DNA methylation across the genome in aged human skeletal muscle tissue and muscle stem cells: The role of HOX genes and physical activity

Turner DC[#1,2,3], Gorski PP[#1,3], Maasar MF[4], Seaborne RA[2,3,6], Baumert P[4,5], Brown AD[2], Kitchen MO[3], Erskine RM[4,7], Dos-Remedios I[8], Voisin S[9], Eynon N[9], Sultanov RI[10], Borisov OV[10,11], Larin AK[10], Semenova EA[10], Popov DV[12], Generozov EV[10], Stewart CE[2], Drust B[14], Owens DJ[2,4], Ahmetov II[$4,10,13], Sharples AP*[1,2,3].

[1] Department for Physical Performance, Norwegian School of Sport Sciences (NiH), Oslo, Norway.

[2] Stem Cells, Ageing and Molecular Physiology Unit, Exercise Metabolism and Adaptation Research Group, Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, United Kingdom.

[3] Institute for Science and Technology in Medicine (ISTM), School of Pharmacy & Bioengineering, Keele University, Staffordshire, United Kingdom.

[4] Exercise Metabolism and Adaptation Research Group, Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, United Kingdom.

[5] Exercise Biology Group, Faculty of Sport and Health Sciences, Technical University of Munich, Munich, Germany.

[6] Centre for Genomics and Child Health, Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom.

[7] Institute of Sport, Exercise and Health, University College London, London, United Kingdom

[8] Orthopedics Department, University Hospitals of the North Midlands, Staffordshire, UK.

[9] Institute for Health and Sport (iHeS), Victoria University, Footscray, Victoria, Australia.

[10] Department of Molecular Biology and Genetics, Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia.

[11] Institute for Genomic Statistics and Bioinformatics, University Hospital Bonn, Bonn, Germany

[12] Laboratory of Exercise Physiology, Institute of Biomedical Problems of the Russian Academy of Sciences, Moscow, Russia.

[13] Laboratory of Molecular Genetics, Kazan State Medical University, Kazan, Russia.

[14] School of Sport, Exercise and Rehabilitation Sciences, College of Life and Environmental Sciences, University of Birmingham.

[#]These authors contributed equally to the work.

*Corresponding author for Skeletal muscle tissue, stem cells and DNA / HOX methylation
a.p.sharples@googlemail.com

[$] Corresponding author for physical activity and HOX methylation
I.Ahmetov@ljmu.ac.uk

Abstract

Skeletal muscle tissue and muscle stem cells show global hypermethylation with aging. However, it is unknown whether the methylomes of differentiating muscle stem cells overlap that of skeletal muscle tissue and how age modulates this overlap. Using the most recent high coverage DNA methylation arrays (850K CpG sites), we compared the methylomes of young and aged skeletal muscle to that of young and aged muscle stem cells at several time points of differentiation (0, 72 hrs, 7 days and 10 days). In line with the existing literature (450K arrays), aged muscle tissue was hypermethylated compared with young tissue. Hypermethylated genes in aged tissue were enriched for; pathways in cancer (that combined; focal adhesion, MAPK signaling, PI3K-Akt-mTOR signaling, p53 signaling, Jak-STAT signaling, TGF-beta and notch signaling), rap1 signaling, axon guidance and hippo signaling. Aged muscle stem cells also demonstrated a hypermethylated profile, enriched for; axon guidance, adherens junction and calcium signaling pathways, particularly at later timepoints of differentiation and myotube formation, that corresponded with reduced morphological differentiation versus young cells, confirmed by reductions in myoD and myogenin. While young muscle stem cells showed less variability in DNA methylation during differentiation, aged muscle stem cells demonstrated extensive and significant changes in DNA methylation at 7 days of differentiation at distinct CpG sites than those identified with age alone. The differentially methylated genes were enriched for; regulation of localisation, regulation of cell communication and regulation of signaling. We also identified CpG sites whose hypermethylated state was retained *in-vitro* in muscle stem cells once isolated from tissue *in-vivo*. These CpGs were located in genes: KIF15, DYRK2, FHL2, MRPS33, ABCA17P. In particular, differential chromosomal region analysis identified 3 regions with 6-8 CpGs in the Homeobox (HOX) transcription factor family of genes that were altered with age in the same direction (hypo or hypermethylated) across both muscle tissue and stem cells. Indeed, HOXD10, HOXD9, HOXD8, HOXA3, HOXC9, HOXB1, HOXB3, HOXC-AS2 and HOXC10 were all significantly differentially hypermethylated in aged tissue. In aged stem cells the same HOX genes (and additionally HOXC-AS3) also displayed the most varied methylation at 7 days of differentiation versus young cells. With HOXD8, HOXC9, HOXB1 and HOXC-AS3 hypermethylated and HOXC10 and HOXC-AS2 hypomethylated. We also determined inverse relationships between mRNA expression and DNA methylation in HOXB1, HOXA3 and HOXC-AS3 in aged cells. Finally, we were able to demonstrate that increased physical activity in young adults was associated with reversing the age-related DNA methylation changes observed in HOXB1, HOXA3 and HOXC10. For the first time, we demonstrated that a large number of HOX genes were differentially epigenetically regulated in aged human skeletal muscle and stem cells and that increased physical activity could reverse the age-related epigenetic changes in the HOX genes.

Introduction

Maintaining skeletal muscle mass and function into older age is fundamental for human health-span and quality of life [1]. Five to ten percent of older humans have sarcopenia [2], which is characterized by reductions in muscle mass and strength [3]. This loss of muscle mass and strength leads to frailty, increased incidence of falls, hospitalization and morbidity [3-10]. Annual costs of fragility are estimated to be 39/32 billion (Euros/USD) for European and USA fragility fractures respectively, with the cost of sarcopenia estimated to be £2 billion in the UK [11]. With an advancing population these costs are only likely to increase with time.

A primary hallmark of ageing is the alteration of the epigenetic landscape. Epigenetics encompasses the interaction between lifestyle/environmental factors and modifications to DNA that affect gene function, without changes to the inherited DNA sequence [12, 13]. DNA methylation is the most studied epigenetic modification, which involves the addition of a covalent methyl group to the 5' position of the pyrimidine ring of a cytosine (5mC). Increased methylation (hypermethylation) to cytosine-guanine (C-G) pairings (CpG sites), especially in CpG-rich regions such as gene promoters, typically leads to reduced capacity for the transcriptional apparatus to bind to these regions, suppressing gene expression [13]. Methylated CpG islands in promoters also lead to a tight compaction of adjacent chromatin via the recruitment of chromatin modifying protein/protein-complexes, further silencing gene expression. In contrast, reduced methylation (hypomethylation) provides a more favorable DNA landscape for the transcriptional apparatus to bind to these regions, as well as more 'relaxed' chromatin, enabling gene expression to occur. Generally, albeit not in every scenario, methylation levels therefore are inversely correlated with gene expression, where hypermethylation is associated with reduced gene expression and hypomethylation with increased gene expression.

DNA methylation in aged skeletal muscle occurs at tissue-specific genes [14]. However, aged muscle also has the smallest overlap with other aged tissue types, suggesting skeletal muscle may be unique in comparison with other tissues in its epigenetic aging processes [14]. Indeed, it has recently been demonstrated that the methylation status of approximately 200 CpG sites can accurately predict chronological age in skeletal muscle tissue [15]. But that this muscle 'clock' only shares 16 of these CpG's with the original 353 CpG pan-tissue Horvath clock [15, 16]. Further, using DNA methylation arrays with coverage of ~450,000 CpG sites [17], Alibegovic *et al.* demonstrated that compared with young human skeletal muscle, aged skeletal muscle is hypermethylated at a considerable number of CpG sites across the genome. This suggests that aging may globally suppress gene expression in skeletal muscle, and therefore may underpin impaired muscle maintenance into older age. However, corresponding gene expression changers were not measured in this particular paper and therefore warrant future investigation. Moreover, our group have demonstrated that when mouse skeletal muscle stem cells were exposed to a high dose inflammatory stress in early proliferative life, the cells retained hypermethylation of muscle specific regulatory factor, MyoD, 30 population doublings later [18]. Suggesting that inflamed proliferative aging in muscle stem cells leads to a retained accumulation of DNA methylation. Finally, it is also worth suggesting that increased physical activity [19] or resistance exercise has been reported to predominantly hypomethylate the genome in young skeletal muscle [20, 21] which is in contrast to the hypermethylation observed with aging, suggestive that

exercise may provide a stimulus to reverse age-related changes in DNA methylation. Albeit, this hypothesis warrants further attention.

Given skeletal muscle fibers are post-mitotic, as they contain terminally differentiated/fused nuclei (myonuclei), repair and regeneration of skeletal muscle tissue is mediated by a separate population of resident stem cells, called satellite cells, that can undergo mitosis. Once activated these cells proliferate and migrate to site of injury to differentiate and fuse with the existing fibers to enable repair. Indeed, altered DNA methylation of target genes have been identified during the differentiation of muscle cells into myotubes *in-vitro* [22]. This has included varied methylation of MyoD [23], Myogenin [24] and Six1 [25]. Furthermore, muscle stem cell derived from aged individuals demonstrate similar proliferative capacity and time to senescence as young adult cells [26, 27], but generally demonstrate impaired differentiation and fusion into myotubes [28-44]. There have been a small number of studies to the contrary, that suggested no effect of age on the differentiation capacity of isolated cells [26, 45, 46]. However, the majority of studies demonstrate an impaired differentiation phenotype in aged cells, yet despite this, there has been only one study that assessed DNA methylation across the genome (450K CpG sites) in aged versus young adult muscle stem cells [26]. This study identified that there was increased (hyper) methylation across the genome in these cells as well as aged tissue [26].

To date, no investigation has compared DNA methylation across the genome during the entire time-course of muscle cell differentiation, particularly using array technology which permits greater coverage (850K CpG arrays) in aged muscle stem cells. Furthermore, to the authors knowledge no one has investigated the differences in gene ontology, associated enriched pathways, differentially methylated chromosomal regions or specific temporal profiles of DNA methylation in aged muscle tissue or across differentiation compared with young adults. Therefore, the objectives of the current study were: 1) To investigate the DNA methylome of human aged and young adult skeletal muscle tissue and primary differentiating muscle derived stem cells over an extensive time-course of; 0 h (30 minutes post transfer to differentiation media), 72 h (hours), 7 d (days) and 10 d of differentiation (using 850K CpG arrays) to identify aged associated epigenetic changes. 2) Identify methylation patterns that are retained *in-vitro* by muscle stem cells from the aging tissue *in-vivo* niche. Finally, 3) to identify if any of the most significant changes observed in DNA methylation with age in muscle tissue and stem cells were also influenced by physical activity levels.

Methods

Skeletal muscle biopsies and primary cell isolations

For young adults (total n = 9, male, age 27 ± 4.4 years), skeletal muscle tissue (~150 mg) was obtained from the vastus lateralis via a conchotome biopsy. Consent and ethical approval were granted for the collection of muscle tissue under NREC, UK approval (16/WM/010) or LJMU, UK local ethics committee approvals [H19/SPS/028 & H15/SPS/031)**.** Six (out of 9) of the young adult's tissue (male, 28 ± 5.3 years) baseline (at rest) array data was derived from Seaborne *et al.* (2018). This is because we used this baseline (at rest) tissue to derived cells (detailed below) for DNA methylation analysis experiments in the present study. We also further added 3 new baseline (rested) young adult tissue biopsies (male, 27 ± 2.5) for new DNA array analyses in the present study. For older adults (n = 5, 2

male/3 female, 83 ± 4 years), tissue biopsies were obtained during elective orthopedic surgeries from University Hospitals of the North Midlands, from the vastus lateralis (knee surgery, n = 2) or gluteus medius muscles (hip surgery, n = 3), under consent and ethical approval 18/WM/0187. DNA and RNA were isolated from these young and aged tissue samples. DNA samples from all 9 young adult and 5 aged adults were analysed for DNA methylation arrays (detailed below), and a subset were analysed for gene expression (young adult n = 4, aged adult n = 5). Primary skeletal muscle cells were derived from a subset of young adult and aged tissue samples, and isolated as per our previous work [20, 21, 47-49]. Briefly, approximately 100 mg biopsy tissue was immediately (~10-30 mins) transferred to a sterile class II biological safety cabinet in pre-cooled (4˚C) transfer media (Hams F-10, 2% hi-FBS, 100 U/ml penicillin, 100 µg/ml streptomycin and 2.5 µg/ml amphotericin-B). Any visible connective and adipose tissue were removed using sterile scalpels and muscle tissue was thoroughly washed 2 × in sterile PBS (containing 100 U/ml penicillin, 100 µg/ml streptomycin, 2.5 µg/ml amphotericin-B). PBS was removed and the muscle tissue was minced in the presence of 0.05% Trypsin/0.02% EDTA and all contents (tissue and trypsin) were transferred to a magnetic stirring platform at 37°C for 10 minutes. The addition of 0.05% Trypsin/0.02% EDTA and titration was repeated on any remaining tissue. The supernatant was collected from both procedures and horse serum (HS) was added at 10% of the total supernatant volume to neutralize the trypsin. The supernatant was centrifuged at 340 g for 5 minutes where the supernatant and cell pellet were both plated in separate pre-gelatinised (0.2% gelatin) T25 flasks containing 7.5 ml fresh pre-heated growth media/GM (GM; Ham's F10 nutrient mix supplemented with 10% hi-NCS, 10% hi-FBS, 100 U/ml penicillin, 100 µg/ml streptomycin, 2.5 µg/ml amphotericin B and 5 mM L-glutamine). Once confluent, cells were trypsinised and reseeded into larger T75's to expand the cell population. Human derived muscle cells (HMDCs) were seeded onto gelatin coated 6 well plates, at a density of 9×10^5 cells/ml in 2 ml of GM for until ~90% confluency was attained (~48 h). GM was removed, and cells underwent 3 × PBS washes before switching to low serum differentiation media (DM; Ham's F10 nutrient mix supplemented with 2% hiFBS, 100 U/ml penicillin, 100 µg/ml streptomycin, 2.5 µg/ml amphotericin B and 5 mM L-glutamine). HMDCs were differentiated for a total of 10 days (d) which received a 1 ml top up at 72 h and 7 d timepoints. Cells were lysed for DNA and RNA at 0 h (30 minutes in DM), 72 h, 7 d and 10 d. All experiments were carried out below passage 10 to prevent senescence. We undertook methylation arrays on DNA isolated from: 0 h young (n = 7), 0 h aged (n = 4), 72 h young (n = 7), 72 h aged (n = 4), 7 d young (n = 6), 7d aged (n = 3), 10 d young (n = 2) and 10 d aged (n= 4). Gene expression was analysed using young (n = 4, 7 d) and aged (n = 3, 7 d) cells.

Myogenicity and morphology measurements

Following isolations, attached single cells were fixed prior to staining and fluorescent immunocytochemistry analysis to determine myogenicity (via desmin positivity) of the isolated young and aged muscle derived cultures. Briefly, approximately 2×10^4 cells were seeded onto 3 × wells of a 6-well plate and were incubated for 24 h in GM. Existing media was removed and cells were washed 3 × in PBS before fixation using the graded methanol/acetone method (50:25:25 TBS:methanol:acetone for 15 minutes followed by 50:50 methanol:acetone) after which cells were permeabilised in 0.2% Triton X-100 and blocked in 5% goat serum (Sigma-Aldrich, UK) in TBS for 30 minutes. Cells were washed 3 × in TBS and incubated overnight (4 ˚C) in 300 µl of TBS (with 2% goat serum and 0.2% Triton X-100) containing primary anti-desmin antibody (1:50; ab15200, Abcam, UK). After overnight

incubation, cells were washed 3 × in TBS and incubated at RT for 3 h in 300 µl of secondary antibody solution (TBS, 2% goat serum and 0.2% Triton X-100) containing the secondary antibody, anti-rabbit TRITC (1:75; T6778, Sigma-Aldrich, UK) to counterstain desmin. Finally, cells were washed again 3 × in TBS, prior to counterstaining nuclei using 300 µl of DAPI solution at a concentration of (300 nM; D1306, Thermo Fisher Scientific, UK) for 30 minutes. Immunostained cells were then visualised using a fluorescent microscope (Nikon, Eclipse Ti-S, Japan or Olympus IX83, Japan) and imaged using corresponding software (Nikon, NIS Elements and Olympus FV10-ASW 4.2). Myoblasts, myotubes and nuclei were visualized using TRITC (Desmin, Excitation: 557 nm, Emission: 576 nm) DAPI (Excitation: 358 nm, Emission: 461 nm) filter cubes. All immunostained samples were imported to Fiji/ImageJ (version 2.0.0) software for subsequent calculations. Note, there was no difference ($p = 0.86$) in myogenic cell proportions in aged (35 ± 5%) versus young cells (34 ± 9%). Therefore, the aged and young muscle derived cells were matched for their desmin positivity (% myoblasts) prior to differentiation experiments and downstream DNA methylation and gene expression analysis. To determine myotube differentiation and formation at each time point, cells were imaged using light microscopy (Olympus, CKX31, Japan) at 0, 72 h, 7 and 10 d of differentiation.

DNA isolation and bisulfite conversion

Prior to DNA isolation, tissue samples were homogenized for 45 seconds at 6,000 rpm × 3 (5 minutes on ice in between intervals) in lysis buffer (180 µl buffer ATL with 20 µl proteinase K) provided in the DNeasy spin column kit (Qiagen, UK) using a Roche Magnalyser instrument and homogenization tubes containing ceramic beads (Roche, UK). Cells were lysed in 180 µl PBS containing 20 µl proteinase K**.** The DNA was then bisulfite converted using the EZ DNA Methylation Kit (Zymo Research, CA, United States) as per manufacturer's instructions.

Infinium MethylationEPIC BeadChip Array

All DNA methylation experiments were performed in accordance with Illumina manufacturer instructions for the Infinium Methylation EPIC BeadChip Array. Methods for the amplification, fragmentation, precipitation and resuspension of amplified DNA, hybridisation to EPIC beadchip, extension and staining of the bisulfite converted DNA (BCD) can be found in detail in our open access methods paper [48]. EPIC BeadChips were imaged using the Illumina iScan® System (Illumina, United States).

DNA methylation analysis, CpG enrichment analysis (GO and KEGG pathways), differentially modified region analysis and Self Organising Map (SOM) profiling

Following DNA methylation quantification via MethylationEPIC BeadChip array, raw .IDAT files were processed using Partek Genomics Suite V.7 (Partek Inc. Missouri, USA) and annotated using the MethylationEPIC_v-1-0_B4 manifest file. We first checked the average detection p-values. The highest average detection p-value for the samples was 0.0023 (Suppl. Figure **1a**), which is well below that recommended in the Oshlack workflow, of 0.01 [50]. We also produced density plots of the raw intensities/signals of the probes (Suppl. Figure **1b**). These demonstrated that all methylated and unmethylated signals were over 11.5 (mean was 11.52 and median was 11.8), and the difference between median methylation and median unmethylated signal was 0.56 [50]. Upon import of the data into

Partek Genomics Suite we removed probes that spanned X and Y chromosomes from the analysis due to having both males and females in the study design, and although the average detection p-value for each samples was on average very low (no higher than 0.0023) we also excluded any individual probes with a detection p-value that was above 0.01 as recommended in [50]. Out of a total of 865,860 probes, removal of the X&Y chromosome probes and those with a detection p-value above 0.01 reduced the probe number to 846,233 (removed 19,627 probes). We also filtered out probes located in known single-nucleotide polymorphisms (SNPs) and any known cross-reactive probes using previously defined SNP and cross-reactive probe lists identified in earlier 850K validation studies [51]. This resulted in a final list of 793,200 probes to be analysed. Following this, background normalisation was performed via functional normalisation (with noob background correction), as previously described [52]. Following functional normalisation, we also undertook quality control procedures of principle component analysis (PCA), density plots by lines as well as box and whisker plots of the normalised data for tissue samples (Suppl. Figures **1c, d, e** respectively) and cell samples (Suppl. Figures 1 **f, g, h** respectively). Any outlier samples were detected using PCA and analysing the normal distribution of β-values. Outliers were then removed if they fell outside 2 standard deviations (SDs) (e.g. Suppl. Figure **1c**) using ellipsoids and or demonstrated different distribution patterns to the samples of the same condition. Only one young adult male tissue sample was removed due to being outside 2 standard deviations outside samples from the same condition (Suppl. Figure **1c**; sample with a strikethrough line). Following normalisation and quality control procedures, we undertook differential methylation analysis by converting β-values to M-values (M-value = log2(β / (1 - β)), as M-values show distributions that are more statistically valid for the differential analysis of methylation levels [53]. We undertook a 1-way ANOVA for comparisons of young and aged skeletal muscle tissue. For primary human muscle cells, we undertook a 2-way ANOVA for age (young adult cells and aged cells) and time (0, 72 h, 7 and 10 d) and also explored the ANOVA main effects for both age and time independently. We also performed planned contrasts within the differentiation time-course in both young adult cells alone and aged cells alone e.g. 0 h vs. 72 h, 0 h vs. 7d 0 h vs. 10 d. As well as planned contrasts for young versus aged cells at each time point of differentiation (0 h vs. 0 h, 72 h vs. 72 h, 7 d vs. 7 d, 10 d vs. 10 d in aged vs. young adult cells respectively). From these ANOVA's and contrasts we then generated CpG lists based on a statistical cut-off that used the False Discovery Rate (FDR) of at least ≤ 0.05. Because in some analysis, an extremely large number of significantly differentially methylated CpG sites were identified (e.g. over 100,000) with an FDR of ≤ 0.05, on these occasions we also provide the data for a more stringent FDR of ≤ 0.01 or 0.001 or FDR ≤ 0.05 'a change (difference in M-value) in methylation greater than 2'. This therefore produced a smaller list of the most highly significant sites and also subsequently enabled sensible pathway analysis depicting the most pertinent and significant CpGs/genes, even if this made the sheer number of significant CpG sites less impressive. We specify when a more stringent FDR than 0.05 was used in the methods text. On these significantly differentially methylated CpG lists we then undertook CpG enrichment analysis within gene ontology pathways and KEGG pathways [54-56] using Partek Genomics Suite and Partek Pathway. We also undertook differentially methylated region analysis (e.g. identifies where several CpG's are differentially methylated within a short chromosomal location/region) using the Bioconductor package DMRcate (DOI: 10.18129/B9.bioc.DMRcate). Finally, in order to plot temporal changes in methylation across the time-course of muscle cell differentiation we undertook Self Organising Map (SOM) profiling of the change in mean methylation within each condition using Partek Genomics Suite.

RNA isolation, primer design & gene expression analysis

Skeletal muscle tissue muscle was homogenised in tubes containing ceramic beads (MagNA Lyser Green Beads, Roche, Germany) and 1 ml Tri-Reagent (Invitrogen, Loughborough, UK) for 45 seconds at 6,000 rpm × 3 (and placed on ice for 5 minutes at the end of each 45 second homogenization) using a Roche Magnalyser instrument (Roche, Germany). Human cells from differentiation experiments were also lysed in 300 µl Tri-Reagent for 5 minutes at RT mechanically dissociated/lysed using a sterile scraper. RNA was then isolated as per Invitrogen's manufacturer's instructions for Tri-reagent as described in our previous work (REFS). Then a one-step RT-PCR reaction (reverse transcription and PCR) was performed using QuantiFastTM SYBR® Green RT-PCR one-step kits on a Rotorgene 3000Q. Each reaction was setup as follows; 4.75 µl experimental sample (7.36 ng/µl totalling 35 ng per reaction), 0.075 µl of both forward and reverse primer of the gene of interest (100 µM stock suspension), 0.1 µl of QuantiFast RT Mix (Qiagen, Manchester, UK) and 5 µl of QuantiFast SYBR Green RT-PCR Master Mix (Qiagen, Manchester, UK). Reverse transcription was initiated with a hold at 50˚C for 10 minutes (cDNA synthesis) and a 5-minute hold at 95˚C (transcriptase inactivation and initial denaturation), before 40-50 PCR cycles of; 95˚C for 10 sec (denaturation) followed by 60˚C for 30 sec (annealing and extension). Primer sequences for genes of interest and reference genes are included in Table 1. All genes demonstrated no unintended targets via BLAST search and yielded a single peak after melt curve analysis conducted after the PCR step above. All relative gene expression was quantified using the comparative Ct ($^{\Delta\Delta}$Ct) method [57]. For human cell differentiation analysis via measurement of myoD and myogenin, a pooled mean Ct for the 0 h young adult control samples were used as the calibrator when comparing aged vs. young adult cells. This approach demonstrated a low % variation in C_t value of 9.5 and 8.5% for myoD and myogenin, respectively. For HOX gene analysis between young and aged tissue and for the 7 d aged cells vs. 7 d young adult cells, the mean Ct of the young adult cells were used as the calibrator. The average, standard deviation and variations in Ct value for the B2M reference gene demonstrated low variation across all samples (mean ± SD, 13.12 ± 0.98, 7.45% variation) for the analysis of myoD and myogenin. For HOX gene analysis, the RPL13a reference gene variation was low in the human tissue (17.77 ± 1.71, 9.6% variation) and cell (15.51 ± 0.59, 3.82% variation) experiments. The average PCR efficiencies of myoD and myogenin were comparable (94.69 ± 8.9%, 9.4% variation) with the reference gene B2M (89.45 ± 3.76%, 4.2% variation). The average PCR efficiencies of the all genes of interest in the tissue analysis of the HOX genes were also comparable (90.87 ± 3.17%, 3.39% variation) with the human reference gene RPL13a (92 ± 2.67%, 2.9% variation). Similarly, for the cell analysis, HOX genes of interest efficiencies were comparable (89.59 ± 4.41%, 4.93% variation 4.93%) with the reference gene RPL13a (89.57 ± 3.55%, 3.97% variation).

Physical Activity and DNA methylation

The human association studies involved 30 physically active and endurance- oriented men of Eastern European descent (age 32.9 ± 9.9 years). The study was approved by the Ethics Committee of the Physiological Section of the Russian National Committee for Biological Ethics and Ethics Committee of the Federal Research and Clinical Center of Physical-

chemical Medicine of the Federal Medical and Biological Agency of Russia. Written informed consent was obtained from each participant. The study complied with the guidelines set out in the Declaration of Helsinki and ethical standards in sport and exercise science research. Physical activity was assessed using questionnaire. Participants were classified as mildly active (1-2 training sessions per week, n=6), moderately active (3-4 training sessions per week, n=8) and highly active (5-7 sessions per week, n=16). DNA methylation of CpG-sites in the muscle fibres of the vastus lateralis was evaluated using Infinium MethylationEPIC BeadChip Array (Illumina, USA). Bisulfite conversion of genomic DNA (500 ng of DNA was taken into the reaction) was performed using the EpiMark® Bisulfite Conversion Kit in accordance with the manufacturer's instructions. For EPIC BeadChips imaging the Illumina iScan® System (Illumina, United States) was used. Preprocessing of raw *.idat files was performed with the R/Bioconductor package RnBeads v 2.0.0. BMIQ algorithm was used for correcting the bias of type-2 probe values. Greedycut algorithm was used for filtering out low-quality probes. All probes with known SNPs and probes on a previously published blacklist of cross-hybridized were also filtered [58]. The final analyses included 796,180 of 868,565 (92%) probes passed QC procedures. QC for samples was performed with Greedycut algorithm from RnBeads package. Outliers were detected by PCA (exclusion criteria: an absolute value of z-score of PC1 or PC2 > 5). All samples passed the QC. The methylation level of each CpG-site after normalization and filtering processes was represented as a β-value ranging from 0 (unmethylated) to 1 (fully methylated) in order to undertake multiple regression (as regression analysis performed better with finite values. Ideally, if values range from 0 to 1 with beta-values satisfying this criteria). Statistical analyses were conducted using PLINK v1.90, R (version 3.4.3) and GraphPad InStat (GraphPad Software, Inc., USA) software. All data are presented as mean (standard deviation). Multiple regression was used for testing associations between the CpG-methylation level and physical activity adjusted for age. P values < 0.05 were considered statistically significant to test the 54 HOX CpG sites identified.

Results

Aged human skeletal muscle tissue is hypermethylated compared with young adult tissue

At an FDR ≤ 0.05 there were 17,994 CpG sites significantly differentially methylated CpGs in aged versus young adult tissue (Suppl. File **1a**), with a more stringent FDR of ≤ 0.01 there were 6,828 CpG sites (Suppl. File **1b**). Out of the most highly significant 6,828 CpG sites in aged muscle, the majority (6,384) were hypermethylated with only 444 hypomethylated versus young adult muscle. A finding that can be visualised via hierarchical clustering of the aged and young adult tissue methylation data (Figure **1a)**. Furthermore, these CpG sites were mainly located in CpG islands (5,018 out of 6,828 CpG's) with the remaining in N-shores (354), S-Shelf (48), S-Shores (341), S-Shelf (69) and 'other' (998). Indeed, 4,976 CpG's out of the 5,018 CpG sites within CpG islands were also significantly hypermethylated in aged compared with young adult tissue (Suppl. File **1c**). After gene ontology analysis, we also confirmed that the majority of the CpG's enriched within the three overarching gene ontology terms: 'Biological process' (Figure **1b**), 'cellular component;' (Figure **1c**) and 'molecular function' (Figure **1d**) demonstrated a predominantly hypermethylated profile in aged compared with young adult muscle. Within the GO terms that include the search term 'muscle', the most significantly enriched was, 'regulation of muscle system process' (Figure

1e, CpG list Suppl. Figure **1e**). Within this GO term was also the term 'regulation of muscle contraction' (CpG list Suppl. file **1f**). Once again, the predominant number of CpG's within these muscle specific gene ontologies demonstrated a hypermethylated versus hypomethylated profile in aged tissue compared with young adult tissue (e.g. 96 out of 110 CpG's and 70 out of 83 CpG's were hypermethylated in 'regulation of muscle system process' and 'regulation of muscle contraction' respectively). Further, enrichment analyses of methylation within KEGG pathways between aged and young adult tissue, identified the top significantly differentially enriched pathways, as: 'Pathways in cancer', 'Rap1 signaling', 'Axon guidance', 'Hippo signaling' (Suppl. file **1g**). These enriched KEGG pathways also displayed predominantly hypermethylated profiles in aged tissue. For example, within the top enriched pathway, 'pathways in cancer', 266 out of 277 CpG's were hypermethylated and only the remaining 11 hypomethylated in aged compared with young adult muscle tissue (Figure **2**; Suppl. File **1h**). Further, following analysis of most significantly differentially methylated CpG loci list for differentially methylated chromosomal regions in aged versus young adult tissue (Suppl. File **1i**), the top 5 regions included: chr8:22422678-22423092 (415 bp) within a CpG island of the SORBS3 gene with 6 CpG's that were all hypermethylated in aged versus young adult tissue. Also, chr6:30653732-30655720 (1989 bp) within a CpG island of the PPPR1 gene contained 8 CpG's that were all hypermethylated compared with young adult muscle. Chr20:13200939-13202437 (1499 bp) within a CpG island of the promoter of gene ISM, contained 8 CpG sites were also all hypermethylated. Similarly, the gene PDE4D1P on Chr1:144931376-144932480 (1105 bp) contained 6 CpG sites that spanned its promoter within a CpG island, once again demonstrating hypermethylation. The only gene demonstrating the opposite direction in methylation status (hypomethylation) in aged tissue was the gene C1orf132, also on chromosome 1 (Chr1:207990896-207991936; 1041 bp), where 5 CpG's within this region were hypomethylated in opposition to the majority of gene regions that were hypermethylated in aged versus young tissue.

Aged primary human muscle stem cells displayed more varied DNA methylation signatures versus young adults during the differentiation time-course

We analysed DNA methylation from differentiating human muscle stem cells at 0 h (30 minutes post dropping to 2% serum), 72 h, 7 and 10 d. Both young and aged adult cells morphologically demonstrated increasing differentiation and myotube formation across the time-course (Figure **3a**). This was confirmed by increases in gene expression in myoD and myogenin in both young adult and aged cells as differentiation progressed (Figure **3b**). Indeed, myotube formation appeared not to be as extensive in the elderly cells, despite the number of starting myogenic cells being the same in both aged and young cells. This reduced capacity for differentiation in the aged cells was confirmed at the molecular level by reduced myoD and myogenin gene expression at 72 h and 10 d in elderly cells compared with the young adult cells (Figure **3b**). Indeed, as well as reduced myogenin in aged cells at 72h, any increases in myogenin gene expression were also delayed in aged cells compared with young cells (Figure **3b**), with this previously been shown to be associated with impairing the fusion process and myotube formation [31, 32]. Indeed, the interaction for a 2-way ANOVA (Age × Time) generated a list of 40,854 CpG sites that were significantly altered with both age and across the time-course of differentiation (FDR ≤ 0.05, Suppl. File **2a**). With a more stringent FDR ≤ 0.01 and ≤ 0.001 there were still 9,938 and 2,020 CpG's significantly differentially methylated respectively in aged cells versus young cells across all time points

(Suppl. File **2b** and **2c** respectively). Overall demonstrating that there were highly significant differences in DNA methylation between aged and young cells during differentiation.

After determining these changes with both age and time. We next wished to assess if there were changes in DNA methylation within each cell type (aged and young adult) across differentiation, to identify if differentiation itself was changing DNA methylation profiles in the aged or young adult cells. In order to do this, we further analysed the main effect for TIME (0, 72 h, 7 d, 10 d) however, this did not demonstrate any significant changes in DNA methylation with an FDR ≤ 0.05. For example, even with a less stringent unadjusted p-value of 0.001 there were only a small number of CpG's (333) considered significant across all differentiation time points (Suppl. File **2d**). This suggested, that DNA methylation was not considerably changing over the time course of muscle stem cell differentiation itself. However, in order to investigate this more closely within aged and young cells at each time point, we performed contrasts across differentiation, first in young adult cells (FDR ≤ 0.05) only, and confirmed there was little to no changes in DNA methylation as the young cells differentiated, with only 1 and 14 CpG's significantly modified at 72 h and 10 d versus baseline levels at 0 h respectively, and with no CpG's significantly different at 7 d vs 0 h (FDR ≤ 0.05). In aged cells similar to young cells, there were no differences observed early after the induction of differentiation in DNA methylation profiles, with no CpG sites significantly different at 72 h vs. 0 h (FDR ≤ 0.05). However, at the later time point of differentiation, 7d, within aged cells, 2,785 CpG's were significant vs. 0 h (Suppl. file **2e**) and 404 CpG's were significantly different at 10 d vs. 0 h (FDR ≤ 0.05; Suppl. File **2f**). Even with a higher stringency (FDR ≤ 0.05 and change/difference in methylation greater than 2) at 7 days of differentiation there were 1,229 significantly different CpG sites versus 0 h in aged cells (Suppl. File **2g**). In this 1,229 CpG list there were balanced number of CpG's that were hypo and hypermethylated (compared with, for example, the aged vs. young tissue analysis above, where aged tissue demonstrated a predominantly profile hypermethylated versus young cells). Indeed, there were slightly more CpG's hypomethylated (676) versus hypermethylated (553) in the aged cells at 7 days of differentiation versus baseline levels at 0 h, where out of these CpG's, 73 were located in islands and 41 associated with promotors. This analysis of time, therefore suggested, that between 0-72 h in both young and aged cells there was little change in DNA methylation profiles. However, by 7 days in aged cells, DNA methylation (both hypo and hypermethylation) was significantly varied, whereas in young cells it remained unchanged. Overall, this may suggest a more dysfunctional differentiation program in aged human muscle cells that is associated with the impaired differentiation and myotube formation we observed morphologically (together with reduced myoD and delayed myogenin expression) in comparison with the young adult cells. Conducting gene ontology analyses at this 7 day time point between young and aged cells (Suppl. File **2h**), it suggested that this varied methylation response in aged cells was enriched in GO terms: 'regulation of localisation', 'regulation of cell communication' and 'regulation of signaling' (Figure **3c**; Suppl. File **2i, j, k** respectively). CpG's within these ontologies also confirmed that there was a similar hypo and hypermethylation profile in aged cells at 7 days. Further, KEGG pathway analysis suggested the top enriched pathways at 7 days in aged cells versus 0 h were: 'axon guidance', cholinergic synapse', 'adrenergic signaling in cardiomyocytes' and 'circadian entrainment' (Suppl. File **2l**). Finally, differential region analysis between young and aged cells at 7 days of differentiation identified two regions on 2 genes that had 5 or more CpG sites significantly differentially methylated (Suppl. File **2m**). The first being Chr3:155394096-155394303 (209 bp) located on gene PLCH1 that had 5 of its CpG's hypomethylated. The

second being a region non-annotated in location Chr2: 119615888-119617128 (1241 bp) that was hypermethylated on 5 CpG's. Suggesting that enriched and varied methylation of these genes occurred at 7 d differentiation in aged cells that was not detected at 7 days in young adult cells.

Aged primary human muscle cells demonstrate hypermethylated signatures versus young adults across differentiation, particularly at 7 days.

We next wished to identify the differences in DNA methylation in aged versus young cells across differentiation at each time point. A one-way ANOVA with the main effect for 'age', that analysed aged cells versus young cells alone (yet included all time points 0, 72 h, 7, 10 d of differentiation), generated a significant differentially methylated (FDR ≤ 0.05) CpG list of 269,898 CpG sites that significantly varied in aged cells versus young cells across all time points. Even with a more stringent cut-off (FDR ≤ 0.01) there were still 159,235 sites significantly modified (Suppl. File **3a**). Increasing the stringency to a change of greater than 2 while keeping an FDR ≤ 0.01 identified 2,719 CpGs with the most highly significant and demonstrated the largest differences between aged and young cells (Suppl. File **3b**). As with the aged versus young skeletal muscle tissue analysis, the majority of these significantly modified CpG sites in aged cells were hypermethylated (2,017 out of 2,719) versus hypomethylated (702 out of 2,719) compared with young cells, visualised in hierarchical clustering of these CpG sites (Figure **4a**). Two hundred and eleven out of these 2,719 CpG's were located in islands and 97 were promoter associated. Gene ontology identified that aged cells demonstrated this significantly enriched hypermethylated profile in GO terms: 'developmental process', 'anatomical structure development', 'anatomical structure morphogenesis' (Suppl. File **3c)** and also in KEGG pathways 'axon guidance', adherens junction', 'calcium signaling', 'focal adhesion' and 'protein digestion and absorption' (Suppl. File **3d).** Analysis of the significantly differentially methylated regions using the 2,719 CpG list above (Suppl. File **3e**), containing the most CpG sites within a given region, identified that a non-coding location on chr12:115134344-115135576 (1233 bp) that contained 13 CpG's that were hypermethylated in aged cells versus young cells. Further, there were 7 CpG's that were hypermethylated in the region of the gene LY6G5C (Chr6:31650736-31651158, 423 bp). There was also region containing 8 CpG's of the HOXC10 gene just upstream of HOXC6 and MIR196 (Chr12:54383692-54385621, 1933 bp) that all demonstrated a hypomethylated profile in aged cells vs. young cells. Interestingly, on the same chromosome just upstream of the HOXC10 gene (Chr12:54376020-54377868, 1849 bp), within the lncRNA HOXC-AS3, there were another 6 CpG's that were hypomethylated in aged cells versus young cells. The concentration of hypomethylated CpG's in aged versus young cells in HOXC genes is interesting given the majority of CpG's, more generally, were hypermethylated in aged cells versus young cells. A finding that is analysed further in the results below.

Given these changes, predominantly favouring hypermethylation (except the HOXC genes identified above) in aged cells versus young cells, we also performed contrasts between aged and young adult cells at each time point of differentiation. When we compared 0 h aged cells with young adult cells at 0 h, there were significant changes in 1,569 CpG sites (FDR ≤ 0.05, with a difference/change greater than 2), this included 738 significantly different CpG sites. Out of these 738 CpG's, 582 were hypermethylated and only 156 hypomethylated in aged versus young cells (Suppl. File **3f**). Gene ontology also confirmed that this

hypermethylated DNA profile in aged cells versus young adult cells at 0 hrs (Suppl. File **3g**) was located in enriched gene ontology terms including: 'cytoskeletal protein binding' (Figure **4b;** Suppl. File **3h**), 'developmental process', 'cell junction' and 'cytoskeleton' and 'actin binding' (Suppl. File **3i, j, k, l** respectively). With KEGG analysis of 0 h aged versus 0 hr young adult cells (Suppl. File **3m**) also suggesting that there were significantly enriched hypermethylated pathways for 'Axon Guidance' (Figure **4c**; CpG List Suppl. File **3n**), 'Insulin secretion', 'Phospholipase D signaling', 'cAMP signaling pathway' and 'Aldosterone synthesis and secretion' (Suppl. Files **3o, p, q, r** respectively).

At 72 h after the induction of differentiation in aged primary human muscle cells versus young adult cells, there were 1,418 significantly differentially methylated CpG sites (FDR ≤ 0.05, with a difference/change greater than 2) there were 645 significant CpG sites between aged and young adult cells (Suppl. File **4a**) again with a greater number hypermethylated (478 out of 645) versus hypomethylated (167 out of 645) in aged vs. young cells at this time point. Once again, as with the GO enrichment in the 0 h analysis above, at 72 h this analysis also identified (Suppl. File **4b**); 'cytoskeletal protein binding' (Suppl. File **4c**), 'actin binding' (Figure **4d**, Suppl. File **4d**), 'cytoskeleton' and 'development process' as most significantly enriched, followed by 'regulation of signaling' (Suppl. Files **4e, f, g** respectively). Again, these GO terms included CpG sites that were predominantly hypermethylated in aged muscle cells at 72 h versus young adult cells at the same time point. Similarly, to the 0 h KEGG pathways analysis, at 72 hrs (Suppl. File **4h**) this analysis also identified: 'Insulin secretion' (Suppl. File **4i**), 'aldosterone synthesis and secretion' (Suppl. File 4j) and 'cAMP signaling pathway' (Figure **4e**, Suppl. Files **4k**) as significantly enriched. Again, demonstrating overall a predominately hypermethylated profile in aged cells at 72 h of differentiation versus young cells at the same timepoint.

For 7 d aged vs. 7 d differentiated young cells (FDR ≤ 0.05, with a difference/change greater than 2) there was the largest number of CpG sites significantly differentially methylated, 5,524 CpG sites (Suppl. File **5a**). This was considerably more than that identified between aged and young cells at the other time points of 0, 72 h (above) and 10 d (below). This included 4,061 (out of 5,524 CpG's) hypermethylated versus only 1,463 (out of 5,524 CpG's) hypomethylated. Gene ontology suggested that this increased hypermethylation in aged cells at 7 d versus young adult cells at 7 d (Suppl. File **5b**), were enriched within terms: 'developmental process' (Figure **5a**, Suppl. File. **5c**) 'anatomical structure morphogenesis' (Suppl. File **5d**), 'neuron part', 'cytoskeletal protein binding' and 'cell junction' (Suppl. File **5d, e, f, g**). KEGG analysis (Suppl. File **5h**) identified: 'Focal adhesion' (Figure **5b**; Suppl. file **5i**), 'adherens junction' (Suppl. File **5j**), 'regulation of actin cytoskeleton' (Suppl. File **5k**), 'cGMP-PKG signaling pathway' Suppl. File **5l**), 'rap1 signaling pathway' (Suppl. File **5m**), as well as the well-known muscle differentiation pathway of 'PI3K-Akt signaling pathway' (Suppl. File **5n**) to be predominantly hypermethylated in aged cells at 7 d versus young cells at the same time point. Analysis of the significantly differentially methylated regions using this 5,524 CpG list at 7 d in aged cells versus 7 d young adult cells (Suppl. File **5o**) also identified that a region of the a gene within the HOX family (identified in earlier analysis above), HOXB1, located in Chr17:46607104-46608568, contained 8 CpG's that were hypermethylated with this 1465 bp region. Also, that a location within Chr3:113160184-113160624 spanning 441 bp also contained 6 CpG's that were hypermethylated spanning the gene WDR52. As well as a gene already identified to be hypermethylated in differentiation in aging cells only at 7 days, the gene LY6G5C, this time spanning a slightly larger region Chr6:31650736-31651362, of

627 bp (vs. Chr6:31650736-31651158, 423 bp in the above analysis), confirming that 6 of its CpG's were hypermethylated at 7 d compared with young adult cells at the same time point. Finally, CTD-2035E11.3 and CTD-2201E18.3/ RPS23P5 located on Chr5:43019772-43020628 (857 bp) and the gene NLRP1 (Chr17:5487376-5488048, 673 bp) were both hypomethylated on 6 of their CpG's within those regions.

For 10 d aged vs. 10 d differentiated young cells with ((FDR ≤ 0.05 with a change greater than 2) there were 288 significantly differentially methylated CpG sites (Suppl. File **6a**). Of which 255 were hypermethylated and only 33 hypomethylated. Gene ontology of the 288 CpG list, suggested that this increased hypermethylation in aged cells at 10 d versus young adult cells at 10 d (Suppl. File **6b**), were enriched within terms 'GDP binding' (Figure **5c**, Suppl. File **6c**), 'phosphotransferase activity', 'positive regulation of antigen receptor', 'mesoderm morphogenesis' and 'actin binding' (Suppl. File **6d, e, f, g** respectively). KEGG analysis (Suppl. File **6h**) identified 'Regulation of actin cytoskeleton' (Figure **5d**; Suppl. File **6i**), 'ErbB signaling pathway' (Suppl. File **6j**) and 'Selenocompound metabolism' (Suppl. File **6k**) to be predominantly hypermethylated in aged cells at 10 days versus young cells at the same time point.

SOM profiling of aged muscle stem cells confirms a varied methylation profile

With the greatest number of significantly differentially methylated CpG's at 7 d in aged cells versus the young. In order to visualise the temporal changes in methylation over the time course of differentiation in aged versus young cells, we conducted SOM profiling analysis of the 2,719 significantly differentially methylated CpG list generated above (those highly significant for 'age', across all time-points of differentiation). SOM analysis averages the methylation values for the group of samples within each condition to identify temporal profile changes in significantly differentially methylated CpG's between aged and young cells over the time-course of differentiation. Confirming the data above, it was also clear using this analysis that aged cells demonstrated a more varied (and perhaps dysfunctional) methylation profile earlier in the time-course of differentiation (between 72 h and 7d) compared with young adult cells (Figure **6a**). When looking at aged cells temporal change individually compared with young cells across the differentiation time-course, out of the 2,719 CpG list, in aged cells 1,504 of them were hypomethylated at 7 days and 956 hypermethylated at 7 days (confirming the aged cells 'time' main effect analysis above) (Figure **6a**). However, in contrast, only a small number of genes demonstrated this altered profile at 7 days in young cells (284 hypo and 110 hyper out of the 2,719 CpG list).

Finally, in order to identify common CpG changes across both tissue and cells and to determine if there were any CpG sites demonstrating similar methylation profiles once isolated *in-vitro* compared with the tissue they were derived from *in-vivo,* the significantly differentially methylated CpG lists from young cells vs. aged cells across all time were overlapped (2,719 CpG list), the 0 h aged vs. 0 h young (738 CpG list), 72 h aged vs. young 72 h (645 CpG list), 7 d aged vs. 7 d young (5,524 CpG list), and the 10 d aged vs 10d young (288 CpG list) (Venn diagram, Figure **6b**). There were 24 genes that were identified across both tissue and cell significantly differentially methylated CpG lists (Suppl. File **7a**). They included 24 CpG's on 16 genes: TSPAN9, RBM22, UBAP1, CAPZB, ZNF549, MBNL2, RMI1, CHRM5, RAB4A, C19orf21, MOBKL1A, ANAPC11, GAS7, PBX1, ELOVL2, FGGY. Furthermore, generating a SOM profile of temporal change in methylation over the time-

course of differentiation in these 24 CpG's, visualised in (Figure 6**c**), also demonstrated that the majority of these CpG's in aged cells at 7 days (16 out of 24 CpG's) demonstrated varied methylated at 7 days, including 11 out of 16 CpG's for annotated genes: TSPAN9, RBM22, UBAP1, CHRM5, C19orf21, ELOVL2, MOBKL1A, ANAPC11, PBX1, CAPZB, FGGY (all detailed in Suppl. File **7a).**

Distinguishing differentiation specific CpG sites in aging cells versus those altered as a consequence of age alone.

The data above suggest that aged cells across all time demonstrate hypermethylation versus young adult cells. Also, that aged cells significantly altered their methylation profiles during differentiation itself, particularly in a large number of CpG's at 7 days of differentiation versus young cells. Therefore, we conducted further analysis on the overlap of differentially methylated CpG's within aging cells at 7 days of differentiation (from the 'time' analysis above) and those that were changed as a consequence of age alone at 7 days (aged cells at 7 days versus young cells at 7 d). This enabled the identification of which methylation sites were altered, but also shared in both aged cell differentiation alone and as a consequence of age. Or, alternatively the sites that were simply changed with age and not differentiation process and vice versa. Indeed, overlapping the aged cells 0 h vs. 7 d (1,229 CpG list) with the 7 d aged vs. 7 d young significant 5,524 CpG list, there were only 334 (206 hypermethylated, 128 hypomethylated) CpG's that were shared (Suppl. File. **8a**). This suggested that differentiation itself modified only 334 CpG's (out of 1,229) in aged cells that were also changed as a consequence of age alone (i.e. in aged vs. young cells at 7 days). The remaining 895 CpG's (1,229 - 334 CpG's; Suppl. File **8b**) were differentiation specific to aged cells. Out of this 895 CpG list, an equal number were hypo and hypermethylated (458 hypo and 437 hypermethylated). Therefore, this overlap analysis also confirmed that data above, where over the time-course of aged cell differentiation itself the methylome is both hypo and hyper methylated on a similar number of CpG's, whereas aging alone predominantly hypermethylates (where out of 5,524 changed at 7 d young vs. aged, 4061 were hypermethylated and only 1,463 hypomethylated). This also suggested that the remaining 5,190 (5,524 minus 334 CpG list, equalling 5,190 CpG's with 3,910 hypermethylated vs. 1,280 hypomethylated) were as a consequence of aging alone (Suppl. File **8c**).

Hypermethylation for a small number of CpG sites is retained in aged muscle cells in-vitro from the in-vivo tissue niche.

We next examined whether methylated profiles were maintained from the tissue once isolated in an *in-vitro* setting. Indeed, 6 CpG's that were identified in the 6,828 significantly differentially methylated CpG list from the aged versus young tissue analysis, as well as highlighted in the 2,719 list 'age' cells (across all time) analysis (Figure **7**). These included: KIF15 (2 CpG's Cg00702638 & cg24888989), DYRK2 (cg09516963), FHL2 (cg22454769), MRPS33 (cg26792755), ABCA17P (cg02331561). Importantly, all of these genes were hypermethylated in the tissue analysis as well as the cells, highlighting that the cells were perhaps retaining their *in-vivo* hypermethylation one these particular CpG's once isolated *in-vitro*. Given that aged cells demonstrated the most varied methylation at 7 days versus young cells. When comparing 7 d aged vs. 7 d young cell list (5,524 list), 4 CpG's (out of the 6 CpG sites identified above) were also identified in the 6,828 tissue CpG list, including:

MGC87042 (cg00394316), C2orf70 (cg23482427), ABCA17P (cg02331561) and cg27209395 (not on an annotated gene). Once more, all of these CpG's (with the exception of C2orf70, cg23482427) were hypermethylated in the tissue analysis as well as the cells. With ABCA17P (cg02331561) highlighted across all gene lists (Figure **6**).

Varied methylation in the HOX family of genes in aging skeletal muscle tissue and stem cells

In the above analyses in aged versus young cells (all time points) there were 8 CpG's with altered methylation within the region of the HOXC10 gene just upstream of HOXC6 and MIR196 (Chr12:54383692-54385621, 1933 bp). Also, on the same chromosome just upstream of the HOXC10 gene (Chr12:54376020-54377868, 1849 bp) within lnRNA HOXC-AS3 there were another 6 CpG's that were altered in aged cells versus young cells. Similarly, using the 5,524 CpG list at 7 d in aged cells versus 7 d young adult cells (Suppl. File **5o**) we also identified that a region of the HOXB1, located in Chr17:46607104-46608568 contained 8 CpG's that were hypermethylated with the 1465 bp region. Therefore, we further analysed all the HOX gene changes in the significantly differentially methylated aged vs. young tissue (6,828 CpG list) and identified that CpG's within HOXD10, HOXD9, HOXD8, HOXA3, HOXC9, HOXB1, HOXB3, HOXC-AS2 and HOXC10 were significantly differentially methylated across all analyses, including the aged vs. young tissue (6828 CpG list; Suppl. File **9a**), aged versus young adult cells (all time) 2,719 CpG list (Suppl. File **9b**) as well as the 7 d aged versus 7 d young cell analysis 5,524 CpG list (Suppl. File **9c**). The timepoint that showed the largest number of significant changes during the time course of differentiation. A Venn diagram analysis depicted the overlap in common gene symbols for these three gene lists (Figure **8a,** Suppl. File **9d**). It was also demonstrated that the majority of these HOX genes were hypermethylated in aged tissue (Suppl. File. **9a**; Figure **8b**). In the cell analysis across all time, these HOX genes also displayed the most varied methylation at 7 days of differentiation in aged cells versus young cells, therefore confirming the varied temporal profile in methylation described above at 7 d was also the case for these HOX genes identified. Finally, when SOM-profiling these 9 HOX genes by symbol (17 CpG's as some HOX genes contained more than one CpG site), over the time-course of differentiation based on the 'age' cells (all-time) 2,719 significantly differentially methylated CpG list. Eight CpG sites across HOX family genes: HOXD8 (cg18448949, cg00035316), HOXB1 (cg04904318, cg02497558, cg22660933, cg10558129), HOXC9 (cg02227188) and HOXB3 (cg09952002) were hypermethylated at 7 days, whereas and 7 CpG's across HOXC10 (cg20402783, cg20403938, cg16898193, cg16791250), HOXB3 (cg23014425, cg04800503), HOXC-AS2 (cg09605287) were hypomethylated at 7 days (Figure **8c;** Suppl. File **9e**). This meant distinct genes were hypermethylated (HOXD8, HOXC9, HOXB1) and hypomethylated (HOXC10, HOXC-AS2) at 7 days in aged cells versus young adult cells, except for one of these genes, HOXB3, that contained 1 CpG that was hypermethylated versus 2 that were hypomethylated.

The HOX genes that changed in both the tissue and cell analysis (by gene symbol) at the mRNA expression level (HOXD8, HOXA3, HOXC9, HOXB1, HOXB3, HOXC-AS2 and HOXC10) were further analysed in order to associate the significant changes in DNA methylation with corresponding gene expression. Interestingly, in aged tissue, despite displaying hypermethylation in all these HOX genes versus young tissue, the genes were not

suppressed at the gene expression level, which may have been expected, yet all significantly elevated (Figure **9a**). However, in the aged cells when analysing these genes at the expression level at 7 days of differentiation (including HOXD8, HOXA3, HOXC9, HOXB1, HOXB3, HOXC-AS2 and HOXC10, as well as HOXC-AS3 identified in the above in the cell analysis only) (Figure **9b**), we identified that there was significantly reduced gene expression in gene HOXB1 (Figure **9b**), that corresponded with HOXB1 hypermethylation. Where in the 5,524 CpG list at 7 d in aged cells versus 7 d young adult cells (Suppl. File **5o**), we previously identified that a region of the HOXB1 located in Chr17:46607104-46608568, contained 8 CpG's that were hypermethylated, as well as HOXB1 Cg's: cg04904318, cg02497558, cg22660933, cg10558129 being identified as hypermethylated in the 2,719 significant main effect for 'age' cell CpG list above (Figure **8c**, Suppl. File **3b**). Further, there was significantly increased gene expression for HOXC-AS3, with this gene identified earlier in the analysis as being with hypomethylated. Indeed, hypomethylation occurred in 6 CpG's in a region upstream of the HOXC10 gene (Chr12:54376020-54377868, 1849 bp) within the HOXC-AS3 gene. Interestingly, HOXC10 also demonstrated an average increase in gene expression, however, it was not statistically significant (Figure **9b**). Finally, HOXA3 also demonstrated significantly reduced expression (Figure **9b**) with corresponding hypermethylation (yet at 10 not 7 days of differentiation) in aged cells (see Figure **8c**). Overall, HOXB1, HOXC-AS3 and HOXA3 demonstrated an inverse relationship with CpG methylation and gene expression in aged versus young cells.

Effect of physical activity on methylation status of HOX family genes

Next, given that aging generally hypermethylated the genome, we have tested the hypothesis that physical activity may be associated with a reverse these epigenetic changes in the *HOX* family of genes. We thus performed a multiple regression analysis using methylation data of 54 sites of *HOX* family genes (those across Suppl. File 9 a,b,c) and the level of physical activity of 30 endurance-trained men. We determined that highly active men had hypomethylated *HOXB1* (cg10558129, $P = 0.032$) and *HOXA3* (cg16406967, $P = 0.038$), *HOXA13* (cg04428700, $P = 0.017$) genes and hypermethylated *HOXC10* gene (cg20403938, $P = 0.045$) compared to less active men (adjusted for age). Given that we identified the opposite trend in aged muscle and cells, where that particularly HOXB1 and HOXA3 were hypermethylated in aged tissue and cells with reduced gene expression in aged cells, and HOXC10 was hypomethylated in aged tissue and cells and the average gene expression increased. These findings suggest that increase exercise is associated with reversed methylation status in these HOX genes compared to age-related changes.

Discussion

In the present study we aimed to investigate the methylome in aged skeletal muscle tissue and differentiating primary muscle stem cells compared with young adults, in order to identify important epigenetically regulated genes in both aging skeletal muscle tissue and stem cells. As with previous studies ([17, 26], and by using more recent, higher coverage, array technology, we identified that aged skeletal muscle tissue demonstrated a considerably hypermethylated profile compared with young adult tissue. We also demonstrated that these hypermethylated profiles in aged tissue were enriched in gene ontology pathways including, 'regulation of muscle system process' and KEGG pathways 'pathways in cancer', a pathway that incorporates previously well described molecular pathways in the regulation of skeletal

muscle such as; focal adhesion, MAPK signaling, PI3K-Akt-mTOR signaling, p53 signaling, Jak-STAT signaling, TGF-beta and Notch signaling, as well as the other significantly enriched pathways of 'rap1 signaling', 'axon guidance', and 'hippo signaling'. This was also the first study to profile DNA methylation over the entire time-course of skeletal muscle differentiation (0, 72 h, 7 and 10 d) using the highest coverage 850K methylation assays. In primary cell cultures, isolated from aged and young adults matched for the proportion of myogenic cells, we identified that (as observed at the tissue level) aged muscle stem cells also demonstrated hypermethylated profiles versus young adult cells. This hypermethylation was enriched in: 'axon guidance', 'adherens junction' and 'calcium signaling' pathways. Furthermore, we identified that the process of cellular differentiation itself did not significantly affect DNA methylation in young cells, however aged cells demonstrated varied methylation profiles particularly at 7 d of differentiation in GO terms: 'regulation of localisation', 'regulation of cell communication', and 'regulation of signaling'. Furthermore, in the majority of different CpG sites that were altered during the process of differentiation alone, aged cells demonstrated significantly hypermethylated profiles during differentiation when compared with young adult cells, again, specifically at 7 d of differentiation including CpG's located within: 'focal adhesion', 'adherens junction' and 'regulation of actin cytoskeleton' pathways, as well as the well-known muscle differentiation pathway, 'PI3K-Akt signaling'. This corresponded with reduced differentiation and myotube formation observed morphologically as well as associated reductions in myoD and myogenin and delayed increases in myogenin gene expression in aged compared with young adult cells. Furthermore, we were able to identify that a small number of CpG sites hypermethylated in aged tissue were also hypermethylated in aged cells, with CpG's located on genes: KIF15, DYRK2, FHL2, MRPS33, ABCA17P. This suggested, that these CpG's retained their methylation status *in-vitro* after being isolated from the *in-vivo* niche. However, it is worth noting that this was a very small number of CpG's compared with the thousands of CpG sites significantly differentially methylated in the aged tissue versus young adult tissue, and those also observed to be significantly different in the aged versus young muscle cells. This was suggestive that the majority of the hypermethylated CpG's observed at the aged tissue level were generally not retained on the same CpG sites in the isolated aged muscle stem cells *in-vitro*. This therefore also perhaps indicates that hypermethylation of DNA within myonuclei (the predominant source of DNA in the tissue) maybe therefore unique to that observed in the isolated aged muscle cells. Indeed, retention of methylation during aging has been previously observed in artificially aged muscle cells [18]. In skeletal muscle tissue, retention of DNA methylation has been observed after skeletal muscle growth, even after a subsequent period of muscle loss, suggesting an epigenetic memory at the DNA methylation level [20, 21, 48]. However, the relative contribution of methylation from myonuclei or satellite cells (or other resident cell types in muscle) to this epigenetic memory effect, and how these retained profiles can last (e.g. past 22 weeks in the Seaborne *et al*., 2018 study) has not been determined. However, interestingly based on the present studies data, we could hypothesise that myonuclear hypermethylation in the tissue with age is quite different to the hypermethylation observed in satellite cells (and daughter progeny). Overall, suggesting that environmental stimuli and aging could affect methylation profiles in the myonuclei differently than those in satellite cells. A hypothesis that requires further investigation, perhaps using single-cell methylome analysis of the different cell populations resident in cells derived from skeletal muscle tissue biopsies.

Importantly, in both tissue and stem cell analysis, we also identified that the homeobox (HOX) family of genes were significantly enriched in differentially methylated region analysis, showing several (e.g. 6-8) CpG's to be methylated within narrow chromosomal regions on these genes, as well as being methylated in the same direction (hyper or hypomethylated) in aged compared with young adults. In particular, we identified: HOXC10 (just upstream of HOXC6) and HOXB1 as having several CpG's differentially methylated. Therefore, closer analysis of all HOX gene associated CpG changes across both tissue and cell differentiation data identified that CpG's located within: HOXD10, HOXD9, HOXD8, HOXA3, HOXC9, HOXB1, HOXB3, HOXC-AS2 and HOXC10 were all significantly differentially methylated across these analyses. In aged tissue the majority of these HOX genes were hypermethylated. In the cell analysis, these HOX genes displayed the most varied methylation at 7 days of differentiation in aged cells versus young cells. Furthermore, distinct HOX genes were hypermethylated (HOXD8, HOXC9, HOXB1) and hypomethylated (HOXC10, HOXC-AS2) at 7 days in aged cells versus young adult cells. In aged cells versus young cells, gene expression analysis also demonstrated an inverse relationship with hypermethylation of HOXB1 and HOXA3 and reduced gene expression, as well as hypomethylation of HOXC-AS3 together with increased gene expression.

This data is interesting given that the HOX genes are evolutionary conserved members of the homeobox superfamily, with 39 HOX genes found in the mammalian genome. They are 'homeotic' genes coding for transcription factors, with a fundamental role in the determination of cellular identity. They were first shown to be important in embryogenesis in drosophila melanogaster (fruit fly) [59]. In muscle they have been described to morphologically identify the hindlimb tissues during development [60-63], but have also been demonstrated to be activated in satellite cells [64, 65], and as markers of hindlimb derived myoblasts [64]. In particular HOXC10, demonstrated 8 CpG's (just upstream of HOXC6 and MIR196; Chr12:54383692-54385621, 1933 bp) that all demonstrated a hypomethylated signature in aged versus young muscle stem cells. There were also 4 CpG sites hypomethylated, particularly at 7 days of differentiation in aged versus young cells. Indeed, HOXC10 has been identified to determine hindlimb identity [60, 62]. Together with HOXC10 hypomethylation, we also demonstrated average (yet not significant) increases in HOXC10 gene expression at 7 days in aged cells versus young cells. Counterintuitively, to our data, previously HOXC10 upregulation has been associated with satellite cell activation in skeletal muscle in response to Roux-en-Y gastric bypass (RYGB) surgery [66], as well as being a marker for hindlimb specific satellite cells. Interestingly, there was lower expression of HOXC10 observed in exercised rats [67], which is perhaps more intuitive with the data provided here, where aged cells demonstrated an increase. However, HOXC10 requires more experimentation to define its mechanistic role in aging skeletal muscle, with HOXC10 and physical exercise being discussed below. Interestingly, the hypomethylation of the HOXC10 (Chr12:54383692-54385621, 1933 bp) occurred in 8 CpG's and on the same chromosome just upstream of the HOXC10 gene (Chr12:54376020-54377868, 1849 bp), within the lncRNA HOXC-AS3, where there were another 6 CpG's that were hypomethylated in aged cells versus young cells (See Suppl. Figure **2** for visualisation of HOXC10 and HOXC-AS3 and their genomic location and methylation). HOXC-AS3 is a long coding RNA (lcRNA) with currently no known data in skeletal muscle, but some literature within cancer and mesenchymal stem cell (MSC) fields [68]. Indeed, MSC's administered with silencing HOXC-AS3 prevented bone loss by enhancing HOXC10 expression [69], and HOXC-AS3 upregulation has also been linked with aggressive cancer [70]. Interestingly, we were able to identify that together with associated

hypomethylation, in a region close to the lcRNA HOXC-AS3, there was significantly increased gene expression of HOXC-AS3 (and an average, yet not significant increase in HOXC10 gene expression) in aged muscle cells at 7 days of differentiation. Given the data above in bone and cancer, HOXC-AS3 upregulation appears to be pro-growth and linked with expression of HOXC10, therefore their increase in the current study maybe hypothesised to be a co-operative and compensatory drive to maintain aged muscle. However, this hypothesis is speculative, and more work needs to be conducted as to the role of HOXC10 and HOXC-AS3 and their potential cooperative mechanisms of action in aged skeletal muscle. We also identified that HOXB1 was hypermethylated with increased gene expression in aged cells at 7 days. HOXB1 has been demonstrated to be hypermethylated in inflamed muscle of children with Juvenile Dermatomyositis (JDM) [71]. This is interesting given aged skeletal muscle is known to be chronically inflamed [72, 73], where we also demonstrate this hypermethylated profile in methylation in aged cells. Finally, HOXA3 was hypermethylated with reduced gene expression in aged cells. However, there is currently little to no work on this gene in skeletal muscle [63] and therefore requires future investigation.

Finally, with aging evoking a hypermethylated signature in tissue and aged muscle stem cells, it was also interesting to speculate that physical exercise, that has been shown to hypomethylate the genome [19-21], could therefore be 'anti-ageing' at the epigenetic level. Indeed, this hypothesis was first supported indirectly in the present study and by previous literature, where the aged tissue analysis in the present study identified the top significantly enriched KEGG pathway as, 'pathways in cancer'. A pathway that incorporates well known pathways important in skeletal muscle, including: Focal adhesion, MAPK, PI3K-Akt, mTOR, p53 signaling, Jak-STAT, TGF-beta and Notch signaling. Furthermore, this pathway was also the top enriched hypomethylated pathway after acute and chronic resistance exercise [21]. Overall, suggesting exercise (resistance exercise in this case) could perhaps reverse the hypermethylated profiles in these pathways in aged muscle. Therefore, in the present study, given that we identified the HOX family of genes to be extensively differentially methylated in aged tissue and stem cells, we went on to determine that higher physical activity levels (endurance exercise in this case) in healthy young adults was associated with changes in HOXB1, HOXC10 and HOXA3 DNA methylation in the opposite direction to the changes we observed in aged tissue and stem cells. This also provided evidence to suggest that increased physical activity could reverse the age-related epigenetic changes in the HOX genes. However, more research into the effect of exercise in an aged population and the changes in HOX methylation status will be required in the future to confirm these findings.

Conclusion

Overall, for the first time, we demonstrate that altered methylation of a large number of the HOX genes are epigenetically differentially regulated in aged human skeletal muscle tissue and during impaired differentiation in aged muscle stem cells. In particular HOXB1, HOXA3 and HOXC-AS3 (and to a certain extent, HOXC10) also demonstrated significant changes in gene expression in aged cells. Importantly, that increased physical activity was associated with reversing age-related changes in HOXB1, HOXA3 and HOXC10 gene methylation.

Acknowledgments

These data were supported by a North Staffordshire Medical Institute (NMSI) grant awarded to Adam P. Sharples (PI), Daniel Turner, Mark Kitchen and Ian Dos-Remedios (Co-I's). This

work was also supported by the UK's Engineering and Physical Sciences Research Council (EPSRC) and the UK Medical Research Council's (MRC) centre for doctoral training, via a studentship awarded to the joint first author Piotr Gorski in the group of Adam P. Sharples (PI). Funds from Keele University and Liverpool John Moores University, UK and The Norwegian School of Sport Sciences, Oslo, Norway also supported the PhD work by Daniel Turner & Piotr Gorski in the group of Adam P. Sharples. Philipp Baumert received a fully-funded Liverpool John Moores PhD scholarship. Mohd Firdaus Maasar received a PhD studentship via the Malasyisan government agency: Majlis Amanah Rakyat (MARA) via Barry Drust, Adam P. Sharples and Dr. Andrew Hulten. The physical activity and DNA methylation study was supported in part by grant from the Russian Science Foundation (Grant No. 17-15-01436: "Comprehensive analysis of the contribution of genetic, epigenetic and environmental factors in the individual variability of the composition of human muscle fibers"; DNA sample collection, genotyping, epigenetic analysis and muscle fibre typing of Russian subjects).

References

1. Sharples, A.P., et al., *Longevity and skeletal muscle mass: the role of IGF signaling, the sirtuins, dietary restriction and protein intake.* Aging Cell, 2015. **14**(4): p. 511-23.
2. Morley, J.E., S.D. Anker, and S. von Haehling, *Prevalence, incidence, and clinical impact of sarcopenia: facts, numbers, and epidemiology-update 2014.* Journal of cachexia, sarcopenia and muscle, 2014. **5**(4): p. 253-259.
3. Hughes, V.A., et al., *Longitudinal muscle strength changes in older adults: influence of muscle mass, physical activity, and health.* Journals of Gerontology Series A-Biological Sciences & Medical Sciences., 2001. **56**(5): p. B209-17.
4. Cruz-Jentoft, A.J., et al., *Sarcopenia: European consensus on definition and diagnosis: Report of the European Working Group on Sarcopenia in Older People.* Age Ageing, 2010. **39**(4): p. 412-23.
5. Hairi, N.N., et al., *Loss of muscle strength, mass (sarcopenia), and quality (specific force) and its relationship with functional limitation and physical disability: the Concord Health and Ageing in Men Project.* J Am Geriatr Soc, 2010. **58**(11): p. 2055-62.
6. Cederholm, T., A.J. Cruz-Jentoft, and S. Maggi, *Sarcopenia and fragility fractures.* Eur J Phys Rehabil Med, 2013. **49**(1): p. 111-7.
7. Cooper, C., et al., *Tools in the assessment of sarcopenia.* Calcif Tissue Int, 2013. **93**(3): p. 201-10.
8. Porter, M.M., A.A. Vandervoort, and J. Lexell, *Aging of human muscle: structure, function and adaptability.* Scandinavian journal of medicine & science in sports, 1995. **5**(3): p. 129-142.
9. Cruz-Jentoft, A.J., et al., *Understanding sarcopenia as a geriatric syndrome.* Curr Opin Clin Nutr Metab Care, 2010. **13**(1): p. 1-7.
10. Beaudart, C., et al., *Health Outcomes of Sarcopenia: A Systematic Review and Meta-Analysis.* PLOS ONE, 2017. **12**(1): p. e0169548.
11. Fuggle, N., et al., *Sarcopenia.* Best practice & research. Clinical rheumatology, 2017. **31**(2): p. 218-242.

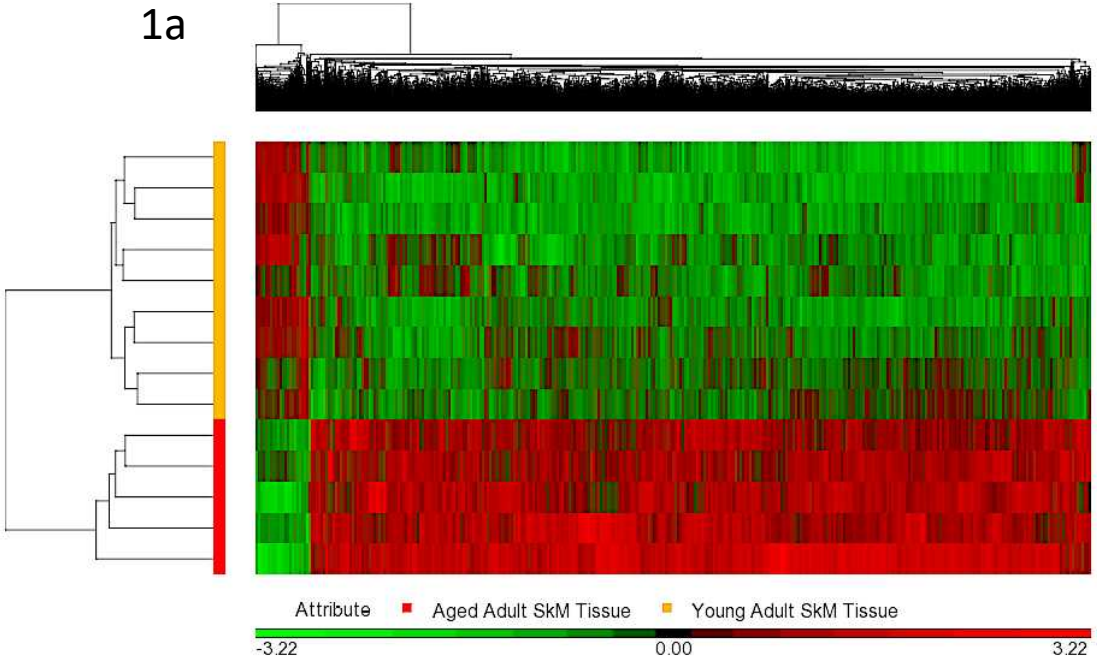
12. Sharples, A.P., C.E. Stewart, and R.A. Seaborne, *Does skeletal muscle have an 'epi'-memory? The role of epigenetics in nutritional programming, metabolic disease, aging and exercise.* Aging Cell, 2016. **15**(4): p. 603-16.
13. Sharples, A.P., R.A. Seaborne, and C.E. Stewart, *Epigenetics of Skeletal Muscle Aging*, in *Epigenetics of Aging and Longevity*, A.M. Vaiserman, Editor. 2018, Academic Press: Boston. p. 389-416.
14. Day, K., et al., *Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape.* Genome Biology, 2013. **14**(9): p. R102.
15. Voisin, S., et al., *An epigenetic clock for human skeletal muscle.* bioRxiv, 2019: p. 821009.
16. Horvath, S., *DNA methylation age of human tissues and cell types.* Genome biology, 2013. **14**(10): p. 3156.
17. Zykovich, A., et al., *Genome-wide DNA methylation changes with age in disease-free human skeletal muscle.* Aging Cell, 2014. **13**(2): p. 360-6.
18. Sharples, A.P., et al., *Skeletal muscle cells possess a 'memory' of acute early life TNF-alpha exposure: role of epigenetic adaptation.* Biogerontology, 2016. **17**(3): p. 603-17.
19. Sailani, M.R., et al., *Lifelong physical activity is associated with promoter hypomethylation of genes involved in metabolism, myogenesis, contractile properties and oxidative stress resistance in aged human skeletal muscle.* Scientific Reports, 2019. **9**(1): p. 3272.
20. Seaborne, R.A., et al., *Human Skeletal Muscle Possesses an Epigenetic Memory of Hypertrophy.* Scientific Reports (Nature) 2018. **8**(1): p. 1898.
21. Turner, D.C., R.A. Seaborne, and A.P. Sharples, *Comparative Transcriptome and Methylome Analysis in Human Skeletal Muscle Anabolism, Hypertrophy and Epigenetic Memory.* Scientific Reports, 2019. **9**(1): p. 4251.
22. Tsumagari, K., et al., *Early de novo DNA methylation and prolonged demethylation in the muscle lineage.* Epigenetics, 2013. **8**(3): p. 317-32.
23. Brunk, B.P., D.J. Goldhamer, and J.C.P. Emerson, *Regulated Demethylation of themyoDDistal Enhancer during Skeletal Myogenesis.* Developmental Biology, 1996. **177**(2): p. 490-503.
24. Fuso, A., et al., *Early demethylation of non-CpG, CpC-rich, elements in the myogenin 5'-flanking region: a priming effect on the spreading of active demethylation.* Cell Cycle, 2010. **9**(19): p. 3965-76.
25. Wu, W., et al., *Molecular characterization, expression patterns and polymorphism analysis of porcine Six1 gene.* Molecular Biology Reports, 2011. **38**(4): p. 2619-2632.
26. Bigot, A., et al., *Age-Associated Methylation Suppresses SPRY1, Leading to a Failure of Re-quiescence and Loss of the Reserve Stem Cell Pool in Elderly Muscle.* Cell Rep, 2015. **13**(6): p. 1172-82.
27. Barberi, L., et al., *Age-dependent alteration in muscle regeneration: the critical role of tissue niche.* Biogerontology, 2013. **14**(3): p. 273-292.
28. Hidestrand, M., et al., *Sca-1-expressing nonmyogenic cells contribute to fibrosis in aged skeletal muscle.* J Gerontol A Biol Sci Med Sci, 2008. **63**(6): p. 566-79.
29. Pietrangelo, T., et al., *Molecular basis of the myogenic profile of aged human skeletal muscle satellite cells during differentiation.* Exp Gerontol, 2009. **44**(8): p. 523-31.

30. Beccafico, S., et al., *Human muscle satellite cells show age-related differential expression of S100B protein and RAGE.* Age (Dordr), 2010.
31. Sharples, A.P., N. Al-Shanti, and C.E. Stewart, *C2 and C2C12 murine skeletal myoblast models of atrophic and hypertrophic potential: relevance to disease and ageing?* J Cell Physiol, 2010. **225**(1): p. 240-50.
32. Sharples, A.P., et al., *Reduction of myoblast differentiation following multiple population doublings in mouse C2 C12 cells: a model to investigate ageing?* J Cell Biochem, 2011. **112**(12): p. 3773-85.
33. Sharples, A.P., et al., *Modelling in vivo skeletal muscle ageing in vitro using three-dimensional bioengineered constructs.* Aging Cell, 2012. **11**(6): p. 986-95.
34. Merritt, E.K., et al., *Heightened muscle inflammation susceptibility may impair regenerative capacity in aging humans.* J Appl Physiol, 2013.
35. Zwetsloot, K.A., et al., *Non-passaged muscle precursor cells from 32-month old rat skeletal muscle have delayed proliferation and differentiation.* Cell Prolif, 2013. **46**(1): p. 45-57.
36. Bigot, A., et al., *Replicative aging down-regulates the myogenic regulatory factors in human myoblasts.* Biol Cell, 2008. **100**(3): p. 189-99.
37. Carlson, M.E. and I.M. Conboy, *Loss of stem cell regenerative capacity within aged niches.* Aging Cell, 2007. **6**(3): p. 371-82.
38. Lancioni, H., et al., *Muscle actin isoforms are differentially expressed in human satellite cells isolated from donors of different ages.* Cell Biol Int, 2007. **31**(2): p. 180-5.
39. Lees, S.J., C.R. Rathbone, and F.W. Booth, *Age-associated decrease in muscle precursor cell differentiation.* Am J Physiol Cell Physiol, 2006. **290**(2): p. C609-15.
40. Lorenzon, P., et al., *Ageing affects the differentiation potential of human myoblasts.* Exp Gerontol, 2004. **39**(10): p. 1545-54.
41. Charge, S.B., A.S. Brack, and S.M. Hughes, *Aging-related satellite cell differentiation defect occurs prematurely after Ski-induced muscle hypertrophy.* Am J Physiol Cell Physiol, 2002. **283**(4): p. C1228-41.
42. Allen, R.E., et al., *Influence of age on accumulation of alpha-actin in satellite-cell-derived myotubes in vitro.* Mech Ageing Dev, 1982. **18**(1): p. 89-95.
43. Kandalla, P.K., et al., *Mechano Growth Factor E peptide (MGF-E), derived from an isoform of IGF-1, activates human muscle progenitor cells and induces an increase in their fusion potential at different ages.* Mech Ageing Dev, 2011. **132**(4): p. 154-62.
44. Corbu, A., et al., *Satellite cell characterization from aging human muscle.* Neurol Res, 2010. **32**(1): p. 63-72.
45. Shefer, G., et al., *Satellite-cell pool size does matter: defining the myogenic potency of aging skeletal muscle.* Developmental Biology, 2006. **294**(1): p. 55-66.
46. Alsharidah, M., et al., *Primary human muscle precursor cells obtained from young and old donors produce similar proliferative, differentiation and senescent profiles in culture.* Aging Cell, 2013.
47. Seaborne, R.A., et al., *UBR5 is a novel E3 ubiquitin ligase involved in skeletal muscle hypertrophy and recovery from atrophy.* J Physiol, 2019. **597**(14): p. 3727-3749.
48. Seaborne, R.A., et al., *Methylome of human skeletal muscle after acute & chronic resistance exercise training, detraining & retraining.* Scientific Data (Nature), 2018. **5**: p. 180213.

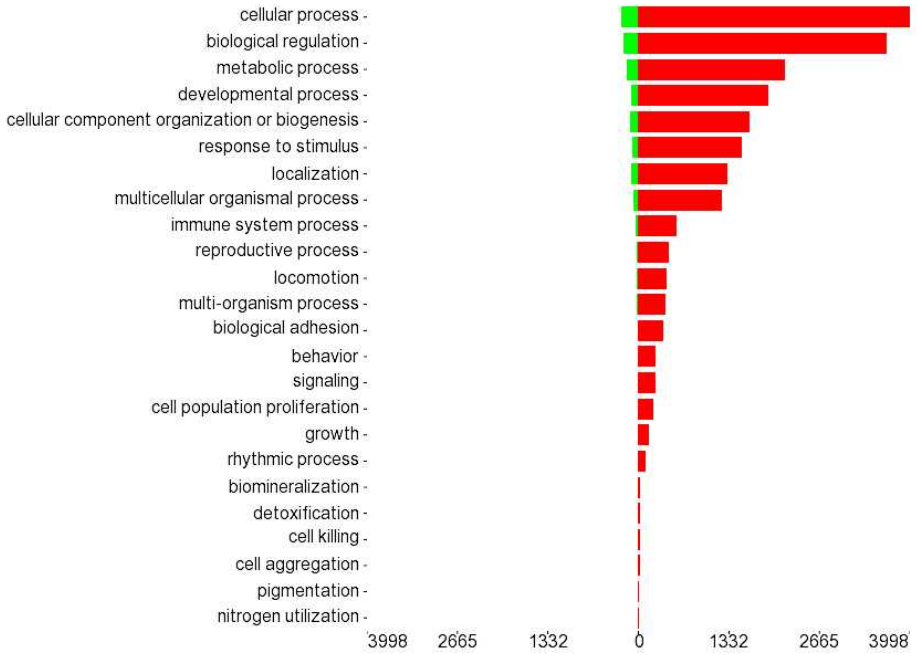
49. Turner, D.C., et al., *Exercising Bioengineered Skeletal Muscle In Vitro: Biopsy to Bioreactor*, in *Methods in Molecular Biology*. 2019. p. 1889; 55-79 (Doi: 10.1007/978-1-4939-8897-6_5).
50. Maksimovic, J., B. Phipson, and A. Oshlack, *A cross-package Bioconductor workflow for analysing methylation array data [version 1; referees: 3 approved, 1 approved with reservations].* F1000Research, 2016. **5**(1281).
51. Pidsley, R., et al., *Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling.* Genome Biology, 2016. **17**(1): p. 208.
52. Maksimovic, J., L. Gordon, and A. Oshlack, *SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips.* Genome Biol, 2012. **13**(6): p. R44.
53. Du, P., et al., *Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis.* BMC Bioinformatics, 2010. **11**(1): p. 587.
54. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic Acids Res, 2000. **28**(1): p. 27-30.
55. Kanehisa, M., et al., *KEGG: new perspectives on genomes, pathways, diseases and drugs.* Nucleic Acids Res, 2017. **45**(D1): p. D353-d361.
56. Kanehisa, M., et al., *KEGG as a reference resource for gene and protein annotation.* Nucleic Acids Res, 2016. **44**(D1): p. D457-62.
57. Schmittgen, T.D. and K.J. Livak, *Analyzing real-time PCR data by the comparative C(T) method.* Nat Protoc, 2008. **3**(6): p. 1101-8.
58. McCartney, D.L., et al., *Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip.* Genomics Data, 2016. **9**: p. 22-24.
59. Shah, N. and S. Sukumar, *The Hox genes and their roles in oncogenesis.* Nat Rev Cancer, 2010. **10**(5): p. 361-71.
60. DeLaurier, A., R. Schweitzer, and M. Logan, *Pitx1 determines the morphology of muscle, tendon, and bones of the hindlimb.* Dev Biol, 2006. **299**(1): p. 22-34.
61. Carpenter, E.M., et al., *Targeted disruption of Hoxd-10 affects mouse hindlimb development.* Development, 1997. **124**(22): p. 4505-14.
62. Logan, M. and C.J. Tabin, *Role of Pitx1 upstream of Tbx4 in specification of hindlimb identity.* Science, 1999. **283**(5408): p. 1736-9.
63. Manley, N.R. and M.R. Capecchi, *Hox group 3 paralogous genes act synergistically in the formation of somitic and neural crest-derived structures.* Dev Biol, 1997. **192**(2): p. 274-88.
64. Choi, H.Y., et al., *Thyroid-Stimulating Hormone Receptor Expression on Primary Cultured Human Extraocular Muscle Myoblasts.* Curr Eye Res, 2018. **43**(12): p. 1484-1488.
65. Seale, P., et al., *Muscle satellite cell-specific genes identified by genetic profiling of MyoD-deficient myogenic cell.* Dev Biol, 2004. **275**(2): p. 287-300.
66. Tamboli, R.A., et al., *Reduction in inflammatory gene expression in skeletal muscle from Roux-en-Y gastric bypass patients randomized to omentectomy.* PLoS One, 2011. **6**(12): p. e28577.
67. Kato, H., et al., *Effect of a 9-week exercise training regimen on expression of developmental genes related to growth-dependent fat expansion in juvenile rats.* Physiol Rep, 2018. **6**(19): p. e13880.

68. Zhang, E., et al., *A novel long noncoding RNA HOXC-AS3 mediates tumorigenesis of gastric cancer by binding to YBX1.* Genome Biol, 2018. **19**(1): p. 154.
69. Li, B., et al., *HOXC10 Regulates Osteogenesis of Mesenchymal Stromal Cells Through Interaction with Its Natural Antisense Transcript lncHOXC-AS3.* Stem Cells, 2019. **37**(2): p. 247-256.
70. Wang, X., et al., *HOXB13 promotes proliferation, migration, and invasion of glioblastoma through transcriptional upregulation of lncRNA HOXC-AS3.* J Cell Biochem, 2019. **120**(9): p. 15527-15537.
71. Wang, M., et al., *Methylation alterations of WT1 and homeobox genes in inflamed muscle biopsy samples from patients with untreated juvenile dermatomyositis suggest self-renewal capacity.* Arthritis and rheumatism, 2012. **64**(10): p. 3478-3485.
72. Greiwe, J.S., et al., *Resistance exercise decreases skeletal muscle tumor necrosis factor alpha in frail elderly humans.* Faseb J, 2001. **15**(2): p. 475-82.
73. Leger, B., et al., *Human sarcopenia reveals an increase in SOCS-3 and myostatin and a reduced efficiency of Akt phosphorylation.* Rejuvenation Res, 2008. **11**(1): p. 163-175B.

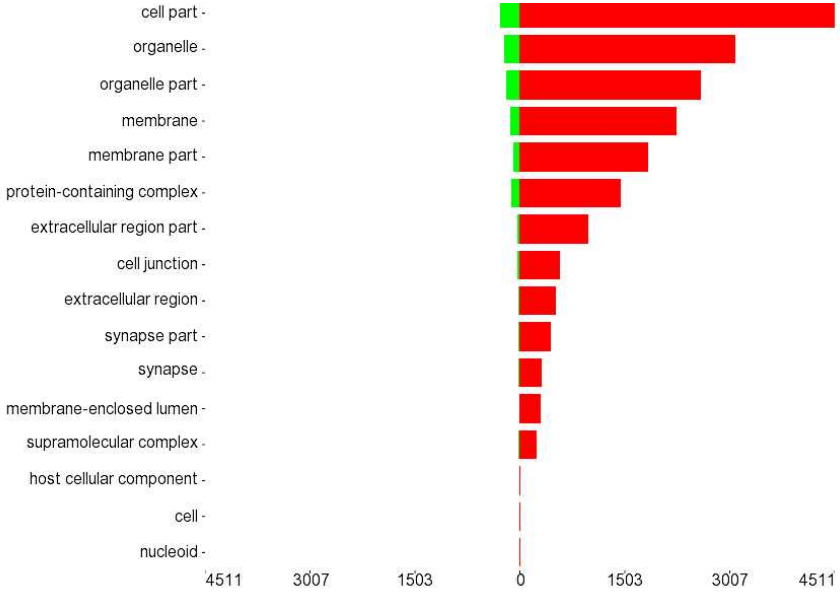
1a



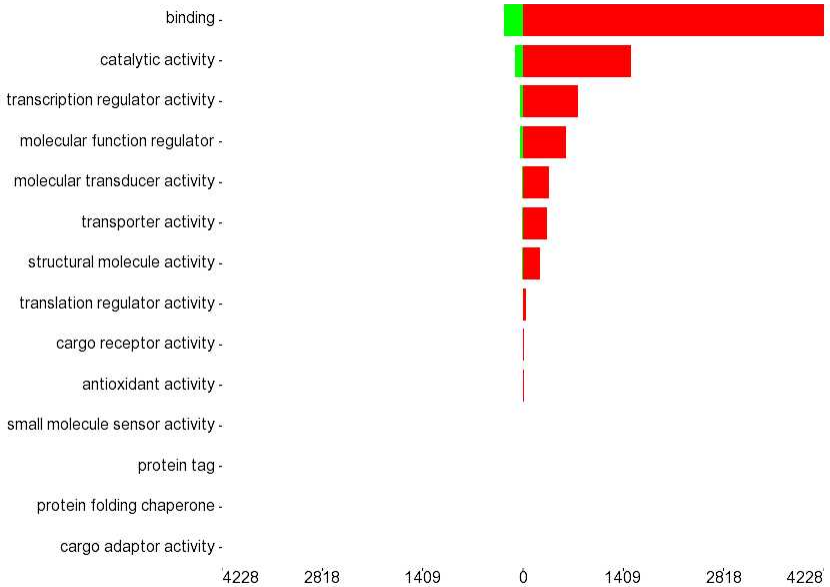
1b ‘Biological process’



1c ‘Cellular component’



1d ‘Molecular function’



1e Regulation of muscle system process

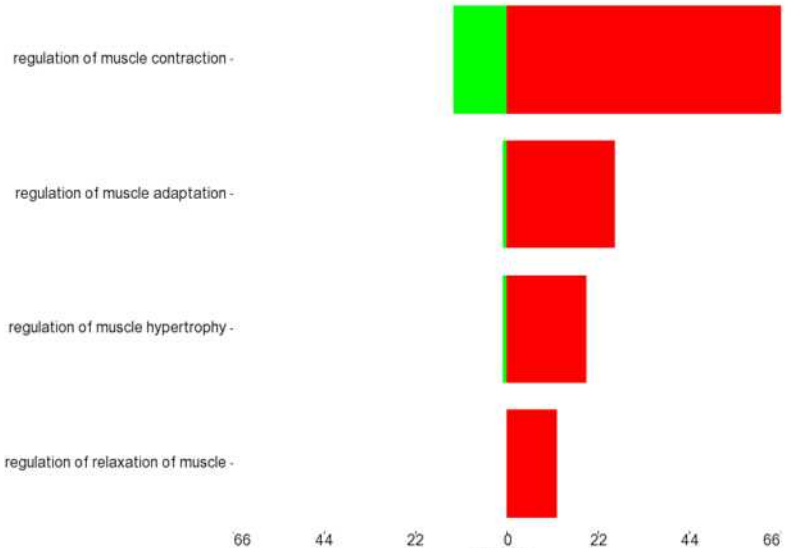
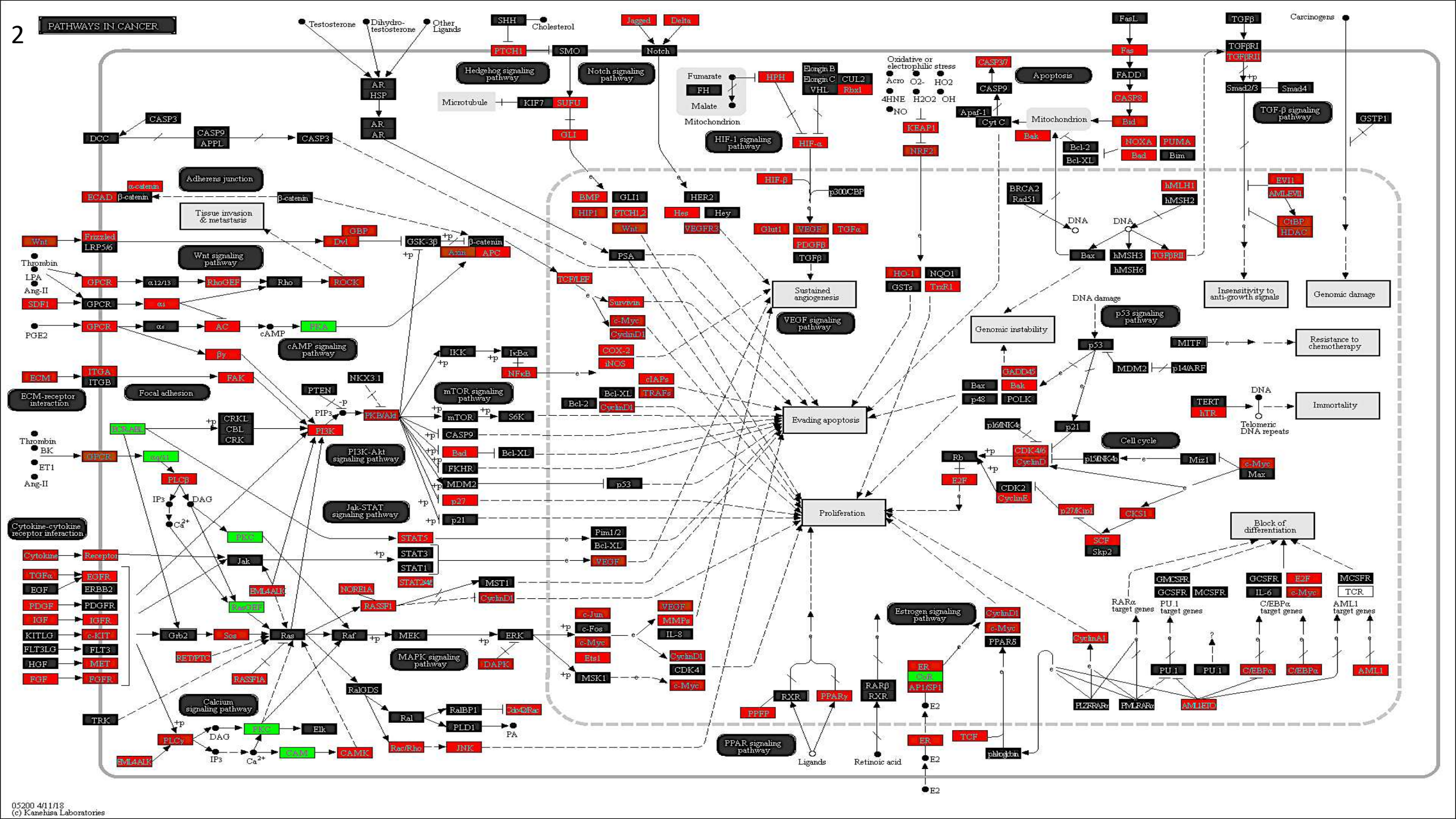


Figure 1. DNA methylation across the genome in aged human skeletal muscle tissue compared with young adults. 1a. Hierarchical clustering heatmap of the significantly differentially methylated CpG sites, suggesting that aged skeletal muscle tissue demonstrates predominantly hypermethylation (RED) compared with young adult's skeletal muscle (GREEN). **1b,c,d,e**. Significantly enriched CpG sites in gene ontology (GO) pathways for overarching GO terms: 'biological process' **(b), '**cellular component' **(c)** and 'molecular function' **(d) ,**and most significantly enriched GO term containing the search word 'muscle'; 'regulation of muscle system process' **(e).** All GO terms demonstrating a predominance of hypermethylation in aged compared with young human skeletal muscle. RED hypermethylated and GREEN hypomethylated CpG's.

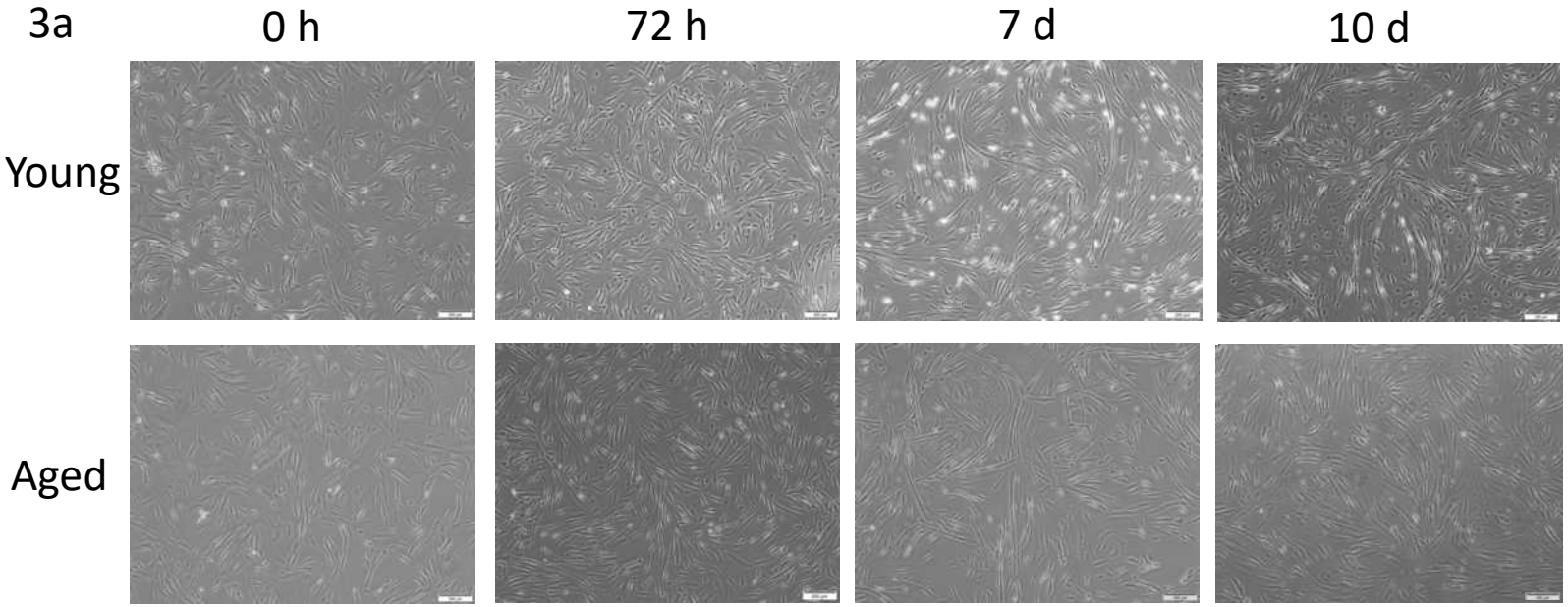
2



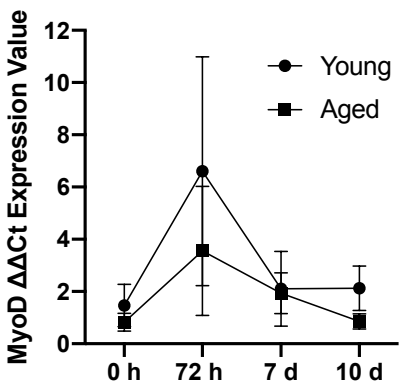
05200 4/11/18
(c) Kanehisa Laboratories

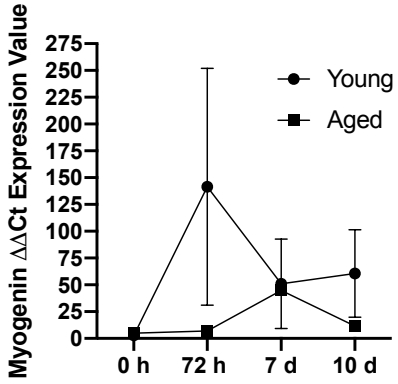
Figure 2. **The top significantly enriched KEGG pathway 'Pathways in Cancer' in aged compared with young adult human skeletal muscle tissue**. This pathway includes well known pathways in skeletal muscle regulation such as: Focal adhesion, MAPK signaling, PI3K-Akt-mTOR signaling, p53 signalling, Jak-STAT signaling, TGF-beta and Notch signalling. In this pathway 266 out of 277 CpG's were hypermethylated (RED) and only the remaining 11 hypomethylated (GREEN) in aged compared young adult muscle tissue. Note, because some genes have multiple CpGs per gene, Partek Pathway (Partek Genomics Suite) used to create this figure, selects the CpG with the lowest p-value (most significant) and uses the beta-value for that particular CpG to colour the pathway diagram. This is accurate therefore in situations where CpG's on the same gene have the same methylation pattern (e.g. all hyper or hypomethylated). However, where multiple CpG's on the same gene have the different methylation pattern (e.g. some hypomethylated and others hypermethylated), only the most significant methylated CpG is chosen and represented in the image. However, in this instance the majority of CpG's (even when there was more than one per gene) were hypermethylated, so in this case depicts an accurate representation of the data. Full CpG lists including the sites that are hypo and hypermethylated in this pathway are included in Suppl. File 1h.

3a


0 h
72 h
7 d
10 d
Young
Aged

3b


MyoD ΔΔCt Expression Value
12
10
8
6
4
2
0
0 h
72 h
7 d
10 d
Young
Aged


Myogenin ΔΔCt Expression Value
275
250
225
200
175
150
125
100
75
50
25
0
0 h
72 h
7 d
10 d
Young
Aged

3c

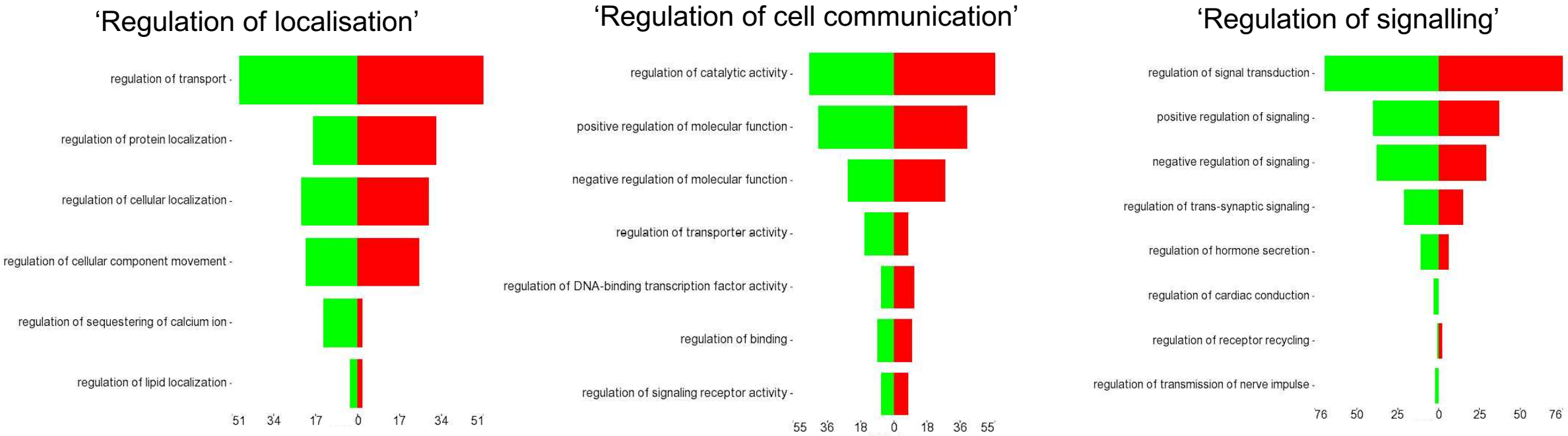
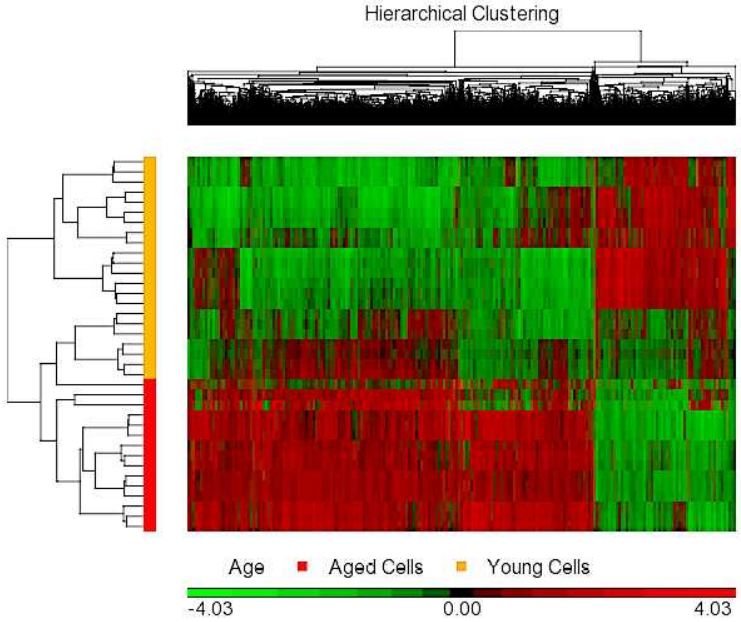

'Regulation of localisation'
regulation of transport
regulation of protein localization
regulation of cellular localization
regulation of cellular component movement
regulation of sequestering of calcium ion
regulation of lipid localization
51 34 17 0 17 34 51
'Regulation of cell communication'
regulation of catalytic activity
positive regulation of molecular function
negative regulation of molecular function
regulation of transporter activity
regulation of DNA-binding transcription factor activity
regulation of binding
regulation of signaling receptor activity
55 36 18 0 18 36 55
'Regulation of signalling'
regulation of signal transduction
positive regulation of signaling
negative regulation of signaling
regulation of trans-synaptic signaling
regulation of hormone secretion
regulation of cardiac conduction
regulation of receptor recycling
regulation of transmission of nerve impulse
76 50 25 0 25 50 76

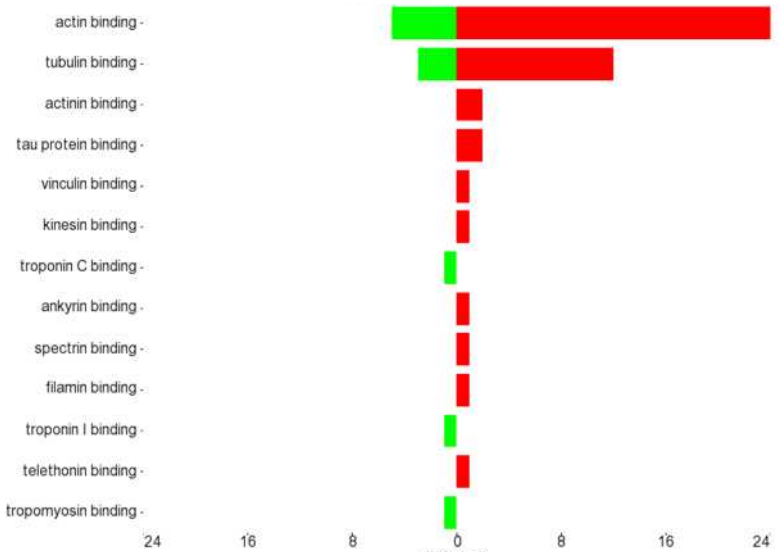
Figure 3. Aged and young adult primary muscle derived stem cells differentiated over 0, 72 h, 7d and 10 d. With corresponding myogenic gene expression and the gene ontology pathways that were enriched across the time-course of differentiation in aged cells. 3a. Light microscope images of aged versus young muscle derived stem cells morphologically depict fewer myotubes in aged versus young cells, particularly at 7 and 10 days. **3b**. MyoD and myogenin gene expression in aged versus young cells over the differentiation time course. Where, in particular, reductions in myogenin were observed at 72 h in aged cells as well as a delayed increase in the upregulation of myogenin compared with young cells. **1c,d,e.** Aged cells also demonstrated significantly altered DNA methylation as they differentiated, whereas young cells did not demonstrate a significant change in methylation as they differentiated. This altered regulation of DNA methylation over the time course of differentiation was most prominent at 7 days and was enriched in GO terms 'regulation of localisation' (**c**), 'regulation of cell communication' (**d**) and 'regulation of signalling' (**e**). RED hypermethylated and GREEN hypomethylated CpG's.

4a



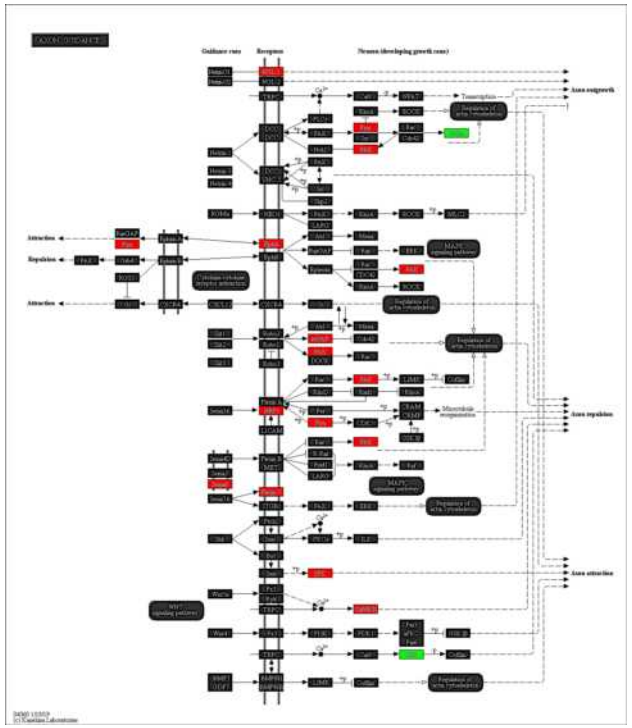
4b

'Cytoskeletal protein binding'
0 h aged vs. 0 h young adult cells



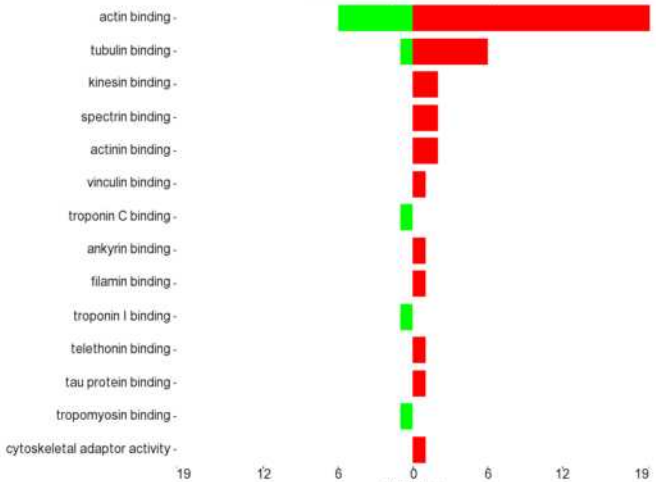
4c

'Axon guidance'
0 h aged vs. 0 h young adult cells



4d

'Cytoskeletal protein binding'
72 h aged vs. 72 h young adult cells



4e

'cAMP signaling'
72 h aged vs. 72 h young adult cells

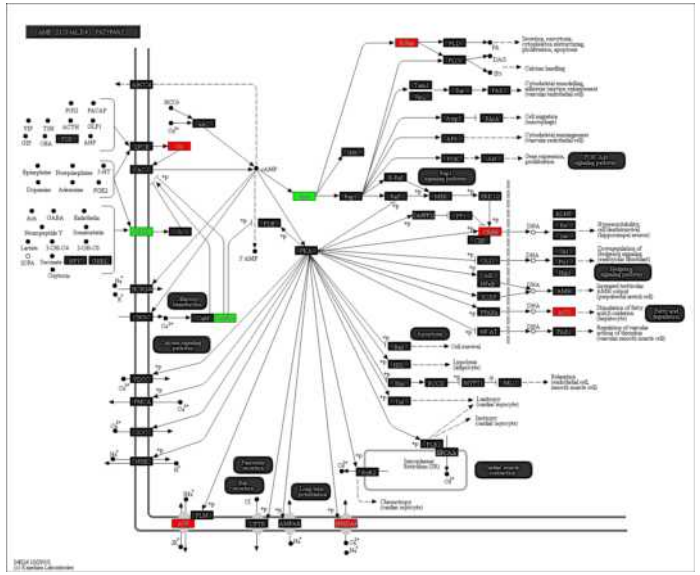
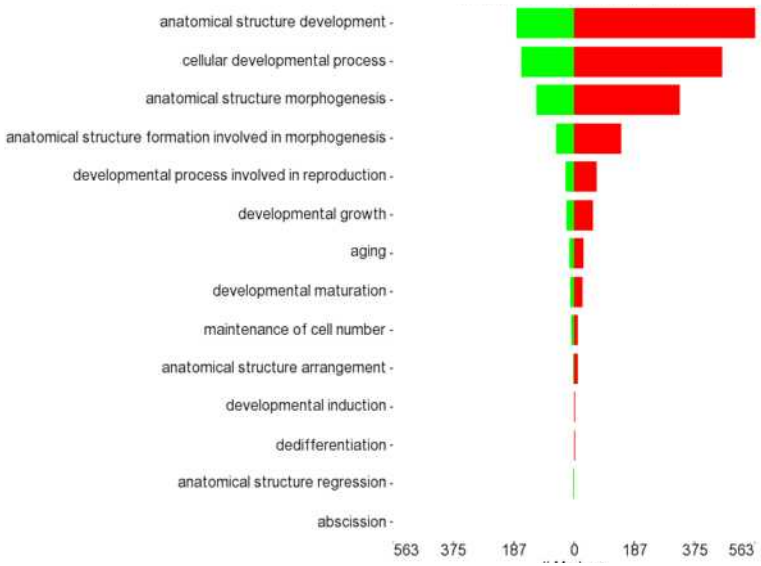


Figure 4. Comparison of DNA methylation in aged versus young muscle stem cells across differentiation. 4a. Hierarchical clustering heatmap of the significantly differentially methylated CpG sites between aged and young muscle stem cells across the entire time-course of differentiation (all time points of 0, 72 h, 7 d and 10 d). **4b.** 0 h aged vs. 0 h young adult muscle stem cells, top significantly enriched GO term 'cytoskeletal protein binding'. **4c.** 0 h aged vs. 0 h young adult, top KEGG pathway 'axon guidance'. **4d.** 72 h aged vs. 72 h young adult top GO term, 'cytoskeletal protein binding', **4e.** 72 h aged vs. 72 h young adult, a top 3 KEGG pathway 'cAMP signaling'. All demonstrating a predominance of hypermethylation in aged compared with young muscle stem cells in these GO terms and KEGG pathways. RED hypermethylated and GREEN hypomethylated CpG's. Note, because some genes have multiple CpGs per gene, Partek Pathway (Partek Genomics Suite) used to create this figure, selects the CpG with the lowest p-value (most significant) and uses the beta-value for that particular CpG to colour the pathway diagram. This is accurate therefore in situations where CpG's on the same gene have the same methylation pattern (e.g. all hyper or hypomethylated). However, where multiple CpG's on the same gene have the different methylation pattern (e.g. some hypomethylated and others hypermethylated), only the most significant methylated CpG is chosen and represented in the image. Full and accurate CpG lists including the sites that are hypo and hypermethylated in these pathways are included in Suppl. File **3n** (for figure 4c) and Suppl. File **4k** (for figure 4e).

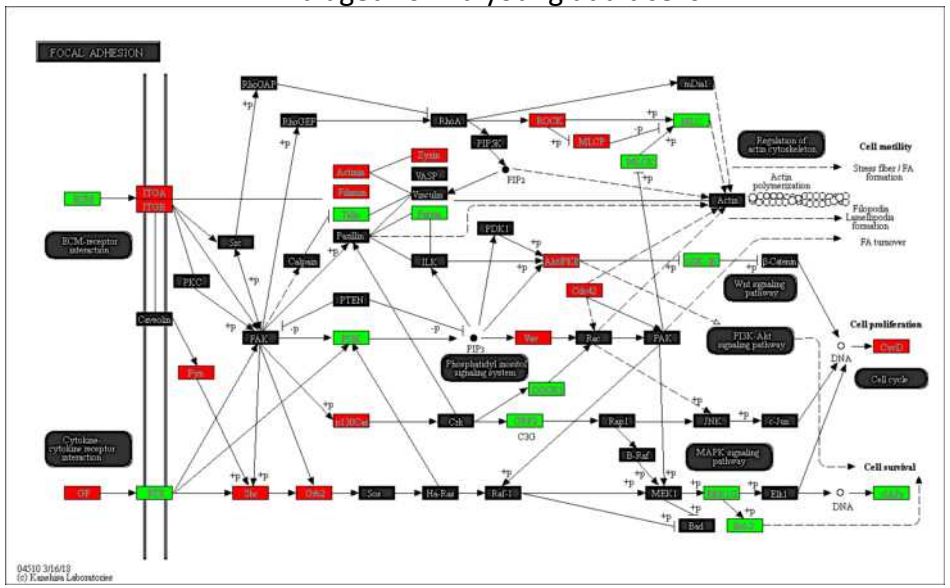
5a

'Developmental process' (incl. anatomical structure development') 7 d aged vs. 7 d young adult cells



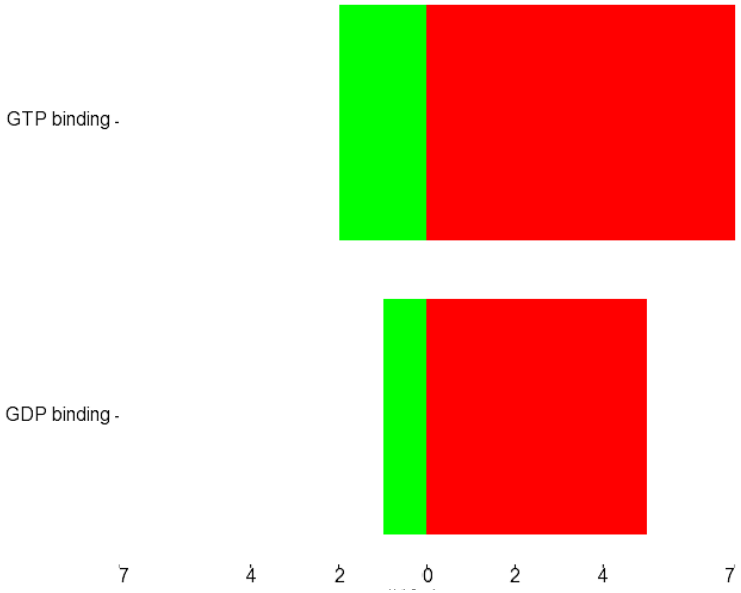
5b

'Focal adhesion'
7 d aged vs. 7 d young adult cells



5c

'GDP binding'
10 d aged vs. 10 d young adult cells



5d

'Regulation of the actin cytoskeleton'
10 d aged vs. 10 d young adult cells

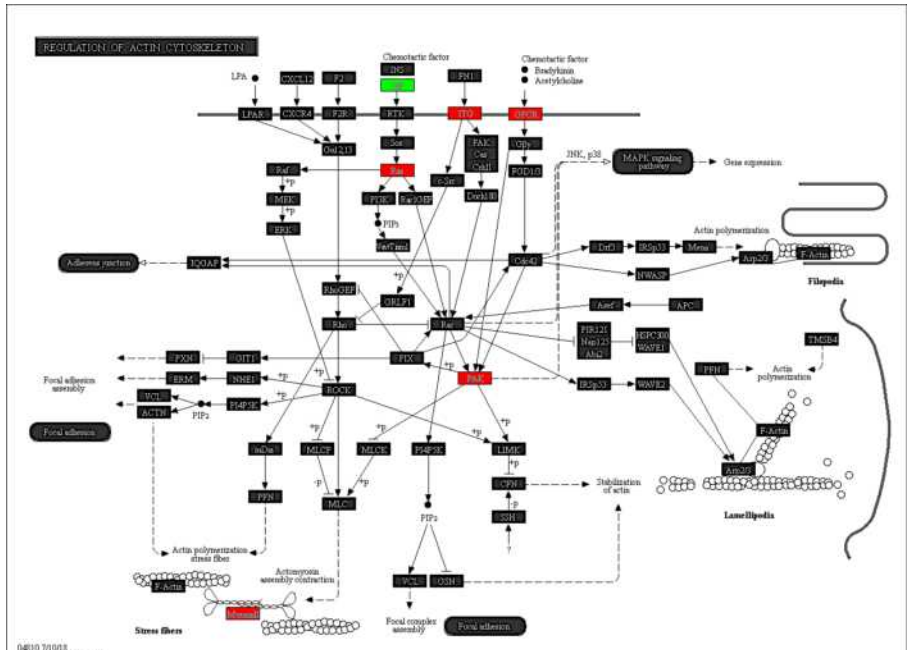
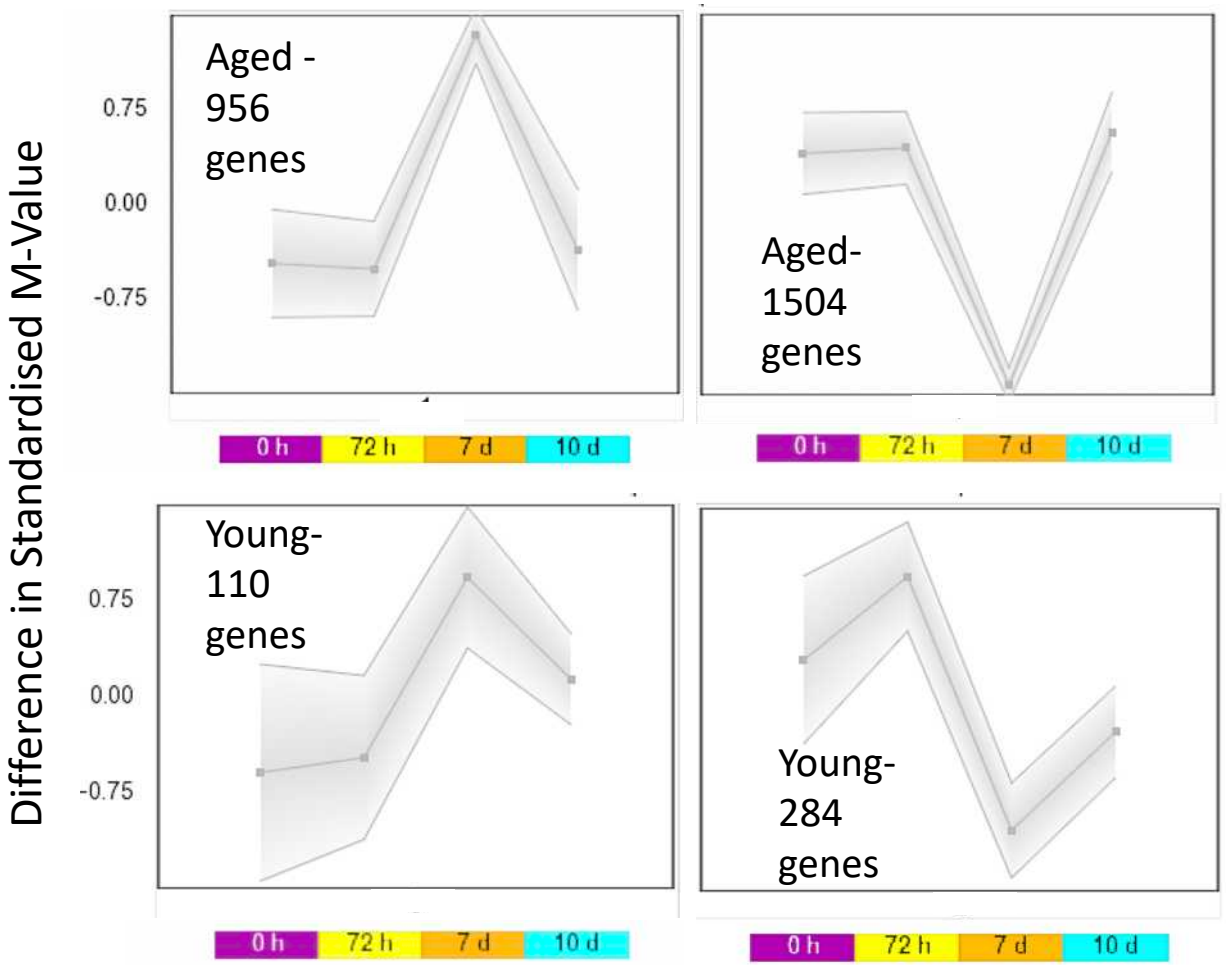
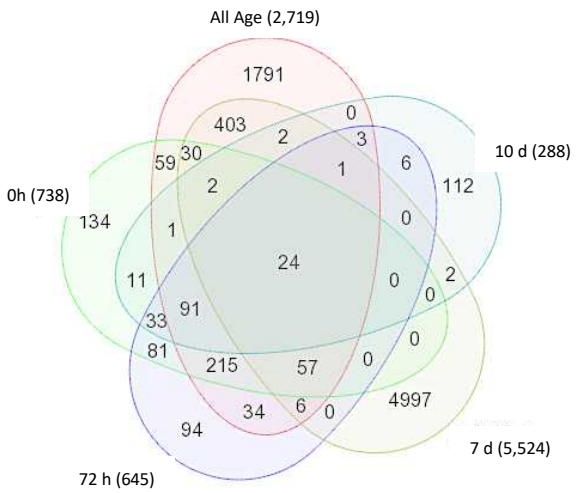


Figure 5. Comparison of DNA methylation in aged versus young muscle stem cells at 7 d and 10 d of differentiation. 5a. 7 d aged vs. 7 d young adult skeletal muscle stem cells, top significantly enriched GO term, 'anatomical structure morphogenesis'. **5b**. 7 d aged vs. 7 d young adult, top KEGG pathway, 'focal adhesion'. **5c**. 10 d aged vs 10 d young adult, top GO term, 'GDP binding'. **5d.** 10 d aged vs. 10 d young adult, top KEGG pathway, 'regulation of the actin cytoskeleton'. All demonstrating a predominance of hypermethylation in aged compared with young muscle stem cells in these GO terms and KEGG pathways. RED hypermethylated and GREEN hypomethylated CpG's. Note, because some genes have multiple CpGs per gene, Partek Pathway (Partek Genomics Suite) used to create this figure, selects the CpG with the lowest p-value (most significant) and uses the beta-value for that particular CpG to colour the pathway diagram. This is accurate therefore in situations where CpG's on the same gene have the same methylation pattern (e.g. all hyper or hypomethylated). However, where multiple CpG's on the same gene have the different methylation pattern (e.g. some hypomethylated and others hypermethylated), only the most significant methylated CpG is chosen and represented in the image. Full and accurate CpG lists including the sites that are hypo and hypermethylated in these pathways are included in Suppl. File **5i** (for figure 5b) and Suppl. File **6i** (for figure 5d).

6a


Difference in Standardised M-Value
0.75
0.00
-0.75
Aged - 956 genes
Aged- 1504 genes
Young- 110 genes
Young- 284 genes
0 h
72 h
7 d
10 d

6b


All Age (2,719)
10 d (288)
0h (738)
7 d (5,524)
72 h (645)
1791
403
112
134
24
91
215
57
4997
94
34
81
33
11
59
30
2
1
0
3
6
2
0
0
0
0
0
0
0
0
0
2
1
6

6c

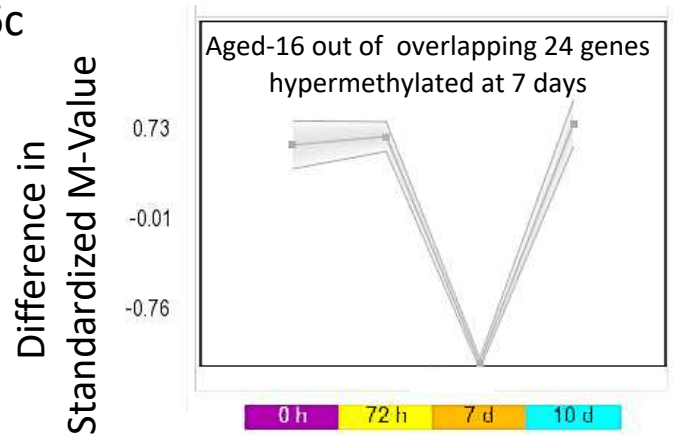
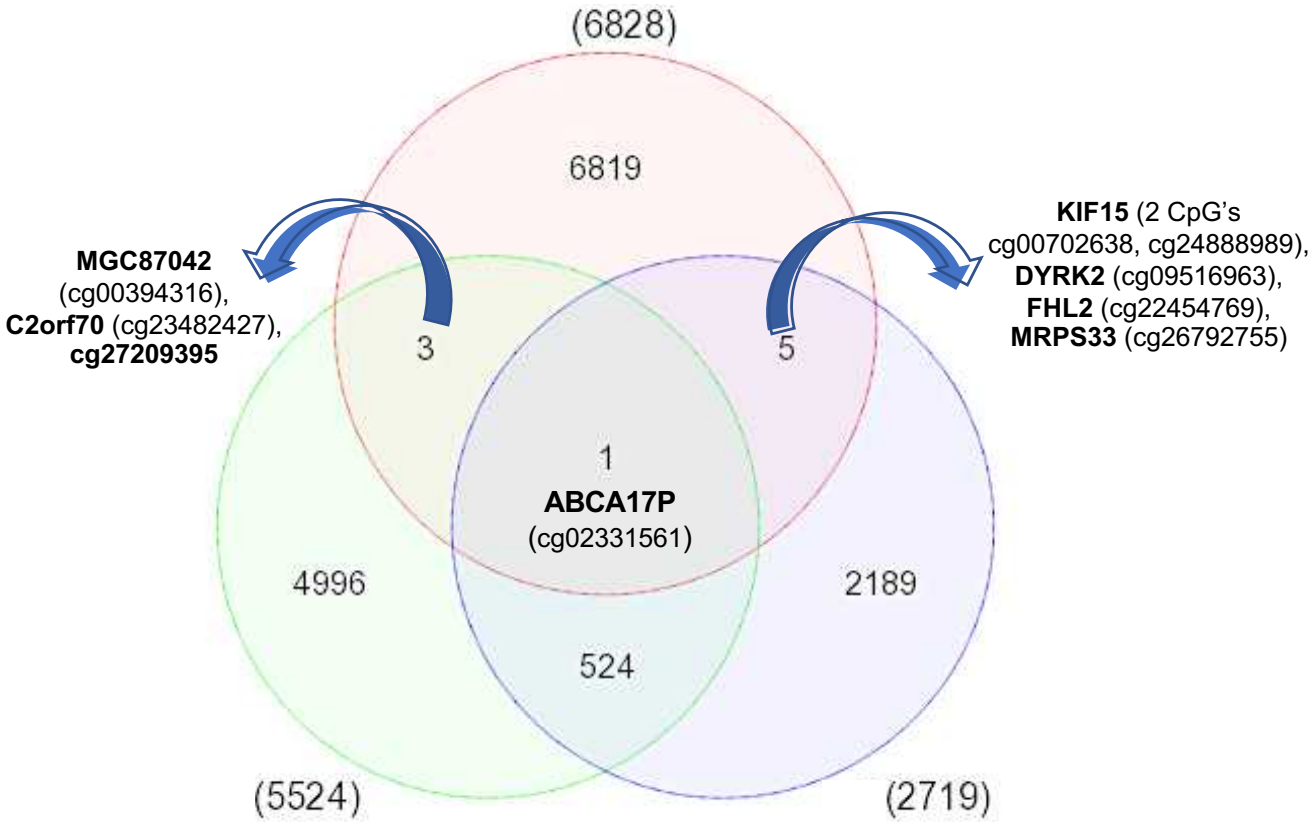

Difference in Standardized M-Value
Aged-16 out of overlapping 24 genes hypermethylated at 7 days
0.73
-0.01
-0.76
0 h
72 h
7 d
10 d

Figure 6. SOM profiling of DNA methylation over the time-course of differentiation in aged muscle stem cells. 6a. Demonstrates that there was a larger number of of hypomethylated and hypermethylated CpG sites in aged muscle stem cells, particularly at 7 days of differentiation, compared with young adult muscle stem cells. Confirming the earlier analysis in figure 3, where aged cells seem to have more varied methylation at 7 days compared with young adult cells. **6b.** Venn Diagram depicting the 24 common CpG sites that were altered at every time point of differentiation between aged and young adult muscle stem cells at each time point of differentiation (0, 72 h, 7d and 10 d). **6c**. As with the above analysis in 6a, SOM profiling identified that 16 out of the 24 CpG's also demonstrated the most varied methylation at 7 days of differentiation in aged cells.

7

Aged versus Young Tissue



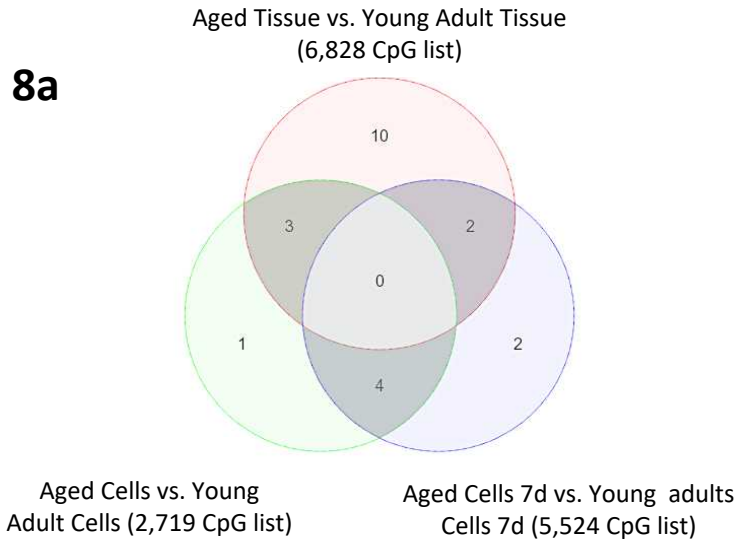
Aged Stem Cells 7 d versus Young Adult Cells 7 d

Aged Stem Cells versus Young Adult Cells (all time)

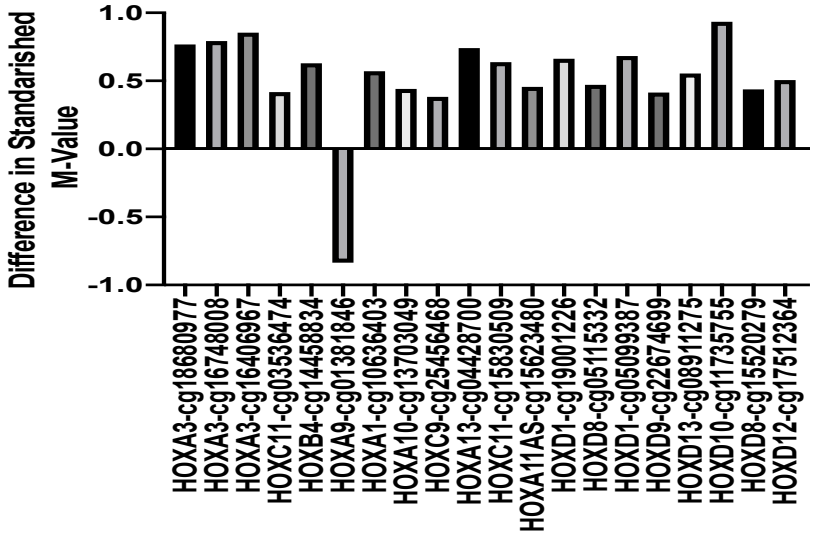
Figure 7. Venn diagram of common overlapping CpG sites significantly differentially methylated between aged skeletal muscle tissue and muscle stem cells. Venn diagram depicts commonly varied CpG sites in aged skeletal muscle tissue that were also varied in aged muscle stem cell analysis. Overlap of the tissue and aged cells included common 6 CpG sites on genes KIF15 (2 CpG's Cg00702638 & cg24888989), DYRK2 (cg09516963), FHL2 (cg22454769), MRPS33 (cg26792755) and ABCA17P (cg02331561). Given that aged cells demonstrated the most varied methylation at 7 days versus young cells. When overlapping the 7 d most significantly differentially methylated CpG lists, 4 CpG's (out of the 6 CpG sites identified above) were also identified, including: MGC87042 (cg00394316), C2orf70 (cg23482427), ABCA17P (cg02331561) and cg27209395 (not on an annotated gene). Once more, all of these CpG's (with the exception of C2orf70, cg23482427) were hypermethylated in the tissue analysis as well as the stem cells. With ABCA17P (cg02331561) highlighted across all CpG lists.

HOX Family Genes

8a



8b



8c

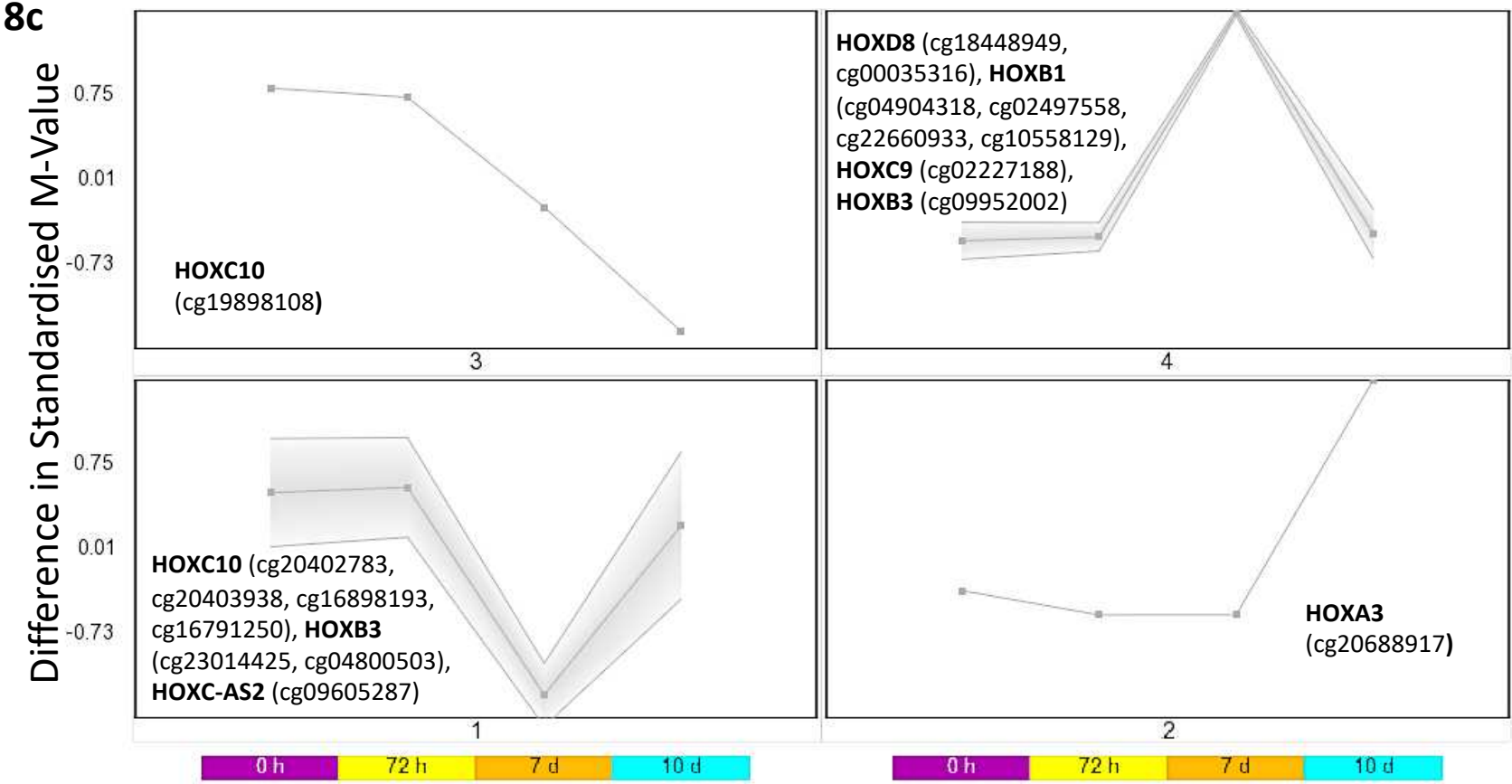


Figure 8. HOX family of genes and their DNA methylation in aged tissue and muscle stem cells. 8a. Venn diagram identifying 9 commonly differentially methylated HOX genes: including HOXD10, HOXD9, HOXD8, HOXA3, HOXC9, HOXB1, HOXB3, HOXC-AS2 and HOXC10 (note this Venn diagram analysis is by 'gene symbol' not 'probe cg' (CpG site), as some HOX genes also had more than 1 CpG per gene symbol, full Cg lists are located in Suppl. Figure 9 a,b,c,d). **8b.** All HOX family genes by CpG site (cg probe) differentially methylated in aged compared with young skeletal muscle tissue, predominantly all demonstrating hypermethylation. **8c.** SOM profiling depicting the temporal regulation of DNA methylation in aged muscle stem cells as they differentiate in CpG's located amongst the HOX family of genes (depicting the 9 HOX genes identified to be changed, by gene symbol, in both the tissue and cells in Figure 8a). The majority of these HOX CpG sites were differentially methylated at 7 days of differentiation in the aged cells.

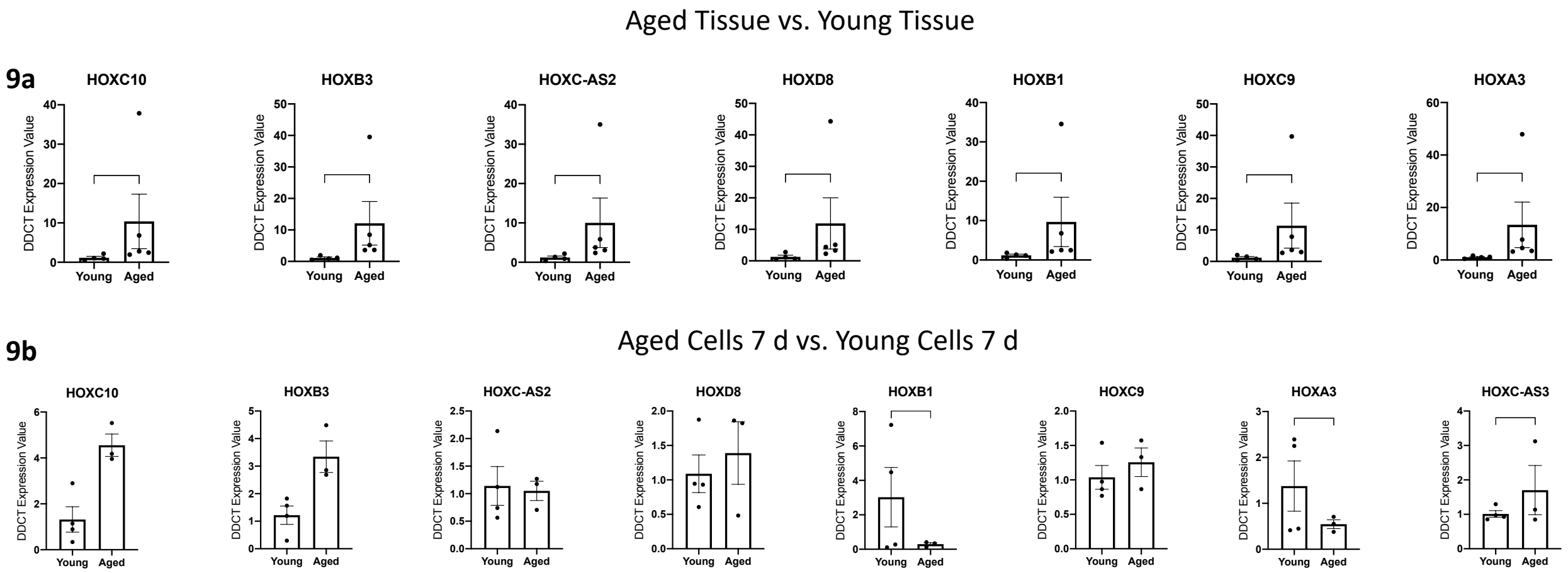


Figure 9. Gene expression of the HOX family of genes in aged compared to young adult skeletal muscle tissue (9a) and muscle stem cells at 7 days of differentiation (9b).