Genome-wide DNA methylation and transcriptome integration reveal distinct sex differences in skeletal muscle.

Shanie Landen[1], Macsue Jacques[1], Danielle Hiam[1], Javier Alvarez[1], Nicholas R Harvey[2,3], Larisa M. Haupt[3], Lyn R, Griffiths[3], Kevin J Ashton[2], Séverine Lamon[4], Sarah Voisin[1*], Nir Eynon[1,5*]

[*] Sarah Voisin and Nir Eynon are co-senior authors

[1] Institute for Health and Sport (iHeS), Victoria University, Melbourne, Australia.

[2] Faculty of Health Sciences & Medicine, Bond University, Gold Coast, QLD 4226, Australia

[3] Queensland University of Technology (QUT), Centre for Genomics and Personalised Health, Genomics Research Centre, School of Biomedical Sciences, Institute of Health and Biomedical Innovation, 60 Musk Ave., Kelvin Grove, Queensland 4059, Australia

[4] Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Geelong, Australia.

[5] Murdoch Children's Research Institute, Melbourne, Australia

Corresponding author:
Professor Nir Eynon
Institute for Health and Sport (iHeS),
Victoria University
PO Box 14428
Melbourne, VIC 8001, Australia.
Tel: (61-3) 9919 5615, Fax: (61-3) 9919 5532, E-mail: Nir.Eynon@vu.edu.au

Abstract

Nearly all human complex traits and diseases exhibit some degree of sex differences, and epigenetics contributes to these differences as DNA methylation shows sex differences in various tissues. However, skeletal muscle epigenetic sex differences remain largely unexplored, yet skeletal muscle displays distinct sex differences at the transcriptome level. We conducted a large-scale meta-analysis of autosomal DNA methylation sex differences in human skeletal muscle in three separate cohorts (Gene SMART, FUSION, and GSE38291), totalling n = 369 human muscle samples (n = 222 males, n = 147 females). We found 10,240 differentially methylated regions (DMRs) at FDR < 0.005, 94% of which were

hypomethylated in males, and gene set enrichment analysis revealed that differentially methylated genes were involved in muscle contraction and metabolism. We then integrated our epigenetic results with transcriptomic data from the GTEx database and the FUSION cohort. Altogether, we identified 326 autosomal genes that display sex differences at both the DNA methylation, and transcriptome levels. Importantly, sex-biased genes at the transcriptional level were overrepresented among the sex-biased genes at the epigenetic level (p-value = 4.6e-13), which suggests differential DNA methylation and gene expression between males and females in muscle are functionally linked. Finally, we validated expression of three genes with large effect sizes (*FOXO3A*, *ALDH1A1*, and *GGT7*) in the Gene SMART cohort with qPCR. *GGT7*, involved in muscle metabolism, displays male-biased expression in skeletal muscle across the three cohorts, as well as lower methylation in males. In conclusion, we uncovered thousands of genes that exhibit DNA methylation differences between the males and females in human skeletal muscle that may modulate mechanisms controlling muscle metabolism and health.

Introduction

Sex differences are evident in nearly all complex traits. Various diseases, including but not limited to cancer, muscular dystrophy, and COVID-19 [1, 2], 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 [3], XY chromosome complements [4-7], genomic imprinting [8]), the hormonal milieu [9, 10], and gene regulation [11], with the latter likely contributing the most [1].

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% [4]. These transcriptomic differences underpin the numerous physiological differences in skeletal muscle between males and females, such as differences in substrate metabolism [12-14]. For example, females oxidize 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 [15], all of which inherently contribute to enhanced fatigue-resistance in female skeletal muscle [16]. As such, females exhibit higher mRNA and protein levels of lipid oxidation-related genes than males [13]. 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-characterized epigenetic modification, and has been shown to differ between males and females in various tissues, such as pancreatic islets [17], blood [18, 19], and more recently in cultured myoblasts and myotubes [20]. While there is ample evidence for transcriptomic sex differences in skeletal muscle [4, 11, 12, 21-23], 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 from another cohort.

Results

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 1**).

FUSION	**Females**, N = 115[1]	**Males**, N = 159[1]	**p-value**[2]
Age (years)	61 (8)	59 (8)	0.2
Health			0.026
Healthy	93 (81%)	109 (69%)	
T2D	22 (19%)	50 (31%)	
[1] Mean (SD); n (%)			
[2] Welch Two Sample t-test; Fisher's exact test			

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

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 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 1**, Supplementary Table 2**)**. Ninety-four percent of DMPs were hypomethylated in males compared with females (**Figure 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 1B**), suggesting that sex explained a substantial

amount of variance at the DMPs.

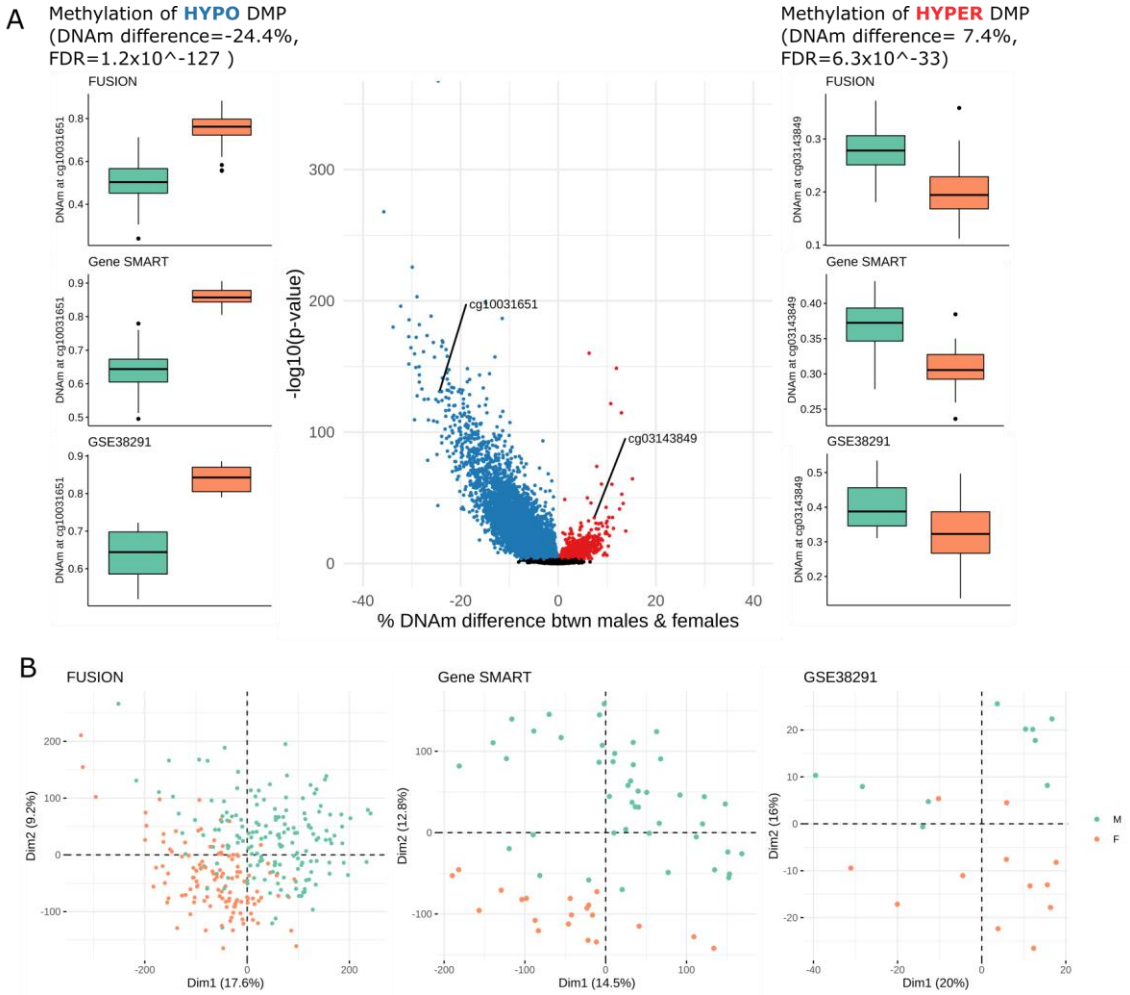


Figure 1. Differentially methylated positions (DMPs) with sex in skeletal muscle. (**A**)Volcano plot of DNA methylation differences between males and females. Each point represents a tested CpG (633,645 in total) and those that appear in color are DMPs at a meta-analysis false discovery rate < 0.005; 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 [24] and type 2 diabetes (T2D) [25]. 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 5). 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 [26]. We compared the distribution of hyper-, hypo-, and non-DMPs among the various chromatin states in human skeletal muscle using the Roadmap Epigenomics project [27]. DMPs were not randomly distributed in the chromatin states (χ^2 p-value < 2.2 x 10^{-16}, **Figure 2A**); specifically, hypo DMPs were enriched in enhancers and depleted in transcription start sites (Supplementary figure 1A), while hyper DMPs were not enriched in any chromatin states given their scarcity. It should be noted that the Roadmap Epigenomics project characterizes 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) [28]. 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 2B**). DMPs were also enriched in CpG island shores and depleted in CpG islands (χ^2 p-value < 2.2e-16) (**Figure 2C**, Supplementary figure 1B).

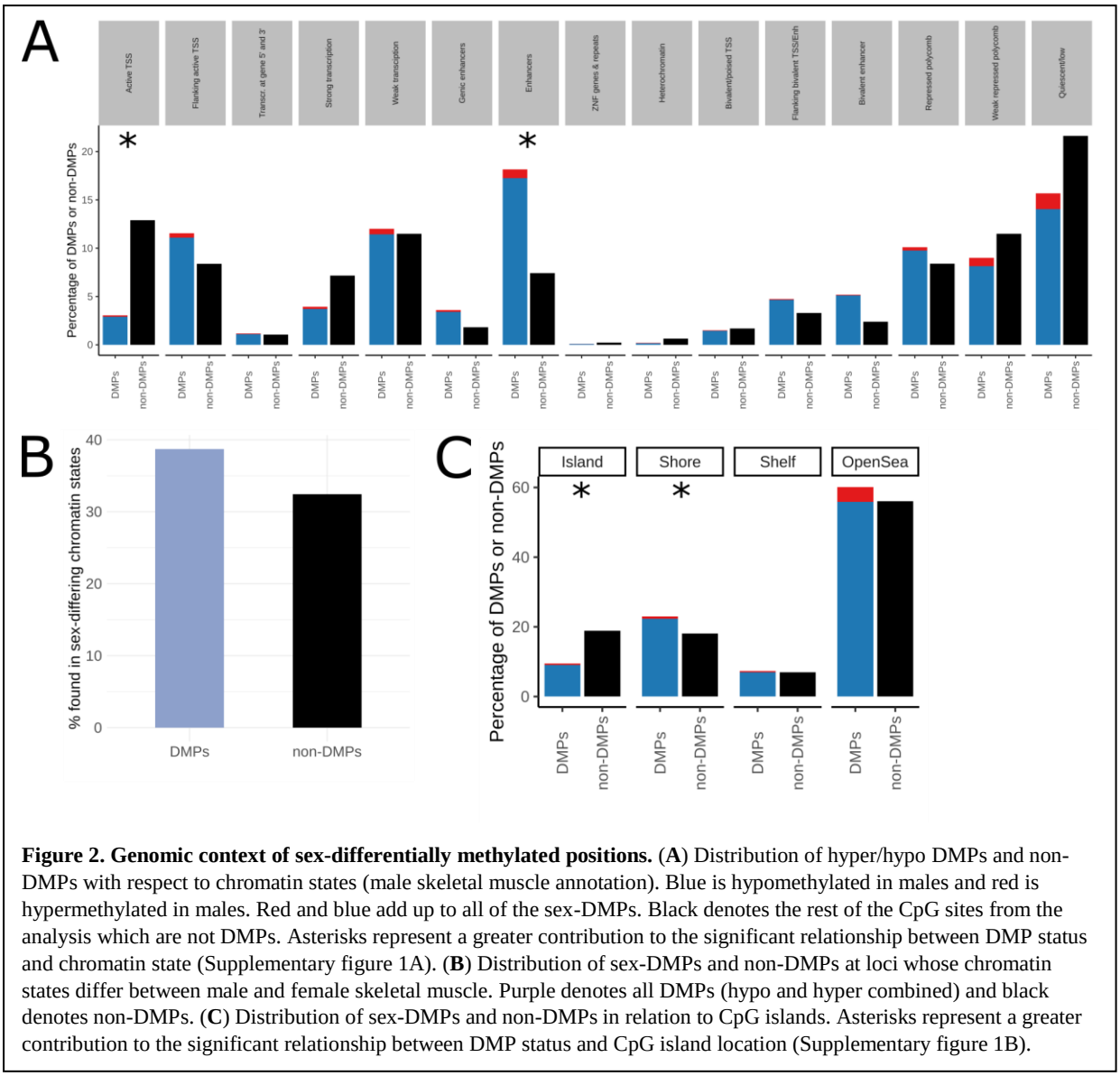


Figure 2. Genomic context of sex-differentially methylated positions. (**A**) Distribution of hyper/hypo DMPs and non-DMPs with respect to chromatin states (male skeletal muscle annotation). Blue is hypomethylated in males and red is hypermethylated in males. Red and blue add up to all of the sex-DMPs. 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 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 1B).

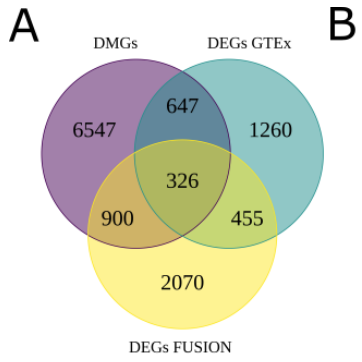
Differentially methylated genes (DMGs) were determined by identifying differentially methylated regions (DMRs), as DMRs remove spatial redundancy (CpG sites ~500 bp apart are typically highly correlated [29]), and may provide more robust and functionally important information than DMPs [30, 31]. We identified 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).

Genes with sex-biased methylation exhibit sex-biased DNA expression 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,** Supplementary table 2), including the gene Gamma-Glutamyltransferase 7 (*GGT7*) (**Figure 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 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).

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

We sought to validate 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 [32]. 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 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*) (**Table 2,** Supplementary table 11, Supplementary figure 4). To assess whether fibre type proportions may contribute to the observed DNA methylation sex differences, we explored whether there were DMRs across myosin heavy chain 6 and 7 (MYH6 and MYH7) between males and females, and identified one DMR in MYH7 (2.3% higher methylation in males, Fisher statistic= 7.2×10^{-6}; Supplementary table 3).



B

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

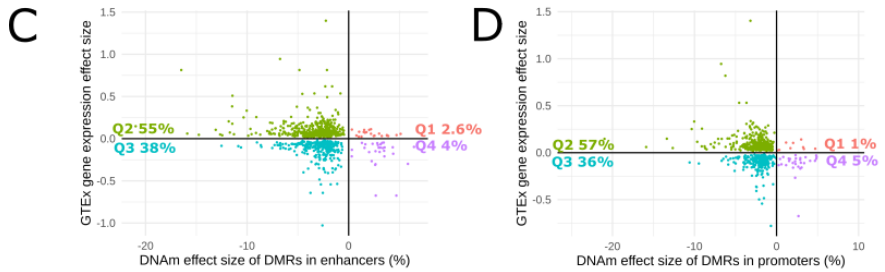


Figure 3. Integration of differentially methylated genes and differentially expressed genes. (**A**) Venn diagram of the overlap between differentially methylated genes (DMGs; derived from DMRs), differentially expressed genes derived from GTEx (DEGs GTEx), and differentially expressed genes derived 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.

	Gene expression effect size between males and females			Mean effect size of DMR showing largest effect size (% DNA methylation difference between males and females)
	GTEx (MASH posterior effect size)	FUSION (Fold change)	Gene SMART (Fold change)	
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

Table 2 Gene expression and DNA methylation differences between males and females for three genes across the cohorts used in the analysis.

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, 28]. We performed GSEA on both the DMRs and DMPs (**Figure 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 10) (FDR < 0.005). However, GSEA on the DMPs revealed enrichment across all three databases (Supplementary tables 5, 7, and 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 2), although the pathway was only significant when analysing the DMPs.

A

	GO	KEGG	Reactome
DMR GSEA	anatomical structure morphogenesis actin filament-based process actin cytoskeleton organization actin cytoskeleton actin binding actin filament binding actin filament organization contractile fiber myofibril skeletal system development muscle system process	none	Muscle contraction
DMP GSEA	anatomical structure morphogenesis actin filament-based process actin cytoskeleton organization actin binding actin cytoskeleton protein metabolic process carbohydrate derivative metabolic process lipid metabolic process muscle contraction muscle tissue development	Inositol phosphate metabolism Metabolic pathways EGFR tyrosine kinase inhibitor resistance MAPK signaling pathway ErbB signaling pathway Rap1 signaling pathway Calcium signaling pathway PI3K-Akt signaling pathway AMPK signaling pathway Focal adhesion Signaling pathways regulating pluripotency of stem cells Regulation of actin cytoskeleton Insulin signaling pathway Thyroid hormone signaling pathway	Transmission across Chemical Synapses Neuronal System Extracellular matrix organization Muscle contraction Cardiac conduction Signaling by Receptor Tyrosine Kinases

B

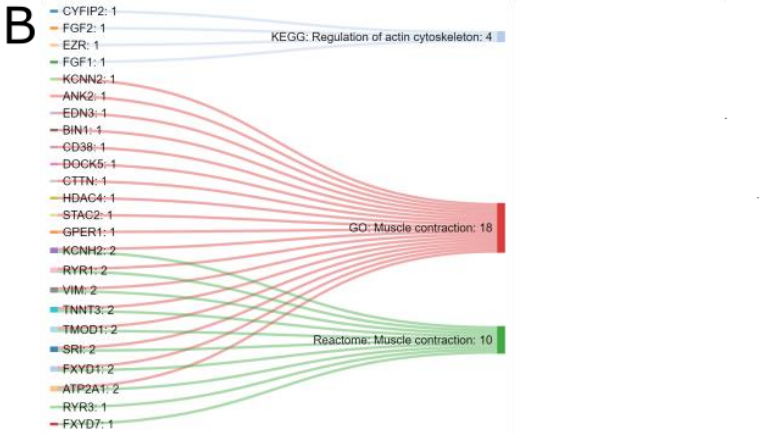


Figure 4. Gene set enrichment analysis of the differentially methylated genes. (**A**) Selected enriched Gene Ontology (GO) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, and Reactome pathways from GSEA of DMRs and DMPs. (**B**) Sankey diagram of muscle contraction-related pathways across 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.

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 [33]. 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 (**Table 2**). 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 5**).

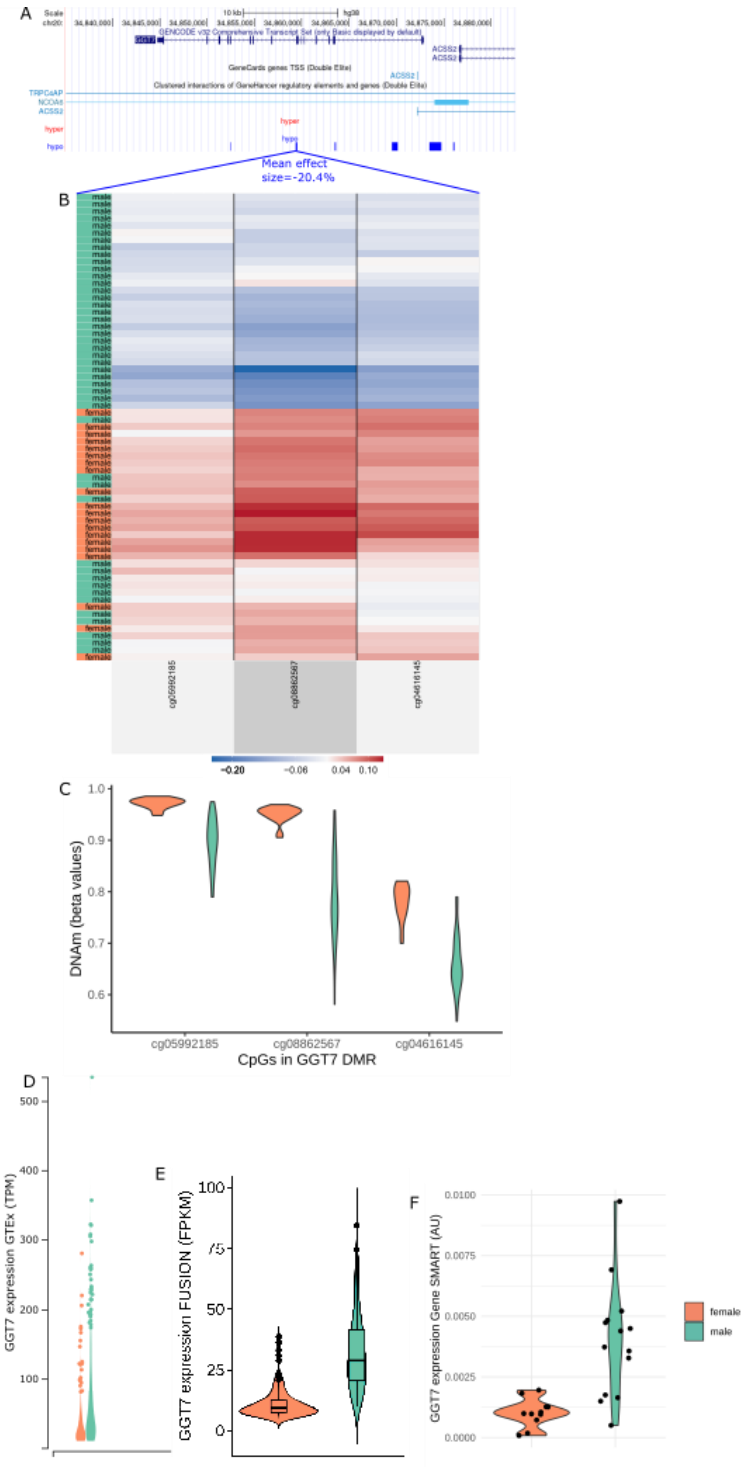


Figure 5. Differential DNA methylation and expression of GGT7 between males and females. (**A**) UCSC gene track of GGT7. From top to bottom: base pair scale in black, GENCODE gene tracks transcript variants in blue, GeneHancer regulatory element annotations in light blue, hyper DMRs tracks in red, hypo 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).

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 validated 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.

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 [34], including blood [35, 36] and pancreatic islets [17]. Uncovering the molecular mechanisms at the root of these epigenetic differences between the sexes was beyond the scope of this paper, but there are few possible explanations. Differences in cell type proportions between the sexes may partly explain our findings [36-38], as type I fibres are hypermethylated compared with type II fibers [39], and as females tend to have a higher proportion of type I fibres than males [15]. The DNA methylation of MYH7 was higher in males across the meta-analysis, potentially indicating lower MYH7 gene expression, which is a unique characteristic of type I slow twitch fibres and not type II fast twitch fibres [40]. This may suggest that males in the meta-analysis may have less type I fibres, which may contribute to some of the observed DNA methylation sex differences. However, neither MYH6 nor MYH7 displayed sex-biased gene expression in the GTEx or FUSION cohorts, although this may be confounded by age as both cohorts were comprised of

older individuals, and older individuals tend to accrue more "hybrid" fibres (co-express more than one myosin heavy chain isoform) with age [41]. Nonetheless, the contribution of differing fibre type proportions on DNA methylation differences between males and females remains to be thoroughly elucidated. 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 [42-44]. 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 [44]. This could be attributed to allele dosage compensation, a female-specific process that silences one of the X chromosomes in a cell [45, 46]. Approximately one-third of genes 'escape' inactivation, remain transcriptionally active in XX cells, [46-48], and have been suggested to affect autosomal DNA methylation via their histone marks [44, 49]. 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 [50, 51]. Finally, sex hormones may contribute to inherent autosomal sex-specific DNA methylation as has been shown in leukocytes [52], but this may only be apparent after taking cellular composition into account [53]. 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 (Rauch et al., 2009, Lister et al., 2009, Ball et al., 2009, Aran et al., 2011, Jones, 2012). 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 [28]. 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 [28]. 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 [4].

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 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 [54]. 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 [55, 56] and in the response to exercise [57]; *DEPTOR* given its role in muscle

glucose sensing which in turn augments insulin action [58]; *GRB10* given that it is imprinted and has been shown to change in methylation with exercise/training [59]; *FOXO3* for its role in ageing, longevity, and regulating the cell cycle [60]; *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 metabolize aldehydes (i.e. alcohol) more efficiently than females [12]; and *GGT7* for its role in antioxidant activity [61]. Of the three genes which were validated across GTEx, FUSION, and Gene SMART, two of them showed consistently higher male expression levels (*GGT7*, *ALDH1A1)* while one showed opposite sex-biased expression (*FOXO3*) 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 [62] and that elderly females have higher skeletal muscle *FOXO3* expression than younger females [63]. 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. 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 [64].

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.

Methods

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 [65], the Finland-United States Investigation of NIDDM Genetics (FUSION) study from the dbGAP repository (phs000867.v1.p1) [66], and the GSE38291 dataset from the Gene Expression Omnibus (GEO) platform [67]. Detailed participant characteristics, study design, muscle collection, data preprocessing, and data analysis specifications for each study are in Supplementary table 1. Briefly, all studies performed biopsies on the *vastus lateralis* muscle, all participants were of Caucasian descent (except 1 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.

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

randomized 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). The genome-wide DNA methylation pattern was analysed with the Infinium MethylationEPIC BeadChip array.

Bioinformatics and statistical analysis of DNA Methylation

PREPROCESSING

The pre-processing of DNA methylation data was performed according to the bioinformatics pipeline developed for the Bioconductor project [68]. Raw methylation data were pre-processed, filtered and normalized across samples. Probes that had a detection p-value of > 0.01, located on X and Y chromosomes or cross-hybridizing, or related to a SNP frequent in European populations, were removed. It is important to note that the list of cross-hybridizing probes was supplied manually [69] as the list supplied to the *ChAMP* package was outdated. Specifically, there are thousands of probes in the Illumina microarrays that cross-hybridize with the X-chromosome and may lead to false discovery of autosomal sex-associated DNA methylation [70]. 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* [71].

STATISTICAL ANALYSIS

We adjusted each EWAS for bias and inflation using the empirical null distribution as implemented in bacon [72]. Inflation and bias in EWAS are caused by unmeasured technical and biological confounding, such as population substructure, batch effects, and cellular heterogeneity [73]. The inflation factor is higher when the expected number of true associations is high (as it is for age); it is also greater for studies with higher statistical power [72]. The results were consistent with the inflation factors and biases reported in an EWAS of age in blood [72]. Results from the independent EWAS were combined using an inverse

variance weighted meta-analysis with METAL [74]. 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 (I^2) > 50% reflects heterogeneity between studies).

To identify DMPs, we used linear models as implemented in the *limma* package in R [75], 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, timepoint and the interaction of sex and timepoint (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 [76] and all CpGs showing an FDR < 0.005 were considered significant [77]. DMRs were identified using the *DMRcate* package [78]. 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 [79] with chromatin states from the Roadmap Epigenomics Project [27] and the latest GeneHancer information [80]. 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 [81] [82].

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 [81, 82], 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. Therefore, 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 [83]. GTEx effect sizes are represented as mash posterior effect sizes [83], 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.

Validation of top genes with qPCR

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 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 Tehcnologies. Ten microliter reactions comprised of SYBR, and optimised concentrations of forward and reverse primers (Supplementary table 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 normalized 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 11). Data are presented as the fold change in males compared to females, using $2^{-\Delta\Delta CT}$.

Acknowledgements

We are grateful for the support of the Australian National Health & Medical Research Council (NHMRC) via Sarah Voisin's Early Career Research Fellowship (APP11577321), and Nir Eynon's Career Development Fellowship (APP1140644). We also thank the Australian Research Council (ARC) for supporting this study (DP190103081 and DP200101830). The Gene SMART was supported by the Collaborative Research Network for Advancing Exercise and Sports Science (201202) from the Department of Education and Training, Australia. This research was also supported by infrastructure purchased with Australian Government EIF Super Science Funds as part of the Therapeutic Innovation Australia- Queensland Node project (LRG). We would like to thank Dr Alba Moreno from Victoria University for her assistance with the qPCR analysis and Dr Andrew Garnham for his expertise performing the muscle biopsies.

Contributions

Conceptualization: S.L., S.V., and N.E. Methodology: S.L., and S.V. Investigation: S.L., S.V. Formal analysis: S.L. Resources: S.L., M.J., D.H., J.A.R., N.R.H., L.M.H., L.R.G., K.J.A., and N.E. Software: S.L. Writing—Original draft: S.L., S.V., D.H., and N.E. Writing—Review & editing: S.L., S.V., D.H., S.Lamon, N.R.H., L.M.H., K.J.A., and N.E. Funding acquisition: S.V., L.R.G., and N.E.

Competing Interests

The authors declare that they have no conflict of interest.

References

1. 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.
2. Takahashi, T., et al., *Sex differences in immune responses that underlie COVID-19 disease outcomes.* Nature, 2020. **588**(7837): p. 315-320.
3. 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.
4. Oliva, M., et al., *The impact of sex on gene expression across human tissues.* Science, 2020. **369**(6509): p. eaba3066.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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.
10. 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.
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. Maher, A.C., et al., *Sex differences in global mRNA content of human skeletal muscle.* PLoS One, 2009. **4**(7): p. e6335.
13. 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.
14. Landen, S., et al., *Genetic and epigenetic sex-specific adaptations to endurance exercise.* Epigenetics, 2019. **14**(6): p. 523-535.
15. 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.
16. Hunter, S.K., *Sex Differences in Human Fatigability: Mechanisms and Insight to.* Acta Physiol (Oxf), 2014. **210**(4): p. 768-89.
17. 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.
18. 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.
19. Liu, J., et al., *A study of the influence of sex on genome wide methylation.* PloS one, 2010. **5**(4): p. e10028.
20. 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.

21. 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.
22. Welle, S., R. Tawil, and C.A. Thornton, *Sex-Related Differences in Gene Expression in Human Skeletal Muscle.* PLoS One, 2008. **3**(1).
23. 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.
24. Voisin, S., et al., *Meta-analysis of genome-wide DNA methylation and integrative OMICs in human skeletal muscle.* bioRxiv, 2020.
25. Davegårdh, C., et al., *DNA methylation in the pathogenesis of type 2 diabetes in humans.* Molecular metabolism, 2018. **14**: p. 12-25.
26. Jones, P.A. and D. Takai, *The Role of DNA Methylation in Mammalian Epigenetics.* Science, 2001. **293**(5532): p. 1068-1070.
27. Kundaje, A., et al., *Integrative analysis of 111 reference human epigenomes.* Nature, 2015. **518**(7539): p. 317-330.
28. 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.
29. 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.
30. 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.
31. 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.
32. 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.
33. 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.
34. 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.
35. Yousefi, P., et al., *Sex differences in DNA methylation assessed by 450 K BeadChip in newborns.* BMC genomics, 2015. **16**(1): p. 911.
36. 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.
37. 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.
38. 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.
39. 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.
40. Stuart, C.A., et al., *Myosin content of individual human muscle fibers isolated by laser capture microdissection.* American Journal of Physiology-Cell Physiology, 2016. **310**(5): p. C381-C389.

41. Deschenes, M.R., *Effects of aging on muscle fibre type and size.* Sports medicine, 2004. **34**(12): p. 809-824.
42. 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.
43. De Vries, G.J., *Minireview: Sex Differences in Adult and Developing Brains: Compensation, Compensation, Compensation.* Endocrinology, 2004. **145**(3): p. 1063-1068.
44. 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.
45. Penny, G.D., et al., *Requirement for Xist in X chromosome inactivation.* Nature, 1996. **379**(6561): p. 131-7.
46. 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.
47. Sidorenko, J., et al., *The effect of X-linked dosage compensation on complex trait variation.* Nature Communications, 2019. **10**(1): p. 3009.
48. Tukiainen, T., et al., *Landscape of X chromosome inactivation across human tissues.* Nature, 2017. **550**(7675): p. 244-248.
49. 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.
50. Trolle, C., et al., *Widespread DNA hypomethylation and differential gene expression in Turner syndrome.* Scientific reports, 2016. **6**(1): p. 1-14.
51. 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.
52. 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.
53. 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.
54. Smith, J., et al., *Promoter DNA hypermethylation and paradoxical gene activation.* Trends in cancer, 2020. **6**(5): p. 392-406.
55. 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.
56. 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.
57. McGee, S.L., et al., *Exercise-induced histone modifications in human skeletal muscle.* The Journal of Physiology, 2009. **587**(24): p. 5951-5958.
58. 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.
59. 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.
60. Stefanetti, R.J., et al., *Recent advances in understanding the role of FOXO3.* F1000Research, 2018. **7**.

61. Bui, T.T., et al., *γ-Glutamyl transferase 7 is a novel regulator of glioblastoma growth.* BMC Cancer, 2015. **15**: p. 225.
62. 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.
63. 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.
64. Baldelli, S., et al., *Glutathione and Nitric Oxide: Key Team Players in Use and Disuse of Skeletal Muscle.* Nutrients, 2019. **11**(10).
65. Yan, X., et al., *The gene SMART study: method, study design, and preliminary findings.* BMC Genomics, 2017. **18**(Suppl 8): p. 821.
66. 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.
67. 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.
68. Tian, Y., et al., *ChAMP: updated methylation analysis pipeline for Illumina BeadChips.* Bioinformatics, 2017. **33**(24): p. 3982-3984.
69. 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.
70. 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.
71. Leek, J.T., et al., *Package 'sva'.* 2014.
72. 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.
73. 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.
74. 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.
75. 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.
76. 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.
77. Benjamin, D.J., et al., *Redefine statistical significance.* Nature Human Behaviour, 2018. **2**(1): p. 6.
78. Peters, T.J., et al., *De novo identification of differentially methylated regions in the human genome.* Epigenetics & chromatin, 2015. **8**(1): p. 6.
79. 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.
80. Fishilevich, S., et al., *GeneHancer: genome-wide integration of enhancers and target genes in GeneCards.* Database, 2017. **2017**.
81. 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.

82. Maksimovic, J., A. Oshlack, and B. Phipson, *Gene set enrichment analysis for genome-wide DNA methylation data.* bioRxiv, 2020.
83. 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.