into the basis of sex differences in the response to exercise, as well as future directions for research in the field in the era of high-throughput technologies and -OMICs.

2.3 Physiological phenotype

2.3.1 Endurance-related phenotypes

Overall, males have higher maximal oxygen uptake (VO_2max; absolute and relative to both lean and total body mass) than females, even when matched for training levels [26-29]. While not without controversy [30], a recent meta-analysis (n = 175; males n = 90; females n = 85) concluded that males show a greater increase in VO_2max relative to body mass following endurance training, regardless of the length of intervention and exercise intensity [31], suggesting that between-studies discrepancies might be due to small sample sizes.

Exercise economy, defined as the amount of energy spent per unit of velocity [32], is a common indicator of endurance performance. Sex differences in exercise economy have been

reported but largely depend on the type of exercise. In swimming, females generally have more suitable technique than males and therefore have better economy relative to body size [33]. In endurance running, most studies agree that males and females have comparable running economy [34], but some studies found that males may be more economical [35]. Lactate threshold is another commonly used measure of endurance performance [36]; it is generally described at the exercise intensity in which the increase in lactate is no longer linear, leading to an accumulation of lactate [37]. One study, which controlled for the menstrual cycle by testing women during anovulation in a relatively small cohort, reported no difference between the sexes in the change in lactate threshold (LT_2, upper limit of LT) following 10 sessions of high intensity-training [38]. Further studies which take menstrual cycle phase into account are needed to compare lactate thresholds between the sexes [39], as some [40, 41], but not all [42-44], report higher blood lactate levels post exercise during the follicular phase signifying less efficient lactate clearance.

2.3.2 Resistance-related phenotypes

Muscle strength underlies many sex differences in exercise performance. Overall, males' upper and lower body strength is greater than females' by 157% and 60% relative to total body mass, respectively [45]; a trend that is observed in recreationally active [46] and trained (matched for training status) [47, 48] males and females. Female lower body strength relative to body mass is also greater than relative upper body strength, a phenomenon absent in males [45, 49]. Despite these potential differences, males and females increase strength to a similar degree following resistance training when expressed in mass-relative terms [50-52], and two studies have reported greater gains in females [53-55]. Underpinning resistance training-induced increases in muscle size and strength is muscle protein turnover [56], which is defined as the balance between muscle protein synthesis and muscle protein degradation. When normalised to lean mass, muscle protein synthesis and degradation rates are similar in

males and females both at rest [57] and in response to training [58], suggesting that endogenous factors prime life-long sex differences in muscle strength that are independent from training.

2.4 Molecular differences between male and female muscle

2.4.1 Morphology of skeletal muscle

Males and females display inherent differences in muscle fibre type, size and distribution. Males have larger fibre cross-sectional area (CSA) that show more type II fibres characteristics, whereas females have smaller fibres that show more type I characteristics [59-61]. Long-term resistance and endurance training alter fibre type proportions and CSA [62]. Resistance training increases CSA of type IIa and IIx fibres [63, 64]. Some [52, 54, 65, 66] but not all [67] studies found similar increases in muscle CSA in response resistance training between the sexes. It has been suggested that the differing findings may be due to resistance training targeting upper or lower body. For example, after a 12-week training intervention targeting the upper body, males, but not females, significantly increased their muscle CSA [67].

High-intensity sprint training resulted in increased type IIx CSA only in females [68]. Endurance training resulted in an increase in type I and decrease type II fibre proportions to a similar degree in both males and females [69]. Studies investigating sex differences in fibre type shifting with training mostly included small sample sizes (< 10 of each sex); therefore, a comprehensive meta-analysis would improve power and consolidate these findings.

At baseline, there are absolute differences in the proportion of fibre types between sexes. Males and females seem to adapt to resistance training to a similar degree, which means that any baseline differences remain evident after the intervention. In endurance or resistance training, exercise intensity, volume and speed of contraction are not only important for fibre type adaptation, but also for sex-specific adaptions.

2.4.2 Neuromuscular phenotypes

Absolute differences in the a) proportion of fibre types and b) CSA of the fibre between sexes result in differences in motor unit activation. Males have larger type II fibres which is associated with recruitment of higher threshold motor units (MU) meaning they exhibit higher action potential amplitudes and can produce a higher contractile force. In contrast, females generally have lower threshold MUs due to the smaller CSA of type II fibres resulting in lower action potential amplitudes and produce a lower contractile force. Therefore, males rely on lower firing rates to modulate contractile force compared to females [70, 71]. These sex difference in neuromuscular activation are important in explaining the apparent differences in fatigue between males and females. Most neuromuscular adaptations in response to exercise occur due to changes in fibres type proportion and CSA [72].

Whether there are sex differences in neuromuscular adaptations in response to exercise training has not been explored extensively. In response to 10-weeks of endurance cycling, there were no differences in neuromuscular adaptations between sexes [70]. No study has investigated whether there are sex differences in neuromuscular adaptation in response to resistance training. Yet there is some evidence after an acute bout of heavy resistance exercise that in males but not females had significant decreases in maximal voluntary neural activation, or the capacity of the nervous systems to fully activate skeletal muscle [73-75]. This suggests a greater impairment in neuromuscular activation in males compared with females after fatiguing exercise [74]. Overall, there may be subtle albeit functionally important sex differences in neuromuscular adaptations to various exercise interventions. However, the available literature is conflicting, most likely because the adaptations are specific to the exercise intensity, velocity, and the targeted muscle groups.

2.4.3 Metabolic adaptations

Sex differences in substrate metabolism during exercise have been well-documented. Females have lower respiratory exchange ratio (RER) during endurance exercise [76, 77], indicating higher beta oxidation. This is associated with higher adipocyte lipolysis, as well as greater intramyocellular lipid content and use [22, 78]. Females also have lower blood glucose appearance and disappearance rates, do not increase muscle glycogen content in response to a high carbohydrate diet (0%), and may spare glycogen stores in muscle during endurance exercise; although the latter has not been consistently reported [22, 76]. During exercise, females oxidize less protein compared with males [79-83]. Specifically, females show lower oxidation of leucine, an amino acid that plays a central role in intracellular signalling during and after exercise [84]. However, the mechanisms underlying the observed sex differences in amino acid metabolism are contradicting, likely due to differences in training status of the participants, specifically the training volumes and intensities, which effect energy balance [24]. Various exercise intensities, types, and durations may all be key in determining sex differences in exercise metabolism.

A seminal study by Carter *et al*. found no sex differences in the maximal activity of key metabolic enzymes involved in β-oxidation, the tricarboxylic acid (TCA) cycle, and the electron transport chain (ETC), neither at baseline nor in response to seven weeks of endurance exercise [59]. This suggests that differences in substrate metabolism during exercise are not due to differences in selected key enzymes within the β-oxidation, TCA and ETC metabolic pathways. It remains to be explored whether other enzymes involved in β-oxidation (i.e., long chain acyl CoA dehydrogenase, enoyl CoA hydratase, keto-thiolase), or the efficiency of these enzymes (K_m), as opposed to the commonly measured maximal activity, are different between the sexes.

Although no sex differences in response to exercise or training have been reported in enzymatic activity, studies investigating candidate muscle proteins have consistently found sex differences in fat oxidation-related proteins, while the literature regarding CHO metabolism-related proteins has some inconsistencies. This implies that fat oxidation is regulated during endurance exercise and that carbohydrate and protein oxidation follow by metabolic demand [24]. There has yet to be a protein-wide (proteomic) study in skeletal muscle aimed at investigating potential sex differences either at baseline or in response to exercise. Due to the nature and complexity of exercise interventions, sample size is often a limiting factor. Larger scale and consortium-based studies [85] are therefore warranted to elucidate sex differences in various aspects of exercise training.

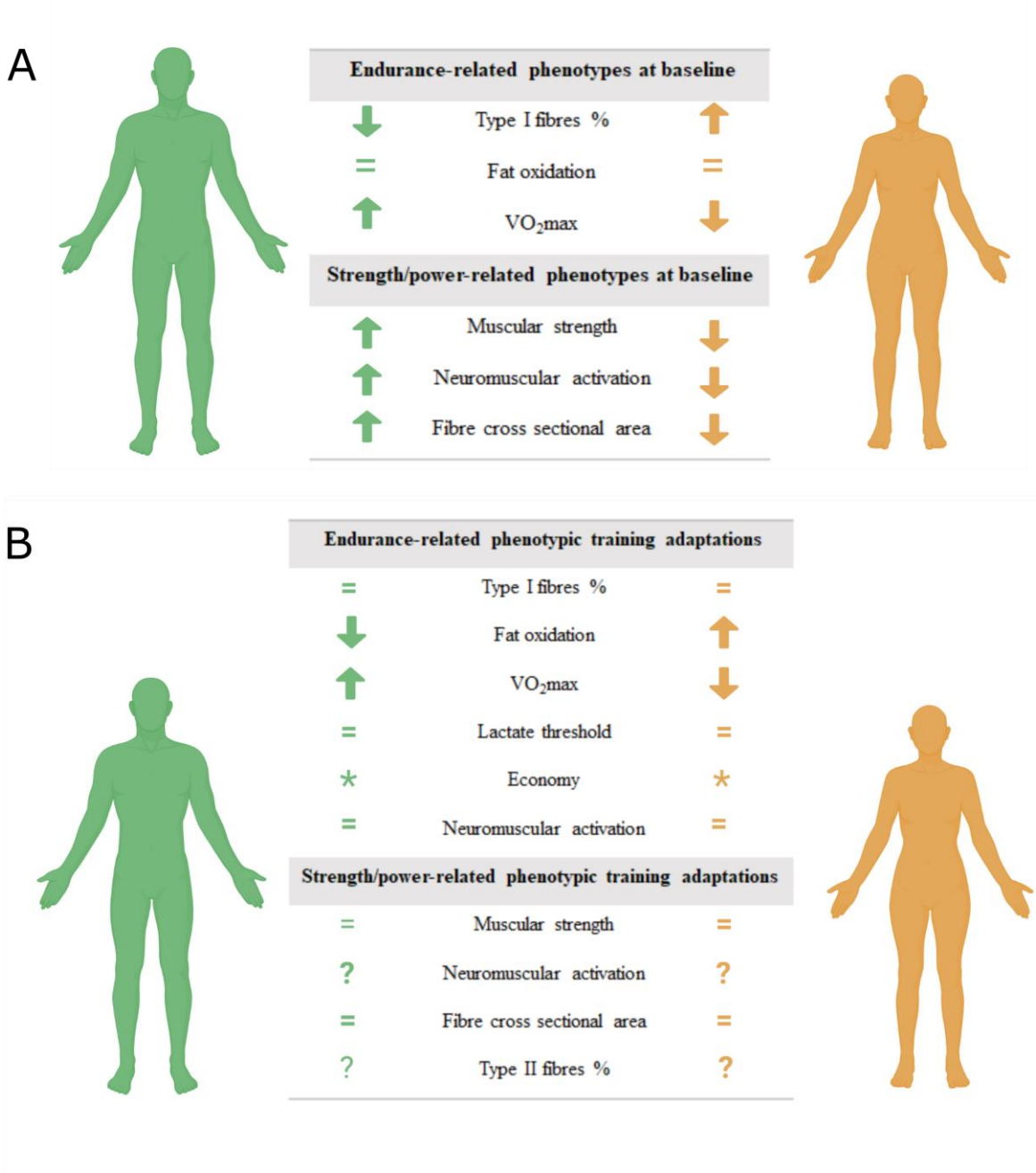


Figure 2.1. Phenotypic sex differences at baseline and after training.

(**A**) Phenotypic sex differences at baseline in endurance- and strength/power- related factors. Direction of arrows indicate which sex generally presents a stronger phenotype. Males denoted by green; females denoted by orange. (**B**) Upwards arrows indicate which sex generally presents a stronger *change* in phenotype following endurance and strength/power training; equal sign indicates an equivalent response to training; question mark indicates that the sex differences is inconclusive or has not been thoroughly studied; asterisk means that it is dependent on modality of exercise. Males denoted by green; females denoted by orange.

2.5 Mechanisms underlying sex differences

2.5.1 Genetics: the influence of sex chromosomes

The sex chromosome complement determines the sexual differentiation of the gonads, the sex hormone milieu, and can directly mediate sex differences independent of sex hormones [86, 87]. For example, the Y chromosome contains the unique Sex-determining Region Y (*SRY*) gene that plays a major role in development of male gonadal tissues but also influences autosomal gene expression in other tissues [88-91]. The non-coding *XIST* gene on the X chromosome controls a female-specific process known as allele dosage compensation in which one X chromosome is silenced [92, 93]. However, approximately one-third of all X chromosome genes (up to 60) 'escape' silencing and remain transcriptionally active in XX cells resulting in sex-biased gene expression [86, 88, 89, 91, 93-95]. Finally, the maternal and paternal X chromosomes carry distinct genomic imprints that silence the expression of specific genes and result in sex-specific differences in gene expression levels [96, 97].

Altogether, these studies show how the sex chromosome complement can modulate autosomal genes with many of these genes involved in critical cellular functions such as energy metabolism [86, 91, 98]. Whether the sex chromosome complement modulates the sex differences in energy substrate utilisation in response to exercise training, is yet to be established [99]. Disentangling the direct effects of sex chromosomes from the effects of sex hormones is difficult in humans; the direct effect of the sex chromosomes in exercise adaptations has yet to be adequately studied in humans.

2.5.2 The hormonal environment: the influence of sex hormones

The hormonal environment is key for exercise adaptations. Lifelong differences in exposure to sex hormones between males and females is determined by the sex chromosome complement composition. Sex steroid hormones are primarily ascribed to reproduction, but

their importance to non-reproductive functions is becoming more apparent [99, 100], including in skeletal muscle [101]. Sex steroid hormones engage through their specific ligand-receptors [9, 102, 103] and influence transcription as well as phenotypes in a tissue- and sex- specific manner [104-108]. Not only can these receptors be differentially expressed between sexes [109], some also show sex-biased gene targeting patterns due to intrinsic differences in sex hormone levels [103, 110]. In addition, males and females are different in their gene regulatory patterns, meaning that the same genes are often regulated by different transcription factors (TFs) or combination of TFs [103]. This can result in differential gene expression or program latent sex differences in gene regulation, which may only become apparent under specific conditions of age, stress, disease or therapeutic treatments. Males and females may therefore be primed to respond differently to physiological stressors [103, 111], such as exercise [2, 10, 12, 24, 77, 78].

The major bioactive sex steroids (testosterone, estradiol and progesterone) derive from a common biosynthesis pathway originating from cholesterol and, while mostly produced by the gonads, are also synthetized by the adrenal glands and a number of peripheral tissues [101]. Both males and females produce androgen hormones (such as testosterone) and ovarian sex hormones (such as estrogens and progestogens), however in amounts that can vary by several orders of magnitude depending on sex and menopausal status.

2.5.2.1 Androgen Hormones

The major androgen hormone, testosterone, exerts its effect through the androgen receptor (AR) and is expressed in male and female muscle cells [109, 112]. In both males [113, 114] and females [115], testosterone directly activates muscle protein synthesis by triggering the Akt/mTOR pathway [116]. In males, it may also inhibit muscle protein degradation pathways [117, 118] and promote the recruitment of mesenchymal pluripotent stem cells into

the myogenic lineage [119], however these effects have not been observed in females [115]. Further, while testosterone levels are consistently associated with muscle mass and strength in males [120, 121], they are not in females [122], suggesting that females rely on other hormones to activate muscle protein synthesis.

In males, circulating testosterone increases with acute bouts of resistance exercise [109, 123, 124] but the effects of long-term resistance training are less clear. While hormonal changes in response to an acute bout of exercise are more challenging to interpret in females due to menstrual variability [125], most studies report no post-resistance exercise surge in female testosterone levels [126]. The exercise-induced testosterone response may facilitate the activation of muscle protein synthesis in males, possibly by increasing AR content [127]. Despite the absence of a post-exercise testosterone peak, the post-exercise upregulation of the AR content is faster in females than in males [109], outlining further, intrinsic sex-specific differences in exercise response. In contrast, several studies reported short-term increases in female testosterone levels after an acute bout of endurance exercise [126], with less conclusive and sometimes conflicting data in males [128, 129].

2.5.2.2 Ovarian Hormones

Estradiol, an estrogen, and progesterone, a progestogen, are the major bioactive ovarian hormones in pre-menopausal females. A few studies investigated the potential role of ovarian hormones on muscle and exercise adaptation in females [101, 126, 130], but, potentially owing to the low natural concentrations in males, male-specific studies remain scarce. Like testosterone, progesterone can activate muscle protein synthesis in post-menopausal females [115]. The effects of estrogens on muscle protein synthesis and hypertrophy are less clear. Estradiol has no effect on muscle protein turnover in post-menopausal female muscle [115, 131] but may nevertheless increase muscle mass when combined with resistance exercise [131,

132]. Ovarian hormones may also help maintain skeletal muscle mass by enhancing the pool of satellite cells [133, 134] or increasing the number of force-generating cross-bridges in the muscle [135]. Rodent models provide further insights into the sex-specific role of ovarian hormones in muscle performance [130]. For example, it is well known that female mice can run more than males in running wheels. Ovariectomy and the resulting suppression of ovarian sex hormone secretion however abolishes this difference [136]. In humans, the well-known sex differences in muscle fatigability and whole-body substrate oxidation have been historically attributed to a higher proportion of type I muscle fibres in females [74]. It is becoming evident that ovarian sex hormones too are modulating substrate utilisation during exercise [24, 77, 78, 130]. Males receiving estradiol showed increased fat oxidation and expression levels of the associated genes, and decreased CHO and protein oxidation [137]. Similarly, suppression and selective sex hormone replacement in females to achieve a high estrogen/low progestogen environment decreased CHO oxidation by reducing hepatic glucose production and the use of muscle glycogen [138]. Aside from modulating substrate utilisation, the estrogen receptor complex can modulate mitochondrial biogenesis and function via regulation of expression of mitochondrial proteins (such as PGC1α), ATP production, reactive oxidative species (ROS), and antioxidant defences [139, 140], adding to the multiple mechanisms through which ovarian hormones may regulate skeletal muscle adaptation in a sex-specific manner.

Consitt *et al.* [126] have previously reviewed the ovarian sex hormone response to different exercise modalities in females. They concluded that acute endurance and, to a certain extent, resistance exercise, may trigger an increase in estradiol concentrations that depend both on the exercise intensity and the phase of the menstrual cycle. Very little data are available for endurance and resistance training, or for progesterone, and primarily in the context of the female athlete amenorrhea [141].

2.5.3 Sex differences in gene expression

Multiple studies [7-11, 142-144] have analysed sex differences in the skeletal muscle transcriptome. These studies found between 60 and 3,000 differentially expressed genes between the sexes, depending on the technique utilized. The Genotype-Tissue Expression (GTEx) database contains over 1,000 mRNA expression profiles in skeletal muscle and identified 13,294 sex-biased genes across all tissues, including 2,866 (2,689 on autosomes) in skeletal muscle [9]. To overcome the low statistical power of individual studies investigating sex differences in the transcriptomic response to exercise and training, two recent, large-scale meta-analyses [5, 6] pooled the results of several studies with different exercise modalities (acute and chronic endurance and resistance exercise). One of these meta-analyses investigated sex differences, and identified 247 genes (across 43 studies, including 739 individuals) whose response to training differed between the sexes and which were primarily involved in chromatin organization [6]. More large-scale studies that not only include both males and females, but also treat sex as a confounder, are required to elucidate the phenotypic consequences of these transcriptomic sex differences in the response to training. A study comparing the skeletal muscle transcriptomes of endurance-trained and untrained males and females found that sex differences are attenuated in trained individuals, suggesting that training makes the transcriptome of males and females more similar [10].

2.5.4 Epigenetics: the influence of DNA methylation, histone modifications, and miRNAs

Epigenetic mechanisms allow for an organism to respond to its environment via changes in gene expression. Epigenetic modifications can be defined as the structural adaptation of chromosomal regions that bring about altered activity states [145, 146]. The main types of epigenetic modifications include DNA methylation, histone modifications, and non-coding RNA including microRNA (miRNA) and long noncoding RNA expression [147].

Epigenetic events up- or down- regulate gene expression and corresponding protein translation, resulting in phenotypical and physiological changes [148].

Epigenetic mechanisms interact with one another to alter the epigenetic state of the genome and establish appropriate gene expression patterns. Given their fundamental role in developmental biology, these epigenetic processes are often evolutionarily conserved across species [149]. Furthermore, their deregulation is associated with diverse developmental phenotypes [150]. Genomic imprinting is an epigenetic process that involves the interaction of DNA methylation and polycomb group protein (PcG) repression which subsequently regulates allele-specific gene expression [151]. We typically inherit two working copies of a gene from each parent; however, close to 100 human genes are imprinted [152], meaning that only one allele is expressed in a parent-specific manner. These loci are conserved among humans, meaning that a maternal locus will always express the inherited maternal allele [97, 153]. Imprinted genes are of great medical significance since they are essential for healthy offspring development, and imprinting dysregulations may lead to metabolic and neurodevelopmental disorders [154, 155].

PcG proteins are chromatin modifiers which typically belong to one of two distinct multi-protein complexes: polycomb repressive complex 1 (PRC1) and 2 (PRC2). Each complex has distinct catalytic activities, nevertheless, both are generally associated with transcription silencing. PRC1 exhibits ubiquitin ligase activity that targets specific histone lysine residues. PRC2 is a methyltransferase that targets a specific histone lysine residue of histone. The trimethylated histone is abundant in facultative heterochromatin (transcriptionally silenced chromatin). PcG complexes play an important role in gene silencing and regulation, and therefore their epigenetic activities are highly involved in mammalian development [150].

The mammalian male and female autosomal epigenomes (DNA methylation, histone modifications, and miRNA) display considerable differences in tissues such as human blood, saliva and skeletal muscle as well as mouse liver and brain [156-164]. Recently, it was suggested that epigenetic modifications influence exercise adaptation [165], and comprehensive reviews have described the potential regulatory effects of epigenetic modifications in the response to exercise training [146, 154, 165-169]. Epigenetic differences may therefore explain some of the sex differences observed in exercise adaptations. Our current understanding of exercise adaptations is based on studies that have mostly investigated only males or grouped males and females together, and have not taken into consideration the potential sex differences in exercise adaptations. Furthermore, there may be sex differences in the epigenetic response to exercise. Since epigenetic changes are associated with health and disease (i.e., cancer and metabolic disorders) [170, 171], and exercise influences epigenetics, epigenetics may be one of the underpinning mechanisms behind the lower disease rate in physically active individuals [172]. Therefore, it is important to elucidate the sex differences in exercise epigenetics.

2.5.4.1 DNA methylation

DNA methylation is the addition by DNA methyltransferase (DNMT) enzymes of a methyl group to the 5' position of a cytosine base. DNA methylation alters protein-protein and protein-DNA interactions, affecting chromatin structure and ultimately increasing or decreasing transcription [173]. DNA methylation is stable through cell divisions, yet dynamic throughout one's lifetime as it is influenced by environmental stimuli (such as exercise training and nutrition) [148, 174]. Previous studies have shown that exercise triggers small (< 10%) and widespread DNA methylation changes in skeletal muscle [175, 176]. To date, two studies have suggested that there may be sex-specific changes in skeletal muscle DNA methylation following exercise, given that sex was a major determinant of variability [176] and that larger

effect sizes were observed in females [154], however these potential sex differences were not further investigated (discussed below). Exercise epigenetics is a new and exciting research field, and we currently have limited knowledge on how epigenetic signals, such as DNA methylation, mediate exercise responses.

A seminal study in 2012 reported lower DNA methylation in specific genes 20 minutes after a bout of high-intensity endurance exercise [174], demonstrating the rapid dynamics of DNA methylation. Potential sex-specific responses were not investigated in this study. However, the rapid demethylation of exercise-responsive genes shows that acute control of DNMT activity during exercise is important for this response. *In vitro* studies suggest that DNMT3B is an important regulator of this gene program [177]. Interestingly, DNMT3B expression in human liver is significantly higher in females than males [178], although it is unclear whether this is also the case in skeletal muscle. While DNMTs are involved in DNA methylation, ten-eleven translocation (TET) enzymes are involved in DNA demethylation. TET enzymes are expressed in human skeletal muscle [179], and given how recently they were discovered, sex differences in skeletal muscle TETs have yet to be investigated. However, one study did not find sex differences in TET expression in mouse hippocampal tissue [180]. Nonetheless, unravelling the dynamics of DNMTs and TETs in both sexes is warranted to reveal the nature of DNA methylation in exercise adaptations.

A recent study is the first to thoroughly investigate DNA methylation sex differences in a skeletal muscle-related human tissue- cultured myoblasts and myotubes (13 men, 13 women). Genome-wide DNA methylation and gene expression (measured with microarrays) were performed on the autosomes and the X chromosomes. Several pathways related to the cell cycle and energy, protein and fatty acid metabolism were enriched in females while pathways mostly related to cell-cell communication (e.g. transforming growth factor-beta, TGF-beta, signalling) were enriched in males. They confirmed the direct DNA methylation effect on gene

expression using a luciferase assay. They found sex differences in both DNA methylation and gene expression for 40 genes in myoblasts (including *LAMP2* and *SIRT1*), 9 in myotubes (*KDM6A*), and 5 in both myoblasts and myotubes (*CREB5*, *RSP4X*, *SYAP1*, *XIST*, *ZRSR2*). Furthermore, this study found more DNA methylation differences during cell differentiation in females compared to males on the autosomes. These intrinsic differences may contribute to the sex-specific differences observed in muscular phenotypes [181]. These findings highlight the importance of taking sex into account in biomedical research, as future medicine will further benefit from such findings. Furthermore, it reinforces the importance of investigating whether sex differences in DNA methylation are also involved in the adaptation to exercise.

A meta-analysis of 16 studies identified 478 loci across several tissues (307 in skeletal muscle) that undergo methylation changes following either acute (one bout) or chronic exercise (walking, cycling, and tai-chi). DNA methylation changed to a larger degree (i.e., larger effect size) in females than males following exercise, suggesting sex differences in the epigenetic response to training [154]. However, the only two studies in the meta-analysis that investigated skeletal muscle comprised of only males [174, 182], causing these conclusions to not be representative of skeletal muscle tissue. Nevertheless, a sex comparison was not the focus of this study, so specific DNA methylation differences between males and females were not investigated. Additional studies have found that long-term exercise is associated with changes in DNA methylation in human skeletal muscle [175, 176]. After 3 months of one-legged knee extensor exercise training in men and women, 4919 loci were differentially methylated in the exercised leg, compared with the control leg. Training and sex were identified as major determinants of variability in methylation on autosomal DNA. Although sex was treated as a confounder, no statistical analysis could performed to determine whether males and females differed in their DNA methylation response to exercise because sex and batch were confounded in the study design, making it impossible to separate batch effect from sex effect [176].

The meta-analysis by Brown also highlights the importance of identifying sex-differences in exercise-induced methylation of genetically imprinted genes [154]. Five loci that underwent DNA methylation changes following training (chronic exercise) were imprinted loci (two loci in skeletal muscle)[154], however sex differences in those genes were not investigated. No study, to date, has investigated whether there are sex-specific differences in DNA methylation changes at imprinted genes following exercise. An editorial on the topic calls for exercise studies to investigate the effect of timing and dosage of maternal exercise on methylation of imprinted genes in offspring. It is currently hypothesised that the dosage of maternal exercise will influence the offspring epigenome in a dose-dependent manner (i.e., positive effects at low/moderate doses and negative effects at high doses) [183]. Since exercise is a gestational stressor (that leads to epigenetic changes in the gamete) and the susceptibility to gestational stressors differs between the sexes [184, 185], it is likely that maternal exercise affects the gamete epigenome differently between the sexes.

2.5.4.2 Histone Modifications

DNA coils around histone proteins for structural and functional reasons. The amino acid residues within histone tails can be modified by acetylation, phosphorylation, methylation, ubiquitination, sumoylation, or ADP ribosylation. These modifications alter histone-DNA interactions and promote recruitment and access of major transcriptional regulators to DNA. [186, 187]. Like many other post-translational modifications, histone modification is a dynamic process and controlled by numerous enzymes that both add and remove these post-translational modifications. For example, histone acetyltransferases (HATs) add acetyl groups to histone lysine residues, which is a common mechanism to induce transcriptional activation. Histone acetylation generally neutralises electrostatic interactions between histones and DNA, which exposes promoter and gene body regions to transcriptional activators, such as RNA polymerase. Conversely, histone deacetylases (HDACs) remove acetyl groups from histone

proteins, resulting in transcriptional silencing. The localisation of HATs and HDACs to particular chromatin regions is highly dependent on DNA bound transcription factors. Reviews have outlined the effect of acute exercise on histone modifications [165, 187, 188]. For example, skeletal muscle contractions induce phosphorylation and nuclear export of the class IIa HDACs, resulting in the relaxation of chromatin regulatory regions in exercise-related genes [189, 190]. Acute exercise typically induces nuclear export of HDACs 4 and 5, causing hyperacetylation of some histone residues. This results in increased glucose transporter type 4 (*GLUT4*) expression, which supports enhanced energy consumption [191, 192]. Histone deacetylation may therefore regulate the response to exercise. Indeed, genetic disruption of the class IIa HDAC corepressor complex induces exercise-like transcriptional and metabolic adaptive responses [193]. Sex-specific differences in the class IIa HDAC signalling and function in response to exercise have been explored in humans, however no differences were observed [194]. The effect of sex hormones on sex-specific histone modifications and transcriptional responses to exercise is an area that is yet to be explored in any detail. Activated estrogen receptors (ERs) regulate gene expression by altering the balance of HAT and HDAC enzymes at specific chromatin regions, resulting in increased histone acetylation and transcriptional activation [195, 196]. Exercise and ERs regulate a number of common gene programs involved in skeletal muscle metabolism [197] but whether there are sex-specific differences in the ER responses to exercise has not been established. Sex-specific differences in substrate utilisation could also impact on histone acetylation responses, with females having a greater reliance on fatty acid oxidation at any particular submaximal power output. It has recently emerged that fatty acids play an important role in providing the acetyl-CoA required for acetylation reactions, with up to 90% of acetylation at specific histone acetylation marks being from carbon derived from fatty acids [198]. The greater reliance on fatty acid oxidation for ATP generation in females could suggest that the availability of free acetyl-CoA for

acetylation reactions is reduced, which in turn would impact gene expression responses. These mechanisms have not yet been investigated in well-controlled studies allowing the analysis of sex-specific responses.

Although histone acetylation is important for transcriptional initiation, a plethora of other histone post-translational modifications play a role in transcriptional responses. Beyond acetylation, there are no studies that have examined histone modifications in response to exercise, yet alone in a sex-specific manner. Understanding the histone modifications evoked by exercise will be important for deciphering sex-specific responses to exercise, as well as understanding interactions with other epigenetic process such as DNA methylation and how they together impact the adaptive response to exercise.

2.5.4.3 MicroRNAs

MiRNAs are derived from double-stranded hairpin loops of about 70 nucleotides, which are cleaved by Dicer protein into single strands of ~22 nucleotides. These small, noncoding RNAs inhibit the translation of specific mRNA targets by either inducing degradation of the mRNA transcript or physically inhibiting the access of translational machinery to the mRNA, ultimately decreasing the expression levels of the targeted mRNA [199-201]. The network dynamics of miRNAs is complex since many miRNAs may work together to repress a certain gene and many genes can be regulated by the same miRNA [202]. Several reviews have summarised the effects of exercise on miRNA expression [169, 203-205]. Briefly, specific miRNAs are upregulated and downregulated with both acute and chronic exercise in humans [206-208]. Russell *et al*. reported an increase in miR-1, -133a, -133b and -181a, as well as key components of the miRNA biogenesis pathways and a decrease in miR-9, -23a, -23b and -31 three hours after a single bout of high-intensity interval endurance exercise in human males [207]. Additionally, they found that after 10 days of training, miR-1 and -29b

were increased, while miR-31 remained decreased (as in the acute testing) [207]. Using reporter assays, this study validated some of the associations of the miRNAs with predicted targets HDAC4 and nuclear respiratory factor 1 (NRF1), both of which are regulated during exercise and are thought to contribute to exercise adaptive responses [192, 209]. Acute and short-term exercise regulate several miRNAs that are potentially involved in the regulation of skeletal muscle regeneration, gene transcription, and mitochondrial biogenesis, suggesting that miRNAs play a role in exercise adaptation. However, no studies investigated the potential differences between males and females in skeletal muscle miRNA activity following exercise. One study investigated the differences in muscle-specific miRNAs, termed myomiR (miR), between males and females at rest [157]. They found sex differences in two (miR-133a and b) of four miRNAs (miR-1, miR-133a, miR-133b, and miR−206) that are crucial for the regulation of skeletal muscle development and function and are known to change following exercise [157]. One study found sex differences in miRNAs in saliva that changed following one bout of long distance running [162]. While those sex-differentially expressed miRNAs are inferred to be involved in fatty acid biosynthesis pathways, targets were not validated. Further research is therefore needed to determine whether miRNA regulation of gene expression contributing to exercise adaptation differs between males and females.

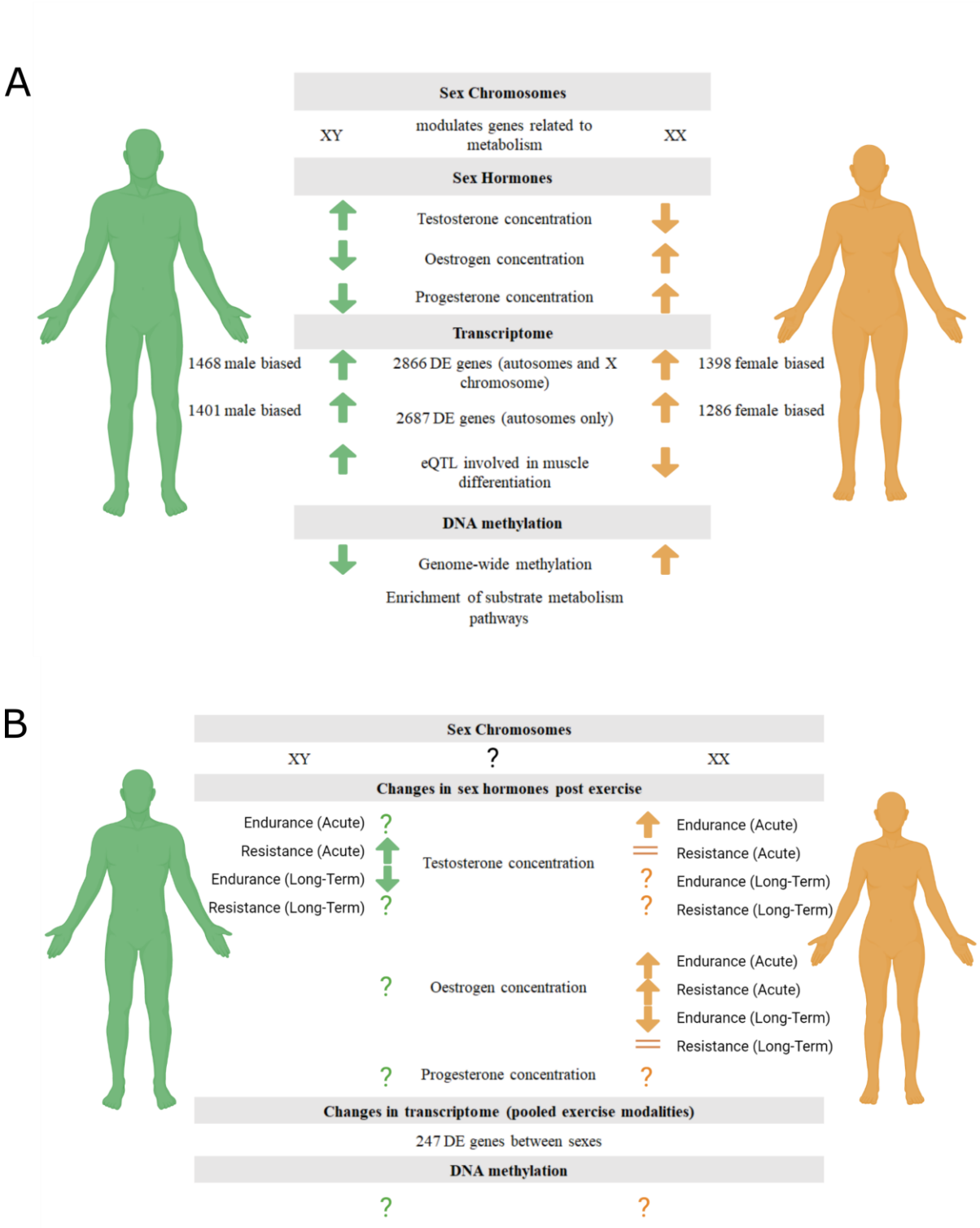


Figure 2.2 Molecular mechanisms underlying sex differences at rest and following training.

(**A**) Sex-specific molecular mechanisms at rest. Direction of arrows indicate which sex generally presents with the higher level. Males denoted by green; females denoted by orange, DE; Differentially Expressed genes; eQTL; expression quantitative trait loci. Number of DE genes reflects skeletal muscle tissue from Genotype-Tissue Expression (GTEx) database (accessed from GTEx portal on 08/26/2020; n = 803; version 8). (**B**) Response of molecular mechanisms after endurance or resistance training. The arrows indicate the direction of *change* following endurance and strength/power training in each sex; question mark indicates that the response is unknown as it has not been thoroughly studied. Males denoted by green; females denoted by orange.

2.5.5 Genomics: the influence of genetic variants

The heritability of VO_2max, a strong indicator of endurance performance, is estimated to be between ~ 22–57%, meaning that ~22-57% of the variability in VO_2max observed in a population can be attributed to genetic variation [210]. Completion of the sequencing of the human genome in 2001 [211] paved the way for DNA sequencing for identification of specific genetic variants correlated with a particular phenotype (e.g., exercise responses/performance outcomes). Since then, various genetic variants that may provide an advantage in exercise performance have been identified; for a detailed review see references [212, 213]. Identifying such genetic variants and their downstream modes of action provide new insight to exercise adaptations. However, it is important to note that athletic ability is a complex trait that is influenced by many aspects and genetic variants, thus making it challenging to identify variants with large effect sizes. Furthermore, common variants typically have small influences on a given trait. A thorough review [214], and a recent commentary [212] on sports genetics highlights the need for larger sample sizes, and both ethnicity-specific and sex-specific analyses to confirm effect sizes of common variants.

To date, two gene variants associated with exercise phenotypes have been substantially replicated in multiple cohorts: alpha-actin-3 (*ACTN3 R577X*) and angiotensin converting enzyme (*ACE I/D*). Both variants were discovered using the candidate gene approach, which is used to find correlations between pre-specified single nucleotide polymorphisms (SNPs) and phenotypes. Most studies found associations between exercise response/performance and the *ACTN3* and *ACE I/D* variants, however, some studies have not. It has been hypothesised that some of the heterogeneity in results is due to sex differences as cohorts are often mixed-sex [215].

ACTN3 encodes the alpha-actin-3 protein that is expressed in the sarcomere of fast glycolytic type II fibres and is important for the generation of explosive power contractions. The substitution of an arginine (R) with a stop codon (X) at the 577 amino acid results in deficiency of the ACTN3 protein (*ACTN3* XX genotype). Most of the studies regarding the association between the *ACTN3* variant and performance report that the RR genotype or the R allele is associated with strength and muscle power [216-218]. Some studies reported sex differences in the genotype-phenotype association of the *R577X* variant [216, 219, 220], for example, Shang *et al.* studied the frequency of RR among endurance athletes and found lower frequency of the RR genotype in female endurance athletes compared with controls, but not in males (18.6% RR in female endurance athletes (n=250) *vs* 33.6% RR in control females (n=450)) [220]. These findings suggest that the X allele may have an advantageous effect on endurance performance in females but not in males. This sex difference could be explained by androgen hormones. Specifically, higher testosterone levels in males could contribute to performance improvements and reduce the relative influence of the *ACTN3* on muscle power, but this hypothesis has not been verified experimentally [216, 219, 221]. However, a study of 486 power athletes and 1,197 controls reported no sex differences in the association of *ACTN3* with performance [217]. Therefore, the *R577X* polymorphism may be contributing to exercise performance differently in males and females, but is not certain at this point.

ACE encodes the central component of the renin–angiotensin system (RAS), angiotensin converting enzyme, which is expressed in skeletal muscle, cardiac muscle, endothelial and kidney epithelial cells [222, 223]. ACE indirectly increases blood pressure by causing blood vessels to constrict. The deletion (termed “D allele”) or insertion (termed “I allele”) of a 287 base pair fragment at location 17q23.3 is a common variant of the gene. The I allele is generally associated with decreased ACE activity and better endurance performance, while the D allele is associated with increased ACE activity and improved muscle strength

[223-231]. One study on Japanese endurance track athletes found associations between the I allele and race distance in men but not in women (12.1% II in short distance male runners *vs* 49.3% II in long distance male runners, n=277 athletes) [229]. Also, one study found that the *D* allele is associated with hypertension in young males but not in young females (n=5014, randomly selected from population); specifically, in DD men the odds of having hypertension increased by a factor of 1.75 compared with II men. Interestingly, this difference was not observed between men and women aged 61-79. [232]. Therefore, the *ACE* I/D genotype-phenotype association may be sex-dependent, however as previously mentioned, many studies either have mixed-sex cohorts [225], only male cohorts [224], or do not have large enough sample sizes to detect potential sex-differences [225, 228].

Genome-wide association studies (GWAS) have emerged as a more effective way to determine the contribution of SNPs to a specific trait or phenotype. As opposed to the candidate-gene approach that is hypothesis-driven, GWAS are unbiased, hypothesis-free, and allow for discovery of novel SNPs and their associated phenotypes. Many exercise GWASs adjust their statistical model for sex [233-235]; however, some recent GWASs found sex differences in the contribution of particular SNPs to exercise phenotypes [236, 237]. Since females may have increased parasympathetic and decreased sympathetic control of heart rate in comparison to males, Ramirez *et al.* studied the association of genotype with the capacity of heart rate response during acute exercise. They identified two SNPs that showed sex-specific associations with the heart rate response to exercise in ~40,000 individuals. Specifically, one locus (*HLA-DRB5/HLA-DRB1*, rs9270779) was only significant in females (after exercise, every additional C allele at rs9270779 was associated with an additional HR change of 0.538 beats/min) while the other locus (*TAF2*, rs60717250) was only significant in males (after exercise, every additional C allele at rs60717250 was associated with an additional HR change of 0.486 beats/min) [237]. However, it is important to note that statistical analyses between the

sexes was not performed, in other words, although significance at a given locus was reached in one sex and not the other, it does not mean that there was statistical significance between the sexes. Another large-scale GWAS (n=195,180) determined the association of 16 SNPs with grip-strength and found no sex differences in individual SNP association with the trait; however, this study found a stronger association between the 16 SNP genetic score and grip strength in males than females (in males every unit increase in genetic score was associated with a 0.2 kg increase in grip strength while in women the increase was only 0.13 kg) [236]. Therefore, it is particularly important to determine SNP contributions to exercise phenotype in a sex-specific manner.

A recent and comprehensive review on the role of sex in genomics of human complex traits brings up important aspects to be taken into consideration in sex-specific genomics. The review proposes three models/mechanisms that contribute to the observed phenotypic sex differences (in human complex traits, specifically epidemiological studies). The first model states that differences in heritability (which SNPs and their effect sizes) contribute to the observed sex differences, however, heritability studies estimate that *only* <5% of the genetic basis of complex traits differ between males and females. The second model states that sex differences in the sex chromosomes have some associations with disease, but alone are unlikely to explain a large proportion of the phenotypic sex differences. Finally, the third model states that sex differences in gene-by-environment interactions are indeed common and are more likely to contribute to the observed sex differences in complex traits [25]. As previously stated, exercise-related phenotypes are complex traits, therefore focusing on the gene-by-environment, or epigenetic, contribution to sex differences will be important for understanding the underlying mechanisms of exercise-related sex differences.

2.6 Conclusions

In humans, there are sex differences in exercise and training responses (**Figure 2.1**). Such differences are underpinned by baseline differences in muscle strength, oxygen consumption, fibre type physiology, and exercise economy. The molecular mechanisms underlying these contributing factors is complex and include the sex chromosome complement, the hormonal milieu, genetics, and epigenetics. To further elucidate molecular adaptations to exercise training, future exercise studies should include both male and female participants [17, 238]. Future studies should not only adjust for sex as a covariate in their statistical analysis, but also be carefully designed to account for factors unique to females, such as the menstrual cycle, contraception or menopause [55, 239], and/or be statistically powered enough to allow for hormone level moderation [55]. Furthermore, conducting comprehensive meta-analyses combining several human studies [6], which generally have small sample sizes, would increase statistical power and shed light on sex differences related to exercise. Skeletal muscle adaptations to exercise training are fibre-type specific [240], and therefore investigating sex differences at the single fibre, as well as single-cell (non-muscle cells such as endothelial cells), levels will provide a deeper understanding of exercise-related sex-differences. Our group is currently conducting the Gene SMART (Skeletal Muscle Adaptive Response to Training) study [241], which aims to elucidate sex differences in response to exercise training by integrating multiple –OMIC layers in a large cohort of males and females. Taking action now will pave the way for a better understanding of the health-promoting molecular changes

induced by physical activity and allow the interpretation of previous and future research through a sex-specific lens.

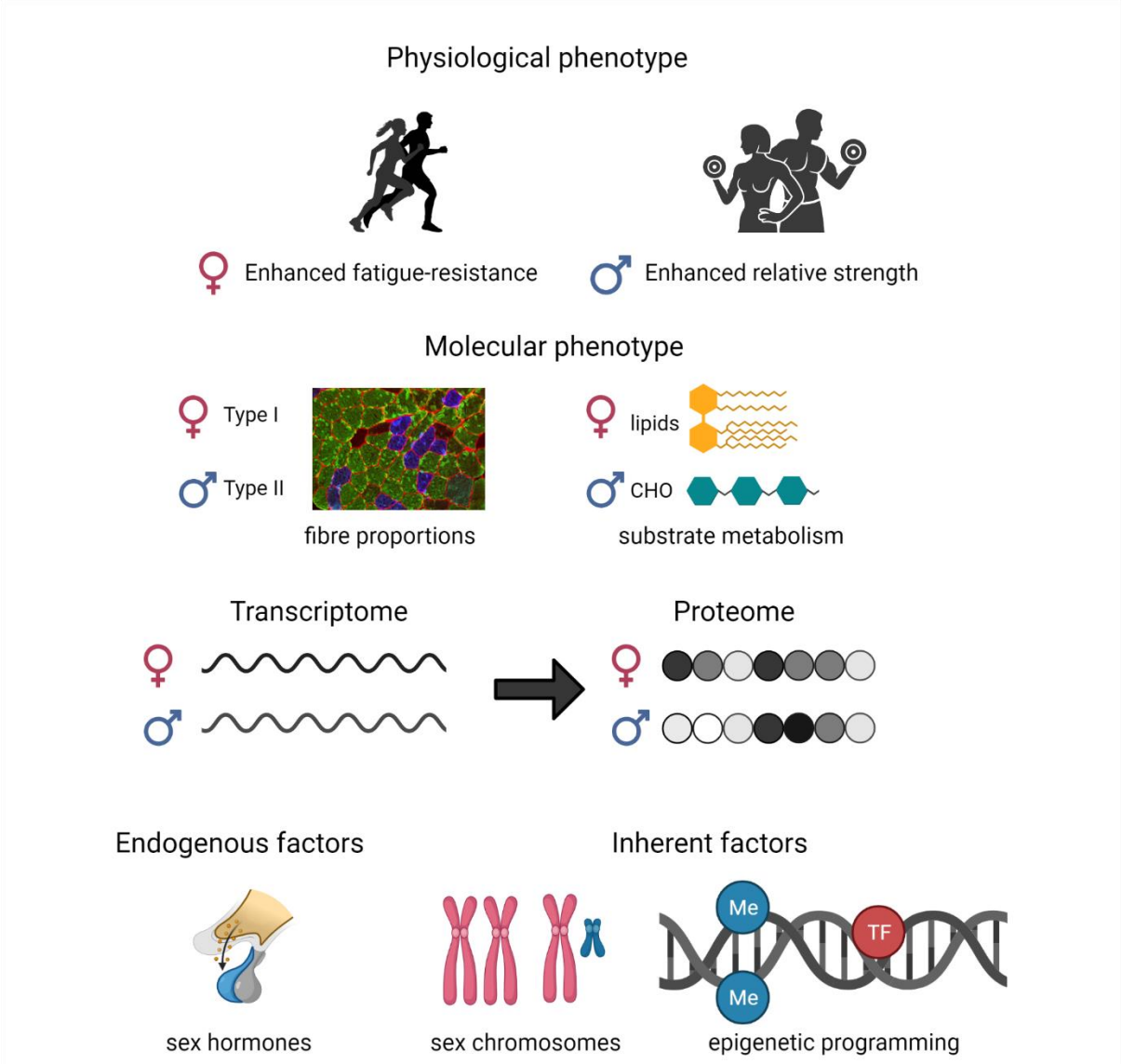


Figure 2.3 Schematic of exercise-related phenotype and molecular sex differences

Females exhibit enhanced fatigue-resistance while males have enhanced muscular strength; females tend to have higher proportions of type I fibres, and tend to oxidize lipids more than carbohydrates during endurance exercise; skeletal muscle transcriptomes differ at baseline and in response to training between the sexes, which likely lead to proteomic sex differences; all of the discussed sex differences arise from a combination of inherent factors such as differences in sex hormone exposure, sex chromosome complement, and epigenetic programming (such as DNA methylation (Me) and transcription factors (TF)).

Chapter 3 : Skeletal muscle methylome and transcriptome integration reveals profound sex differences related to muscle function and substrate metabolism

VICTORIA UNIVERSITY

THE NEW WAY TO DO UNI

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS

This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

Title of Paper/Journal/Book:	Skeletal muscle methylome and transcriptome integration reveals profound sex differences related to muscle function and substrate metabolism. Submitted; under review.
Surname:	Landen
First name:	Shanie
Institute:	Institute for Health and Sport
Candidate's Contribution (%):	68
Accepted and in press:	x
Published:	
Status:	
Date:	
Date:	2021

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

Shanie Landen Digitally signed by Shanie Date: 2021.05.08 16:26:55 +10'00'	2021
Signature	**Date**

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:

1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

PO Box 14428, Melbourne
Vic 8001, Australia
+61 3 9919 6100

3. There are no other authors of the publication according to these criteria;
4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and
5. The original data will be held for at least five years from the date indicated below and is stored at the following **location(s)**:

Name(s) of Co-Author(s)	Contribution (%)	Nature of Contribution	Signature	Date
Javier Alvarez Romero	3%	Data collection	[redacted]	08/08/2021
Danielle Hiam	5%	Data collection, writing and editing of manuscript		08/08/2021
Macsue Jacques	3%	Data collection		08/08/2021
Séverine Lamon	2%	Editing of manuscript		08/08/2021
Nir Eynon	5%	Editing of manuscript and final approval		08/08/2021
Sarah Voisin	10%	Guided student through statistical analyses and editing of manuscript		08/08/2021
Lyn R Griffiths	1%	Support in data collection		08/08/2021
Kevin J Ashton	1%	Editing of manuscript		08/08/2021
Larisa M Haupt	1%	Data collection and editing of manuscript		08/08/2021
Nicholas R Harvey	1%	Data collection and editing of manuscript		08/08/2021

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

3.1 Introduction

Sex differences are evident in nearly all complex traits. Various diseases, including but not limited to cancer, muscular dystrophy, and COVID-19 [25, 242], display sex differences in prevalence, onset, progression, or severity. To improve treatment for such diseases, it is crucial to uncover the molecular basis for the sex differences and their consequences on organ function. Sexually differentiated traits and phenotypes stem from a combination of factors, including genetics (gene variants-by-sex interactions [243], XY chromosome complements [9, 244-246], genomic imprinting [247]), the hormonal milieu [104, 248], and gene regulation [11], with the latter likely contributing the most [25].

Recently, a large-scale study from the Genotype-Tissue Expression (GTEx) consortium unravelled mRNA expression differences between the sexes that are not driven by sex chromosomes, across all tissues. Skeletal muscle was particularly divergent between the sexes, as gene expression profiles in this tissue could predict sex with high specificity ≥ 90%, and sensitivity ≥ 98% [9]. These transcriptomic differences underpin the numerous physiological differences in skeletal muscle between males and females, such as differences in substrate metabolism [2, 12, 142]. For example, females oxidise more lipids and less carbohydrates and amino acids during endurance exercise, and albeit depending on training status, tend to have a higher proportion of type I (slow-twitch) muscle fibres [249], all of which inherently contribute to enhanced fatigue-resistance in female skeletal muscle [250]. As such, females exhibit higher mRNA and protein levels of lipid oxidation-related genes than males [2]. Interestingly, the top gene set corresponding to sex-biased genes in the GTEx study corresponded to targets of the epigenetic writer polycomb repressive complex 2 (PRC2) and its associated epigenetic mark (H3K27me3). This suggests that the sex-specific deposition of epigenetic marks may be the source of sex differences in gene expression.

Epigenetics is a system of gene regulation that influences gene expression and is modulated by the genetic sequence and environmental stimuli. DNA methylation is currently the best-characterised epigenetic modification, and has been shown to differ between males and females in various tissues, such as pancreatic islets [251], blood [158, 252], and more recently in cultured myoblasts and myotubes [181]. While there is ample evidence for transcriptomic sex differences in skeletal muscle [7-9, 11, 142, 143], it is unclear whether sex differences exist in the DNA methylome of skeletal muscle tissue, and to what extent. Epigenome-wide association studies (EWAS) are ideal for investigating the impact of sex on genome-wide DNA methylation when addressing both the basis and translational aspect of sex differences. Therefore, we performed a large-scale EWAS meta-analysis to explore sex differences in the DNA methylome of human skeletal muscle tissue, using three datasets from our own laboratory and open-access databases (n = 369 individuals; 217 males, 152 females). We established a list of robust DNA methylation (CpG) sites and regions showing DNA methylation differences between males and females, and explored their genomic context. We then integrated them with sex-biased gene expression from the GTEx, and inferred the potential downstream effects on skeletal muscle function. Lastly, we confirmed our findings with transcriptomic data from one cohort used in the meta-analysis and targeted qPCR (*FOXO3A*, *ALDH1A1*, and *GGT7*) from another cohort.

3.2 Results

3.2.1 Males show profound genome-wide autosomal hypomethylation compared with females in human skeletal muscle

The DNA methylation meta-analysis was conducted on 369 individuals from three datasets (217 males, 152 females). We focused exclusively on the 22 autosomes to eliminate the confounding effect of sex differences in the sex chromosome complement where X-

chromosome inactivation takes place exclusively in females. All of the Gene SMART cohort individuals were apparently healthy, while the FUSION cohort individuals were either healthy or diagnosed with type 2 diabetes mellitus, and the GSE38291 cohort individuals included monozygotic twins discordant for type 2 diabetes mellitus (**Table 3.1**).

FUSION	**Females**, N = 120[1]	**Males**, N = 162[1]	**p-value**[2]
Age (years)	61 (8)	59 (8)	0.12
Health			0.057
Healthy	95 (79%)	111 (69%)	
T2D	25 (21%)	51 (31%)	
Gene SMART	**Females**, N = 20[1]	**Males**, N = 45[1]	**p-value**[2]
Age (years)	35 (7)	32 (8)	0.10
Health			
Healthy	20 (100%)	45 (100%)	
GSE38291	**Females**, N = 12[1]	**Males**, N = 10[1]	**p-value**[2]
Age (years)	66 (9)	70 (4)	0.15
Health			
Healthy	6 (50%)	5 (50%)	
T2D	6 (50%)	5 (50%)	

[1] Mean (SD); n (%)

[2] Welch Two Sample t-test

Table 3.1 Characteristics of individuals in each data set included in the DNA methylation meta-analysis. Statistics shown for differences between males and females.

We found 56,813 differentially methylated positions (DMPs, single CpG sites) between males and females, spread across the 22 autosomes, at a stringent meta-analysis False Discovery Rate (FDR) < 0.005 (**Figure 3.1**, Supplementary table 3.2). Ninety-four percent of DMPs were hypomethylated in males compared with females (**Figure 3.1A**). On average, the magnitude of DNA methylation differences between males and females was +2.8% (hyper DMPs) and -3.5% (hypo DMPs), with the largest effect sizes reaching +15.2% and -35.7%. In each of the three cohorts, participants did not cluster according to sex when including the whole autosomal methylome, but they did cluster according to sex when only focusing on the 56,813

DMPs (**Figure 3.1B**), suggesting that sex explained a substantial amount of variance at the DMPs.

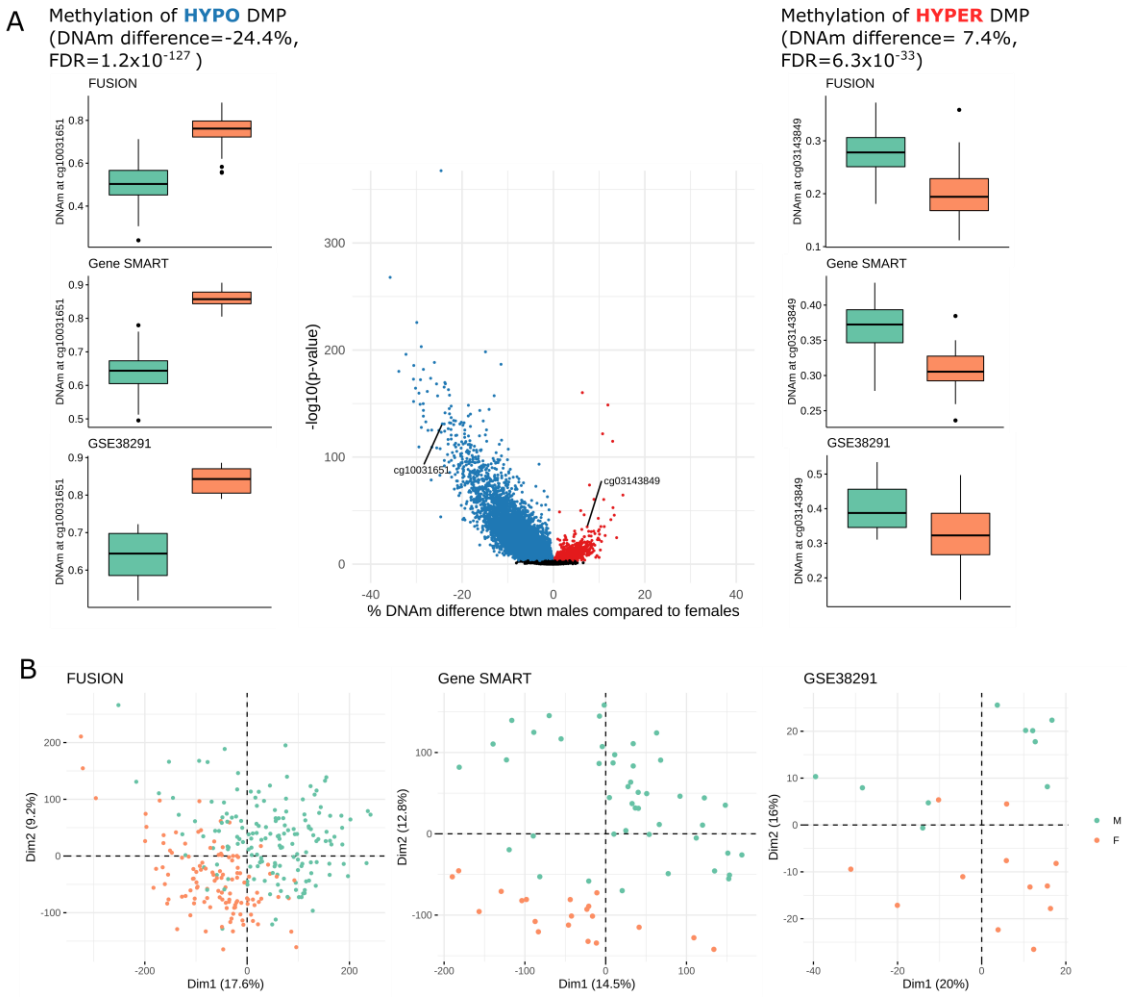


Figure 3.1 Differentially methylated positions (DMPs) with sex in skeletal muscle.

(A)Volcano plot of DNA methylation differences between males and females. Each point represents a tested CpG (633,645 in total) and those that appear in color are DMPs at a meta-analysis false discovery rate < 0.005; red DMPs are hypermethylated in males compared with females; blue DMPs are hypomethylated in males compared with females. The x-axis represents the amount of DNA methylation difference between the sexes and the y-axis represents statistical significance (higher = more significant). Two DMPs that were present in all three studies and showed the largest effect size are labeled with the respective CpG and boxplots of β-values from each study appear to the right (hyper DMP) and left (hypo DMP). (B) Principal component analysis plots of the methylation values at the DMPs; each point on the graph represents an individual; males denoted in green, females denoted in orange.

Each data set had a unique study design that required adjustment for factors known to affect DNA methylation, such as age [253] and type 2 diabetes (T2D) [254]. We adjusted each dataset for these factors, but noted that sex was associated with T2D in the FUSION dataset,

meaning that male participants from the FUSION cohort more commonly had T2D than females. Therefore, it is possible that the sex-related signal capture in this dataset was partially confounded by T2D. We repeated the meta-analysis excluding T2D participants from the FUSION cohort, but results remained unchanged (Supplementary figure 3.4). This confirms that our results are not confounded by T2D.

Since the effect of DNA methylation on gene expression depends on the genomic context, we explored the genomic locations of the DMPs to gain insights into their potential function [149]. We compared the distribution of hyper-, hypo-, and non-DMPs among the various chromatin states in human skeletal muscle using the Roadmap Epigenomics project [255]. DMPs were not randomly distributed in the chromatin states (χ^2 p-value < 2.2 x 10^{-16}, **Figure 3.2A**); specifically, hypo DMPs were enriched in enhancers and depleted in transcription start sites (Supplementary figure 3.1A), while hyper DMPs were not enriched in any chromatin states given their scarcity. It should be noted that the Roadmap Epigenomics project characterises both male and female skeletal muscle chromatin states regions, and there are 536 regions across 369 unique genes where male and female chromatin states differ (across many tissues including skeletal muscle) [256]. Therefore, we performed the chromatin state enrichment analysis on both the male and female chromatin state annotation in skeletal muscle, which yielded equivalent findings. We next determined whether the DNA methylation sex differences are enriched in regions in which the corresponding chromatin state displays sex differences. DMPs were indeed enriched in loci whose chromatin states differ between males and females: 38.7 % of DMPs vs. 32.4% of non-DMPs are in chromatin states that differ between males and females, which means that the odds of a DMP being located in a sex-differing chromatin state increased by a factor of 1.3 compared with a non-DMP. (OR = 0.76, 95% confidence interval = 0.75-0.77, Fisher test p-value < 2.2e-16) (**Figure 3.2B**). DMPs were

also enriched in CpG island shores and depleted in CpG islands (χ^2 p-value < 2.2e-16) (**Figure 3.2C**, Supplementary figure 3.1B).

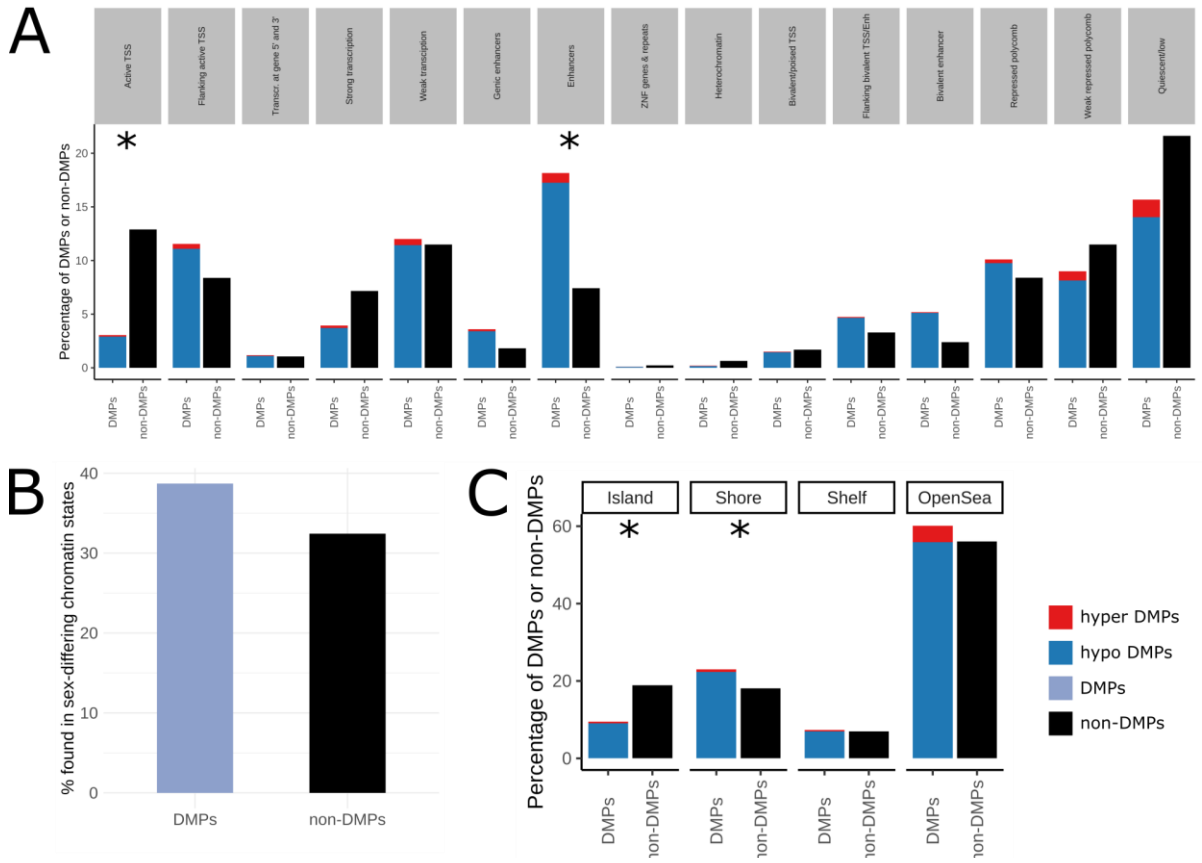


Figure 3.2 Genomic context of sex-differentially methylated positions.

(**A**) Distribution of hyper/hypo DMPs and non-DMPs with respect to chromatin states (male skeletal muscle annotation). Blue is hypomethylated in males and red is hypermethylated in males. Red and blue add up to all of the sex-DMPs. Black denotes the rest of the CpG sites from the analysis which are not DMPs. Asterisks represent a greater contribution to the significant relationship between DMP status and chromatin state (Supplementary figure 3.1A). (**B**) Distribution of sex-DMPs and non-DMPs at loci whose chromatin states differ between male and female skeletal muscle. Purple denotes all DMPs (hypo and hyper combined) and black denotes non-DMPs. (**C**) Distribution of sex-DMPs and non-DMPs in relation to CpG islands. Asterisks represent a greater contribution to the significant relationship between DMP status and CpG island location (Supplementary figure 3.1B).

Differentially methylated genes (DMGs) were determined by identifying differentially methylated regions (DMRs), as DMRs remove spatial redundancy (CpG sites ~500 bp apart are typically highly correlated [257]), and may provide more robust and functionally important information than DMPs [258, 259]. We identified 10,240 DMRs (Stouffer, harmonic mean of the individual component FDRs (HMFDR), and Fisher p-value < 0.005). These DMRs were

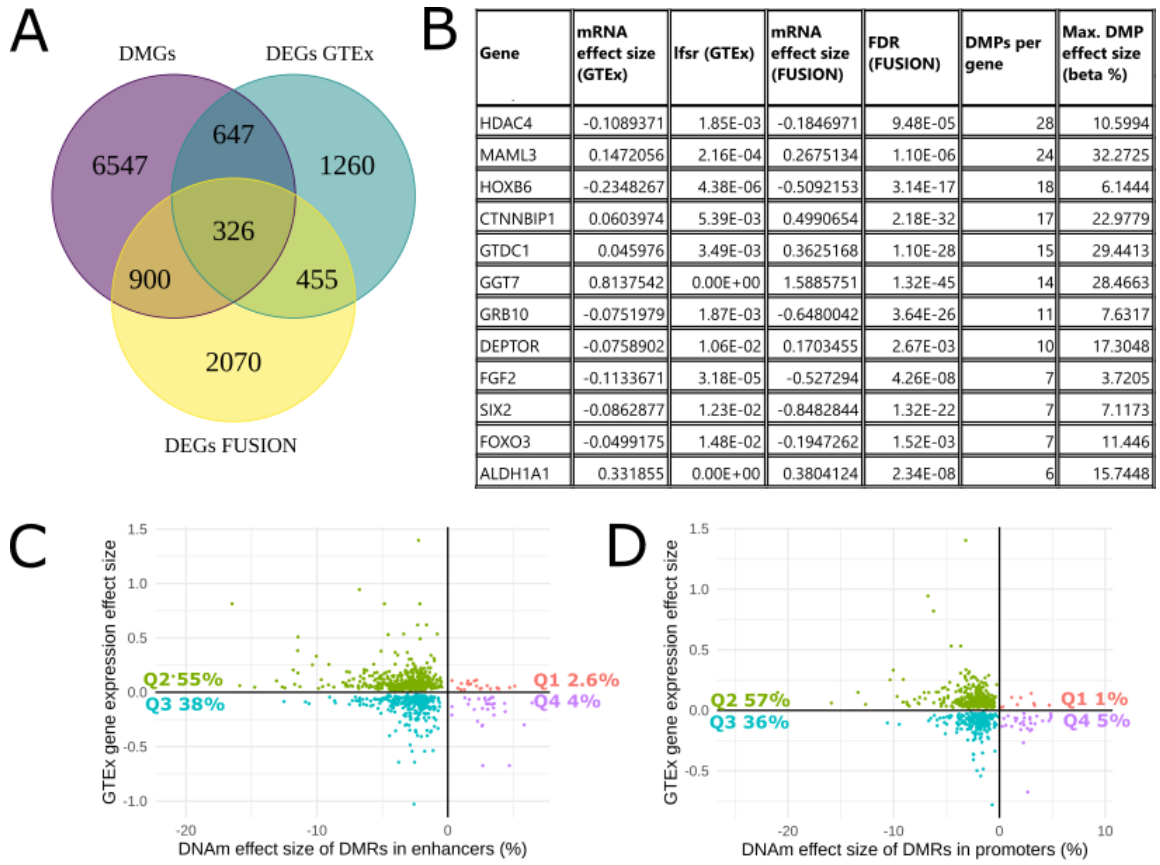
annotated to 8,420 unique autosomal genes (including non-coding genes) (Supplementary table 3.3).

3.2.2 Genes with sex-biased methylation exhibit sex-biased DNA methylation in human skeletal muscle

To gain insights into the potential downstream effects of sex-biased DNA methylation on gene expression, we integrated results from the EWAS meta-analysis of sex with genes whose mRNA expression levels are known to differ between males and females. We used version 8 of the Genotype-Tissue Expression (GTEx) database which contains 803 RNA-sequencing profiles in human skeletal muscle (n = 543 males and n = 260 females). There were 2,689 sex-differentially expressed genes (DEGs) on the autosomes in skeletal muscle (accessed from GTEx portal on 08/26/2020). Of the 2,689 DEGs, 973 (~36%) were in common with DMGs from our cohorts (**Figure 3.3,** Supplementary table 3.2), including the gene Gamma-Glutamyltransferase 7 (*GGT7*) (**Figure 3.5**). We confirmed an enrichment of DMRs across sex-biased genes (hypergeometric test p-value = 4.6e-13), suggesting that the overlap between sex-differentially methylated genes and sex-differentially expressed genes is larger than what would be expected by chance alone. To gain insight on the relationship between DNA methylation and gene expression of sex-biased genes, we assessed the direction of correlation between DMRs that are annotated to either promoter (TssA and TssAFlnk) or enhancer (Enh and EnhG) regions and their given gene expression (**Figure 3.3C-D**). Sixty-two and 59 % of DMRs in promoter and enhancer regions, respectively, were inversely correlated with gene expression (from GTEx transcriptome data, similar results were yielded with the FUSION transcriptome data). The inverse correlation between DNA methylation at both promoter and enhancer regions with gene expression was more than would be expected to occur by random chance (10,000 random permutations; p-value <0.0001 and p-value = 0.0009, respectively; Supplementary figure 3.3).

3.2.3 Validation of GTEx sex-biased genes in the cohorts used for methylation analysis

We sought to confirm the sex-biased gene expression obtained from GTEx in a subset of the samples used for methylation analysis since the DMGs and DEGs analyses were obtained from different muscle groups (the DMGs of the current study are from the *vastus lateralis* while the GTEx DEGs are from the *gastrocnemius*). Although both are skeletal muscle tissue from the leg, there may be differences in muscle phenotypes in differing muscle groups [260]. Analysis of RNA sequencing data from the FUSION cohort revealed 3,751 autosomal genes with sex-biased expression (FDR < 0.005). The FDR threshold we chose for the FUSION gene expression data was more stringent than the GTEx local false sign rate threshold (lfsr < 0.05), yet, ~34% of the genes which were both DEGs in GTEx and DMGs were also DEGs in the FUSION cohort, totalling 326 genes (hereinto referred to as `overlapping genes`) (**Figure 3.3A**). Given that both the GTEx and FUSION cohorts include participants of relatively older ages, we sought to confirm the mRNA levels in the younger cohort in the analysis (the Gene SMART) for three genes that displayed sex differences at both the mRNA and DNA methylation levels (*GGT7, FOXO3, and ALDH1A1*) (**Figure 3.6,** Supplementary table 3.11, Supplementary table 3.12).



Gene	mRNA effect size (GTEx)	lfsr (GTEx)	mRNA effect size (FUSION)	FDR (FUSION)	DMPs per gene	Max. DMP effect size (beta %)
HDAC4	-0.1089371	1.85E-03	-0.1846971	9.48E-05	28	10.5994
MAML3	0.1472056	2.16E-04	0.2675134	1.10E-06	24	32.2725
HOXB6	-0.2348267	4.38E-06	-0.5092153	3.14E-17	18	6.1444
CTNNBIP1	0.0603974	5.39E-03	0.4990654	2.18E-32	17	22.9779
GTDC1	0.045976	3.49E-03	0.3625168	1.10E-28	15	29.4413
GGT7	0.8137542	0.00E+00	1.5885751	1.32E-45	14	28.4663
GRB10	-0.0751979	1.87E-03	-0.6480042	3.64E-26	11	7.6317
DEPTOR	-0.0758902	1.06E-02	0.1703455	2.67E-03	10	17.3048
FGF2	-0.1133671	3.18E-05	-0.527294	4.26E-08	7	3.7205
SIX2	-0.0862877	1.23E-02	-0.8482844	1.32E-22	7	7.1173
FOXO3	-0.0499175	1.48E-02	-0.1947262	1.52E-03	7	11.446
ALDH1A1	0.331855	0.00E+00	0.3804124	2.34E-08	6	15.7448

Figure 3.3 Integration of differentially methylated genes and differentially expressed genes.

(**A**) Venn diagram of the overlap between differentially methylated genes (DMGs; derived from DMRs), differentially expressed genes derived from GTEx (DEGs GTEx), and differentially expressed genes derived from FUSION (DEGs FUSION) between males and females. (**B**) Subset of 12 genes with consistently large effect sizes or of biological relevance to skeletal muscle. (**C**) Correlation between the effect sizes of DMRs in enhancer regions and the effect sizes of gene expression of the relative annotated gene (for GTEx sex-biased genes). Quadrant percentages indicate the percentage DMRs/DEGs that fall into each quadrant. (**D**) Correlation between the effect sizes of DMRs in promoter regions and the effect sizes of gene expression of the relative annotated gene (for GTEx sex-biased genes). Quadrant percentages indicate the percantage DMRs/DEGs that fall into each quadrant.

3.2.4 Gene set enrichment analysis of differentially methylated regions

We next performed Gene set enrichment analysis (GSEA) on the DMGs, as GSEA using epigenomic features may reveal distinct enriched pathways that may not display gene expression differences [11, 256]. We performed GSEA on both the DMRs and DMPs (**Figure 3.4**). GSEA on the DMRs revealed enrichment of several Gene Ontology (GO) terms, one Reactome pathway ("muscle contraction"), but no Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Supplementary table 3.10) (FDR < 0.005). However, GSEA on the DMPs

revealed enrichment across all three databases (Supplementary tables 3.5, 3.7, and 3.9). Most of the enriched GO terms are biological process (BP) terms, many of which relate to anatomical structure development as well as many muscle-related processes. Nine-hundred and twenty-five genes of the 1,407 genes involved in KEGG metabolic pathways were differentially methylated, representing many aspects of substrate metabolism (Supplementary figure 3.2), although the pathway was only significant when analysing the DMPs.

A

	GO			KEGG			Reactome		
	GO term	sig genes/total	FDR	KEGG pathway	sig genes/total	FDR	Reactome pathway	sig genes/total	FDR
DMR GSEA	anatomical structure morphogenesis	1110/2619	$1.5x10^{-9}$	none at FDR < 0.005			Muscle contraction	103/200	0.003
	anatomical structure development	2146/5656	$2.1x10^{-9}$						
	developmental process	2283/6103	$1.1x10^{-8}$						
	movement of cell or subcellular component	851/2061	$3.8x10^{-8}$						
	cell differentiation	1544/4007	$3.8x10^{-8}$						
	multicellular organism development	1965/5192	$4.1x10^{-8}$						
	locomotion	744/1795	$4.1x10^{-8}$						
	system development	1781/4657	$4.1x10^{-8}$						
	actin filament-based process	356/727	$5.9x10^{-8}$						
	multicellular organismal process	2551/7207	$7.4x10^{-8}$						
DMP GSEA	anatomical structure morphogenesis	1851/2619	$2.8x10^{-20}$	Metabolic pathways	925/1407	$1.0x10^{-7}$	Signaling by Receptor Tyrosine Kinases	345/455	$8.8x10^{-7}$
	cytoplasm	6654/10520	$1.2x10^{-18}$	Focal adhesion	160/193	$2.0x10^{-5}$	Extracellular matrix organization	225/292	$7.4x10^{-6}$
	anatomical structure development	3701/5656	$4.9x10^{-18}$	Rap1 signaling pathway	165/207	$1.2x10^{-4}$	Neuronal System	288/388	$8.4x10^{-5}$
	developmental process	3960/6103	$4.0x10^{-17}$	Calcium signaling pathway	175/227	$2.1x10^{-4}$	Muscle contraction	155/200	$1.2x10^{-4}$
	cell development	1489/2084	$4.0x10^{-17}$	MAPK signaling pathway	214/282	$2.4x10^{-4}$	Cardiac conduction	107/135	$2.4x10^{-3}$
	cell junction	1301/1781	$1.3x10^{-16}$	Hippo signaling pathway	124/153	$2.4x10^{-4}$	Transmission across Chemical Synapses	188/255	$4.9x10^{-3}$
	system development	3068/4657	$1.7x10^{-16}$	AMPK signaling pathway	96/117	$3.8x10^{-4}$			
	multicellular organism development	3396/5192	$2.4x10^{-16}$	Cushing syndrome	122/153	$3.8x10^{-4}$			
	regulation of signaling	2341/3484	$8.5x10^{-16}$	Pathways in cancer	361/507	$3.8x10^{-4}$			
	nervous system development	1600/2258	$4.5x10^{-15}$	Insulin signaling pathway	104/131	$5.0x10^{-4}$			

B

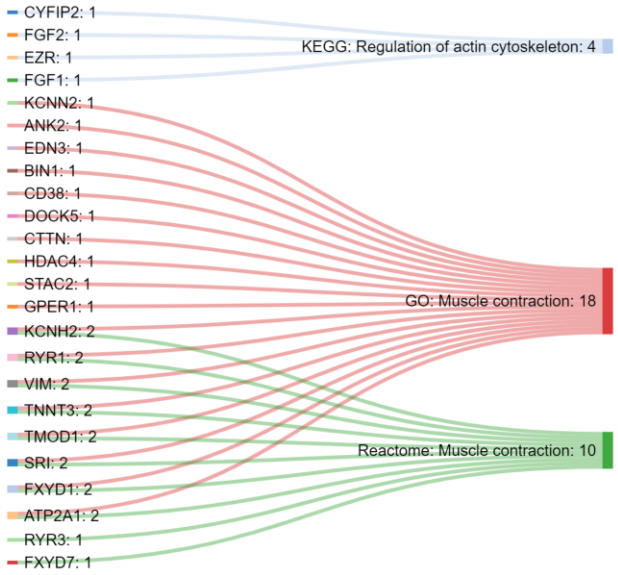


Figure 3.4 Gene set enrichment analysis of the differentially methylated genes.

(**A**) Selected enriched Gene Ontology (GO) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, and Reactome pathways from GSEA of DMRs and DMPs. (**B**) Sankey diagram of muscle contraction-related pathways across the three GSEA databases tested and genes within those pathways that were both differentially methylated and expressed (in GTEx and FUSION) between males and females. Numbers next to pathways denote the number of enriched genes in the pathway; numbers next to genes denote the number of pathways (from the ones displayed) that the gene belongs to.

3.2.5 DNA methylation and gene expression of GGT7, FOXO3 and ALDH1A1 consistently differ between males and females in human skeletal muscle

Three-hundred twenty-six genes exhibited differential methylation in the meta-analysis and differential expression among the GTEx and FUSION cohorts, termed `overlapping genes`. Of those genes, we tested three for gene expression levels, *GGT7,* Forkhead Box O3 (*FOXO3*), and Aldehyde Dehydrogenase 1 Family Member A1 (*ALDH1A1*), in the younger cohort included in the DNA methylation analysis (Gene SMART) given the effect that age has on skeletal muscle gene expression [261]. These three genes showed a large effect size in gene expression and DNA methylation, displayed moderate gene expression levels in skeletal muscle relative to other tissues, and/or contained numerous DMPs and DMRs (**Figure 3.6**, Supplementary table 3.12). The direction of sex-biased expression was consistent for *GGT7* and *ALDH1A1* across GTEx, FUSION, and Gene SMART cohorts (GTEx lfsr < $2.2e^{-16}$; FUSION FDR= $2.3e^{-8}$, Gene SMART p-value= 0.03), while the direction was opposite for *FOXO3* (FUSION and GTEx *FOXO3* expression lower in males, Gene SMART *FOXO3* expression higher in males (GTEx lfsr = 0.01; FUSION FDR= 0.001, Gene SMART p-value= 0.002)). As a specific example of the extent of sex differences across the different layers of analysis, *GGT7* displays male-biased expression in skeletal muscle (GTEx lfsr < $2.2e^{-16}$; FUSION FDR= $1.3e^{-45}$, Gene SMART p-value= 0.0003) as well as lower methylation in males at DMPs and DMRs annotated to *GGT7* (max DMR: Fisher p-value <0.00^{-15}, max beta value effect size=-28.5%, mean beta value effect size=-20.4%) (**Figure 3.5**).

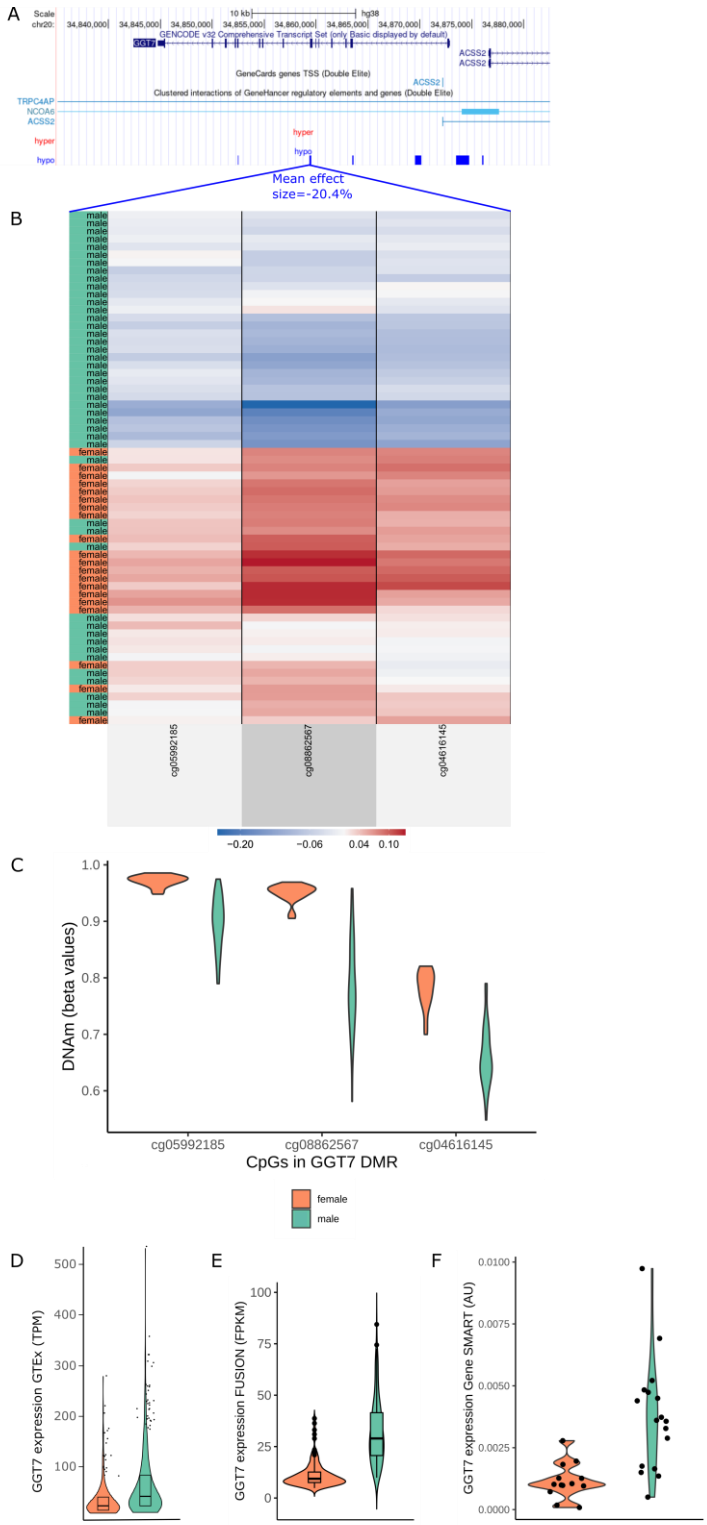


Figure 3.5 Differential DNA methylation and expression of GGT7 between males and females.

(**A**) UCSC gene track of GGT7. From top to bottom: base pair scale in black, GENCODE gene tracks transcript variants in blue, GeneHancer regulatory element annotations in light blue, hyper DMRs tracks in red, hypo DMRs tracks in blue. (**B**) Heatmap of the Gene SMART study (beta values adjusted for all confounders except sex) across the 3 CpGs included in the GGT7 hypo DMR selected in blue lines and labeled with mean DMR effect size (n=65). Each row represents an individual; green denotes males and orange denotes females; ordered by similarity to other individuals. Each column corresponds to a CpG in the DMR, ordered by genomic location

and corresponding to 5C. Blue denotes hypomethylation; tred denotes hypermethylation. (**C**) Distribution of DNA methylation (beta values) in males and females, for the three CpGs in the DMR, matching 5B (n = 65). (**D**) GGT7 RNAseq expression (TPM- transcripts per million) in males and females of the GTEx (adapted from GTEx portal, n = 803). (**E**) GGT7 RNAseq expression in the FUSION males and females (FPKM- fragments per kilobase of transcript per million) (n = 274). (**F**) GGT7 qPCR expression in a subset of Gene SMART males and females (Arbitrary Units; $2^{-\Delta Ct}$) (n = 25).

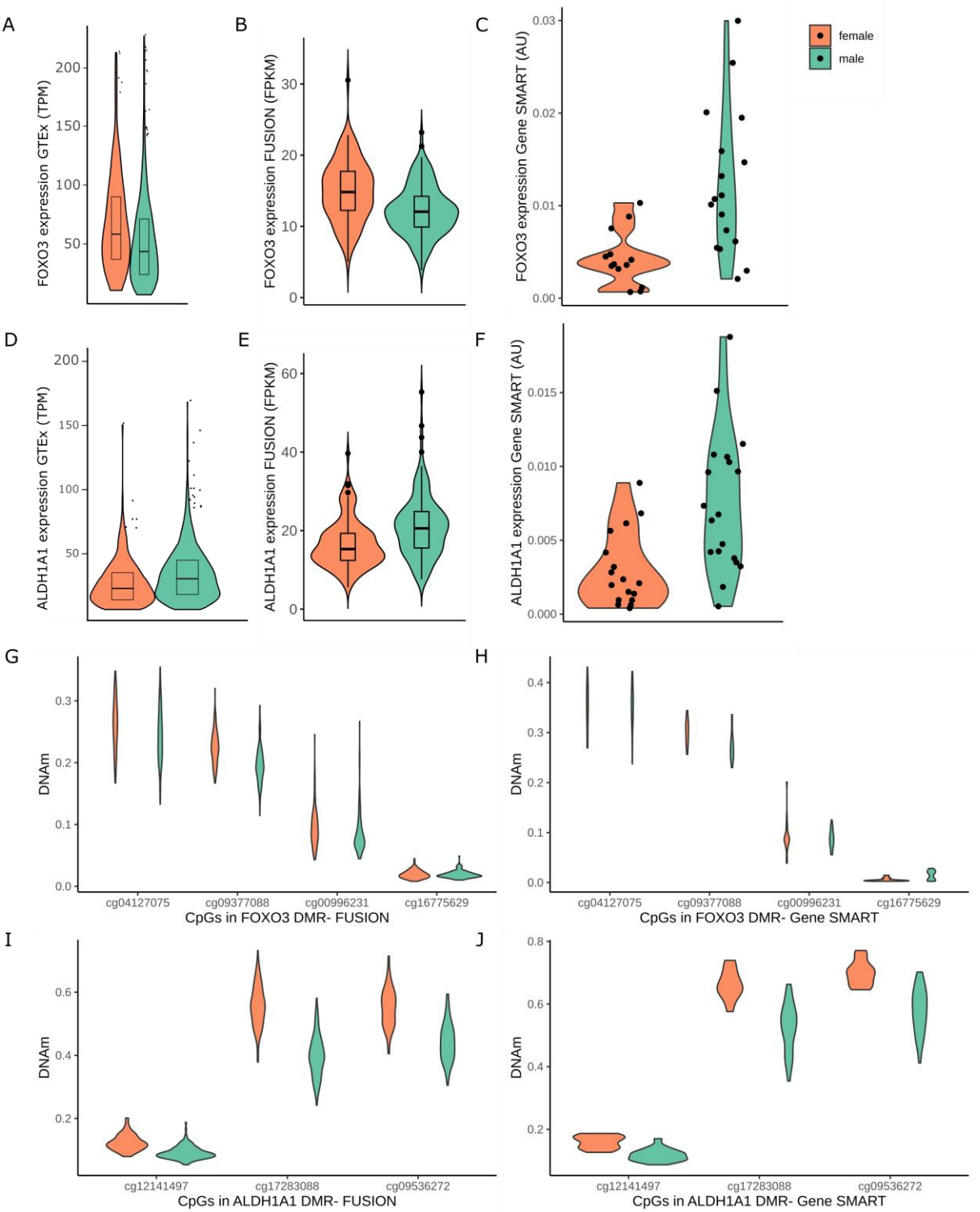


Figure 3.6 Gene expression and DNA methylation for FOXO3 and ALDH1A1 across cohorts.

Distribution of FOXO3 expression in males and females from (**A**) the GTEx (RNA-seq) (**B**) the FUSION cohort (RNA-seq) (**C**) the Gene SMART cohort (qPCR). Distribution of ALDH1A1 expression in males and females from (**D**) the GTEx (RNA-seq) (**E**) the FUSION cohort (RNA-seq) (**F**) the Gene SMART cohort (qPCR). Distribution of methylation of CpGs in FOXO3 DMR in males and females from (**G**) the FUSION cohort and (**H**) the Gene SMART cohort. Distribution of methylation of CpGs in ALDH1A1 DMR in males and females from (**I**) the FUSION cohort and (**J**) the Gene SMART cohort.

3.3 Discussion

We conducted a large-scale meta-analysis of DNA methylation differences between males and females in skeletal muscle, and integrated them with transcriptomic data. We revealed that males display profound genome-wide hypomethylation compared with females. We then showed that many sex-biased genes found in GTEx also exhibit sex-biased DNA methylation, which was partially confirmed in the FUSION cohort. We then assessed the gene expression (qPCR) levels of three genes with large DNA methylation and expression differences between the sexes across cohorts, and confirmed the higher gene expression in males of *GGT7* and *ALDH1A1*. Finally, we showed that the DMGs are overwhelmingly involved in muscle contraction, as well as other metabolic and anatomical structure-related pathways.

We identified 56,813 sex-differentially methylated autosomal sites in skeletal muscle, representing ~10% of the tested CpG sites. Similarly, there were 2,689 sex-differentially expressed autosomal genes in skeletal muscle as identified using GTEx cohort, representing ~13% of the expressed genes. In the present study, the overwhelming majority (94%) of the DMPs were hypomethylated in males. Interestingly, global autosomal hypomethylation in males has been observed in various other tissues [156], including blood [262, 263] and pancreatic islets [251]. There are a few possible explanations for the molecular mechanisms at the root of these epigenetic differences between the sexes (investigated in Chapter 4). Differences in cell type proportions between the sexes may partly explain our findings [263-265], as type I fibres are hypermethylated compared with type II fibers [266], and as females

tend to have a higher proportion of type I fibres than males [249]. Although not well-understood, the sex chromosome complement may also influence autosomal DNA methylation patterns. In cultured fibroblasts, the presence of sex-determining region Y (*SRY*) is associated with lower autosomal methylation levels [88, 89, 267]. Additionally, a higher number of the X chromosomes, in the absence of *SRY*, leads to increased methylation levels at a specific sex-differentially methylated autosomal region [267]. This could be attributed to allele dosage compensation, a female-specific process that silences one of the X chromosomes in a cell [92, 93]. Approximately one-third of genes 'escape' inactivation, remain transcriptionally active in XX cells, [93-95], and have been suggested to affect autosomal DNA methylation via their histone marks [267, 268]. Moreover, females with Turner syndrome (partially/fully missing one X) and monosomy X have lower global methylation than XX females, but higher than XY males [269, 270]. Finally, sex hormones may contribute to inherent autosomal sex-specific DNA methylation as has been shown in leukocytes [271], but this may only be apparent after taking cellular composition into account [272]. The effect of sex hormones on DNA methylation in skeletal muscle has yet to be explored.

The relationship between DNA methylation and gene expression is complex; DNA methylation at promoters, enhancers, and 1st exons is generally believed to enhance gene silencing, while DNA methylation at gene bodies can sometimes be associated with increased gene expression [149, 273-276]. Using a permutation test, we showed that DNA methylation differences between the sexes at promoters and enhancers were more often associated with lower gene expression than would be expected by chance alone. DNA methylation differences between the sexes were also particularly prominent in chromatin states that are known differ between males and females. This suggests that DNA methylation differences between males and females reflect alterations in chromatin activity, and differential epigenetic states and expression are likely functionally connected. In line with this, chromatin states that differ

between the sexes have been shown to be enriched for sex-biased genes across various tissues, including skeletal muscle [256]. However, it is not yet possible to assess whether the relationship reflects correlation or direct causality. There is still debate around whether epigenomic features drive regulatory processes or are merely a consequence of transcription factor binding [256]. A recent study analysing sex differences in regulatory networks in the GTEx database identified that many transcription factors (TF) have sex-biased targeting patterns [11]. Further supporting the effect of TF on sex-biased gene expression, another recent study also on the GTEx database found enrichment of TF binding sites in the promoters of sex-biased genes [9].

We identified 326 genes with consistent differential skeletal muscle DNA methylation and expression across 1,172 individuals altogether (369 individuals from three cohorts for DNA methylation and 1,077 individuals from two cohorts for gene expression). Although we utilized stringent Stouffer, Fisher, and HMFDR thresholds, we did not set an effect size threshold, which may be the reason we identified an overwhelming 8,420 sex-differentially methylated autosomal genes. Although we found profound global DNA hypomethylation in males, of the overlapping genes there were equivalent numbers of genes over- and under-expressed in males compared with females for both GTEx and FUSION. Indeed, hypermethylation is not always associated with decreased gene expression [277]. The substantial overlap between differentially methylated genes and differentially expressed genes highlights many genes that may be of interest for their roles in muscle-related processes. We focused on three of these genes that displayed a large DNA methylation difference between males and females, are highly expressed in skeletal muscle, or play a role in skeletal muscle function: *HDAC4* given its role in neurogenic muscle atrophy [278, 279] and in the response to exercise [192]; *DEPTOR* given its role in muscle glucose sensing which in turn augments insulin action [280]; *GRB10* given that it is imprinted and has been shown to change in

methylation with exercise/training [154]; *FOXO3* for its role in ageing, longevity, and regulating the cell cycle [281]; *ALDH1A1* for its role in aldehyde oxidation and because sex differences in skeletal muscle mRNA levels have been reported, suggesting that males might be able to metabolise aldehydes (i.e. alcohol) more efficiently than females [142]; and *GGT7* for its role in antioxidant activity [282]. Of these three genes (which were validated across GTEx, FUSION, and Gene SMART), FOXO3 and GGT7 have also been reported to exhibit differential methylation between male and female myoblasts as well as myotubes [181]. *GGT7* and *ALDH1A1* showed consistently higher expression levels in males while *FOXO3* showed opposite sex-biased expression in the young versus the old cohorts. *FOXO3* expression was lower in males in the older cohorts (GTEx and FUSION), and higher in males in the younger cohort (Gene SMART). Other studies have shown that males have higher *FOXO3* expression in young skeletal muscle [283] and that elderly females have higher skeletal muscle *FOXO3* expression than younger females [284]. While *FOXO3* skeletal muscle gene expression differs between males and females, it seems that the direction is opposite in young and old individuals, which emphasizes the caution that should be used when interpreting sex differences across a large age range of individuals. Interestingly, FOXO3 was hypomethylated in skeletal muscle with age in a recent study from our group [285]. The promoter, 1st exon, and gene body of *GGT7* were hypomethylated in males and males had higher *GGT7* expression. *GGT7* is highly expressed in skeletal muscle and metabolises glutathione, which is a ubiquitous "master antioxidant" that contributes to cellular homeostasis. Efficient glutathione synthesis and high levels of glutathione-dependent enzymes are characteristic features of healthy skeletal muscle and are also involved in muscle contraction regulation [286].

In conclusion, we showed that the DNA methylation of hundreds of genes differs between male and female human skeletal muscle. Integration of the DNA methylome and transcriptome, as well as gene expression validation, identify sex-specific genes associated

with muscle metabolism and function. Uncovering the molecular basis of sex differences across different tissues will aid in the characterization of muscle phenotypes in health and disease. The effects of upstream drivers on sex differences in the muscle methylome, such as transcription factors, the XY chromosomes, hormones, and cell type differences still need to be explored. Molecular mechanisms that display sex differences in skeletal muscle may help uncover novel targets for therapeutic interventions.

3.4 Methods

3.4.1 Datasets

We conducted a meta-analysis of three independent epigenome-wide association studies (EWAS) of sex including the Gene Skeletal Muscle Adaptive Response to Training **(**SMART) study from our lab [241], the Finland-United States Investigation of NIDDM Genetics (FUSION) study from the dbGAP repository (phs000867.v1.p1) [287], and the GSE38291 dataset from the Gene Expression Omnibus (GEO) platform [288]. Detailed participant characteristics, study design, muscle collection, data preprocessing, and data analysis specifications for each study are in Supplementary table 3.1. Briefly, all studies performed biopsies on the *vastus lateralis* muscle, all participants were of Caucasian descent (except one individual of mixed Caucasian/aboriginal decent), and included either healthy or healthy and T2D individuals aged 18-80 years. The Gene SMART study was approved by the Victoria University human ethics committee (HRE13-223) and written informed consent was obtained from each participant. NIH has approved our request [#96795-2] for the dataset general research use in the FUSION tissue biopsy study.

3.4.2 DNA Extraction and Methylation Method – Gene SMART study samples

Genomic DNA was extracted from the samples using the AllPrep DNA/RNA MiniKit (Qiagen, 80204) following the user manual guidelines. Global DNA methylation profiling was

generated with the Infinium MethylationEPIC BeadChip Kit (Queensland University of Technology and Diagenode, Austria). The first batch contained only males and were randomised for timepoint and age. The second batch contained males and females and samples were scrambled on the chips to ensure randomness when correcting for batch effect (i.e. old and young males and females included on each chip across all time points). The genome-wide DNA methylation pattern was analysed with the Infinium MethylationEPIC BeadChip array.

3.4.3 Bioinformatics and statistical analysis of DNA Methylation

3.5.3.1 Preprocessing

The pre-processing of DNA methylation data was performed according to the bioinformatics pipeline developed for the Bioconductor project [289]. Raw methylation data were pre-processed, filtered and normalised across samples. Probes that had a detection *p-v*alue of > 0.01, located on X and Y chromosomes or cross-hybridising, or related to a SNP frequent in European populations, were removed. It is important to note that the list of cross-hybridising probes was supplied manually [290] as the list supplied to the *ChAMP* package was outdated. Specifically, there are thousands of probes in the Illumina microarrays that cross-hybridise with the X-chromosome and may lead to false discovery of autosomal sex-associated DNA methylation [291]. The BMIQ algorithm was used to correct for the Infinium type I and type II probe bias. β-values were corrected for both batch and position in the batch using *ComBat* [292].

3.5.3.2 Statistical analysis

We adjusted each EWAS for bias and inflation using the empirical null distribution as implemented in *bacon* [293]. Inflation and bias in EWAS are caused by unmeasured technical and biological confounding, such as population substructure, batch effects, and cellular heterogeneity [294]. The inflation factor is higher when the expected number of true

associations is high; it is also greater for studies with higher statistical power [293]. The results were consistent with the inflation factors and biases reported in an EWAS in blood [293]. Results from the independent EWAS were combined using an inverse variance weighted meta-analysis with METAL [295]. We used METAL since it does not require all DNA methylation datasets to include every CpG site on the HumanMethylation arrays. For robustness, we only included CpGs present in at least 2 of the 3 cohorts (633,645 CpGs). We used a fixed effects (as opposed to random effects) meta-analysis, assuming one true effect size of sex on DNA methylation, which is shared by all the included studies. Nevertheless, Cochran's Q-test for heterogeneity was performed to test whether effect sizes were homogeneous between studies (a heterogeneity index [I2] > 50% reflects heterogeneity between studies).

To identify DMPs, we used linear models as implemented in the *limma* package in R [296], using the participants' ID as a blocking variable to account for the repeated measures design (for twin [GSE38291] and duplicate samples [Gene SMART], using DuplicateCorrelation). The main sources of variability in methylation varied depending on the cohort and were adjusted for in the linear model accordingly. For the Gene SMART study, we adjusted the linear models for age, batch (2017 vs 2019), sex, and time point (before and after four weeks of high-intensity interval training). For the FUSION study, we adjusted the linear models for age, sex, BMI, smoking status, and OGTT status. For the GSE38291 study, we adjusted the linear models for age, sex, and diabetes status. All results were adjusted for multiple testing using the Benjamini and Hochberg correction [297] and all CpGs showing an FDR < 0.005 were considered significant [298]. DMRs were identified using the *DMRcate* package [299]. DMRs with Stouffer, Fisher, and harmonic mean of the individual component FDRs (HMFDR) statistics < 0.005 were deemed significant. Effect sizes are reported as mean differences in DNA methylation (%) between the sexes.

Next, we integrated a comprehensive annotation of Illumina HumanMethylation arrays [300] with chromatin states from the Roadmap Epigenomics Project [255] and the latest GeneHancer information [301]. DMPs that were annotated to two differing chromatin states were removed for simplicity and because there were very few such DMPs. GSEA on KEGG and GO databases was performed on DMRs and DMPs using the *goregion* and *gometh* (*gsameth* for Reactome) functions in the *missMethyl* R package [302] [303].

3.4.4 Integration of DNA Methylation and Gene Expression

The Genotype-Tissue Expression (GTEx) Project sex-biased data was downloaded from the GTEx Portal on 08/26/2020 and filtered for skeletal muscle samples. The enrichment of DMG for GTEx DEGs was done by supplying the list of sex-biased genes to the *gsameth* function in the *missMethyl* R package [302, 303], which performs a hypergeometric test, taking into account biases due to the number of CpG sites per gene and the number of genes per probe on the EPIC array. Caution should be taken when interpreting the number of DMPs reported per DMG. The analysis for direction of correlation between DNA methylation and gene expression was performed by randomly shuffling DNA methylation effect sizes and performing 10,000 permutations to assess how often a negative correlation occurs. This analysis was performed for both GTEx and FUSION transcriptome data and yielded similar results; data presented reflect results from the integration of differential methylation with differential GTEx expression. Significance reported for GTEx sex-biased genes is represented as the local false sign rate (lfsr) which is analogous to FDR [304]. GTEx effect sizes are represented as mash posterior effect sizes [304], in which positive values indicate male-biased genes and negative values indicate female-biased genes. FUSION and Gene SMART gene expression significance statistics are represented as FDR and p-value, respectively, and effect sizes as fold changes for both cohorts.

3.4.5 Validation of top genes with qPCR – Gene SMART study samples

Skeletal muscle previously stored at –80°C was lysed with the RLT buffer Plus buffer (Qiagen) and beta-mercaptoethanol using the TissueLyser II (Qiagen, Australia). DNA was extracted using the AllPrep DNA/RNA Mini Kit following the manufacturer guidelines (Qiagen, Australia). RNA yield and purity were assessed using the spectrophotometer (NanoDrop One, Thermofisher). RNA was reverse transcribed to cDNA using a commercially available iScript Reverse Transcriptase supermix (cat #1708841) and C1000 Touch Thermal Cycler (Bio-Rad, Hercules, CA, USA). Complementary DNA samples were stored at −20°C until further analysis. Quantitative real-time PCR was performed using SYBR Green Supermix (Bio-Rad, Hercules, CA) and gene-specific primers (listed in Supplementary table 3.11). Primers were either adapted from existing literature or designed using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) to include all splice variants, and were purchased from Integrated DNA Technologies. Ten microliter reactions comprised of SYBR, and optimised concentrations of forward and reverse primers (Supplementary table 3.11 for primer conditions), nuclease free water and 8 ng of cDNA were run in triplicate using an automated pipetting system (epMotion M5073, Eppendorf, Hamburg, Germany), with no-template negative controls on a 384-well plate in a thermo-cycler (QuantStudio 12K Flex System, ThermoFisher Scientific, Australia). Gene expression was normalised to the geometric mean expression of the two most stable housekeeping genes, as determined by Ref finder, TATAA-box binding protein (TPB), and 18s rRNA, which did not differ between sexes (Supplementary table 3.11). Data are presented as the fold change in males compared to females, using $2^{-\Delta\Delta CT}$.

The dataset generated and analysed during the current study are available in the GEO repository, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE171140.

Chapter 4 : Biological factors contributing to sex differences in the DNA methylome of human skeletal muscle

4.1 Introduction

It is becoming increasingly evident that there are many sex differences that should be considered in medical and basic research' study design, data analyses, results and interpretation of findings. Nonetheless, potential sex differences are often not taken into consideration in human and animal studies [305]. Many scientific journals now require peer-reviewed papers to abide by the sex and gender equity in research (SAGER) guidelines [306], which outline the rationale and ways in which sex should be accounted for in research across several disciplines.

Several studies have identified that there are sex differences in the transcriptome across various human tissues, with skeletal muscle being among one of the tissues with the most sex-biased gene expression [8, 9, 11]. Epigenetic modifications are the main contributors to observed, sex-specific phenotypic differences in complex traits and diseases [25]. DNA methylation, the most studied form of epigenetic modification, displays sex differences in multiple tissues [181, 251, 252, 262] including, in our recent findings, in human skeletal muscle (Landen *et al.* under review; Chapter 3 in thesis). Specifically, we have identified 8,420 genes that exhibit DNA methylation differences between males and females in human skeletal muscle.

Potential biological drivers of autosomal DNA methylation sex differences include the X and Y chromosomes [88, 89, 307], fibre and cell type proportions [249, 263-265], sex hormones [271], and genetic variants [243]. Human skeletal muscle is comprised of three main fibre types, slow twitch type 1 fibres (oxidative), fast twitch type 2A fibres (intermediate oxidative and glycolytic), and fast twitch type 2X fibres (glycolytic). Besides the less frequently occurring hybrid fibres, each fibre type expresses a unique myosin heavy chain

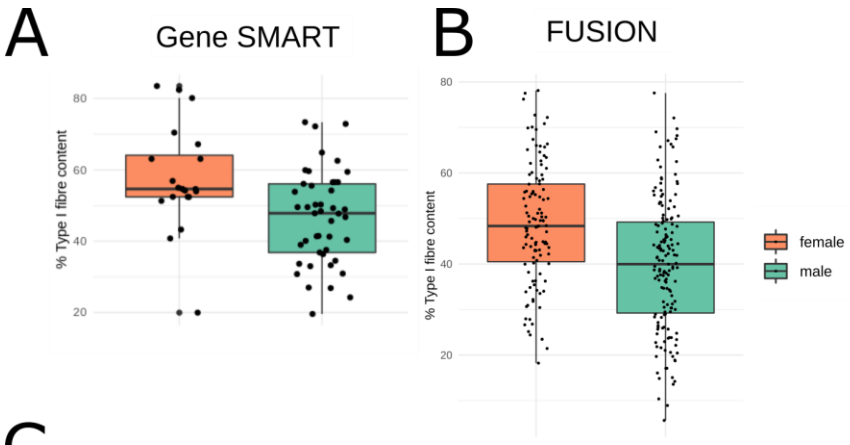
isoform [308], and as such, DNA methylation patterns vary between type 1 and 2A fibres [266]. Human fibre type proportions show significant variability between individuals. Moreover, females matched for age and training status show higher type 1 fibre proportions than males [249].

Males and females are exposed to dimorphic levels of sex steroid hormones throughout life, which may result in inherent cellular differences [309]. For instance, estrogen and androgen receptors can act as transcription factors (TFs), meaning that when bound by the corresponding hormone, they will enter the nucleus, bind specific sites on the DNA (transcription factor binding sites (TFBSs), and ultimately alter chromatin accessibility and gene expression [310]. Ovarian hormones estrogen and progesterone fluctuate throughout the menstrual cycle and effect cellular function accordingly [311].

The present study aimed to investigate the intrinsic biological factors (e.g. fibre type, circulating sex hormones, and TFBSs) driving sex-specific differences in DNA methylation in human skeletal muscle. We assessed type I fibre proportions in 65 healthy human skeletal muscle samples from the Gene SMART cohort (45 males and 20 females) and 274 healthy/T2D human skeletal muscle samples from the FUSION cohort (159 males and 115 females), and investigated whether type I (slow-twitch) fibre proportions were associated with DNA methylation at the loci exhibiting sex-biased DNA methylation. To address the intricate question of the effect of sex hormone levels on genome-wide autosomal DNA methylation, we assessed the impact of blood hormone levels on skeletal muscle DNA methylation patterns and whether cyclic ovarian hormone levels in the blood acutely affect DNA methylation patterns in females during the early follicular phase. Finally, we investigated whether sex- differentially methylated loci are enriched for known imprinted genes and TFBSs, specifically those that respond to hormone-related TFs.

4.2 Results

The investigation of intrinsic biological factors was performed on a subset of the cohorts used for the meta-analysis in Chapter 3 (data available for each dataset in Supplementary Table 4.1). Females in both the Gene SMART and FUSION cohorts had higher proportions of type I fibres (**Figure 4.1A-B**). Gene SMART females had higher estrogen and sex hormone-binding globulin (SHBG) levels, as well as lower free testosterone and testosterone levels than males (**Figure 4.1C,** Supplementary figure 4.2C-D).



C

	Females, N = 20[1]	**Males**, N = 44[1]	**p-value**[2]
Testosterone	1 (0)	20 (6)	<0.001
SHBG	73 (33)	33 (14)	<0.001
Testosterone (free)	9 (4)	492 (160)	<0.001
Estrogen	218 (215)	103 (24)	0.027

[1] Mean (SD)
[2] Welch Two Sample t-test

Figure 4.1 Fibre type distributions (Gene SMART and FUSION) and sex hormone levels (Gene SMART) of males and females.

Percent of type I fibres in the (**A**) Gene SMART cohort (20 females, 45 males) as determined by immunohistochemistry; p-value = 0.0001; 59% in females versus 47% in males. (**B**) FUSION cohort (115 females, 159 males) as determined by RNA-seq; p-value = 2.06 x 10-7; 48.6% in females versus 39.9% in males. (**C**) Concentrations of circulating testosterone (nmol/L), free testosterone (pmol/L), sex hormone-binding globulin (SHBG) (nmol/L), and estrogen (pmol/L) in males and females from the Gene SMART cohort. Includes 20 females and 44 males (missing data for one male included in the DNA methylation analysis) before the exercise training intervention. Values are represented as means with standard deviations in parenthesis; t-test performed in sex comparison.

4.2.1 Muscle fibre type proportions were associated with differential methylation at loci exhibiting sex-biased DNA methylation

Males typically show a greater proportion of type II muscle fibres compared with females [249], and type II fibres exhibit hypomethylation compared to type I fibres [266]. Therefore, we hypotheise that the observed DNA methylation sex differences, specifically the hypomethylation in males, may be a result of differing fibre type distributions between males and females. We first estimated type I fibre proportions in the Gene SMART cohort via immunohistochemistry (Supplementary figure 4.1B) and the FUSION cohort via RNA-seq (see "Methods"). In both the Gene SMART and FUSION cohorts, females had higher proportions of type I fibres than males (**Figure 4.1A-B**). We could not directly add fibre type proportions to the linear model as a covariate, since fibre type proportions are not a *confounder* (i.e. a factor that influences both sex and DNA methylation independently), but may be a direct *downstream effect* of sex, in turn affecting DNA methylation. Adding fibre type proportions in the model would therefore distort the association between sex and DNA methylation. To overcome this issue, we stratified the cohorts by sex, added fibre type proportions to the model as a covariate and identified DNA methylation patterns associated with fibre type proportions. We then meta-analysed the results to find CpGs robustly associated with fibre type proportions across both cohorts and all sexes (see "Methods"). We identified 16,275 CpGs associated with fibre type proportions (Supplementary figure 4.1A, Supplementary table 4.4). When restricting the analysis to the loci exhibiting sex-biased DNA methylation, 8,805 (15.5%) of those were associated with fibre type proportions (FDR < 0.005). Effect sizes ranged from -0.28% to +0.30% DNA methylation difference per % increase in type I fibre content (**Figure 4.2A,** Supplementary table 4.1).

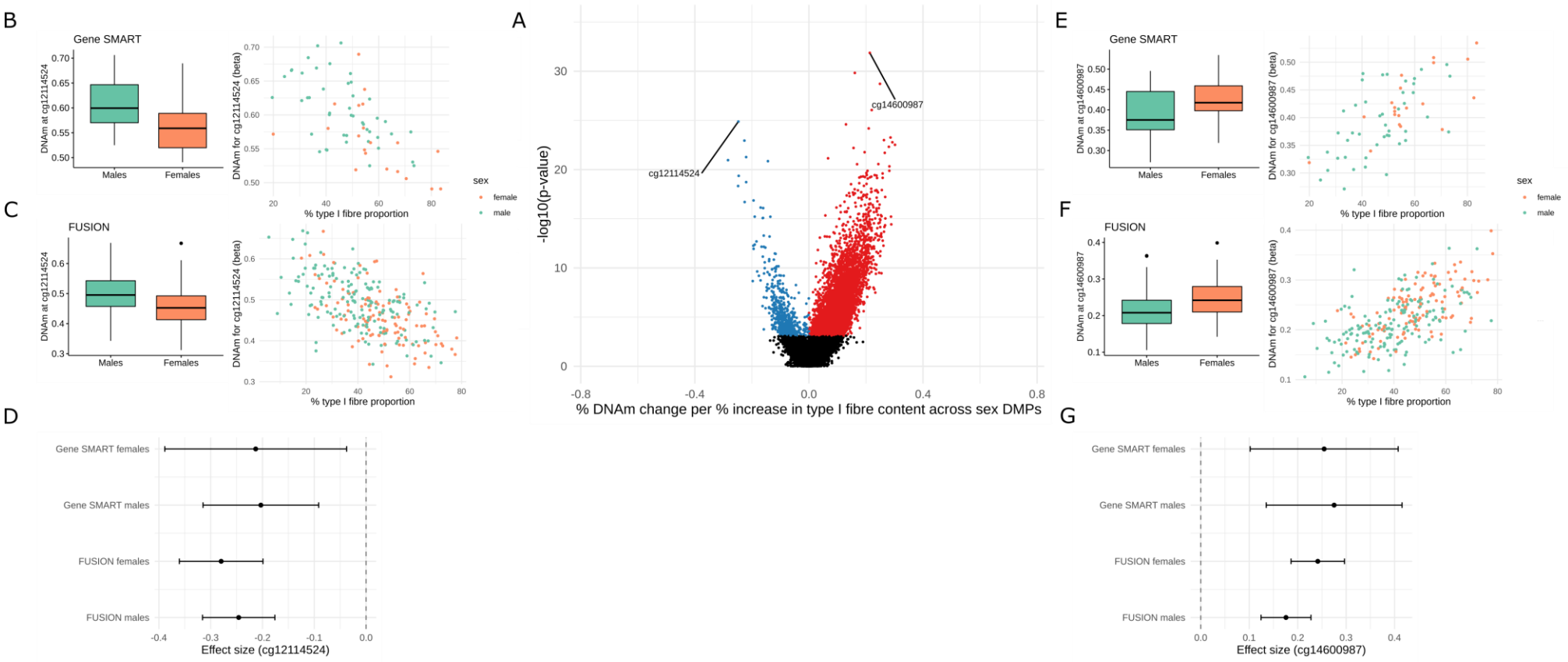


Figure 4.2 Fibre type-related DNA methylation loci across sex-biased DNA methylation loci.

(**A**) Meta-analysis effect size (x-axis) and meta-analysis significance (y-axis) for the 56,813 tested sex-biased CpGs. Hypomethylated (blue) and hypermethylated (red) point represent differentially methylated positions (DMPs) at false discovery rate (FDR) < 0.005. One hyper- and one hypo- DMP which showed the largest effect sizes are labeled with the respective CpG; with boxplots of β-values per sex and scatter plots of β-values relative to type I fibre proportion from the Gene SMART (**B,C**) and FUSION (**E,F**) cohorts. Females are represented in orange and males in green. (**D,G**) Forest plots for the given CpG, showing effect size and confidence intervals for each sex in each study.

4.2.2 Circulating sex hormone levels were not associated with methylation at loci exhibiting sex-biased DNA methylation

We then aimed to determine whether circulating sex hormone levels underlie the observed DNA methylation sex differences. We analysed estrogen (as estradiol, E2), testosterone (T), free testosterone (Free T), and sex-hormone binding globulin (SHBG) levels using mass-spectrometry (Free T derived from calculation) of blood serum in males and females in the Gene SMART cohort. Males and females significantly differed in all four hormone levels (**Figure 4.1C**, Supplementary figure 4.2C-D). To avoid collinearity with sex, we separated males and females for the association between sex-differential DNA methylation and circulating hormones. We assessed whether each of the four hormone levels was associated with DNA methylation across all of the CpGs and across the sex-DMPs in each sex by adjusting the linear model for a given hormone. In both males and females, circulating free testosterone, testosterone, estrogen, and SHBG levels were not highly associated with DNA methylation (less than five DMPs; FDR < 0.005) of neither all of the CpGs tested nor the sex-DMPs previously identified (Supplementary figure 4.3A-B).

4.2.3 Circulating ovarian hormones are not associated with differential methylation in female skeletal muscle

To limit the potentially confounding effect of fluctuating ovarian hormone levels on DNA methylation, female muscle biopsies were collected in the early follicular phase of the menstrual cycle (days 1-7) and blood serum were tested for follicle stimulating hormone (FSH), luteinising hormone (LH), and progesterone (as well as E2 as previously mentioned). Given the intricate fluctuations of the ovarian hormones throughout the menstrual cycle, a principal component analysis (PCA) was conducted using the four ovarian hormones and the first two principal components (PC) were included in the linear model (see “Methods”; Supplementary figure 4.3C). PC1 was not associated with differential methylation of any CpGs, while PC2

was associated with methylation of very few (8) CpGs (Supplementary figure 4.3A). This suggests that variations in ovarian hormone levels in the early follicular phase did not confound our results. Although, the study was designed to minimize the fluctuations of cyclic ovarian hormones (FSH, LH, estrogen, progesterone), and not to investigate the effects of cyclic hormones on genome-wide methylation.

4.2.4 Sex-differentially methylated loci are enriched for hormone-related transcription factor binding sites and not for imprinted genes

Given the role of transcription factors (TFs) in regulating chromatin accessibility and thus effecting downstream gene expression [312], as well as the recent studies identifying sex differences in TF targeting patterns [9, 11]; we next tested whether the meta-analysis DMPs were enriched for the experimentally validated binding sites (TFBSs) of 268 TFs from 518 different cell and tissue types [313, 314]. The DMPs were enriched for the binding sites of 41 TFs (p-value < 0.005, **Figure 4.3,** Supplementary table 4.2), including hormone-related TFs such as androgen (*AR*), estrogen (*ESR1*), and glucocorticoid (*NR3C1*) receptors.

DNA methylation is the most important epigenetic modification involved in genomic imprinting. Imprinting is the epigenetic marking of the parental genomes with respect to their parental origin in an allele-specific manner, and is required for proper gene regulation during development and in differentiated tissues [315]. Sex differences in expression levels of imprinted genes have been reported in mice embryo [316], suggesting that imprinted genes may display sex differences which could underlie sex-biased gene regulation. Therefore, we sought to investigate whether genomic imprinting contributes to the observed sex differences in DNA methylation. Sex-DMPs were not enriched for the 180 tested imprinted genes (imprinted genes across all tissues, accessed from GTEx portal; FDR = 0.4) [300].

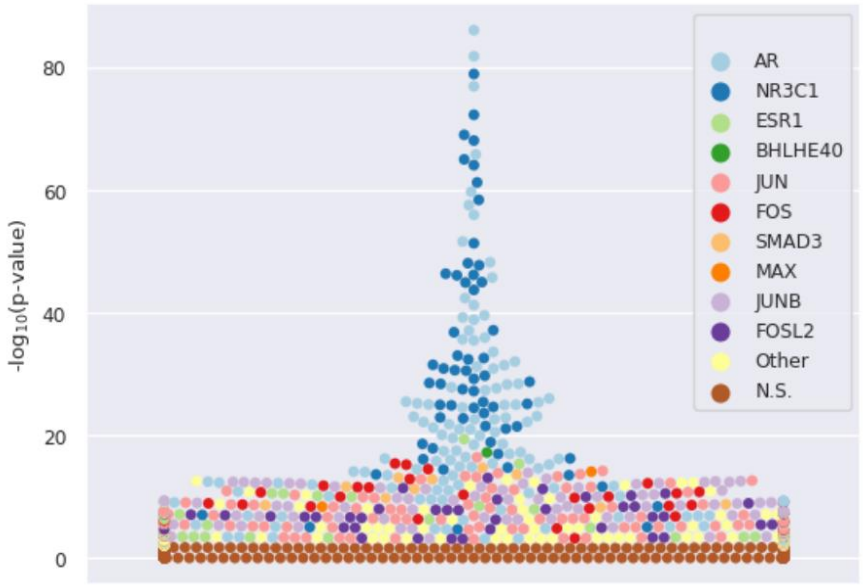


Figure 4.3 Bee swarm plot of transcription factors (TFs) corresponding to enriched transcription factor binding sites (TFBSs).

Enrichment of TFBSs (-($\log_{10}$(p-value) using Fisher's exact tests) on the y-axis for differentially methylated positions (DMPs) according to UniBind [2]. The names of the top 10 enriched TFs are denoted by the colour key; brown denotes non-significant TFs. The various data sets for the same TFs are graphed with the corresponding colour.

4.3 Discussion

We investigated a number of intrinsic, biological factors that may explain the observed DNA methylation sex differences in skeletal muscle; including fibre type proportions, blood hormone levels, and known transcription factor binding sites. We found that a 16% of sex-biased DNA methylation loci in skeletal muscle were attributed to fibre type differences. We then showed that blood serum estrogen, testosterone, free testosterone, and SHBG levels are not associated with sex-differential DNA methylation. Lastly, we report an enrichment of TFBSs among the differentially methylated loci, corresponding to 41 known TFs, with top TFs responding to sex hormones.

The large-scale meta-analysis on DNA methylation sex differences we previously conducted (Chapter 3 in this thesis) revealed that ~94% of differences displayed hypomethylation in males. Our hypotheses as to the factors contributing to the differential methylation included, sex differences in fibre type proportions, cell type proportions, exposure

to differing levels of sex hormones (the effect on TFBSs as well as other consequences), and the X chromosome complement. Assessing the potential contribution of differing cell types and X chromosome complement was beyond the scope of this study. As outlined in Chapter 3, X chromosome methylation affects certain regions of autosomal methylation, nonetheless the impact of differing X chromosome complement on autosomal sex differences across the genome has yet to be elucidated.

We hypothesised that differences in fibre type proportions between sexes may partly explain our findings [263, 264, 317], as studies report that type I fibres are hypermethylated compared with type II fibres [266], and as females tend to have a higher proportion of type I fibres than males [249]. Consistent with this, we observed that females had higher proportions of type I muscle fibres than males and that type I fibre content was mostly associated with DNA hypermethylation. Importantly, 16% of the loci exhibiting sex-biased DNA methylation were also associated with fibre type proportions. This suggests that at those CpGs, differences in DNA methylation between the sexes is due to the inherent sex differences in fibre type proportions. Nonetheless, the vast majority of the loci that exhibit sex-biased DNA methylation (84%; 48,008 CpGs) differ regardless of the sex differences in fibre type proportions. A recent study on the FUSION cohort, adjusted for fibre type proportions and found that it explains a substantial portion of the variability in DNA methylation for many metabolic phenotypes of interest [287]. Skeletal muscle DNA methylation analyses are performed on whole muscle due to the cost and technical limitations of isolating muscle cell types. Differing non-muscle cell types may be present in a muscle biopsy sample and it is currently unknown how much of the muscle DNA methylation profile may actually be representing other cell types [318-320]. Bioinformatics deconvolution methods have not yet been developed for bulk skeletal muscle DNA methylation. Considering that the DNA methylation differences between cell types are large [265], future studies should aim at determining DNA methylation patterns of the different

muscle fibre and cell types so that bulk muscle DNA methylation data can be adjusted for the appropriate cell and fibre proportions.

None of the circulating sex hormones were associated with differential methylation across all CpGs, nor across the sex-DMPs in males or females. However, the range of each hormone within each sex may not be large enough to draw out the effect of varying levels of the various sex hormones tested on the methylome. In the current study, hormone levels were measured from blood while DNA methylation was measured from skeletal muscle. DNA methylation patterns are highly tissue-specific [321, 322] and sex hormone levels in the circulation are not necessarily correlated with those intramuscularly. Moreover, intramuscular, and not circulating, sex hormone levels may be correlated with muscular function [323, 324]. A recent review emphasizes the importance of measuring intramuscular sex hormone levels when assessing muscle-related properties in females [325].

The enrichment of hormone-related TFBS among the sex-DMPs suggests that lifelong exposure to differing hormone levels significantly contributes to the observed sex differences in skeletal muscle DNA methylation. In Unibind, ChIP-seq data in skeletal muscle was limited to one TF (CTCF), so the enrichment of TFBSs among sex DMPs may have limited functional significance in skeletal muscle. Nonetheless, many of the TFs that showed strong enrichment in the present study, such as *AR* [326, 327], *ESR1* [327], and *SMAD3 [6]* are expressed in skeletal muscle and have important roles in muscle phenotype. Two recent studies leveraging the GTEx database identified sex differences in TF targeting patterns across several human tissues, including skeletal muscle, which contribute to sex-biased gene regulatory networks [11] and gene expression [9]. Differences in sex hormone levels between developing males and females are already evident *in utero* [328], making it challenging to design an experiment in humans that disentangles the effect of long-term hormonal exposure from biological sex, and other related factors, on cell function. Studies have utilised menopausal females [329] and

transgender people [330] receiving hormone replacement therapy (HRT) to investigate the influence of long term exposure to sex hormones on various phenotypes and risk of diseases. For example, HRT for one postmenopausal monozygotic twin and not the other has positive effects on regulation of muscle contraction and myonuclei organization, suggesting that estrogen has direct effects on muscle function [331]. Nevertheless, uncovering the genomic regions that display sex-differential methylation as well as contain hormone-responsive TFBSs, provides insight on which genomic regions, hormones, and TFs are discerning male and female skeletal muscle.

There is an imbalance of males and females included in biological research, and it is commonly stated that the fluctuating hormones throughout the female menstrual cycle and the variability that this may add to a given study is the main reason for the exclusion of females [18, 332]. In the present study, we therefore collected muscle biopsies during the early follicular phase (see "Methods") and assessed whether the hormones that fluctuate are associated with DNA methylation in females. There were no associations between the individual hormones (LH, FSH, progesterone, estrogen) and methylation in females. Moreover, no differential methylation was associated with the first PCA, and very few DMPs with the second PCA in females. Therefore, fluctuations in cyclic ovarian hormones during the early follicular phase did not affect the skeletal muscle DNA methylome in females. The current study was designed to minimise the potential effect of cyclic hormones on DNA methylation, and was not purposely designed to disentangle the relationship between the cyclic hormones and DNA methylation. To better understand the transient changes in DNA methylation due to acute hormonal changes, future studies should include females at every stage of the menstrual cycle in order to maximize the range of each hormone. Few studies have assessed the effect of cyclic hormones on DNA methylation. For example, Saare *et al.* found an association between menstrual cycle phase and DNA methylation in the endometrium, a tissue highly sensitive to

cyclic hormone levels [333]. However, to our knowledge, no study has investigated the effect of cyclic hormones on skeletal muscle DNA methylation in any candidate genes or genome-wide.

In conclusion, we uncovered important biological factors underlying sex-specific skeletal muscle DNA methylation. Uncovering the molecular basis of sex differences across different tissues will aid in the characterization of muscle phenotypes in health and disease. The effects of other upstream drivers on sex differences in the muscle methylome, such as non-muscle cell type, the XY chromosomes, and genetic variants still need to be explored. Molecular mechanisms that display sex differences in skeletal muscle may help uncover novel targets for therapeutic interventions.

4.4 Methods

4.4.1 Participants

The analysis of intrinsic biological factors was performed on the Gene SMART and FUSION cohorts utilised in Chapter 3. Detailed participant characteristics, study design, muscle collection, data preprocessing, and data analysis specifications for each study are in Supplementary table 3.1. Briefly, the portion of the Gene SMART (Skeletal Muscle Adaptive Response to Training) study cohort in this analysis included 20 females and 45 males aged 18-45 years with Caucasian ancestry, apparently healthy, not taking medications, and with a BMI of 18-35 [334]. This study was approved by the Human Ethics Research Committee at Victoria University (HRE13-223), and all participants provided written informed consent. Each participant was given an individualised diet 48 hours prior to the biopsy, according to the current Australian National Health & Medical Research Council (NHMRC) guidelines, to standardise diet across the participants and minimise the effects of this potentially confounding factor (15-20% protein, 50-55% carbohydrates, <30% fat and <10% saturated fat). Participants

were asked to abstain from alcohol, caffeine, or any food item not prescribed in the diet was during those 48 hours. This thesis involved adding the female cohort to the already existing male cohort, of which most of the data was collected prior to the commencement of this thesis.

The Finland-United States Investigation of NIDDM Genetics (FUSION) study from the dbGAP repository (phs000867.v1.p1) [287], included participants aged 20-77 which were either healthy or had T2D.

4.4.2 Controlling for the female menstrual cycle – Gene SMART study samples

Various contraceptives have different dosage, administration patterns, and different hormone combinations causing variability in metabolism and gene expression [117], therefore only females not taking any form of hormonal contraceptives were recruited for the Gene SMART study. Furthermore, to minimise the effect of fluctuating hormone levels, females were required to have a regular menstrual cycle (27-35 days), and all samples were aimed to be collected during the early follicular phase (day 1-day 8 of cycle), with few exceptions due to logistics. Estrogen, progesterone, follicle stimulating hormone (FSH), and luteinizing hormone (LH) were measured in blood serum. Given the intricate fluctuations of the ovarian hormones (Supplementary figure 4.2A), these four hormones were combined into a principal component (PC) analysis and the first two PCs, which explained the majority of the variability (Supplementary figure 4.3C), were both added into the linear model. We assessed whether the first two PCs were associated with DNA methylation across all of the CpGs in females. The linear model was of the form:

$$DNAm \sim time + age + PC1 + PC2$$

4.4.3 Blood serum hormones and analysis – Gene SMART study samples

The hormone assays were completed in the accredited pathology laboratory at Monash Health, Australia. Estradiol (E2) and Progesterone assays are competitive binding

immunoenzymatic assays performed on the Unicel DXI 800 system (Beckman Coulter). FSH assay is based on Microparticle Enzyme Immunoassay (MEIA) and is carried out on the Unicel DXI 800 system (Beckman Coulter). The LH and sex-hormone binding globulin (SHBG) assays were performed using a sequential two-step immunoenzymatic ("sandwich") assay carried out on a Unicel DXI 800 (Beckman Coulter). Testosterone was measured using the HPLC–tandem mass spectrometry method using a liquid sample extraction (AB Sciex Triple Quad 5500 liquid chromatography–tandem mass spectrometry). Free testosterone was calculated by the Södergard free testosterone calculation (36). One male from the Gene SMART cohort had missing hormone levels. For this individual, missing values were imputed with the *mice* package in R [335]. Hormone levels were compared between sexes using Welch's two sample t-test using baseline hormone values. To investigate whether sex hormones are associated with DNA methylation of the sex-DMPs, we included each sex hormone as a covariate in a linear model, separately. To avoid collinearity with sex, we separated males and females for this portion of the analysis. We assessed whether each of the four hormone levels (estrogen, testosterone, free testosterone, SHBG) was associated with DNA methylation across all of the CpGs and across the sex-DMPs in each sex by adjusting the linear model for a given hormone. The linear models for the association between the sex hormones (estrogen, testosterone, free testosterone, and SHBG) and DNA methylation in each sex followed the form:

$$DNAm \sim time + age + sex\ hormone\ levels + batch\ (relevant\ in\ males)$$

4.4.4 Muscle Biopsy

Muscle biopsies were sampled from the *vastus lateralis* muscle, using a suction-modified Bergström needle, under local anaesthesia of the skin and fascia (1% Xylocaine). The

muscle samples were cleaned of excess blood, fat, and connective tissue and then flash-frozen in liquid nitrogen and stored in -80ºC (see Chapter 3 “Methods” in the thesis).

4.4.5 Fibre types: meta-analysis and derivation from immunohistochemistry and RNA-seq

To assess whether fibre type proportions differed between males and females in the Gene SMART and FUSION cohorts we used a beta regression model [336] using the *betareg* package in R. We then included type I fibre ratio as a covariate in the linear models. Two females from the Gene SMART cohort had missing type I fibre proportions. For these two individuals missing values were imputed with the *mice* package in R [335]. Although proportions in type II fibres (type IIA and type IIX) and hybrid fibres (combination of the differing MHC forms) effect DNA methylation profiles, we only used type I proportions in our analysis. This is because the available methods to estimate fibre type proportions in our cohorts do not accurately measure hybrid fibres, and moreover, more confidently distinguish between type I and type II fibres than between type IIA and type IIX.

Myosin heavy chain is currently the best available marker for fibre typing [308]. Gene SMART muscle sections were frozen in optimum-cutting temperature (OCT) medium by holding the sample with OCT in liquid-nitrogen cooled Isopentane until frozen. Samples were stored in -80°C until they were sectioned at 8µM with a cryostat. The IHC protocol was performed as is described elsewhere [337]. Briefly, sections were blocked in 4% goat serum (Invitrogen). Primary antibodies BA-F8 (MHCI), BF-35 (MHCIIA), and 6H1 (MHCIIX) were purchased from DSHB, Iowa. Secondary antibodies goat anti-mouse IgG2b 350, goat anti-mouse IgG1 488, and goat anti-mouse IgM 555 were purchased from Invitrogen. Some samples were fixed in 4% paraformaldehyde in PBS for other analyses and for those samples an antigen-retrieval protocol consisting of a 10 min incubation at 50° C of Proteinase K diluted in milliQ

(1:1000) and subsequent 1 min washes was performed before the IHC. Imaging was performed on the Olympus BX51.

To determine type I fibre proportions in the FUSION cohort we followed the validated method as performed by the original study on the FUSION cohort [287]. Briefly, we derived type I fibre proportions from the RNA-seq expression data (TPMs) for type I (MYH7), type IIA (MYH2), and type IIX (MYH1). We calculated the ratio of MYH7 out of the total. We then included this ratio in the linear models.

To determine whether the inherent sex differences in fibre type proportions underlie the sex differences in DNA methylation we separated the males and females of the Gene SMART and FUSION cohorts and performed a meta-analysis on the four groups (FUSION females, FUSION males, Gene SMART females, Gene SMART males). Given that females displayed significantly higher type I fibre proportions than males in both cohorts, we could not simply include type I fibre content in a linear model performed on a mixed sex cohort as two issues would arise: i) collinearity of fibre type with sex ii) differences in fibre type proportions may be a downstream effect of sex. Dividing the cohorts by sex, conducting a meta-analysis, and selecting the sex-biased DMPs and performing an FDR adjustment among those cites allowed us to address whether fibre type proportion is associated with DNA methylation at sex-biased DNA methylation loci. The fibre type meta-analysis was performed with the same methodology of the sex meta-analysis as described in the "Methods" of Chapter 3; utilising the *bacon* R package and METAL software. The linear model for the association of type I fibre proportion and DNA methylation followed the forms:

A) Gene SMART:

$$DNAm \sim time + age + type\ I\ fibre\ proportion\ (as\ decimal) + batch\ (relevant\ in\ males)$$

B) FUSION:

$$DNAm \sim age + BMI + smoke\ status + ogtt\ status + type\ I\ fibre\ proportion\ (as\ decimal)$$

4.4.6 DNA Extraction and Methylation – Gene SMART study samples

Genomic DNA was extracted from the samples using the AllPrep DNA/RNA MiniKit (Qiagen, 80204) following the user manual guidelines. Global DNA methylation profiling was generated with the Infinium MethylationEPIC BeadChip Kit (Queensland University of Technology and Diagenode, Austria). Males and females of different ages and time points were scrambled on the chips to ensure randomness when correcting for batch effect (i.e. old and young males and females across all time points included on each chip).

4.4.7 Bioinformatics Analysis of DNA Methylation

The pre-processing of DNA methylation was explained in Chapter 3 "Methods." To identify DMPs associated with fibre type and circulating hormone levels, we used linear models as implemented in the *limma* package in R [296], using the participants' ID as a blocking variable to account for the repeated measures design (for twin and duplicate samples, using DuplicateCorrelation). Since type I fibre proportion, estrogen, free testosterone, SHBG and testosterone levels were significantly different between males and females (**Figure 4.1**; Supplementary figure 4.2C-D), we were unable to avoid collinearity with sex and therefore separated the sexes to address these questions as explained above. All tests were adjusted for multiple testing using the Benjamini and Hochberg correction [297] and adjusted p-values with $p < 0.005$ were considered significant [298]. Effect sizes are reported as mean differences in methylation (%) per percentage increase in type I fibre proportion. The list of imprinted genes across all tissues was accessed from GTEx portal. The enrichment of imprinted genes among sex-DMPs was performed by supplying the list of imprinted genes to the *gsameth* function in

the *missMethyl* R package [295, 296], which performs a hypergeometric test, taking into account biases due to the number of CpG sites per gene and the number of genes per probe on the EPIC array.

4.4.8 Enrichment of TFBSs

Enrichment of TFBSs among the identified DMRs was performed using the enrichment analysis tool in http://unibind.uio.no/ which utilizes the *runLOLA* function of the R package *LOLA* [313].

Chapter 5 : Sex-specific DNA methylation in human skeletal muscle following high intensity interval training and lifelong physical activity

5.1 Introduction

Regular exercise is one of the most cost-effective and accessible ways to improve and maintain health, with evident benefits across many tissues and diseases [338]. Thus, there is much interest in understanding how physical activity promotes health at the molecular level [339]. Both a single acute bout of exercise and exercise training induce epigenetic changes in skeletal muscle, the most energy-demanding tissue during exercise [4]. Various modalities of exercise training induce changes in the skeletal muscle methylome [4], transcriptome [340], proteome [240], and subsequent physiology [341], ultimately promoting health benefits. However, much of our understanding of molecular adaptations to exercise is limited to studies where the majority of participants were male or sex was not taken into account [18, 19]. Sex modulates myriads of biological processes, and therefore uncovering potential sex differences in molecular adaptations to exercise training may improve clinical practice.

During endurance exercise, males tend to oxidize more carbohydrates and proteins while females tend to oxidize more fatty acids [24]. However, these metabolic differences have not been observed at the enzymatic level, as key enzymes involved in β-oxidation, citric acid cycle, and electron-transport chain have similar activity levels in males and females, both at baseline and in response to endurance training [59]. Strong basal differences in the skeletal muscle transcriptome, specifically in metabolic processes, have been observed between the sexes [7-9, 11, 143]. Beyond baseline differences between sexes, a recent meta-analysis has investigated sex differences in transcriptomic *response* to training, and found 247 genes differentially expressed between the sexes following training. Nonetheless, exercise training

studies including males and females are very limited in this meta-analysis, and more studies investigating multiple layers of gene regulation are required to better understand the sex differences in the molecular response to exercise training. The skeletal muscle methylome is responsive to exercise training [340] and displays stark basal differences between the sexes [13] (chapter three), but whether it shows different responses to exercise training in males and females has yet to be investigated.

To address this, we investigated sex differences in physiological (maximum oxygen consumption (VO_2max), lactate threshold (LT), and peak power output (PP)), and epigenetic (genome-wide skeletal muscle DNA methylation) responses to four weeks of high intensity-interval training (HIIT). We also investigated whether baseline fitness levels, which reflect lifelong physical activity trajectories, were associated with similar or distinct DNA methylation patterns in males and females. Finally, we investigated the distribution of DNA methylation changes in functional regions of the human genome (CpG islands, chromatin states), and characterised the putatively affected genes and pathways using statistical enrichment methods.

5.2 Results

5.2.1 Four weeks of HIIT lead to small fitness improvements that are similar in both sexes

Twenty females and 45 males from the Gene SMART study completed four weeks of HIIT. The three fitness parameters measured before and after training include PP, VO_2max, and LT. Males had higher absolute values than females for all parameters (**Table 5.1**, Supplementary figure 5.1). Both males and females had VO_2max averages (48.6 and 44.1 mL/min/kg, respectively) slightly above those reported in the healthy general population for the corresponding age groups (35-45 for males and 30-40 for females; mL/min/kg) [26]. After four weeks of HIIT, both males and females showed substantial improvements in PP and LT,

and neither males nor females showed substantial improvements in VO_2max or fitness z-scores (average of the three fitness parameters, see "Methods") (**Table 5.1**, Supplementary figure 5.1). There were no sex differences in the degree of response to four weeks of HIIT of any of the fitness measurements (time:sex interaction).

	Males + females (combined)				
	PRE	4WP	p-value (time)	p-value (sex)	p-value (time:sex)
VO2max (mL/min/kg)	47.18	47.81	0.14	0.015	0.706
Lactate Threshold (watts/kg)	2.54	2.74	6.93×10^{-12}	0.009	0.981
Peak power (watts/kg)	3.65	3.87	1.14×10^{-12}	0.001	0.945
Fitness z score	-0.001	0.005	0.826	0.953	0.961

Table 5.1 Cardiorespiratory fitness parameters of males and females before and after four week of high-intensity interval training (HIIT).

Average values P-values for males, females, and males and females combined were analysed with linear regression.

5.2.2 Four weeks of HIIT lead to small DNAm changes that are similar in both sexes

We identified 1,261 CpGs whose methylation changed after four weeks of HIIT at a stringent false discovery rate (FDR) threshold < 0.005 (**Figure 5.1A, Figure 5.2**). A majority (80%) of the HIIT DMPs were hypermethylated following the HIIT intervention. These DMPs clustered into 28 DMRs, which were annotated to 29 distinct genes (Supplementary table 5.1). Furthermore, the DMPs were enriched for genes previously reported to display DNA methylation and transcriptional changes after three months of leg-extensor training [340] (hypergeometric test p-value = 4.6×10^{-7}), such as *SMAD3*, as well as seven Gene Ontology (GO) terms, all of which related to skeletal muscle function, such as actin binding, sarcomere, and contractile fibre (FDR < 0.005). However, the DMPs were not enriched for any Reactome pathways.

We found no sex-specific DNA methylation changes after four weeks of HIIT (sex-by-training interaction) at a stringent false discovery rate (FDR) threshold < 0.005. A global examination of all the statistical tests performed genome-wide did not reveal an inflation of near zero p-values, suggesting that males and females do not differ in their epigenetic response to four weeks of HIIT (**Figure 5.1B**).

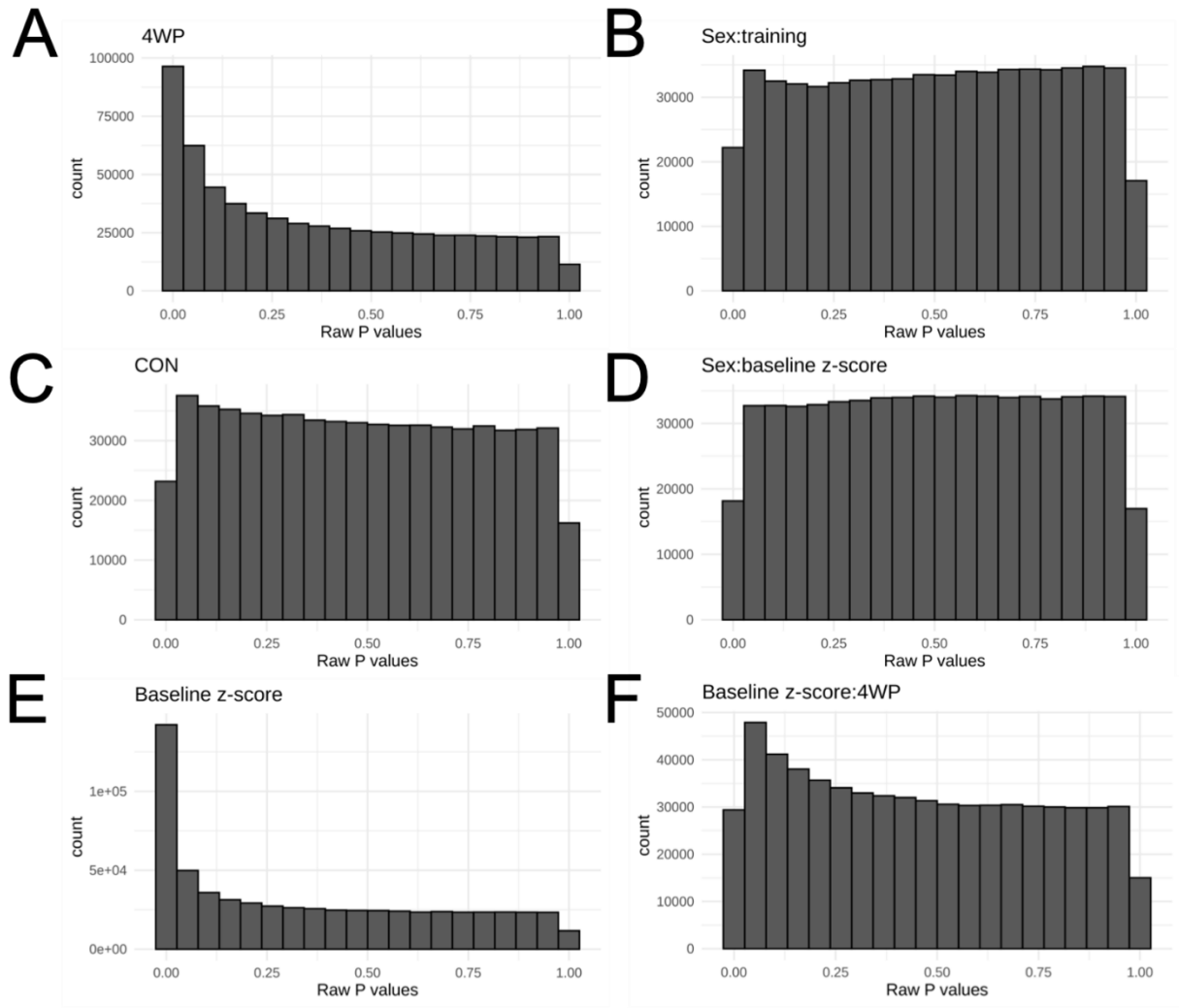


Figure 5.1 Histogram of p-values for DNA methylation for all tested CpGs.

(A) P-value histogram for the time point following four weeks of high intensity interval training (HIIT) ("4WP") relative to before the HIIT time point ("PRE"), model DNAm ~ sex + time + batch + age + baseline z. **(B)** P-value histogram for the interaction of sex and the time point following four weeks of HIIT relative to before the HIIT time point, model DNAm ~ sex * time + batch + age + baseline z. **(C)** P-value histogram for the control time point ("CON"; one month control period) relative to the PRE time point (before starting the HIIT intervention), model DNAm ~ sex + time + batch + age + baseline z. **(D)** P-value histogram for the interaction of sex and baseline fitness z-score, model DNAm ~ time + batch + age + baseline z * sex. **(E)** P-value histogram for baseline fitness z-score, model DNAm ~ time + batch + age + baseline z + sex. **(F)** P-value histogram for the

interaction of baseline fitness z-score and the time point following four weeks of HIIT relative to before the HIIT time point, model DNAm ~ sex + batch + age + baseline z * time.

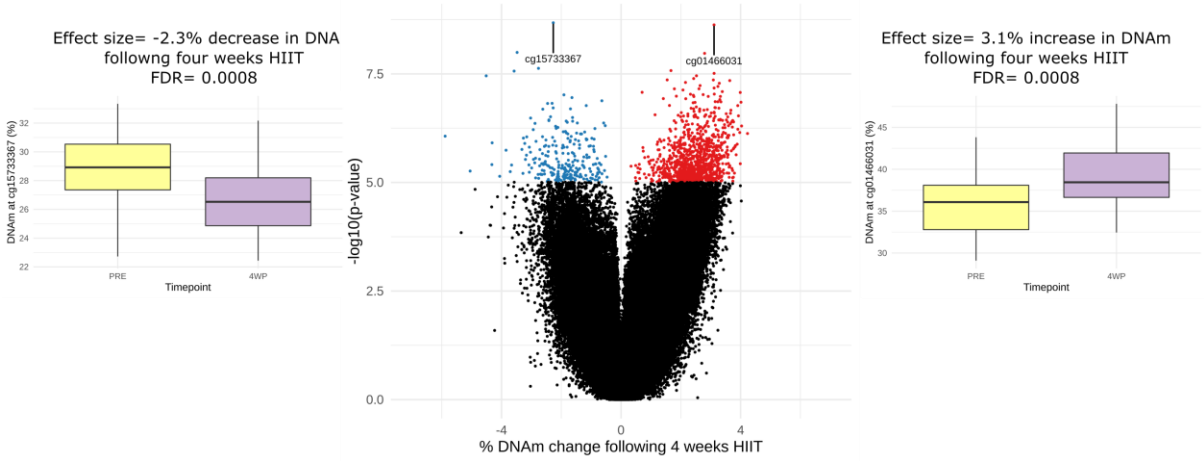


Figure 5.2 Differentially methylated positions (DMPs) after four weeks of high-intensity interval training (HIIT).

Volcano plot showing the effect of four weeks of HIIT on the 641,715 tested CpGs. The linear model was adjusted for sex, baseline fitness z-score, age, and batch. The 1,261 DMPs at a false discovery rate (FDR) < 0.005 are displayed in colors, with red dots denoting hypermethylated DMPs, and blue dots hypomethylated DMPs. Boxplot on the right represents the DNA methylation levels before and after HIIT of the hypermethylated DMP pointed to; boxplot on the left represents the DNA methylation levels before and after HIIT of the hypomethylated DMP pointed to. Yellow denotes methylation values at rest before the training intervention ("PRE"); purple denotes methylation values at rest after the training intervention ("4WP").

5.3.3 Cardiorespiratory fitness is associated with distinct DNA methylation signatures that are independent of sex

We then assessed whether baseline levels of cardiorespiratory fitness were associated with specific epigenetic signatures in skeletal muscle. We found 64,341 DMPs associated with baseline fitness (FDR < 0.005), with moderate-to-large effect sizes (-5.8% to 6.9% DNA methylation difference per unit of fitness z-score (**Figure 5.1E, Figure 5.3A**). Given that fitness z-scores ranged from -2.0 to +2.4, a CRF-associated CpG differed up to ~30% between a fitter and less fit individual. The DMPs were mostly hypermethylated (74%) in fitter individuals, overrepresented outside of CpG islands, in enhancers and in regions flanking active promoters, while underrepresented in CpG islands and active promoters (**Figure 5.3B-**

C; Supplementary figure 5.2) (Chi^2 p-value < 2.2e-16). DMPs clustered into 8,585 differentially methylated regions (DMRs) located in 6,559 unique genes (Differentially Methylated Genes, DMGs) (Supplementary table 5.2), including *GRB10* and *HDAC4*, genes which displayed sex-specific DNA methylation at baseline unrelated to CRF (chapter three). DMGs were involved in two Reactome pathways (muscle contraction, FDR = 0.0018; regulation of lipid metabolism by PPARα, FDR = 0.047) (Supplementary table 5.3), as well as several skeletal muscle-related GO terms such as actin filament-based process, myofibril, and muscle contraction (Supplementary table 5.4). DMRs were enriched for genes previously shown to be differentially expressed in endurance-trained vs untrained muscle [10] (hypergeometric test p-value = 4.6×10^{-9}), such as *MIPEP* and *CKMT2*. We did not detect sex-specific DNA methylation patterns associated with baseline fitness (sex-by-fitness interaction) (**Figure 5.1D**). To address the potential limitation of differing fitness levels among individuals in our cohort, we assessed whether DNA methylation changes were associated with four weeks of HIIT and baseline fitness z-scores (time-by-baseline fitness z-score interaction). Although we found no significant loci (FDR < 0.005), we did observe an inflation of near zero p-values (**Figure 5.1F**), suggesting that baseline fitness may have affected the degree of training-induced DNA methylation changes but that we were underpowered to detect it.

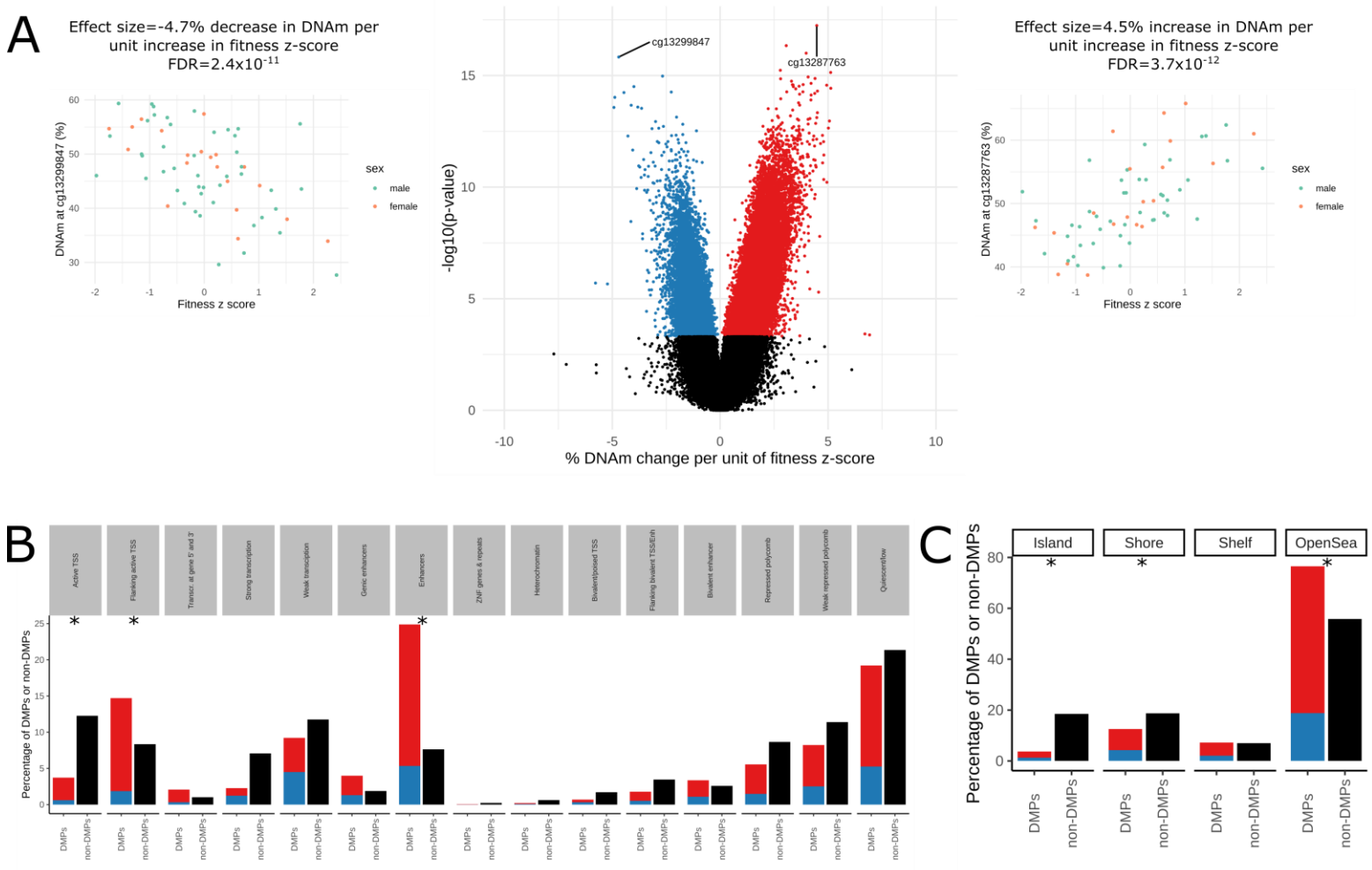


Figure 5.3 Differentially methylated positions (DMPs) associated with baseline fitness z-scores.

(A) Volcano plot showing the effect of baseline fitness z-score on the 641,715 tested CpGs. The linear model was adjusted for sex, baseline fitness z-score, age, batch, and the sex-by-baseline fitness z-score interaction. The 64,341 DMPs at a false discovery rate (FDR) < 0.005 are displayed in colors, with red dots denoting hypermethylated DMPs, and blue dots hypomethylated DMPs. Dotplot on the right represents the DNA methylation levels versus baseline fitness z-scores of the hypermethylated DMP pointed to; dotplot on the left represents the DNA methylation levels versus baseline fitness z-scores of the hypomethylated DMP pointed to. Orange denotes females and green denotes males. **(B)** Percentage of DMPs and non-DMPs associated with baseline fitness z-scores that are annotated to each chromatin state as determined by the Roadmap epigenome project; colours correspond to A; asterisks denote chromatin states that significantly contribute to the chi2 test (Supplementary figure 5.2A). **(C)** Percentage of DMPs and non-DMPs associated with baseline fitness z-scores that are annotated to CpG islands, shores, shelves, and open sea; colours correspond to A and B; asterisks denote island locations that significantly contribute to the chi2 test (Supplementary figure 5.2B).

5.3 Discussion

We investigated whether genome-wide DNA methylation in skeletal muscle has a sex-specific response to HIIT, and whether males and females display distinct DNA methylation signatures of baseline cardiorespiratory fitness in muscle. Four weeks of HIIT induced modest changes in the DNA methylome with a stringent FDR< 0.005, and we detected no sex-specific response. Cardiorespiratory fitness at baseline was associated with widespread DNA methylation changes in the muscle methylome, but these were independent of sex. Fitness-associated genes were involved in biological processes relevant for skeletal muscle function, such as muscle contraction and various metabolic pathways.

Four weeks might be relatively short training intervention in particular for our moderately-trained participants, nonetheless, we detected small overall changes in the muscle methylome after HIIT, most of which were hypermethylated. A handful of studies have reported DNA methylation changes in skeletal muscle after short-term (< 6 months) resistance or endurance exercise training [182, 340, 342]. Conversely to our results, existing studies reported equal global fractions of hypo- and hyper- methylation following training [182, 340, 342]. Two of these studies contained only males [182, 342], and in the only study containing females [340], sex was confounded with batch and therefore could not be statistically taken into account. Thus, the disparity in fractions of global hyper- and hypo- methylation between our study and those in the literature may be due to the inclusion of both sexes and the subsequent statistical adjustment. One study failed to detect changes in the muscle methylome after HIIT/resistance/combined training [343], but their analysis was restricted to promoters and to DNA methylation changes > 5%. In the present study, fitness-related DMPs were depleted across active promoters, which is consistent with enrichment reported by Lindholm *et al.* among enhancers, gene bodies, and intergenic regions [340]. In addition, we, as well as others [340, 342], detected only modest (<6%) effect sizes with training, suggesting that

Robinson *et al*. were unable to detect exercise-induced changes because of their stringent effect size threshold and limited genome coverage. Of the genes associated with differential methylation following the HIIT intervention in our study, *SMAD3* has recently been shown via meta-analysis, as a central regulator in transcriptomic networks in response to acute exercise [6]. Altogether this suggests that DNA methylation changes in *SMAD3* could underlie its role as a regulator of exercise training molecular responses in males and females, however direct causality cannot be determined.

No study, to date investigated potential sex differences in exercise responses, as the abovementioned studies included only a limited number of males and females, or were restricted to male participants. The present study was therefore the first to investigate potential sex differences in epigenetic response to HIIT in a substantial cohort of males (n = 45) and females (n = 20). However, we found no evidence of a sex-specific response to HIIT at the epigenetic level, and this is in spite of the marked sex differences in the muscle methylome at baseline (chapter three in thesis). Lindholm *et al*. also reported that muscle methylome clustered by sex and training time point, but their DNA methylation assay was performed on separate batches for males and females. Batch effects in the Illumina arrays can significantly confound results and samples should be strategically positioned as it is often not possible to remove technical signal when batches are confounded with variables of interest [344]. Similar results were reported at the transcriptional level in a single cohort (12 males and 11 females), with no differences between the sexes after training despite baseline differences [340]. However, utilizing a meta-analysis (409 males and 310 females) Amar *et al.* detected sex differences in the transcriptomic response to training. Thus, it cannot be excluded that additional cohorts could provide sufficient power to detect sex differences in the DNA methylome response to training. Altogether, our findings indicate that short-term HIIT alters

the male and female skeletal muscle methylomes similarly. However, the paucity of studies on the topic means that it is too early to draw any firm conclusions.

In our cross-sectional sample of healthy individuals, cardiorespiratory fitness (CRF) reflects years, if not lifelong patterns of regular physical activity. Individuals with higher CRF displayed distinct DNA methylation patterns, with no sex-specific differences. This suggests that life-long physical activity induces similar changes in the male and female muscle methylome, which is consistent with the lack of sex-specific response to HIIT we observed. CRF-associated regions were mostly hypermethylated, and were enriched in enhancers and regions flanking active promoters while depleted in active promoters. In contrast, Sailani *et al*. found that lifelong physical activity was associated with promoter hypomethylation in older healthy men [345]; discrepancies between our results and theirs may be due to the differences in genomic coverage owing to utilization of a different DNA methylation technique (sequencing), as well as age and sex of participants. In a recent study, untrained male and untrained female transcriptomes showed more significant differences than trained male and trained female transcriptomes, suggesting that endurance training shifts male and female transcriptomes to be more similar to one another [10]. Interestingly, genes identified by Chapman *et al*. as differentially expressed with lifelong training were overrepresented among the genes that were differentially methylated with CRF in the present study. This indicates that exercise training may trigger both DNA methylation and mRNA expression changes at specific genes, such as *MIPEP* and *CKMT2*, however the direction of causality cannot be concluded.

MIPEP, mitochondrial intermediate peptidase, is highly expressed in human skeletal muscle and plays a critical role in the oxidative phosphorylation system (OXPHOS), the oxidation of substrates to generate ATP in the mitochondria [346]. DNA methylation of *MIPEP* was associated with CRF, regardless of sex, and was different between the sexes at baseline (chapter three), thus shedding light on a gene highly active in skeletal muscle that retains basal

sex differences despite changes with physical activity. *CKMT2*, creatine kinase S-type, is an important factor in myogenesis differentiation as well as plays a central role in energy transduction in tissues with large energy demands, such as skeletal muscle, by catalysing the transfer of phosphate between ATP and phosphogens such as creatine kinase [347]. *CKMT2* protein levels have been reported to increase following 12 weeks of endurance training in both type 1 and type 2 skeletal muscle fibres [240]. In the current study, *CKMT2* DNA methylation was associated with CRF, but did not exhibit sex differences at baseline, and may be central in exercise adaptations in skeletal muscle, regardless of sex. Finally, we found that genes that were differentially methylated with CRF were involved in pathways and biological processes that are highly relevant to skeletal muscle function during exercise, indicating that years of regular exercise training may favourably shape the skeletal muscle DNA methylome so as to promote health.

Participants in our human cohort ranged from sedentary to recreationally active, to exceptionally active (one male and one female outliers). This heterogeneity in baseline fitness levels may limit our ability to detect changes in the DNA methylome following four weeks of HIIT, as changes in physiological and molecular measures may differ in magnitude depending on the exercise training history of the individual. To address this potential limitation, we assessed whether DNA methylation changes associated with four weeks of HIIT were also associated with baseline fitness z-scores (time-by-baseline fitness z-score interaction). Although we found no significant loci, the inflation of near zero p-values suggests that there may be an association between baseline fitness levels and level of training response but that we were underpowered to detect it. In contrast, a heterogeneous cohort might be advantageous as it better reflects the general population.

Short-term exercise training induced less changes in the muscle methylome, while CRF, which represents lifelong patterns of physical activity, was associated with marked

epigenetic signatures. This suggests that exercise-induced epigenetic changes are dose-dependent (i.e. training over longer periods of time leads to more pronounced epigenetic changes in muscle), and although the current study had a sample size significantly larger than previous studies with human muscles (typically n=7-10), larger sample sizes, utilising multi-site studies and initiatives, and open access data sharing, are required to detect the small shifts in the methylome that can be achieved with short-term exercise training.

In conclusion, we showed that males and females display similar DNA methylation changes in response to four weeks of HIIT as well as in DNA methylation signatures of CRF, despite profound differences in muscle DNA methylation at baseline. Genes whose DNA methylation was associated with CRF were involved in pathways related to muscle contraction and metabolism, suggesting that lifelong physical activity shapes the regulatory landscape of entire pathways relevant for muscle function. We uncovered multiple genes whose methylation levels were associated with CRF, such as *MIPEP* and *CKMT2*. More research is required to elucidate sex differences in the DNA methylomic response to exercise training as this study was the first and only to date to investigate this question. Furthermore, given the transient and rapid nature of chromatin organization changes, this may be an important avenue for future exercise physiology research.

5.4 Methods

5.4.1 Muscle Biopsy and Blood Sampling

See chapter 3 and 4 of thesis. Briefly, muscle biopsies were taken from the *vastus lateralis* muscle after an overnight fast. Intravenous blood was drawn immediately after the biopsy. Muscle was flash-frozen in liquid nitrogen and stored in -80º C. Biopsies were also taken immediately and three hours after the first training session, this portion of the analysis was used for research questions not included in this thesis (Hiam et al., *unpublished*).

5.5.2 Study Design and Exercise Protocol

An overview of the exercise protocol used in the Gene SMART (Skeletal Muscle Adaptive Response to Training) study has been previously published [334] (**Figure 5.3**). The training intervention consisted of four weeks of a control period (six of the females and 1 of the males had control periods which were analysed for DNA methylation), followed by four weeks of high-intensity interval training (HIIT) performed on a cycle ergometer.

Participants were asked to refrain from strenuous exercise before the testing days. Familiarisation graded-exercise (GXT) and VO_2max tests were performed prior to starting the study. Then, participants were additionally tested at three time points: before the control period, before the first training session, and after the training period. Testing at each time point was performed in duplicates and the average between two tests was used for analysis unless the coefficient of variance was above 10% and in that case the maximum value was used. The peak power output (PP) and lactate threshold (LT) were determined from the GXTs. PP, or the supramaximal power output, is reported as 105% of the maximum power reached in the last stage of the GXT as is commonly used [348]. GXTs were performed on an electronically-braked cycle-ergometer (Lode-excalibur sport, Groningen, the Netherlands) and consisted of 4-min stages separated by 30-s rest periods until exhaustion. The test started at 50 and 60 watts and was increased by 25 and 30 watts in each subsequent stage, for females and males, respectively. Participants that were particularly untrained had lower starting watts and smaller increases in watts per stage (20 watt increases each stage; starting at 25 watts). Starting and stage-increments watts were determined and adjusted after the familiarization GXT and kept consistent for all of the GXTs for that participant throughout the trial. Capillary blood samples were taken at rest, after each completed stage, and immediately following exhaustion, and were analysed by a YSI 2300 STAT Plus system (Yellow Springs, Ohio, USA). During the GXT the LT was calculated by the modified DMAX method, which is determined by the point on the

polynomial or exponential regression curve that yields the maximum perpendicular distance to the straight line connecting the first increase in lactate concentration above resting value and the final lactate point [349]. LT was measured using the modified D-max exponential method in the females and the polynomial method in the males. The two methods yielded similar results, but the exponential method yielded less variance between two replicated tests and was therefore implicated when females began to be recruited for the study. In any case, LT measurements are scaled per sex and therefore the differing methods used do not impact conclusions. VO_2max was obtained following the GXT after resting for seven minutes, as VO_2max is likely to be achieved following a previous maximum effort, known as priming [350]. VO_2max was determined using a calibrated Quark CPET metabolic system (COSMED, Rome, Italy). Participants wore the Cosmed face mask and we collected VO_2 at stationary for 2 min, while exercising for 3 min at the intensity of the first stage of GXT (25, 60, or otherwise determined), and during exercise to exhaustion at 105% of peak watts measured during the last stage of the previous GXT. VO_2max was considered the highest value in 1 min obtained during the test. The HIIT phase commenced 48–72 h after the last baseline exercise test.

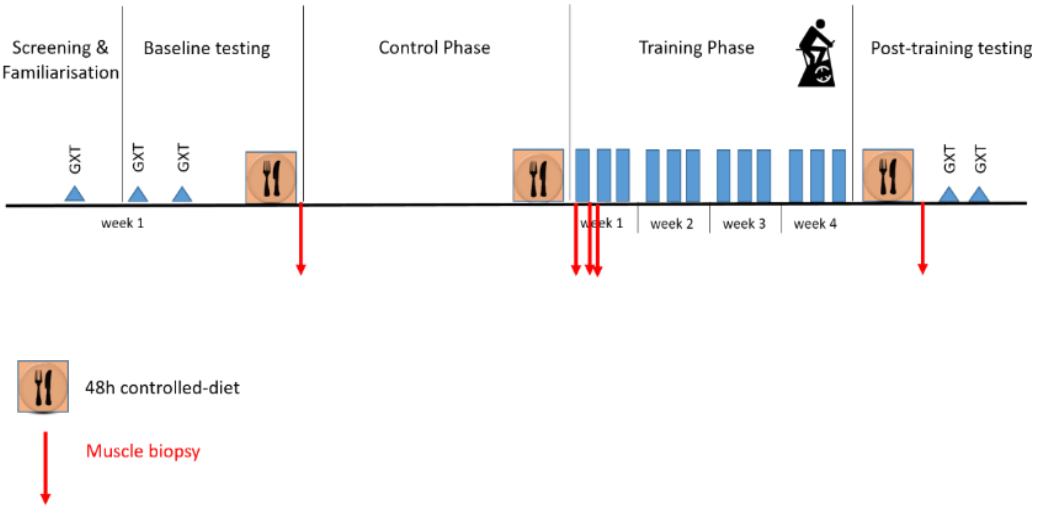


Figure 5.3 Gene SMART study design.

Blue arrows denote graded-exercise testing (GXT) and maximum oxygen consumption (VO2max) testing.; blue bars indicate one of the 12 HIIT sessions; cutlery indicate a 48 hour control diet and red arrows denote muscle biopsies and blood sampling.

Participants performed 12 HIIT sessions over the four weeks (three sessions/week). All training sessions were completed on an electronically braked cycle ergometer (Velotron, Racer Mate Inc, Seattle, USA) and were preceded by a five minute warm up. Each session consisted of six to twelve two minute intervals performed at different resistances. The training resistances for each participant were calculated according to the LT and PP determined via the GXTs. Specifically, 40, 50, 60, and 70 % were used in the following equation to determine training load:

$$LT + 40\% * (PP - LT)$$

The training intensity and duration (number of two minute repetitions) became progressively higher and longer throughout the 12 sessions, with the last session being a tapering session approaching the final GXTs and VO_2max tests. Muscle biopsies, blood sampling, GXTs and training sessions were separated by a minimum of 48h to avoid overtraining and acute training effects.

We combined the physiological fitness measurements (PP, VO_2max, and LT) to obtain a comprehensive representation of cardiorespiratory fitness. This fitness z-score was calculated by averaging all of the z scores of each physiological measurement, which were each scaled per sex. The calculation for each z score:

$$Z = \frac{x - \mu}{\sigma}$$

Z	=	z score
x	=	observed value
μ	=	mean of the sample
σ	=	standard deviation of the sample

Each measurement was scaled separately because of the differing units of each measurement. Measurements were scaled per sex given the significant differences in all three measurements

between the sexes (absolute and relative to body size). The sex comparison of physiological measurements and z-scores, before and after the intervention, were analysed using a linear model of the form:

$$Physiological\ measurement\ (z\ score, PP, VO_2max, LT) \sim sex + time + age$$

5.5.3 Controlling for diet

As explained in the "Methods" in Chapter 4, each participant was provided with individualised, pre-packaged meals for the 48 h prior to the resting muscle biopsies. The energy content of the provided meals was calculated using the Mifflin St-Jeor equation and each participant's body mass (BM), height and age [351]. The content of the diets were constructed based on the current National Health and Medical Research Council (NHMRC) guidelines. Participants were provided with a post-training and post-testing snack consisting of protein (0.3 g·kg−1 BM) and carbohydrates (0.3 g·kg−1 BM) [352]. Participants were asked to refrain from alcohol and caffeine during the dietary control period, which is 48 h prior to each resting biopsy. Outside of the dietary-control period they were asked to continue with their normal exercise and dietary habits.

5.5.4 Participants and control of confounders

Females with a regular menstrual cycle (26-35 days)[353] not taking hormonal contraceptives were recruited in order to obtain a homogenous cohort, as different contraceptives have different dosage, administration patterns, and different hormone combinations causing variability in metabolism and gene expression [354]. For consistency and to control for the potential effects of hormonal fluctuations during the female menstrual cycle, all biopsies were performed during the early follicular phase (days 1-7 of cycle).

Participants (total of six females and one male) served as their own controls as they underwent four weeks of a control period prior to starting the training, this was done in order to assess whether DNA methylation fluctuates with regular lifestyle (diet, sleep, exercise, etc.) in the absence of the exercise training intervention (**Figure 5.2C**).

5.5.5 DNA Extraction and Methylation

As described in the "Methods" of Chapters 3 and 5, Genomic DNA was extracted from the samples using the AllPrep DNA/RNA MiniKit (Qiagen, 80204) following the user manual guidelines. Global DNA methylation profiling was generated with the Infinium MethylationEPIC BeadChip Kit (Queensland University of Technology and Diagenode, Austria). The first batch contained only males and were randomised for timepoint and age. The second batch contained males and females and samples were scrambled on the chips to ensure randomness when correcting for batch effect (i.e. old and young males and females across all time points included on each chip). The genome-wide DNA methylation pattern was analysed with the Infinium MethylationEPIC BeadChip array.

5.5.6 Bioinformatics Analysis

Pre-processing was performed as described in "Methods" of Chapters 3 and 4. We adjusted the EWAS for bias and inflation using the empirical null distribution as implemented in *bacon* [293]. Inflation and bias in EWAS are caused by unmeasured technical and biological confounding, such as population substructure, batch effects, and cellular heterogeneity [294]. The inflation factor is higher when the expected number of true associations is high; it is also greater for studies with higher statistical power [293]. The results were consistent with the inflation factors and biases reported in an EWAS in blood [293].

To identify DMPs, we used linear models as implemented in the *limma* package in R [296], using the participants' ID as a blocking variable to account for the repeated measures

design. All results were adjusted for multiple testing using the Benjamini and Hochberg correction [297] and all CpGs showing an FDR < 0.005 were considered significant for the association of DNA methylation with baseline fitness [298]. When no DMPs were detected at FDR < 0.005, we examined the histogram of p-values to evaluate whether results were truly negative or whether we were underpowered. CRF-associated DMRs were identified using the *DMRcate* package [299]. DMRs with Stouffer, Fisher, and harmonic mean of the individual component FDRs (HMFDR) statistics < 0.005 were deemed significant. Effect sizes are reported as mean differences in DNA methylation beta values (%).

We integrated a comprehensive annotation of Illumina HumanMethylation arrays [300] with chromatin states from the Roadmap Epigenomics Project [255] and the latest GeneHancer information [301]. Baseline fitness-DMPs that were annotated to two differing chromatin states were removed for simplicity and because there were very few such DMPs. GSEA on Reactome and GO databases was performed on DMRs using the *goregion* (for GO) and *gsameth* (for Reactome) functions in the *missMethyl* R package [302] [303].

The linear models used to address the DNA methylation questions in this chapter include:

(A) Effect of four weeks of HIIT:

$$DNAm \sim sex + time + batch + age + baseline\ z$$

(B) Interaction of sex and time following four weeks of HIIT:

$$DNAm \sim sex * time + batch + age + baseline\ z$$

(C) Effect of control month:

$$DNAm \sim sex + time + batch + age + baseline\ z$$

(D) Interaction of sex and baseline fitness z-score:

$$DNAm \sim time + batch + age + baseline\,z * sex$$

(E) Effect of baseline fitness:

$$DNAm \sim time + batch + age + baseline\,z + sex$$

(F) Interaction of baseline fitness and time following four weeks of HIIT:

$$DNAm \sim sex + batch + age + baseline\,z * time$$

Chapter 6 : General discussion, contribution to knowledge, limitations, and future research

6.1 General discussion

Although medical research aimed at enhancing health is constantly on the rise, physical activity remains the most effective way to improve health and prevent various diseases. Advancing the understanding of gene regulation in skeletal muscle of both males and females following exercise training provides new, comprehensive, and valuable insights to discover new therapeutic treatment targets, as well as to inform and improve future biomedical research.

In Chapter 1, we introduced the general outline of each chapter. In Chapter 2 we discussed the sex differences associated with exercise training by summarising the existing literature on the genetic, epigenetic, molecular, phenotypic, and structural sex differences in skeletal muscle. There is convincing evidence that skeletal muscle displays a plethora of sex differences across various levels from genotype to phenotype, both at baseline and in response to exercise training. Despite the current knowledge on skeletal muscle sex differences, however, the underlying molecular mechanisms remain largely unexplored. There is a large body of evidence suggesting that sex differences in skeletal muscle might exist at the epigenetic level, as epigenetic sex differences have been observed in other tissues [158, 251, 252], and there are many transcriptomic sex differences in skeletal muscle [7-11, 142]. Given the evidence, we recommend that more research be undertaken to elucidate whether there are sex differences in the molecular response to exercise training.

The overall aim of this thesis was to investigate whether there are sex-specific DNA methylation changes associated with four weeks of high intensity exercise training. To explore this, the thesis was separated into three experimental chapters:

I. Do baseline ('pre-training') DNA methylation patterns differ between male and female skeletal muscle? (Chapter 3)

Several studies have identified a plethora of sex differences in the skeletal muscle transcriptome [7-11, 142, 143], suggesting that sex differences in muscle phenotypes have a molecular basis; however, these mechanisms are poorly understood. In Chapter 3 we showed that the skeletal muscle DNA methylome displays profound differences between males and females, and that the sex-differentially methylated regions were enriched for genes with known sex-biased expression, suggesting that differential methylation and expression are functionally linked. Additionally, differentially methylated regions were enriched for substrate metabolism, as well as muscle contraction pathways. This is corroborated by results from transcriptomic studies, which report that skeletal muscle female-biased genes are enriched for pathways involved in fatty acid metabolism while male-biased genes are enriched for pathways involved in protein catabolism [143]. Together, these findings suggest that the observed phenotypic skeletal muscle sex differences, such as substrate metabolism and contractility, are associated with DNA methylation patterns in human muscle.

II. Which biological factors underlie the basal DNA methylation sex differences? (Chapter 4)

To explore the influence of biological factors (e.g., fibre type distribution, circulating and lifelong exposure to hormones) on the observed sex-specific DNA methylation patterns, we assessed the association between these factors and regions that exhibited sex-specific DNA methylation. Sex differences in fibre type proportions and sex hormones have been thought to contribute to many of the sex differences in muscle phenotypes, such as metabolism [2, 355]. In Chapter 4, the lack of association between circulating sex hormones (estrogen, testosterone, free testosterone, and SHBG) and DNA methylation at the sex-DMPs in each sex (analysed separately) suggests that acute levels of sex hormones do not explain the sex differences in

DNA methylation. Rather, the lifelong exposure to sex hormones and their cumulative molecular effects likely influence DNA methylation and bring about sex-differential methylation, as sex-differentially methylated regions were enriched for hormone-responsive TFBSs (for TFs estrogen, androgen, and glucocorticoid receptors). The literature partially corroborates that gene regulation differs between the sexes due to hormone-related TFs; one study found that sex-biased gene expression across many tissues is enriched for hormone-responsive TFBSs [9], while another using the same data found that those same sex-biased genes were not enriched for hormone-related TF binding motifs across multiple tissues. Nonetheless, both of these studies find large differences between the sex-specific regulatory networks of differing tissues, highlighting the importance of elucidating the effect of sex and the implicated biological factors in each independent tissue [11]. Lastly, given the differing methylation profiles of differing fibre types [266, 287], we analysed muscle fibre type proportions and treated type I fibre proportion as a covariate in the linear model. These results revealed that type I fibre proportion is associated with DNA methylation at several CpG sites, which were primarily hypermethylated, as has been corroborated in the literature [266]. Furthermore, since females (in our study and as reported in the literature [249]) tend to have higher type I fibre proportions, and the majority of sex-specific DNA methylation displayed hypermethylation in females, it was of great interest to determine whether the profound hypermethylation in females across the genome was owing to the hypermethylation patterns of type I fibres. Despite this, fibre type proportion was associated with 16% sex-specific methylation sites, suggesting that the vast majority of differences in genome-wide methylation in males and females is not an artefact of the higher type I proportions observed in females. Altogether, the results from this chapter reveal that circulating hormone levels do not explain the sex differences in skeletal muscle DNA methylation, although lifelong hormone exposure and subsequent gene regulatory action do partially explain them.

III. Are there sex differences in the DNA methylome response to exercise training? (Chapter 5)

A handful of studies have reported that exercise training induces changes to the skeletal muscle DNA methylome which are then associated with gene expression and protein changes [4, 138, 297, 301]. However, these results either included males only or did not account for sex in their statistical model and pooled males and females together. In this thesis, by training males and females, and treating sex as a confounder, it was possible to investigate whether there are sex differences in the DNA methylome response to training. We found that four weeks of HIIT elicited small DNA methylation changes with no differential methylation between the sexes with training. This agrees with the literature as exercise training induces relatively small but widespread changes in DNA methylation [4]. One study on the transcriptomic response to training did not find any sex differences [314]; however, a meta-analysis on the transcriptomic response to exercise training reports sex differences in 247 genes after training [6]. This example from the literature emphasizes the strength of meta-analyses, especially in investigating healthy individuals in which the effect of the intervention may be relatively small and thus, require a larger sample size to detect significance. Unfortunately, we were not able to conduct a meta-analysis of the interaction of sex and training in muscle because the only other study that included females arranged male and female samples on separate batches [297]. This makes it impossible to disentangle sex effect from batch effect, which is reported to be substantial in DNA methylation arrays [303]. Findings from this chapter suggest that fitter individuals have different DNA methylation patterns than less fit individuals at rest, and that these DNA methylation patterns are similar between males and females. These findings are in line with what is currently known about the molecular profiles of trained and untrained individuals; a recent study comparing the transcriptomes of trained and untrained individuals reported that the differences between trained males and females were reduced when compared

to the differences between untrained males and females [10]. Altogether, the findings from this chapter indicate that lifelong fitness alters the DNA methylome considerably, and does so similarly in males and females, while a four week training intervention alters the DNA methylome slightly, and does so similarly in males and females. Nonetheless, replication studies as well as larger sample sizes are needed to further explore the question of whether training induces different DNA methylation changes between male and female skeletal muscle.

6.2 Contribution to knowledge

The work of this thesis has significantly contributed to the body of knowledge by showing, for the first time, that the human skeletal muscle DNA methylome displays profound differences between males and females, and identifying thousands of genes that display sex-differential methylation. Furthermore, by leveraging the GTEx database, we identified hundreds of genes with both sex-differential expression and DNA methylation in skeletal muscle. By integrating genome-wide sex-biased DNA methylation and expression in skeletal muscle, we shed light on distinct molecular sex differences in skeletal muscle. We then showed that intrinsic biological factors, such as fibre type proportions and hormone-related transcription factor activity, are associated with DNA methylation sex differences. Lastly, we performed the first comparison of male and female skeletal muscle methylomes i) after exercise training and ii) with lifelong patterns of physical activity, and find that both four weeks of HIIT and lifelong physical activity modifies the methylome similarly between the sexes.

Collectively, this work provides multiple important contributions to the literature. It exemplifies a tightly-controlled human exercise training study that included both males and females, in which the results were not greatly confounded by ovarian hormone fluctuations, the common justification for studies to research only males [18]. It provides a comprehensive picture of the genes distinguishing male and female skeletal muscle. Additionally, it contributes to the general growing knowledge on biological sex differences, an aspect of biology

commonly overlooked and understudied – which addressed can drive discovery and not addressed can lead to undesired ramifications and misinterpretation of results [1, 25, 356, 357]. Furthermore, it provides a more thorough understanding of the exercise training-induced molecular changes that occur in humans, with greater potential for future discovery of key pathways and genes involved.

6.3 Limitations

We endeavoured to minimise shortcomings as much possible in our study design and research methodologies. Nevertheless, as with every human exercise study, there are inevitable research limitations. In this section, I will highlight the most important limitations I have encountered during my PhD research.

One limitation was the lack of control of the participants' lifestyle outside of the lab (e.g., diet, sleep, physical activity), which may confound the effect of training and affect the DNA methylome. To mitigate these effects, we limited the effect on molecular analyses by providing participants with consistent 48-hour diets (according to the NHMRC guidelines) and required them to refrain from strenuous physical activity for 48 hours prior to biopsies, as well as to fast 12 hours before the muscle biopsy. We also performed the muscle biopsies consistently at the same time of the day to mitigate the influence of circadian cycles.

There were limitations related to the participants which reduced broad applicability of our findings to the general cohort. Given the reported effect of ethnic groups on DNA methylation [358], we limited our analysis to Caucasian individuals in order to obtain a homogenous cohort to increase power to detect significant differences. This means that our results are less applicable to other ethnic groups. Furthermore (again, to achieve a more homogenous cohort to increase power) we included only females not taking hormonal

contraceptives, which reduces the applicability of our results across the whole female population.

Importantly, we recruited participants of varying fitness levels, ranging from rather sedentary to moderately trained (VO_2max: 29 to 70.5 for males and 29.5 to 67.1 for females; mL/kg/min); this could either pose as a limitation or an advantage. It may be a limitation as individuals who were already training may have not been as affected by the HIIT intervention, therefore displaying less epigenetic differences, making it harder to detect the effect of training. Although DNA methylation exercise training studies have most commonly used FDR thresholds of 0.05 [359] [340, 342], we preferred to be more stringent (FDR < 0.005). On the other hand, including individuals across different fitness and training levels makes the conclusions of this thesis more applicable to the general population. Another limitation in detecting changes with training was the length of the intervention. Our training sessions were high-intensity and we applied progressive stimuli. However, we hypothesise that a substantial remodelling of the skeletal muscle methylome might be more detectable with a training intervention longer than four weeks.

Another limitation was the use of a bulk muscle sample and not being able to take differing cell populations into consideration. Bulk tissue contains other cell types which display differing DNA methylation patterns [263-265]. Bulk skeletal muscle may contain endothelial, immune, mesenchymal stem, satellite, and fibroadipogenic progenitor cells, as well as pericytes and fibroblasts [319]. There is currently no deconvolution algorithm established to extract cell type proportions from DNA methylation data, as is available for blood [360] and saliva [361]. We attempted to address this limitation by measuring fibre type proportions, as muscle fibres also display differing methylation profiles [266]. Although myosin heavy chain is currently the best available marker for fibre typing [308], this approach still had a few limitations. It has been reported that the variability in fibre proportions between different muscle pieces within

the same biopsy and between different samplings of the same muscle is diminished when counting 150 muscle fibres, representing the most accurate estimation of fibre proportions (when utilising IHC) [362]. We attempted as much as possible to count at least 150 fibres per sample, but this was not always achieved. Nevertheless, we only included data that had at least 100 fibres, as that has been the recommendation from the literature until recently [362, 363]. In addition, we only investigated the associations with type I fibre proportions. Accounting for differing distributions of type IIA, type IIX, and hybrid fibres would further elucidate the role of fibre type proportions in skeletal muscle DNA methylation and associated sex differences.

One of the limitations with regards to the hormone-related analysis was that we measured circulating hormone levels as opposed to intramuscular hormone levels. We were recently made aware (after analysing blood hormone levels) of the lack of association between blood and intramuscular hormone levels [325]. Given that we assessed DNA methylation in skeletal muscle, it may have been more relevant to measure intramuscular hormone levels. Furthermore, given the sexually dimorphic nature of the sex hormones, we had a problem of collinearity with sex in our linear model (similar problem to fibre type proportions), therefore we split the males and females for this portion of the analysis. Given the relatively small range of sex hormone levels within healthy individuals of the same sex, the effect of circulating sex hormones may not have been detectable in our cohort. Lastly, the enrichment of TFBSs from the Unibind ChIP-seq database mostly contained data from other tissues besides skeletal muscle. It has been reported that differential gene expression is differentially targeted by TFs, depending on the tissue. Therefore, relying on TF data that has been curated from different tissues than skeletal muscle is not ideal for the investigation included in this thesis. Altogether, in the investigation of the biological factors underlying the DNA methylation baseline sex differences (Chapter 2), it was challenging to conclude the basis of DNA methylation sex differences with confidence.

6.4 Future research

Often research projects provide insights that lead to more new questions. To inform future research, I will highlight some of the research areas which this thesis has shed light on, and that if explored, may provide interesting and valuable findings:

I. Integration of findings with other –omics layers to reveal phenotypic effects

The DNA methylation patterns observed in males and females both in response to exercise training and lifelong physical activity are only one piece of the story. To deeply investigate whether there are sex differences in the molecular adaptations to exercise training and the downstream ramifications of epigenetic differences, other –omics layers should be integrated. The field would benefit greatly from integrating the transcriptome, the proteome/phosphoproteome, miRNAs, histone modifications, chromatin accessibility, and transcription factor binding, among others. Integrating these skeletal muscle-specific -omics layers will provide several layers of information and more statistical power, which will drive discovery of sex-specific pathways. Given the time allocated for my PhD I did not have time to integrate multiple –omics layers with my exercise training data. In the last few months of my PhD, the meta-analysis of the skeletal muscle transcriptomic response to training was published [6]. I hope to integrate my findings from Chapter 5 with this powerful meta-analysis to reveal whether the transcriptomic sex differences Amar *et al.* observed following training can be observed at the DNA methylation level, shedding light on the molecular mechanisms at play.

Also, while we did integrate the DNA methylation findings at *baseline* with other published transcriptome data (Chapter 3), it would be even more powerful to integrate these findings with other available skeletal muscle –omic data using a meta-analysis approach. Lastly, the findings from this thesis would be strengthened by analyzing the DNA methylome

results with skeletal muscle-specific TFBSs, as this thesis utilised the currently available data on TFBSs, which is derived from other tissues.

II. Meta-analysis of DNA methylation sex differences across several tissues

This thesis focused on skeletal muscle; however, during the preparation of this thesis multiple excellent quality studies have brought my attention to the sex differences in gene regulation across several tissues [9, 11]. Integrating sex-biased epigenetic gene regulation across several tissues will provide greater insight which may be relevant to more research fields. Furthermore, although skeletal muscle is the largest tissue functionally involved during exercise, exercise training also effects peripheral tissues, such as adipose tissue [364], which then engage in cross-talk with other tissues in the body [365] and ultimately results in enhanced health. Thus, future research should investigate the effect that exercise training has on various tissues besides skeletal muscle.

III. The effect of sex hormones on gene regulation and phenotype

It is understood that sex hormones influence physiology and pathophysiology, both genomically and non-genomically [366]. Genomically, sex hormones can bind receptors which then function as TFs and affect transcription. Non-genomically, sex hormones affect molecular function, for example, estrogen effects vascular function acutely and in the longer term. Elucidating the multitude of effects that sex hormones have on systemic function will greatly contribute to our understanding of physiology in both healthy and diseased populations. However, it is often challenging to disentangle the effect of sex hormones and biological sex, as some sex hormones display sexual dimorphism (limitation from Chapter 4). For that reason, the field would gain useful knowledge from experiments designed to separate the effect of sex from the effect of sex hormones. A recent study has been able to investigate the effect of sex on gene regulation in cultured cells without the acute effect of sex hormones on molecular function [181]; nonetheless, lifelong exposure to differing sex hormone levels and their effect

on gene regulation, which could be retained in cell culture, poses a more challenging question to address. We propose an idea, which if ethically feasible, would address this question. Comparing the epigenomes of embryonic stem cells from XX and XY embryos would eliminate the effect of lifelong exposure to sex hormones and more simply address the effect of biological sex on the epigenome. Another way to address the effect of long-term exposure to sex hormones is to compare the epigenomes of individuals undergoing hormone replacement therapy, a project which our lab is currently undertaking. Although in this study design the effects of life long exposure to hormones (especially during developmental life stages) likely cannot be totally reversed, the effects of exposure to sex hormones over a long period of time can nonetheless be investigated *in vivo*.

IV. The effect of the sex chromosomes on molecular and physical sex differences

The X and Y chromosomes are inherent drivers of many observed sex differences. To elucidate the multitude of effects they exert on molecular and physical characteristics, their effect on autosomal genes should be explored. By studying individuals with sex chromosome complement disorders, studies have found that sex chromosomes affect DNA methylation at specific autosomal genes; however, this has not yet been fully explored on a genome-wide scale. Also, to more thoroughly understand biological sex differences, the effect of X-chromosome inactivation needs to be further explored. There is a need for novel bioinformatics methods to compare the sex differences in DNA methylation of the X chromosome as it is currently bioinformatically challenging. given the unequal dosage between males and females [25]. The field of biological sex differences would benefit from further research on the X and Y chromosomes and their impact on gene regulation and function.

V. Sex differences in the effect of gene variants on DNA methylation: methylation QTLs

Common genetic variants may impact DNA methylation at a specific site. Such loci are termed methylation quantitative trait loci (meQTL). Sex differences in meQTLs mean that they may act in one sex but not in the other, or they could be shared by both sexes but with differing effect sizes, or allelic distributions in males or females. This could ultimately result in differing gene expression levels of a specific gene between males and females. Very few studies have investigated sex differences in meQTLs; however, many studies have investigated expression QTLs (eQTLs), and have reported very few sex-stratified eQTLs. This suggests that the effect of genetic variants on DNA methylation and expression are likely not highly sex-stratified and do not greatly contribute to sex differences in complex traits [25]. Given this, the time allocated for completing this thesis, and the sample size needed to perform such analysis with adequate power, we did not investigate the potential sex differences in the effect of genetic variants on DNA methylation.

VI. The effect of cell types

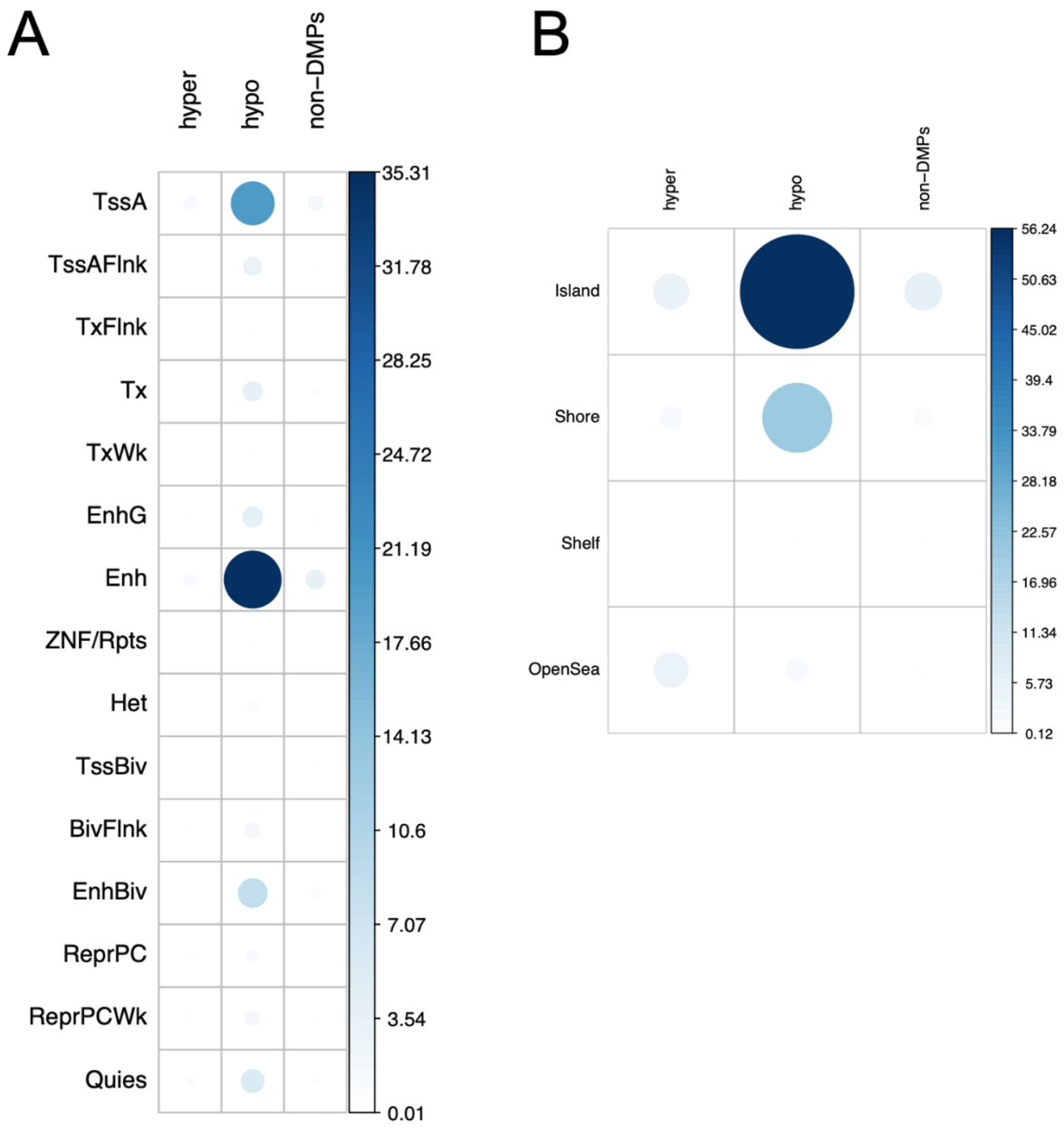
As previously discussed in Chapter 4 and in the limitations above, there may be non-muscle cells within a bulk muscle tissue sample which are characterised by differing DNA methylation patterns. To confirm that the findings from this thesis are indeed skeletal muscle-specific, cell sorting and cell-specific DNA methylation would need to be conducted. This is not only very costly, but is also very challenging from a methods point of view, as muscle cells are multinucleated and are therefore challenging to separate to a single-cell level. Future work from our lab is intending to address this limitation both in general and in relation to sex differences.

VII. Larger and more diverse cohorts (meta-analysis) for a longer intervention of various exercise training modalities

To extend the exercise training findings (Chapter 5) to a larger portion of the population, as well as increase power to detect this magnitude of effect sizes, larger and more diverse cohorts of males and females are required. This would allow for a more powerful meta-analysis to be performed. In addition, a longer intervention would bring about larger and more detectable effect sizes [154]. Lastly, the effect of various training modalities, not just HIIT, would be essential in elucidating the effects of general physical activity across the sexes.

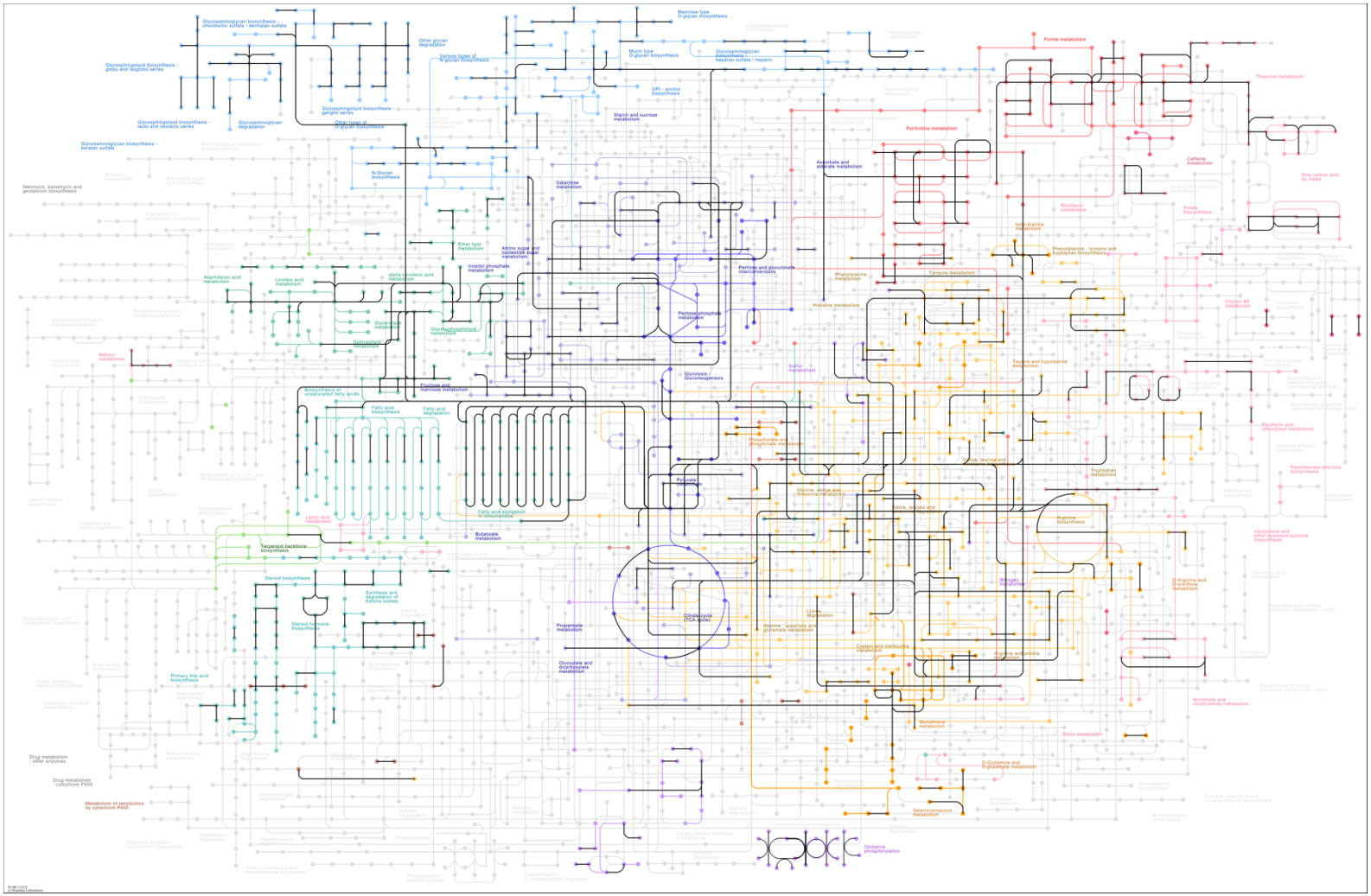
Appendix: supplementary figures and tables

Supplementary figures



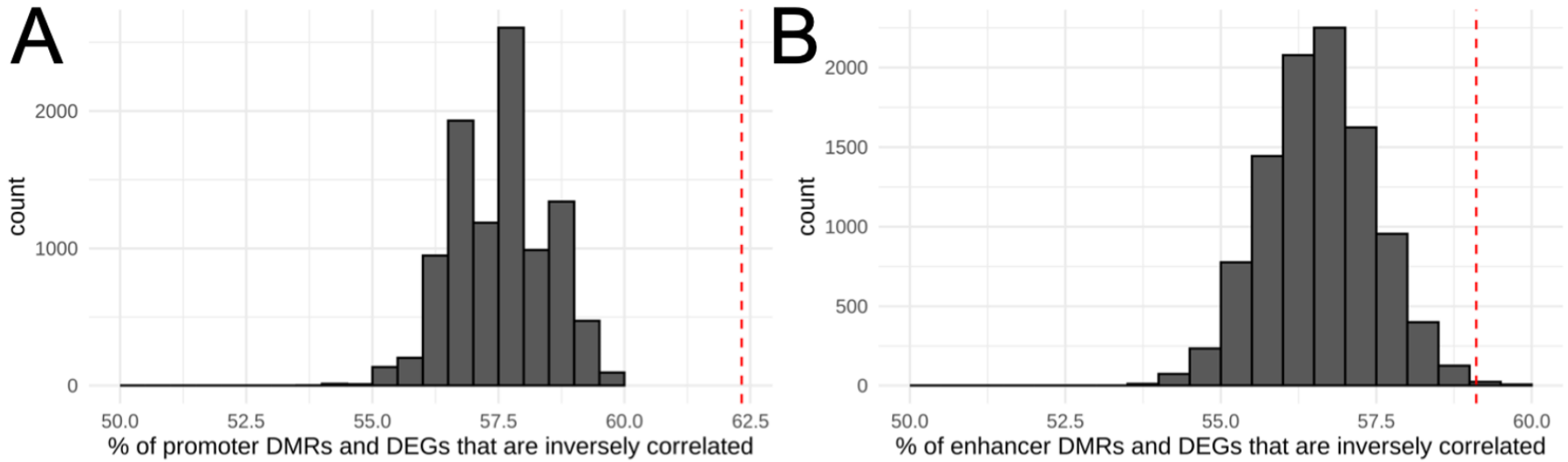
Supplementary Figure 3.1 Correlation plots of chi^2 tests of genomic locations.

(A) Correlation plot of the percent contributions to the chi2 test for chromatin states in hyper-, hypo-, and non-DMPs. This plot is using the male chromatin state annotation in skeletal muscle but the female chromatin state annotation revealed equivalent results. Darker blue indicates a greater contribution to the significant relationship between DMP status and chromatin state. (B) Correlation plot of the percent contributions to the chi2 test for CGI status of hyper-, hypo-, and non-DMPs. Darker blue indicates a greater contribution to the significant relationship between DMP status and CGI status.



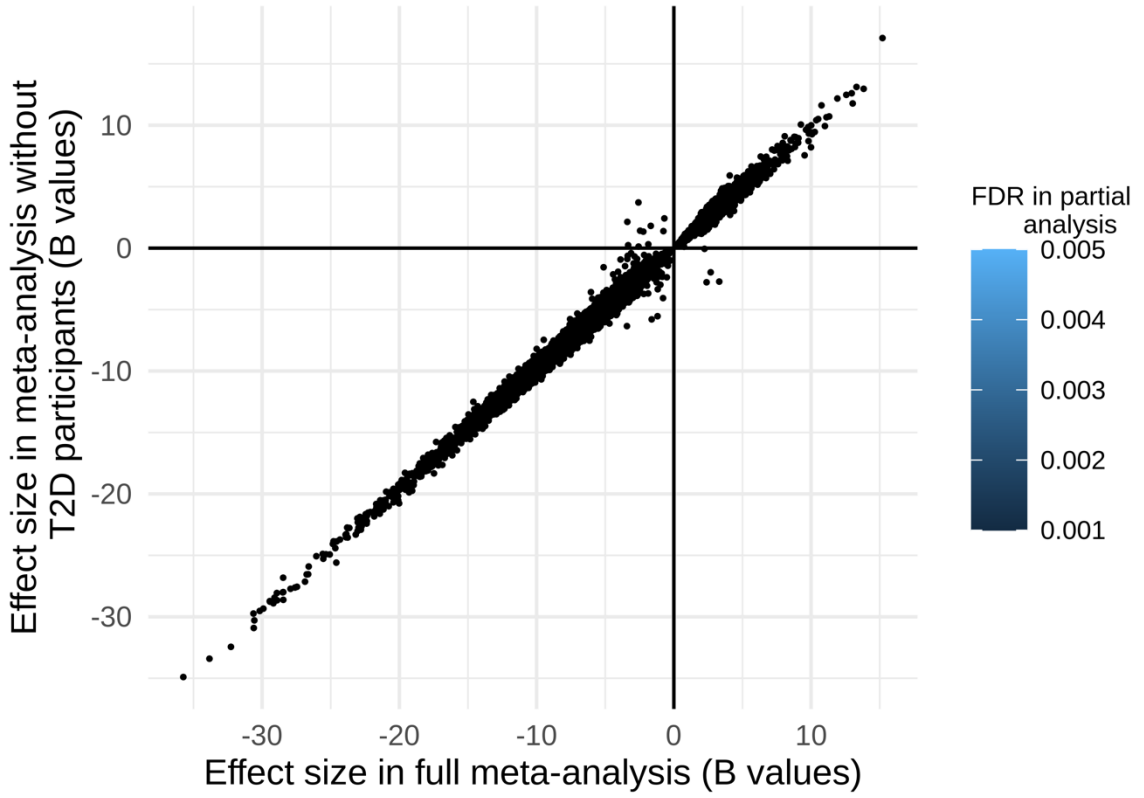
Supplementary Figure 3.2 Differentially methylated genes within the KEGG metabolic pathways map (hsa01100).

Genes and components of the pathways which display differential methylation between males and females at baseline are outlined in black, KEGG GSEA using DMPs.



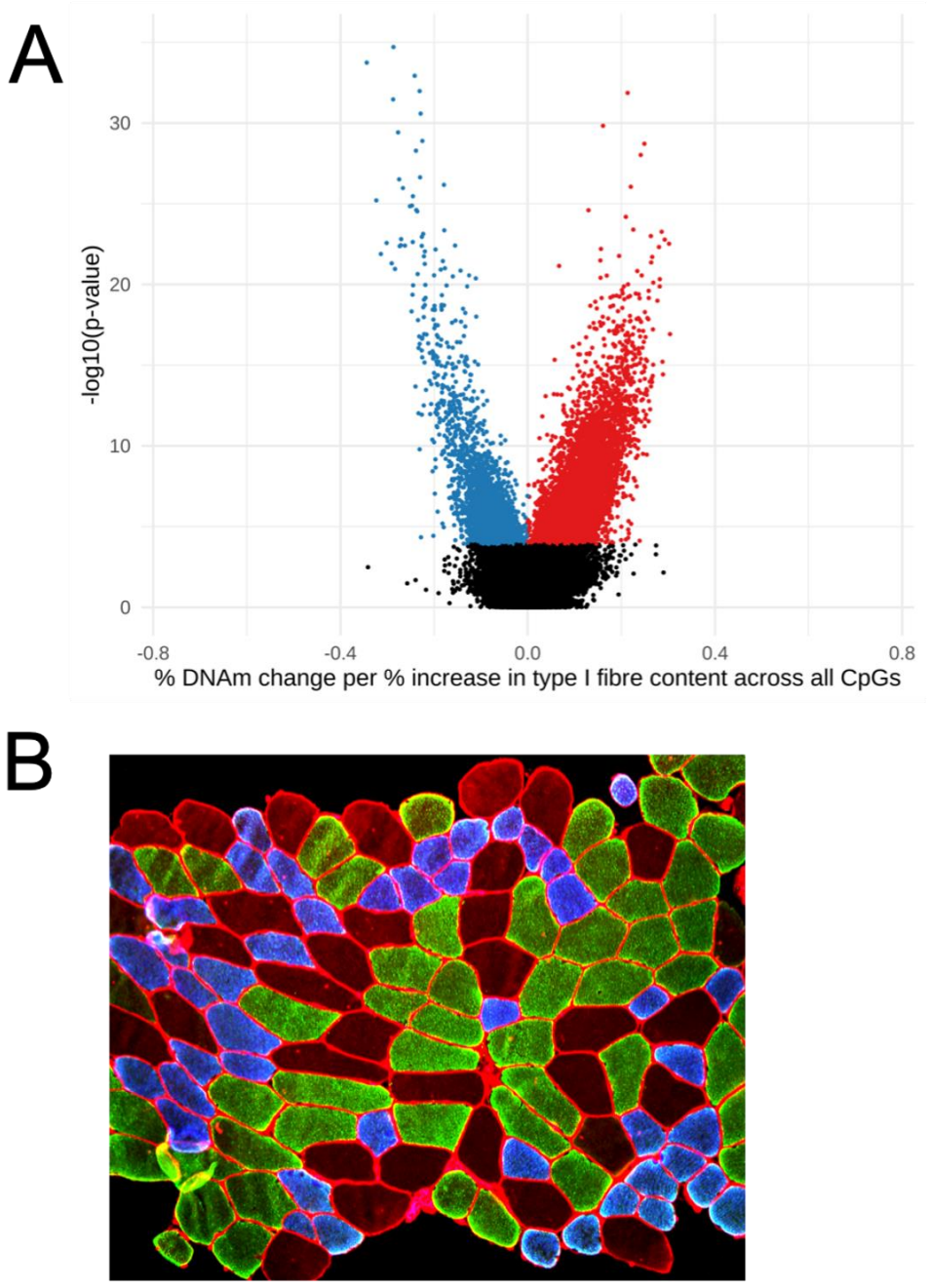
Supplementary Figure 3.3 Distribution of the 10,000 random permutations for a negative correlation between DNA methylation and gene expression.

(A) Histogram of 10,000 random permutations of DMRs annotated to promoter regions and correlation with GTEx gene expression. Effect sizes of DMRs were randomly shuffled and the resulting correlation with gene expression is plotted. Red dashed line indicated the real percentage of promoter DMRs that are negatively correlated with gene expression. (B) Histogram of 10,000 random permutations of DMRs annotated to enhancer regions and correlation with GTEx gene expression. Effect sizes of DMRs were randomly shuffled and the resulting correlation with gene expression is plotted. Red dashed line indicated the real percentage of enhancer DMRs that are negatively correlated with gene expression.



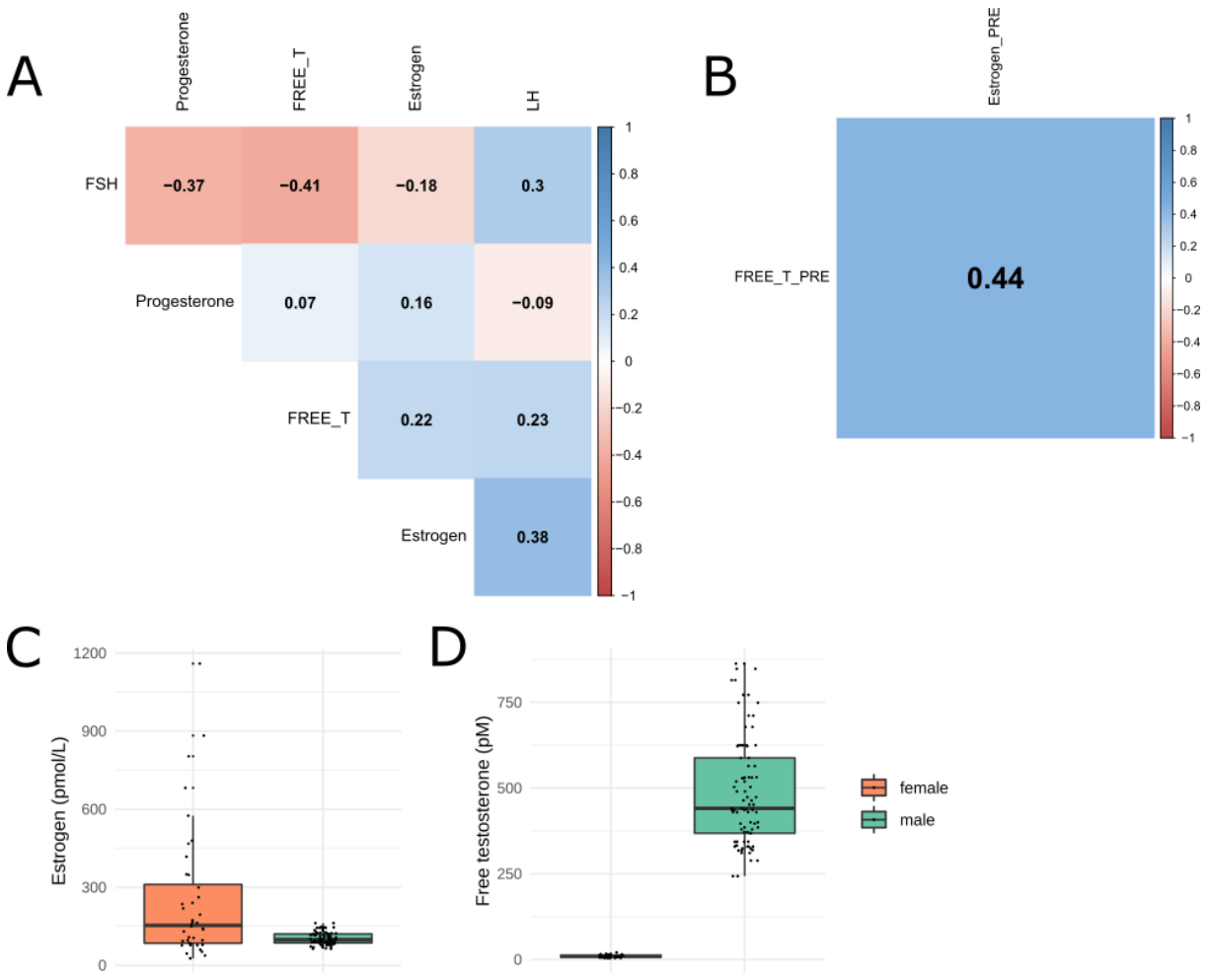
Supplementary Figure 3.4 Comparison of results from the full meta-analysis and from a meta-analysis excluding T2D participants in FUSION.

Each point is one of the 56,813 differentially methylated positions (DMPs) discovered in the full meta-analysis. To compare results from the full and partial meta-analysis we plotted the effect size (B value percentages) in the full meta-analysis (x-axis) against the effect size of the partial meta-analysis (y-axis). To show whether DMPs remained significant in the partial meta-analysis, we coloured points according to the false discovery rate (FDR) in the partial meta-analysis.



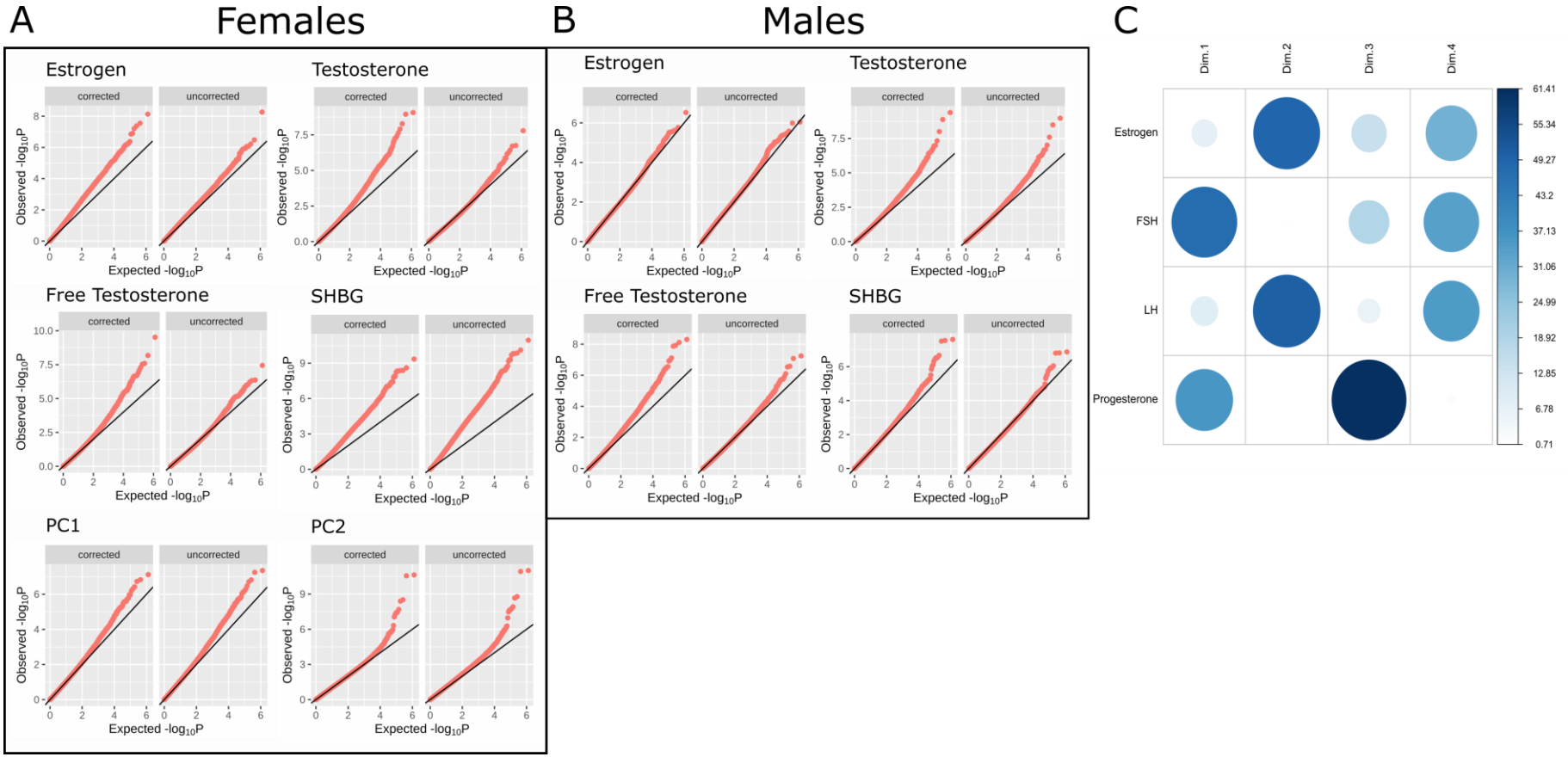
Supplementary Figure 4.1 Fibre type proportion analysis.

(A) Differentially methylated positions (DMPs) with type I fibre proportion across all CpGs conducted with a meta-analysis of males and females, separately, from the Gene SMART and FUSION cohorts. Volcano plot of DNA methylation changes per percent increase in type I fibre content (expressed at percentage of beta value). Each point represents a tested CpG (665,904 in total) and those that appear in color are DMPs at a false discovery rate < 0.005; red DMPs are hypermethylated in type I fibres; blue DMPs are hypomethylated in type I fibres. The x-axis represents the amount of DNA methylation difference with increasing type I fibre content and the y-axis represents statistical significance (higher = more significant). (B) Fibre type Immunohistochemistry myosin heavy chain staining of skeletal muscle section. Example of cross-sectional fibres of one participant. Blue fibres indicate type I, green indicate type IIa, and red indicate type IIx; cell membrane staining in red. A minimum of 100 fibres counted per person for approximation of fibre type proportions.



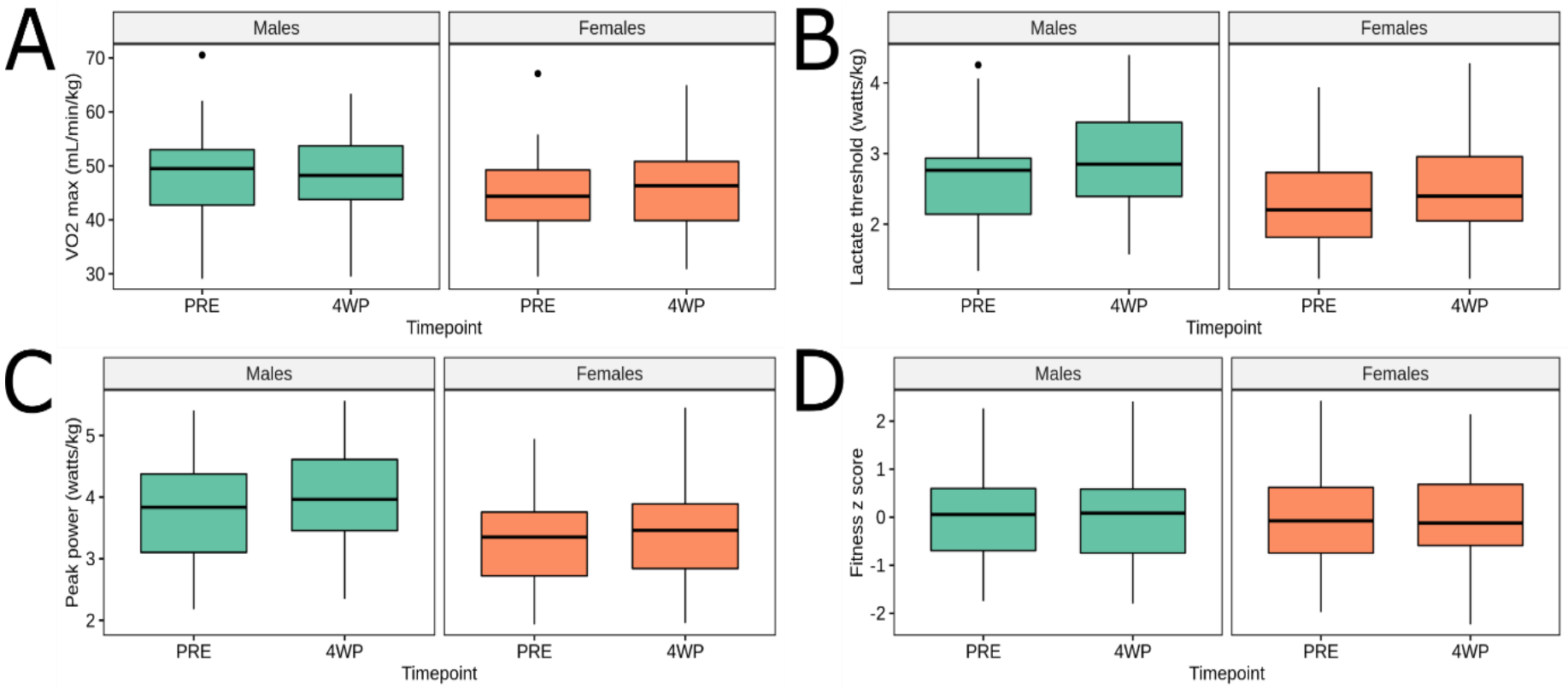
Supplementary Figure 4. 2 Circulating hormone levels of males and females in the Gene SMART study.

Correlation between hormones levels in **(A)** females and **(B)** males. Blue hues indicate a positive correlation and red hues indicate a negative correlation. Hormone levels measures from blood serum; FSH- follicle stimulating hormone, LH- luteinising hormone, Free T- free testosterone (calculated from sex hormone binding globulin and testosterone levels), estrogen- estradiol 2. Circulating **(C)** estrogen and **(D)** free testosterone levels in Gene SMART males and females.



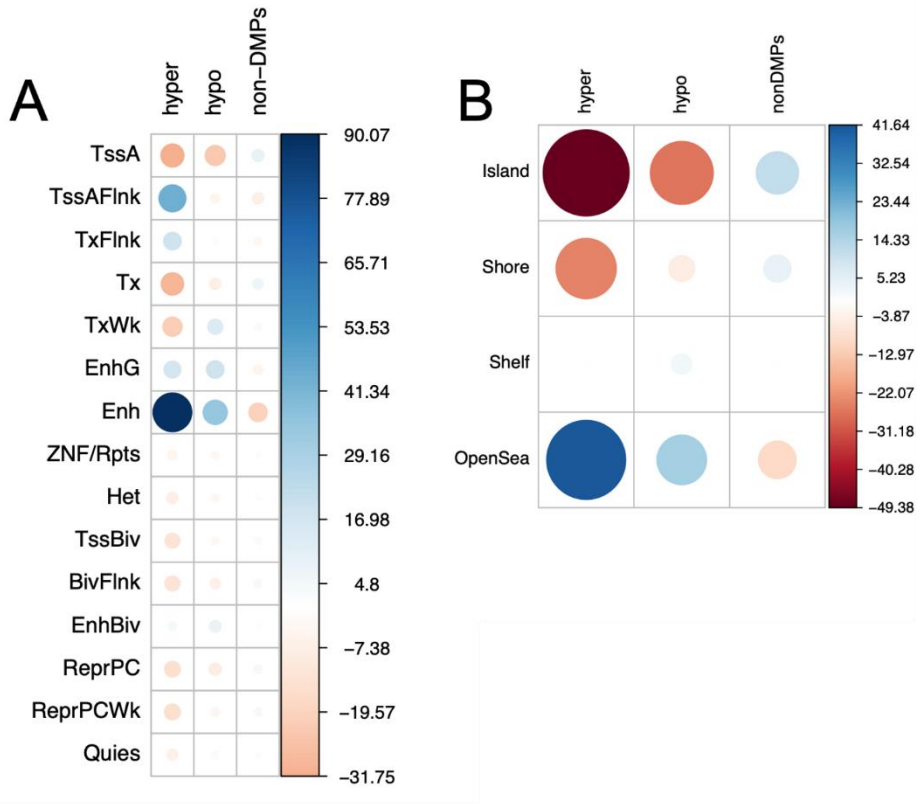
Supplementary Figure 4. 3 Association between DNA methylation and circulating hormones in males and females from the Gene SMART study.

Each Epigenome-wide association study (EWAS) was corrected for with BACON in R [1], as labelled "corrected." The uncorrected quantile-quantile (QQ) plots correspond to the unadjusted p-values from the EWAS. QQ plots are not showing a separation of the observed from the expected when all points are on or near the middle line between the x-axis and the y-axis; meaning that not many P values are more significant than expected under the null hypothesis. (**A**) QQ plots of all CpGs in DNA methylation analysis in females for the four hormones (estrogen, testosterone, free testosterone, and SHBG) and the first two principal components of the ovarian hormones (from follicle stimulating hormone, leutinizing hormone, estrogen, and progesterone). (**B**) QQ plots of all CpGs in DNA methylation analysis in males for the four hormones (estrogen, testosterone, free testosterone, and SHBG). (**C**) Contributors to the dimensions of the principal component analysis (PCA) of the ovarian hormones in females used in the linear model in A. Larger circle indicates a larger contribution to the given dimension of the PCA.



Supplementary Figure 5. 1 Boxplots of fitness before and after four weeks of high-intensity interval training (HIIT) in males and females.

(A) Relative VO2max (mL/min/kg) in males and females. (B) Relative lactate threshold (LT) (watts in which LT was reached/kg) in males and females. (C) Relative peak power (PP) (watts/kg) in males and females. (D) Fitness z scores in males and females. Green denotes males; orange denotes females.



Supplementary Figure 5.2 Correlation plots of residuals from the chi2 test for baseline fitness-DMPs enriched among the differing (A) Roadmap Epigenome project chromatin states and (B) CpG island locations.

Blue denotes enrichment and red denotes depletion.

Supplementary tables

Supplementary tables which contained several rows are presented as screen shots for ease of viewing the thesis. Full length supplementary tables can be provided upon request.

***Supplementary table 3.1*- Study descriptions. Description of participants, study design, muscle collection, and data preprocessing/analysis.**

	Participant characteristics						Study design and muscle collection								Data preprocessing						Data analysis
Dataset ID	**n**	**Health status**	**Age (mean ± SD)**	**Age range (min-max)**	**% male**	**Ethnicity**	**Design**	**Muscle**	**Time of biopsy collection**	**Control diet before biopsy**	**Exercise before biopsy**	**Fasted before biopsies**	**Stimulus**	**Timepoints**	**Array**	**Data preprocessed by us**	**Probes left after filtering**	**Sample filtering**	**Adjustment of Type I and Type II**	**Batch correction**	**Covariates included in the model**
FUSION	282	Healthy/ T2D	59.4 ± 7.:	20-77	54%	Caucasian	Cross-sectional	*Vastus lateralis*	Morning	No	No strenuous exercise for 1 day	Yes	N/A	N/A	HMEPIC	YES	656,914	YES (39 removed)	YES	YES	Age, BMI, diabetes status, smoking status
Gene SMART	130	Healthy	32± 8.1	18-45	80%	Caucasian + one mixed Aborigial /Caucasian	Pre/post	*Vastus lateralis*	7am-10am	2 days control diet	No strenuous exercise for 2 days	Yes	HIIT training	Pre, post (after acute HIIE, after 4, 8 and 12 weeks of HIIT, after a control	HMEPIC	YES	641,715	YES (2 removed)	YES	YES	Age, timepoint, batch, random intercept
GSE38291	22	Healthy/ T2D (twins)	68 ± 8	53-80	45%	Caucasian	Cross-sectional	*Vastus lateralis*	Unknown	Unknown	Unknown	Unknown	N/A	N/A	HM27	YES	26,129	YES (0 removed)	N/A (no need for the HM27 array)	NO (no information)	Age, diabetes, random intercept

***Supplementary table 3.2*- sex-DMPs. Differentially methylated positions between males and females in the meta-analysis FDR < 0.005. Corresponding chromosome, genomic location, annotated genes, male and female chromatin states from the Roadmaps Epigenomics Project, and genes annotated by GeneHancer. Positive effect size indicates higher DNA methylation in males compared to females.**

CpG	Effect	StdErr (M val	P.value	Direction in	Number of s	t_stat	BetaEffectSi	FDR	Chromosom	CpG start	CpG end	probe_stran	Annotated G	transcriptTy	transcriptID	distToTSS	CGI	CGIposition	Male Chrom	Female Chro	Genes uniqu
cg11099555	-0.3688	0.056	4.378E-11	--?	1	-6.5857143	-6.0668	2.28E-09	chr7	136698939	136698941	-	AC009264.1	antisense	ENST000005	465334	CGI:chr7:136869107-13687		15_Quies	15_Quies	AC009264.1
cg11405300	-0.3064	0.0473	9.038E-11	--?	1	-6.4778013	-5.2384	4.4797E-09	chr11	133952151	133952153	-	IGSF9B;IGSF	protein_cod	ENST000003	4603;65;4834	CGI:chr11:13	N_Shelf	12_EnhBiv	2_TssAFlnk	IGSF9B
cg13921251	-0.1593	0.0398	0.00006173	--?	1	-4.0025126	-2.7422	0.00089003	chr9	35036080	35036082	-					CGI:chr9:35071759-350728		3_TxFlnk	3_TxFlnk	
cg03983220	-0.3989	0.0554	5.738E-13	--?	1	-7.200361	-4.7502	3.9324E-11	chr16	85534800	85534802	+					CGI:chr16:85517776-85518		7_Enh	7_Enh	GSE1
cg14764176	-0.2409	0.0352	7.937E-12	--?	1	-6.84375	-3.6164	4.6404E-10	chr17	54172600	54172602	+					CGI:chr17:53757289-53757		15_Quies	15_Quies	
cg00908004	-0.1755	0.0392	7.512E-06	--?	1	-4.4770408	-2.9676	0.00013826	chr13	113178280	113178282	+	PCID2;PCID2;	protein_cod	ENST000002	30389;30431;	CGI:chr13:113152974-1131		4_Tx	4_Tx	PCID2
cg16313875	-0.1636	0.0422	0.0001074	--?	1	-3.8767773	-2.8922	0.00144343	chr17	73014022	73014024	+	SLC39A11;SL	protein_cod	ENST000002	78671;78666;	CGI:chr17:73164783-73166		7_Enh	15_Quies	SLC39A11
cg08819647	-0.1496	0.0362	0.00003517	--?	1	-4.1325967	-2.6836	0.00054295	chr8	22228927	22228929	+	PHYHIP;PHYH	protein_cod	ENST000003	3174;3414;-2	CGI:chr8:22221587-222218		13_ReprPC	13_ReprPC	PHYHIP
cg07884152	0.1243	0.029	0.00001855	++?	1	4.2862069	1.7087	0.00030856	chr3	50348328	50348330	-	NPRL2;NPRL2	protein_cod	ENST000002	2763;2453;23	CGI:chr3:503	N_Shelf	6_EnhG	6_EnhG	
cg06993307	-0.2605	0.0593	0.00001111	--?	1	-4.3929174	-2.7092	0.00019579	chr16	54935138	54935140	+	CTD-3032H1	lincRNA	ENST000005	225	CGI:chr16:54	S_Shore	10_TssBiv	10_TssBiv	CTD-3032H12
cg12423493	-0.1173	0.0328	0.0003491	--?	1	-3.5762195	-1.8873	0.00402719	chr11	64240340	64240342	+	FKBP2;FKBP2	protein_cod	ENST000003	-664;-1408;-8	CGI:chr11:64	N_Shore	2_TssAFlnk	2_TssAFlnk	FKBP2
cg22074675	-0.1236	0.0312	0.00007449	--?	1	-3.9615385	-0.9769	0.00104924	chr11	68121205	68121207	+	CHKA;CHKA;C	protein_cod	ENST000002	-1;239;-377;-	CGI:chr11:68	Island	1_TssA	1_TssA	CHKA;CTD-26
cg02990368	-0.1789	0.0361	7.223E-07	--?	1	-4.9556787	-1.2402	1.7001E-05	chr1	204151520	204151522	-	ETNK2;ETNK2	protein_cod	ENST000003	374;483;-92	CGI:chr1:204	Island	11_BivFlnk	10_TssBiv	ETNK2
cg06480224	-0.2106	0.0248	2.018E-17	--?	1	-8.4919355	-3.6914	2.3419E-15	chr2	202149015	202149017	+	KIAA2012;KIA	protein_cod	ENST000004	75726;13299;	CGI:chr2:202171328-20217		7_Enh	7_Enh	
cg24215175	-0.1481	0.0393	0.0001609	--?	1	-3.7684478	-2.4125	0.00205365	chr22	40179564	40179566	-	TNRC6B;TNR	protein_cod	ENST000003	134747;1622	CGI:chr22:40044848-40045		2_TssAFlnk	2_TssAFlnk	TNRC6B
cg00232091	-0.493	0.0492	1.285E-23	--?	1	-10.020325	-8.4692	2.5915E-21	chr20	44553710	44553712	+	PKIG;PKIG;PK	protein_cod	ENST000003	21794;21801;	CGI:chr20:44531698-44532		2_TssAFlnk	7_Enh	PKIG
cg27596068	-0.2986	0.0385	8.244E-15	---	2	-7.7558442	-3.7767	7.1227E-13	chr11	75561255	75561257	+	SERPINH1;SE	protein_cod	ENST000003	-831;-1115;-9	CGI:chr11:75	N_Shore	7_Enh	7_Enh	SERPINH1
cg10137821	-0.2675	0.034	3.824E-15	--?	1	-7.8676471	-4.6858	3.4399E-13	chr3	50230776	50230778	+	GNAI2;GNAI2	protein_cod	ENST000002	0;4484;4039;-	CGI:chr3:502	S_Shelf	15_Quies	7_Enh	GNAI2
cg03431108	-0.1788	0.0497	0.000319	--?	1	-3.5975855	-1.208	0.00372046	chr8	40343480	40343482	+	CTA-392C11.1	lincRNA;linc	ENST000005	-117;-78	CGI:chr8:40897645-408978		15_Quies	15_Quies	CTA-392C11.1
cg22563246	-0.2806	0.0576	1.098E-06	--?	1	-4.8715278	-1.8	2.4751E-05	chr3	187702216	187702218	-	RP11-211G3.	antisense;a	ENST000004	-150;-97;341	CGI:chr3:187	S_Shelf	7_Enh	7_Enh	
cg04835995	-0.1453	0.034	0.00001973	--?	1	-4.2735294	-2.597	0.00032584	chr8	101137609	101137611	-	RN7SKP249	misc_RNA	ENST000003	-350	CGI:chr8:101126365-10112		7_Enh	2_TssAFlnk	RN7SKP249
cg21635588	-0.2358	0.0498	2.164E-06	--?	1	-4.7349398	-3.9937	4.5514E-05	chr6	154301563	154301565	+	CNKSR3;IPCE	protein_cod	ENST000006	208551;5520	CGI:chr6:154509426-15451		15_Quies	15_Quies	IPCEF1;CNKSF
cg23146310	-0.143	0.0355	0.00005551	--?	1	-4.028169	-2.2976	0.00081131	chr2	239746964	239746966	+					CGI:chr2:239728304-23972		13_ReprPC	14_ReprPCW	
cg27318087	-0.212	0.0382	2.943E-08	--?	1	-5.5497382	-3.7799	9.3521E-07	chr16	12803037	12803039	+	CPPED1;CPPE	protein_cod	ENST000002	814;980;814;(	CGI:chr16:12	N_Shore	2_TssAFlnk	2_TssAFlnk	CPPED1
cg03350299	-0.1405	0.0336	0.00002919	--?	1	-4.1815476	-1.9944	0.00046097	chr2	21044087	21044089	-	APOB;APOB;/	protein_cod	ENST000002	-14;-125;-142	CGI:chr2:210	Island	13_ReprPC	14_ReprPCW	APOB
cg06911753	-0.1397	0.0244	1.099E-08	--?	1	-5.7254098	-2.4849	3.7983E-07	chr11	47589683	47589685	+	C1QTNF4	protein_cod	ENST000003	4976	CGI:chr11:47	Island	12_EnhBiv	12_EnhBiv	
cg10825710	-0.221	0.0538	0.00004035	--?	1	-4.1078067	-3.6743	0.00061319	chr12	1629170	1629172	+	WNT5B;WNT	protein_cod	ENST000003	12114;-27;54	CGI:chr12:16	N_Shore	13_ReprPC	13_ReprPC	WNT5B
cg10543500	-1.0813	0.0949	4.704E-30	--?	1	-11.394099	-10.8617	1.4829E-27	chr13	44161399	44161401	-	SMIM2;SMIN	protein_cod	ENST000004	-142;-544;17	CGI:chr13:44141737-44142		7_Enh	7_Enh	
cg18334620	-0.209	0.0445	2.648E-06	--?	1	-4.6966292	-3.5941	5.4516E-05	chr2	71601162	71601164	-	DYSF;DYSF;D	protein_cod	ENST000002	147440;1344	CGI:chr2:71560300-715607		4_Tx	4_Tx	DYSF
cg05796704	-0.1959	0.0375	1.713E-07	--?	1	-5.224	-3.2027	4.6339E-06	chr1	11654160	11654162	-	FBXO2;FBXO2	protein_cod	ENST000003	522;1625;37;-	CGI:chr1:116	Island	7_Enh	2_TssAFlnk	FBXO2;FBXO4
cg25088991	-0.1293	0.0284	5.277E-06	--?	1	-4.5528169	-1.9294	0.00010095	chr7	955490	955492	+	ADAP1;ADAP	protein_cod	ENST000002	-793;-84;-140	CGI:chr7:953	S_Shore	13_ReprPC	13_ReprPC	ADAP1;COX19
cg19342711	-0.2514	0.057	0.00001019	--?	1	-4.4105263	-1.6426	0.0001814	chr17	17074366	17074368	+	MPRIP;MPRII	protein_cod	ENST000003	31517;31606;	CGI:chr17:17042070-17043		2_TssAFlnk	7_Enh	MPRIP
cg05686323	-0.2626	0.0407	1.112E-10	--?	1	-6.4520885	-4.3839	5.4264E-09	chr6	31865384	31865386	+	SLC44A4;SLC	protein_cod	ENST000002	13617;13662;	CGI:chr6:318	S_Shelf	14_ReprPCW	14_ReprPCW	SLC44A4
cg20366545	-0.1891	0.0293	1.024E-10	--?	1	-6.4539249	-2.8157	5.0345E-09	chr9	121676145	121676147	-	DAB2IP;DAB	protein_cod	ENST000002	109043;1090	CGI:chr9:121698688-12169		7_Enh	7_Enh	
cg13712686	-0.2736	0.0482	1.35E-08	--?	1	-5.6763485	-4.1626	4.5796E-07	chr11	77046484	77046486	-					CGI:chr11:77	S_Shore	12_EnhBiv	14_ReprPCW	
cg25757345	-0.1463	0.0331	0.00001	--?	1	-4.4199396	-2.6136	0.00017842	chr5	171537458	171537460	-					CGI:chr5:171456596-17145		14_ReprPCW	15_Quies	
cg16508068	-0.1349	0.0314	0.00001765	--?	1	-4.2961783	-2.3141	0.00029521	chr10	92691978	92691980	-	HHEX;HHEX;I	protein_cod	ENST000002	2027;158;-54	CGI:chr10:92	Island	10_TssBiv	10_TssBiv	HHEX
cg10893604	-0.2511	0.0438	9.984E-09	--?	1	-5.7328767	-4.2953	3.4777E-07	chr8	28118859	28118861	+	ELP3;ELP3;ELI	protein_cod	ENST000002	26021;25792;	CGI:chr8:28092953-280933		5_TxWk	7_Enh	ELP3
cg10950615	-1.0688	0.0704	4.558E-52	--?	1	-15.181818	-14.698	4.1616E-49	chr2	121562217	121562219	-	CLASP1;CLAS	protein_cod	ENST000002	87259;87259;	CGI:chr2:121530044-12153		7_Enh	7_Enh	
cg25410220	-0.1349	0.033	0.00004429	--?	1	-4.0878788	-2.3653	0.00066569	chr7	69801822	69801824	-	AUTS2;AUTS	protein_cod	ENST000003	202489;2022	CGI:chr7:69597388-696000		15_Quies	15_Quies	AUTS2
cg05193430	-0.1138	0.03	0.0001494	--?	1	-3.7933333	-0.4109	0.00192439	chr14	67619524	67619526	+	ARG2;ARG2;/	protein_cod	ENST000002	-274;-456;-41	CGI:chr14:67	N_Shore	2_TssAFlnk	2_TssAFlnk	ARG2
cg25897519	-0.5675	0.0596	1.726E-21	--?	1	-9.5218121	-4.3692	2.9172E-19	chr8	41569895	41569897	+	RP11-360L9.4	antisense	ENST000005	8473	CGI:chr8:415	S_Shelf	14_ReprPCW	14_ReprPCW	RP11-360L9.4

***Supplementary table 3.3*- sex-DMRs. Differentially methylated regions between males and females in the meta-analysis Stouffer, HMFDR, and Fisher <0.005. Corresponding chromosome, genomic location, width of DMR, number of CpGs in DMR, statistics (Stouffer, Harmonic mean false discovery rate (HMFDR), and Fisher statistic), maximum and mean DMR effect sizes, and annotated genes. Positive effect size indicates higher DNA methylation in males compared to females.**

	Chromosome	CpG start	CpG end	width	no.cpgs	min_smooth	Stouffer	HMFDR	Fisher	Maximum DMR Effect Size (%)	Mean DMR Effect Size (%)	Annotated genes
1	chr12	67647537	67649018	1482	9	0	0	0	0	-24.593	-3.569133333	DYRK2
2	chr2	144237447	144237880	434	6	0	2.194E-288	2.52E-139	0	-29.4413	-14.7521	GTDC1
3	chr3	46580101	46580445	345	4	0	1.146E-256	1.186E-145	0	-28.4953	-17.05945	LRRC2;TDGF1
4	chr7	129404714	129405193	480	3	0	1.001E-201	1.245E-262	0	-35.7397	-18.95393333	AHCYL2
5	chr2	26704268	26704824	557	3	0	0	1.89E-193	0	-33.837	-26.2961	KCNK3
6	chr20	34859382	34859547	166	3	0	0	2.782E-134	0	-28.4663	-20.42366667	GGT7
7	chr1	9900407	9900594	188	3	0	2.49E-303	6.567E-130	1.796E-305	-21.4233	-15.8935	CTNNBIP1
8	chr1	206555705	206557837	2133	16	4.277E-161	1.104E-143	2.7605E-67	1.967E-300	-14.5423	-4.56025	RASSF5;EIF2D
9	chr10	112907053	112907263	211	2	2.227E-248	4.565E-281	2.15E-145	5.808E-280	-29.1787	-26.45135	
10	chr4	140096161	140096639	479	3	5.164E-207	1.185E-238	3.295E-191	3.602E-274	-32.2725	-20.97283333	MAML3
11	chr15	41476675	41477277	603	3	4.547E-247	2.507E-191	7.857E-130	9.212E-251	-23.0285	-13.90296667	RTF1
12	chr2	3470080	3470397	318	3	8.318E-278	4.542E-167	2.927E-165	2.239E-250	-23.7641	-15.22666667	TRAPPC12
13	chr1	40392249	40392389	141	3	3.033E-269	1.237E-166	6.729E-128	8.743E-243	-24.5633	-15.9123	RLF, SMAP2;SMAP2
14	chr21	33022784	33024826	2043	11	3.135E-119	1.497E-218	9.0386E-43	4.654E-232	-5.837	-3.433218182	AP000282.2;OLIG2
15	chr8	47353285	47353349	65	2	2.849E-236	4.337E-211	1.533E-177	2.89E-228	-28.4566	-18.7743	SPIDR
16	chr16	664967	665202	236	5	1.349E-246	4.218E-177	1.07E-102	8.215E-225	-20.0036	-10.53046	WDR90;MSLN
17	chr4	10022697	10023074	378	4	2.294E-240	1.61E-170	3.6424E-92	5.42E-223	-19.84	-13.0901	SLC2A9
18	chr6	33277021	33279731	2711	51	5.362E-261	1.4342E-62	3.3658E-74	2.651E-222	-12.2542	-2.055272549	B3GALT4;RPS18, B3GALT4
19	chr1	22652487	22653245	759	6	1.757E-225	3.415E-184	1.228E-79	1.741E-216	-15.7569	-7.879583333	C1QB
20	chr14	93918787	93918902	116	3	1.472E-218	1.865E-215	1.9809E-83	7.391E-215	-18.2553	-14.67893333	FAM181A-AS1;FAM181A
21	chr20	57519468	57519563	96	3	1.669E-226	2.107E-181	1.859E-126	2.869E-214	-21.8281	-13.19176667	CTCFL
22	chr6	32128746	32129729	984	15	2.008E-261	6.37E-134	3.245E-67	7.598E-212	-13.1137	-4.536333333	FKBPL;ATF6B
23	chr6	7142344	7142404	61	3	2.679E-217	6.189E-213	2.1248E-78	3.907E-211	-14.7835	-12.58543333	RREB1
24	chr12	52494661	52494695	35	2	3.955E-203	2.286E-199	2.43E-116	1.899E-199	-16.1931	-15.9973	KRT6A
25	chr8	28348187	28348760	574	2	7.21E-111	8.739E-199	9.022E-110	3.887E-198	-19.291	-15.97215	ZNF395;FBXO16
26	chr1	117847246	117847515	270	2	4.114E-189	2.913E-179	5.398E-148	1.982E-192	-30.6081	-18.12615	
27	chr1	155131092	155131573	482	3	1.558E-185	1.947E-165	7.144E-108	1.016E-190	-20.7222	-13.21176667	EFNA1
28	chr2	128445302	128445834	533	4	9.391E-162	5.881E-157	5.291E-101	2.378E-183	-16.7628	-9.151775	RNA5SP103
29	chr22	50205931	50207066	1136	10	7.662E-192	4.049E-148	3.3652E-39	1.631E-182	-16.2475	-8.39769	SELENOO;LOC105373095
30	chr8	42377095	42378736	1642	11	1.985E-149	5.06E-161	4.4376E-42	2.497E-182	-13.255	-6.312363636	DKK4
31	chr11	68934424	68934822	399	3	4.566E-228	1.5173E-74	1.075E-183	4.886E-182	-26.0489	-9.0011	IGHMBP2
32	chr17	81527601	81528926	1326	10	3.415E-134	2.868E-163	1.1923E-41	1.132E-181	-9.2058	-4.84327	FSCN2
33	chr12	6375052	6378510	3459	20	1.737E-153	2.264E-106	2.7157E-61	4.482E-181	-13.8479	-3.69003	LTBR;SCNN1A
34	chr3	129819692	129819707	16	2	1.448E-183	2.51E-173	8.17E-129	2.515E-180	-20.1548	-11.8752	TMCC1
35	chr15	100559078	100559286	209	2	1.992E-228	3.6806E-92	3.39E-182	3.666E-180	-11.4739	-5.6187	RP11-526I2.1;PRKXP1
36	chr1	17622130	17623108	979	6	1.011E-174	4.229E-171	1.4919E-47	8.323E-180	-11.8363	-6.8379	ARHGEF10L
37	chr15	60681547	60681591	45	2	2.157E-180	3.962E-133	2.519E-164	4.554E-175	-23.6434	-12.7029	RORA-AS2;RORA
38	chr8	100336130	100336815	686	10	6.549E-157	2.378E-163	4.9896E-26	4.26E-169	-11.0365	-7.61418	KB-1991G8.1;RNF19A

***Supplementary table 3.4*- Overlapping genes. Genes which displayed sex-biased gene expression in GTEx and FUSION as well as sex-biased DNA methylation (according to DMRs) in the meta-analysis. Corresponding chromosome, ensemble gene ID, gene name, GTEx mash posterior effect size, GTEx local false sign rate threshold, FUSION mRNA fold change, FUSION mRNA FDR, and number of DMPs per gene. Positive effect sizes indicate higher DNA methylation or expression in males compared to females.**

	Chromosome	Ensembl gene ID	Gene name	GTEx mRNA mash Posterior Effect Size	GTEx mRNA lfsr	FUSION mRNA Fold Change	FUSION mRNA FDR	Number of DMPs per gene
1	15	ENSG00000258484	SPESP1	0.896763442	0	1.04900802	4.35997E-16	8
2	4	ENSG00000168785	TSPAN5	0.944846577	0	1.599302605	1.05889E-29	15
3	16	ENSG00000177508	IRX3	1.481956989	0	2.549910054	5.60727E-51	3
4	13	ENSG00000225083	GRTP1-AS1	1.299341166	0	4.078484761	1.18978E-87	4
5	7	ENSG00000105983	LMBR1	0.220908416	0	0.244482982	6.19758E-09	4
6	18	ENSG00000101542	CDH20	1.403773886	0	2.009229944	4.16723E-28	8
7	20	ENSG00000131067	GGT7	0.813754203	0	1.588575131	1.32339E-45	14
8	9	ENSG00000165092	ALDH1A1	0.331854946	0	0.380412402	2.33559E-08	6
9	4	ENSG00000185818	NAT8L	-0.541925172	1.77332E-33	-0.91386131	1.59302E-27	6
10	3	ENSG00000188001	TPRG1	-0.674253368	1.23146E-29	-1.566755476	4.45704E-28	22
11	21	ENSG00000184221	OLIG1	-1.028171623	4.95482E-22	-0.979984683	8.86816E-08	9
12	2	ENSG00000162998	FRZB	-0.778889767	4.2392E-18	-0.457797035	0.004826006	2
13	5	ENSG00000080709	KCNN2	-0.49715601	5.14593E-18	-0.258500855	0.000225302	5
14	2	ENSG00000116035	VAX2	-0.642296	5.91753E-18	-1.192211862	9.08215E-19	15
15	13	ENSG00000165566	AMER2	2.104263097	1.11022E-16	3.652850828	1.24073E-44	5
16	5	ENSG00000247516	MIR4458HG	0.231980584	2.22045E-16	0.283203153	2.87866E-05	11
17	16	ENSG00000103034	NDRG4	0.50883442	3.33067E-16	0.609049364	4.71692E-20	7
18	20	ENSG00000124177	CHD6	0.203461143	4.10783E-15	0.112915249	0.001038032	10
19	6	ENSG00000080007	DDX43	0.819225109	1.15685E-13	0.811779123	1.06193E-06	7
20	4	ENSG00000218336	TENM3	0.235961663	3.0147E-12	-0.181430105	0.000877683	13
21	12	ENSG00000187109	NAP1L1	-0.124675911	5.94037E-12	-0.160750883	0.000399233	3
22	2	ENSG00000196141	SPATS2L	0.176863742	1.19545E-11	0.232405682	2.97368E-10	18
23	11	ENSG00000154114	TBCEL	0.114897005	1.3362E-11	0.335324085	2.57142E-11	6
24	15	ENSG00000062524	LTK	0.619469281	5.81468E-11	0.558470421	0.000756301	9
25	4	ENSG00000145362	ANK2	-0.19343973	1.22831E-10	-0.18165562	9.35168E-06	11
26	11	ENSG00000134569	LRP4	-0.266795696	3.20603E-10	-0.62138249	2.23691E-21	3
27	7	ENSG00000055118	KCNH2	-0.399401851	4.53005E-10	-0.831584172	1.38265E-16	7
28	8	ENSG00000182759	MAFA	0.279412348	1.1143E-09	0.39407019	8.43081E-05	7
29	15	ENSG00000224078	SNHG14	0.171342585	1.18999E-09	0.745109151	5.93038E-40	26
30	13	ENSG00000068650	ATP11A	0.123651006	1.92848E-09	0.303864831	1.14534E-08	16
31	11	ENSG00000109927	TECTA	0.18627004	2.92436E-09	0.477044223	3.52898E-14	1
32	2	ENSG00000115840	SLC25A12	-0.10432204	6.05716E-09	0.150651404	0.000397709	4
33	17	ENSG00000136448	NMT1	-0.05914162	6.3039E-09	-0.096912018	8.32748E-05	4
34	5	ENSG00000069011	PITX1	-0.484148965	6.80739E-09	-0.677146167	2.17144E-06	6
35	5	ENSG00000055163	CYFIP2	-0.303654847	4.2316E-08	-0.270436514	0.00061429	15
36	2	ENSG00000075340	ADD2	-0.26890806	6.52172E-08	-0.430069679	2.58787E-12	3
37	2	ENSG00000071051	NCK2	0.093075437	7.15991E-08	0.137063293	0.001249407	14

***Supplementary table 3.5*- GO (sex-DMPs). Gene Ontology terms identified with GSEA using the differentially methylated positions. Type of GO term- biological process (BP), molecular function (MF), and cellular component (CC), name of GO term, N represents the total number of genes in the GO term, DE represents the number of differentially methylated genes in the GO term, and SigGenesInSet are the differentially methylated genes in the GO term.**

	ONTOLOGY	TERM	N	DE	P.DE	FDR	SigGenesInSet
GO:0009653	BP	anatomical structure morphogenesis	2619	1851	1.2233E-24	2.762E-20	ADA,CDH2,AKT3,NR2E3,AE
GO:0005737	CC	cytoplasm	10520	6654	1.019E-22	1.1504E-18	A1BG,ADA,CDH2,AKT3,AC(
GO:0048856	BP	anatomical structure development	5656	3701	6.517E-22	4.9047E-18	ADA,CDH2,AKT3,NR2E3,AE
GO:0032502	BP	developmental process	6103	3960	7.1954E-21	4.0089E-17	ADA,CDH2,AKT3,NR2E3,AE
GO:0048468	BP	cell development	2084	1489	8.8779E-21	4.0089E-17	CDH2,NR2E3,ABI1,FOXO6,I
GO:0030054	CC	cell junction	1781	1301	3.5487E-20	1.3354E-16	ADA,CDH2,ABI1,CDH5,GJC1
GO:0048731	BP	system development	4657	3068	5.21E-20	1.6805E-16	ADA,CDH2,AKT3,NR2E3,AE
GO:0007275	BP	multicellular organism development	5192	3396	8.3515E-20	2.357E-16	ADA,CDH2,AKT3,NR2E3,AE
GO:0023051	BP	regulation of signaling	3484	2341	3.3747E-19	8.4661E-16	ADA,CDH2,AKT3,CDH3,TAI
GO:0007399	BP	nervous system development	2258	1600	1.9804E-18	4.4713E-15	CDH2,AKT3,NR2E3,ABI1,FC
GO:0010646	BP	regulation of cell communication	3444	2307	4.8352E-18	9.9244E-15	ADA,CDH2,AKT3,CDH3,TAI
GO:0022008	BP	neurogenesis	1557	1134	2.5525E-17	4.8024E-14	CDH2,NR2E3,ABI1,FOXO6,I
GO:0008092	MF	cytoskeletal protein binding	935	710	3E-17	5.2102E-14	CDH2,ABI1,CDH3,KIF28P,B(
GO:0000902	BP	cell morphogenesis	988	756	1.2098E-16	1.9511E-13	CDH2,ABI1,CDH3,CDH4,CD
GO:0030029	BP	actin filament-based process	727	558	2.0336E-16	3.061E-13	ABI1,KCNE3,HCN4,DNAJB6
GO:0042995	CC	cell projection	2083	1461	2.1717E-16	3.0646E-13	ADA,CDH2,ABI1,KCNE3,IQ(
GO:0003824	MF	catalytic activity	5296	3389	2.8679E-16	3.8089E-13	ADA,AKT3,NAALADL1,ACO
GO:0051179	BP	localization	6293	4017	3.2824E-16	4.1172E-13	A1BG,ADA,CDH2,AKT3,CDI
GO:0051239	BP	regulation of multicellular organismal process	3105	2062	4.5519E-16	5.4092E-13	ADA,CDH2,AKT3,KCNE3,CD
GO:0030154	BP	cell differentiation	4007	2622	5.0936E-16	5.7501E-13	ADA,CDH2,NR2E3,ABI1,CD
GO:0050793	BP	regulation of developmental process	2556	1724	8.9731E-16	9.6474E-13	ADA,CDH2,AKT3,CDH3,SRA
GO:0048522	BP	positive regulation of cellular process	5274	3404	1.1966E-15	1.2281E-12	ADA,CDH2,AKT3,NR2E3,CD
GO:0012505	CC	endomembrane system	4160	2711	1.4018E-15	1.3188E-12	A1BG,CDH2,ABI1,TRAPPC3
GO:0048869	BP	cellular developmental process	4081	2663	1.3603E-15	1.3188E-12	ADA,CDH2,AKT3,NR2E3,AE
GO:0048699	BP	generation of neurons	1459	1060	1.7625E-15	1.5917E-12	CDH2,NR2E3,ABI1,FOXO6,I
GO:0120025	CC	plasma membrane bounded cell projection	2009	1407	2.1328E-15	1.8521E-12	ADA,CDH2,ABI1,KCNE3,IQ(
GO:0016043	BP	cellular component organization	5979	3859	2.5444E-15	2.1276E-12	CDH2,AKT3,CDKN2B-AS1,A
GO:0030182	BP	neuron differentiation	1306	958	3.4452E-15	2.778E-12	CDH2,NR2E3,ABI1,FOXO6,I
GO:0035556	BP	intracellular signal transduction	2707	1818	4.7589E-15	3.705E-12	ADA,CDH2,AKT3,TANK,PD(
GO:0009966	BP	regulation of signal transduction	2998	1995	6.8866E-15	5.1829E-12	ADA,CDH2,AKT3,CDH3,TAI
GO:0032989	BP	cellular component morphogenesis	750	580	1.2924E-14	9.2692E-12	CDH2,ABI1,CDH4,RANBP9,
GO:0051128	BP	regulation of cellular component organization	2235	1533	1.3137E-14	9.2692E-12	CDH2,FOXO6,BCL2L11,CDH
GO:0043168	MF	anion binding	2620	1776	1.3659E-14	9.3453E-12	AKT3,KIF28P,SBK3,HCN4,SI
GO:0019899	MF	enzyme binding	2105	1452	2.1925E-14	1.4144E-11	CDH2,TANK,TRAPPC3L,HDA
GO:0030036	BP	actin cytoskeleton organization	633	487	2.1782E-14	1.4144E-11	ABI1,DNAJB6,USH1C,ARPC
GO:0048518	BP	positive regulation of biological process	5798	3697	2.7517E-14	1.7257E-11	ADA,CDH2,AKT3,NR2E3,CD
GO:0070161	CC	anchoring junction	737	555	3.9549E-14	2.4133E-11	CDH2,CDH5,GJC1,PTPRU,A

***Supplementary table 3.6*- GO (sex-DMRs). Gene Ontology terms identified with GSEA using the differentially methylated regions. Type of GO term- biological process (BP), molecular function (MF), and cellular component (CC), name of GO term, N represents the total number of genes in the GO term, DE represents the number of differentially methylated genes in the GO term, and SigGenesInSet are the differentially methylated genes in the GO term.**

	ONTOLOGY	TERM	N	DE	P.DE	FDR	SigGenesInSet
GO:0009653	BP	anatomical structure morphogenesis	2619	1110	6.5317E-14	1.4747E-09	AKT3,NR2E3,ABI1,CDH4,CD
GO:0048856	BP	anatomical structure development	5656	2146	1.9031E-13	2.1485E-09	AKT3,NR2E3,ABI1,FOXO6,E
GO:0032502	BP	developmental process	6103	2283	1.445E-12	1.0875E-08	AKT3,NR2E3,ABI1,FOXO6,E
GO:0006928	BP	movement of cell or subcellular component	2061	851	8.1103E-12	3.8213E-08	AKT3,CDH4,CDH5,MIR1290
GO:0030154	BP	cell differentiation	4007	1544	8.4626E-12	3.8213E-08	NR2E3,ABI1,FOXO6,EPOP,S
GO:0007275	BP	multicellular organism development	5192	1965	1.2122E-11	4.0897E-08	AKT3,NR2E3,ABI1,FOXO6,E
GO:0040011	BP	locomotion	1795	744	1.3514E-11	4.0897E-08	AKT3,CDH4,CDH5,MIR1290
GO:0048731	BP	system development	4657	1781	1.4491E-11	4.0897E-08	AKT3,NR2E3,ABI1,FOXO6,E
GO:0030029	BP	actin filament-based process	727	356	2.3625E-11	5.9266E-08	ABI1,DNAJB6,GJC1,ARPC1E
GO:0032501	BP	multicellular organismal process	7207	2551	3.9509E-11	7.4336E-08	AKT3,NR2E3,ABI1,FOXO6,E
GO:0048468	BP	cell development	2084	893	3.6033E-11	7.4336E-08	NR2E3,ABI1,FOXO6,SH2B3
GO:0048869	BP	cellular developmental process	4081	1562	3.9325E-11	7.4336E-08	AKT3,NR2E3,ABI1,FOXO6,E
GO:0030036	BP	actin cytoskeleton organization	633	313	2.2437E-10	3.8967E-07	ABI1,DNAJB6,ARPC1B,ABI2
GO:0015629	CC	actin cytoskeleton	488	248	2.9666E-10	4.7842E-07	ARPC1B,ADAM8,ACTR1B,M
GO:0003779	MF	actin binding	408	218	3.8411E-10	5.7816E-07	ARPC1B,MAEA,CORO2B,NI
GO:0048870	BP	cell motility	1605	652	8.7339E-10	1.16E-06	AKT3,CDH5,MIR1290,PTPR
GO:0051674	BP	localization of cell	1605	652	8.7339E-10	1.16E-06	AKT3,CDH5,MIR1290,PTPR
GO:0009887	BP	animal organ morphogenesis	1038	467	1.4521E-09	1.8214E-06	AKT3,NR2E3,ABI2,ZMPSTE2
GO:0048513	BP	animal organ development	3407	1301	1.5682E-09	1.8635E-06	AKT3,NR2E3,ABI1,SH2B3,D
GO:0048646	BP	anatomical structure formation involved in morphogenesis	1132	482	1.7895E-09	2.0202E-06	AKT3,ABI1,CDH5,ADAM8,C
GO:0007010	BP	cytoskeleton organization	1251	543	3.6946E-09	3.9722E-06	ABI1,CDH5,PARP3,RANBP9
GO:0009888	BP	tissue development	1971	778	4.0901E-09	4.1975E-06	ABI1,CDH5,GJC1,KRTAP29-
GO:0016477	BP	cell migration	1455	593	5.6513E-09	5.5476E-06	AKT3,CDH5,MIR1290,PTPR
GO:0023051	BP	regulation of signaling	3484	1345	1.1141E-08	1.0481E-05	AKT3,SH2B3,ARHGEF33,CD
GO:0051239	BP	regulation of multicellular organismal process	3105	1191	1.737E-08	1.5687E-05	AKT3,FOXO6,SH2B3,CDH4,
GO:0023052	BP	signaling	6155	2187	2.491E-08	2.1631E-05	AKT3,NR2E3,ABI1,FOXO6,S
GO:0010646	BP	regulation of cell communication	3444	1325	3.0972E-08	2.59E-05	AKT3,SH2B3,ARHGEF33,CD
GO:0071944	CC	cell periphery	5229	1853	4.2732E-08	3.4457E-05	NAALADL1,ABI1,SH2B3,CD
GO:0048522	BP	positive regulation of cellular process	5274	1937	4.7958E-08	3.7338E-05	AKT3,NR2E3,ABI1,FOXO6,S
GO:0048518	BP	positive regulation of biological process	5798	2102	5.7154E-08	4.3014E-05	AKT3,NR2E3,ABI1,FOXO6,S
GO:0072359	BP	circulatory system development	1117	466	6.423E-08	4.678E-05	AKT3,CDH5,GJC1,ADAM8,C
GO:0042995	CC	cell projection	2083	857	8.2738E-08	5.7392E-05	ABI1,IQCJ-SCHIP1,CDH8,CE
GO:0051270	BP	regulation of cellular component movement	1025	430	8.3884E-08	5.7392E-05	AKT3,CDH5,MIR1290,PTPR
GO:0007154	BP	cell communication	6181	2184	9.7832E-08	6.4966E-05	AKT3,NR2E3,ABI1,FOXO6,S
GO:0007155	BP	cell adhesion	1352	564	1.2742E-07	7.9911E-05	SH2B3,CDH4,CDH5,DNAJB6
GO:0035295	BP	tube development	1063	445	1.2447E-07	7.9911E-05	AKT3,CDH5,GJC1,ADAM8,C
GO:0022610	BP	biological adhesion	1358	566	1.3603E-07	8.3006E-05	SH2B3,CDH4,CDH5,DNAJB6

***Supplementary table 3.7*- KEGG (sex-DMPs). Kyoto Encyclopedia of Genes and Genomes pathways identified with GSEA using the differentially methylated positions. Description of KEGG pathway, N represents the total number of genes in the KEGG pathway, DE represents the number of differentially methylated genes in the KEGG pathway, and SigGenesInSet are the differentially methylated genes in the KEGG pathway.**

	Description	N	DE	P.DE	FDR	SigGenesInSet
path:hsa01100	Metabolic pathways	1407	925	2.9946E-10	1.0211E-07	ADA,ACOT8,GNPDA1,ABH
path:hsa04510	Focal adhesion	193	160	1.167E-07	1.9897E-05	AKT3,PAK4,LAMC3,MYL9,V
path:hsa04015	Rap1 signaling pathway	207	165	1.0679E-06	0.00012139	AKT3,RASGRP2,RAPGEF3,V
path:hsa04020	Calcium signaling pathway	227	175	2.5064E-06	0.00021367	TRDN,ADCY1,ADCY2,ADCY
path:hsa04010	MAPK signaling pathway	282	214	3.6941E-06	0.00023941	AKT3,RASGRP1,RASGRP2,C
path:hsa04390	Hippo signaling pathway	153	124	4.2126E-06	0.00023941	PATJ,APC2,YAP1,YWHAQ,F
path:hsa04152	AMPK signaling pathway	117	96	9.1206E-06	0.00037582	AKT3,CAMKK2,CFTR,RAB1(
path:hsa04934	Cushing syndrome	153	122	9.9148E-06	0.00037582	CDK4,CDK6,CDKN1A,CDKN
path:hsa05200	Pathways in cancer	507	361	9.919E-06	0.00037582	AKT3,BCL2L11,FRAT1,RASG
path:hsa04910	Insulin signaling pathway	131	104	1.453E-05	0.00049546	AKT3,SORBS1,SH2B2,PPAR
path:hsa04750	Inflammatory mediator regulation of TRP channels	97	81	3.2367E-05	0.00091975	ADCY1,ADCY2,ADCY3,ADC
path:hsa04935	Growth hormone synthesis, secretion and action	115	94	3.0458E-05	0.00091975	AKT3,ADCY1,ADCY2,ADCY3
path:hsa04919	Thyroid hormone signaling pathway	117	96	4.2515E-05	0.00111521	AKT3,RCAN2,NCOA2,PLCD
path:hsa00562	Inositol phosphate metabolism	69	59	6.8786E-05	0.00167542	CDIPT,PLCD3,PIP5KL1,PIKF
path:hsa04810	Regulation of actin cytoskeleton	207	158	7.959E-05	0.00180934	ARPC5,ARPC4,ARPC1B,ACT
path:hsa04310	Wnt signaling pathway	154	120	0.0001379	0.00289904	FRAT1,APC2,WIF1,FZD10,C
path:hsa04927	Cortisol synthesis and secretion	62	53	0.00015303	0.00289904	ADCY1,ADCY2,ADCY3,ADC
path:hsa04961	Endocrine and other factor-regulated calcium reabsorption	51	45	0.00014505	0.00289904	ADCY6,ADCY9,AP2M1,AP2
path:hsa04960	Aldosterone-regulated sodium reabsorption	36	33	0.00022806	0.00409304	NEDD4L,SFN,HSD11B2,INS
path:hsa01521	EGFR tyrosine kinase inhibitor resistance	77	65	0.00024895	0.00423472	AKT3,BCL2L11,EGF,EGFR,EI
path:hsa04151	PI3K-Akt signaling pathway	336	239	0.00026079	0.00423472	AKT3,BCL2L11,C8orf44-SGI
path:hsa04012	ErbB signaling pathway	80	67	0.00035326	0.00481843	AKT3,CDKN1A,CDKN1B,PA
path:hsa04925	Aldosterone synthesis and secretion	94	76	0.00033383	0.00481843	ADCY1,ADCY2,ADCY3,ADC
path:hsa05100	Bacterial invasion of epithelial cells	76	63	0.00034455	0.00481843	ARPC5,ARPC4,ARPC1B,ACT
path:hsa05225	Hepatocellular carcinoma	161	122	0.00034459	0.00481843	AKT3,FRAT1,CDK4,CDK6,CI
path:hsa04550	Signaling pathways regulating pluripotency of stem cells	139	106	0.00036787	0.00482478	AKT3,COMMD3-BMI1,APC2
path:hsa04360	Axon guidance	175	136	0.00056725	0.00716413	PLXNC1,CDK5,PAK4,SEMA3
path:hsa04728	Dopaminergic synapse	128	99	0.00061643	0.00750721	AKT3,GNB5,ADCY5,COMT,(
path:hsa04926	Relaxin signaling pathway	126	96	0.00065761	0.00773255	AKT3,GNB5,ADCY1,ADCY2,
path:hsa04014	Ras signaling pathway	223	163	0.00070414	0.0080037	AKT3,RASGRP1,RASGRP2,F
path:hsa04066	HIF-1 signaling pathway	105	81	0.00075828	0.00802223	AKT3,CDKN1A,CDKN1B,EG
path:hsa05132	Salmonella infection	239	166	0.00079987	0.00802223	AKT3,ABI1,ARPC5,ARPC4,A
path:hsa05205	Proteoglycans in cancer	199	148	0.00077015	0.00802223	AKT3,CDKN1A,VAV3,FRS2,
path:hsa05226	Gastric cancer	147	111	0.00077829	0.00802223	AKT3,FRAT1,CDH17,CDKN1
path:hsa04928	Parathyroid hormone synthesis, secretion and action	105	83	0.00094386	0.00919593	CDKN1A,GNA13,ADCY1,AD
path:hsa05210	Colorectal cancer	84	67	0.00101241	0.00958973	AKT3,BCL2L11,CDKN1A,AP
path:hsa04670	Leukocyte transendothelial migration	107	82	0.00108912	0.01003754	CDH5,MYL9,RAPGEF3,VAV

***Supplementary table 3.8*- KEGG (sex-DMRs). Kyoto Encyclopedia of Genes and Genomes pathways identified with GSEA using the differentially methylated regions. Description of KEGG pathway, N represents the total number of genes in the KEGG pathway, DE represents the number of differentially methylated genes in the KEGG pathway, and SigGenesInSet are the differentially methylated genes in the KEGG pathway.**

	Description	N	DE	P.DE	FDR	SigGenesInSet
path:hsa04550	Signaling pathways regulating pluripotency of stem cells	139	72	0.00015373	0.025243	AKT3,COMMD3-BMI1,APC
path:hsa04934	Cushing syndrome	153	80	0.00022208	0.025243	CDK6,CDKN1A,CDKN1B,AP
path:hsa05200	Pathways in cancer	507	226	9.5009E-05	0.025243	AKT3,RASGRP1,CDK6,RASC
path:hsa04925	Aldosterone synthesis and secretion	94	54	0.00030211	0.02575482	ADCY1,ADCY2,ADCY3,ADC
path:hsa04020	Calcium signaling pathway	227	107	0.00197697	0.0964287	TRDN,ADCY1,ADCY2,ADCY
path:hsa04927	Cortisol synthesis and secretion	62	36	0.0021813	0.0964287	ADCY1,ADCY2,ADCY3,ADC
path:hsa04950	Maturity onset diabetes of the young	26	16	0.00184293	0.0964287	RFX6,NR5A2,GCK,HHEX,MI
path:hsa05206	MicroRNAs in cancer	291	102	0.00226226	0.0964287	COMMD3-BMI1,CDK6,CDK
path:hsa04015	Rap1 signaling pathway	207	100	0.00461531	0.11605435	AKT3,RASGRP2,RAPGEF3,R
path:hsa04510	Focal adhesion	193	97	0.00369326	0.11605435	AKT3,LAMC3,MYL9,CHAD,(
path:hsa04540	Gap junction	86	44	0.0052757	0.11605435	TUBB3,ADCY1,ADCY2,ADCY
path:hsa04916	Melanogenesis	101	51	0.00443962	0.11605435	ADCY1,ADCY2,ADCY3,ADC
path:hsa04933	AGE-RAGE signaling pathway in diabetic complications	95	49	0.00327644	0.11605435	AKT3,CDKN1B,COL1A1,CO
path:hsa04972	Pancreatic secretion	95	45	0.00514599	0.11605435	ADCY1,ADCY2,CFTR,ADCY3
path:hsa05100	Bacterial invasion of epithelial cells	76	41	0.00520066	0.11605435	ARPC1B,MAD2L2,ARPC1A,
path:hsa05224	Breast cancer	145	69	0.0057857	0.11605435	AKT3,CDK6,CDKN1A,APC2,
path:hsa05225	Hepatocellular carcinoma	161	76	0.00569837	0.11605435	AKT3,CDK6,CDKN1A,APC2,
path:hsa04670	Leukocyte transendothelial migration	107	52	0.00625079	0.1184177	CDH5,MYL9,RAPGEF3,CLDN
path:hsa04330	Notch signaling pathway	52	28	0.00899027	0.13601733	RBPJL,CTBP1,CTBP2,JAG1,I
path:hsa04390	Hippo signaling pathway	153	74	0.00821621	0.13601733	PATJ,APC2,RASSF1,FZD10,
path:hsa04910	Insulin signaling pathway	131	61	0.00917419	0.13601733	AKT3,SORBS1,SH2B2,PPAR
path:hsa04918	Thyroid hormone synthesis	73	37	0.00864225	0.13601733	ADCY1,ADCY2,ADCY3,ADC
path:hsa04935	Growth hormone synthesis, secretion and action	115	58	0.00786654	0.13601733	AKT3,ADCY1,ADCY2,ADCY3
path:hsa04921	Oxytocin signaling pathway	151	73	0.00977877	0.13894005	CDKN1A,MYL9,CAMKK2,AD
path:hsa04810	Regulation of actin cytoskeleton	207	95	0.01322424	0.16742101	ARPC1B,ABI2,APC2,MYL9,E
path:hsa04928	Parathyroid hormone synthesis, secretion and action	105	53	0.01342187	0.16742101	CDKN1A,GNA13,ADCY1,AD
path:hsa04979	Cholesterol metabolism	50	24	0.01374718	0.16742101	OSBPL5,CYP27A1,ABCA1,P
path:hsa05163	Human cytomegalovirus infection	213	91	0.01227909	0.16742101	AKT3,CDK6,CDKN1A,GNA1
path:hsa01521	EGFR tyrosine kinase inhibitor resistance	77	41	0.01671519	0.18999599	AKT3,EGF,EGFR,EIF4E,ERBE
path:hsa04931	Insulin resistance	105	50	0.01649158	0.18999599	AKT3,PPARGC1A,SLC27A3,
path:hsa04152	AMPK signaling pathway	117	56	0.0175286	0.19130725	AKT3,CAMKK2,CFTR,PPAR(
path:hsa05215	Prostate cancer	94	46	0.01795259	0.19130725	AKT3,CDKN1A,CDKN1B,CR
path:hsa05414	Dilated cardiomyopathy	92	47	0.02111628	0.21820157	ADCY1,ADCY2,ADCY3,ADC
path:hsa05165	Human papillomavirus infection	319	134	0.02224275	0.22308165	AKT3,PATJ,CDK6,CDKN1A,
path:hsa04964	Proximal tubule bicarbonate reclamation	21	12	0.02532784	0.24676555	SLC38A3,SLC25A10,GLS,AC
path:hsa04066	HIF-1 signaling pathway	105	48	0.03131437	0.26619624	AKT3,CDKN1A,CDKN1B,EG

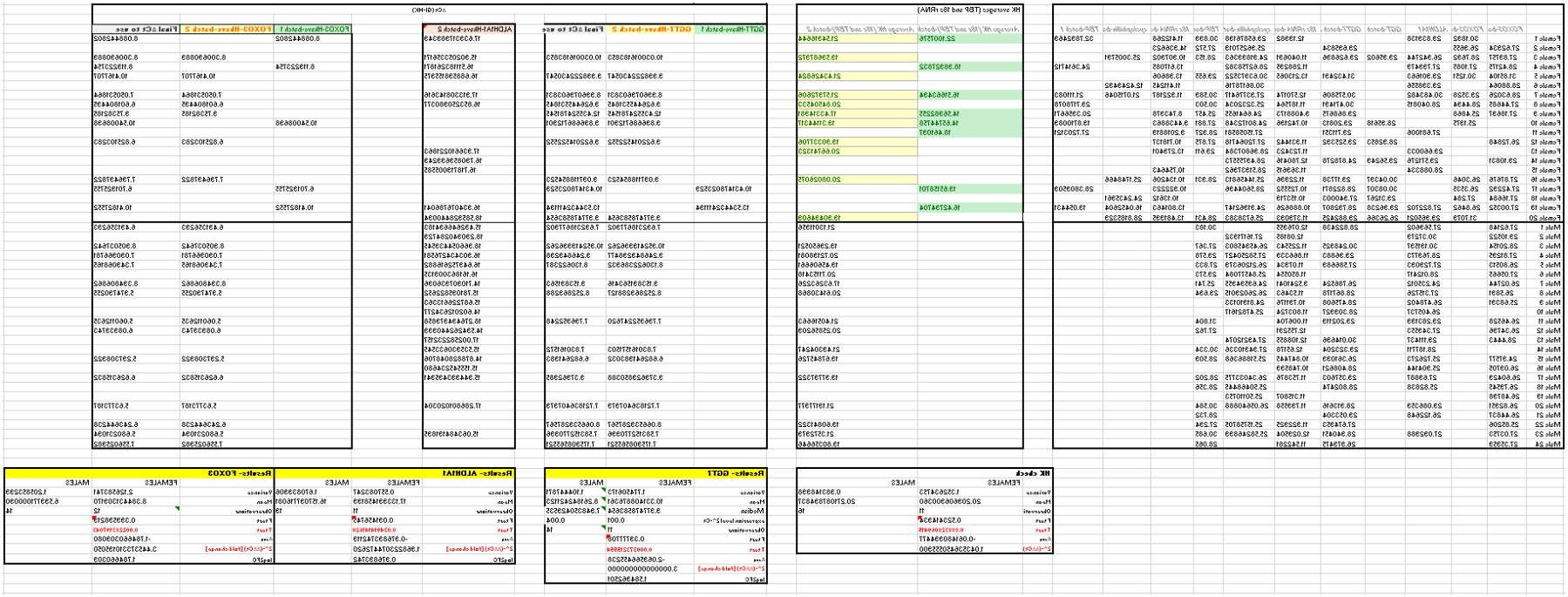
***Supplementary table 3.9*- Reactome (sex-DMPs). Reactome pathways identified with GSEA using the differentially methylated positions associated with sex. Description of Reactome pathway, N represents the total number of genes in the Reactome pathway, DE represents the number of differentially methylated genes in the Reactome pathway, and SigGenesInSet are the differentially methylated genes in the Reactome pathway.**

	Description	N	DE	P.DE	FDR	SigGenesInSet
R-HSA-9006934	Homo sapiens: Signaling by Receptor Tyrosine Kinases	455	345	5.0766E-09	8.7825E-07	AKT3,ABI1,SH2B3,CDH5,RA
R-HSA-1474244	Homo sapiens: Extracellular matrix organization	292	225	8.5503E-08	7.396E-06	ADAM8,ADAM10,LAMC3,F
R-HSA-112316	Homo sapiens: Neuronal System	388	288	1.4551E-06	8.3909E-05	HCN4,GJC1,ABCC9,KCNK7,
R-HSA-397014	Homo sapiens: Muscle contraction	200	155	2.8219E-06	0.00012205	KCNE3,ABCC9,KCNK7,AKAI
R-HSA-5576891	Homo sapiens: Cardiac conduction	135	107	6.9404E-05	0.00240137	KCNE3,ABCC9,KCNK7,AKAI
R-HSA-112315	Homo sapiens: Transmission across Chemical Synapses	255	188	0.00017123	0.00493714	CACNG2,TUBB3,UNC13B,C/
R-HSA-1630316	Homo sapiens: Glycosaminoglycan metabolism	118	91	0.00031488	0.00743888	ABCC5,GPC6,UST,ST3GAL6,
R-HSA-194315	Homo sapiens: Signaling by Rho GTPases	363	257	0.00034399	0.00743888	ABI1,ARHGEF33,ARPC5,AR
R-HSA-194138	Homo sapiens: Signaling by VEGF	103	83	0.00050256	0.00869426	AKT3,ABI1,CDH5,ABI2,WAS
R-HSA-194840	Homo sapiens: Rho GTPase cycle	125	99	0.00050226	0.00869426	ARHGEF33,FAM13A,NET1,I
R-HSA-112314	Homo sapiens: Neurotransmitter receptors and postsynaptic signal transmission	192	141	0.00066009	0.01032403	CACNG2,TUBB3,CAMKK2,G
R-HSA-166520	Homo sapiens: Signaling by NTRKs	99	79	0.00071612	0.01032403	CDK5,FRS2,ADCYAP1R1,AP
R-HSA-5663202	Homo sapiens: Diseases of signal transduction	360	252	0.00093044	0.01238195	AKT3,HDAC5,BCL2L11,AKA
R-HSA-425407	Homo sapiens: SLC-mediated transmembrane transport	239	170	0.00128444	0.01587207	SLC12A7,SLC26A1,SLC22A7
R-HSA-1500931	Homo sapiens: Cell-Cell communication	118	90	0.00148495	0.01605601	CDH2,CDH3,CDH4,CDH5,CI
R-HSA-73887	Homo sapiens: Death Receptor Signalling	129	98	0.00140694	0.01605601	BCL2L11,ARHGEF33,NET1,V
R-HSA-196854	Homo sapiens: Metabolism of vitamins and cofactors	181	126	0.00186416	0.0189706	GPC6,NAMPT,GPHN,SLC19
R-HSA-2219528	Homo sapiens: PI3K/AKT Signaling in Cancer	98	76	0.00244589	0.02143733	AKT3,CDKN1A,CDKN1B,NR
R-HSA-425393	Homo sapiens: Transport of inorganic cations/anions and amino acids/oligopeptides	103	80	0.00253146	0.02143733	SLC12A7,SLC26A1,SLC38A3
R-HSA-71387	Homo sapiens: Metabolism of carbohydrates	278	189	0.00260222	0.02143733	GNPDA1,ABCC5,GPC6,UST,
R-HSA-983712	Homo sapiens: Ion channel transport	175	128	0.00231749	0.02143733	C8orf44-SGK3,ATP9A,SGK2
R-HSA-1474228	Homo sapiens: Degradation of the extracellular matrix	134	99	0.00275814	0.02168901	ADAM8,ADAM10,CAPN9,M
R-HSA-76002	Homo sapiens: Platelet activation, signaling and aggregation	247	172	0.0030883	0.02322938	A1BG,RASGRP1,LHFPL2,RA
R-HSA-196849	Homo sapiens: Metabolism of water-soluble vitamins and cofactors	117	83	0.00656237	0.04366503	NAMPT,GPHN,SLC19A2,MT
R-HSA-199418	Homo sapiens: Negative regulation of the PI3K/AKT network	107	80	0.00653086	0.04366503	AKT3,NRG3,FRS2,PIK3AP1,
R-HSA-6811558	Homo sapiens: PI5P, PP2A and IER3 Regulate PI3K/AKT Signaling	100	75	0.00606183	0.04366503	NRG3,FRS2,PIK3AP1,NRG4
R-HSA-1483257	Homo sapiens: Phospholipid metabolism	191	136	0.00735923	0.0471536	LPCAT3,PEMT,CDIPT,AGPA
R-HSA-6811442	Homo sapiens: Intra-Golgi and retrograde Golgi-to-ER traffic	197	133	0.01216919	0.07518819	TMED7-TICAM2,KIF20A,AC
R-HSA-948021	Homo sapiens: Transport to the Golgi and subsequent modification	181	122	0.01460228	0.08711017	TMED7-TICAM2,PREB,ACTF
R-HSA-9006936	Homo sapiens: Signaling by TGF-beta family members	99	70	0.01519613	0.08763099	CDK9,FSTL3,STUB1,FST,FST
R-HSA-373760	Homo sapiens: L1CAM interactions	110	81	0.01869816	0.10434779	RANBP9,TUBB3,CHL1,LYPLA
R-HSA-5173105	Homo sapiens: O-linked glycosylation	107	78	0.0198551	0.10600554	MUC12,SPON2,SPON1,POM
R-HSA-983231	Homo sapiens: Factors involved in megakaryocyte development and platelet production	144	99	0.02022071	0.10600554	SH2B3,KIF20A,CDK5,TUBB3
R-HSA-195258	Homo sapiens: RHO GTPase Effectors	248	165	0.02443868	0.11294389	ABI1,ARPC5,ARPC4,ARPC1
R-HSA-195721	Homo sapiens: Signaling by WNT	261	178	0.02676705	0.11294389	FRAT1,PSMD14,KAT5,GNB5
R-HSA-199977	Homo sapiens: ER to Golgi Anterograde Transport	150	100	0.02494054	0.11294389	TMED7-TICAM2,PREB,ACTF
R-HSA-3247509	Homo sapiens: Chromatin modifying enzymes	183	128	0.02518966	0.11294389	CDK4,MCRS1,TADA3,NCOA

***Supplementary table 3.10*- Reactome (sex-DMRs). Reactome pathways identified with GSEA using the differentially methylated regions. Description of Reactome pathway, N represents the total number of genes in the Reactome pathway, DE represents the number of differentially methylated genes in the Reactome pathway, and SigGenesInSet are the differentially methylated genes in the Reactome pathway.**

	Description	N	DE	P.DE	FDR	SigGenesInSet
R-HSA-397014	Homo sapiens: Muscle contraction	200	103	1.5781E-05	0.00273013	KCNK7,TRDN,MYL9,SORBS
R-HSA-1474244	Homo sapiens: Extracellular matrix organization	292	138	0.00031499	0.0272467	ADAM8,LAMC3,P3H3,ADAI
R-HSA-5576891	Homo sapiens: Cardiac conduction	135	68	0.00233597	0.10103068	KCNK7,TRDN,NKX2-5,FGF1
R-HSA-9006934	Homo sapiens: Signaling by Receptor Tyrosine Kinases	455	202	0.0017744	0.10103068	AKT3,ABI1,SH2B3,CDH5,RA
R-HSA-194138	Homo sapiens: Signaling by VEGF	103	55	0.00329298	0.11393712	AKT3,ABI1,CDH5,ABI2,BAIA
R-HSA-1474228	Homo sapiens: Degradation of the extracellular matrix	134	64	0.0053537	0.15436488	ADAM8,ADAMTS8,COL1A1
R-HSA-194840	Homo sapiens: Rho GTPase cycle	125	64	0.0080302	0.19846071	ARHGEF33,NET1,GNA13,AH
R-HSA-6785807	Homo sapiens: Interleukin-4 and Interleukin-13 signaling	103	45	0.01523086	0.32936744	CDKN1A,AKT1,F13A1,FGF2
R-HSA-112316	Homo sapiens: Neuronal System	388	167	0.02191037	0.37904948	GJC1,KCNK7,TUBB3,CAMKI
R-HSA-73887	Homo sapiens: Death Receptor Signalling	129	60	0.02038392	0.37904948	ARHGEF33,NET1,GNA13,AH
R-HSA-112314	Homo sapiens: Neurotransmitter receptors and postsynaptic signal transmission	192	83	0.0382666	0.38120157	TUBB3,CAMKK2,GNB5,ADC
R-HSA-112315	Homo sapiens: Transmission across Chemical Synapses	255	111	0.02826986	0.38120157	TUBB3,CAMKK2,GNB5,ADC
R-HSA-1500931	Homo sapiens: Cell-Cell communication	118	55	0.03187647	0.38120157	CDH4,CDH5,CDH8,CDH13,C
R-HSA-194315	Homo sapiens: Signaling by Rho GTPases	363	149	0.03716246	0.38120157	ABI1,ARHGEF33,ARPC1B,A
R-HSA-202733	Homo sapiens: Cell surface interactions at the vascular wall	121	51	0.02804653	0.38120157	MERTK,PROCR,COL1A1,CX
R-HSA-416476	Homo sapiens: G alpha (q) signalling events	204	78	0.03966259	0.38120157	RASGRP1,RASGRP2,RGS19,
R-HSA-71387	Homo sapiens: Metabolism of carbohydrates	278	107	0.03793683	0.38120157	UST,ST3GAL6,CSPG5,B3GN
R-HSA-76002	Homo sapiens: Platelet activation, signaling and aggregation	247	101	0.03927305	0.38120157	A1BG,RASGRP1,LHFPL2,RA
R-HSA-1483257	Homo sapiens: Phospholipid metabolism	191	81	0.04507191	0.41039164	PEMT,STARD10,CHKB,MGL

***Supplementary table 3.11*- PCRs. Results from qPCR of *GGT7, FOXO3,* and *ALDH1A1* in Gene SMART cohort. Blank cells indicate sample that had a standard deviation of greater than 1 Ct between triplicates. Housekeeping genes tested are Cyclophillin, 18s rRNA, and TBP. Primer sequences and PCR conditions used.**



Gene	Sequence (5'->3')		Annealing Temperature	F&R concentrations	Purchased from
GGT7	Forward primer	CCTTGTGTTTGGGTATCGTGG	58	10uM	Integrated DNA technologies
	Reverse primer	TTCGTCGGATGTCATGTACCA			
FOXO3	Forward primer	TGGTTTGAACGTGGGGAACT	58	10uM	Integrated DNA technologies
	Reverse primer	TGTCAGTTTGAGGGTCTGCT			
ALDH1A1	Forward primer	GGCCCTCAGATTGACAAGGA	60	5uM	Integrated DNA technologies
	Reverse primer	ATGATTTGCTGCACTGGTCC			
TBP	Forward primer	CAGTGACCCAGCAGCATCACT	58	10uM	Integrated DNA technologies
	Reverse primer	AGGCCAAGCCCTGAGCGTAA			
18s rRNA	Forward primer	CTTAGAGGGACAAGTGGCG	58	10uM	Integrated DNA technologies
	Reverse primer	GGACATCTAAGGGCATCACA			

***Supplementary table 3.12*- Gene expression and DNA methylation differences between males and females for three genes across the cohorts used in the analysis.**

	Gene expression effect size between males and females			
	GTEx (MASH posterior effect size)	FUSION (Fold change)	Gene SMART (Fold change)	Mean effect size of DMR showing largest effect size (% DNA methylation difference between males and females)
GGT7	0.81	1.6	3.0	-20.4
FOXO3	-0.04	-0.2	3.4	-1.9
ALDH1A1	0.33	0.4	2.0	-10

***Supplementary table 4.1*- Data available for each of the datasets included in the DNA methylation meta-analysis. Immunohistochemistry (IHC); sex hormone-binding globulin (SHBG), free testosterone (Free T), testosterone (T), estradiol (E2)**

	DNA methylation	**mRNA expression**	**Fibre proportions**	**Circulating hormones**
FUSION	EPIC	Transcriptomics (RNA-seq)	Derived from RNA-seq	x
Gene SMART	EPIC	3 genes (qPCR)	Derived from IHC	SHBG, Free Testosterone, Testosterone, Estradiol
GSE38291	27K	x	x	x

***Supplementary table 4.2*- List of transcription factors (TFs) included in analysis for enrichment of transcription factor binding sites (TFBSs) among differentially methylated positions (DMPs). The current UniBind database tests a total of 268 unique TFs from 518 different cell types.**

userSet	dbSet	collection	pValueLog	oddsRatio	pValue	support	rnkPV	rnkOR	rnkSup	maxRnk	meanRnk	b	c	d	descriptio	cellType	tissue	antibody	treatment	dataSourc	filename	size
1	32	AR	86.06216818	4.130761	8.66626E-87	344	1	138	16	138	51.7	899	56339	608140	AR HEK29	HEK293 (embryonic kidney)	NULL	AR	10 nM R1	GTRD	EXP03287	85213
1	156	AR	81.81043301	4.162766	1.54727E-82	324	2	135	19	135	52	840	56359	608199	AR PC3(bc	PC3(bone mets prostate adenocarcinoma)	NULL	AR	R1881	GEO	GSE54110	60594
1	3091	NR3C1	78.91466377	4.469229	1.21713E-79	290	3	114	34	114	50.3	700	56393	608339	NR3C1 U2	U2OS (osteosarcoma)	BTO:0001	NR3C1	GLUCC	Array Expr	ERP00708	44490
1	33	AR	76.90089422	3.942277	1.25634E-77	323	4	156	20	156	60	884	56360	608155	AR HEK29	HEK293 (embryonic kidney)	NULL	AR	10 nM R1	GTRD	EXP03287	76249
1	3098	NR3C1	72.25690657	5.295997	5.53469E-73	226	5	81	79	81	55	460	56457	608579	NR3C1 em	embryonic kidney	NULL	NR3C1	dexameth	GTRD	EXP03221	31363
1	3154	NR3C1	68.98854148	3.796618	1.02674E-69	302	6	171	29	171	68.7	858	56381	608181	NR3C1 U2	U2OS (osteosarcoma)	BTO:0001	NR3C1	SHHIC5	GEO	GSE65847	50363
1	3155	NR3C1	68.06148104	7.639468	8.67998E-69	161	7	35	294	294	112	227	56522	608812	NR3C1 U2	U2OS (osteosarcoma)	BTO:0001	NR3C1	SHNS	GEO	GSE65847	15065
1	163	AR	65.7999573	2.480988	1.58505E-66	540	8	498	2	498	169	2352	56143	606687	AR LHSAR	LHSAR (prostate epithelial cells)	NULL	AR	LACZ	GEO	GSE56288	118921
1	3121	NR3C1	64.98476252	4.866651	1.03571E-65	219	9	91	94	94	64.7	485	56464	608554	NR3C1 A5	A549 (lung carcinoma)	BTO:0000	NR3C1	100nM de	GTRD	EXP04027	28099
1	3120	NR3C1	64.04689803	5.515904	8.9764E-65	193	10	59	147	147	72	377	56490	608662	NR3C1 A5	A549 (lung carcinoma)	BTO:0000	NR3C1	100nM de	GTRD	EXP04027	22096
1	3119	NR3C1	61.22366743	4.747005	5.97493E-62	211	11	98	113	113	74	479	56472	608560	NR3C1 A5	A549 (lung carcinoma)	BTO:0000	NR3C1	100nM de	GTRD	EXP04022	24200
1	164	AR	59.70767676	2.701949	1.9603E-60	420	12	401	8	401	140	1678	56263	607361	AR PLHSA	PLHSAR	NA	AR	FOXA1_H	GEO	GSE56288	165694
1	3118	NR3C1	58.38607724	5.101732	4.11077E-59	188	13	87	168	168	89.3	397	56495	608642	NR3C1 A5	A549 (lung carcinoma)	BTO:0000	NR3C1	100nM de	GTRD	EXP04022	21901
1	146	AR	57.53509188	5.93092	2.91681E-58	163	14	52	274	274	113	296	56520	608743	AR DU145	DU145 (prostate carcinoma)	BTO:0001	AR	ARQ6540X	GEO	GSE47987	17439
1	149	AR	55.94460358	8.306696	1.13605E-56	125	15	25	560	560	200	162	56558	608877	AR DU145	DU145 (prostate carcinoma)	BTO:0001	AR		GEO	GSE47987	11076
1	160	AR	51.58661252	3.443206	2.59052E-52	253	16	222	55	222	97.7	792	56430	608247	AR VCaP (	VCaP (prostate carcinoma)	BTO:0003	AR	SHPIAS1_I	GEO	GSE56086	83992
1	3097	NR3C1	51.30858243	4.89829	4.9138E-52	171	17	90	236	236	114	376	56512	608663	NR3C1 em	embryonic kidney	NULL	NR3C1	dexameth	GTRD	EXP03221	20349
1	7	AR	48.19741898	4.308703	6.34718E-49	182	18	125	197	197	113	455	56501	608584	AR VCaP (	VCaP (prostate carcinoma)	BTO:0003	AR	AR stimul	GTRD	EXP00040	43913
1	3095	NR3C1	48.05838476	4.651239	8.74209E-49	168	19	102	248	248	123	389	56515	608650	NR3C1 LN	LNCaP (prostate carcinoma)	BTO:0001	NR3C1	100 nM DH	GTRD	EXP03050	29935
1	3092	NR3C1	47.71820251	4.382697	1.91336E-48	177	20	120	210	210	117	435	56506	608604	NR3C1 IM	IMR90 (lung fibroblasts)	NULL	NR3C1		Array Expr	ERP00709	28830
1	3090	NR3C1	46.34477475	4.008013	4.5209E-47	189	21	148	163	163	111	508	56494	608531	NR3C1 IM	IMR90 (lung fibroblasts)	NULL	NR3C1	GLUCC	Array Expr	ERP00708	28403
1	3094	NR3C1	46.15779462	4.627493	6.95353E-47	162	22	103	283	283	136	377	56521	608662	NR3C1 LN	LNCaP (prostate carcinoma)	BTO:0001	NR3C1	100 nM De	GTRD	EXP03050	24559
1	3102	NR3C1	46.07866491	6.103095	8.34325E-47	127	23	50	540	540	204	224	56556	608815	NR3C1 A5	A549 (lung carcinoma)	BTO:0000	NR3C1	etoh	GTRD	EXP03623	11080
1	161	AR	45.71610141	2.967517	1.92264E-46	275	24	326	41	326	130	999	56408	608040	AR VCaP (	VCaP (prostate carcinoma)	BTO:0003	AR		GEO	GSE56086	88405
1	3152	NR3C1	44.96948767	6.291617	1.07278E-45	121	25	48	601	601	225	207	56562	608832	NR3C1 LN	LNCaP (prostate carcinoma)	BTO:0001	NR3C1	1F5_SIFO	GEO	GSE30623	14366
1	3122	NR3C1	44.90898859	4.365333	1.23314E-45	167	26	121	253	253	133	412	56516	608627	NR3C1 A5	A549 (lung carcinoma)	BTO:0000	NR3C1	100nM de	GTRD	EXP04027	17961
1	3115	NR3C1	43.70353043	6.327784	1.97911E-44	117	27	47	639	639	238	199	56566	608840	NR3C1 A5	A549 (lung carcinoma)	BTO:0000	NR3C1	100nM de	GTRD	EXP03979	11362
1	213	AR	42.36519823	3.426144	4.31322E-43	208	28	224	115	224	122	654	56475	608385	AR VCaP (	VCaP (prostate carcinoma)	BTO:0003	AR	R1881_4H	GEO	GSE84432	55427
1	99	AR	41.23125427	4.110906	5.87145E-42	163	29	139	274	274	147	427	56520	608612	AR LTAD (	LTAD (long term androgen deprivation)	NULL	AR	Treatmen	GTRD	EXP04945	41997
1	205	AR	39.54509787	4.976998	2.85038E-40	129	30	89	519	519	213	279	56554	608760	AR VCaP (	VCaP (prostate carcinoma)	BTO:0003	AR	SH1_R188	GEO	GSE79128	22606
1	207	AR	39.2091216	4.662572	6.17843E-40	136	31	101	460	460	197	314	56547	608725	AR VCaP (	VCaP (prostate carcinoma)	BTO:0003	AR	SH2_R188	GEO	GSE79128	25393
1	212	AR	38.84485747	3.250553	1.42936E-39	204	32	252	120	252	135	676	56479	608363	AR VCaP (	VCaP (prostate carcinoma)	BTO:0003	AR	R1881_30	GEO	GSE84432	53687
1	3111	NR3C1	37.12333802	11.07326	7.52769E-38	70	33	9	1391	1391	478	68	56613	608971	NR3C1 A5	A549 (lung carcinoma)	BTO:0000	NR3C1	100nM de	GTRD	EXP03968	4624
1	21	AR	37.06841844	3.305747	8.54243E-38	190	34	239	159	239	144	619	56493	608420	AR VCaP (	VCaP (prostate carcinoma)	BTO:0003	AR	100 nM DH	GTRD	EXP03050	59512
1	3139	NR3C1	36.7992457	3.155845	1.58765E-37	201	35	279	125	279	146	686	56482	608353	NR3C1 HA	HASM2 (Human airway smooth muscle 2)	NULL	NR3C1	Treatmen	GTRD	EXP04965	39106
1	206	AR	35.92634044	4.837189	1.18484E-36	120	36	92	606	606	245	267	56563	608772	AR VCaP (	VCaP (prostate carcinoma)	BTO:0003	AR	SH2_DHT	GEO	GSE79128	23313

***Supplementary table 4.3*- List of differentially methylated positions (DMPs) associated with type I fibre content (FDR < 0.005) across the sex-DMPs. Corresponding chromosome, genomic location, annotated genes, and genes annotated by GeneHancer. Positive effect size indicates higher DNA methylation in type I fibres compared to type II fibres.**

	CpG	Effect	StdErr	P.value	Direction	HetISq	HetChiSq	HetDf	HetPVal	t_stat	BetaEffectSize	FDR	CpG_beg	CpG_end	genesUniq
cg14600987	cg14600987	1.7133	0.1441	1.349E-32	++++	0	1.679	3	0.6417	11.88966	0.213404	7.6641E-28	1931447	1931449	TNNT3
cg22495058	cg22495058	1.1989	0.1043	1.472E-30	++++	0	0.134	3	0.9874	11.4947267	0.160886	4.1814E-26	11612246	11612248	GREB1
cg08198488	cg08198488	1.5055	0.1336	1.9E-29	++++	0	2.971	3	0.3961	11.2687126	0.249098	3.5982E-25	63041592	63041594	TPM1
cg18877271	cg18877271	1.426	0.1331	8.783E-27	++++	0	1.137	3	0.7681	10.7137491	0.220154	1.2475E-22	1931407	1931409	TNNT3
cg12114524	cg12114524	-1.4655	0.1401	1.303E-25	----	0	1.095	3	0.7784	-10.4603854	-0.247423	1.4805E-21	201421665	201421667	TNNI1
cg21026396	cg21026396	1.7007	0.1636	2.527E-25	++++	43.7	5.326	3	0.1494	10.3954768	0.129867	2.3928E-21	63049500	63049502	TPM1;RP11-244F12.3
cg05408442	cg05408442	1.3935	0.1352	6.513E-25	++++	0	0.338	3	0.9527	10.3069527	0.209616	5.286E-21	24365429	24365431	STPG1
cg26946015	cg26946015	1.7567	0.1739	5.451E-24	++++	0	0.331	3	0.954	10.1017826	0.286036	3.8711E-20	66144515	66144517	
cg21108554	cg21108554	1.5693	0.1563	1.004E-23	++++	0	0.319	3	0.9564	10.0403071	0.263031	6.3378E-20	150364238	150364240	RP4-584D14.5
cg13435855	cg13435855	-1.3377	0.1334	1.125E-23	----	0	2.695	3	0.441	-10.0277361	-0.226265	6.3915E-20	201421077	201421079	TNNI1
cg09993645	cg09993645	1.7676	0.1769	1.668E-23	++++	0	0.825	3	0.8434	9.99208592	0.29255	8.6149E-20	184823646	184823648	FAM129A
cg20980993	cg20980993	2.0115	0.2025	3.01E-23	++++	18.3	3.672	3	0.2991	9.93333333	0.301804	1.4251E-19	43785721	43785723	BLVRA
cg10423149	cg10423149	1.6732	0.1693	4.799E-23	++++	0	1.766	3	0.6224	9.88304784	0.280315	2.0973E-19	62179590	62179592	PTPRG
cg06535121	cg06535121	1.6028	0.1626	6.302E-23	++++	58.2	7.18	3	0.06637	9.85731857	0.156415	2.5574E-19	26088488	26088490	
cg10628683	cg10628683	1.2309	0.1261	1.706E-22	++++	0	0.947	3	0.814	9.76130056	0.195207	6.4615E-19	45823881	45823883	TNNC2
cg07629519	cg07629519	1.6563	0.17	1.979E-22	++++	0	0.905	3	0.8242	9.74294118	0.265875	7.0271E-19	86282859	86282861	REEP1
cg13081262	cg13081262	-1.3465	0.1397	5.343E-22	----	0	0.842	3	0.8394	-9.6385111	-0.220162	1.7856E-18	23433654	23433656	MYH7
cg24325991	cg24325991	1.3637	0.1419	7.095E-22	++++	0	1.575	3	0.6652	9.61028894	0.067097	2.2394E-18	184974778	184974780	FAM129A
cg00501105	cg00501105	-2.0235	0.2115	1.095E-21	----	11	3.37	3	0.338	-9.56737589	-0.283723	3.2742E-18	22037048	22037050	CDKN2B-AS1
cg15155209	cg15155209	-1.2281	0.1287	1.389E-21	----	0	0.356	3	0.9492	-9.54234654	-0.143789	3.9457E-18	14460768	14460770	SLC6A6
cg03418649	cg03418649	1.516	0.1601	2.755E-21	++++	0	1.935	3	0.5859	9.46908182	0.243421	7.3108E-18	1924737	1924739	TNNT3
cg13952079	cg13952079	1.2783	0.135	2.831E-21	++++	0	1.265	3	0.7375	9.46888889	0.168831	7.3108E-18	237223327	237223329	
cg15871435	cg15871435	1.9256	0.2045	4.701E-21	++++	0	0.826	3	0.8433	9.41613692	0.282115	1.1612E-17	23549393	23549395	KLHL29
cg20115418	cg20115418	1.6995	0.1826	1.327E-20	++++	0	2.342	3	0.5045	9.30722892	0.205369	3.0156E-17	197838125	197838127	PLCL1
cg17417147	cg17417147	1.9397	0.2084	1.323E-20	++++	0	2.194	3	0.5331	9.30758157	0.281906	3.0156E-17	66565532	66565534	CTSF
cg17147440	cg17147440	1.3489	0.1458	2.209E-20	++++	1.8	3.055	3	0.3832	9.25171468	0.201403	4.8269E-17	63045152	63045154	TPM1
cg19615017	cg19615017	1.5796	0.171	2.482E-20	++++	0	0.734	3	0.8651	9.2374269	0.258344	5.2226E-17	30631966	30631968	TGFBR2
cg05388410	cg05388410	1.4183	0.1542	3.582E-20	++++	0	2.324	3	0.508	9.19779507	0.202838	6.8838E-17	63041355	63041357	TPM1
cg02785006	cg02785006	1.5055	0.1637	3.635E-20	++++	0	1.798	3	0.6154	9.19670128	0.230599	6.8838E-17	100542905	100542907	RP11-151A6.4;GGACT
cg09156538	cg09156538	1.1979	0.1301	3.432E-20	++++	0	2.868	3	0.4124	9.20753267	0.202576	6.8838E-17	55904359	55904361	IL31RA
cg07358846	cg07358846	1.4394	0.1567	4.027E-20	++++	0	2.607	3	0.4563	9.18570517	0.237039	7.3802E-17	111951872	111951874	KCND3
cg20404622	cg20404622	-1.5055	0.164	4.299E-20	----	0	2.049	3	0.5623	-9.17987805	-0.24632	7.6325E-17	75073958	75073960	AC111186.1
cg04052013	cg04052013	1.1348	0.1238	4.912E-20	++++	0	0.637	3	0.8879	9.16639742	0.189852	8.4565E-17	1938015	1938017	TNNT3
cg19642007	cg19642007	1.294	0.1413	5.183E-20	++++	0	2.005	3	0.5715	9.15782024	0.218633	8.6031E-17	1923233	1923235	TNNT3
cg19387862	cg19387862	1.0813	0.1181	5.3E-20	++++	0	0.708	3	0.8713	9.15580017	0.182449	8.6031E-17	50428617	50428619	SPIB
cg21410048	cg21410048	1.4561	0.1591	5.693E-20	++++	0	1.016	3	0.7973	9.15210559	0.196303	8.9843E-17	23541928	23541930	KLHL29
cg27412039	cg27412039	1.328	0.1452	6.042E-20	++++	36.7	4.74	3	0.1919	9.14600551	0.214343	9.2774E-17	1926370	1926372	TNNT3

***Supplementary table 4.4*- List of differentially methylated positions (DMPs) associated with type I fibre content (FDR < 0.005) across all tested CpGs. Corresponding chromosome, genomic location, annotated genes, and genes annotated by GeneHancer. Positive effect size indicates higher DNA methylation in type I fibres compared to type II fibres.**

	CpG	Effect	StdErr	P.value	Direction	HetISq	HetChiSq	HetDf	HetPVal	t_stat	BetaEffectSize	FDR	CpG_beg	CpG_end	genesUniq
cg11801374	cg11801374	-0.4839	0.1083	7.945E-06	----	0	2.203	3	0.5314	-4.46814404	-6.2541	0.00061298	32009890	32009892	CCM2L
cg11405300	cg11405300	0.8337	0.135	6.557E-10	++++	0	0.479	3	0.9235	6.17555556	13.8643	2.6272E-07	133952151	133952153	IGSF9B
cg03983220	cg03983220	0.6672	0.1625	0.00004033	++++	0	1.189	3	0.7556	4.10584615	7.5176	0.00214487	85534800	85534802	
cg10133462	cg10133462	-0.7587	0.1709	8.983E-06	----	0	2.229	3	0.5262	-4.43943827	-7.188	0.000674	1311324	1311326	MAEA
cg07281764	cg07281764	0.6651	0.1583	0.00002666	++++	32.4	4.436	3	0.218	4.20151611	10.9671	0.00156524	54167122	54167124	SMUG1;RP11-834C11.8
cg06440804	cg06440804	-0.4796	0.0974	8.46E-07	----	0	1.492	3	0.6841	-4.92402464	-7.4781	0.00010221	76145326	76145328	RP11-449J21.5;RNF157
cg04067992	cg04067992	-1.0572	0.1775	2.591E-09	----	0	1.931	3	0.5869	-5.95605634	-10.3819	8.5414E-07	110313076	110313078	COL4A2
cg10543500	cg10543500	1.755	0.2615	1.923E-11	++++	34.8	4.6	3	0.2035	6.71128107	16.4125	1.2935E-08	44161399	44161401	SMIM2;SMIM2-AS1
cg25356611	cg25356611	-0.5735	0.1026	2.253E-08	----	7.6	3.245	3	0.3553	-5.58966862	0.8303	5.3013E-06	1668295	1668297	
cg23212050	cg23212050	1.2167	0.2409	4.385E-07	++++	14.4	3.506	3	0.32	5.05064342	13.044	6.007E-05	9390098	9390100	
cg22028621	cg22028621	0.5943	0.1473	0.00005454	++++	0	1.105	3	0.7758	4.03462322	10.1351	0.00271094	150274586	150274588	
cg04118910	cg04118910	0.6527	0.1446	6.393E-06	++++	0	1.33	3	0.722	4.51383126	10.4563	0.00051815	114109694	114109696	RASA3-IT1;RASA3
cg13473691	cg13473691	0.4434	0.1087	0.00004518	++++	22.6	3.877	3	0.275	4.07911684	7.6038	0.00234074	135779803	135779805	KCNT1
cg03052074	cg03052074	0.5498	0.1143	1.513E-06	++++	0	2.688	3	0.4423	4.81014873	7.7538	0.00016348	40969518	40969520	ELF1
cg06875255	cg06875255	0.6579	0.1297	3.909E-07	++++	0	1.028	3	0.7944	5.07247494	10.7453	5.4628E-05	71560696	71560698	DYSF
cg07242890	cg07242890	-0.5635	0.1267	8.613E-06	----	0	1.087	3	0.7802	-4.44751381	-7.5944	0.00065272	55534935	55534937	ERC2
cg08048699	cg08048699	-0.4284	0.1109	0.000112	---+	52.6	6.333	3	0.0965	-3.86293959	-5.2781	0.00467975	57528831	57528833	TIMM10;RN7SL605P
cg12902237	cg12902237	0.5251	0.1221	0.000017	++++	0	1.143	3	0.7668	4.3005733	8.4844	0.0011055	49765004	49765006	
cg25262105	cg25262105	0.8915	0.1953	5.014E-06	++++	0	2.732	3	0.4348	4.56477215	11.9708	0.0004268	2319505	2319507	LLfos-48D6.2
cg02052154	cg02052154	0.4787	0.104	4.204E-06	++++	0	1.607	3	0.6578	4.60288462	6.7579	0.00037025	128077663	128077665	RUVBL1-AS1;RUVBL1
cg16911296	cg16911296	0.9836	0.1691	6.01E-09	++++	0	1.604	3	0.6584	5.81667652	16.3796	1.7333E-06	23691041	23691043	KLHL29
cg09803359	cg09803359	0.6336	0.1578	0.0000596	++++	0	2.742	3	0.4331	4.01520913	10.09	0.00289735	125177535	125177537	NSMCE2
cg06699201	cg06699201	-0.4732	0.1189	0.00006869	----	0	2.039	3	0.5644	-3.97981497	-6.3809	0.00322961	9081793	9081795	A2M
cg00366716	cg00366716	0.9527	0.2422	0.00008339	++++	33.3	4.499	3	0.2124	3.93352601	0.1138	0.00374139	128021438	128021440	SND1
cg05376491	cg05376491	-0.6468	0.122	1.157E-07	----	0	1.582	3	0.6635	-5.30163934	-7.6441	2.0116E-05	99340443	99340445	
cg01591190	cg01591190	0.7988	0.1434	2.563E-08	++++	5.5	3.176	3	0.3653	5.57043236	12.436	5.8812E-06	25348029	25348031	
cg21069500	cg21069500	0.5489	0.135	0.00004774	++++	0	2.025	3	0.5672	4.06592593	8.8752	0.00244465	139670573	139670575	CXXC5
cg26312108	cg26312108	1.3769	0.2339	3.95E-09	++++	0	1.79	3	0.6172	5.88670372	16.4832	1.2116E-06	95989729	95989731	RP11-759A24.3
cg07651766	cg07651766	0.6747	0.1478	4.968E-06	++++	0	2.046	3	0.5629	4.56495264	11.4678	0.00042337	222887083	222887085	DISP1;RP11-455P21.3
cg00854448	cg00854448	0.6007	0.1235	1.152E-06	++++	0	0.522	3	0.9141	4.86396761	8.3488	0.00013149	205319748	205319750	NUAK2
cg06287058	cg06287058	0.6094	0.1371	8.761E-06	++++	61	7.702	3	0.05259	4.44493071	3.8424	0.00066145	5957448	5957450	RANBP3
cg24418589	cg24418589	1.0297	0.1496	5.785E-12	++++	67.6	9.256	3	0.02607	6.88302139	16.0308	4.4742E-09	28029349	28029351	TTC28
cg02694847	cg02694847	0.5459	0.1113	9.292E-07	++++	0	1.262	3	0.7383	4.9047619	9.0613	0.00011018	44062472	44062474	DBNL
cg21456212	cg21456212	-0.7929	0.18	0.00001053	----	0	2.266	3	0.519	-4.405	-6.334	0.000763	112705521	112705523	ATP11A
cg14384960	cg14384960	0.8073	0.19	0.00002137	++++	0	0.354	3	0.9495	4.24894737	10.4035	0.00131787	145886138	145886140	ZNF827
cg01085421	cg01085421	-0.6852	0.1531	7.579E-06	----	0	1.788	3	0.6177	-4.47550621	-5.4241	0.00059083	55055517	55055519	

***Supplementary table 5.1*- List of differentially methylated regions (DMRs) associated with training (time) with sex as a covariate (FDR < 0.05). Corresponding chromosome, genomic location, annotated genes, and genes annotated by GeneHancer. Positive effect size indicates higher DNA methylation in post four weeks of HIIT compared to pre.**

	logFC	AveExpr	t	P.Value	adj.P.Val	B	BetaEffectSize	CpG	CpG_chrm	CpG_beg	CpG_end	probe_stran	genesUniq	geneNames	transcriptTy	transcriptID	distToTSS	CGI	CGIposition	E107	E108	GeneHancer	genesUniq_with_enh
1	0.83615867	-5.8120822	6.14326576	8.8297E-09	0.00544306	8.89935585	0.556302593	cg00360072	chr2	239401022	239401024	-	HDAC4;AC062017.1	AC062017.1;A	antisense;a	ENST000004(	-414;-783;-72	CGI:chr2:239	Island	1_TssA	5_TxWk	HDAC4	HDAC4;AC062017.1
2	-0.5910438	4.27216999	-6.0092478	1.6964E-08	0.00544306	8.3492456	-1.210125808	cg06153087	chr12	50213593	50213595	-	LIMA1	LIMA1;LIMA1	protein_cod	ENST000003(	69953;69913;	CGI:chr12:50	166943-50167	5_TxWk	5_TxWk		LIMA1
3	0.61010741	-5.2840215	5.80897178	4.4377E-08	0.0057139	7.53886202	0.753853268	cg22607959	chr5	55307615	55307617	-	DHX29;SKIV2L2	DHX29;DHX2	protein_cod	ENST000002!	107;79;-75;5	CGI:chr5:553	Island	1_TssA	1_TssA	DHX29	DHX29;SKIV2L2
4	-0.7960734	-5.1459076	-5.7810416	5.0675E-08	0.0057139	7.42700642	-0.754122935	cg04091209	chr15	75451771	75451773	-	SIN3A	SIN3A;SIN3A	protein_cod	ENST000003(	4012;-67;-18	CGI:chr15:75	Island	1_TssA	1_TssA	MAN2C1, SN	SIN3A
5	0.67733011	-4.8188657	5.77570443	5.1974E-08	0.0057139	7.40566524	0.625334486	cg07283778	chr18	712794	712796	+	ENOSF1	ENOSF1;ENO	protein_cod	ENST000002!	-118;-250;-13	CGI:chr18:71	Island	1_TssA	1_TssA		ENOSF1
6	0.87225544	-5.6306769	5.74078913	6.1318E-08	0.0057139	7.26631628	0.597649363	cg01582104	chr1	229558848	229558850	-	ABCB10	ABCB10	protein_cod	ENST000003(	-153	CGI:chr1:229	Island	1_TssA	1_TssA	NUP133, ABC	ABCB10
7	0.51031361	-5.2708055	5.71539995	6.9127E-08	0.0057139	7.16527632	0.569934513	cg07674153	chr14	80955255	80955257	+	TSHR;CEP128	CEP128;CEP1	protein_cod	ENST000000!	4262;4229;42	CGI:chr14:80	S_Shore	1_TssA	2_TssAFlnk,	TSHR	TSHR;CEP128
8	-0.7428651	-5.8350257	-5.7090332	7.1233E-08	0.0057139	7.13997728	-0.452569303	cg05086993	chr1	6554603	6554605	-	TAS1R1;NOL9	NOL9;NOL9;T	protein_cod	ENST000003	-68;-103;-578	CGI:chr1:655	Island	1_TssA	1_TssA		TAS1R1;NOL9
9	0.85260436	-5.8374206	5.67182422	8.4859E-08	0.00605061	6.99243521	0.516015829	cg26178523	chr17	73232625	73232627	+	FAM104A;C17orf80	C17orf80;C17	protein_cod	ENST000002!	392;-12;3;-37	CGI:chr17:73	Island	1_TssA	1_TssA	C17orf80	FAM104A;C17orf80
10	-0.3600419	4.39233376	-5.6362752	1.0025E-07	0.0062016	6.85197574	-0.999460108	cg15626797	chr5	171645508	171645510	+						CGI:chr5:171	630067-17163	14_ReprPCW	14_ReprPCW		
11	-0.4941071	4.07749666	-5.6237264	1.0631E-07	0.0062016	6.80251129	-1.35336772	cg02735047	chr3	63906381	63906383	+	ATXN7	ATXN7;ATXN	protein_cod	ENST000002!	41821;41824;	CGI:chr3:639	11716-639128	5_TxWk	5_TxWk		ATXN7
12	-0.3831135	4.46322753	-5.5458136	1.5278E-07	0.00723386	6.49679194	-0.893202502	cg16168329	chr1	41162550	41162552	-	SCMH1	SCMH1;SCMH	protein_cod	ENST000003	79560;-1069;	CGI:chr1:412	41285-412426	5_TxWk	5_TxWk		SCMH1
13	0.64588067	-6.0397275	5.54359273	1.5436E-07	0.00723386	6.4881129	0.354664492	cg24259629	chr3	45842303	45842305	+	LZTFL1	LZTFL1;LZTFL	protein_cod	ENST000002!	-137;-174;73	CGI:chr3:458	Island	1_TssA	2_TssAFlnk		LZTFL1
14	0.86884489	-5.4893436	5.53881785	1.5782E-07	0.00723386	6.46946	0.749019953	cg22311009	chr9	96854232	96854234	-	ZNF782	ZNF782;ZNF7	protein_cod	ENST000002!	21342;45;-49	CGI:chr9:968	Island	1_TssA	1_TssA		ZNF782
15	0.62883208	-6.1646428	5.49950754	1.8928E-07	0.00809751	6.31624651	0.352643747	cg11943176	chr5	141692093	141692095	-						CGI:chr5:141	Island	14_ReprPCW	2_TssAFlnk		
16	-0.5145603	5.00295064	-5.4794897	2.0758E-07	0.0083253	6.23846802	-0.76764336	cg04296321	chr6	158078102	158078104	-	SYNJ2	SYNJ2;SYNJ2	protein_cod	ENST000003!	96215;422;79	CGI:chr6:158	086687-15808	4_Tx	4_Tx		SYNJ2
17	-0.3750726	-5.2185524	-5.4648823	2.2201E-07	0.00838037	6.18181534	-0.535672101	cg14836864	chr1	45500317	45500319	+	MMACHC;CCDC163P	CCDC163P;CC	protein_cod	ENST000006	-269;-238;-34	CGI:chr1:454	Island	1_TssA	2_TssAFlnk	NASP, MMAC	MMACHC;CCDC163P
18	0.70932227	-6.1003544	5.41531305	2.7867E-07	0.00993498	5.99022619	0.381274378	cg02471784	chr12	10674306	10674308	+	STYK1	STYK1;STYK1	protein_cod	ENST000000	-14;-276;-648	CGI:chr12:10	S_Shore	11_BivFlnk	11_BivFlnk		STYK1
19	0.3876963	-4.829209	5.39562997	3.049E-07	0.00996588	5.91443377	0.648338931	cg26612385	chr12	76031890	76031892	-	PHLDA1;RP11-290L1.2	PHLDA1;PHLI	protein_cod	ENST000002(	2042;-286;-1	CGI:chr12:76	Island	2_TssAFlnk	2_TssAFlnk	PHLDA1	PHLDA1;RP11-290L1.2
20	-0.4036545	4.24558393	-5.3915692	3.106E-07	0.00996588	5.89881733	-1.07024875	cg00902895	chr13	112771441	112771443	+	ATP11A	ATP11A;ATP1	protein_cod	ENST000003	81112;81112;	CGI:chr13:11	S_Shore	5_TxWk	5_TxWk		ATP11A
21	0.45890611	-5.3767918	5.35067488	3.7416E-07	0.01066592	5.74194087	0.448084195	cg14580600	chr22	39319780	39319782	-	AL022326.1;RPL3;SNO	AL022326.1;R	miRNA;prote	ENST000006	-633;-9;-1166	CGI:chr22:39	Island	1_TssA	1_TssA	SNORD83A, I	AL022326.1;RPL3;SNORD43
22	0.49981737	2.7513182	5.34738918	3.7978E-07	0.01066592	5.72936728	3.41993186	cg21427017	chr4	152633770	152633772	+	TMEM154	TMEM154;TN	protein_cod	ENST000003(	46395;7655	CGI:chr4:152	534956-15253	15_Quies	15_Quies		TMEM154
23	-0.5656323	3.45733685	-5.3368474	3.9839E-07	0.01066592	5.68905744	-2.019861812	cg26026991	chr9	8711855	8711857	-	PTPRD;RP11-134K1.2	PTPRD;PTPR(	protein_cod	ENST000003!	22091;22094;	CGI:chr9:871	N_Shore	15_Quies	15_Quies		PTPRD;RP11-134K1.2
24	0.60245856	-5.686781	5.33656316	3.989E-07	0.01066592	5.68797125	0.426516408	cg06651286	chr3	197750554	197750556	+	KIAA0226;FYTTD1	FYTTD1;FYTT	protein_cod	ENST000002(	804;676;1001	CGI:chr3:197	Island	1_TssA	1_TssA	FYTTD1, LRC	KIAA0226;FYTTD1
25	0.58216089	-5.6945954	5.30642675	4.5722E-07	0.01155161	5.57300181	0.390717889	cg08891904	chr4	986819	986821	-	SLC26A1;DGKQ;IDUA	DGKQ;IDUA;I	protein_cod	ENST000005!	76;-178;-225;	CGI:chr4:991	138-991907	1_TssA	1_TssA	SLC26A1, DG	SLC26A1;DGKQ;IDUA
26	0.60761281	-5.9059199	5.30125506	4.6803E-07	0.01155161	5.55331129	0.39800457	cg13592003	chr8	144901429	144901431	-	ZNF250	ZNF250;ZNF2	protein_cod	ENST000002!	32;2;-6;739;2	CGI:chr8:144	Island	1_TssA	1_TssA	RPL8, ZNF25(	ZNF250
27	-0.6066114	4.97151995	-5.2848897	5.0391E-07	0.0119765	5.49107861	-0.860277767	cg22889316	chr6	5180822	5180824	-	LYRM4	LYRM4;LYRM	protein_cod	ENST000003!	80117;35916;	CGI:chr6:513	4938-5135561	15_Quies	15_Quies		LYRM4
28	0.46757515	-4.7632444	5.26942355	5.4027E-07	0.01208648	5.43237239	0.787960712	cg00926502	chr13	102799198	102799200	-	KDELC1;BIVM	BIVM;BIVM;I	protein_cod	ENST000002!	149;149;-120	CGI:chr13:10	Island	1_TssA	2_TssAFlnk		KDELC1;BIVM
29	0.63390168	-4.7464772	5.26699712	5.4621E-07	0.01208648	5.42317162	0.884530696	cg00091986	chr11	46380631	46380633	+	MDK	MDK;MDK;M	protein_cod	ENST000003!	-437;-1080;-1	CGI:chr11:46	Island	2_TssAFlnk	5_TxWk		MDK
30	0.21324529	3.16634044	5.23333103	6.3537E-07	0.01313466	5.2957795	1.343164698	cg01375719	chr3	184581188	184581190	+	EPHB3;EIF2B5	EIF2B5;EPHB	protein_cod	ENST000004(	445812;1940	CGI:chr3:184	N_Shelf	4_Tx	13_ReprPC		EPHB3;EIF2B5
31	0.55007187	-6.2335405	5.22875632	6.4853E-07	0.01313466	5.27850726	0.328031832	cg20097268	chr1	110407989	110407991	-	LAMTOR5-AS1;LAMTC	LAMTOR5;LA	protein_cod	ENST000002!	-47;-295;-318	CGI:chr1:110	S_Shore	1_TssA	2_TssAFlnk	SLC16A4	LAMTOR5-AS1;LAMTOR5
32	0.46501243	-5.3227073	5.22654906	6.5498E-07	0.01313466	5.27017687	0.52688372	cg10535858	chr12	57088765	57088767	-	NAB2;TMEM194A	NAB2;NAB2;I	protein_cod	ENST000003(	-129;-329;-38	CGI:chr12:57	Island	1_TssA	1_TssA	STAT6, NEMF	NAB2;TMEM194A
33	0.40179247	-5.5700011	5.19798323	7.4423E-07	0.0142435	5.16256145	0.443327046	cg09922736	chr9	134068889	134068891	-	BRD3	BRD3;BRD3	protein_cod	ENST000003(	-872;-354	CGI:chr9:134	S_Shore	2_TssAFlnk	1_TssA	BRD3	BRD3
34	0.8100833	-6.4916995	5.19486576	7.5466E-07	0.0142435	5.15083901	0.319151413	cg21038223	chr3	49021552	49021554	-	MIR425;MIR191;DALF	DALRD3;DAL	protein_cod	ENST000003!	-301;-494;73	CGI:chr3:490	Island	1_TssA	1_TssA	ENSG000002	MIR425;MIR191;DALRD3;NDUFAF3
35	0.38034963	-4.2659389	5.17870678	8.1105E-07	0.01454274	5.0901468	0.944471134	cg16169361	chr8	8893898	8893900	+	MFHAS1	MFHAS1	protein_cod	ENST000002!	-253	CGI:chr8:889	Island	1_TssA	1_TssA		MFHAS1
36	-0.4891083	-5.4961066	-5.165052	8.6188E-07	0.01454274	5.0389512	-0.492645854	cg21951594	chr4	127632991	127632993	-	INTU	INTU;INTU;I	protein_cod	ENST000003!	59;34;25;972	CGI:chr4:127	622876-12762	1_TssA	2_TssAFlnk		INTU
37	0.60612299	-4.3938002	5.16192628	8.7394E-07	0.01454274	5.02724385	1.121017191	cg09022552	chr16	28950797	28950799	+	NFATC2IP	NFATC2IP;NF	protein_cod	ENST000003(	-140;-210;-10	CGI:chr16:28	Island	1_TssA	1_TssA	NFATC2IP	NFATC2IP
38	0.96727842	-6.1985639	5.16027868	8.8037E-07	0.01454274	5.02107453	0.485112339	cg03055492	chr2	63841358	63841360	+	UGP2	UGP2;UGP2;(	protein_cod	ENST000003!	-352;418;-96	CGI:chr2:638	Island	1_TssA	2_TssAFlnk	UGP2	UGP2
39	0.51292976	-5.3320274	5.1593952	8.8383E-07	0.01454274	5.01776691	0.583804107	cg10198124	chr17	61863704	61863706	-	BRIP1	BRIP1;BRIP1	protein_cod	ENST000002!	-183;-360	CGI:chr17:61	Island	1_TssA	1_TssA	BRIP1	BRIP1
40	-0.383542	-4.2029381	-5.1335072	9.9144E-07	0.01549268	4.92100186	-1.200471467	cg19809977	chr1	36386070	36386072	+	STK40	STK40;STK40	protein_cod	ENST000003	-186;-182;-17	CGI:chr1:363	Island	1_TssA	1_TssA	SH3D21, ENS	STK40
41	-0.4255981	3.80724511	-5.1298866	1.0075E-06	0.01549268	4.90749303	-1.286114107	cg00132291	chr2	61878012	61878014	-	CCT4;AC107081.5	AC107081.5;(	antisense;p	ENST000004(	9580;10792;-	CGI:chr2:618	88310-618888	4_Tx	4_Tx		CCT4;AC107081.5
42	-0.5466363	-5.7498705	-5.1284317	1.014E-06	0.01549268	4.90206612	-0.444871097	cg18846554	chr14	55272064	55272066	-	FBXO34;RP11-665C16	FBXO34;FBX(	protein_cod	ENST000003!	761;-90;641;-	CGI:chr14:55	Island	1_TssA	1_TssA	LOC1053705(	FBXO34;RP11-665C16.6
43	0.80742196	-5.9696402	5.0964693	1.1679E-06	0.01652353	4.78308913	0.41602705	cg02496111	chr16	3106777	3106779	+	RP11-473M20.11	RP11-473M2(	sense_overl	ENST000005!	13	CGI:chr16:31	Island	1_TssA	1_TssA	ZNF205	RP11-473M20.11
44	0.46462662	-3.2032822	5.09420634	1.1796E-06	0.01652353	4.77468314	2.256449412	cg01068014	chr13	91347970	91347972	+	MIR17HG	MIR17HG;MI	processed_t	ENST000004(	150;28;150	CGI:chr13:91	Island	1_TssA	1_TssA		MIR17HG

***Supplementary table 5.2*- List of differentially methylated regions (DMRs) associated with cardiorespiratory fitness (CRF) with sex as a covariate (FDR < 0.005). Corresponding chromosome, genomic location, annotated genes, and genes annotated by GeneHancer. Positive effect size indicates higher DNA methylation with higher fitness z-score compared to lower.**

	seqnames	start	end	width	strand	no.cpgs	min_smooth	Stouffer	HMFDR	Fisher	maxdiff	meandiff	genes annotated
1	chr13	35474754	35477611	2858	*	18	1.0056E-78	1.7425E-54	2.6121E-05	1.1959E-48	2.94986487	1.94189422	NBEA;MAB21L1
2	chr17	5499828	5501187	1360	*	11	5.5648E-62	5.7747E-35	1.6195E-05	1.6092E-31	2.90680161	1.93127256	NLRP1
3	chr11	128685043	128686684	1642	*	11	1.5464E-45	1.2938E-33	6.4046E-05	7.6277E-30	2.26416627	1.78606345	FLI1;RP11-744N12.3
4	chr3	128491156	128493429	2274	*	11	2.0958E-37	1.7271E-32	5.3105E-05	6.3494E-29	2.02332241	1.60350241	GATA2-AS1;GATA2;DNAJB8
5	chr21	45454328	45455585	1258	*	10	3.3454E-52	4.9986E-32	2.9719E-05	1.0468E-28	2.04532971	1.65603082	COL18A1
6	chr2	33133991	33134620	630	*	11	1.4063E-52	4.6442E-32	5.6843E-05	1.8094E-28	2.78082765	1.73174382	LTBP1
7	chr7	27167698	27170654	2957	*	12	1.0978E-37	6.8358E-32	0.00011555	3.8446E-28	2.39701035	1.71257004	HOXA10-HOXA9;HOXA9;HOXA10-AS;MIR196B;HOXA10
8	chr16	2210858	2213059	2202	*	12	4.5553E-36	6.2109E-31	0.00028943	5.158E-27	1.83875812	1.52310167	BRICD5;RP11-304L19.8;PGP
9	chr3	35678016	35679905	1890	*	14	1.0676E-40	2.6075E-31	0.00055977	5.7091E-27	2.77252659	2.14797316	ARPP21
10	chr19	35138117	35139746	1630	*	12	3.0416E-41	2.0713E-30	0.0003558	1.7057E-26	2.75789927	1.66773366	CTD-2527I21.4;CTD-2527I21.5;FXYD1;LGI4
11	chr6	1605677	1608431	2755	*	10	5.9512E-29	3.8668E-29	0.00012252	9.6054E-26	2.41930689	1.93406966	FOXCUT
12	chr19	13013789	13015173	1385	*	8	1.0651E-46	1.3487E-27	1.6503E-05	5.4102E-25	2.41658318	2.07509536	CTC-239J10.1;NFIX
13	chr21	29130409	29131153	745	*	8	6.469E-47	2.5955E-27	2.1103E-05	1.1278E-24	2.74561034	2.1946465	MAP3K7CL
14	chr11	1867947	1871657	3711	*	12	1.9115E-23	9.4582E-28	0.00020689	3.4413E-24	2.48668961	1.61130132	LSP1
15	chr16	1558792	1561971	3180	*	11	7.6419E-27	1.0367E-26	0.00023649	1.9333E-23	-1.79431771	-1.36771753	IFT140;TMEM204
16	chr11	72012928	72014543	1616	*	10	1.8894E-36	6.0147E-26	0.00018442	7.7528E-23	1.94314502	1.67079865	RP11-849H4.4;NUMA1
17	chr3	187735074	187737217	2144	*	10	2.7946E-33	1.4537E-25	0.000178	1.3066E-22	2.37174555	1.79707276	BCL6
18	chr5	159104123	159106072	1950	*	8	3.5947E-32	3.1332E-25	6.2605E-05	1.5142E-22	2.85822266	1.9878136	RP11-175K6.1;LINC02202;RP11-175K6.2
19	chr16	85286252	85287275	1024	*	9	1.4578E-35	7.6803E-24	0.0001962	4.8443E-21	2.43833402	1.65210308	
20	chr8	144502754	144504153	1400	*	12	5.0859E-34	1.2769E-24	0.00103817	5.0302E-21	2.34284506	1.70938049	CTD-2517M22.14;GPT
21	chr5	66958487	66959191	705	*	8	3.1385E-37	1.9364E-23	0.00012994	8.3501E-21	2.65140382	2.11682674	MAST4
22	chr19	46495528	46496049	522	*	8	1.5123E-38	4.404E-23	3.937E-05	9.7654E-21	1.997698	1.33574978	PPP5D1;CTB-158D10.3;PNMAL2
23	chr7	135748791	135749041	251	*	7	1.149E-37	7.6968E-23	5.3415E-05	1.6317E-20	2.78780536	2.23993849	FAM180A
24	chr13	113044093	113044701	609	*	9	2.3132E-36	5.758E-23	0.00019882	2.8343E-20	2.24279386	1.71054356	MCF2L
25	chr11	57540389	57542443	2055	*	9	4.2464E-28	1.0883E-22	0.00020804	4.9577E-20	3.99715138	2.18862533	SMTNL1
26	chr6	129861009	129861405	397	*	7	2.1014E-35	7.3821E-22	8.9086E-05	1.4539E-19	2.51823624	2.30003674	TMEM244
27	chr7	5428565	5429903	1339	*	6	1.5152E-32	2.9472E-21	2.8598E-05	2.7987E-19	2.30792725	2.12169227	RP11-1275H24.1;RP11-1275H24.3
28	chr6	30752029	30752706	678	*	6	5.2763E-37	4.8539E-21	1.9942E-05	2.9509E-19	2.21520773	1.94174754	HLA-K, HLA-J, TRIM39, ABCF1, C6orf136, HCG20, VARS2, G
29	chr1	203085847	203087068	1222	*	8	1.7598E-31	1.7663E-21	0.00024454	5.8475E-19	2.2718292	1.95258968	MYOG
30	chr10	49297806	49299084	1279	*	9	1.3277E-31	1.1178E-21	0.00045134	6.2306E-19	2.22152572	1.90073724	C10orf71-AS1;C10orf71
31	chr7	112480674	112482161	1488	*	9	1.6767E-29	2.9567E-21	0.00031987	1.1315E-18	2.50609721	2.017168	LSMEM1;IFRD1
32	chr19	3369479	3370245	767	*	6	1.2925E-35	2.3063E-20	1.7789E-05	1.3896E-18	2.19961264	1.7678545	NFIC;AC005514.2
33	chr11	1838607	1839004	398	*	8	9.0172E-33	4.8129E-21	0.00029948	1.5417E-18	2.78263009	1.99678594	LSP1, TNNI2
34	chr8	51809096	51809951	856	*	8	9.8445E-32	2.022E-20	0.0001749	4.2956E-18	2.7933837	1.7553003	PXDNL
35	chr16	58501188	58501651	464	*	6	6.3773E-33	4.5376E-20	5.1361E-05	4.3749E-18	2.09242416	1.59181021	NDRG4
36	chr4	94995818	94996334	517	*	5	4.51E-34	9.4098E-20	1.7999E-05	4.8061E-18	2.65687382	2.21560691	BMPR1B
37	chr17	40364400	40365229	830	*	6	2.0936E-32	7.5088E-20	7.0956E-05	7.2804E-18	-1.77130676	-1.37472512	CTD-2267D19.3;GJD3

***Supplementary table 5.3*- Reactome (CRF-DMRs). Reactome pathways identified with GSEA using the differentially methylated regions associated with cardiorespiratory fitness (CRF). Description of Reactome pathway, N represents the total number of genes in the Reactome pathway, DE represents the number of differentially methylated genes in the Reactome pathway, and SigGenesInSet are the differentially methylated genes in the Reactome pathway.**

	Description	N	DE	P.DE	FDR	SigGenesInSet
R-HSA-397014	Homo sapiens: Muscle contraction	189	113	1.0564E-05	0.00188038	ABCC9,KCNK7,SORBS3,TRDN,SORBS1,MYL12A,CORIN,AHCYL1,SCN11A,MYL6B,DES,
R-HSA-1483206	Homo sapiens: Glycerophospholipid biosynthesis	114	64	0.00182395	0.04732262	PLA2G4B,LPCAT3,PEMT,AGPAT2,STARD10,CHKA,MGLL,OSBPL5,OSBPL8,OSBPL10,P
R-HSA-194840	Homo sapiens: Rho GTPase cycle	124	77	0.00142545	0.04732262	FAM13A,NET1,DLC1,VAV3,RALBP1,AKAP13,CHN1,ARAP1,FGD4,ARHGAP42,A2M,AF
R-HSA-1989781	Homo sapiens: PPARA activates gene expression	114	64	0.001473	0.04732262	NR1H3,CDK8,CARM1,NCOA2,PPARGC1A,PLIN2,PPARGC1B,CPT2,CREBBP,CYP4A11,
R-HSA-2219528	Homo sapiens: PI3K/AKT Signaling in Cancer	98	60	0.001861	0.04732262	NRG3,FRS2,CHUK,PIK3AP1,CREB1,KLB,HBEGF,EGF,EGFR,ERBB2,ERBB4,AKT1,AKT2,E
R-HSA-400206	Homo sapiens: Regulation of lipid metabolism by PPARalpha	116	65	0.00165787	0.04732262	NR1H3,CDK8,CARM1,NCOA2,PPARGC1A,PLIN2,PPARGC1B,CPT2,CREBBP,CYP4A11,
R-HSA-76002	Homo sapiens: Platelet activation, signaling and aggregation	247	128	0.00152732	0.04732262	A1BG,LHFPL2,RASGRP2,ABCC4,RAPGEF3,FAM3C,VAV3,VTI1B,RAPGEF4,MGLL,CLU,(
R-HSA-1500931	Homo sapiens: Cell-Cell communication	118	69	0.00250567	0.05575118	CDH2,CDH3,CDH4,CDH5,CDH8,CDH10,CDH11,CDH12,CDH13,CDH15,CDH17,CDH18,F
R-HSA-1483257	Homo sapiens: Phospholipid metabolism	191	103	0.00329221	0.06511253	PLA2G4B,LPCAT3,PEMT,AGPAT2,STARD10,CHKA,MGLL,OSBPL5,OSBPL8,OSBPL10,P
R-HSA-6811558	Homo sapiens: PI5P, PP2A and IER3 Regulate PI3K/AKT Signaling	100	58	0.00405941	0.07225745	NRG3,FRS2,PIK3AP1,KLB,HBEGF,EGF,EGFR,ERBB2,ERBB4,AKT1,ESR1,ESR2,FGF1,FGF
R-HSA-1474244	Homo sapiens: Extracellular matrix organization	292	155	0.00487637	0.07800155	ADAM10,LAMC3,CRTAP,FBLN5,MMP24,ADAMTS8,ADAMTS5,EMILIN1,CAPN11,CAP
R-HSA-194138	Homo sapiens: Signaling by VEGF	104	63	0.00550052	0.07800155	CDH5,VAV3,BAIAP2,AHCYL1,NCKAP1,CRK,CTNNA1,CTNNB1,DOCK1,AKT1,AKT2,PTI
R-HSA-5663202	Homo sapiens: Diseases of signal transduction by growth factor r	370	187	0.00569674	0.07800155	HDAC5,BCL2L11,ADAM10,CDK8,CNKSR1,ATG7,NRG3,FRS2,FAM114A2,CDC37,DUSP
R-HSA-114608	Homo sapiens: Platelet degranulation	115	57	0.00710532	0.08769223	A1BG,LHFPL2,ABCC4,FAM3C,VTI1B,CLU,CD109,EGF,A2M,ALB,F5,F13A1,ALDOA,END
R-HSA-199418	Homo sapiens: Negative regulation of the PI3K/AKT network	107	61	0.00738979	0.08769223	NRG3,FRS2,PIK3AP1,KLB,HBEGF,EGF,EGFR,ERBB2,ERBB4,AKT1,AKT2,ESR1,ESR2,FGF
R-HSA-5576891	Homo sapiens: Cardiac conduction	124	70	0.01101685	0.11575883	ABCC9,KCNK7,TRDN,CORIN,AHCYL1,SCN11A,DMPK,HIPK1,FGF11,FGF12,KCNE4,WV
R-HSA-76005	Homo sapiens: Response to elevated platelet cytosolic Ca2+	120	59	0.01105562	0.11575883	A1BG,LHFPL2,ABCC4,FAM3C,VTI1B,CLU,CD109,EGF,A2M,ALB,F5,F13A1,ALDOA,END
R-HSA-112316	Homo sapiens: Neuronal System	386	198	0.02679964	0.25107032	HCN4,ABCC9,KCNK7,FLOT1,KCNMB2,TUBB3,TUBB4A,UNC13B,ARL6IP5,PDLIM5,CAM
R-HSA-5173105	Homo sapiens: O-linked glycosylation	106	58	0.02556985	0.25107032	CHST4,SPON2,SPON1,POMT1,B3GNT2,ADAMTS13,ADAMTS8,ADAMTS5,ADAMTS6,
R-HSA-983712	Homo sapiens: Ion channel transport	175	91	0.02924674	0.260296	C8orf44-SGK3,ATP9A,SGK2,TCIRG1,TRDN,ATP8A1,SLC17A3,RIPK3,WWP1,CLCN1,CL
R-HSA-373760	Homo sapiens: L1CAM interactions	110	61	0.03514444	0.29527473	RANBP9,TUBB3,TUBB4A,SCN11A,LYPLA2,CLTC,CSNK2A1,NCAN,AP2A2,AP2B1,DLG1
R-HSA-381426	Homo sapiens: Regulation of Insulin-like Growth Factor (IGF) trar	117	54	0.03649463	0.29527473	CDH2,ENAM,ADAM10,PRSS23,MGAT4A,CP,VCAN,AFP,DMP1,ALB,F5,FBN1,PCSK9,H
R-HSA-9006925	Homo sapiens: Intracellular signaling by second messengers	295	146	0.04234372	0.32770354	HDAC5,STUB1,CAMKK2,NRG3,AHCYL1,ADCY2,FRS2,ADCY3,WWP2,ADCY5,ADCY7,CH
R-HSA-8957275	Homo sapiens: Post-translational protein phosphorylation	100	47	0.0444496	0.3296679	CDH2,ENAM,ADAM10,PRSS23,MGAT4A,CP,VCAN,AFP,DMP1,ALB,F5,FBN1,PCSK9,H
R-HSA-112315	Homo sapiens: Transmission across Chemical Synapses	254	130	0.05296837	0.37713476	TUBB3,TUBB4A,UNC13B,ARL6IP5,CAMKK2,ADCY2,CPLX1,ADCY3,ADCY5,ADCY7,CHR
R-HSA-8978868	Homo sapiens: Fatty acid metabolism	168	72	0.07190269	0.49225685	ACOT8,ACAA2,SLC27A3,SLC27A2,ACOT7,CYP2U1,CPT1B,CPT2,CRAT,ACSM6,CYP4A1

***Supplementary table 5.4*- GO (CRF-DMRs). Gene Ontology terms identified with GSEA using the differentially methylated regions associated with cardiorespiratory fitness (CRF). Type of GO term- biological process (BP), molecular function (MF), and cellular component (CC), name of GO term, N represents the total number of genes in the GO term, DE represents the number of differentially methylated genes in the GO term, and SigGenesInSet are the differentially methylated genes in the GO term.**

	ONTOLOGY	TERM	N	DE	P.DE	FDR	SigGenesInSet
GO:0030029	BP	actin filament-based process	771	460	3.7998E-16	8.586E-12	HCN4,DNAJB6,ARPC2,SORBS3,TENM1,CDK5,KLHL41,CORO2B,DLC1,RAPGEF3,TESK2,BAIAP2,CAP2,NEBL,SORBS1,PDLIM5,AVIL,CFL2,NCKAP1,IQGAP2,TRIOE
GO:0043292	CC	contractile fiber	232	157	5.0303E-15	5.6833E-11	DNAJB6,ABCC9,MYZAP,STUB1,KLHL41,NEBL,PDLIM5,CFL2,FERMT2,LDB3,AHNAK2,SYNPO,FBXO32,C10orf71,MYOM3,XIRP2,KLHL40,ABRA,MYL6B,CRYAB,CSI
GO:0030016	CC	myofibril	221	150	3.2302E-14	2.433E-10	DNAJB6,ABCC9,MYZAP,STUB1,KLHL41,NEBL,PDLIM5,CFL2,FERMT2,LDB3,AHNAK2,SYNPO,FBXO32,C10orf71,MYOM3,XIRP2,KLHL40,ABRA,CRYAB,CSRP2,CTI
GO:0030036	BP	actin cytoskeleton organization	675	402	1.8177E-13	1.0268E-09	DNAJB6,ARPC2,SORBS3,TENM1,CDK5,KLHL41,CORO2B,DLC1,RAPGEF3,TESK2,BAIAP2,CAP2,NEBL,SORBS1,PDLIM5,AVIL,CFL2,NCKAP1,IQGAP2,TRIOBP,CIT,
GO:0030017	CC	sarcomere	201	137	3.2141E-13	1.4525E-09	DNAJB6,ABCC9,MYZAP,STUB1,KLHL41,NEBL,PDLIM5,CFL2,FERMT2,LDB3,AHNAK2,SYNPO,FBXO32,C10orf71,MYOM3,XIRP2,KLHL40,ABRA,CRYAB,CSRP2,CTI
GO:0003012	BP	muscle system process	452	260	5.869E-13	2.2103E-09	ADA,HCN4,SORBS3,KLHL41,TRDN,SORBS1,PDLIM5,MYL12A,PPP1R13L,NMU,PPARGC1A,AKAP13,CHRM3,CHRNA1,CHRNB1,CHRNB4,CHRND,CHRNE,CHRNG
GO:0006936	BP	muscle contraction	350	207	6.4099E-12	1.9508E-08	ADA,HCN4,SORBS3,KLHL41,TRDN,SORBS1,MYL12A,PPP1R13L,NMU,CHRM3,CHRNA1,CHRNB1,CHRNB4,CHRND,CHRNE,CHRNG,CHUK,CKMT2,CLCN1,LTB4R,
GO:0022610	BP	biological adhesion	1410	722	6.9068E-12	1.9508E-08	ADA,CDH2,CDH3,BCL2L11,SH2B3,CDH4,CDH5,DNAJB6,CDH8,PTPRU,TSPAN32,CDH10,GPC6,EDIL3,CDH11,CDH12,ARPC2,CDH13,CDH15,CDH17,PLXNC1,CDH1
GO:0007155	BP	cell adhesion	1404	719	7.9018E-12	1.9839E-08	ADA,CDH2,CDH3,BCL2L11,SH2B3,CDH4,CDH5,DNAJB6,CDH8,PTPRU,TSPAN32,CDH10,GPC6,EDIL3,CDH11,CDH12,ARPC2,CDH13,CDH15,CDH17,PLXNC1,CDH1
GO:0003779	MF	actin binding	414	261	2.2407E-11	5.0631E-08	ARPC2,MAEA,CORO2B,CAP2,NEBL,SORBS1,PDLIM5,AVIL,CFL2,IQGAP2,FERMT2,TRIOBP,LDB3,FAM107A,SYNPO,FMNL2,PHACTR3,FGD4,ADSS1,WHAMM,SP
GO:0031674	CC	I band	135	99	3.807E-11	7.8202E-08	DNAJB6,MYZAP,STUB1,NEBL,PDLIM5,CFL2,FERMT2,LDB3,AHNAK2,SYNPO,FBXO32,C10orf71,XIRP2,KLHL40,CRYAB,CSRP2,CTNNB1,DES,SYNPO2,SMTNL1,AL
GO:0051270	BP	regulation of cellular component movement	1075	551	5.047E-11	9.5035E-08	ADA,CDKN2B-AS1,HDAC5,HCN4,CDH5,C8orf44-SGK3,PTPRU,CDH13,PLXNC1,ADAM10,CDK5,CDK6,TRIB1,SPRY2,SEMA3A,DLC1,ATP8A1,ADARB1,SEMA6C,SE
GO:0051239	BP	regulation of multicellular organismal process	3234	1486	8.176E-11	1.3196E-07	ADA,CDH2,CDH3,MIR365A,MIR365B,BTNL10,HDAC5,BCL2L11,SH2B3,CDH4,HCN4,CDH5,ABCC9,NR1H3,GPC6,UST,ENAM,PLXNC1,LPCAT3,ADAM10,CDK5,NU
GO:2000145	BP	regulation of cell motility	992	508	7.8315E-11	1.3196E-07	ADA,CDKN2B-AS1,HDAC5,CDH5,C8orf44-SGK3,PTPRU,CDH13,PLXNC1,ADAM10,CDK5,CDK6,TRIB1,SPRY2,SEMA3A,DLC1,ATP8A1,ADARB1,SEMA6C,SEMA6B
GO:0040012	BP	regulation of locomotion	1031	523	2.5421E-10	3.8295E-07	ADA,CDKN2B-AS1,HDAC5,CDH5,C8orf44-SGK3,PTPRU,CDH13,PLXNC1,ADAM10,CDK5,CDK6,TRIB1,SPRY2,SEMA3A,DLC1,ATP8A1,ADARB1,SEMA6C,SEMA6B
GO:0030334	BP	regulation of cell migration	933	479	2.7165E-10	3.8364E-07	ADA,CDKN2B-AS1,HDAC5,CDH5,C8orf44-SGK3,PTPRU,CDH13,PLXNC1,ADAM10,CDK5,TRIB1,SPRY2,SEMA3A,DLC1,ATP8A1,ADARB1,SEMA6C,SEMA6B,SEMA
GO:0008092	MF	cytoskeletal protein binding	936	513	3.4562E-10	4.594E-07	BCL2L11,ARPC2,FARP1,SORBS3,CDK5,STUB1,MAEA,APC2,CORO2B,NDRG1,TUBGCP3,CAP2,NEBL,ARFGEF2,SORBS1,PDLIM5,CENPF,RGS14,AVIL,CFL2,KIF1C,
GO:0030018	CC	Z disc	124	91	7.0473E-10	8.8467E-07	DNAJB6,MYZAP,STUB1,NEBL,PDLIM5,CFL2,LDB3,AHNAK2,SYNPO,FBXO32,C10orf71,XIRP2,CRYAB,CSRP2,CTNNB1,DES,SYNPO2,FHL2,FHL3,FKBP1A,PALLD,F
GO:0071944	CC	cell periphery	5111	2216	8.0081E-10	9.5238E-07	ADA,CDH2,NAALADL1,CDH3,ZBTB42,BTNL10,SH2B3,CDH4,HCN4,CDH5,TOM1,SH2D3C,RANBP9,IQCJ-SCHIP1,ABCC5,CDH8,ABCC9,PTPRU,TSPAN32,ATP9A,C
GO:0023051	BP	regulation of signaling	3479	1610	9.5302E-10	1.0767E-06	ADA,CDH2,CDH3,BCL2L11,SH2B3,CDH5,RANBP9,IQCJ-SCHIP1,C8orf44-SGK3,ABCC9,NR1H3,PTPRU,GPC6,CDH11,TSPAN5,PPIF,RAD50,HIPK3,CDH13,TRAP1,
GO:0006928	BP	movement of cell or subcellular component	2123	1023	1.1402E-09	1.2268E-06	ADA,CDH2,CDKN2B-AS1,HDAC5,CDH4,HCN4,CDH5,RANBP9,C8orf44-SGK3,PTPRU,GPC6,ARPC2,CDH13,PLXNC1,ADAM10,CDK5,CDK6,TRIB1,SPRY2,NET1,AP
GO:0005886	CC	plasma membrane	5001	2165	1.2708E-09	1.3053E-06	ADA,CDH2,NAALADL1,CDH3,ZBTB42,BTNL10,SH2B3,CDH4,HCN4,CDH5,TOM1,SH2D3C,RANBP9,IQCJ-SCHIP1,ABCC5,CDH8,ABCC9,PTPRU,TSPAN32,ATP9A,C
GO:0070161	CC	anchoring junction	797	438	1.6089E-09	1.5807E-06	CDH2,CDH3,CDH5,CDH8,PTPRU,CDH10,CDH11,CDH12,ARPC2,CDH13,CDH18,SORBS3,TNK2,ADAM10,PATJ,FLOT1,CNKSR1,PAK4,CORO2B,DLC1,NDRG1,BAIA
GO:0009653	BP	anatomical structure morphogenesis	2689	1307	1.7169E-09	1.6165E-06	ADA,CDH2,HDAC5,BCL2L11,CDH4,CDH5,RANBP9,DNAJB6,CDH8,CDH10,GPC6,CDH11,UST,CDH12,ARPC2,ENAM,CDH13,FRY,PLXNC1,CDH18,FARP1,ADAM10
GO:0035023	BP	regulation of Rho protein signal transduction	79	63	1.9417E-09	1.755E-06	LPAR6,FLOT1,NET1,DLC1,AKAP13,COL3A1,ABRA,ARHGAP42,PRAG1,ABCA1,LPAR1,FLCN,CCDC125,EPS8,F2RL1,F2RL2,KANK1,ARHGEF18,ABL1,KANK2,ABL2,
GO:0042383	CC	sarcolemma	133	97	2.9065E-09	2.526E-06	CDH2,FLOT1,PEMT,CIB2,BVES,AHNAK2,CLCN1,COL6A2,COL6A3,DAG1,DES,DLG1,DTNA,FGF6,FLNC,SYNM,NCSTN,CORO1C,ALOX12,STAC3,GHRHR,BIN1,GO
GO:0015629	CC	actin cytoskeleton	492	285	3.3024E-09	2.7637E-06	CDH2,MYZAP,ARPC2,ACTR1A,FLOT1,SPRY2,MAEA,APC2,APBB3,CORO2B,DLC1,RAPGEF3,BAIAP2,CAP2,CIB2,NEBL,SORBS1,PDLIM5,MYL12A,DCTN6,AVIL,CC
GO:0010646	BP	regulation of cell communication	3441	1586	3.8723E-09	3.1249E-06	ADA,CDH2,CDH3,BCL2L11,SH2B3,CDH5,RANBP9,IQCJ-SCHIP1,C8orf44-SGK3,NR1H3,PTPRU,GPC6,CDH11,TSPAN5,PPIF,RAD50,HIPK3,CDH13,TRAP1,FAM13A
GO:0032879	BP	regulation of localization	2731	1266	5.1749E-09	4.0321E-06	ADA,CDH2,CDKN2B-AS1,CDH3,HDAC5,HCN4,CDH5,DNAJB6,MRLN,C8orf44-SGK3,SAE1,NR1H3,PTPRU,GPC6,KCNK7,PPIF,SGK2,CDH13,PLXNC1,LPCAT3,TENI
GO:0035556	BP	intracellular signal transduction	2683	1247	5.4258E-09	4.0867E-06	ADA,CDH2,PLA2G4B,BCL2L11,SH2B3,SH2D3C,RANBP9,IQCJ-SCHIP1,C8orf44-SGK3,MYZAP,PPIF,SGK2,RAD50,HIPK3,CDH13,TRAP1,FAM13A,LPAR6,SORBS3,
GO:0040011	BP	locomotion	1850	886	7.6827E-09	5.5999E-06	ADA,CDH2,CDKN2B-AS1,HDAC5,CDH4,CDH5,RANBP9,C8orf44-SGK3,PTPRU,GPC6,ARPC2,CDH13,PLXNC1,ADAM10,CDK5,CDK6,TRIB1,SPRY2,NET1,APC2,PA
GO:0016477	BP	cell migration	1497	723	1.9059E-08	1.3458E-05	ADA,CDH2,CDKN2B-AS1,HDAC5,CDH5,C8orf44-SGK3,PTPRU,GPC6,CDH13,PLXNC1,ADAM10,CDK5,TRIB1,SPRY2,NET1,APC2,PAK4,LAMC3,SEMA3A,DLC1,ATI
GO:0007166	BP	cell surface receptor signaling pathway	2898	1320	2.7104E-08	1.8495E-05	ADA,CDH2,CDH3,BTNL10,BCL2L11,SH2B3,CDH5,RANBP9,C8orf44-SGK3,NR1H3,PTPRU,TSPAN32,GPC6,TSPAN5,ARPC2,CDH13,CDH17,PLXNC1,TNK2,ADAM1
GO:0030054	CC	cell junction	1990	1007	2.7829E-08	1.8495E-05	ADA,CDH2,CDH3,CDH5,IQCJ-SCHIP1,CDH8,PTPRU,CDH10,GPC6,MYZAP,CDH11,CDH12,ARPC2,CDH13,CDH15,CDH17,CDH18,FARP1,SORBS3,TNK2,ADAM10,
GO:0000902	BP	cell morphogenesis	992	553	4.3463E-08	2.806E-05	CDH2,CDH4,CDH5,RANBP9,CDH8,CDH10,CDH11,UST,CDH12,ARPC2,ENAM,FRY,PLXNC1,CDH18,FARP1,ADAM10,CDK5,FLOT1,ZMPSTE24,LAMC3,SEMA3A,TU
GO:0006941	BP	striated muscle contraction	168	104	4.9487E-08	3.1061E-05	HCN4,KLHL41,PPP1R13L,CHRNA1,CHRNB1,CHRND,CHUK,JSRP1,ADORA1,ADRA1B,CCN2,GRK2,DLG1,DMPK,DSP,DTNA,FGF12,ALDOA,GPD1L,NEDD4L,SYNM
GO:0032970	BP	regulation of actin filament-based process	388	223	5.1152E-08	3.1239E-05	HCN4,ARPC2,SORBS3,TENM1,CDK5,CORO2B,DLC1,RAPGEF3,BAIAP2,AVIL,CFL2,NCKAP1,IQGAP2,TRIOBP,FAM107A,AKAP13,SYNPO,ARAP1,FRMD6,WHAM

References

1. Tannenbaum, C., et al., *Sex and gender analysis improves science and engineering.* Nature, 2019. **575**(7781): p. 137-146.
2. Tarnopolsky, M., *Sex differences in exercise metabolism and the role of 17-beta estradiol.* Medicine & Science in Sports & Exercise, 2008. **40**(4): p. 648-654.
3. McGee, S.L. and M. Hargreaves, *Epigenetics and exercise.* Trends in Endocrinology & Metabolism, 2019. **30**(9): p. 636-645.
4. Jacques, M., et al., *Epigenetic changes in healthy human skeletal muscle following exercise–a systematic review.* Epigenetics, 2019. **14**(7): p. 633-648.
5. Pillon, N.J., et al., *Transcriptomic profiling of skeletal muscle adaptations to exercise and inactivity.* Nature communications, 2020. **11**(1): p. 1-15.
6. Amar, D., et al., *Time trajectories in the transcriptomic response to exercise-a meta-analysis.* Nature Communications, 2021. **12**(1): p. 1-12.
7. Welle, S., R. Tawil, and C.A. Thornton, *Sex-Related Differences in Gene Expression in Human Skeletal Muscle.* PLoS One, 2008. **3**(1).
8. Gershoni, M. and S. Pietrokovski, *The landscape of sex-differential transcriptome and its consequent selection in human adults.* BMC Biology, 2017. **15**(1): p. 7.
9. Oliva, M., et al., *The impact of sex on gene expression across human tissues.* Science, 2020. **369**(6509): p. eaba3066.
10. Chapman, M.A., et al., *Skeletal muscle transcriptomic comparison between long-term trained and untrained men and women.* Cell reports, 2020. **31**(12): p. 107808.
11. Lopes-Ramos, C.M., et al., *Sex Differences in Gene Expression and Regulatory Networks across 29 Human Tissues.* Cell Reports, 2020. **31**(12): p. 107795.
12. Landen, S., et al., *Genetic and epigenetic sex-specific adaptations to endurance exercise.* Epigenetics, 2019. **14**(6): p. 523-535.
13. Landen, S., et al., *Genome-wide DNA methylation and transcriptome integration reveal distinct sex differences in skeletal muscle.* bioRxiv, 2021.
14. Hurtado, A.M., et al., *Female subsistence strategies among Ache hunter-gatherers of eastern Paraguay.* Human Ecology, 1985. **13**(1): p. 1-28.
15. Bliege, et al., *Why Women Hunt*

Risk and Contemporary Foraging in a Western Desert Aboriginal Community. Current Anthropology, 2008. **49**(4): p. 655-693.

16. Khramtsova, E.A., L.K. Davis, and B.E. Stranger, *The role of sex in the genomics of human complex traits.* Nat Rev Genet, 2019. **20**(3): p. 173-190.
17. Bhargava, A., et al., *Considering Sex as a Biological Variable in Basic and Clinical Studies: An Endocrine Society Scientific Statement.* Endocrine Reviews, 2021. **42**(3): p. 219-258.
18. Costello, J.T., F. Bieuzen, and C.M. Bleakley, *Where are all the female participants in Sports and Exercise Medicine research?* European Journal of Sport Science, 2014. **14**(8): p. 847-851.
19. Hagstrom, A.D., et al., *Sex bias in cohorts included in sports medicine research.* Sports Medicine, 2021: p. 1-6.
20. Knowles, O.E., et al., *Resistance training and skeletal muscle protein metabolism in eumenorrheic females: Implications for researchers and practitioners.* Sports Medicine, 2019. **49**(11): p. 1637-1650.
21. Maughan, R., J.S. Watson, and J. Weir, *Strength and cross-sectional area of human skeletal muscle.* The Journal of physiology, 1983. **338**(1): p. 37-49.

22. Carter, S.L., C. Rennie, and M.A. Tarnopolsky, *Substrate utilization during endurance exercise in men and women after endurance training.* Am J Physiol Endocrinol Metab, 2001. **280**(6): p. E898-907.
23. Horton, T.J., et al., *Fuel metabolism in men and women during and after long-duration exercise.* Journal of Applied Physiology, 1998. **85**(5): p. 1823-1832.
24. Tarnopolsky, M.A., *Sex differences in exercise metabolism and the role of 17-beta estradiol.* Med Sci Sports Exerc, 2008. **40**(4): p. 648-54.
25. Khramtsova, E.A., L.K. Davis, and B.E. Stranger, *The role of sex in the genomics of human complex traits.* Nature Reviews Genetics, 2018: p. 1.
26. Rapp, D., et al., *Reference values for peak oxygen uptake: cross-sectional analysis of cycle ergometry-based cardiopulmonary exercise tests of 10 090 adult German volunteers from the Prevention First Registry.* BMJ open, 2018. **8**(3): p. e018697.
27. Kenney, W.L., J.H. Wilmore, and D.L. Costill, *Physiology of sport and exercise*. 2015: Human kinetics.
28. Komi, P.V. and J. Karlsson, *Skeletal muscle fibre types, enzyme activities and physical performance in young males and females.* Acta Physiol Scand, 1978. **103**(2): p. 210-8.
29. Sharma, H.B. and J. Kailashiya, *Gender difference in aerobic capacity and the contribution by body composition and Haemoglobin concentration: a study in young Indian National Hockey Players.* Journal of clinical and diagnostic research: JCDR, 2016. **10**(11): p. CC09.
30. Senefeld, J.W., et al., *Comment on:" Sex Dimorphism of VO2max Trainability: A Systematic Review and Meta-analysis"/Reply.* Sports Medicine, 2020. **50**(5): p. 1047-1050.
31. Diaz-Canestro, C. and D. Montero, *Sex Dimorphism of VO2max Trainability: A Systematic Review and Meta-analysis.* Sports Medicine, 2019. **49**(12): p. 1949-1956.
32. Losnegard, T., D. Schäfer, and J. Hallén, *Exercise economy in skiing and running.* Frontiers in physiology, 2014. **5**: p. 5.
33. Holfelder, B., N. Brown, and D. Bubeck, *The influence of sex, stroke and distance on the lactate characteristics in high performance swimming.* PLoS One, 2013. **8**(10): p. e77185.
34. Bunc, V. and J. Heller, *Energy cost of running in similarly trained men and women.* European journal of applied physiology and occupational physiology, 1989. **59**(3): p. 178-183.
35. Bransford, D.R. and E.T. Howley, *Oxygen cost of running in trained and untrained men and women.* Medicine and science in sports, 1977. **9**(1): p. 41-44.
36. Londeree, B.R., *Effect of training on lactate/ventilatory thresholds: a meta-analysis.* Med Sci Sports Exerc, 1997. **29**(6): p. 837-43.
37. Ghosh, A.K., *Anaerobic threshold: its concept and role in endurance sport.* The Malaysian journal of medical sciences: MJMS, 2004. **11**(1): p. 24.
38. Hoffmann, S., et al., *Comparison of training responses and performance adaptations in endurance-trained men and women performing high-intensity interval training.* J Sports Sci, 2021. **39**(9): p. 1010-1020.
39. McNulty, K.L., et al., *The effects of menstrual cycle phase on exercise performance in eumenorrheic women: a systematic review and meta-analysis.* Sports medicine, 2020: p. 1-15.
40. Jurkowski, J.E., et al., *Effects of menstrual cycle on blood lactate, O2 delivery, and performance during exercise.* J Appl Physiol Respir Environ Exerc Physiol, 1981. **51**(6): p. 1493-9.

41. McCracken, M., B. Ainsworth, and A.C. Hackney, *Effects of the menstrual cycle phase on the blood lactate responses to exercise.* Eur J Appl Physiol Occup Physiol, 1994. **69**(2): p. 174-5.
42. Smekal, G., et al., *Menstrual cycle: no effect on exercise cardiorespiratory variables or blood lactate concentration.* Med Sci Sports Exerc, 2007. **39**(7): p. 1098-106.
43. McNulty, K.L., et al., *The Effects of Menstrual Cycle Phase on Exercise Performance in Eumenorrheic Women: A Systematic Review and Meta-Analysis.* Sports Med, 2020. **50**(10): p. 1813-1827.
44. Bemben, D.A., P.C. Salm, and A.J. Salm, *Ventilatory and blood lactate responses to maximal treadmill exercise during the menstrual cycle.* J Sports Med Phys Fitness, 1995. **35**(4): p. 257-62.
45. Monteiro, E.R., et al., *Male relative muscle strength exceeds females for bench press and back squat.* Journal of Exercise Physiology Online, 2016. **19**: p. 79+.
46. Weber, C.L., M. Chia, and O. Inbar, *Gender differences in anaerobic power of the arms and legs-a scaling issue.* Medicine and science in sports and exercise, 2006. **38**(1): p. 129.
47. Bishop, P., K. Cureton, and M. Collins, *Sex difference in muscular strength in equally-trained men and women.* Ergonomics, 1987. **30**(4): p. 675-687.
48. Monteiro, E.R., et al., *Male relative muscle strength exceeds females for bench press and back squat.* J. Exerc. Physiol. Online, 2016. **19**: p. 79-86.
49. Roberts, B.M., G. Nuckols, and J.W. Krieger, *Sex Differences in Resistance Training: A Systematic Review and Meta-Analysis.* J Strength Cond Res, 2020. **34**(5): p. 1448-1460.
50. Staron, R.S., et al., *Skeletal muscle adaptations during early phase of heavy-resistance training in men and women.* Journal of Applied Physiology, 1994. **76**(3): p. 1247-1255.
51. Gentil, P., et al., *Comparison of upper body strength gains between men and women after 10 weeks of resistance training.* PeerJ, 2016. **4**: p. e1627.
52. Abe, T., et al., *Time course for strength and muscle thickness changes following upper and lower body resistance training in men and women.* Eur J Appl Physiol, 2000. **81**(3): p. 174-80.
53. Kell, R.T., *The influence of periodized resistance training on strength changes in men and women.* J Strength Cond Res, 2011. **25**(3): p. 735-44.
54. O'Hagan, F.T., et al., *Response to resistance training in young women and men.* Int J Sports Med, 1995. **16**(5): p. 314-21.
55. Knowles, O.E., et al., *Resistance Training and Skeletal Muscle Protein Metabolism in Eumenorrheic Females: Implications for Researchers and Practitioners.* Sports Med, 2019. **49**(11): p. 1637-1650.
56. Russell, A.P., *Molecular regulation of skeletal muscle mass.* Clinical and Experimental Pharmacology and Physiology, 2010. **37**(3): p. 378-384.
57. Fujita, S., et al., *Basal muscle intracellular amino acid kinetics in women and men.* Am J Physiol Endocrinol Metab, 2007. **292**(1): p. E77-83.
58. West, D.W., et al., *Sex-based comparisons of myofibrillar protein synthesis after resistance exercise in the fed state.* J Appl Physiol (1985), 2012. **112**(11): p. 1805-13.
59. Carter, S., et al., *Changes in skeletal muscle in males and females following endurance training.* Canadian journal of physiology and pharmacology, 2001. **79**(5): p. 386-392.
60. Staron, R.S., et al., *Fiber type composition of the vastus lateralis muscle of young men and women.* J Histochem Cytochem, 2000. **48**(5): p. 623-9.

61. Simoneau, J.A. and C. Bouchard, *Human variation in skeletal muscle fiber-type proportion and enzyme activities.* American Journal of Physiology-Endocrinology and Metabolism, 1989. **257**(4): p. E567-E572.
62. Wilson, J.M., et al., *The effects of endurance, strength, and power training on muscle fiber type shifting.* The Journal of Strength & Conditioning Research, 2012. **26**(6): p. 1724-1729.
63. Paddon-Jones, D., et al., *Adaptation to chronic eccentric exercise in humans: the influence of contraction velocity.* Eur J Appl Physiol, 2001. **85**(5): p. 466-71.
64. Jansson, E., B. Sjödin, and P. Tesch, *Changes in muscle fibre type distribution in man after physical training. A sign of fibre type transformation?* Acta Physiol Scand, 1978. **104**(2): p. 235-7.
65. Roth, S.M., et al., *Muscle Size Responses to Strength Training in Young and Older Men and Women.* Journal of the American Geriatrics Society, 2001. **49**(11): p. 1428-1433.
66. Miller, R.M., D.A. Bemben, and M.G. Bemben, *The influence of sex, training intensity, and frequency on muscular adaptations to 40 weeks of resistance exercise in older adults.* Exp Gerontol, 2021. **143**: p. 111174.
67. Hubal, M.J., et al., *Variability in muscle size and strength gain after unilateral resistance training.* Med Sci Sports Exerc, 2005. **37**(6): p. 964-72.
68. Esbjörnsson Liljedahl, M., et al., *Different responses of skeletal muscle following sprint training in men and women.* Eur J Appl Physiol Occup Physiol, 1996. **74**(4): p. 375-83.
69. Howald, H., et al., *Influences of endurance training on the ultrastructural composition of the different muscle fiber types in humans.* Pflügers Archiv, 1985. **403**(4): p. 369-376.
70. Sontag, S.A., et al., *Endurance training alters motor unit activation strategies for the vastus lateralis, yet sex-related differences and relationships with muscle size remain.* Eur J Appl Physiol, 2021. **121**(5): p. 1367-1377.
71. Trevino, M.A., et al., *Sex-related differences in muscle size explained by amplitudes of higher-threshold motor unit action potentials and muscle fibre typing.* Acta Physiol (Oxf), 2019. **225**(4): p. e13151.
72. Jenkins, N.D.M., et al., *Increases in motor unit action potential amplitudes are related to muscle hypertrophy following eight weeks of high-intensity exercise training in females.* Eur J Sport Sci, 2020: p. 1-11.
73. Häkkinen, K., *Neuromuscular fatigue and recovery in male and female athletes during heavy resistance exercise.* Int J Sports Med, 1993. **14**(2): p. 53-9.
74. Ansdell, P., et al., *Sex differences in fatigability following exercise normalised to the power–duration relationship.* The Journal of physiology, 2020. **598**(24): p. 5717-5737.
75. Ansdell, P., et al., *Contraction intensity and sex differences in knee-extensor fatigability.* J Electromyogr Kinesiol, 2017. **37**: p. 68-74.
76. Tarnopolsky, M., et al., *Carbohydrate loading and metabolism during exercise in men and women.* Journal of applied Physiology, 1995. **78**(4): p. 1360-1368.
77. Tarnopolsky, M.A., *Gender differences in substrate metabolism during endurance exercise.* Can J Appl Physiol, 2000. **25**(4): p. 312-27.
78. Tarnopolsky, M.A., et al., *Influence of endurance exercise training and sex on intramyocellular lipid and mitochondrial ultrastructure, substrate use, and mitochondrial enzyme activity.* American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 2007. **292**(3): p. R1271-R1278.

79. Lamont, L.S., A.J. McCullough, and S.C. Kalhan, *Gender differences in leucine, but not lysine, kinetics.* Journal of applied physiology, 2001. **91**(1): p. 357-362.
80. Lamont, L.S., A.J. McCullough, and S.C. Kalhan, *Gender differences in the regulation of amino acid metabolism.* Journal of Applied Physiology, 2003. **95**(3): p. 1259-1265.
81. Phillips, S.M., et al., *Gender differences in leucine kinetics and nitrogen balance in endurance athletes.* Journal of Applied Physiology, 1993. **75**(5): p. 2134-2141.
82. Riddell, M.C., et al., *Substrate utilization during exercise performed with and without glucose ingestion in female and male endurance-trained athletes.* International journal of sport nutrition and exercise metabolism, 2003. **13**(4): p. 407-421.
83. McKenzie, S., et al., *Endurance exercise training attenuates leucine oxidation and BCOAD activation during exercise in humans.* American Journal of Physiology-Endocrinology And Metabolism, 2000. **278**(4): p. E580-E587.
84. Layman, D.K., *Role of leucine in protein metabolism during exercise and recovery.* Can J Appl Physiol, 2002. **27**(6): p. 646-63.
85. Pitsiladis, Y.P., et al., *Athlome Project Consortium: a concerted effort to discover genomic and other "omic" markers of athletic performance.* Physiol Genomics, 2016. **48**(3): p. 183-90.
86. Zore, T., M. Palafox, and K. Reue, *Sex differences in obesity, lipid metabolism, and inflammation-A role for the sex chromosomes?* Mol Metab, 2018. **15**: p. 35-44.
87. Taylor, R.W., et al., *Sex Differences in Regional Body Fat Distribution From Pre- to Postpuberty.* Obesity, 2010. **18**(7): p. 1410-1416.
88. Wijchers, P.J., et al., *Sexual Dimorphism in Mammalian Autosomal Gene Regulation Is Determined Not Only by Sry but by Sex Chromosome Complement As Well.* Developmental Cell, 2010. **19**(3): p. 477-484.
89. De Vries, G.J., *Minireview: Sex Differences in Adult and Developing Brains: Compensation, Compensation, Compensation.* Endocrinology, 2004. **145**(3): p. 1063-1068.
90. Link, J.C., et al., *Metabolic impact of sex chromosomes.* Adipocyte, 2013. **2**(2): p. 74-79.
91. Arnold, A.P., *The end of gonad-centric sex determination in mammals.* Trends Genet, 2012. **28**(2): p. 55-61.
92. Penny, G.D., et al., *Requirement for Xist in X chromosome inactivation.* Nature, 1996. **379**(6561): p. 131-7.
93. Carrel, L. and H.F. Willard, *X-inactivation profile reveals extensive variability in X-linked gene expression in females.* Nature, 2005. **434**(7031): p. 400-4.
94. Sidorenko, J., et al., *The effect of X-linked dosage compensation on complex trait variation.* Nature Communications, 2019. **10**(1): p. 3009.
95. Tukiainen, T., et al., *Landscape of X chromosome inactivation across human tissues.* Nature, 2017. **550**(7675): p. 244-248.
96. Arnold, A.P., *A general theory of sexual differentiation.* J Neurosci Res, 2017. **95**(1-2): p. 291-300.
97. Abramowitz, L.K. and M.S. Bartolomei, *Genomic imprinting: recognition and marking of imprinted loci.* Current opinion in genetics & development, 2012. **22**(2): p. 72-78.
98. Raznahan, A., et al., *Sex-chromosome dosage effects on gene expression in humans.* Proceedings of the National Academy of Sciences, 2018. **115**(28): p. 7398-7403.
99. Varlamov, O., C.L. Bethea, and C.T. Roberts, Jr., *Sex-specific differences in lipid and glucose metabolism.* Front Endocrinol (Lausanne), 2014. **5**: p. 241.

100. Pihlajamaa, P., B. Sahu, and O.A. Jänne, *Determinants of Receptor- and Tissue-Specific Actions in Androgen Signaling.* Endocr Rev, 2015. **36**(4): p. 357-84.
101. Alexander, S.E., A.C. Pollock, and S. Lamon, *The effect of sex hormones on skeletal muscle adaptation in females.* Eur J Sport Sci, 2021: p. 1-11.
102. Fornes, O., et al., *JASPAR 2020: update of the open-access database of transcription factor binding profiles.* Nucleic Acids Res, 2020. **48**(D1): p. D87-d92.
103. Lopes-Ramos, C.M., et al., *Sex Differences in Gene Expression and Regulatory Networks across 29 Human Tissues.* Cell Rep, 2020. **31**(12): p. 107795.
104. Pihlajamaa, P., et al., *Tissue-specific pioneer factors associate with androgen receptor cistromes and transcription programs.* The EMBO journal, 2014. **33**(4): p. 312-326.
105. Rajan, P., et al., *Identification of novel androgen-regulated pathways and mRNA isoforms through genome-wide exon-specific profiling of the LNCaP transcriptome.* PLoS One, 2011. **6**(12): p. e29088.
106. Wu, Y., et al., *Identification of androgen response elements in the insulin-like growth factor I upstream promoter.* Endocrinology, 2007. **148**(6): p. 2984-93.
107. Hevener, A.L., et al., *The impact of ERα action on muscle metabolism and insulin sensitivity - Strong enough for a man, made for a woman.* Mol Metab, 2018. **15**: p. 20-34.
108. Wiik, A., et al., *Expression of both oestrogen receptor alpha and beta in human skeletal muscle tissue.* Histochem Cell Biol, 2009. **131**(2): p. 181-9.
109. Vingren, J.L., et al., *Effect of resistance exercise on muscle steroid receptor protein content in strength-trained men and women.* Steroids, 2009. **74**(13-14): p. 1033-9.
110. Mayne, B.T., et al., *Large Scale Gene Expression Meta-Analysis Reveals Tissue-Specific, Sex-Biased Gene Expression in Humans.* Front Genet, 2016. **7**: p. 183.
111. Lopes-Ramos, C.M., et al., *Gene Regulatory Network Analysis Identifies Sex-Linked Differences in Colon Cancer Drug Metabolism.* Cancer Research, 2018. **78**(19): p. 5538-5547.
112. Ekenros, L., et al., *Expression of sex steroid hormone receptors in human skeletal muscle during the menstrual cycle.* Acta Physiol (Oxf), 2017. **219**(2): p. 486-493.
113. Sheffield-Moore, M., et al., *Short-term oxandrolone administration stimulates net muscle protein synthesis in young men.* J Clin Endocrinol Metab, 1999. **84**(8): p. 2705-11.
114. Urban, R.J., et al., *Testosterone administration to elderly men increases skeletal muscle strength and protein synthesis.* Am J Physiol, 1995. **269**(5 Pt 1): p. E820-6.
115. Smith, G.I., et al., *Testosterone and progesterone, but not estradiol, stimulate muscle protein synthesis in postmenopausal women.* J Clin Endocrinol Metab, 2014. **99**(1): p. 256-65.
116. Basualto-Alarcón, C., et al., *Testosterone signals through mTOR and androgen receptor to induce muscle hypertrophy.* Med Sci Sports Exerc, 2013. **45**(9): p. 1712-20.
117. Zhao, W., et al., *Expression of the muscle atrophy factor muscle atrophy F-box is suppressed by testosterone.* Endocrinology, 2008. **149**(11): p. 5449-60.
118. Zhao, W., et al., *Testosterone protects against dexamethasone-induced muscle atrophy, protein degradation and MAFbx upregulation.* J Steroid Biochem Mol Biol, 2008. **110**(1-2): p. 125-9.
119. Herbst, K.L. and S. Bhasin, *Testosterone action on skeletal muscle.* Current Opinion in Clinical Nutrition & Metabolic Care, 2004. **7**(3): p. 271-277.
120. Bhasin, S., et al., *Testosterone dose-response relationships in healthy young men.* Am J Physiol Endocrinol Metab, 2001. **281**(6): p. E1172-81.

121. Mouser, J.G., P.D. Loprinzi, and J.P. Loenneke, *The association between physiologic testosterone levels, lean mass, and fat mass in a nationally representative sample of men in the United States.* Steroids, 2016. **115**: p. 62-66.
122. Alexander, S.E., et al., *Total testosterone is not associated with lean mass or handgrip strength in pre-menopausal females.* Sci Rep, 2021. **11**(1): p. 10226.
123. Hakkinen, K. and A. Pakarinen, *Acute hormonal responses to heavy resistance exercise in men and women at different ages.* Int J Sports Med, 1995. **16**(8): p. 507-13.
124. Kraemer, W.J., et al., *Acute hormonal responses to heavy resistance exercise in younger and older men.* Eur J Appl Physiol Occup Physiol, 1998. **77**(3): p. 206-11.
125. Cano Sokoloff, N., M. Misra, and K.E. Ackerman, *Exercise, Training, and the Hypothalamic-Pituitary-Gonadal Axis in Men and Women.* Frontiers of hormone research, 2016. **47**: p. 27-43.
126. Consitt, L.A., J.L. Copeland, and M.S. Tremblay, *Endogenous anabolic hormone responses to endurance versus resistance exercise and training in women.* Sports Med, 2002. **32**(1): p. 1-22.
127. Spiering, B.A., et al., *Elevated endogenous testosterone concentrations potentiate muscle androgen receptor responses to resistance exercise.* J Steroid Biochem Mol Biol, 2009. **114**(3-5): p. 195-9.
128. Webb, M.L., et al., *Serum testosterone concentration during two hours of moderate intensity treadmill running in trained men and women.* Endocr Res, 1984. **10**(1): p. 27-38.
129. Bae, Y.J., et al., *Unraveling the steroid hormone response in male marathon runners: Correlation of running time with aldosterone and progesterone.* J Steroid Biochem Mol Biol, 2019. **195**: p. 105473.
130. Campbell, S.E. and M.A. Febbraio, *Effects of ovarian hormones on exercise metabolism.* Curr Opin Clin Nutr Metab Care, 2001. **4**(6): p. 515-20.
131. Hansen, M., et al., *Effects of estrogen replacement and lower androgen status on skeletal muscle collagen and myofibrillar protein synthesis in postmenopausal women.* J Gerontol A Biol Sci Med Sci, 2012. **67**(10): p. 1005-13.
132. Dam, T.V., et al., *Transdermal Estrogen Therapy Improves Gains in Skeletal Muscle Mass After 12 Weeks of Resistance Training in Early Postmenopausal Women.* Frontiers in Physiology, 2021. **11**(1691).
133. Collins, B.C., et al., *Estrogen Regulates the Satellite Cell Compartment in Females.* Cell Rep, 2019. **28**(2): p. 368-381.e6.
134. Kahlert, S., et al., *Effects of estrogen on skeletal myoblast growth.* Biochem Biophys Res Commun, 1997. **232**(2): p. 373-8.
135. Moran, A.L., et al., *Estradiol replacement reverses ovariectomy-induced muscle contractile and myosin dysfunction in mature female mice.* J Appl Physiol (1985), 2007. **102**(4): p. 1387-93.
136. Krentzel, A.A., et al., *Temporal and bidirectional influences of estradiol on voluntary wheel running in adult female and male rats.* Horm Behav, 2020. **120**: p. 104694.
137. Hamadeh, M.J., M.C. Devries, and M.A. Tarnopolsky, *Estrogen supplementation reduces whole body leucine and carbohydrate oxidation and increases lipid oxidation in men during endurance exercise.* J Clin Endocrinol Metab, 2005. **90**(6): p. 3592-9.
138. D'Eon, T.M., et al., *Regulation of exercise carbohydrate metabolism by estrogen and progesterone in women.* Am J Physiol Endocrinol Metab, 2002. **283**(5): p. E1046-55.
139. Ventura-Clapier, R., et al., *Estrogens, Estrogen Receptors Effects on Cardiac and Skeletal Muscle Mitochondria.* Front Endocrinol (Lausanne), 2019. **10**: p. 557.

140. Ikeda, K., K. Horie-Inoue, and S. Inoue, *Functions of estrogen and estrogen receptor signaling on skeletal muscle.* J Steroid Biochem Mol Biol, 2019. **191**: p. 105375.
141. De Cree, C., *Sex steroid metabolism and menstrual irregularities in the exercising female. A review.* Sports Med, 1998. **25**(6): p. 369-406.
142. Maher, A.C., et al., *Sex differences in global mRNA content of human skeletal muscle.* PLoS One, 2009. **4**(7): p. e6335.
143. Lindholm, M.E., et al., *The human skeletal muscle transcriptome: sex differences, alternative splicing, and tissue homogeneity assessed with RNA sequencing.* Faseb j, 2014. **28**(10): p. 4571-81.
144. Liu, D., et al., *Skeletal muscle gene expression in response to resistance exercise: sex specific regulation.* BMC Genomics, 2010. **11**: p. 659.
145. Bird, A., *Perceptions of epigenetics.* Nature, 2007. **447**: p. 396.
146. Voisin, S., et al., *Exercise training and DNA methylation in humans.* Acta Physiol (Oxf), 2015. **213**(1): p. 39-59.
147. Handy, D.E., R. Castro, and J. Loscalzo, *Epigenetic modifications: basic mechanisms and role in cardiovascular disease.* Circulation, 2011. **123**(19): p. 2145-56.
148. Jaenisch, R. and A. Bird, *Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals.* Nat Genet, 2003. **33 Suppl**: p. 245-54.
149. Jones, P.A. and D. Takai, *The Role of DNA Methylation in Mammalian Epigenetics.* Science, 2001. **293**(5532): p. 1068-1070.
150. Piunti, A. and A. Shilatifard, *The roles of Polycomb repressive complexes in mammalian development and cancer.* Nature Reviews Molecular Cell Biology, 2021. **22**(5): p. 326-345.
151. Weaver, J.R. and M.S. Bartolomei, *Chromatin regulators of genomic imprinting.* Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 2014. **1839**(3): p. 169-177.
152. El Khattabi, L.A., et al., *A genome-wide search for new imprinted genes in the human placenta identifies DSCAM as the first imprinted gene on chromosome 21.* European Journal of Human Genetics, 2019. **27**(1): p. 49.
153. Reik, W. and J. Walter, *Genomic imprinting: parental influence on the genome.* Nature Reviews Genetics, 2001. **2**: p. 21.
154. Brown, W.M., *Exercise-associated DNA methylation change in skeletal muscle and the importance of imprinted genes: a bioinformatics meta-analysis.* Br J Sports Med, 2015. **49**(24): p. 1567-78.
155. Kong, A., et al., *Parental origin of sequence variants associated with complex diseases.* Nature, 2009. **462**(7275): p. 868-74.
156. McCarthy, N.S., et al., *Meta-analysis of human methylation data for evidence of sex-specific autosomal patterns.* BMC Genomics, 2014. **15**(1): p. 981.
157. Nielsen, S., et al., *Muscle specific miRNAs are induced by testosterone and independently upregulated by age.* Frontiers in Physiology, 2013. **4**: p. 394.
158. Liu, J., et al., *A study of the influence of sex on genome wide methylation.* PloS one, 2010. **5**(4): p. e10028.
159. El-Maarri, O., et al., *Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males.* Human genetics, 2007. **122**(5): p. 505-514.
160. Boks, M.P., et al., *The relationship of DNA methylation with age, gender and genotype in twins and healthy controls.* PloS one, 2009. **4**(8): p. e6767.
161. Guo, L., et al., *A comprehensive analysis of miRNA/isomiR expression with gender difference.* PloS one, 2016. **11**(5): p. e0154955.

162. Hicks, S.D., et al., *Distance running alters peripheral microRNAs implicated in metabolism, fluid balance, and myosin regulation in a sex-specific manner.* Physiological genomics, 2018.
163. Tsai, H.-W., P.A. Grant, and E.F. Rissman, *Sex differences in histone modifications in the neonatal mouse brain.* Epigenetics, 2009. **4**(1): p. 47-53.
164. Sugathan, A. and D.J. Waxman, *Genome-wide analysis of chromatin states reveals distinct mechanisms of sex-dependent gene regulation in male and female mouse liver.* Molecular and cellular biology, 2013. **33**(18): p. 3594-3610.
165. Howlett, K.F. and S.L. McGee, *Epigenetic regulation of skeletal muscle metabolism.* Clin Sci (Lond), 2016. **130**(13): p. 1051-63.
166. Denham, J., et al., *Exercise: putting action into our epigenome.* Sports Med, 2014. **44**(2): p. 189-209.
167. Pareja-Galeano, H., F. Sanchis-Gomar, and J.L. García-Giménez, *Physical Exercise and Epigenetic Modulation: Elucidating Intricate Mechanisms.* Sports Medicine, 2014. **44**(4): p. 429-436.
168. Ntanasis-Stathopoulos, J., et al., *Epigenetic regulation on gene expression induced by physical exercise.* J Musculoskelet Neuronal Interact, 2013. **13**(2): p. 133-146.
169. Zacharewicz, E., S. Lamon, and A.P. Russell, *MicroRNAs in skeletal muscle and their regulation with exercise, ageing, and disease.* Front Physiol, 2013. **4**: p. 266.
170. Panzeri, I. and J.A. Pospisilik, *Epigenetic control of variation and stochasticity in metabolic disease.* Mol Metab, 2018.
171. Kanwal, R., K. Gupta, and S. Gupta, *Cancer epigenetics: an introduction.* Methods Mol Biol, 2015. **1238**: p. 3-25.
172. Fogelholm, M., *Physical activity, fitness and fatness: relations to mortality, morbidity and disease risk factors. A systematic review.* Obesity reviews, 2010. **11**(3): p. 202-221.
173. Bird, A.P. and A.P. Wolffe, *Methylation-Induced Repression— Belts, Braces, and Chromatin.* Cell, 1999. **99**(5): p. 451-454.
174. Barrès, R., et al., *Acute Exercise Remodels Promoter Methylation in Human Skeletal Muscle.* Cell Metabolism, 2012. **15**(3): p. 405-411.
175. Nitert, M.D., et al., *Impact of an Exercise Intervention on DNA Methylation in Skeletal Muscle From First-Degree Relatives of Patients With Type 2 Diabetes.* Diabetes, 2012. **61**(12): p. 3322-3332.
176. Lindholm, M.E., et al., *An integrative analysis reveals coordinated reprogramming of the epigenome and the transcriptome in human skeletal muscle after training.* Epigenetics, 2014. **9**(12): p. 1557-69.
177. Barrès, R., et al., *Non-CpG methylation of the PGC-1α promoter through DNMT3B controls mitochondrial density.* Cell metabolism, 2009. **10**(3): p. 189-198.
178. Xiao, Y., et al., *Age and gender affect DNMT3a and DNMT3b expression in human liver.* Cell Biol Toxicol, 2008. **24**(3): p. 265-72.
179. Tsumagari, K., et al., *Early de novo DNA methylation and prolonged demethylation in the muscle lineage.* Epigenetics, 2013. **8**(3): p. 317-332.
180. Hadad, N., et al., *Absence of genomic hypomethylation or regulation of cytosine-modifying enzymes with aging in male and female mice.* Epigenetics & chromatin, 2016. **9**(1): p. 30.
181. Davegårdh, C., et al., *Sex influences DNA methylation and gene expression in human skeletal muscle myoblasts and myotubes.* Stem Cell Research & Therapy, 2019. **10**(1): p. 26.

182. Nitert, M.D., et al., *Impact of an exercise intervention on DNA methylation in skeletal muscle from first-degree relatives of patients with type 2 diabetes.* Diabetes, 2012. **61**(12): p. 3322-3332.
183. Chalk, T.E. and W. M Brown, *Exercise epigenetics and the fetal origins of disease.* Epigenomics, 2014. **6**(5): p. 469-472.
184. Gilbert, J.S. and M.J. Nijland, *Sex differences in the developmental origins of hypertension and cardiorenal disease.* American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 2008. **295**(6): p. R1941-R1952.
185. Intapad, S., et al., *Sex differences in the developmental origins of cardiovascular disease.* Physiology, 2014. **29**(2): p. 122-132.
186. Li, B., M. Carey, and J.L. Workman, *The role of chromatin during transcription.* Cell, 2007. **128**(4): p. 707-19.
187. Hargreaves, M., *Exercise and Gene Expression.* Prog Mol Biol Transl Sci, 2015. **135**: p. 457-69.
188. McGee, S.L. and M. Hargreaves, *Histone modifications and exercise adaptations.* Journal of Applied Physiology, 2011. **110**(1): p. 258-263.
189. Smith, J.A., et al., *CaMK activation during exercise is required for histone hyperacetylation and MEF2A binding at the MEF2 site on the Glut4 gene.* American Journal of Physiology-Endocrinology and Metabolism, 2008. **295**(3): p. E698-E704.
190. McKinsey, T.A., et al., *Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation.* Nature, 2000. **408**(6808): p. 106.
191. McGee, S.L. and M. Hargreaves, *Exercise and skeletal muscle glucose transporter 4 expression: molecular mechanisms.* Clinical And Experimental Pharmacology & Physiology, 2006. **33**(4): p. 395-399.
192. McGee, S.L., et al., *Exercise-induced histone modifications in human skeletal muscle.* The Journal of Physiology, 2009. **587**(24): p. 5951-5958.
193. Gaur, V., et al., *Disruption of the class IIa HDAC corepressor complex increases energy expenditure and lipid oxidation.* Cell reports, 2016. **16**(11): p. 2802-2810.
194. Vissing, K., et al., *Effect of sex differences on human MEF2 regulation during endurance exercise.* Am J Physiol Endocrinol Metab, 2008. **294**(2): p. E408-15.
195. Sun, J.-M., H.Y. Chen, and J.R. Davie, *Effect of estradiol on histone acetylation dynamics in human breast cancer cells.* Journal of Biological Chemistry, 2001.
196. Kim, M.Y., S.J. Hsiao, and W.L. Kraus, *A role for coactivators and histone acetylation in estrogen receptor alpha-mediated transcription initiation.* The EMBO journal, 2001. **20**(21): p. 6084-6094.
197. Ribas, V., et al., *Skeletal muscle action of estrogen receptor alpha is critical for the maintenance of mitochondrial function and metabolic homeostasis in females.* Sci Transl Med, 2016. **8**(334): p. 334ra54.
198. McDonnell, E., et al., *Lipids reprogram metabolism to become a major carbon source for histone acetylation.* Cell reports, 2016. **17**(6): p. 1463-1472.
199. Humphreys, D.T., et al., *MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function.* Proc Natl Acad Sci U S A, 2005. **102**(47): p. 16961-6.
200. Pillai, R.S., et al., *Inhibition of translational initiation by Let-7 MicroRNA in human cells.* Science, 2005. **309**(5740): p. 1573-6.
201. Guo, H., et al., *Mammalian microRNAs predominantly act to decrease target mRNA levels.* Nature, 2010. **466**(7308): p. 835-840.
202. Tonevitsky, A.G., et al., *Dynamically regulated miRNA-mRNA networks revealed by exercise.* BMC Physiology, 2013. **13**(1): p. 9.

203. Russell, A.P. and S. Lamon, *Exercise, Skeletal Muscle and Circulating microRNAs.* Prog Mol Biol Transl Sci, 2015. **135**: p. 471-96.
204. Kirby, T.J. and J.J. McCarthy, *MicroRNAs in skeletal muscle biology and exercise adaptation.* Free Radic Biol Med, 2013. **64**: p. 95-105.
205. McCarthy, J.J., *The MyomiR Network in Skeletal Muscle Plasticity.* Exercise and sport sciences reviews, 2011. **39**(3): p. 150-154.
206. Nielsen, S., et al., *Muscle specific microRNAs are regulated by endurance exercise in human skeletal muscle.* The Journal of Physiology, 2010. **588**(20): p. 4029-4037.
207. Russell, A.P., et al., *Regulation of miRNAs in human skeletal muscle following acute endurance exercise and short-term endurance training.* J Physiol, 2013. **591**(18): p. 4637-53.
208. Mueller, M., et al., *Different molecular and structural adaptations with eccentric and conventional strength training in elderly men and women.* Gerontology, 2011. **57**(6): p. 528-538.
209. Baar, K., et al., *Adaptations of skeletal muscle to exercise: rapid increase in the transcriptional coactivator PGC-1.* The FASEB journal, 2002. **16**(14): p. 1879-1886.
210. Zadro, J.R., et al., *The Beneficial Effects of Physical Activity: Is It Down to Your Genes? A Systematic Review and Meta-Analysis of Twin and Family Studies.* Sports Medicine - Open, 2017. **3**(1): p. 4.
211. International Human Genome Sequencing, C., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**: p. 860.
212. Eynon, N., et al., *Preface: genomics and biology of exercise is undergoing a paradigm shift.* BMC Genomics, 2017. **18**(Suppl 8): p. 825.
213. Wang, G., et al., *Genomics of elite sporting performance: what little we know and necessary advances*, in *Advances in genetics*. 2013, Elsevier. p. 123-149.
214. Mattsson, C.M., et al., *Sports genetics moving forward: lessons learned from medical research.* Physiological genomics, 2016. **48**(3): p. 175-182.
215. Ma, F., *The Association of Sport Performance with ACE and ACTN3 Genetic Polymorphisms: A Systematic Review and Meta-Analysis.* 2013. **8**(1).
216. Clarkson, P.M., et al., *ACTN3 genotype is associated with increases in muscle strength in response to resistance training in women.* Journal of Applied Physiology, 2005. **99**(1): p. 154-163.
217. Druzhevskaya, A.M., et al., *Association of the ACTN3 R577X polymorphism with power athlete status in Russians.* European journal of applied physiology, 2008. **103**(6): p. 631-634.
218. Eynon, N., et al., *Genes for elite power and sprint performance: ACTN3 leads the way.* Sports Med, 2013. **43**(9): p. 803-17.
219. Yang, N., et al., *ACTN3 Genotype Is Associated with Human Elite Athletic Performance.* The American Journal of Human Genetics, 2003. **73**(3): p. 627-631.
220. Shang, X., et al., *Association between the ACTN3 R577X polymorphism and female endurance athletes in China.* International journal of sports medicine, 2010. **31**(12): p. 913-916.
221. MacArthur, D.G. and K.N. North, *A gene for speed? the evolution and function of α-actinin-3.* Bioessays, 2004. **26**(7): p. 786-795.
222. Williams, A.G., et al., *The ACE gene and muscle performance.* Nature, 2000. **403**: p. 614.
223. Yan, X., et al., *The ACE I/D gene variant predicts ACE enzyme content in blood but not the ACE, UCP2 and UCP3 protein content in human skeletal muscle in the Gene SMART study.* J Appl Physiol (1985), 2018.

224. Montgomery, H.E., et al., *Human gene for physical performance.* Nature, 1998. **393**: p. 221.
225. Myerson, S., et al., *Human angiotensin I-converting enzyme gene and endurance performance.* Journal of Applied Physiology, 1999. **87**(4): p. 1313-1316.
226. Puthucheary, Z., et al., *The ACE Gene and Human Performance.* Sports Medicine, 2011. **41**(6): p. 433-448.
227. Woods, D.R., S.E. Humphries, and H.E. Montgomery, *The ACE I/D Polymorphism and Human Physical Performance.* Trends in Endocrinology & Metabolism, 2000. **11**(10): p. 416-420.
228. Charbonneau, D.E., et al., *ACE Genotype and the Muscle Hypertrophic and Strength Responses to Strength Training.* Med Sci Sports Exerc, 2008. **40**(4): p. 677-83.
229. Min, S.-K., et al., *Is there a gender difference between ACE gene and race distance?* Applied Physiology, Nutrition, and Metabolism, 2009. **34**(5): p. 926-932.
230. Gómez-Gallego, F., et al., *The I allele of the ACE gene is associated with improved exercise capacity in women with McArdle disease.* British journal of sports medicine, 2008. **42**(2): p. 134-140.
231. Ma, F., et al., *The Association of Sport Performance with ACE and ACTN3 Genetic Polymorphisms: A Systematic Review and Meta-Analysis.* PLoS ONE, 2013. **8**(1): p. e54685.
232. Higaki, J., et al., *Deletion allele of angiotensin-converting enzyme gene increases risk of essential hypertension in Japanese men: the Suita Study.* Circulation, 2000. **101**(17): p. 2060-2065.
233. Timmons, J.A., et al., *Using molecular classification to predict gains in maximal aerobic capacity following endurance exercise training in humans.* Journal of applied physiology, 2010. **108**(6): p. 1487-1496.
234. Bouchard, C., et al., *Genomic predictors of the maximal O2 uptake response to standardized exercise training programs.* Journal of Applied Physiology, 2011. **110**(5): p. 1160-1170.
235. Verweij, N., Y.J. van de Vegte, and P. van der Harst, *Genetic study links components of the autonomous nervous system to heart-rate profile during exercise.* Nat Commun, 2018. **9**(1): p. 898.
236. Willems, S.M., et al., *Large-scale GWAS identifies multiple loci for hand grip strength providing biological insights into muscular fitness.* Nat Commun, 2017. **8**: p. 16015.
237. Ramirez, J., et al., *Thirty loci identified for heart rate response to exercise and recovery implicate autonomic nervous system.* Nat Commun, 2018. **9**(1): p. 1947.
238. Arnegard, M.E., et al., *Sex as a Biological Variable: A 5-Year Progress Report and Call to Action.* J Womens Health (Larchmt), 2020. **29**(6): p. 858-864.
239. Costello, J.T., F. Bieuzen, and C.M. Bleakley, *Where are all the female participants in Sports and Exercise Medicine research?* Eur J Sport Sci, 2014. **14**(8): p. 847-51.
240. Deshmukh, A., et al., *Deep muscle-proteomic analysis of freeze-dried human muscle biopsies reveals fiber type-specific adaptations to exercise training.* Nature communications, 2021. **12**(1): p. 1-15.
241. Yan, X., et al., *The gene SMART study: method, study design, and preliminary findings.* BMC Genomics, 2017. **18**(Suppl 8): p. 821.
242. Takahashi, T., et al., *Sex differences in immune responses that underlie COVID-19 disease outcomes.* Nature, 2020. **588**(7837): p. 315-320.
243. Rawlik, K., O. Canela-Xandri, and A. Tenesa, *Evidence for sex-specific genetic architectures across a spectrum of human complex traits.* Genome biology, 2016. **17**(1): p. 166.

244. Zore, T., M. Palafox, and K. Reue, *Sex differences in obesity, lipid metabolism, and inflammation—A role for the sex chromosomes?* Molecular metabolism, 2018. **15**: p. 35-44.
245. Arnold, A.P., *Y chromosome's roles in sex differences in disease.* Proceedings of the National Academy of Sciences, 2017. **114**(15): p. 3787-3789.
246. Arnold, A.P., X. Chen, and Y. Itoh, *What a difference an X or Y makes: sex chromosomes, gene dose, and epigenetics in sexual differentiation*, in *Sex and gender differences in pharmacology*. 2013, Springer. p. 67-88.
247. Golden, L.C., et al., *Parent-of-origin differences in DNA methylation of X chromosome genes in T lymphocytes.* Proceedings of the National Academy of Sciences, 2019. **116**(52): p. 26779-26787.
248. Varlamov, O., C.L. Bethea, and C.T. Roberts Jr, *Sex-specific differences in lipid and glucose metabolism.* Frontiers in endocrinology, 2015. **5**: p. 241.
249. Haizlip, K.M., B.C. Harrison, and L.A. Leinwand, *Sex-Based Differences in Skeletal Muscle Kinetics and Fiber-Type Composition.* Physiology (Bethesda), 2015. **30**(1): p. 30-9.
250. Hunter, S.K., *Sex Differences in Human Fatigability: Mechanisms and Insight to.* Acta Physiol (Oxf), 2014. **210**(4): p. 768-89.
251. Hall, E., et al., *Sex differences in the genome-wide DNA methylation pattern and impact on gene expression, microRNA levels and insulin secretion in human pancreatic islets.* Genome Biol, 2014. **15**(12): p. 522.
252. Singmann, P., et al., *Characterization of whole-genome autosomal differences of DNA methylation between men and women.* Epigenetics & chromatin, 2015. **8**(1): p. 1-13.
253. Voisin, S., et al., *Meta-analysis of genome-wide DNA methylation and integrative OMICs in human skeletal muscle.* bioRxiv, 2020.
254. Davegårdh, C., et al., *DNA methylation in the pathogenesis of type 2 diabetes in humans.* Molecular metabolism, 2018. **14**: p. 12-25.
255. Kundaje, A., et al., *Integrative analysis of 111 reference human epigenomes.* Nature, 2015. **518**(7539): p. 317-330.
256. Yen, A. and M. Kellis, *Systematic chromatin state comparison of epigenomes associated with diverse properties including sex and tissue type.* Nature communications, 2015. **6**(1): p. 1-13.
257. Guo, S., et al., *Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA.* Nature genetics, 2017. **49**(4): p. 635-642.
258. VanderKraats, N.D., et al., *Discovering high-resolution patterns of differential DNA methylation that correlate with gene expression changes.* Nucleic acids research, 2013. **41**(14): p. 6816-6827.
259. Schlosberg, C.E., N.D. VanderKraats, and J.R. Edwards, *Modeling complex patterns of differential DNA methylation that associate with gene expression changes.* Nucleic acids research, 2017. **45**(9): p. 5100-5111.
260. Miller, A.E.J., et al., *Gender differences in strength and muscle fiber characteristics.* European journal of applied physiology and occupational physiology, 1993. **66**(3): p. 254-262.
261. Su, J., et al., *A novel atlas of gene expression in human skeletal muscle reveals molecular changes associated with aging.* Skeletal muscle, 2015. **5**(1): p. 1-12.
262. Yousefi, P., et al., *Sex differences in DNA methylation assessed by 450 K BeadChip in newborns.* BMC genomics, 2015. **16**(1): p. 911.
263. Inoshita, M., et al., *Sex differences of leukocytes DNA methylation adjusted for estimated cellular proportions.* Biology of sex differences, 2015. **6**(1): p. 1-7.

264. Bauer, M., *Cell-type-specific disturbance of DNA methylation pattern: a chance to get more benefit from and to minimize cohorts for epigenome-wide association studies.* International Journal of Epidemiology, 2018. **47**(3): p. 917-927.
265. Suzuki, M.M. and A. Bird, *DNA methylation landscapes: provocative insights from epigenomics.* Nature Reviews Genetics, 2008. **9**(6): p. 465-476.
266. Begue, G., et al., *DNA methylation assessment from human slow-and fast-twitch skeletal muscle fibers.* Journal of Applied Physiology, 2017. **122**(4): p. 952-967.
267. Ho, B., et al., *X chromosome dosage and presence of SRY shape sex-specific differences in DNA methylation at an autosomal region in human cells.* Biol Sex Differ, 2018. **9**(1): p. 10.
268. Grafodatskaya, D., et al., *Multilocus loss of DNA methylation in individuals with mutations in the histone H3 lysine 4 demethylase KDM5C.* BMC Med Genomics, 2013. **6**: p. 1.
269. Trolle, C., et al., *Widespread DNA hypomethylation and differential gene expression in Turner syndrome.* Scientific reports, 2016. **6**(1): p. 1-14.
270. Sharma, A., et al., *DNA methylation signature in peripheral blood reveals distinct characteristics of human X chromosome numerical aberrations.* Clinical epigenetics, 2015. **7**(1): p. 1-15.
271. Hiam, D., et al., *Epigenetic reprogramming of immune cells in women with PCOS impact genes controlling reproductive function.* The Journal of Clinical Endocrinology & Metabolism, 2019. **104**(12): p. 6155-6170.
272. Iwasaki, M., et al., *Association of postmenopausal endogenous sex hormones with global methylation level of leukocyte DNA among Japanese women.* BMC cancer, 2012. **12**(1): p. 323.
273. Aran, D., et al., *Replication timing-related and gene body-specific methylation of active human genes.* Human molecular genetics, 2011. **20**(4): p. 670-680.
274. Lister, R., et al., *Human DNA methylomes at base resolution show widespread epigenomic differences.* nature, 2009. **462**(7271): p. 315-322.
275. Rauch, T.A., et al., *A human B cell methylome at 100− base pair resolution.* Proceedings of the National Academy of Sciences, 2009. **106**(3): p. 671-678.
276. Ball, M.P., et al., *Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells.* Nature biotechnology, 2009. **27**(4): p. 361-368.
277. Smith, J., et al., *Promoter DNA hypermethylation and paradoxical gene activation.* Trends in cancer, 2020. **6**(5): p. 392-406.
278. Mielcarek, M., et al., *HDAC4 as a potential therapeutic target in neurodegenerative diseases: a summary of recent achievements.* Frontiers in cellular neuroscience, 2015. **9**: p. 42.
279. Pigna, E., et al., *HDAC4 preserves skeletal muscle structure following long-term denervation by mediating distinct cellular responses.* Skeletal muscle, 2018. **8**(1): p. 1-16.
280. Meng, Z.-X., et al., *Glucose sensing by skeletal myocytes couples nutrient signaling to systemic homeostasis.* Molecular cell, 2017. **66**(3): p. 332-344. e4.
281. Stefanetti, R.J., et al., *Recent advances in understanding the role of FOXO3.* F1000Research, 2018. **7**.
282. Bui, T.T., et al., *γ-Glutamyl transferase 7 is a novel regulator of glioblastoma growth.* BMC Cancer, 2015. **15**: p. 225.
283. Skelly, L.E., et al., *Effect of sex on the acute skeletal muscle response to sprint interval exercise.* Experimental physiology, 2017. **102**(3): p. 354-365.
284. Welle, S., et al., *Skeletal muscle gene expression profiles in 20–29 year old and 65–71 year old women.* Experimental gerontology, 2004. **39**(3): p. 369-377.

285. Voisin, S., et al., *Meta-analysis of genome-wide DNA methylation and integrative omics of age in human skeletal muscle.* Journal of Cachexia, Sarcopenia and Muscle, 2021. **12**(4): p. 1064-1078.
286. Baldelli, S., et al., *Glutathione and Nitric Oxide: Key Team Players in Use and Disuse of Skeletal Muscle.* Nutrients, 2019. **11**(10).
287. Taylor, D.L., et al., *Integrative analysis of gene expression, DNA methylation, physiological traits, and genetic variation in human skeletal muscle.* Proceedings of the National Academy of Sciences, 2019. **116**(22): p. 10883-10888.
288. Ribel-Madsen, R., et al., *Genome-wide analysis of DNA methylation differences in muscle and fat from monozygotic twins discordant for type 2 diabetes.* PloS one, 2012. **7**(12): p. e51302.
289. Tian, Y., et al., *ChAMP: updated methylation analysis pipeline for Illumina BeadChips.* Bioinformatics, 2017. **33**(24): p. 3982-3984.
290. Pidsley, R., et al., *Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling.* Genome biology, 2016. **17**(1): p. 1-17.
291. Chen, Y.-a., et al., *Cross-reactive DNA microarray probes lead to false discovery of autosomal sex-associated DNA methylation.* The American Journal of Human Genetics, 2012. **91**(4): p. 762-764.
292. Leek, J.T., et al., *Package 'sva'.* 2014.
293. van Iterson, M., E.W. van Zwet, and B.T. Heijmans, *Controlling bias and inflation in epigenome-and transcriptome-wide association studies using the empirical null distribution.* Genome biology, 2017. **18**(1): p. 1-13.
294. Leek, J.T., et al., *Tackling the widespread and critical impact of batch effects in high-throughput data.* Nature Reviews Genetics, 2010. **11**(10): p. 733-739.
295. Willer, C.J., Y. Li, and G.R. Abecasis, *METAL: fast and efficient meta-analysis of genomewide association scans.* Bioinformatics, 2010. **26**(17): p. 2190-2191.
296. Smyth, G.K., *limma: Linear Models for Microarray Data*, in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, R. Gentleman, et al., Editors. 2005, Springer New York: New York, NY. p. 397-420.
297. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the royal statistical society. Series B (Methodological), 1995: p. 289-300.
298. Benjamin, D.J., et al., *Redefine statistical significance.* Nature human behaviour, 2018. **2**(1): p. 6-10.
299. Peters, T.J., et al., *De novo identification of differentially methylated regions in the human genome.* Epigenetics & chromatin, 2015. **8**(1): p. 6.
300. Zhou, W., P.W. Laird, and H. Shen, *Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes.* Nucleic acids research, 2017. **45**(4): p. e22-e22.
301. Fishilevich, S., et al., *GeneHancer: genome-wide integration of enhancers and target genes in GeneCards.* Database, 2017. **2017**.
302. Phipson, B., J. Maksimovic, and A. Oshlack, *missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform.* Bioinformatics, 2016. **32**(2): p. 286-288.
303. Maksimovic, J., A. Oshlack, and B. Phipson, *Gene set enrichment analysis for genome-wide DNA methylation data.* bioRxiv, 2020.
304. Urbut, S.M., et al., *Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions.* Nature genetics, 2019. **51**(1): p. 187-195.

305. Mamlouk, G.M., et al., *Sex bias and omission in neuroscience research is influenced by research model and journal, but not reported NIH funding.* Frontiers in neuroendocrinology, 2020. **57**: p. 100835.
306. Heidari, S., et al., *Sex and gender equity in research: rationale for the SAGER guidelines and recommended use.* Research integrity and peer review, 2016. **1**(1): p. 1-9.
307. Ho, B., et al., *X chromosome dosage and presence of SRY shape sex-specific differences in DNA methylation at an autosomal region in human cells.* Biology of sex differences, 2018. **9**(1): p. 1-10.
308. Schiaffino, S. and C. Reggiani, *Fiber types in mammalian skeletal muscles.* Physiol Rev, 2011. **91**(4): p. 1447-531.
309. Nugent, B.M., et al., *Hormonal programming across the lifespan.* Hormone and metabolic research= Hormon-und Stoffwechselforschung= Hormones et metabolisme, 2012. **44**(8): p. 577.
310. Gertz, J., et al., *Distinct properties of cell-type-specific and shared transcription factor binding sites.* Molecular cell, 2013. **52**(1): p. 25-36.
311. KRZYCH, U., et al., *Effects of sex hormones on some T and B cell functions, evidenced by differential immune expression between male and female mice and cyclic pattern of immune responsiveness during the estrous cycle in female mice.* American Journal of Reproductive Immunology, 1981. **1**(2): p. 73-77.
312. Voss, T.C. and G.L. Hager, *Dynamic regulation of transcriptional states by chromatin and transcription factors.* Nature Reviews Genetics, 2014. **15**(2): p. 69-81.
313. Puig, R.R., et al., *UniBind: maps of high-confidence direct TF-DNA interactions across nine species.* bioRxiv, 2020.
314. Gheorghe, M., et al., *A map of direct TF–DNA interactions in the human genome.* Nucleic acids research, 2019. **47**(4): p. e21-e21.
315. Paulsen, M. and A.C. Ferguson-Smith, *DNA methylation in genomic imprinting, development, and disease.* The Journal of pathology, 2001. **195**(1): p. 97-110.
316. Faisal, M., H. Kim, and J. Kim, *Sexual differences of imprinted genes' expression levels.* Gene, 2014. **533**(1): p. 434-438.
317. Suzuki, M. and A. Bird, *Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9: 465-476.* Nature reviews. Genetics, 2008. **9**: p. 465-76.
318. Rubenstein, A.B., et al., *Single-cell transcriptional profiles in human skeletal muscle.* Scientific reports, 2020. **10**(1): p. 1-15.
319. De Micheli, A.J., et al., *A reference single-cell transcriptomic atlas of human skeletal muscle tissue reveals bifurcated muscle stem cell populations.* Skeletal muscle, 2020. **10**(1): p. 1-13.
320. Domenig, S.A., A.S. Palmer, and O. Bar-Nur, *Stem Cell-Based and Tissue Engineering Approaches for Skeletal Muscle Repair*, in *Organ Tissue Engineering*. 2020, Springer International Publishing. p. 1-62.
321. Pai, A.A., et al., *A genome-wide study of DNA methylation patterns and gene expression levels in multiple human and chimpanzee tissues.* PLoS Genet, 2011. **7**(2): p. e1001316.
322. Zhou, J., et al., *Tissue-specific DNA methylation is conserved across human, mouse, and rat, and driven by primary sequence conservation.* BMC genomics, 2017. **18**(1): p. 1-17.
323. Pöllänen, E., et al., *Differential influence of peripheral and systemic sex steroids on skeletal muscle quality in pre-and postmenopausal women.* Aging cell, 2011. **10**(4): p. 650-660.

324. Pöllänen, E., et al., *Intramuscular sex steroid hormones are associated with skeletal muscle strength and power in women with different hormonal status.* Aging Cell, 2015. **14**(2): p. 236-248.
325. Alexander, S.E., A.C. Pollock, and S. Lamon, *The effect of sex hormones on skeletal muscle adaptation in females: Influence of sex hormones on female muscle physiology.* European Journal of Sport Science, 2021(just-accepted): p. 1-27.
326. Sinha-Hikim, I., et al., *Androgen receptor in human skeletal muscle and cultured muscle satellite cells: up-regulation by androgen treatment.* The Journal of Clinical Endocrinology & Metabolism, 2004. **89**(10): p. 5245-5255.
327. Ghanim, H., et al., *Diminished androgen and estrogen receptors and aromatase levels in hypogonadal diabetic men: reversal with testosterone.* Eur J Endocrinol, 2018. **178**(3): p. 277-283.
328. Meakin, A.S., et al., *Let's Talk about Placental Sex, Baby: Understanding Mechanisms That Drive Female-and Male-Specific Fetal Growth and Developmental Outcomes.* International Journal of Molecular Sciences, 2021. **22**(12): p. 6386.
329. Skinner, B.D., et al., *A systematic review and meta-analysis examining whether changing ovarian sex steroid hormone levels influence cerebrovascular function.* Frontiers in physiology, 2021. **12**.
330. Kotamarti, V.S., et al., *Risk for Venous Thromboembolism in Transgender Patients Undergoing Cross-Sex Hormone Treatment: A Systematic Review.* The Journal of Sexual Medicine, 2021.
331. Qaisar, R., et al., *Hormone replacement therapy improves contractile function and myonuclear organization of single muscle fibres from postmenopausal monozygotic female twin pairs.* J Physiol, 2013. **591**(9): p. 2333-44.
332. Aguiar, A.S., et al., *The exercise sex gap and the impact of the estrous cycle on exercise performance in mice.* Scientific reports, 2018. **8**(1): p. 1-8.
333. Saare, M., et al., *The influence of menstrual cycle and endometriosis on endometrial methylome.* Clinical epigenetics, 2016. **8**(1): p. 1-10.
334. Eynon, N. 2018; Available from: www.vu.edu.au/speed-gene.
335. Van Buuren, S. and K. Groothuis-Oudshoorn, *Multivariate imputation by chained equations in RJ Stat*. Softw.
336. Ferrari, S. and F. Cribari-Neto, *Beta regression for modelling rates and proportions.* Journal of applied statistics, 2004. **31**(7): p. 799-815.
337. Bloemberg, D. and J. Quadrilatero, *Rapid determination of myosin heavy chain expression in rat, mouse, and human skeletal muscle using multicolor immunofluorescence analysis.* PloS one, 2012. **7**(4): p. e35273.
338. Fiuza-Luces, C., et al., *Exercise is the real polypill.* Physiology, 2013.
339. Sanford, J.A., et al., *Molecular transducers of physical activity consortium (MoTrPAC): Mapping the dynamic responses to exercise.* Cell, 2020. **181**(7): p. 1464-1474.
340. Lindholm, M.E., et al., *An integrative analysis reveals coordinated reprogramming of the epigenome and the transcriptome in human skeletal muscle after training.* Epigenetics, 2014. **9**(12): p. 1557-1569.
341. Hughes, D.C., S. Ellefsen, and K. Baar, *Adaptations to endurance and strength training.* Cold Spring Harbor perspectives in medicine, 2018. **8**(6): p. a029769.
342. Seaborne, R.A., et al., *Human skeletal muscle possesses an epigenetic memory of hypertrophy.* Scientific reports, 2018. **8**(1): p. 1-17.
343. Robinson, M.M., et al., *Enhanced protein translation underlies improved metabolic and physical adaptations to different exercise training modes in young and old humans.* Cell metabolism, 2017. **25**(3): p. 581-592.

344. Price, E.M. and W.P. Robinson, *Adjusting for batch effects in DNA methylation microarray data, a lesson learned.* Frontiers in genetics, 2018. **9**: p. 83.
345. Sailani, M.R., et al., *Lifelong physical activity is associated with promoter hypomethylation of genes involved in metabolism, myogenesis, contractile properties and oxidative stress resistance in aged human skeletal muscle.* Scientific reports, 2019. **9**(1): p. 1-11.
346. Fiorenza, M., et al., *High-intensity exercise training enhances mitochondrial oxidative phosphorylation efficiency in a temperature-dependent manner in human skeletal muscle: implications for exercise performance.* The FASEB Journal, 2019. **33**(8): p. 8976-8989.
347. Stelzer, G., et al., *The GeneCards suite: from gene data mining to disease genome sequence analyses.* Current protocols in bioinformatics, 2016. **54**(1): p. 1.30. 1-1.30. 33.
348. Sawyer, B.J., et al., *Supra-Versus Submaximal Cycle Ergometer Verification of VO2max in Males and Females.* Sports, 2020. **8**(12): p. 163.
349. Bishop, D., et al., *Relationship between plasma lactate parameters and muscle characteristics in female cyclists.* Medicine and science in sports and exercise, 2000. **32**(6): p. 1088-1093.
350. Gordon, D., et al., *The incidence of plateau at VO(2max) is affected by a bout of prior-priming exercise.* Clin Physiol Funct Imaging, 2012. **32**(1): p. 39-44.
351. Mifflin, M.D., et al., *A new predictive equation for resting energy expenditure in healthy individuals.* The American journal of clinical nutrition, 1990. **51**(2): p. 241-247.
352. Burke, L., *Fasting and recovery from exercise.* British Journal of Sports Medicine, 2010. **44**(7): p. 502-508.
353. Mihm, M., S. Gangooly, and S. Muttukrishna, *The normal menstrual cycle in women.* Anim Reprod Sci, 2011. **124**(3-4): p. 229-36.
354. Godsland, I.F., et al., *The Effects of Different Formulations of Oral Contraceptive Agents on Lipid and Carbohydrate Metabolism.* New England Journal of Medicine, 1990. **323**(20): p. 1375-1381.
355. Lundsgaard, A.M. and B. Kiens, *Gender Differences in Skeletal Muscle Substrate Metabolism – Molecular Mechanisms and Insulin Sensitivity.* Front Endocrinol (Lausanne), 2014. **5**.
356. Klein, S.L., et al., *Opinion: Sex inclusion in basic research drives discovery.* Proc Natl Acad Sci U S A, 2015. **112**(17): p. 5257-8.
357. Bartz, D., et al., *Clinical advances in sex-and gender-informed medicine to improve the health of all: a review.* JAMA Internal Medicine, 2020. **180**(4): p. 574-583.
358. Galanter, J.M., et al., *Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures.* elife, 2017. **6**: p. e20532.
359. Turner, D.C., et al., *DNA methylation across the genome in aged human skeletal muscle tissue and muscle-derived cells: the role of HOX genes and physical activity.* Scientific reports, 2020. **10**(1): p. 1-19.
360. You, C., et al., *A cell-type deconvolution meta-analysis of whole blood EWAS reveals lineage-specific smoking-associated DNA methylation changes.* Nature communications, 2020. **11**(1): p. 1-13.
361. Zheng, S.C., et al., *A novel cell-type deconvolution algorithm reveals substantial contamination by immune cells in saliva, buccal and cervix.* Epigenomics, 2018. **10**(7): p. 925-940.

362. Nederveen, J.P., et al., *Variability in skeletal muscle fibre characteristics during repeated muscle biopsy sampling in human vastus lateralis.* Applied Physiology, Nutrition, and Metabolism, 2020. **45**(4): p. 368-375.
363. Lexell, J. and C. Taylor, *Variability in muscle fibre areas in whole human quadriceps muscle: how to reduce sampling errors in biopsy techniques.* Clinical Physiology, 1989. **9**(4): p. 333-343.
364. Rönn, T. and C. Ling, *Effect of exercise on DNA methylation and metabolism in human adipose tissue and skeletal muscle.* Epigenomics, 2013. **5**(6): p. 603-605.
365. Whitham, M. and M.A. Febbraio, *Redefining Tissue Crosstalk via Shotgun Proteomic Analyses of Plasma Extracellular Vesicles.* Proteomics, 2019. **19**(1-2): p. e1800154.
366. Mendelsohn, M.E., *Genomic and nongenomic effects of estrogen in the vasculature.* Am J Cardiol, 2002. **90**(1a): p. 3f-6f.