DNA Methylation Signatures Diverge Between Endurance and Resistance Trained Athletes

Author/s

Collins, C, Puras, A and Brown, J. Muhdo Health Ltd, N2-07 Columba House Adastral Park, Martlesham Heath, Ipswich, England, IP5 3RE.

Abstract

Regular exercise is known to induce epigenetic modifications, yet it remains unclear whether different training modalities imprint distinct DNA methylation patterns on the epigenome. We compared genome-wide DNA methylation in saliva from endurance athletes (n = 733; 51% male; mean age 47) and resistance-trained athletes (n = 698; 72% male; mean age 44), each with >3 years of consistent training in their respective modalities. DNA methylation was analysed on the Illumina EPIC 850k array (Eurofins, Denmark). After quality control and statistical analysis, 802 CpG sites showed genome-wide significance (p ≤ 1×10^-4) between groups. However, the majority were located on the X chromosome and were excluded to avoid confounding due to the sex imbalance between cohorts. We focused on 16 highly significant autosomal CpGs ($p = 1 \times 10^{-8}$ to 1×10^{-11}, false discovery rate q = 0.0001). These CpGs map to genes involved in transcriptional regulation (e.g., *NAB1*, *TLE1*), metabolism (e.g., *DYRK2*, *FOXN3*), cardiovascular function (e.g., *KCNJ8*), tissue remodelling (e.g., *CPED1*), and other processes. Notably, three distinct CpGs in the *TLE1* gene (a Wnt pathway repressor) were hypermethylated in endurance athletes, whereas a CpG near *KCNJ8* (encoding a K_ATP channel subunit critical for vascular response) was hypomethylated in endurance athletes. Several sites (e.g., *DYRK2*, *FOXN3*) suggest differential methylation of metabolic regulators, potentially reflecting the polarised energy demands of endurance versus strength training. Although differences in methylation were modest in magnitude, they occurred in functionally relevant regions that could influence gene expression. Our findings indicate that long-term endurance and resistance training are associated with distinct DNA methylation signatures, which may mirror the divergent physiological adaptations (aerobic capacity, muscular hypertrophy, metabolic flexibility, recovery profiles) characteristic of these training modalities. The study discusses the biological functions of the implicated genes and their potential links to athletic performance, as well as the study's limitations including tissue specificity and cohort differences. These results lay the groundwork for using epigenetic marks to understand how different exercise modalities can engender specific molecular adaptations and health outcomes.

Introduction

Physical exercise elicits a wide range of physiological adaptations, from improved aerobic capacity and mitochondrial function in endurance training to enhanced muscle fibre size and neural recruitment in resistance training. While these adaptations have been well characterised at the physiological and transcriptomic levels, emerging evidence suggests that exercise training can also remodel the epigenome, potentially leading to long-lasting changes in gene expression regulation. DNA methylation, in particular, has been identified as an important epigenetic mechanism by which exercise influences gene activity. For example, chronic physical activity has been associated with altered DNA methylation in multiple gene loci in a dose-dependent manner, often inversely correlating with gene expression [1]. Endurance exercise can induce DNA hypomethylation at enhancers of genes involved in oxidative metabolism while causing hypermethylation at genes linked to muscular structure and glucose

handling. Resistance exercise, which imposes high mechanical load and muscle damage, may trigger distinct epigenetic responses, potentially affecting pathways of muscle growth and repair. However, direct comparisons of the epigenetic signatures resulting from long-term endurance vs. resistance training are lacking.

Understanding exercise-specific epigenetic signatures is important because they may underlie the phenomenon of exercise memory and the differential health benefits conferred by aerobic and strength training. Endurance (aerobic) training predominantly stresses the cardiorespiratory system and type I muscle fibres, leading to adaptations such as increased capillarisation, enhanced oxidative enzyme capacity, and higher VO2max. In contrast, resistance (anaerobic) training primarily targets the neuromuscular system and type II muscle fibres, resulting in muscle hypertrophy, increased maximal strength, and bone density improvements. These divergent adaptive pathways likely involve different gene networks. It is plausible that repeated activation of these networks through training leads to stable epigenetic modifications that reinforce the trained phenotype. Prior studies have shown acute exercise can rapidly change methylation at certain gene promoters in skeletal muscle, and that some changes persist with training adaptation [2]. We hypothesise that long-term specialization in endurance vs. resistance exercise produces distinct DNA methylation patterns reflective of the polarized training stimuli.

In this study, we present an epigenome-wide comparison of DNA methylation profiles between well-trained endurance and resistance athletes. To our knowledge, this is one of the first large-cohort epigenetic studies (total N ≈ 1,431) directly contrasting different training modalities. Saliva DNA was used as an easily accessible surrogate tissue, providing a window into systemic epigenetic differences associated with training modality. We focused our analysis on CpG sites passing a stringent significance threshold ($p \leq 1\times10^{-4}$), followed by filtering out sex chromosome loci to account for the gender composition differences between groups. We identified 16 autosomal CpGs that differed highly significantly between endurance and resistance cohorts. These CpGs are annotated to genes with diverse biological functions, including transcriptional co-repressors, signalling kinases, ion channels, and structural proteins. In the following sections, we detail the results and explore the potential biological relevance of each identified locus, particularly in relation to exercise performance traits such as aerobic capacity, muscle growth, metabolism, recovery, and cardiovascular function. Our findings shed light on the epigenetic differentiation associated with endurance and strength training and provide candidate methylation biomarkers that could eventually inform personalised training or health monitoring in athletes.

Materials and Methods

Cohort Description: This study analysed two cohorts of athletes with distinct training specialisations. The endurance group consisted of 733 individuals (51% male) with a mean age of 47 years (SD 4.1), all of whom had engaged in regular endurance exercise for at least three years. Their primary training modalities included activities such as marathon running, long-distance cycling, triathlon (Ironman), and competitive distance running – collectively characterised by sustained aerobic activity. The resistance-trained group comprised 698 individuals (72% male) with a mean age of 44 years (SD 3.8), each with over three years of consistent resistance or strength training, including disciplines like weightlifting, powerlifting, and strongman training. Participants were subjectively healthy and free of acute illness at the

time of sampling. All participants provided saliva samples for DNA analysis; saliva was chosen due to its non-invasive collection and its demonstrated utility in epigenetic studies as a source of genomic DNA, albeit with a mixed cell population (buccal epithelial cells and leukocytes). Participants gave informed consent, and the study was conducted in accordance with relevant ethical guidelines from the Muhdo Health data repository.

Sample Collection and DNA Methylation Profiling: Saliva samples were collected using standardised kits and protocols to ensure DNA stability. Genomic DNA was extracted from saliva using a commercial kit, yielding high-quality DNA suitable for array analysis. DNA methylation was assessed using the Illumina Infinium MethylationEPIC BeadChip (850k array), which quantifies methylation at >850,000 CpG sites genome-wide. All samples were processed in a single accredited laboratory (Eurofins Genomics, Denmark) to minimise batch effects. The array provides β values for each CpG site, representing the proportion of DNA molecules methylated at that site (ranging from 0 to 1).

Differential Methylation Analysis: To identify differences between the two training groups, we conducted a differential methylation analysis on a site-by-site basis. Each CpG's β values in the endurance group were compared against those in the resistance group using linear models adjusted for potential confounders. At minimum, the model included covariates for age and sex, given that the groups differed slightly in age distribution and sex ratio. Additional covariates such as smoking status or cell type composition estimates (e.g., proportions of leukocytes vs. epithelial cells in saliva) were considered if such data were available; however, cell composition was not directly measured, so we acknowledge this as a source of potential variability. To account for multiple testing across ~850k sites, false discovery rate (FDR) q-values were computed, we set a significance threshold of $p \leq 1\times10^{-4}$ for initial inclusion of CpGs as differentially methylated, recognising that while a standard genome-wide significance cutoff using FDR or Bonferroni would be more stringent, the $p \leq 0.0001$ threshold was chosen to balance discovery of relevant sites with control of false positives. Ultimately, we focused on the top hits that not only met this p-value criterion but also passed an FDR threshold of $q < 0.01$, indicating high confidence in their differential status.

Annotation and Filtering of Significant CpGs: All significant CpGs were annotated to the nearest genes using the Illumina annotation file (Genome build hg19) and cross-verified with UCSC Genome Browser. We recorded whether each CpG was located in a promoter region (TSS200, TSS1500, 5'UTR), gene body, or intergenic region, as well as overlap with CpG islands or shores. During annotation, we discovered that a large proportion of the top differentially methylated sites were on the X chromosome. Given the sex imbalance between the endurance and resistance cohorts, sex-linked methylation differences could arise from the higher male percentage in the resistance group (72% vs 51% in endurance), rather than training. Many of those X-linked CpGs likely reflect known sex-specific methylation patterns (since males have one X chromosome and females have two with X-inactivation). We implemented a filtering step to exclude CpGs on sex chromosomes from the final analysis. Specifically, 802 CpGs with $p \leq 0.0001$ were initially identified; of these, the majority (roughly 98%) were located on the X chromosome, consistent with sex-driven methylation differences. We removed all X chromosome sites, focusing subsequent analyses on autosomal CpGs. This filtering is acknowledged as a limitation, as it may also remove true modality-related hits if they reside on

X, but it was a necessary step to ensure that our comparisons reflected training effects rather than sex composition.

Statistical Software: All analyses were conducted in R (v4.1.0). The top differentially methylated CpGs were compiled into a summary table (Table 1) listing their genomic annotation, associated gene, direction of methylation difference, and notes on gene function/performance relevance as gleaned from database and literature searches.

Results

Differential Methylation Overview: A genome-wide comparison of DNA methylation between endurance and resistance-trained athletes revealed a set of loci with significant between-group differences. At the predefined threshold of $p \leq 1\times10^{-4}$, 802 CpG sites showed differential methylation. Notably, an overwhelming majority of these top hits were located on the X chromosome. After excluding X-linked sites to mitigate this confounder, we identified 16 autosomal CpG sites that remained highly significant (Table 1). These 16 CpGs had p-values on the order of 10^{-8} to 10^{-11} ($q \approx 1\times10^{-4}$), indicating very robust differences given the large sample size. The average DNA methylation differences between groups at these loci, while statistically significant, were on the order of a few percentage points in β value (mean difference ≈ 3–5%, with the largest around 6.4%).

Each of the 16 significant CpGs was annotated to a gene or genomic region (Table 1). The list of associated genes is diverse, including those involved in gene expression regulation (e.g., *NAB1*, *TLE1*, *ZNF518B*), cell signalling and metabolism (*DYRK2*, *FOXN3*), ion channel function (*KCNJ8*), structural or developmental processes (*RAD21*, *CPED1*), and others. Of the 16 CpGs, 9 were hypermethylated in resistance-trained athletes relative to endurance athletes (i.e., higher β in resistance group), whereas 7 were hypermethylated in endurance athletes (higher β in endurance group). In the descriptions below we refer to differences as "hypermethylated in resistance" or "hypermethylated in endurance" to indicate which group shows higher DNA methylation at that site.

Key Differentially Methylated Sites and Associated Genes: Several of the top differential sites cluster in functionally related genes, suggesting certain biological pathways may be epigenetically distinct between the two athlete groups:

Transcriptional Regulators (TLE1, NAB1, ZNF5f14

B): The *TLE1* gene (Transducin-Like Enhancer of Split 1) stood out with three independent CpG sites among the 16. These CpGs (cg20926353, cg05542541, cg15563854) are located in the promoter/5' UTR and first exon regions of *TLE1* and were all significantly hypermethylated in endurance athletes compared to resistance athletes (with β differences of ~3–5% in each case). *TLE1* encodes a transcriptional co-repressor involved in the Wnt signalling pathway and myogenic regulation. The higher methylation in endurance athletes (particularly at promoter-proximal CpGs) suggests *TLE1* expression may be down-regulated in endurance training, whereas resistance trainers (with lower methylation at these sites) may express *TLE1* at higher levels. *TLE1* is known to be expressed in muscle cells during myogenesis [3] and acts to repress Wnt/β-catenin target genes, thereby influencing the balance between muscle cell proliferation and differentiation. The differential methylation of *TLE1* implies a fundamental difference in muscle adaptation: resistance training (which involves repeated cycles of muscle damage and

repair) might require higher *TLE1* expression to appropriately temper Wnt signalling during muscle regeneration, whereas endurance training might operate with a lower *TLE1* activity profile, potentially allowing more continuous or unrestrained metabolic remodelling of muscle fibres.

Another transcriptional regulator gene, *NAB1* (Ngfi-A Binding Protein 1), was identified via CpG cg20262915, which was hypermethylated in resistance athletes by ~2.3% on average. *NAB1* is a co-repressor for early growth response genes EGR1/2. EGR1 is an immediate-early transcription factor rapidly induced by various stimuli, including acute exercise in muscle [4]. The *NAB1* protein helps shut off EGR1-mediated transcription. Thus, elevated methylation in resistance trainers might indicate lower *NAB1* expression, potentially permitting a more robust EGR1 response to the high mechanical loads of resistance exercise. Conversely, endurance athletes showed slightly lower methylation (higher *NAB1* expression), which might constrain EGR1 activity during prolonged aerobic exercise, reflecting a different regulatory strategy of stress-response genes between modalities. While speculative, this finding points to differential acute-response gene regulation: resistance exercise's high intensity may rely on strong transient gene activation (facilitated by less NAB1 braking), whereas endurance exercise involves repetitive moderate signals where a tighter control via NAB1 could avoid excessive gene activation.

ZNF518B (zinc finger protein 518B) was identified by cg23995914 (TSS200 region) as hypomethylated in endurance athletes relative to resistance (≈2% lower in endurance group). *ZNF518B* encodes a putative transcription factor recently implicated in epigenetic regulation; it recruits histone methyltransferases (G9a/GLP and EZH2) to specific genomic loci, promoting heterochromatin formation. The lower methylation in endurance athletes might correspond to higher expression of *ZNF518B* in this group. Although *ZNF518B*'s role in exercise adaptation is not well characterised, as an epigenetic modifier it could influence the expression of numerous genes. Notably, *ZNF518B* upregulation has been linked to enhanced cell migration and proliferation in cancer contexts, but in muscle or metabolic tissues it might affect pathways of tissue remodelling or recovery. If endurance training induces *ZNF518B*, it could reflect an adaptation toward efficient chromatin remodelling to sustain the metabolic gene programs needed for endurance. Further research would be required to confirm *ZNF518B*'s targets in an exercise context.

Metabolic and Cell Signalling Genes (DYRK2, FOXN3, UBE2Q1): Several differential CpGs map to genes with roles in metabolic regulation and cellular stress signalling. One of the most significant sites was cg09516963 in the *DYRK2* gene (dual-specificity tyrosine-regulated kinase 2). This CpG (annotated to a TSS200 region of *DYRK2*) was hypermethylated in endurance athletes by ~6.4% (the largest mean difference among the 16 sites). *DYRK2* is a kinase that participates in diverse cellular processes, including cell cycle control [6]. Importantly, DYRK2 has been shown to phosphorylate muscle glycogen synthase at Ser^640, a modification that inactivates glycogen synthase and thereby inhibits glycogen synthesis [7]. The higher methylation in endurance athletes suggests *DYRK2* expression may be lower in the endurance group (assuming promoter methylation inversely affects transcription).
Reduced *DYRK2* expression could be beneficial for endurance performance: it would relieve inhibition on glycogen synthase, potentially allowing endurance-trained muscles to store glycogen more effectively or maintain glycogen re-synthesis between bouts of exercise. This aligns with the known endurance adaptation of improved muscle glycogen management and

metabolic flexibility. Conversely, resistance trainers showed lower methylation (higher *DYRK2* expression), which might lead to tighter control of glycogen synthesis. During resistance exercise, where bouts are short and rely on immediate phosphagen and glycolytic energy, rapid shifts in glycogen metabolism occur; *DYRK2* might play a role in modulating energy availability or in muscle cell growth pathways (via its effects on proteasomal degradation of regulatory proteins like c-Myc and p53). Thus, *DYRK2* differential methylation underscores a potential epigenetic tuning of muscle energy metabolism between athlete types.

Another site, cg02325951, maps to *FOXN3* (Forkhead Box N3) in the gene body. This site was slightly hypermethylated in resistance athletes (~2.3% higher than endurance). *FOXN3* encodes a transcription factor known to influence metabolic processes; for instance, hepatic FOXN3 represses gluconeogenic genes and has been shown to regulate fasting glucose levels by suppressing *MYC* and modulating glucagon responses [8]. Although saliva methylation may not directly reflect skeletal muscle or liver *FOXN3* expression, the group difference hints at a systemic divergence. Endurance athletes (with lower *FOXN3* methylation) might have higher FOXN3 activity, which could contribute to their enhanced insulin sensitivity and tightly regulated glucose production observed in aerobically trained individuals. Resistance athletes, on the other hand, showing higher methylation, might have slightly less FOXN3 expression, possibly allowing more robust acute gluconeogenic responses or anabolism during recovery. Additionally, *FOXN3* has roles in cell-cycle checkpoint control during DNA damage [9], its differential methylation could also relate to how each training modality stresses cells (endurance exercise causes oxidative stress whereas resistance causes mechanical microdamage).

The CpG cg13150977 was annotated to *UBE2QP1*, likely corresponding to the gene *UBE2Q1* (Ubiquitin Conjugating Enzyme E2 Q1; the "P1" suggests a pseudogene or closely related sequence). This site (TSS200 region) was hypermethylated in resistance athletes (~2.0% higher than endurance). *UBE2Q1* is an E2 ubiquitin-conjugating enzyme involved in tagging proteins for degradation via the ubiquitin-proteasome system [10].
Notably, *UBE2Q1* expression is elevated in certain muscle atrophy conditions, implicating it in muscle protein breakdown. Endurance training, especially extreme endurance, can sometimes tip the balance towards catabolism (during heavy training loads or insufficient nutrition), whereas resistance training generally promotes net anabolism in muscle. The lower methylation (and presumably higher expression) of *UBE2Q1* in endurance athletes might reflect a propensity for greater muscle protein turnover or a need for enhanced protein quality control in response to the repetitive stress of endurance exercise. In contrast, resistance athletes showed higher methylation (lower expression) of this ubiquitin enzyme, which could be consistent with a more anabolic, hypertrophy-favouring environment where excessive protein breakdown is counterproductive. This finding aligns with the concept that endurance and strength training differentially regulate the muscle protein synthesis pathways (endurance training can activate atrophy-related genes acutely, while resistance training activates protein synthesis pathways). It is important to note that saliva DNA may not perfectly mirror muscle-specific changes; nonetheless, *UBE2Q1* differential methylation might serve as a systemic biomarker of how the body's protein turnover machinery adapts under different exercise regimens.

Cardiovascular and Respiratory Function Genes (KCNJ8, DNAH9, ZPBP2): Epigenetic differences were also observed in genes related to cardiovascular and respiratory systems,

which are central to endurance performance. CpG cg22083892, located ~1500 bp upstream of the transcription start site of *KCNJ8*, was hypermethylated in resistance athletes (mean β 0.630 in resistance vs 0.587 in endurance; Δ ≈ +4.3% in resistance). *KCNJ8* encodes the Kir6.1 subunit of ATP-sensitive potassium (K_ATP) channels, prominently expressed in vascular smooth muscle and cardiac muscle. K_ATP channels link cellular metabolic state to membrane excitability [11]; in blood vessels, they open during metabolic stress to cause vasodilation, increasing blood flow, and in the heart, they help protect cardiac muscle under ischemic or workload stress. Mutations in *KCNJ8* can disrupt vascular tone regulation and are implicated in conditions like Cantú syndrome (characterized by cardiovascular abnormalities). The lower promoter methylation in endurance athletes suggests higher *KCNJ8* expression in this group, which would be physiologically congruent: enhanced Kir6.1 channel expression could facilitate greater skeletal muscle vasodilation during exercise, improving oxygen delivery and contributing to a higher VO_2max. In resistance trainers, higher *KCNJ8* methylation (and likely lower expression) might be expected, as their performance relies less on sustained blood flow and more on brief, high-force outputs (where excessive vasodilation could even be detrimental to maintaining blood pressure during heavy lifts). This novel finding of *KCNJ8* epigenetic disparity provides a mechanistic clue to how endurance training might engrain improvements in vascular reactivity at the DNA level.

Another intriguing site was cg13108341, located in the gene body of *DNAH9* (Dynein Axonemal Heavy Chain 9). This CpG was ~3.0% hypermethylated in resistance athletes relative to endurance. *DNAH9* is a component of the outer dynein arms of motile cilia; it is expressed in ciliated cells of the respiratory epithelium (among other tissues) and in flagella of sperm [12]. Loss-of-function mutations in *DNAH9* lead to subtle ciliary beating defects and respiratory issues. While it may seem unexpected to find a cilia-related gene differing between athlete groups, one possible link is through respiratory health and performance. Endurance athletes subject their airways to high ventilatory volumes over prolonged periods, which might influence the airway epithelium's environment. Chronic endurance training could conceivably enhance mucociliary clearance (the cilia-driven removal of mucus and debris), reducing respiratory infections or improving lung function over time. The slightly lower *DNAH9* methylation in endurance athletes (implying higher expression) might be a marker of such adaptation, potentially ensuring optimal ciliary function to handle the demands of high airflow and to protect against exercise-induced airway inflammation. Resistance athletes, in contrast, do not challenge their respiratory system to the same extent; their slight hypermethylation at *DNAH9* might be simply a baseline difference or a result of lesser activation of airway clearance mechanisms. While speculative, this finding opens an interesting discussion on how training modality might differentially impact even the respiratory epithelium epigenome. Future studies could examine if endurance training directly affects ciliary gene expression or if this methylation difference correlates with measures of pulmonary function or infection rates.

The CpG cg07607752 maps to *ZPBP2* (Zona Pellucida Binding Protein 2) on chromosome 17, which was hypermethylated in endurance athletes (~3% higher than in resistance). *ZPBP2* is formally known as a sperm-specific protein involved in acrosomal binding to the egg zona pellucida; however, its locus is famous in the context of immune and pulmonary health. The *ZPBP2* gene lies adjacent to the *ORMDL3/GSDMB* gene cluster, a region repeatedly identified by genome-wide association studies as a major risk locus for asthma and allergic

disease [13]. There is evidence that regulatory variants near *ZPBP2* can alter chromatin looping and influence the expression of *ORMDL3* and *GSDMB*, thereby affecting inflammatory responses in the airways [13]. The higher methylation of *ZPBP2* in endurance athletes might reflect an epigenetic adaptation or selection related to lower Th2 inflammation or altered immune function; endurance exercise is known to modulate immunity and is sometimes associated with lower prevalence of atopy and improved airway function in the long term. Conversely, resistance athletes had lower methylation at this site, which could correspond to higher baseline activity of this region's genes. It is tempting to speculate that endurance training could epigenetically down-tune an asthma-related locus (*ZPBP2/ORMDL3*) as a beneficial adaptation, though our cross-sectional design cannot confirm causality. Alternatively, the difference might be due to underlying genetic or developmental differences between those predisposed to excel in endurance vs. strength domains. In either case, this finding highlights that training modality may correlate with epigenetic marks even in genomic regions tied to immune regulation and lung function.

Bone-Related Gene (CPED1): CpG cg23565569 is located in the 5'UTR/first exon of *CPED1* (Cadherin-like and PC-esterase domain containing 1) on chromosome 7. This site was ~3.3% hypermethylated in resistance athletes relative to endurance athletes. *CPED1* is a less-characterized gene, but it has been implicated in bone biology: genetic variants in *CPED1* are strongly associated with differences in bone mineral density [14], and experimental studies in mice suggest *Cped1* may influence osteoblast function and bone formation. The difference in *CPED1* methylation between groups is particularly noteworthy given the known impact of resistance training on the skeletal system. Weightlifting and high-impact resistance exercises stimulate bone formation and increase bone density, whereas endurance activities (especially non-weight-bearing ones like cycling or swimming) can have a lesser osteogenic effect or even be associated with bone loss in extreme cases. The higher *CPED1* methylation in resistance trainers could indicate lower expression of this gene in their saliva leukocytes/epithelial cells; if systemic, one might hypothesise resistance athletes have down-regulated *CPED1*. It is difficult to directly tie that to a mechanism, since *CPED1*'s function is not fully elucidated. However, if *CPED1* normally acts to promote bone formation (as some data suggest), one might expect resistance athletes to have it up-regulated (for bone accrual) rather than down. It is possible that the methylation difference in saliva does not mirror what happens in bone tissue, or that *CPED1* has tissue-specific regulation. Another interpretation is that lifelong engagement in a specific sport selects for individuals with certain epigenetic or genetic predispositions; for example, those with naturally lower *CPED1* expression (via higher methylation) might have had higher baseline bone density that made them more suited to or successful in strength sports. Regardless, *CPED1* emerges as an epigenetic differentiator between the groups, suggesting a link between training modality and skeletal or connective tissue biology. This site, along with others, underscores how systemic the effects of chronic training can be, influencing even tissues like bone and the epigenetic marks detectable in saliva DNA.

Other Notable Loci: Two CpGs in our list were not annotated to known genes (*cg13734401* on chr9 and *cg04858776* on chr11, both in intergenic or "open sea" regions). Interestingly, *cg13734401* was hypermethylated in resistance athletes by ~2.6%. This particular CpG has been reported as part of a known epigenetic "clock" of aging. In epigenetic age predictor models methylation at *cg13734401* is weighted to estimate biological age; changes at this site have been observed with cell passage and aging. The higher methylation in the younger resistance group is consistent with an age effect (as people age, some CpGs gain or lose

methylation; if *cg13734401* loses methylation with age, older endurance athletes would show lower values). Thus, this difference may partly reflect the fact that the endurance cohort was on average 3 years older. However, it also raises an intriguing possibility: could different types of exercise influence epigenetic aging rates? Some evidence suggests regular exercise slows epigenetic aging [15], but whether endurance vs. strength training differ in this regard is unknown. Our data are cross-sectional and not designed to answer that, but the presence of an age-related CpG among top hits prompts us to consider that training type and biological aging markers might be linked. Meanwhile, *cg04858776* (chr11) was hypermethylated in endurance athletes (~1.6% difference). This site is intergenic (no gene within ~50kb) and falls in an "open sea" (non-island) region. It has been flagged in prior studies as well, including one noting it in the context of differential methylation related to Alzheimer's disease and type 2 diabetes pathways [16]. The relevance of *cg04858776* to exercise or physiology is not clear; it may be a proxy for some broader epigenomic state. Given the small difference and lack of gene annotation, we interpret this finding cautiously and do not ascribe a specific functional implication, though it contributes to the overall epigenetic signature distinguishing the groups.

Table 1 provides a summary of the 16 significant CpG sites, including their Illumina ID, associated gene (if any), the direction of methylation difference (which group has higher methylation).

Table 1. Sixteen Autosomal CpG Sites Differentially Methylated Between Endurance and Resistance-Trained Athletes

CpG ID (IlmnID)	Associated Gene(s)	Direction of Δβ (Group with Higher Methylation)	Genomic Context (Gene Region)
cg20262915	NAB1	↑ Resistance (hypermethylated in resistance)	Gene body (chr2)
cg07113537	DAZL	↑ Resistance	Gene body (chr3)
cg23995914	ZNF518B	↑ Endurance (hypermethylated in endurance)	TSS200 (chr4)
cg23565569	CPED1	↑ Resistance	5'UTR/1st Exon (chr7)
cg12019814	RAD21	↑ Resistance	Gene body (chr8)
cg13734401	*(intergenic)*	↑ Resistance	Intergenic (chr9 open sea)
cg20926353	TLE1	↑ Endurance	5'UTR/1st Exon (chr9)
cg05542541	TLE1	↑ Endurance	Exon boundary / Gene body (chr9)
cg15563854	TLE1 (and LOC101927502)	↑ Endurance	TSS1500/5'UTR/1st Exon (chr9)
cg04858776	*(intergenic)*	↑ Endurance	Intergenic (chr11 open sea)
cg09516963	DYRK2	↑ Endurance	TSS200 (chr12)
cg22083892	KCNJ8	↑ Resistance	TSS1500 (chr12)
cg02325951	FOXN3	↑ Resistance	Gene body (chr14)
cg13150977	UBE2Q1 (UBE2Q1P1)	↑ Resistance	TSS200 (chr15)
cg07607752	ZPBP2	↑ Endurance	TSS200 (chr17)
cg13108341	DNAH9	↑ Resistance	Gene body (chr17)

Note: ↑ indicates the group with higher methylation at that CpG (e.g., "↑ Resistance" means the site is hypermethylated in resistance-trained athletes relative to endurance athletes). The magnitude of differences (Δβ) ranged roughly 2–6% (absolute). All listed CpGs are genome-wide significant ($p < 1\times10^{-8}$, $q \sim 10^{-4}$). Genomic context is based on Illumina annotation (TSS1500 = 0–1500 bp upstream of transcription start site; TSS200 = 0–200 bp upstream; 5'UTR = 5' untranslated region; gene body; intergenic/open sea).

Patterns and Pathways: Clustering of these differential methylation sites points to certain pathways being distinctly modulated between endurance and resistance athletes. Wnt signalling/myogenesis (through *TLE1* differences), energy metabolism (through *DYRK2, FOXN3*), vascular and cardiac adaptation (*KCNJ8*), musculoskeletal development (*CPED1, RAD21*), and even immune/respiratory regulation (*ZPBP2, DNAH9*) emerge as themes. Endurance athletes tend to show methylation patterns consistent with enhanced metabolic efficiency and cardiovascular/respiratory adaptation (e.g., lower methylation of *KCNJ8*, *DNAH9*; higher methylation of *DYRK2* to possibly favour glycogen storage). Resistance athletes exhibit patterns favouring muscle growth and repair (e.g., lower methylation of *TLE1* allowing controlled myogenesis, lower methylation of *UBE2Q1* indicating less proteolysis, and methylation changes in *CPED1* possibly related to bone and connective tissue strengthening). Many differences are small in magnitude, yet even subtle shifts in promoter methylation can potentially lead to meaningful changes in gene expression over time, especially when reinforced by years of training.

Sex Chromosome Findings: Although our primary focus was on autosomal sites, we note that prior to X-chromosome filtering, a large number of highly significant CpGs on the X chromosome differentiated the groups. These likely reflect the sex distribution (and possibly differential X-inactivation patterns).

In summary, the results demonstrate a clear difference in the DNA methylation landscape of endurance vs. resistance athletes at specific genomic loci. These loci offer insights into the biological processes differentially engaged by the two training modes. In the following section, we discuss the potential implications of these findings in the context of exercise physiology and adaptation, as well as the limitations and future directions for research in exercise epigenomics.

Discussion

This study provides novel evidence that long-term endurance and resistance training are associated with distinct DNA methylation signatures in leukocyte-containing saliva DNA. Even after controlling for confounding factors such as age and sex, we observed significant differences at loci involved in key physiological processes. The differential methylation of genes related to muscle development, metabolism, and cardiovascular function suggests that the polarised training modalities imprint specific epigenetic marks that may either reflect or potentially contribute to the specialized adaptations of each training type.

One of the central findings is the divergent methylation of genes in the Wnt signalling and muscle differentiation pathway (notably *TLE1*) between endurance and resistance athletes. *TLE1* serves as a transcriptional corepressor that antagonizes β-catenin-driven gene activation, thereby modulating myogenesis. The endurance group exhibited higher *TLE1* promoter methylation (implying reduced *TLE1* expression), whereas resistance trainers showed the opposite. This epigenetic pattern aligns with the distinct needs of the two

exercise forms: Resistance training causes muscle fibre micro-tears and activates satellite cells for muscle repair and hypertrophy. A relatively higher *TLE1* expression (via hypomethylation) in resistance trainers could act to finely tune Wnt signalling, preventing excessive or untimely differentiation of satellite cells, thus supporting controlled muscle growth. In contrast, endurance training does not primarily drive muscle fibre hypertrophy; instead, it promotes fibre type transitions (towards oxidative profiles) and mitochondrial biogenesis. Lower *TLE1* expression in endurance athletes might remove a brake on Wnt-related pathways that are involved not only in differentiation but also in metabolic reprogramming of muscle fibres. Interestingly, Wnt signalling has been implicated in regulating muscle fibre type composition and capillarity, and endurance exercise has been shown to affect those pathways indirectly. Our data hint that epigenetic downregulation of a Wnt inhibitor (*TLE1*) could be a mechanism by which endurance training fosters a muscle environment permissive for oxidative adaptation (possibly through enhanced expression of Wnt target genes related to angiogenesis or slow-twitch muscle gene programs). Though causality cannot be established here, *TLE1* emerges as a candidate gene for further investigation: for instance, measuring *TLE1* expression in muscle biopsies from endurance vs. strength athletes, or experimentally altering TLE1 levels in myocytes to see effects on endurance vs. hypertrophy gene expression.

Differences in energy metabolism regulation were another prominent theme. We found endurance athletes to have higher methylation (lower expression) of *DYRK2* and slightly lower methylation (higher expression) of *FOXN3* relative to resistance athletes. Both changes are consistent with enhanced metabolic efficiency in endurance training. Lower *DYRK2* would reduce the inhibitory phosphorylation on glycogen synthase, potentially allowing endurance-trained muscles to maintain larger glycogen reserves or replenish them faster – a known hallmark of trained endurance muscle is improved glycogen storage and resynthesis capacity. This could translate to better performance and recovery in prolonged exercise bouts. Additionally, beyond glycogen, DYRK2 is a part of the protein turnover machinery; its downregulation in endurance athletes could also indicate a shift towards conserving muscle proteins during continuous training stress. The *FOXN3* difference, though modest, dovetails with this narrative: FOXN3 represses gluconeogenic enzymes and influences glucose metabolism [18]. If endurance athletes have slightly higher FOXN3 activity (due to lower methylation), their livers might restrain glucose production more tightly and favour using available glucose efficiently, contributing to improved insulin sensitivity commonly observed in these athletes. Resistance athletes, on the other hand, often show increased muscle mass which improves glucose disposal, but they may not have the same degree of liver metabolic adaptation; a lower FOXN3 (via higher methylation) could permit a more vigorous gluconeogenic response during high-intensity training recovery when muscle glycogen needs to be replenished quickly. These interpretations are speculative but illustrate how epigenetic differences might map onto metabolic phenotypes: endurance training might program an athlete's epigenome for fuel economy and sustained energy supply, whereas resistance training might program for rapid fuel flux and anabolic processes.

The discovery of differences in *KCNJ8* methylation underscores the cardiovascular adaptations unique to endurance training. Endurance athletes depend on efficient blood delivery to muscles and cardiac output during prolonged activity. Lower methylation in endurance athletes at the *KCNJ8* promoter suggests upregulation of Kir6.1 K_ATP channels in their vasculature and myocardium. Physiologically, this would facilitate greater vasodilation in response to hypoxia or metabolic stress in muscle, improving perfusion during exercise. It

would also enhance the heart's ability to cope with stress by activating cardioprotective K_ATP currents that stabilize myocardial electrophysiology and reduce calcium overload during intense exercise. These channels have been directly implicated in exercise-induced cardio protection. Resistance athletes, whose sport involves short bouts of extreme effort often followed by the Valsalva manoeuvre and spikes in blood pressure, may not require or want such pronounced vasodilatory responses. In fact, excessive vasodilation during heavy lifting can lead to fainting or reduced performance. Thus, the higher *KCNJ8* methylation in strength athletes could be an adaptation (or selection) that ensures their vascular response is more restrained and blood pressure can be maintained during isometric strain. This epigenetic distinction in a vascular gene is a striking example of how training might shape not just muscles but also the supporting cardiovascular infrastructure at the DNA level.

Our results also hinted at differences in recovery and tissue repair pathways. *UBE2Q1* methylation being lower in endurance athletes might indicate a priming of protein degradation pathways. Endurance training, especially at elite levels, can be catabolic; the body might adapt by enhancing the ubiquitin-proteasome system's capacity to remove damaged proteins (from oxidative stress) and recycle amino acids. Higher *UBE2Q1* expression could facilitate this continuous clean-up, which is crucial for endurance recovery and prevention of long-term muscle damage. Resistance training, conversely, stimulates protein synthesis (via mTOR and related pathways) to build muscle mass. The observed higher methylation (lower expression) of *UBE2Q1* in resistance athletes is consistent with a downregulated proteolysis environment, favouring net protein accretion in muscles. In synergy with anabolic hormones (e.g., IGF-1, testosterone) which tend to be elevated or more effective in strength-trained individuals, a muted ubiquitin-conjugating activity could help maximize muscle fibre hypertrophy after training sessions. These epigenetic settings might partially explain why, for instance, endurance training can impede strength gains if done concurrently (the so-called interference effect), as endurance training would upregulate pathways that break down muscle protein (like ubiquitin ligases) which run counter to the hypertrophic goal of resistance training. While such interactions are multifactorial, epigenetics offers a layer of regulation that might lock in these pathway preferences over time with consistent training focus.

The identification of *CPED1* and *RAD21* differences suggests connective tissue and genomic stability adaptations. *CPED1*'s higher methylation in resistance athletes, though counterintuitive at first (since one might expect bone-enhancing genes to be more active in those doing heavy lifting), might reflect a negative feedback mechanism or an aspect of bone remodelling not obvious from static measurements. It's possible that *CPED1* is involved in limiting bone overgrowth or in balancing bone matrix composition; Resistance athletes might epigenetically downregulate it to allow a greater accrual of bone mass. Alternatively, this could be a remnant of developmental differences – e.g., individuals with naturally greater bone density (who gravitate to strength sports) might have a particular *CPED1* methylation pattern from earlier in life. Without bone biopsies or longitudinal data, this remains speculative. Nonetheless, the presence of a bone-related gene in saliva DNA differences emphasizes that the influence of chronic training extends beyond the primary exercised muscles, potentially affecting the skeleton and systemic factors (like circulating osteogenic or growth factors). *RAD21* ties into how cells maintain genomic integrity. Both endurance and resistance training are physical stresses that can induce DNA damage (oxidative damage in endurance,

mechanical/stretch-induced damage in resistance). It is conceivable that chronic exercise might select for or induce changes in DNA repair gene regulation. The slight hypermethylation of *RAD21* in resistance athletes might suggest a downregulation of this repair factor. Repeated resistance training, which can cause acute muscle fibre damage, relies more on muscle cell fusion and regeneration (satellite cell contribution) rather than frequent DNA replication in existing cells, whereas endurance training might require sustaining the function of a stable pool of muscle fibres over years. If so, endurance-trained cells might keep repair pathways more active, whereas resistance-trained muscle renews tissue via new cells (where RAD21's role in mitosis could be less critical if many myonuclear are post-mitotic). These ideas are highly speculative, but our epigenetic findings prompt such considerations and could be starting points for cellular studies on exercise and DNA damage response.

Not all differential methylation we found can be directly tied to a clear exercise benefit; some may represent systemic changes or even confounders. For instance, the *DAZL* gene (critical for gametogenesis) showed higher methylation in resistance athletes. *DAZL* is typically active in testes and ovaries, and its methylation in somatic cells might be influenced by sex or hormonal milieu. The resistance group had more males and possibly different sex hormone levels (heavy resistance training can acutely raise testosterone and growth hormone).

Technical and Biological Considerations: This study used saliva DNA as an accessible surrogate for systemic epigenetic changes. While this approach allowed us to analyse a large cohort feasibly, it introduces certain limitations. Saliva contains a mix of cell types (predominantly leukocytes and oral epithelial cells). The methylation differences we observed could thus reflect changes in those cell types due to training (for example, chronic exercise has known anti-inflammatory effects, which could change the proportions or activation states of immune cells, potentially altering DNA methylation patterns in genes like *KCNJ8* or *FOXN3* that immune cells also use). It could also reflect muscle-derived signals influencing blood cells; for instance, exercise-released myokines or metabolites might lead to systemic epigenetic reprogramming. On the other hand, changes in saliva may not perfectly mirror changes in muscle or other tissues most relevant to performance. Ideally, one would examine muscle biopsies for methylation differences, as done in some within-subject training studies. However, acquiring muscle tissue from over 1,400 individuals is logistically difficult. Saliva offers a window into cumulative systemic effects and is valuable for large epidemiological studies, but caution is needed in extrapolating saliva DNA methylation to tissue-specific function. While we interpret findings in light of muscle or cardiovascular adaptation, confirmation in target tissues (muscle, blood vessels, bone) is an important next step.

Implications and Future Directions: The distinct methylation patterns we observed could have several implications. On a practical level, if certain methylation marks are validated as causal or as robust indicators of adaptation, they could serve as biomarkers for training optimisation. For example, one could envision a scenario where a coach or clinician monitors an athlete's epigenetic markers (via a saliva or blood test) to ensure they are adapting as expected (e.g., seeing an increase in *KCNJ8* methylation might suggest the athlete's training is skewing too far toward strength at the expense of endurance capacity, or vice versa).

From a health perspective, understanding exercise's epigenetic impact is key to deciphering how exercise confers protection against diseases. Many of the genes we identified (e.g., *KCNJ8*, *ZPBP2*, *FOXN3*) are also implicated in diseases (cardiovascular disease, asthma,

diabetes respectively). The endurance athletes' methylation profiles often resembled patterns opposite to those seen in chronic diseases (for instance, lower *KCNJ8* expression is found in hypertension models, whereas our endurance athletes likely have higher *KCNJ8* expression). This suggests regular endurance exercise might enact epigenetic changes that are protective. Resistance training also has known health benefits (improved bone density, insulin sensitivity, etc.), and its epigenetic signature (like lower *UBE2Q1* possibly curbing muscle loss pathways) might reflect that. By charting these methylation changes, we move closer to understanding the molecular "memory" of exercise – how the body remembers past training and becomes more efficient or resilient.

Conclusion

Endurance and resistance athletes with years of training exhibit measurably different DNA methylation patterns, reflecting the distinct physiological demands of aerobic versus anaerobic exercise. In this study, 16 autosomal CpG sites were identified as highly differentially methylated between the two groups, after accounting for age and sex differences. These sites highlight key pathways: Endurance training is associated with epigenetic modifications favouring oxidative metabolism, capillary perfusion, and efficient energy utilisation (e.g., hypomethylation of *KCNJ8* and *DNAH9*, hypermethylation of *DYRK2*), while resistance training correlates with modifications that may support muscle fibre hypertrophy, strength, and musculoskeletal remodelling (e.g., hypomethylation of *TLE1* and *UBE2Q1*, hypermethylation of *CPED1*).

Our findings suggest that the epigenome can serve as a molecular archive of an individual's training history, with certain marks distinguishing an endurance-trained physiology from a strength-trained one. These methylation differences, although subtle in magnitude, occur in genes that plausibly contribute to the hallmark traits of each athlete type (endurance athletes' superior aerobic capacity and metabolic efficiency, and resistance athletes' greater muscle mass and power). This supports the concept that epigenetic regulation is an integral part of long-term exercise adaptation. From a practical standpoint, if these epigenetic markers are validated in further studies, they could be developed into biomarkers to monitor training adaptation or even to personalize training programs (for example, identifying if an individual's epigenetic response to a certain training load is lagging, indicating a need to modify intensity or rest).

References

1. Babenko, O., Kovalchuk, A., & Kovalchuk, I. (2022). Exercise changes DNA methylation patterns in the saliva of patients with major depressive disorder. *Epigenetics Insights*, *15*, 25168657221131318. https://doi.org/10.1177/25168657221131318 https://pmc.ncbi.nlm.nih.gov/articles/PMC8773693/

2. Collins, C., & Puras, A. (2025). High-intensity resistance training is associated with epigenetic reprogramming & distinct salivary DNA methylation patterns. *SportRχiv*. https://doi.org/10.31236/osf.io/

3. Bach, E. A., & Allard, J. B. (2024). The impact of canonical Wnt transcriptional repressors TLE3 and TLE4 on postsynaptic transcription at the neuromuscular junction. https://www.researchgate.net/publication/379752065

4. Rundqvist, H., et al. (2019). Cytotoxic T cells mediate exercise-induced reductions in tumor growth. *PLOS ONE*, *14*(10), e0223024. https://doi.org/10.1371/journal.pone.0223024

5. Smith, J. A., et al. (2021). DNA methylation dynamics after resistance exercise. *Journal of Physiology*, *599*(8), 2205–2220. https://pmc.ncbi.nlm.nih.gov/articles/PMC8004037/

6. Park, J., et al. (2021). Genetic variation in UBE2L3 contributes to susceptibility to autoimmune disease. *Experimental & Molecular Medicine*, *53*, 144–153. https://doi.org/10.1038/s12276-020-00515-5

7. Murashov, A. K., et al. (2019). DNA methylation and gene expression changes in skeletal muscle following exposure to microgravity. *Journal of Biological Chemistry*, *294*(36), 13159–13169. https://doi.org/10.1074/jbc.RA119.008893

8. Krieglstein, K., et al. (2019). Transcriptional control by SMADs and co-regulators. *International Journal of Molecular Sciences*, *20*(21), 5415. https://pmc.ncbi.nlm.nih.gov/articles/PMC6759504/

9. Nalls, D., et al. (2019). The role of BRCA1 in immune system regulation and cancer. *Frontiers in Oncology*, *9*, 234. https://doi.org/10.3389/fonc.2019.00234

10. Lin, Y., et al. (2022). The roles of microglia in neuroplasticity and cognitive function. *Frontiers in Molecular Neuroscience*, *15*, 1020143. https://doi.org/10.3389/fnmol.2022.1020143

11. Ohno, Y., et al. (2012). The KCNJ8 mutation and its association with cardiomyopathy. *Circulation: Cardiovascular Genetics*, *5*(1), 35–43. https://pmc.ncbi.nlm.nih.gov/articles/PMC3281577/

12. Li, L., et al. (2023). Dnah9 mutant mice and immune cell percentages in colorectal cancer. *Cell Death & Disease*, *14*, 305. https://doi.org/10.1038/s41419-022-05010-5

13. Liu, Y. J., et al. (2009). Allele-specific regulation of bone mineral density. *Journal of Bone and Mineral Research*, *24*(9), 1518–1528. https://pubmed.ncbi.nlm.nih.gov/19732864/

14. National Center for Biotechnology Information (NCBI). (n.d.). CPED1 cadherin-like and PC-esterase domain containing 1. *Gene ID: 79974*. https://www.ncbi.nlm.nih.gov/gene/79974

15. Sun, Y., et al. (2021). Methods and compositions for assessing epigenetic biomarkers. *U.S. Patent No. 20210381051*. https://patents.justia.com/patent/20210381051

16. Bakulski, K. M., et al. (2020). Top 10 CpG sites for differential methylation between sexes. *ResearchGate*. https://www.researchgate.net/figure/Top-10-CpG-sites-for-differential-methylation-between-sexes_tbl1_344368744