



**VICTORIA UNIVERSITY**  
MELBOURNE AUSTRALIA

*Methylome and proteome integration in human skeletal muscle uncover group and individual responses to high-intensity interval training.*

This is the Published version of the following publication

Jacques, Macsue, Landen, Shanie, Romero, Javier Alvarez, Hiam, Danielle, Schittenhelm, Ralf B, Hanchapola, Iresha, Shah, Anup D, Voisin, Sarah and Eynon, Nir (2023) Methylome and proteome integration in human skeletal muscle uncover group and individual responses to high-intensity interval training. *The FASEB Journal*, 37 (10). ISSN 0892-6638

The publisher's official version can be found at  
<https://faseb.onlinelibrary.wiley.com/doi/10.1096/fj.202300840RR>  
Note that access to this version may require subscription.

Downloaded from VU Research Repository <https://vuir.vu.edu.au/47140/>

## RESEARCH ARTICLE

# Methylome and proteome integration in human skeletal muscle uncover group and individual responses to high-intensity interval training

Macsue Jacques<sup>1</sup>  | Shanie Landen<sup>1</sup>  | Javier Alvarez Romero<sup>1</sup> | Danielle Hiam<sup>1,2</sup>  | Ralf B. Schittenhelm<sup>3</sup>  | Iresha Hanchapola<sup>3</sup> | Anup D. Shah<sup>3</sup>  | Sarah Voisin<sup>1,4</sup>  | Nir Eynon<sup>1,5</sup> 

<sup>1</sup>Institute for Health and Sport (iHeS), Victoria University, Melbourne, Victoria, Australia

<sup>2</sup>Institute of Nutrition and Health Sciences, Deakin University, Melbourne, Victoria, Australia

<sup>3</sup>Monash Proteomics & Metabolomics Facility, Monash University, Melbourne, Victoria, Australia

<sup>4</sup>Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

<sup>5</sup>Australian Regenerative Medicine Institute, Monash University, Melbourne, Victoria, Australia

## Correspondence

Nir Eynon, Australian Regenerative Medicine Institute, Monash University, Clayton, Victoria 3800, Australia.  
Email: [nir.eynon@monash.edu](mailto:nir.eynon@monash.edu)

## Funding information

Department of Education and Training | Australian Research Council (ARC), Grant/Award Number: DP190103081; DHAC | National Health and Medical Research Council (NHMRC), Grant/Award Number: APP1140644, APP1194159 and APP1157732

## Abstract

Exercise is a major beneficial contributor to muscle metabolism, and health benefits acquired by exercise are a result of molecular shifts occurring across multiple molecular layers (i.e., epigenome, transcriptome, and proteome). Identifying robust, across-molecular level targets associated with exercise response, at both group and individual levels, is paramount to develop health guidelines and targeted health interventions. Sixteen, apparently healthy, moderately trained ( $\text{VO}_2 \text{max} = 51.0 \pm 10.6 \text{ mL min}^{-1} \text{ kg}^{-1}$ ) males (age range = 18–45 years) from the Gene SMART (Skeletal Muscle Adaptive Responses to Training) study completed a longitudinal study composed of 12-week high-intensity interval training (HIIT) intervention. Vastus lateralis muscle biopsies were collected at baseline and after 4, 8, and 12 weeks of HIIT. DNA methylation (~850 CpG sites) and proteomic (~3000 proteins) analyses were conducted at all time points. Mixed models were applied to estimate group and individual changes, and methylome and proteome integration was conducted using a holistic multilevel approach with the mixOmics package. A total of 461 proteins significantly changed over time (at 4, 8, and 12 weeks), whilst methylome overall shifted with training only one differentially methylated position (DMP) was significant ( $\text{adj.}p\text{-value} < .05$ ). *K*-means analysis revealed cumulative protein changes by clusters of proteins that presented similar changes over time. Individual responses to training were observed in 101 proteins. Seven proteins had large effect-sizes  $>0.5$ , among them are two novel exercise-related proteins, *LYRM7* and *EPN1*. Integration analysis showed bidirectional relationships between the methylome and proteome. We showed a significant influence of HIIT on the epigenome and more so on the proteome in human

**Abbreviations:** DEG, differentially expressed genes; DEP, differentially expressed proteins; DMP, differentially methylated position; DMR, differentially expressed regions; GXTs, graded exercise test; HIIT, high intensity interval training; LT, lactate threshold; SMART, skeletal muscle adaptive responses; sPLS, partial least squares;  $\text{VO}_{2\text{max}}$ , maximal oxygen uptake;  $W_{\text{peak}}$ , peak power output.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *The FASEB Journal* published by Wiley Periodicals LLC on behalf of Federation of American Societies for Experimental Biology.

muscle, and uncovered groups of proteins clustering according to similar patterns across the exercise intervention. Individual responses to exercise were observed in the proteome with novel mitochondrial and metabolic proteins consistently changed across individuals. Future work is required to elucidate the role of these proteins in response to exercise.

#### KEYWORDS

DNA methylation, epigenetics, exercise, proteomics, skeletal muscle

## 1 | INTRODUCTION

Exercise training confers many health benefits, including increased cardiorespiratory fitness, and a reduction in the risk for chronic diseases (e.g., type 2 diabetes, cardiovascular disease).<sup>1–4</sup> These benefits are underpinned by changes in genes and protein expression, leading to positive shifts in cell metabolism.<sup>5–9</sup> Identifying exercise-enhanced molecules is crucial for the development of targeted therapies to treat metabolic diseases.

High-throughput unbiased OMICs, including genomics, epigenomics, transcriptomics, and proteomics, are now widely used to identify exercise-induced pathways in healthy and diseased populations.<sup>10,11</sup> While genomic and transcriptomic control of exercise adaptations have been well described in both athletes and the general population,<sup>11–14</sup> epigenetics<sup>15–18</sup> and untargeted proteomics<sup>19,20</sup> studies are lacking.<sup>17,21</sup> Epigenetics shape cellular identity and lead to downstream change in gene expression, and can be modulated by environmental stimuli. The best-characterized epigenetic modification is DNA methylation,<sup>22</sup> which involves the addition of a methyl group (CH<sub>3</sub>) to cytosines in the DNA.<sup>23</sup> DNA methylation levels are altered in virtually all diseases, including diseases that are mitigated by exercise training (e.g., metabolic syndrome<sup>24</sup>). Whether exercise training improves muscle function via a remodeling of epigenetic patterns is still not well known, with only handful of studies suggesting DNA methylation patterns change following an exercise training intervention.<sup>17,18,25–31</sup> Furthermore, it remains currently unknown whether changes in DNA methylation translate to a change in protein expression levels.

Cellular processes are controlled by the activity of proteins, and high-throughput proteomics offers great potential to uncover the molecular mechanisms underlying exercise-induced adaptations.<sup>8</sup> Proteins are the end product of gene expression, and may be regarded as closest to physiological and functional profiles of cells and tissues.<sup>32</sup> For example, a recent time series analysis of multi-omics in blood revealed four different clusters of proteins that change following an acute bout of exercise (including proteins involved in inflammatory response, immune response, oxidative stress, etc.).<sup>6</sup> A handful of studies have

investigated changes in specific proteins in response to long-term exercise in skeletal muscle<sup>31–34</sup> and found that exercise increases the abundance of mitochondrial proteins,<sup>31,33</sup> and regulates proteins involved in excitation and contraction coupling.<sup>33</sup>

Individuals respond differently to similar exercise programs, suggesting that exercise training prescription may need to be tailored to individual genetic profile, sex, and other intrinsic factors that modulate exercise responses.<sup>4,35–37</sup> However, most exercise training interventions only assess exercise responses at the group level. Many exercise training studies that attempted to quantify inter-individual variability (i.e., variability observed between individuals) suffer from poor study designs that failed to quantify intra-individual variation (i.e., variability observed within same individual).<sup>38</sup> Several study designs are available to study individual response to training,<sup>37–41</sup> but only a handful of studies have properly implemented them,<sup>40,42–44</sup> and no study has explored individual responses in omics profile data to exercise training.

To gain a better understanding of group and individual responses to exercise training, we profiled 16 healthy men before and after a 12-week HIIT intervention, at two OMIC levels (DNA methylation and protein expression). Individuals were sampled at regular intervals (at baseline, 4, 8, and 12 weeks) during the course of the HIIT intervention, allowing for the quantification of individual responses over time. Finally, we integrated the two OMIC layers together and uncovered pathways that are consistently altered following HIIT at the methylome and proteome levels.

## 2 | METHODS

### 2.1 | Participants

Twenty participants from the Gene SMART (Skeletal Muscle Adaptive Responses to Training) study<sup>45</sup> commenced the study, and 16 participants completed 12 weeks of high-intensity interval training (HIIT). From the 20 initially recruited participants, 19 completed 4 weeks of HIIT (1 dropout), of these 18 completed

8 weeks (1 dropout), and 16 completed the full 12 weeks of HIIT (1 dropout and 1 exclusion due to inconsistent results (i.e., duplicate tests (represented by the blue arrows in Figure 1) provided more than 10% difference). The participants' demographic information is included in Table S1.

Participants were apparently healthy, moderately trained males ( $VO_{2max}$   $51.0 \pm 10.6 \text{ mL min}^{-1} \text{ kg}^{-1}$ ), aged 18–45 years old. The study was approved by the Victoria University Human Ethics Committee (HRE13-223 and HRE21-122) and written informed consent was obtained from each participant. Participants were excluded from the study if they had a past history of definite or possible coronary heart disease, significant chronic or recurrent respiratory condition, significant neuromuscular, major musculoskeletal problems interfering with the ability to cycle, uncontrolled endocrine and metabolic disorders or diabetes requiring insulin, and other therapies.<sup>45</sup>

## 2.2 | Study design

Participants were tested at baseline and after 4, 8, and 12 weeks of HIIT. In brief, HIIT sessions were conducted three times a week with 2-min intervals interspaced by 1-min rest. Intensity of training started at 60% of lactate threshold and increased by 10% each week every 4 weeks (i.e., first week 60%, fourth week 90%). To ensure progression, training intensity was re-adjusted every 4 weeks based on the newly determined Peak Power output ( $W_{peak}$ ) and Lactate Threshold (LT) from the Graded Exercise Test (GXTs), and training was repeated as previously described (i.e., starting with the new 60% lactate threshold and progressing to 90% of lactate threshold). These tests also allowed for monitoring individual progress of participants for the longitudinal analysis of training adaptations. To increase accuracy in measurement and to reduce biological day-to-day variability in participants' performance, physiological measures of fitness ( $W_{peak}$ , LT, and  $VO_{2max}$ ) were assessed from two GXTs conducted at each time point (Figure 1).

## 2.3 | Muscle biopsies

Biopsies were conducted in the morning 48 h after last exercise test. A controlled diet for 48 h prior to the muscle biopsies was provided to the participants, according to the guidelines of the Australian National Health & Medical Research Council (NHMRC). Muscle biopsies were taken by an experienced medical doctor using a Bergstrom needle from the vastus lateralis muscle of the participants' dominant leg, following local anesthesia (2 mL, 1% Lidocaine (Lignocaine)). The needle was inserted in the participant's leg and manual suction was applied for muscle collection. Excess blood from biopsy was removed using an absorbent sheet, and muscle was then immediately frozen in liquid nitrogen and stored at  $-80^\circ$  until further analyses. Muscle biopsies were collected at four timepoints (Pre, 4, 8, and 12 weeks) each in vertical line order for each timepoint, and sent for comprehensive analyses of DNA methylation and proteome (Figure 1).

## 2.4 | DNA extraction and DNA methylation analyses

DNA was extracted using the Qiagen All prep DNA/RNA kit (Cat/ID: 80204) according to manufacturer's guidelines. In brief, ~15 mg of muscle samples was homogenized and separated by a column into genomic DNA and RNA, and then eluted separately. RNA was stored at  $-80^\circ$  for future analyses while dsDNA concentrations were estimated with the Invitrogen Qubit Fluorometer. DNA methylation analyses were carried by the Illumina Infinium Methylation EPIC array (<https://www.illumina.com/products/by-type/microarray-kits/infinium-methylation-epic.html>) according to manufacturer's protocols. Briefly, the DNA input amount was 500 ng for bisulfite conversion. QC of bisulfite conversion was carried out by MPS (methylation-specific PCR) on specific regions. Once QC was evaluated, whole genome amplification was conducted followed by

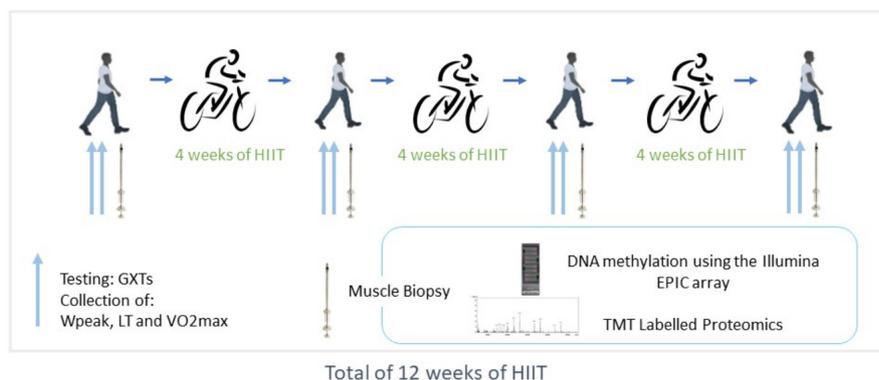


FIGURE 1 Study design.

Total of 12 weeks of HIIT

array hybridization, and single base extension before the array was scanned.

## 2.5 | Pre-processing and statistical analyses

Pre-processing of the DNA methylation data was conducted using the ChAMP analysis pipeline<sup>46</sup> in the R statistical software (<http://www.r-project.org>). Data were imported into R as IDAT file. Probes were filtered based on detection  $p$ -value  $>.01$  (default value for the `champ.load` function) and were filtered out if  $<3$  beads in at least 5% of samples, missing  $\beta$ -values, non-CpG probes, or if probes aligned to multiple locations in the dataset. To refine the dataset, probes associated with single nucleotide polymorphisms (SNPs) related to the “EUR” population as described in Zhou et al.<sup>47</sup> were excluded, along with probes situated on the sex chromosomes. This decision was driven by a pivotal aspect of our study, which involved the replication of the high-intensity interval training (HIIT) regimen after a wash-out period exceeding 1 year. Given that this repetition was only feasible with male participants, due to the ongoing initial training phase among female participants, an assortment of both male and female samples was present across different experimental batches. In light of this, the exclusion of sex chromosome-related probes was imperative to prevent potential bias during the batch adjustment phase, ensuring the integrity of subsequent analyses.  $\beta$ -values were then obtained as follows:

$$\beta - \text{value} = \frac{\text{intensity of the methylated allele}}{\text{intensity of the unmethylated allele} + \text{intensity of the methylated allele} + 100}$$

A  $\beta$ -mixture quantile normalization method was conducted to normalize the data for Type I and Type II probes, generated from the HMEPIC array. We then performed a singular value decomposition (SVD) analysis to identify variations in each individual dataset. Because chip position as well as batch were significantly identified as sources of variation, data were converted to  $M$ -values and then normalized for batch and chip position using the `ComBat` function in the `sva` package. Finally, a final quality check (QC) was performed on the data to ensure all sources of unwanted variability were accounted for.

Following normalization and QC procedures, we undertook differential methylation analysis by converting  $\beta$ -values to  $M$ -values since those present distributions that are more appropriate for statistical test and differential analysis of methylation levels<sup>48</sup>:

$$M - \text{value} = \log^2(\beta / (1 - \beta))$$

To identify exercise-induced timeline changes in DNA methylation, we used linear mixed models using the `lmerTest`<sup>49</sup> package. This part of the analyses included male participants who completed the longer (12 weeks) intervention. The models were built as follows:

$$\text{DNA methylation} = \text{Timepoint} + \text{Fitness} + \text{Age} + \text{random (ID)} \\ + \text{random (ID * Timepoint)}$$

where outcome was the DNA methylation probes, timepoint (PRE, 4WP, 8WP, and 12WP), covariates in this case were age and baseline fitness, and ID was individual identification for each participant. Individual changes were accessed by the random effect of the model. DMPs associated with either fixed (group level) or random (individual level) effects were considered significant if adjusted  $p$ -value  $<.05$ . If significant DMPs were identified, we proceeded with the analyses by identifying if such positions were associated with specific regions (DMRs: clusters of DMPs that presented consistent DNA methylation change with fitness measure), for this, we used the `DMRcate`<sup>50</sup> package. To identify fitness-associated pathways, we performed a gene set enrichment analysis using two methods: 1—The `champ.ebGSEA` function which does not consider pre-determined significance of DMPs, but instead performs its own analyses across multiple genes identified by DNA methylation probes and phenotype selected; 2—The `GOMeth` methodology which requires pre-specified significant positions.<sup>51</sup> All pathways with an adjusted  $p$ -value  $<.05$  were consid-

ered significant and  $p$ -values obtained from both analyses were adjusted using the Benjamini and Hochberg method (BH).

Plots were created utilizing the following packages: `ggplot2`,<sup>52</sup> `ggpubr`,<sup>53</sup> `complexHeatmaps`,<sup>54</sup> and `FactorMiner`.<sup>55</sup> All analyses were performed using R software version 4.0.2.

### 2.5.1 | Proteomic extraction and analyses

Muscle tissue was lysed in 300- $\mu$ L SDS solubilization buffer (5% SDS, 50mM TEAB, pH 7.55), heated at 95°C for 10 min, and then probe-sonicated before measuring the protein concentration using the BCA method. A total protein amount of 100  $\mu$ g (suspended in 50  $\mu$ L) was used for each sample for subsequent analyses. The lysed samples were denatured and alkylated by adding TCEP

(Tris(2-carboxyethyl) phosphine hydrochloride) and CAA (2-Chloroacetamide) to a final concentration of 10 and 40 mM, respectively, and the mixture was incubated at 55°C for 15 min. Sequencing grade trypsin was added at an enzyme-to-protein ratio of 1:50 and incubated overnight at 37°C after the proteins were trapped using S-Trap mini columns (Profi). Tryptic peptides were eluted from the columns using (i) 50 mM TEAB, (ii) 0.2% formic acid, and (iii) 50% acetonitrile, 0.2% formic acid. The fractions were pooled, concentrated in a vacuum concentrator, and reconstituted in 40  $\mu$ L 200 mM HEPES, pH 8.5. Using a Pierce Quantitative Colorimetric Peptide Assay Kit (Thermo Scientific), equal amounts of peptides for each sample were labeled with the TMTpro 16plex reagent set (Thermo Scientific) according to the manufacturer's instructions, this labeling strategy was implemented to minimize channel leakage. Individual samples were then pooled and high-pH RP-HPLC was used to fractionate each pool into 12 fractions and were acquired individually by LC-MS/MS to maximize the number of peptide and protein identifications.

Using a Dionex UltiMate 3000 RSLCnano system equipped with a Dionex UltiMate 3000 RS autosampler, an Acclaim PepMap RSLC analytical column (75  $\mu$ m  $\times$  50 cm, nanoViper, C18, 2  $\mu$ m, 100  $\text{\AA}$ ; Thermo Scientific), and an Acclaim PepMap 100 trap column (100  $\mu$ m  $\times$  2 cm, nanoViper, C18, 5  $\mu$ m, 100  $\text{\AA}$ ; Thermo Scientific), the tryptic peptides were separated by increasing concentrations of 80% acetonitrile (ACN)/0.1% formic acid at a flow of 250 nL/min for 158 min and analyzed with an Orbitrap Fusion Tribrid mass spectrometer (ThermoFisher Scientific). The instrument was operated in data-dependent acquisition mode to automatically switch between full scan MS1 (in Orbitrap), MS2 (in ion trap), and MS3 (in Orbitrap) acquisition. Each survey full scan (380–1580 m/z) was acquired with a resolution of 120 000, an AGC (automatic gain control) target of 50%, and a maximum injection time of 50 ms. Dynamic exclusion was set to 60 s after one occurrence. Keeping the cycle time fixed at 2.5 s, the most intense multiply charged ions ( $z \geq 2$ ) were selected for MS2/MS3 analysis. MS2 analysis used CID fragmentation (fixed collision energy mode, 30% CID Collision Energy) with a maximum injection time of 150 ms, a “rapid” scan rate, and an AGC target of 40%. Following the acquisition of each MS2 spectrum, an MS3 spectrum has been acquired from multiple MS2 fragment ions using Synchronous Precursor Selection. The MS3 scan was acquired in the Orbitrap after HCD collision with a resolution of 50 000 and a maximum injection time of 250 ms.

The raw data files were analyzed with Proteome Discoverer (Thermo Scientific) to obtain quantitative MS3 reporter ion intensities.

## 2.5.2 | Proteomic pre-processing and analysis in the Gene SMART cohort

A total of 3368 proteins were identified in this study. The data pre-processing and downstream statistical analysis were performed using the R statistical software (<http://www.r-project.org>). Before normalization, proteomic data were filtered for high-confidence protein observations, and contaminants and proteins that were only identified by a single peptide or proteins not identified/quantified consistently across the experiment were removed. Two samples (timepoint PRE) were removed during the pre-processing analysis due to low protein count. Remaining missing values were imputed using MNAR method, assuming the missing values were due to low expression for such proteins, and then normalized using the VSN method. Both imputations and VSN were conducted in the *DEP* package.<sup>56</sup> Batch effects were normalized with the internal referencing scaling (IRS) method<sup>57</sup> by the use of reference channels. After normalization steps, 2365 proteins were included in subsequent analyses.

Pre-processed proteomic data (described above) was used in this analysis. Protein expression (logFC) was regressed against baseline fitness ( $VO_{2max}$ ), age, and timepoint (PRE, 4WP, 8WP, and 12WP) as covariates:

$$\text{Proteome} = \text{Timepoint} + \text{Fitness} + \text{Age} + \text{random (ID)} \\ + \text{random (ID} * \text{Timepoint)}$$

Individual changes were accessed by the random effect of the model. Results with an FDR < 0.05 were deemed significant. Group changes were further analyzed and classified based on similar trajectories over time using K-means method with the *factoextra* package.<sup>58</sup> Overrepresentation analysis (ORA) of differentially expressed proteins was performed using the *clusterProfiler* package,<sup>59</sup> using the reactome database as gene sets, and all tested proteins as the background.

## 2.6 | PCR and proteomic validations

Based on our proteomic results, we have then selected the top proteins presenting the largest effect sizes and that based on the literature could be of interest to further explore in the exercise context to perform transcriptome analysis to investigate if protein expression aligned with transcript expression at the same timepoints.

RT-PCR for *USP2*, *AMPD3*, and *SOD3* was performed on muscle biopsy. RNA was extracted using the AllPrep DNA/RNA FFPE Kit (#80234 Qiagen). Ten nanograms of RNA was then diluted into 50  $\mu$ L and reverse transcription was conducted using the iScript™ Reverse Transcription

Gene (5' → 3')	Reverse sequence	Forward sequence
<i>SOD3</i>	CTCCGCCGAGTCAGAGTTG	ATGCTGGCGCTACTGTGTTC
<i>AMPD3</i>	TGGGAGAGTAGACCTTGTGCT	AGTGAGCTGCGTGACCTGTA
<i>USP2</i>	CTCCACATCTGTCGGCCTTTC	GGGCTCCATAACGAGGTGAAC

TABLE 1 Primer information for each gene used in the validation step.

Supermix for RT-qPCR (Bio-Rad) with a thermomixer. Primers for *USP2*, *AMPD3*, and *SOD3* used for this experiment are described below (Table 1). RT-PCR was conducted using the QuantumStudio-7 (Bio-Rad). mRNA expression levels were quantified by real-time PCT using SYBR green fluorescence. Cycle threshold (Ct) values were normalized to a housekeeping gene, *Cyclophilin (PPIA, peptidyl-prolyl cis-trans isomerase A)* Accession no. NM021130.4.<sup>60</sup> Samples were analyzed in triplicates and data were manually curated. In cases where samples yielded a standard deviation >0.4, the divergent sample was removed. Manually curated data were used in the analysis. Gene expression (logFC) was regressed against time and age as covariates:

$$\text{Gene Expression} = \text{Timepoint} + \text{Age} + \text{random (ID)} \\ + \text{random (ID} * \text{Timepoint)}$$

To validate the findings obtained through mass spectrometry-based proteomics, a subset of the samples collected at baseline were subjected to a repeat analysis using the same mass spectrometry protocol. This subset was carefully selected as representative of the larger sample cohort. Subsequently, a correlation analysis was performed to assess the consistency and reproducibility of the proteomic profiles between the two independent measurements. Encouragingly, a high correlation (>.8) was observed between the results from the initial mass spectrometry analysis and the repeated measurements, suggesting robustness and reliability of the identified protein markers (e.g., see Figure S1). This step strengthens the confidence in the identified proteins and their differential expression, thus reinforcing the credibility of our proteomic findings.

## 2.7 | OMIC integration Analysis

Integration between DNA methylation and proteomic data was performed using a holistic approach where the 10 000 closest to significant CpGs as recommended by the package, and all identified proteins were included in the analysis. Integration was conducted with the mixOmics package, using a multilevel approach to account for repeated measures.<sup>61,62</sup> Here, we utilized a multilevel approach of the sparse partial least squares analysis (sPLS). The sPLS algorithm implemented in the mixOmics

package selects variables to include in the analysis based on their relevance to the response variables while promoting sparsity. The selection process aims to identify a subset of variables most strongly associated with the response while minimizing the number of included variables to improve interpretability and reduce overfitting. The sPLS algorithm incorporates variable selection through the use of regularization techniques, such as L1-penalty (Lasso) or Elastic Net, which encourage some of the regression coefficients to be exactly zero. By enforcing sparsity, the algorithm automatically identifies and excludes variables that have negligible or no association with the response. The specific implementation of sPLS in the mixOmics package follows the principles previously outlined.<sup>61,63</sup>

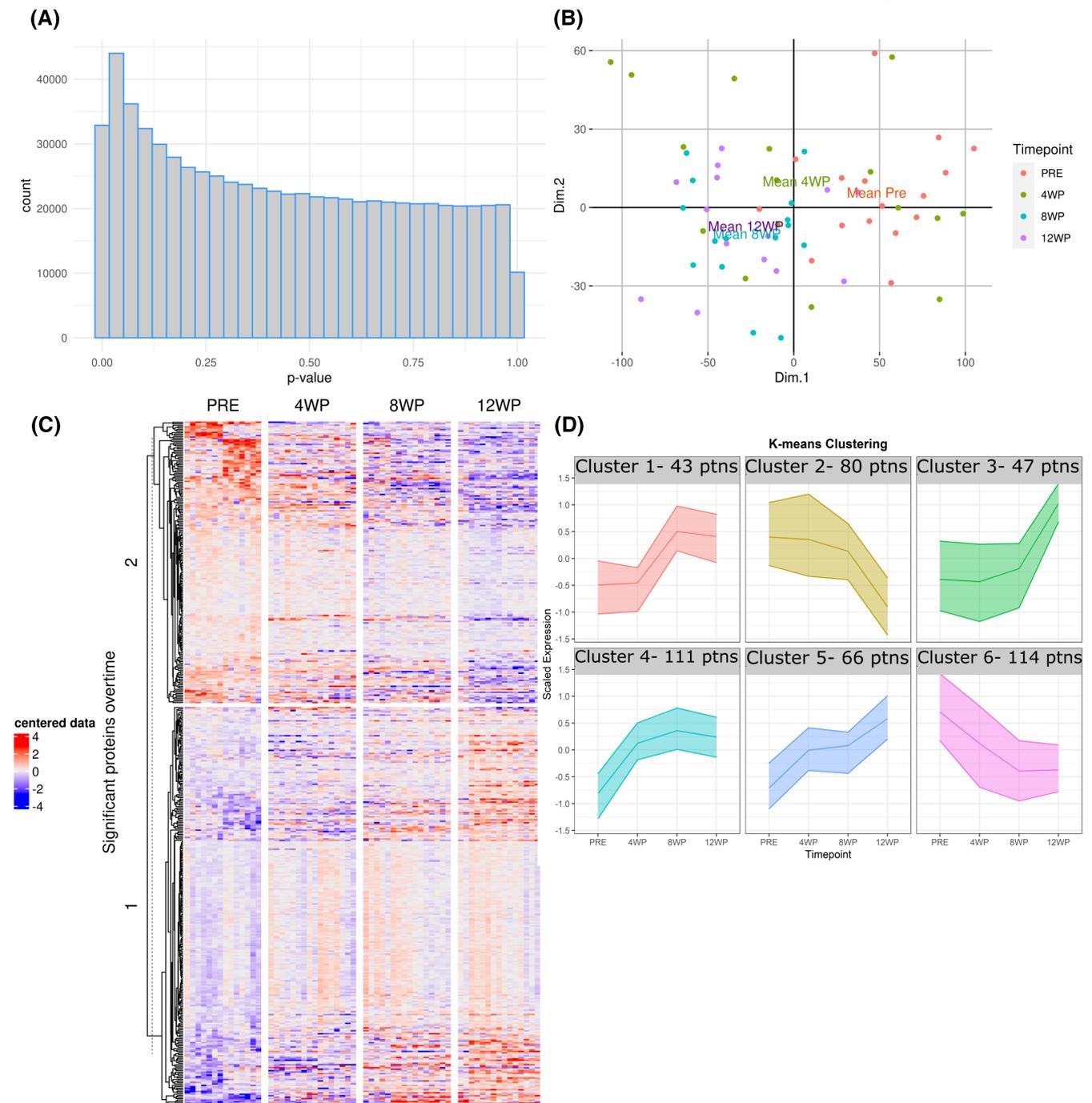
## 3 | RESULTS

### 3.1 | The muscle methylome and proteome of the group are altered over the course of a 12-week HIIT intervention

First, we investigated exercise-induced changes in the methylome and proteome as a dose-response (0 (baseline), 4, 8, and 12 weeks of HIIT) (Figure 1). We identified only one CpG site significantly altered by HIIT (cg23669611, FDR=0.043). However, a closer look at all statistical tests performed revealed an inflation of *p*-values toward zero (Figure 2A), suggesting that widespread changes in the methylome did occur, but we were underpowered to detect them. This was confirmed by our residual analysis where a significant shift (*p*-value < .05) in the methylome was observed over time (Figure 2B).

Changes in the proteome were much clearer. Of the 2635 proteins included in the analysis, 461 changed over the course of the HIIT intervention (Figure 2C and Table S2). Among those, 26 are involved in NADH dehydrogenase complex (Complex I), 10 with ATP synthase, and 18 were mitochondrial ribosomal proteins. Unlike mitochondrial ribosomal proteins, the expression of nuclear ribosomal proteins decreased over time. Similarly, proteins involved in eukaryotic translation decreased over time.

Next, we assessed whether the 461 significantly changing proteins exhibit specific trends over time (i.e., proteins that changed after 4 weeks, or after 8 weeks, etc.). We performed *K*-means clustering analyses with the 461 proteins



**FIGURE 2** DNA methylation and proteomics changes over the course of the exercise intervention. (A) Histogram of inflation of  $p$ -values based on DNA methylation results from mixed model, and  $y$ -axis represents number of CpGs. (B) PCA plot of DNA methylation residuals changes over time. Each point represents a different participant methylation profile. Means are highlighted as text. Significant changes over time were observed in the methylome (Estimate:  $-26.151$ , SD:  $5.108$ ,  $p$ -value =  $.00004$ ). (C) Hierarchical clustering Heatmap of significant proteins presenting changes over time in the group level ( $n = 461$ ,  $p < .05$ ). Note that timepoint pre contains only 14 samples given two being removed during the pre-processing analysis. (D) K-means clustering plot representing trajectory change of significant proteins over time. Each cluster is represented by a group of proteins that present similar changes during the exercise intervention. The number on top of each individual plot represents the number of proteins that belong to each cluster. For example, in cluster 1 out of the 461 significant proteins, 43 showed positive changes between 4 and 8 weeks.

which were divided into six clusters (Figure 2D and Table S3). This clustering revealed coordinated timeline changes with some proteins increasing only between 4WP

and 8WP (Cluster 1) while others increased only between 8WP and 12WP (Cluster 3). Next, we performed a pathway enrichment analysis, and based on all differentially

expressed proteins DEP 14 pathways, all related to mitochondrial function (Figure 3A), were over-represented (DEP) (FDR < 0.05). Interestingly, when the pathway enrichment analysis was performed based on the identified temporal clusters, only cluster 4 (increased from PRE to 8WP) and cluster 5 (increased from PRE to 4WP and 8WP to 12WP) yielded over-represented pathways (Figure 3B,C), all related to mitochondria.

### 3.2 | Individual responses in DNA methylation and protein expression over the course of a 12-week HIIT intervention

The repeated testing protocol allowed to delineate individual progress curves by removing the “noise” around these curves (i.e., intra-individual variability). Similar to results at the group level, we identified little individual variability in DNA methylation response, but widespread individual variability in protein responses. One CpG (cg06587054 located in *RPSAP31*; *PARL*, FDR = 0.001) and 101 proteins (Figure 4A—significant proteins only—and Table S4) showed significant individual variability in response to HIIT. Fifty two of these proteins also changed at the group level over time (Table S2). Seven of the proteins that presented significant individual changes also had large effect sizes >0.5 at the group level (Figure 4B), including: USP2, a protein known to be involved in biological rhythms (i.e., circadian clock) and aging,<sup>64</sup> and differentiation of myoblasts to myotubes<sup>65</sup>; TRIM28, a chromatin regulator that mediates repression of many transcriptional factors<sup>66</sup>; SOD3, an antioxidant enzyme that catalyzes the conversion of superoxide radicals, protecting tissues from oxidative stress<sup>66</sup>; RHOT1, a mitochondrial GTPase involved in mitochondrial trafficking<sup>66</sup>; LYRM7 which works as a chaperone of the inner mitochondrial complex III (the main enzyme complex in the mitochondrial respiratory chain) stabilizing this matrix prior to its translocation and insertion into the late CIII dimeric intermediate within the mitochondrial inner membrane<sup>66</sup>; EPN1 which is involved in endocytosis, and loss of function of this gene is associated with reduced tumor growth and progression of cancer<sup>66</sup>; and AMPD3 which has a critical role in balancing energy charge and nucleotide metabolism.<sup>67,68</sup> Most of these proteins have been previously reported to be expressed in skeletal muscle<sup>69–73</sup> and for some associated with exercise.<sup>69–72,74–77</sup> However, SOD3 had only been studied in rodent muscle, and this is the first time that this gene has been confirmed to be induced by exercise training in vivo *human* skeletal muscle. TRIM28 is novel in the exercise context in humans but has been previously associated with

skeletal muscle size regulation in mice,<sup>73</sup> whilst EPN1 and LYRM7 have not been previously associated with either skeletal muscle or exercise response. No pathways were over-represented among the 101 proteins showing individual training responses (FDR > 0.05).

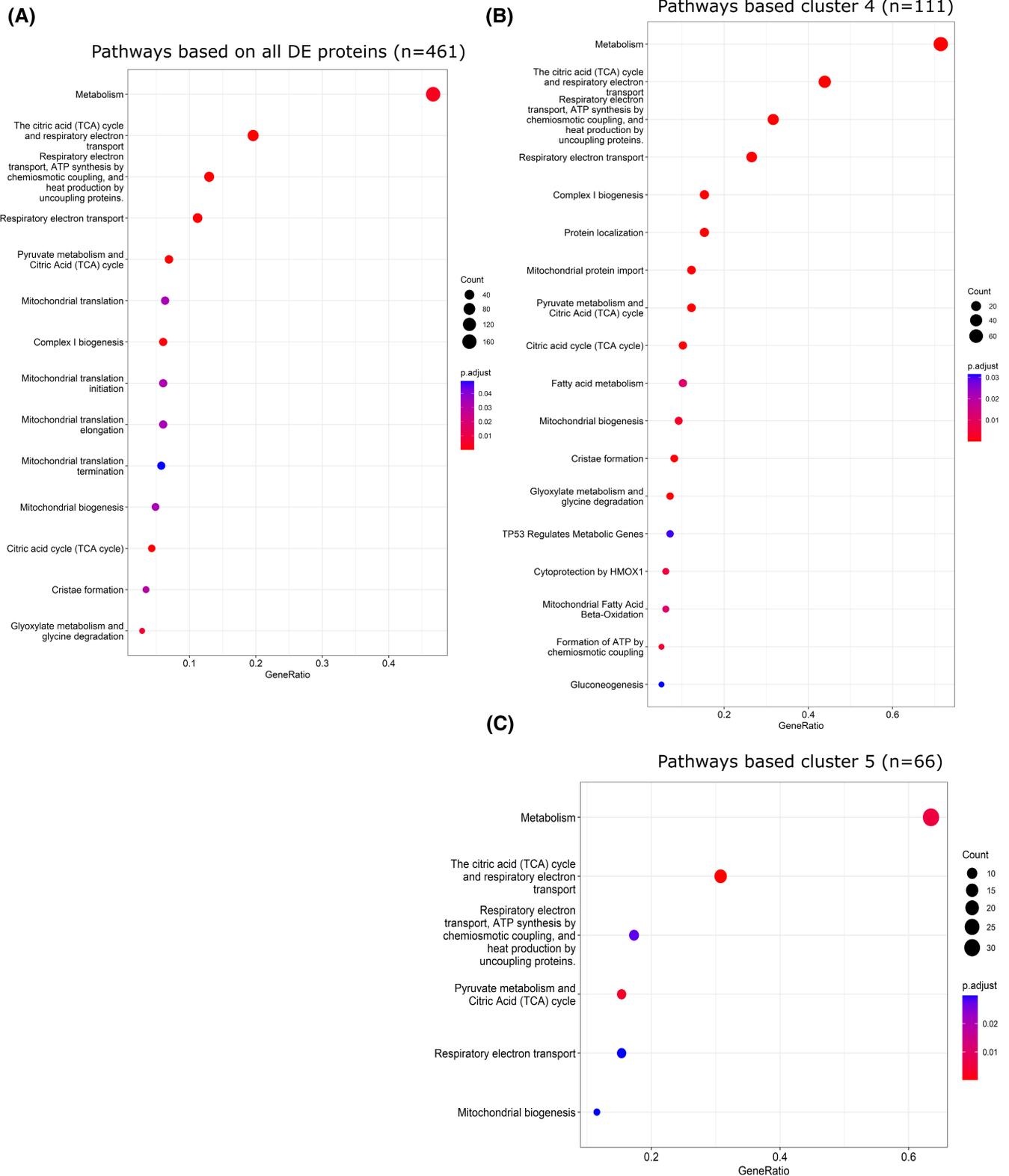
### 3.3 | Transcriptome validation

Protein changes can reflect changes in mRNA expression levels. To further investigate the proteins that showed significant individual changes (AMPD3, SOD3, EPN1, USP2, LYRM7, RHOT1, and TRIM28), we selected the top three with the largest group analysis effect sizes (AMPD3, SOD3, and USP2), and analyzed them at the transcriptional level using real-time PCR analysis in the same participants. We did not detect mRNA expression changes over time ( $p$ -value > .05) at the group level, but we detected significant individual changes ( $p$ -value = .003) for USP2 but transcript changes were not correlated with proteome changes (Pearson  $R$  = -.12,  $p$ -value = .36).

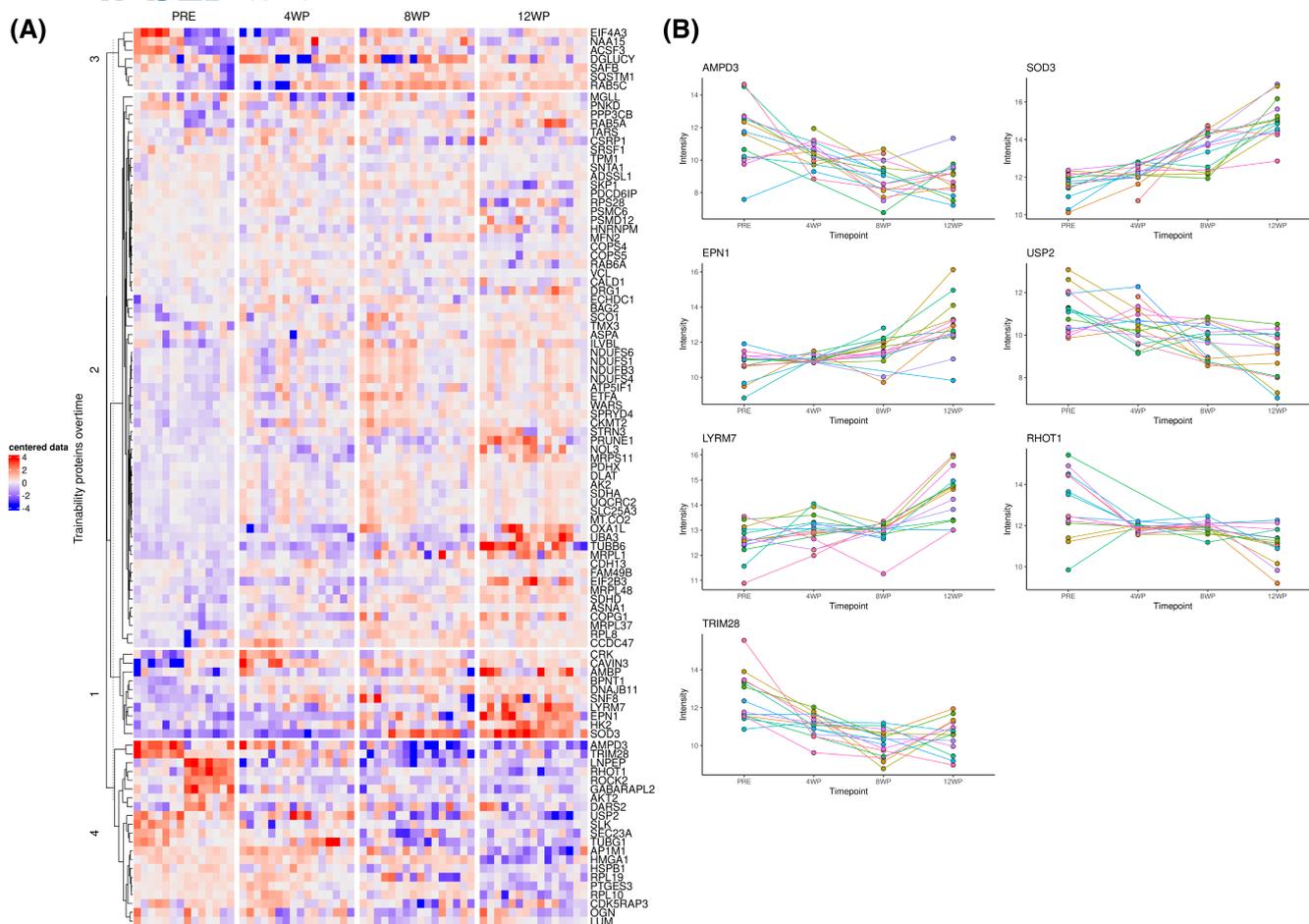
This is unsurprising as transcriptome changes may be occurring at different timepoints than those collected (i.e., after 1 h of exercise bout, after 3 h of exercise bout). Based on a comprehensive meta-analysis by Amar et al. (2021) who investigated exercise-induced transcriptome changes in skeletal muscle,<sup>78</sup> AMPD3 and USP2 expression increased while RHOT1 expression decreased after an acute bout of exercise training. After long-term exercise, *TRIM28*, *RHOT1*, and *USP2* transcripts were down-regulated,<sup>78</sup> in line with our protein expression results. Only USP2 was significant at the individual level based on our analysis. We can only speculate that the other proteins would be true at the individual level if a larger cohort was studied as the transcriptome results derived from the extrameta tool<sup>78</sup> are only based on group changes.

### 3.4 | Integration of DNA methylation and proteome based on individual response

To further investigate the relationship between the methylome and the proteome, we applied a holistic and unbiased multilevel analysis using the mixOmics package.<sup>61,62</sup> The multilevel analysis accounts for variation observed between individuals with subtle differences which otherwise would be masked by individual variation. After accounting for repeated measures, we observed a clear shift from baseline up to 8WP for all individuals when methylome and proteome were integrated. This indicates the clear contribution of both omics to



**FIGURE 3** Over-representation analyses (ORA) of differentially expressed proteins and protein clusters extracted from *K*-mean analysis. (A) Dotplot of significant reactome pathways for all DE proteins ( $n=461$ ). (B) Dotplot of significant reactome pathways for proteins cluster 4 ( $n=111$ ). (C) Dotplot of significant reactome pathways for proteins cluster 5 ( $n=66$ ). In A, B, and C, the size of the dots represents the number of genes assigned to each pathway, and the color of dots represents significance based on adjusted  $p$ -value of  $<.05$ .

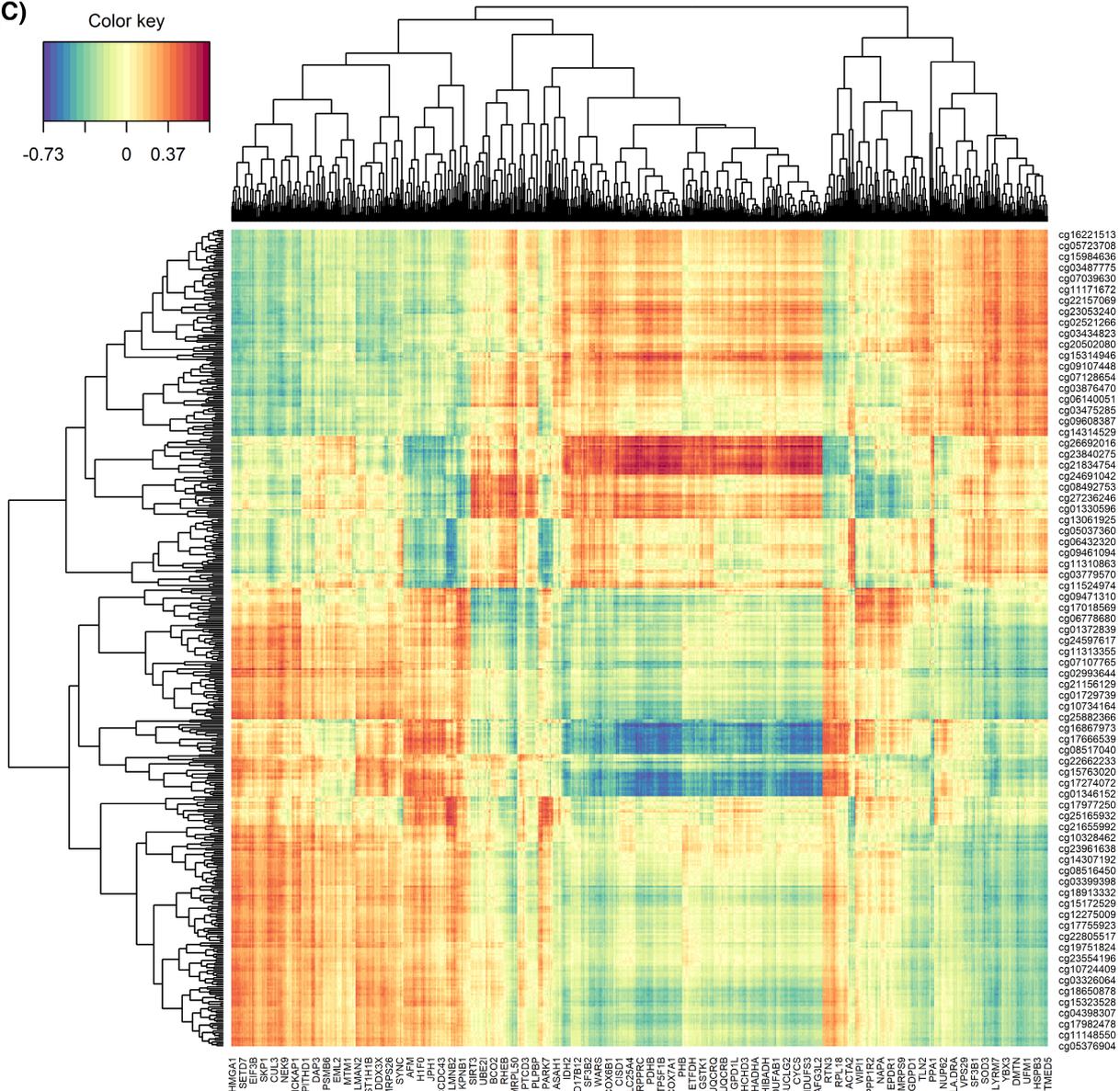
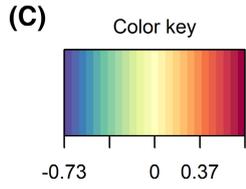
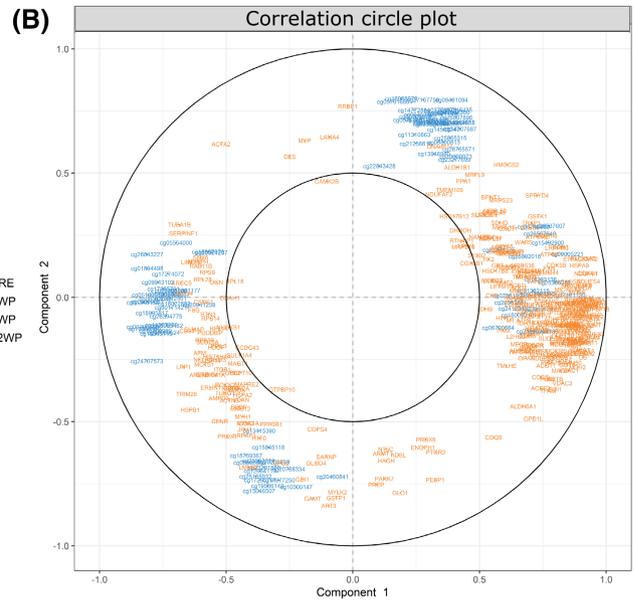
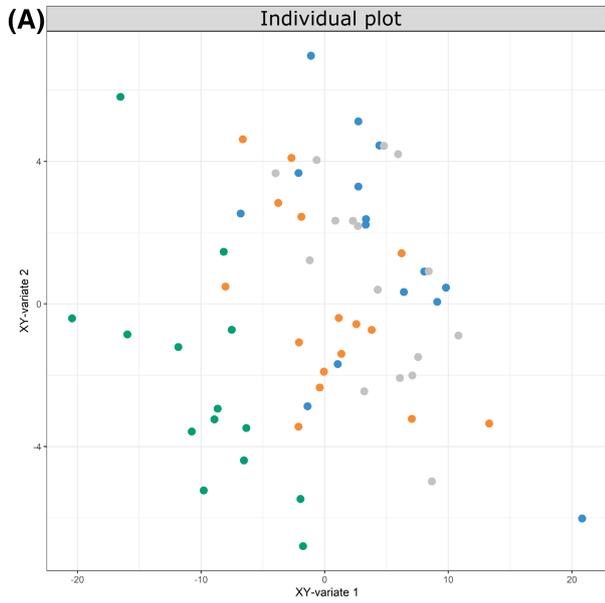


**FIGURE 4** Trainability proteins. (A) Heatmap with the 101 proteins where individual training response was estimated as significant ( $\text{adj.}p < .05$ ). (B) Out of the 101 proteins, seven have shown a change in effect size  $>0.5$  and are represented by a scatterplot with individual points for each participant. Each participant is represented in a different color and color is unique to participant.

individual changes over time (Figure 5A). Sparse partial least squares multivariate (sPLSm) analysis (a powerful algorithm to deal with omics data) revealed that 261 proteins were strongly positively correlated with 45 CpGs, and the same proteins were also strongly negatively correlated with 45 different CpGs (Figure 5B,C) as shown in the middle section of Figure 5C. Many of the 261

proteins are related to mitochondrial functioning (i.e., UQCRC1, IDH3B, ATP5PB, NDUFA4, etc.), mitochondrial calcium ion transport (i.e., PHB, AFG3L2, VDAC1, PHB2, VDAC2, VDAC3, LATM1), mitochondrial fatty acid beta-oxidation (i.e., ACADS, MMUT, ACAA2, ECHS1, HADHA, HADHB, etc.), mitochondrial protein import (i.e., IDH3G, TIMM44, LDHD, FH, CHCHD3,

**FIGURE 5** Multilevel integration of methylome and proteome. (A) Individual plot based on methylome and proteome integration data after controlling for repeated measures. Samples are projected into the space spanned by the averaged components of both methylome (x) and proteome (y). Each dot represents an individual and different colors represent different timepoints. This plot shows that samples within the same timepoint tend to cluster and shift slightly over time as indicated by the colors. (B) Correlation circle plot derived from the sparse partial least squares multivariate (sPLSm) analysis performed on the methylome (blue) and proteome (orange) data after controlling for repeated measures. This plot represents the relationship between methylome and proteome markers, where points close together represent markers that are similar to each other. Markers located on the right side of component 1 represent positive correlations between variables while those on the left represent a negative correlation between variables. This plot highlights the contributing variables that together explain the covariance between the two datasets. The middle circle represents a cutoff = 0.5 which highlights only the variables with stronger contributions to each component found in the outer circle. (C) Clustered Image Map from the sPLSm after controlling for repeated measures. The plot displays the similarity values between the methylome (rows) and proteins (columns) variables selected across two dimensions and clustered with a complete Euclidean distance method. Cell colors represent the correlation between methylome and proteome (i.e., blue negative correlation and red positive correlation).



CS, etc.), respiratory electron transport (NDUFV1/2/3, NDUFA4/6/7/8/9/10/12/13, SDHA/B, COX5A/B, etc.), and metabolism pathways (i.e., GLUT1, SCO1, SDHD, UQCRC, SLC15A1, ATP5PO, ACSL1, etc.).

## 4 | DISCUSSION

We found that intensified exercise training-induced changes in the methylome and, to a greater extent, the proteome in human skeletal muscle. Over 400 proteins significantly changed after exercise, and *K*-means analysis revealed clear clustering highlighting timely changes exhibited by each group of proteins. Utilizing repeated testing approach, we uncovered over 100 proteins associated with individual responses to exercise, and seven proteins had a large effect size of  $>0.5$ . Only one DNA methylation loci (DMP) was significantly associated with individual responses to exercise (cg06587054 which is annotated to the *PARL*—presenilin-associated rhomboid-like gene located on chr3:183884873-183884875).

We found no association between DNA methylation levels and their annotated protein levels. At the group level analyses, only one DMP decreased in methylation after 12 weeks of exercise. This DMP, cg23669611, is located within the Reticulon 4 receptor-like 1 gene (*RTN4RL1*) on chromosome 17 (position: 1998389-1998391). According to the human protein atlas, the protein associated with *RTN4RL1* is known to protect motor neurons against apoptosis.<sup>66</sup> However, the role of *RTN4RL1* in skeletal muscle and its response to exercise remain unknown. Surprisingly, in our proteomic analyses, we did not find an association between *RTN4RL1* and exercise responses. Instead, we observed a significant decrease in expression of the associated protein,<sup>79</sup> *RTN4* (Reticulon 4), after exercise (effect size:  $-0.07$ , adjusted *p*-value: .05). *RTN4* and *RTN4RL1* interact together and *RTN4RL1* is involved in mediating the inhibitory effects of *RNT4*,<sup>124</sup> which may explain the observed decrease in *RTN4* expression after exercise. There could be several explanations for the lack of association between DNA methylation and protein levels. One explanation is that the regulation of gene and protein levels involve complex interactions and mechanisms beyond the direct influence of DNA methylation. Other factors, such as transcriptional regulation, post-transcriptional modifications, or protein degradation processes, may play a role in modulating protein levels independent of DNA methylation changes. Additionally, it is important to consider that our understanding of the relationship between DNA methylation and protein expression is still evolving. The intricate interplay between epigenetic modifications, gene regulation, and protein synthesis is a complex area of research, and further

studies are needed to unravel the underlying mechanisms involved.

At the individual level analyses, we also found decreased methylation levels in one DMP annotated to the *PARL* gene (cg06587054). *PARL* has a pivotal role in the quality control and maintenance of the mitochondria steady state.<sup>80,81</sup> Interestingly, *PARL* knockout mice have been shown to suffer from progressive multisystem disease characterized by progressive muscle and immune organ atrophy.<sup>81,82</sup> We report a consistent individual decrease in methylation; thus, it is possible that decreased DNA methylation in this DMP following exercise provides a protective effect against atrophy. This DMP (cg06587054) is located in a CpG island in an active transcription chromatin state; thus, we hypothesize that exercise may induce decrease in methylation of this gene, which, in turn, will lead to increase in gene and protein expression. Further targeted investigation in larger cohorts in all omics is warranted to confirm causality and elucidate exercise effects in skeletal muscle *PARL* expression.

We found many proteins significantly associated with exercise at both the group and individual level analyses. Group changes highlight modifications in mitochondrial-related pathways with most proteins increasing in expression in an exercise dose-dependent manner, a concept that has been well described by the literature.<sup>43,83–88</sup> Interestingly, we found that mitochondrial ribosomal proteins increased with exercise while nuclear ribosomal proteins decreased with exercise. Both mitochondrial and nuclear genome encode components of the mitochondrion's ribosome.<sup>89</sup> Ribosomal proteins are encoded in the nuclear genome and then imported to the mitochondria where they assemble with the two rRNAs to form mitochondrial ribosomes, which are responsible for translation of 13 mitochondrial mRNAs.<sup>89</sup> Previous study showed that ribosomes are likely to increase in response to resistance training and induce hypertrophy.<sup>90</sup> The exercise applied in the current study was endurance (HIIT) and is not likely to induce significant hypertrophy<sup>91–93</sup>; hence, this might be one of the explanations for the observed results. Interestingly, our results show a decrease in *TRIM28* after training, and given that *TRIM28* overexpression has been recently found to regulate muscle size,<sup>73</sup> the decrease in expression observed corroborates with the decrease of ribosomal proteins, given decrease of both is associated with non-hypertrophy after training. However, in-depth analysis including skeletal muscle RNA content, ribosomal biogenesis, and targeted comparison between mitochondrial ribosomes and nuclear ribosomes in response to different types of exercise is warranted to confirm these findings.

In the present study, we investigated individual differences in responses to exercise and their associations with changes in the proteome and methylome. We identified

seven proteins, with large effect sizes, significantly associated with individual responses to exercise. Among these proteins, two novel exercise-related proteins, LYRM7 and EPN1, stood out with the largest effect size. LYRM7, also known as Complex III assembly factor, has been previously linked to mitochondrial complex III defects<sup>94</sup> and associated with lactic acidosis, muscle weakness, and exercise intolerance.<sup>95</sup> Interestingly, we showed a time-dependent increase in LYRM7 protein expression. This suggests that higher expression of LYRM7 gene and protein may have a positive influence on skeletal muscle function and could be a candidate for further functional and molecular analyses to explore its potentially protective role in skeletal muscle. EPN1, a major component of clathrin-mediated endocytosis,<sup>96</sup> plays a crucial role in the uptake of extracellular molecules and their transport into the intracellular region through endosomes.<sup>97</sup> The exact role of EPN1 protein in exercise adaptation and skeletal muscle functioning remains to be investigated.

Additionally, three out of the seven identified exercise-related proteins (SOD3, RHOT1, and LYRM7) were previously associated with exercise in other tissues, indicating potential systemic or cross-tissue effects. SOD3, a redox enzyme involved in reducing the potential toxicity of reactive oxygen species,<sup>98</sup> has been reported to increase expression after exercise in plasma exosomes in both mice and humans.<sup>74</sup> This suggests a potential systemic cardioprotective effect of SOD3 released from skeletal muscle into circulation.<sup>74</sup> RHOT1, a mitochondrial GTPase involved in mitochondrial trafficking,<sup>99</sup> has decreased expression after exercise in rat liver,<sup>75</sup> and similarly, we observed decreased RHOT1 expression in a different tissue, skeletal muscle, following exercise. Furthermore, a recent study on the effects of exercise training in the human proteome and acetylome identified novel exercise-training regulated proteins, including glutamyl-tRNA synthase (QARS) and rab GDP dissociation inhibitor alpha (GDI1),<sup>33</sup> associated with insulin-stimulated glucose uptake. Interestingly, our analysis substantiated these findings, as we observed significant changes in the same direction for QARS and GDI1 after exercise (QARS – ES: 0.04, adj.*p*-value: .05, GDI1 – ES: –0.06, adj.*p*-value: .013). This cross-study replication strengthens the potential importance of these proteins in insulin regulation following exercise and calls for future functional validation studies to investigate the underlying mechanisms.

We employed an unbiased integration approach to explore subtle differences in exercise responses across multiple biological layers, including the epigenome and proteome. Through this integration analysis, we identified a set of proteins that exhibited both positive and negative associations with specific CpG sites. These findings

hold significant biological meaning and warrant further investigation in future replication and mechanistic studies. Remarkably, among the 261 proteins highlighted in our results, we observed a positive correlation between certain proteins and CpG sites located within genes involved in essential biological processes. For instance, we found proteins positively correlated with CpGs located in genes related to fatty acid metabolism, such as COMMD9 (involved in LDL regulation),<sup>100</sup> ARHGAP42 (associated with hypertension),<sup>101</sup> APOBEC1 (linked to weight loss and muscle development),<sup>102</sup> PTPRT, PTPRN2 (implicated in obesity),<sup>103–106</sup> and DACT1, DIO2, RPTOR, and PLEKHM3 (involved in muscle myogenesis and hypertrophy).<sup>107–111</sup> Conversely, these proteins were negatively correlated with CpG sites located in genes such as KCNMA1, NOTCH1, CAMKK2 (related to skeletal muscle regeneration, proliferation, and differentiation),<sup>112–117</sup> DHRS3, DGKG, LPIN1, WNT5A (associated with obesity, metabolic syndrome, and lipid regulation),<sup>118–120</sup> LRRC2 (a mediator of mitochondrial and cardiac function),<sup>121</sup> and PDE4A (linked to diabetes).<sup>122,123</sup> These associations provide valuable insights into the intricate relationships between specific genes, proteins, and exercise response. Based on the outcomes of this integration analysis, we strongly recommend prioritizing future investigations into the relationship between these genes/proteins. Such studies hold the potential to uncover critical mechanisms associated with exercise response across multiple biological layers, shedding light on the intricate interplay between exercise, gene regulation, and protein expression.

## 5 | CONCLUSIONS

Collectively, we found a significant influence of exercise in multiple omic layers, with particular strong effect of high-intensity exercise on the proteome, in a dose-dependent manner. Novel proteins, that have not been previously associated with either skeletal muscle or exercise response, consistently changed at the group level and across individuals (individual response), highlighting the robust effect these proteins have on exercise responses. Furthermore, we were able to replicate the association between recently identified novel proteins, and exercise responses. We note that a limitation of this study is the relatively small sample size masking some of the effects the methylome has on exercise. The effects of exercise are systemic. In fact, some proteins observed in our study may affect other tissues. Hence, systemic investigations including multiple tissue types following exercise will help to shed light on the benefits that exercise exerts in humans. Future mechanistic and validation studies are required to validate our novel findings.

## AUTHOR CONTRIBUTIONS

Macsue Jacques, Nir Eynon, and Sarah Voisin conceived and designed the research, Macsue Jacques, Shanie Landen, Javier Alvarez Romero, and Danielle Hiam performed the research and acquired the data. Ralf B. Schittenhelm and Iresha Hanchapola, analyzed the proteomic data. Macsue Jacques, Anup D. Shah, and Sarah Voisin, performed statistical analyses. All authors were involved in drafting and revising the manuscript.

## ACKNOWLEDGMENTS

We would like to express gratitude to all of our participants and their efforts during the intervention making this study possible. This work was supported by Nir Eynon's NHMRC Career Development Fellowship (APP1140644), and NHMRC Investigator Grant (APP1194159), as well as Sarah Voisin's NHMRC Early Career Fellowship (APP1157732). The Gene SMART study is also supported by an Australian Research Council (ARC) Discovery Project Grant (DP190103081). This study used BPA-enabled (Bioplatforms Australia)/NCRIS-enabled (National Collaborative Research Infrastructure Strategy) infrastructure located at the Monash Proteomics and Metabolomics Facility. Open access publishing facilitated by Monash University, as part of the Wiley - Monash University agreement via the Council of Australian University Librarians.

## DISCLOSURES

The authors declare that they have no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The methylation data that support the findings of this study are openly available at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE171140>, reference number GSE171140. The physiological data are available upon request due to ethical reasons.

## ORCID

Macsue Jacques  <https://orcid.org/0000-0002-4337-7022>

Shanie Landen  <https://orcid.org/0000-0002-7658-030X>

Danielle Hiam  <https://orcid.org/0000-0003-0135-329X>

Ralf B. Schittenhelm  <https://orcid.org/0000-0001-8738-1878>

Anup D. Shah  <https://orcid.org/0000-0002-3793-1998>

Sarah Voisin  <https://orcid.org/0000-0002-4074-7083>

Nir Eynon  <https://orcid.org/0000-0003-4046-8276>

## REFERENCES

1. Bassuk SS, Manson JE. Epidemiological evidence for the role of physical activity in reducing risk of type 2 diabetes and cardiovascular disease. *J Appl Physiol (1985)*. 2005;99(3):1193-1204.
2. Blair SN, Kohl HW 3rd, Paffenbarger RS Jr, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA*. 1989;262(17):2395-2401.
3. LaMonte MJ, Barlow CE, Jurca R, Kampert JB, Church TS, Blair SN. Cardiorespiratory fitness is inversely associated with the incidence of metabolic syndrome: a prospective study of men and women. *Circulation*. 2005;112(4):505-512.
4. Vellers HL, Kleeberger SR, Lightfoot JT. Inter-individual variation in adaptations to endurance and resistance exercise training: genetic approaches towards understanding a complex phenotype. *Mamm Genome*. 2018;29(1-2):48-62.
5. Ash GI, Kostek MA, Lee H, et al. Glucocorticoid receptor (NR3C1) variants associate with the muscle strength and size response to resistance training. *PLoS One*. 2016;11(1):e0148112.
6. Contrepois K, Wu S, Moneghetti KJ, et al. Molecular choreography of acute exercise. *Cell*. 2020;181(5):1112-1130.e16.
7. Denham J, Marques FZ, O'Brien BJ, Charchar FJ. Exercise: putting action into our epigenome. *Sports Med*. 2014;44(2):189-209.
8. Deshmukh AS. Proteomics of skeletal muscle: focus on insulin resistance and exercise biology. *Proteomes*. 2016;4(1):6.
9. Hoffman NJ. Omics and exercise: global approaches for mapping exercise biological networks. *Cold Spring Harb Perspect Med*. 2017;7(10):a029884.
10. Alvarez-Romero J, Voisin S, Eynon N, Hiam D. Mapping robust genetic variants associated with exercise responses. *Int J Sports Med*. 2021;42(1):3-18.
11. Pillon NJ, Gabriel BM, Dollet L, et al. Transcriptomic profiling of skeletal muscle adaptations to exercise and inactivity. *Nat Commun*. 2020;11(1):470.
12. Harvey NR, Voisin S, Dunn PJ, et al. Genetic variants associated with exercise performance in both moderately trained and highly trained individuals. *Mol Genet Genomics*. 2020;295(2):515-523.
13. Sarzynski MA, Bouchard C. World-class athletic performance and genetic endowment. *Nat Metab*. 2020;2(9):796-798.
14. Williams CJ, Li Z, Harvey N, et al. Genome wide association study of response to interval and continuous exercise training: the Predict-HIIT study. *J Biomed Sci*. 2021;28(1):37.
15. Ehlert T, Simon P, Moser DA. Epigenetics in sports. *Sports Med*. 2013;43(2):93-110.
16. Howlett KF, McGee SL. Epigenetic regulation of skeletal muscle metabolism. *Clin Sci (Lond)*. 2016;130(13):1051-1063.
17. Jacques M, Hiam D, Craig J, Barres R, Eynon N, Voisin S. Epigenetic changes in healthy human skeletal muscle following exercise—a systematic review. *Epigenetics*. 2019;14(7):633-648.
18. Voisin S, Eynon N, Yan X, Bishop DJ. Exercise training and DNA methylation in humans. *Acta Physiol (Oxf)*. 2015;213(1):39-59.
19. Hingst JR, Bruhn L, Hansen MB, et al. Exercise-induced molecular mechanisms promoting glycogen supercompensation in human skeletal muscle. *Mol Metab*. 2018;16:24-34.
20. Parker BL, Burchfield JG, Clayton D, et al. Multiplexed temporal quantification of the exercise-regulated plasma peptidome. *Mol Cell Proteomics*. 2017;16(12):2055-2068.
21. Molendijk J, Parker BL. Proteome-wide systems genetics to identify functional regulators of complex traits. *Cell Syst*. 2021;12(1):5-22.
22. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. *Nat Rev Genet*. 2016;17(8):487-500.

23. Weinhold B. Epigenetics: the science of change. *Environ Health Perspect.* 2006;114(3):A160-A167.
24. van Dijk SJ, Molloy PL, Varinli H, Morrison JL, Muhlhausler BS, members of EpiSCOPE. Epigenetics and human obesity. *Int J Obes (Lond).* 2015;39(1):85-97.
25. Seaborne RA, Strauss J, Cocks M, et al. Methylome of human skeletal muscle after acute & chronic resistance exercise training, detraining & retraining. *Sci Data.* 2018;5:180213.
26. Lindholm ME, Marabita F, Gomez-Cabrero D, et al. An integrative analysis reveals coordinated reprogramming of the epigenome and the transcriptome in human skeletal muscle after training. *Epigenetics.* 2014;9(12):1557-1569.
27. Nitert MD, Dayeh T, Volkov P, et al. Impact of an exercise intervention on DNA methylation in skeletal muscle from first-degree relatives of patients with type 2 diabetes. *Diabetes.* 2012;61(12):3322-3332.
28. Ronn T, Volkov P, Davegardh C, et al. A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue. *PLoS Genet.* 2013;9(6):e1003572.
29. Bajpeyi S, Covington JD, Taylor EM, Stewart LK, Galgani JE, Henagan TM. Skeletal muscle PGC1alpha -1 nucleosome position and -260 nt DNA methylation determine exercise response and prevent ectopic lipid accumulation in men. *Endocrinology.* 2017;158(7):2190-2199.
30. Barres R, Yan J, Egan B, et al. Acute exercise remodels promoter methylation in human skeletal muscle. *Cell Metab.* 2012;15(3):405-411.
31. Robinson MM, Dasari S, Konopka AR, et al. Enhanced protein translation underlies improved metabolic and physical adaptations to different exercise training modes in young and old humans. *Cell Metab.* 2017;25(3):581-592.
32. Srisawat K, Shepherd SO, Lisboa PJ, Burniston JG. A systematic review and meta-analysis of proteomics literature on the response of human skeletal muscle to obesity/type 2 diabetes mellitus (T2DM) versus exercise training. *Proteomes.* 2017;5(4):30.
33. Hostrup M, Lemminger AK, Stocks B, et al. High-intensity interval training remodels the proteome and acetylome of human skeletal muscle. *Elife.* 2022;31(11):e69802.
34. Deshmukh AS, Steenberg DE, Hostrup M, et al. Deep muscle-proteomic analysis of freeze-dried human muscle biopsies reveals fiber type-specific adaptations to exercise training. *Nat Commun.* 2021;12(1):304.
35. Selected body measurements of children 6–11 years. *Vital Health Stat.* 1973;11(123):1-48.
36. Roberts MD, Haun CT, Mobley CB, et al. Physiological differences between low versus high skeletal muscle hypertrophic responders to resistance exercise training: current perspectives and future research directions. *Front Physiol.* 2018;9:834.
37. Ross R, Goodpaster BH, Koch LG, et al. Precision exercise medicine: understanding exercise response variability. *Br J Sports Med.* 2019;53(18):1141-1153.
38. Chrzanowski-Smith OJ, Piatrikova E, Betts JA, Williams S, Gonzalez JT. Variability in exercise physiology: can capturing intra-individual variation help better understand true inter-individual responses? *Eur J Sport Sci.* 2020;20(4):452-460.
39. Hecksteden A, Kraushaar J, Scharhag-Rosenberger F, Theisen D, Senn S, Meyer T. Individual response to exercise training—a statistical perspective. *J Appl Physiol (1985).* 2015;118(12):1450-1459.
40. Hecksteden A, Pitsch W, Rosenberger F, Meyer T. Repeated testing for the assessment of individual response to exercise training. *J Appl Physiol (1985).* 2018;124(6):1567-1579.
41. Voisin S, Jacques M, Lucia A, Bishop DJ, Eynon N. Statistical considerations for exercise protocols aimed at measuring trainability. *Exerc Sport Sci Rev.* 2019;47(1):37-45.
42. Bonafiglia JT, Ross R, Gurd BJ. The application of repeated testing and monoexponential regressions to classify individual cardiorespiratory fitness responses to exercise training. *Eur J Appl Physiol.* 2019;119(4):889-900.
43. Jacques M, Landen S, Alvarez Romero J, et al. Individual physiological and mitochondrial responses during 12 weeks of intensified exercise. *Physiol Rep.* 2021;9(15):e14962.
44. Jacques M, Landen S, Romero JA, et al. Implementation of multiple statistical methods to estimate variability and individual response to training. *Eur J Sport Sci.* 2022;23(4):588-598.
45. Yan X, Eynon N, Papadimitriou ID, et al. The gene SMART study: method, study design, and preliminary findings. *BMC Genomics.* 2017;18(Suppl 8):821.
46. Morris TJ, Butcher LM, Feber A, et al. ChAMP: 450k Chip analysis methylation pipeline. *Bioinformatics.* 2014;30(3):428-430.
47. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res.* 2017;45(4):e22.
48. Du P, Zhang X, Huang CC, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics.* 2010;11:587.
49. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Softw.* 2017;82(13):1-26.
50. Peters TJ, Buckley MJ, Statham AL, et al. De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin.* 2015;8:6.
51. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics.* 2016;32(2):286-288.
52. Ginestet C. ggplot2: elegant graphics for data analysis. *J R Stat Soc Ser A.* 2011;1(174):245-246.
53. Kassambara A. ggpubr: 'ggplot2' Based Publication Ready Plots. 2018, R package version 0.1.7.
54. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics.* 2016;32(18):2847-2849.
55. Lê S, Josse J, Husson F. FactoMineR: An R package for multivariate analysis. *J Stat Softw.* 2008;25(1):1-18.
56. Zhang X, Smits AH, van Tilburg GB, Ovaa H, Huber W, Vermeulen M. Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat Protoc.* 2018;13(3):530-550.
57. Plubell DL, Wilmarth PA, Zhao Y, et al. Extended multiplexing of tandem mass tags (TMT) labeling reveals age and high fat diet specific proteome changes in mouse epididymal adipose tissue. *Mol Cell Proteomics.* 2017;16(5):873-890.
58. Kassambara A, Mundt F. factoextra: extract and visualize the results of multivariate data analyses. 2020 [R package factoextra version 1.0.7].
59. Guangchuang Y, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284-287.
60. Kuang J, Yan X, Genders AJ, Granata C, Bishop DJ. An overview of technical considerations when using quantitative real-time

- PCR analysis of gene expression in human exercise research. *PLoS One*. 2018;13(5):e0196438.
61. Rohart F, Gautier B, Singh A, Le Cao KA. mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol*. 2017;13(11):e1005752.
  62. Liquet B, Le Cao KA, Hocini H, Thiebaut R. A novel approach for biomarker selection and the integration of repeated measures experiments from two assays. *BMC Bioinformatics*. 2012;13:325.
  63. Le Cao KA, Boitard S, Besse P. Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics*. 2011;12:253.
  64. Voisin S, Harvey NR, Haupt LM, et al. An epigenetic clock for human skeletal muscle. *J Cachexia Sarcopenia Muscle*. 2020;11(4):887-898.
  65. Hashimoto M, Saito N, Ohta H, et al. Inhibition of ubiquitin-specific protease 2 causes accumulation of reactive oxygen species, mitochondria dysfunction, and intracellular ATP decrement in C2C12 myoblasts. *Physiol Rep*. 2019;7(14):e14193.
  66. Ponten F, Schwenk JM, Asplund A, Edqvist PH. The human protein atlas as a proteomic resource for biomarker discovery. *J Intern Med*. 2011;270(5):428-446.
  67. Norman B, Mahnke-Zizelman DK, Vallis A, Sabina RL. Genetic and other determinants of AMP deaminase activity in healthy adult skeletal muscle. *J Appl Physiol (1985)*. 1998;85(4):1273-1278.
  68. Cheng J, Morisaki H, Toyama K, Ikawa M, Okabe M, Morisaki T. AMPD3-deficient mice exhibit increased erythrocyte ATP levels but anemia not improved due to PK deficiency. *Genes Cells*. 2012;17(11):913-922.
  69. Dunlap KR, Laskin GR, Waddell DS, et al. Aerobic exercise-mediated changes in the expression of glucocorticoid responsive genes in skeletal muscle differ across the day. *Mol Cell Endocrinol*. 2022;550:111652.
  70. Fisher AG, Seaborne RA, Hughes TM, et al. Transcriptomic and epigenetic regulation of disuse atrophy and the return to activity in skeletal muscle. *FASEB J*. 2017;31(12):5268-5282.
  71. Hong S, Zhou W, Fang B, et al. Dissociation of muscle insulin sensitivity from exercise endurance in mice by HDAC3 depletion. *Nat Med*. 2017;23(2):223-234.
  72. Kilic-Erkek O, Caner V, Abban-Mete G, Baris IC, Bor-Kucukatay M. Determination of the pathways of potential muscle damage and regeneration in response to acute and long-term swimming exercise in mice. *Life Sci*. 2021;272:119265.
  73. Steinert ND, Potts GK, Wilson GM, et al. Mapping of the contraction-induced phosphoproteome identifies TRIM28 as a significant regulator of skeletal muscle size and function. *Cell Rep*. 2021;34(9):108796.
  74. Abdelsaid K, Sudhahar V, Harris RA, et al. Exercise improves angiogenic function of circulating exosomes in type 2 diabetes: role of exosomal SOD3. *FASEB J*. 2022;36(3):e22177.
  75. Chen L, Yang Y, Wang Y, et al. Proteomic response of the rat liver in differential swimming modes. *Clin Exp Pharmacol Physiol*. 2018;45(6):581-590.
  76. Kusuyama J, Makarewicz NS, Albertson BG, et al. Maternal exercise-induced SOD3 reverses the deleterious effects of maternal high-fat diet on offspring metabolism through stabilization of H3K4me3 and protection against WDR82 carbonylation. *Diabetes*. 2022;71(6):1170-1181.
  77. Ookawara T, Haga S, Ha S, et al. Effects of endurance training on three superoxide dismutase isoenzymes in human plasma. *Free Radic Res*. 2003;37(7):713-719.
  78. Amar D, Lindholm ME, Norrbom J, Wheeler MT, Rivas MA, Ashley EA. Time trajectories in the transcriptomic response to exercise - a meta-analysis. *Nat Commun*. 2021;12(1):3471.
  79. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49(D1):D605-D612.
  80. Civitarese AE, MacLean PS, Carling S, et al. Regulation of skeletal muscle oxidative capacity and insulin signaling by the mitochondrial rhomboid protease PARL. *Cell Metab*. 2010;11(5):412-426.
  81. Spinazzi M, De Strooper B. PARL: the mitochondrial rhomboid protease. *Semin Cell Dev Biol*. 2016;60:19-28.
  82. Cipolat S, Rudka T, Hartmann D, et al. Mitochondrial rhomboid PARL regulates cytochrome c release during apoptosis via OPA1-dependent cristae remodeling. *Cell*. 2006;126(1):163-175.
  83. Fealy CE, Grevendonk L, Hoeks J, Hesselink MKC. Skeletal muscle mitochondrial network dynamics in metabolic disorders and aging. *Trends Mol Med*. 2021;27:1033-1044.
  84. Huertas JR, Casuso RA, Agustin PH, Cogliati S. Stay fit, stay young: mitochondria in movement: the role of exercise in the new mitochondrial paradigm. *Oxid Med Cell Longev*. 2019;2019:7058350.
  85. Liang X, Liu L, Fu T, et al. Exercise inducible lactate dehydrogenase B regulates mitochondrial function in skeletal muscle. *J Biol Chem*. 2016;291(49):25306-25318.
  86. Lopez-Otin C, Galluzzi L, Freije JMP, Madeo F, Kroemer G. Metabolic control of longevity. *Cell*. 2016;166(4):802-821.
  87. Perry CG, Lally J, Holloway GP, Heigenhauser GJ, Bonen A, Spriet LL. Repeated transient mRNA bursts precede increases in transcriptional and mitochondrial proteins during training in human skeletal muscle. *J Physiol*. 2010;588(Pt 23):4795-4810.
  88. Smith HJ, Sharma A, Mair WB. Metabolic communication and healthy aging: where should we focus our energy? *Dev Cell*. 2020;54(2):196-211.
  89. Sylvester JE, Fischel-Ghodsian N, Mougey EB, O'Brien TW. Mitochondrial ribosomal proteins: candidate genes for mitochondrial disease. *Genet Med*. 2004;6(2):73-80.
  90. Stec MJ, Mayhew DL, Bamman MM. The effects of age and resistance loading on skeletal muscle ribosome biogenesis. *J Appl Physiol (1985)*. 2015;119(8):851-857.
  91. Callahan MJ, Parr EB, Hawley JA, Camera DM. Can high-intensity interval training promote skeletal muscle anabolism? *Sports Med*. 2021;51(3):405-421.
  92. Joannis S, Gillen JB, Bellamy LM, et al. Evidence for the contribution of muscle stem cells to nonhypertrophic skeletal muscle remodeling in humans. *FASEB J*. 2013;27(11):4596-4605.
  93. Joannis S, McKay BR, Nederveen JP, et al. Satellite cell activity, without expansion, after nonhypertrophic stimuli. *Am J Physiol Regul Integr Comp Physiol*. 2015;309(9):R1101-R1111.
  94. Invernizzi F, Tigano M, Dallabona C, et al. A homozygous mutation in LYRM7/MZM1L associated with early onset encephalopathy, lactic acidosis, and severe reduction of mitochondrial complex III activity. *Hum Mutat*. 2013;34(12):1619-1622.
  95. Hempel M, Kremer LS, Tsiakas K, et al. LYRM7—associated complex III deficiency: a clinical, molecular genetic, MR tomographic, and biochemical study. *Mitochondrion*. 2017;37:55-61.

96. Kyung JW, Bae JR, Kim DH, Song WK, Kim SH. Epsin1 modulates synaptic vesicle retrieval capacity at CNS synapses. *Sci Rep*. 2016;6:31997.
97. Pilecka I, Banach-Orlowska M, Miaczynska M. Nuclear functions of endocytic proteins. *Eur J Cell Biol*. 2007;86(9):533-547.
98. Wang Y, Branicky R, Noe A, Hekimi S. Superoxide dismutases: dual roles in controlling ROS damage and regulating ROS signaling. *J Cell Biol*. 2018;217(6):1915-1928.
99. Fransson A, Ruusala A, Aspenstrom P. Atypical Rho GTPases have roles in mitochondrial homeostasis and apoptosis. *J Biol Chem*. 2003;278(8):6495-6502.
100. Fedoseienko A, Wijers M, Wolters JC, et al. The COMMD family regulates plasma LDL levels and attenuates atherosclerosis through stabilizing the CCC complex in endosomal LDLR trafficking. *Circ Res*. 2018;122(12):1648-1660.
101. Carney EF. Hypertension: role of ARHGAP42 in hypertension. *Nat Rev Nephrol*. 2017;13(3):134.
102. Sato Y, Probst HC, Tatsumi R, Ikeuchi Y, Neuberger MS, Rada C. Deficiency in APOBEC2 leads to a shift in muscle fiber type, diminished body mass, and myopathy. *J Biol Chem*. 2010;285(10):7111-7118.
103. Feng X, Scott A, Wang Y, et al. PTPRT regulates high-fat diet-induced obesity and insulin resistance. *PLoS One*. 2014;9(6):e100783.
104. Fernandez-Carrion R, Sorli JV, Coltell O, et al. Sweet taste preference: relationships with other tastes, liking for sugary foods and exploratory genome-wide association analysis in subjects with metabolic syndrome. *Biomedicine*. 2021;10(1):79.
105. Lee S. The association of genetically controlled CpG methylation (cg158269415) of protein tyrosine phosphatase, receptor type N2 (PTPRN2) with childhood obesity. *Sci Rep*. 2019;9(1):4855.
106. Selvanayagam T, Walker S, Gazzellone MJ, et al. Genome-wide copy number variation analysis identifies novel candidate loci associated with pediatric obesity. *Eur J Hum Genet*. 2018;26(11):1588-1596.
107. Carmody C, Ogawa-Wong AN, Martin C, et al. A global loss of Dio2 leads to unexpected changes in function and fiber types of slow skeletal muscle in male mice. *Endocrinology*. 2019;160(5):1205-1222.
108. Ignacio DL, Silvestre DH, Anne-Palmer E, et al. Early developmental disruption of type 2 deiodinase pathway in mouse skeletal muscle does not impair muscle function. *Thyroid*. 2017;27(4):577-586.
109. Contriciani RE, da Veiga FC, Do Amaral MJ, et al. Dact1 is expressed during chicken and mouse skeletal myogenesis and modulated in human muscle diseases. *Comp Biochem Physiol B Biochem Mol Biol*. 2021;256:110645.
110. Zhang X, Yuan S, Li H, et al. The double face of miR-320: cardiomyocytes-derived miR-320 deteriorated while fibroblasts-derived miR-320 protected against heart failure induced by transverse aortic constriction. *Signal Transduct Target Ther*. 2021;6(1):69.
111. Baraldo M, Nogara L, Dumitras GA, et al. Raptor is critical for increasing the mitochondrial proteome and skeletal muscle force during hypertrophy. *FASEB J*. 2021;35(12):e22031.
112. Maquod F, Cetrone M, Mele A, Tricarico D. Molecular structure and function of big calcium-activated potassium channels in skeletal muscle: pharmacological perspectives. *Physiol Genomics*. 2017;49(6):306-317.
113. Bubak MP, Stout K, Tomtschik J, et al. Notch, Numb and Numb-like responses to exercise-induced muscle damage in human skeletal muscle. *Exp Physiol*. 2022;107:800-806.
114. Fujimaki S, Seko D, Kitajima Y, et al. Notch1 and Notch2 coordinately regulate stem cell function in the quiescent and activated states of muscle satellite cells. *Stem Cells*. 2018;36(2):278-285.
115. Ma L, Li C, Lian S, et al. ActivinA activates Notch1-Shh signaling to regulate proliferation in C2C12 skeletal muscle cells. *Mol Cell Endocrinol*. 2021;519:111055.
116. Qu Z, Liu A, Liu C, et al. Theaflavin promotes mitochondrial abundance and glucose absorption in myotubes by activating the CaMKK2-AMPK signal axis via calcium-ion influx. *J Agric Food Chem*. 2021;69(29):8144-8159.
117. Ye C, Zhang D, Zhao L, et al. CaMKK2 suppresses muscle regeneration through the inhibition of myoblast proliferation and differentiation. *Int J Mol Sci*. 2016;17(10):1695.
118. Reue K. The lipin family: mutations and metabolism. *Curr Opin Lipidol*. 2009;20(3):165-170.
119. Deisenroth C, Itahana Y, Tollini L, Jin A, Zhang Y. p53-inducible DHRS3 is an endoplasmic reticulum protein associated with lipid droplet accumulation. *J Biol Chem*. 2011;286(32):28343-28356.
120. Ng MC, Tam CH, So WY, et al. Implication of genetic variants near NEGR1, SEC16B, TMEM18, ETV5/DGKG, GNPDA2, LIN7C/BDNF, MTCH2, BCDIN3D/FAIM2, SH2B1, FTO, MC4R, and KCTD15 with obesity and type 2 diabetes in 7705 Chinese. *J Clin Endocrinol Metab*. 2010;95(5):2418-2425.
121. McDermott-Roe C, Leleu M, Rowe GC, et al. Transcriptome-wide co-expression analysis identifies LRRC2 as a novel mediator of mitochondrial and cardiac function. *PLoS One*. 2017;12(2):e0170458.
122. Hanna R, Nour-Eldine W, Saliba Y, et al. Cardiac phosphodiesterases are differentially increased in diabetic cardiomyopathy. *Life Sci*. 2021;283:119857.
123. Arcaro CA, Assis RP, Oliveira JO, et al. Phosphodiesterase 4 inhibition restrains muscle proteolysis in diabetic rats by activating PKA and EPAC/Akt effectors and inhibiting FoxO factors. *Life Sci*. 2021;278:119563.
124. Oertle T, van der Haar ME, Bandtlow CE, et al. Nogo-A inhibits neurite outgrowth and cell spreading with three discrete regions. *J Neurosci*. 2003;23(13):5393-5406.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Jacques M, Landen S, Romero JA, et al. Methylome and proteome integration in human skeletal muscle uncover group and individual responses to high-intensity interval training. *The FASEB Journal*. 2023;37:e23184. doi:[10.1096/fj.202300840RR](https://doi.org/10.1096/fj.202300840RR)