From the department of Physiology & Pharmacology
Karolinska Institutet, Stockholm, Sweden

THE SEARCH FOR HUMAN SKELETAL MUSCLE MEMORY

Exercise effects on the transcriptome and epigenome

Maléne Lindholm

Karolinska Institutet

Stockholm 2015

Published by Karolinska Institutet.
Printed by Eprint AB 2015
Cover illustration: *Elina Antilla*

ISBN 978-91-7676-060-4

THE SEARCH FOR HUMAN SKELETAL MUSCLE MEMORY

Exercise effects on the transcriptome and epigenome

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Maléne Lindholm

Principal Supervisor:
Professor Carl Johan Sundberg
Karolinska Institutet
Department of Physiology & Pharmacology
Division of Molecular Exercise Physiology

Co-supervisor(s):
Professor Tomas J Ekström
Karolinska Institutet
Department of Clinical Neuroscience
Center for Molecular Medicine

PhD. Helene Rundqvist
Karolinska Institutet
Department of Cell and Molecular Biology

Associate professor Helene Fischer
Karolinska Institutet
Department of Laboratory Medicine
Division of Clinical Physiology

Opponent:
Professor Frank W Booth
University of Missouri
Department of Medical Pharmacology and Physiology

Examination Board:
Professor Fawzi Kadi
Örebro University
School of Health and Medical Sciences

Professor Anna Krook
Karolinska Institutet
Department of Physiology and Pharmacology
Division of Integrative Physiology

Associate Professor Rickard Sandberg
Karolinska Institutet
Department of Cell and Molecular Biology

To my family

"If we could give every individual the right amount of nourishment and exercise, not too little and not too much, we would have found the safest way to health"

Hippocrates, ~ 450 BC

ABSTRACT

Regular physical activity is an environmental stimulus that is highly associated to many health benefits, while physical inactivity is detrimental for health and physical function. Regular exercise training is used in the prevention and treatment of a large number of disease conditions, including obesity, type II diabetes, cardiovascular disease and cancer, and reduces the risk for premature death. Most tissues adapt to exercise training, not least skeletal muscle tissue, which is highly plastic. The local adaptation of muscle is important not only for muscle function but also the health effects of training that affect the whole body. The cellular adaptations in skeletal muscle are driven by extra- and intracellular signals arising from the exercise stimulus, for example changes in shear stress, oxygen tension, energy levels, pH and temperature. Ultimately, these cellular perturbations lead to gene expression and protein alterations that improve muscle function. Thus, it is important from a clinical, as well as basic science perspective to understand the regulation of skeletal muscle gene activity and how activity changes contribute to the many health benefits of a physically active lifestyle.

The understanding of training-induced changes in gene expression and the underlying mechanisms have progressed extensively over the past 20 years. Still, many key mechanisms remain to be investigated. The overall purpose of this thesis was to investigate the influence of epigenetic mechanisms, *i.e.* DNA methylation and post-translational modifications of histones, on endurance training adaptation. Epigenetic mechanisms are important for cellular memory. Thus, another objective was to investigate if there were any residual intrinsic memory effects of previous endurance training, and if that could induce different responses to a repeated training period after detraining.

The results in this thesis are based on skeletal muscle biopsies from the *vastus lateralis*, taken before and after three months, or six weeks, of endurance training, or at rest in elite athletes and sedentary individuals. In the first study, the baseline skeletal muscle transcriptome was investigated. Studies using repeated skeletal muscle sampling regularly assume that potential changes are due to the intervention and not inherent variability between samples. The results showed, using global RNA sequencing analysis, that tissue homogeneity was remarkably high within a muscle and in the corresponding muscle of the contralateral leg of an individual, while the transcriptome difference between male and female skeletal muscle was substantial. This study also found 23 000 isoforms expressed in skeletal muscle at baseline, together with almost 2500 previously unannotated, novel transcripts, out of which at least five were protein-coding.

The transcriptome changes induced by three months of one-legged knee extension training were very significant. Over 3000 isoforms were found to be differentially expressed, as well as 34 of the novel transcripts discovered at baseline. The one-legged training regime meant that the other leg was included as an intraindivdual control leg, which was exposed to the same other environmental factors such as diet, stress, sleep etc. We found that the training

response of the trained leg was very specific, although significant but markedly smaller changes occurred also in the untrained leg. At the protein level, a specific investigation of HIF (hypoxia inducible factor) was performed. HIF is activated by acute exercise, but was hypothesized to be attenuated by long-term training due to its inhibitory effect on mitochondrial energy production. A comparison of skeletal muscle from elite athletes with normally active individuals, showed that the negative regulators of HIF were higher in the elite athletes, indicating a reduced HIF activity in that group. This was supported by similar findings in a six-week bicycle training study.

Three months of endurance training induced changes in DNA methylation at almost 5000 specific sites across the human skeletal muscle genome that were associated to functionally relevant transcriptional changes. Many of these changes occurred in regulatory enhancer regions and the differentially methylated sites were associated to transcription factor binding sites for myogenic regulatory factors (increases in methylation) and the ETS family (decreases in methylation). Six weeks of bicycle training showed a strong trend towards a global downregulation of trimethylation of histone H3, lysine 27, previously described as a dynamic and predominantly inhibitory modification. The specific genes potentially affected by this histone modification in response to training are currently being analyzed using chromatin immunoprecipitation followed by sequencing.

After the initial three months of one-legged endurance training, a subset of the subjects came back after nine months of detraining and performed a second three-month training period. This time, they trained both legs in the exact same way as one leg was trained in the first period. One leg had thus been previously well-trained, while the other was previously untrained. Potential residual effects were investigated by comparing biopsies obtained from both legs before starting the second training period. At the transcriptome level, there were no indications of remaining effects, although the exertion perceived in the first training session of period 2 was lower in the previously trained leg. Repeated training induced similar changes physiologically and at the global transcriptome level between the two legs. There were specific differences in the gene activity changes between the legs, but with the current approach, we found no overall significant differences in the response to a repeated training period.

Collectively, the results in this thesis show that endurance exercise training induced associated changes in the epigenome and transcriptome of human skeletal muscle. The data included an in-depth analysis of the human skeletal muscle transcriptome at baseline and how it changes in response to repeated endurance training periods, with no detectable muscle memory of previous training at the transcriptome level. The results contribute to a better understanding of the molecular pathways involved in physiological adaptation to endurance training and can potentially be used to describe how training prevents disease development and different dysfunctions.

LIST OF SCIENTIFIC PAPERS

This thesis is based on the following papers, referred to in the text by their Roman numerals:

I. **Lindholm ME**, Huss M, Solnestam BW, Kjellqvist S, Lundeberg J, Sundberg CJ. *The human skeletal muscle transcriptome: sex differences, alternative splicing, and tissue homogeneity assessed with RNA sequencing.* FASEB J. 2014 Oct;28(10):4571-81

II. **Lindholm ME***, Marabita F*, Gomez-Cabrero D, Rundqvist H, Ekström TJ, Tegnér J, Sundberg CJ. *An integrative analysis reveals coordinated reprogramming of the epigenome and the transcriptome in human skeletal muscle after training.* Epigenetics. 2014 Dec 2;9(12):1557-69.

III. **Lindholm ME***, Saini A*, Fischer H, Öst A, Alenius M, Castro J, Auer G, Ekström TJ, Rundqvist H, Sundberg CJ. *Effects of endurance training on histone 3 lysine 27 trimethylation in human skeletal muscle.* In manuscript.

IV. **Lindholm ME***, Giacomello S*, Werne Solnestam B, Fischer H, Kjellqvist S, Huss M, Sundberg CJ. *The impact of endurance training on human skeletal muscle memory, global isoform expression and novel transcripts.* In manuscript.

V. **Lindholm ME**, Fischer H, Poellinger L, Johnson RS, Gustafsson T, Sundberg CJ, Rundqvist H. *Negative regulation of HIF in skeletal muscle of elite endurance athletes: a tentative mechanism promoting oxidative metabolism.* Am J Physiol Regul Integr Comp Physiol. 2014 Aug 1;307(3):R248-55**.**

Publications by the author, which are not included in the thesis:

Lindholm ME, Rundqvist H. *Skeletal muscle HIF-1 and exercise.* Exp. Physiol. 2015. Epub ahead of print.

Lindskog C, Linné J, Fagerberg L, Hallström BM, Sundberg CJ, **Lindholm M**, Huss M, Kampf C, Choi H, Liem DA, Ping P, Väremo L, Mardinoglu A, Nielsen J, Larsson E, Pontén F, Uhlén M. *The human cardiac and skeletal muscle proteomes defined by transcriptomics and antibody-based profiling.* BMC Genomics 2015 Jun 25;16:475.

Marabita F, Almgren M, **Lindholm ME**, Ruhrmann S, Fagerström-Billai F, Jagodic M, Sundberg CJ, Ekström TJ, Teschendorff AE, Tegnér J, Gomez-Cabrero D. *An evaluation of analysis pipelines for DNA methylation profiling using the Illumina HumanMethylation450 BeadChip platform.* Epigenetics 2013 Mar;8(3):333-46.

CONTENTS

LIST OF ABBREVIATIONS

ADP	Adenosine diphosphate
AMPK	AMP-activated kinase
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
β-HAD	3-hydroxyacyl-CoA dehydrogenase
BMIQ	Beta mixture quantile dilation
BSA	Bovine serum albumin
CaMKII	Ca^{2+}-calmodulin-dependent protein kinase II
cDNA	Complementary DNA
ChIP	Chromatin Immunoprecipitation
COX	Cytochrome *c* oxidase
CpG	Cytosine-phosphate-Guanine
CPM	Counts per million
CREB	cAMP response element-binding protein
CS	Citrate synthase
DAVID	Database for annotation, visualization and integrated discovery
DEG	Differentially expressed gene
DMOG	Dimethyloxallyl glycine
DMP	Differentially methylated position
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
ECL	Enhanced chemiluminescence
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EPO	Erythropoietin
eRNA	Enhancer RNA
FDR	False discovery rate
FIH	Factor inhibiting HIF

FPKM	Fragments per kilobase of exon per million mapped fragments
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GLUT-4	Glucose transporter 4
GO	Gene ontology
HDAC	Histone deacetylase
HIF-1α	Hypoxia inducible factor 1α
HSMM	Human skeletal muscle myoblast
LDH	Lactate dehydrogenase
lncRNA	Long non-coding RNA
LUMA	Luminometric methylation assay
MAPK	Mitogen-activated protein kinase
5-mC	5-methylcytosine
5-hmC	5-hydroxymethylcytosine
miRNA	Micro RNA
MDH	Malate dehydrogenase
MDS	Multidimensional scaling
MEF	Myocyte enhancer factor
MRF	Myogenic regulatory factor
mRNA	Messenger RNA
NAD^+	Nicotinamide adenine dinucleotide, oxidized
NADH	Nicotinamide adenine dinucleotide, reduced
NRF-1	Nuclear respiratory factor 1
OPLS	Orthogonal projections to latent structures by means of partial least squares
OPLS-DA	OPLS-Discriminant analysis
ORF	Open reading frame
PCA	Principal component analysis
PDHc	Pyruvate dehydrogenase complex
PDK-1	Pyruvate dehydrogenase kinase 1
PGC1α	PPAR gamma coactivator 1 alpha

PHD	Prolyl hydroxylase
PPAR	Peroxisome proliferator-activated receptor
PVDF	Polyvinylidene difluoride
qPCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
ROS	Reactive oxygen species
RPE	Rate of perceived exertion
rRNA	Ribosomal RNA
SDS	Sodium Dodecyl Sulfate
siRNA	Small interfering RNA
SIRT6	Sirtuin 6
snoRNA	Small nucleolar RNA
SNP	Single nucleotide polymorphism
snRNA	Small nuclear RNA
TET	Ten-eleven-translocation
tRNA	Transfer RNA
VEGF	Vascular endothelial growth factor
VO_2-peak	Peak oxygen uptake

1 BACKGROUND

The genome is the foundation of all life. In humans, it works as a unique hardware that is set when the egg and the sperm fuse together. The genome holds an immense amount of complex information that, in interaction with environmental cues, governs development, growth, function and disease (Lander et al., 2001; Venter et al., 2001). The genomic sequence is virtually the same in all of our cells. Despite this, there is a vast diversity among cells in the human body. Differences in gene activation determine cellular structure and function and also allow cells to specifically respond to different stimuli from the environment. One of our most plastic tissues is skeletal muscle, which readily adapts to changes in functional demands. Physical activity is an environmental stimulus that is associated to many health benefits, while physical inactivity is detrimental for health and skeletal muscle function (Booth et al., 2012). Regular physical activity is now considered fundamental for prevention, management and treatment of many of our most common chronic diseases (Haskell et al., 2007; Neufer et al., 2015). Much of the effect of training on muscle is made possible through regulation of gene activity. It is important from a clinical, as well as basic science perspective to understand the activity of the genome in skeletal muscle and how the activity is altered to induce the many health benefits of a physically active lifestyle.

1.1 THE GENOME AND EPIGENOME

1.1.1 Genes and their splice variants

The human haploid genome comprises an astonishing 3 billion bases (International Human Genome Sequencing, 2004). The estimated number of protein-coding genes within the genome has decreased from early estimates of 25 000 and 30 000 at the end of the Human Genome Project (Lander et al., 2001; Venter et al., 2001) to approximately 20 000 today (Clamp et al., 2007). The genes are transcribed into RNAs, *e.g.* messenger RNAs (mRNAs). The mRNAs are subsequently translated into proteins, a process known as the central dogma of molecular biology (Crick, 1970). Transcription is initiated at the promoter region of a gene through interaction between regulatory transcription factors and RNA Polymerase II. In most cases, genes are first transcribed into pre-mRNA and subsequently modified into mRNA that is exported to the cytoplasm for translation. However, the number of functional proteins in the human body is higher than the number of genes (Kim et al., 2014). One mechanism that greatly increases the complexity of the transcriptome is alternative splicing, that allows for different mRNA combinations from exons (coding parts of the gene), effectively producing several mRNA isoforms from the same gene. Approximately 95% of multiexon genes are believed to undergo alternative splicing (Pan et al., 2008), which can occur through several different mechanisms and often is tissue-specific (Wang et al., 2008). Transcription can also start from different promoters for the same gene (Ayoubi and Van De Ven, 1996) and include alternative 3' polyadenylation sites (addition of multiple adenine nucleotides) resulting in additional isoforms (Proudfoot et al., 2002). Through regulation of gene transcription and translation, a cell can thus adjust its functional proteome in accordance to its needs.

1.1.2 The non-coding genome

The protein-coding part of the genome is only about 1.2% (Consortium, 2012). The rest of the DNA was previously considered as "junk DNA" (Pennisi, 2012), but in recent years the understanding of the role and function of the non-coding part of the genome has increased enormously. Over 80% is now considered to have some regulatory role (Consortium, 2012) and up to 75% can be transcribed (Djebali et al., 2012). Non-coding RNAs include highly abundant transfer RNA (tRNA) and ribosomal RNA (rRNA), both enriched in the cytoplasm of cells and involved in translation of mRNA to protein (Djebali et al., 2012). Also present in the cytoplasm are microRNAs (miRNAs), very small (~22 nt) inhibitors of translation (Bartel, 2004). Regulatory RNAs in the nucleus include long non-coding RNAs (lncRNAs) that are less abundant but more cell-type specific than mRNAs, small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) (Djebali et al., 2012). Enhancers are short regulatory regions of DNA that have been known to bind transcription factors and influence gene activity for some time (Khoury and Gruss, 1983). Recently, enhancers have also been shown to be transcribed into RNA (eRNAs), their level of transcription correlating to that of nearby genes (De Santa et al., 2010; Kim et al., 2010). All of these different RNA transcripts together constitute the transcriptome, which is continuously proving to be more and more complex. The function of the vast majority of the transcriptome, including transcribed genes, is seemingly auto-regulation.

1.1.3 Epigenetics and chromatin

For transcription to occur, accessibility to the DNA is crucial. The DNA double helix is wrapped around octamers of histone proteins, called nucleosomes. DNA together with histones is referred to as chromatin, which in humans is intricately organized into 2x23 chromosomes (22 different autosomes plus the two different sex chromosomes X and Y) (Allis et al., 2007). A major mechanism regulating the chromatin structure, and thus accessibility to the DNA, is epigenetics, which literally means above our genes. The term was first used by Conrad Waddington in 1942 to describe the relationship between genotypes and phenotypes (Waddington, 1942). In 1996, it was defined by Riggs and colleagues as "the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence", while a more modern definition was proposed by Bird in 2007; "the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states" (Bird, 2007). The definition by Bird removes the concept of inheritance, which is an important development as epigenetic changes occur also in non-proliferating cells. The two most common epigenetic modifications are methylation of DNA and post-translational modifications of histone proteins.

1.1.3.1 DNA methylation

DNA methylation is an essential determinant of genomic stability and a key mechanism in the development and differentiation of cells (Jaenisch and Bird, 2003). In humans, methylation changes are associated to the normal aging process (Issa, 2014) as well as numerous different diseases, including many types of cancer, different imprinting disorders

(Robertson, 2005), rheumatoid arthritis (Liu et al., 2013), metabolic disease (Kirchner et al., 2013), cardiovascular disease (Kim et al., 2013b) and Alzheimer's (Bennett et al., 2015).

DNA methylation is the process where DNA methyltransferases (DNMTs) add a methyl group (CH_3) to position 5 of a cytosine (C) base in the genome, producing 5-methylcytosine or 5-mC. It most commonly occurs at C's that are 5' of a guanine (G) base, referred to as a CpG (Cytosine-phosphate-Guanine) sequence (Jaenisch and Bird, 2003), although it can occur also at non-CpG sites (Ramsahoye et al., 2000). DNA methylation regulates transcription through direct inhibition of transcription factor binding to the DNA or through methyl-binding proteins (*e.g.* MeCP2 and MDB1-4) that specifically bind to methylated DNA and for example recruit histone-modifying enzymes. DNA methylation in promoters is associated to repression of gene transcription, while in gene bodies the association appears to be the opposite (Ball et al., 2009) although the effect on transcription is not fully understood (Jjingo et al., 2012; Kulis et al., 2012). In human somatic cells, 70-80% of all CpG sequences are methylated and 5-mC accounts for approximately 1% of all bases in our DNA (Bird, 2002). The CpG dinucleotide sequence is under-represented in the genome due to deamination that may occur for 5-mC, transforming it to thymine (T), and the average GC content of the whole genome is about 41%. However, there are regions of the DNA with higher GC content and CpG frequency called CpG islands, commonly unmethylated and located in the promoters of genes (Lander et al., 2001). A methylated CpG island correlates with transcriptional repression (Bird, 2002). Hypermethylation of CpG islands is known to occur in cancer, which is also associated with a global hypomethylation (Feinberg and Vogelstein, 1983; Robertson, 2005). While promoters and CpG islands were in focus for many years, more distant regions have caught increasing interest. In 2009, Irizarry et al. showed that methylation changes in so called CpG island shores (±2kb from a CpG island) occurred more frequently in colon cancer compared to changes in CpG islands, and were also highly correlated to gene expression changes (Irizarry et al., 2009). The terminology has been expanded to include CpG island shelves (2-4kb from the CpG island) (Bibikova et al., 2011) and "open sea", >4kb from a CpG island (Sandoval et al., 2011).

Another modification that has been known for years, but recently has received a lot of attention, is 5-hydroxymethylcytosine (5-hmC). A family of enzymes known as TET (Ten-Eleven-Translocation) converts 5-mC into 5-hmC, which is a step towards demethylation (Dahl et al., 2011). Interestingly, 5-hmC is associated with transcriptional activation and the levels appear to be more tissue specific than for 5-mC (Munzel et al., 2011).

1.1.3.2 Histone modifications

Histone proteins are very small, and positively charged in order to attract the negatively charged DNA. Each nucleosome is made up of an octamer of histone proteins, consisting of two molecules each of histone H2A, H2B, H3 and H4 (although variants exist), with 147 base pairs of DNA wrapped around it. Each histone has a tail, protruding from the core region, that can be post-translationally modified to regulate how compact the chromatin is (Kouzarides, 2007). Changes in structure and/or electrostatic forces alter the contact between DNA and histones and thereby the accessibility for transcription factor binding and transcription. Histone tails can be modified in many different ways. Lysine acetylation results in a more open chromatin formation as it opposes the positive charge of lysine, while deacetylation has the opposite effect. Other modifications include for example methylation (which can occur as mono- di- or trimethylation), phosphorylation, ubiquitination and SUMOylation, and depending on position in the tail, the effect on transcription differs. The complexity of the histone modifications has led to the concept of a histone code (Allis et al., 2007). Similarly to DNA methylation, histone modifications are also associated to many critical cellular functions, including cell differentiation, and diseases such as cancer, mental disorders and rheumatoid arthritis (for review see (Hirst and Marra, 2009; Lennartsson and Ekwall, 2009). Figure 1 illustrates the organization of DNA into chromatin.

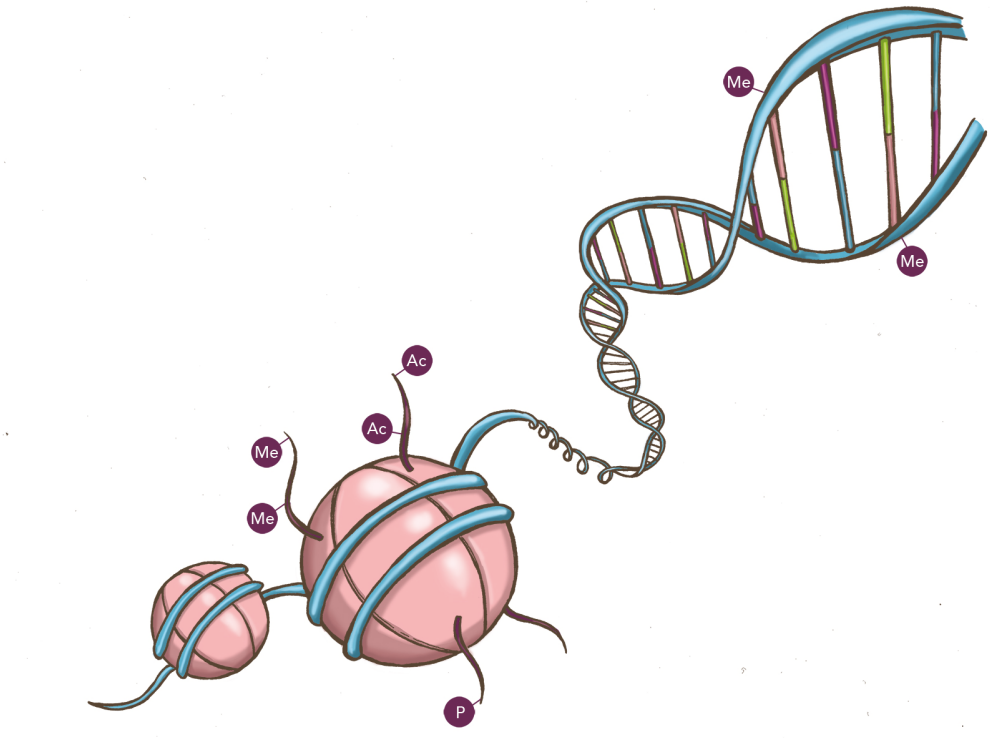

Figure 1. Basic chromatin structure. The ball-shaped structures illustrate the nucleosomes, with the DNA double helix wrapped around octamers of histone proteins. Post-translational modifications (Ac for acetylation, Me for methylation and P for phosphorylation) mark the protruding histone tails. Methyl marks are also present on cytosines of the DNA. *Illustration: Elina Anttila.*

1.2 SKELETAL MUSCLE BIOLOGY

Skeletal muscles execute all voluntary movements of our body, from locomotion to speech, breathing and facial expressions. Signals from the nervous system activate skeletal muscle fibers to contract and create mechanical force. Depending on age, gender and size, skeletal muscle normally makes up 40-45% of the body weight, by far the most abundant tissue in healthy humans.

1.2.1 Skeletal muscle structure and function

Human skeletal muscle fibers are large, multinucleated cells that are highly organized into bundles. In humans there are three different types of fibers; type I, type IIA and type IIX. The type I fibers are slow-twitch oxidative fibers that can produce relatively low amount of force over a long time. Type IIA and IIX are fast-twitch, more glycolytic fibers that have higher force production but lower endurance. With increasing contractile force, type IIA are recruited first and type IIX last. For a recent description of the different fiber type characteristics, see (Egan and Zierath, 2013). Each fiber consists of multiple myofibrils, where the smallest contractile unit is the sarcomere, containing the two main myofilaments; actin and myosin. These ultimately produce the mechanical force. Each myofibril is surrounded by sarcoplasmic reticulum that stores Ca^{2+} for release upon activation of the muscle. Ca^{2+} is the key that unlocks the contractile proteins and thus initiates the contractile process. Going from rest to full activity can increase energy demand >100-fold in a muscle. The main energy molecule of human cells is ATP (adenosine triphosphate). During short-lasting and intense exercise, the first source of energy for contraction comes directly from ATP that is hydrolyzed. Concurrently, chemical energy from creatine phosphate is transferred to form new ATP molecules. For work performed up to 90 seconds, anaerobic lactate formation provides most of the ATP needed. Over longer work periods, ATP is primarily produced by mitochondria, which is highly abundant in skeletal muscle fibers, especially Type I. Energy for virtually all chemical reactions in the body is derived from the breakdown of ATP into ADP (adenosine diphosphate) and free phosphate. To produce ATP through aerobic metabolism, the muscle needs oxygen and nutrient (glucose and fatty acids) supply via a network of capillaries that are interspersed in the fiber bundles (Windhorst and Mommaerts, 1996).

1.3 SKELETAL MUSCLE ADAPTATION TO ENDURANCE TRAINING

Physical activity is a highly potent environmental stimulus that induces adaptation in almost the entire human body. Regular exercise training improves and prevents a large number of disease conditions and reduces the risk for premature death (Neufer et al., 2015). Endurance training positively influences for example the cardiovascular system (Whyte and Laughlin, 2010), the endocrine system (Henriksson, 1995), the nervous system (van Praag, 2009), and even promotes a better mental health (Fox, 1999). Recent data also shows that endurance exercise can attenuate aging of the skin (Crane et al., 2015). Skeletal muscle is very plastic and quickly adapts to cope with the altered physiological demands associated with exercise.

The local adaptation of muscle is important not only for muscle function but also the health effects of training that affect the entire human body.

1.3.1 Cellular adaptation

An acute bout of endurance exercise transiently disrupts the cellular homeostasis. Repeated bouts of exercise, *i.e.* training, then lead to cumulative effects that slowly adapt the tissue (Hawley, 2002). Skeletal muscle cellular and tissue adaptations are well established and some examples are illustrated in Figure 2. An early adaptation seen with endurance training is increased capillarization (angiogenesis) that improves the supply of oxygen and nutrients and removal of waste products to/from the skeletal muscle. The transport capacity for glucose and lipids improves with training. There is an increased abundance of the main glucose transporter GLUT-4, resulting in improved insulin sensitivity, as well as fatty acid transporters in the skeletal muscle cell and its mitochondria. The improved transport ability is also accompanied by an increased storage of lipids and glycogen within the skeletal muscle fibers. The enzyme activity for breakdown of fatty acids also increases, as does the overall mitochondrial density (mitochondrial biogenesis), resulting in a higher aerobic capacity for oxidation of both lipids and carbohydrates (Coffey and Hawley, 2007; Egan and Zierath, 2013).

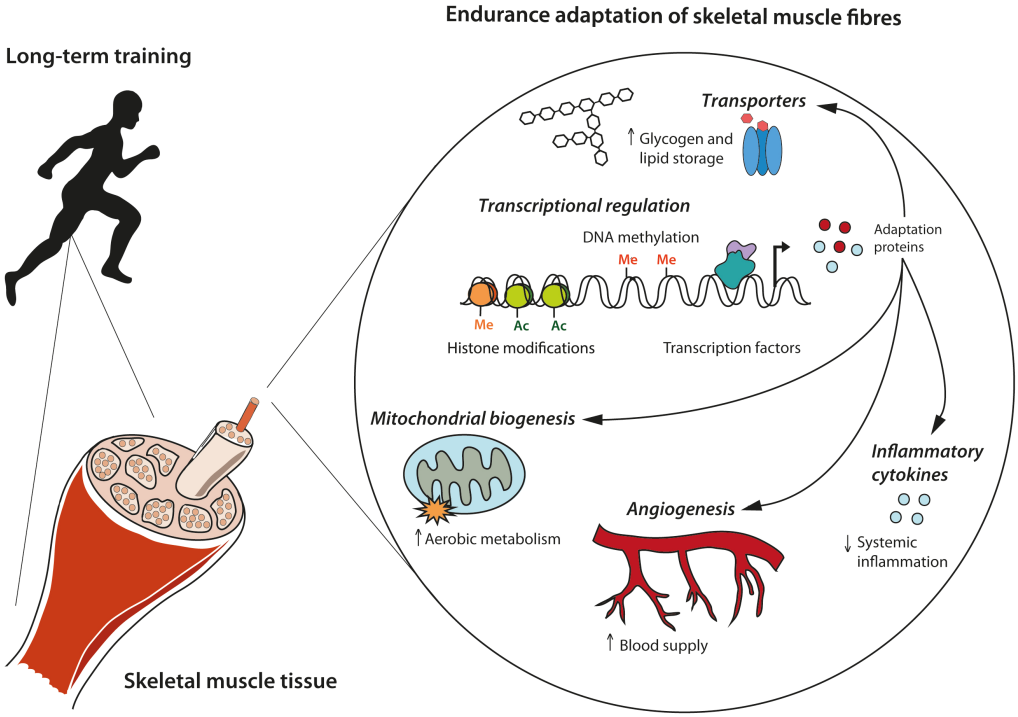

Figure 2. Skeletal muscle endurance-induced adaptation. Examples of known skeletal muscle tissue adaptations that occur in response to endurance training in humans. *Illustration: Susanna Appel.*

Together with the increase in supply of oxygen to the working muscles, these adaptations greatly improve the aerobic metabolic capacity of skeletal muscle to produce the high amounts of ATP required for endurance exercise. Fiber type changes that occur with training include a decrease in the proportion of type IIX fibers, while type IIA increases. Type I fibers, which are mainly recruited during low-intensity endurance training, may increase somewhat in size. Although the degree of fiber type transition is very limited, the amount of slow, oxidative myosin ATPases and oxidative enzymes in all fibers increases (Saltin and Gollnick, 2011). Whether transitions between fast-twitch and slow-twitch fibers can occur with training is not fully understood (Booth et al., 2010). Repeated contractions require propagation of action potentials along the sarcolemma, which in turns necessitates restoration of ion distribution across the cell membrane. Endurance training increases the ATPase activity of ion pumps, resulting in a faster restoration of the ion distribution (Green, 2000). The pH buffer capacity also improves (Hawley, 2002), which is beneficial because of the increased lactic acid production that occurs at higher exercise intensities. Apart from the local adaptation in skeletal muscle, exercise also stimulates many different signaling molecules that can exit skeletal muscle fibers and induce systemic effects on different tissues. An example that is shown in Figure 2 is the anti-inflammatory effect of long-term training. The adaptation processes of skeletal muscle were described early by (Holloszy and Booth, 1976) and recently reviewed by (Baar, 2014; Egan and Zierath, 2013).

1.3.2 Molecular adaptation mechanisms

The cellular adaptations are driven by extracellular and intracellular signals arising from the exercise stimulus. Extracellular signals include for example the increased nervous stimulation and changes in hormone levels and shear stress (Blauuw et al., 2013). A majority of the cellular changes are, however, endogenous and include for example altered temperature, pH, oxygen tension, ROS (reactive oxygen species) production, ATP/ADP ratio and levels of free Ca^{2+} (Booth and Thomason, 1991; Egan and Zierath, 2013; Saltin and Gollnick, 2011). These perturbations of the intracellular milieu initiate a remodeling process in skeletal muscle fibers by altering the activity and abundance of many different proteins. Ultimately, these changes lead to transcriptional regulation through activation of transcription factors or repressors, cofactors for these processes, or regulators of translation. AMPK (AMP-activated protein kinase) is activated when the ATP levels in the cell decrease, resulting in suppression of several anabolic pathways including glycogenesis, and stimulation of catabolic processes, *e.g.* glycogenolysis. Transcription factors activated by AMPK include MEF2 (Myocyte Enhancer Factor 2) and NRF-1 (Nuclear Respiratory Factor 1) that are important for mitochondrial biogenesis. CaMKII (Ca^{2+}-calmodulin-dependent protein kinase II) is an enzyme that is induced by exercise and activates several transcriptional regulators, including CREB (cAMP responsive element-binding protein), MEF2 and HDACs (Histone deacetylases) (Egan and Zierath, 2013). Mitogen-activated protein kinases (MAPKs) are also activated by endurance training (Widegren et al., 2000), the p38 MAPK, for example, has together with AMPK been suggested to increase the nuclear abundance of the cofactor PGC1α (Peroxisome proliferator-activated receptor (PPAR) gamma coactivator 1α) with training (Little et al., 2010). PGC1α

has repeatedly been described as a master regulator of mitochondrial biogenesis and regulates a number of downstream transcription factors and thereby expression of many different genes (Gidlund et al., 2015; Olesen et al., 2010). The oxygen tension in skeletal muscle cells is sensed by prolyl hydroxylases (PHDs). A reduced oxygen tension inactivates PHDs (Epstein et al., 2001), which would result in the acute activation of hypoxia inducible factor 1 that has been observed with endurance exercise (Ameln et al., 2005). This system is discussed further below under section 1.3.5. Associated to the oxygen tension is the redox balance, which is mainly reflected by the ratio between NAD^+ and NADH (the oxidized and reduced form of the electron carrier Nicotinamide adenine dinucleotide). Sirtuins are a group of protein deacetylases that are NAD^+-dependent and involved in the regulation of metabolism (Schwer and Verdin, 2008). Both SIRT1 and SIRT3 have been implicated in endurance training adaptation by affecting mitochondrial biogenesis, potentially through activation of PGC1α. SIRT3 is located in mitochondria and could thus influence function directly. For an overview of the complex role of SIRT1 and SIRT3 in exercise adaptation, see (White and Schenk, 2012).

The above mentioned factors represent a selection of some of the important factors known to be affected by endurance training in skeletal muscle. All these factors act directly or indirectly as transcription factors, regulating gene expression and thereby protein composition and activity. The induced changes, in turn, lead to the functional adaptation of skeletal muscle.

1.3.3 Exercise-induced transcriptional regulation

The transcriptome is the complete set of coding and non-coding RNA transcripts in a cell at a given point in time. A typical mammalian cell expresses between 11 000 and 13 000 genes (Ramskold et al., 2009). By increasing or decreasing transcription of these genes, cells can adjust different functions in response to environmental changes. Both acute endurance exercise and training are known to induce specific gene expression changes in human skeletal muscle (Ameln et al., 2005; Gidlund et al., 2015; Gustafsson et al., 1999; Kraniou et al., 2000; Norrbom et al., 2010; Ookawara et al., 2002; Pilegaard et al., 2003; Rullman et al., 2007; Russell et al., 2003; Short et al., 2005; Wadley et al., 2001; Wallace et al., 2011), something that has been studied extensively during many years. At the global level, an acute bout of endurance exercise in four subjects changed the expression of 126 genes (Mahoney et al., 2005), while a later study found 938 genes to be differentially expressed in twelve subjects performing a one-legged exercise bout (Catoire et al., 2012). In 2005, Timmons et al. conducted the first transcriptome-wide longitudinal study on eight sedentary individuals that performed six weeks of training. Using the Affymetrix platform (with 63000 probe sets), almost 500 genes were differentially expressed and several enriched processes were identified as contributors to the physiological remodeling (Timmons et al., 2005). With additional subjects, the differentially expressed genes subsequently increased to ~800, referred to as the training-induced transcriptome (Keller et al., 2011; Timmons et al., 2010). Regulation of gene

expression is therefore well established as an important mechanism behind skeletal muscle adaptation to endurance training.

Other tissues are not as well studied in humans, likely due to difficulties in acquiring tissue. Global changes in adipose tissue gene expression with six months of endurance training have been reported (Rönn et al., 2013), while cardiac muscle mainly has been studied in animals (Ellison et al., 2012).

1.3.4 Epigenetics and exercise

Epigenetic modifications regulate gene expression and may therefore be an important mechanism for skeletal muscle adaptation to training. There are a few studies that have associated the levels of physical activity in different cohorts to overall methylation level in leukocytes (Luttropp et al., 2013; White et al., 2013). In skeletal muscle, the epigenomic pattern changes over time, in part due to differences in lifestyle (Fraga et al., 2005). In response to an acute 60-min cycling exercise bout, it has been shown that global histone 3 acetylation at lysine 36 increases and histone deacetylases (HDAC) 4 and 5 are translocated out of the myonuclei (McGee et al., 2009). The HDAC 4 and 5 enzymes belong to the class II HDACs, which are enriched in striated muscle and have been implicated in the regulation of muscle gene expression, *e.g.* by repressing oxidative genes (McGee and Hargreaves, 2011). Transient specific changes in DNA methylation at a few promoter CpG sites also occur with one bout of endurance exercise in human skeletal muscle (Barres et al., 2012). In the context of type II diabetes, a six-month exercise intervention has been shown to alter DNA methylation in skeletal muscle (Nitert et al., 2012). For review, see (Voisin et al., 2015).

Epigenetics has also emerged as a mechanism for cellular memory (Bird, 2007).Transmission of cell memory is, however, complex and not fully understood but histone modifications and other epigenetic mechanisms are involved (Ng and Gurdon, 2008). Histone modification patterns can for example establish gene regulatory regions in an active or a poised state (Kimura, 2013). That environmental exposures at a young age can alter epigenetic marks and give rise to a different phenotype also later in life has been studied in mice for maternal nurturing of the offspring (Bird, 2007). This type of memory that is passed on from one cell to the next is highly interesting biologically. Although skeletal muscle cells do not divide, it is possible that environmentally-induced epigenetic changes (displayed as an altered gene activity) can remain for a long period of time. Exercise training-induced adaptation of human skeletal muscle is to a large extent expected to be lost upon detraining (Bishop et al., 2014). However, there are some indications of potential small residual effects. Resistance training increases the number of myonuclei per skeletal muscle fiber, and in mice, this effect is maintained some time after detraining (Bruusgaard et al., 2010). In humans, previous administration of androgenic anabolic steroids has induced an increase in the number of myonuclei, which may persist after cessation (Yu et al., 2014). These few studies indicate that possible intrinsic memory mechanisms may play a role after previous adaptation to physiological or pharmacological interventions in skeletal muscle. Despite the very limited

scientific evidence, there is still a notion in the popular science literature that once a person has been fit, it is easier to become fit again after a period of detraining.

1.3.5 HIF-1 AND EXERCISE

Hypoxia Inducible Factor 1α (HIF-1α) is a transcription factor that is an important mediator of cellular adaptation to hypoxia (Wang and Semenza, 1993). Target genes of HIF-1 promote improved oxygen delivery to cells, through for example erythropoietin (EPO) and vascular endothelial growth factor (VEGF), and improve anaerobic glycolysis through induction of glucose transporters and glycolytic enzymes, including lactate dehydrogenase (LDH) (Semenza, 1999). This feature made HIF-1α into an interesting candidate for training adaptation and it is activated following an acute bout of exercise (Ameln et al., 2005). However, after four weeks of unilateral training, HIF-1α transiently increased only in the untrained leg when exposed to acute exercise (Lundby et al., 2006). Also, another response gene of HIF-1 is pyruvate dehydrogenase kinase 1 (PDK-1) that inhibits the pyruvate dehydrogenase complex from converting pyruvate from glycolysis into acetyl-CoA that is required for further aerobic catabolism in the citric acid cycle and eventually the electron transport chain (Kim et al., 2006; Papandreou et al., 2006). By inhibiting aerobic energy production, HIF-1α would rather prevent further training adaptation. In humans, patients with the autosomal recessive disorder Chuvash polycythemia have constantly elevated HIF-1 activity, which leads to increased lactate accumulation and reduced muscle pH in response to exercise (Formenti et al., 2010). This is also supported by studies in mice with a skeletal muscle-specific HIF-1 deletion, that have many features typically associated with trained muscle, especially regarding mitochondrial function (Mason et al., 2004; Mason et al., 2007b).

HIF-1 is constitutively expressed and continuously degraded in normoxia, while it is stabilized and activated during hypoxia (Semenza, 1999). Yet, HIF-1 controls basal expression of its target genes also under normoxia (Xia et al., 2009). Regulation of HIF-1 activity occurs at multiple levels by adjustment of the rate of degradation or the transactivation capacity. Negative regulation of HIF-1α mainly occurs through hydroxylation of critical proline residues by PHDs, which mark the protein for degradation through the ubiquitin-proteasome pathway (Ivan et al., 2001; Jaakkola et al., 2001). There are two inhibitors of the transactivation capacity of HIF-1. Factor Inhibiting HIF (FIH) hydroxylates an asparagine residue that prevents binding to the coactivator CBP/p300 (Lando et al., 2002; Mahon et al., 2001). An epigenetic corepressor of HIF-1 is Sirtuin 6 (SIRT6), a histone 3 lysine 9 deacetylase (Michishita et al., 2008) that specifically targets glycolytic genes in favor of mitochondrial activity (Zhong et al., 2010).

2 AIMS

Understanding the molecular mechanisms behind adaptation to endurance training is highly important. It may have implications for our future lifestyle, for example through optimization of training interventions to promote the largest effect in each individual and by artificially inducing some of the beneficial effects of training in patients unable to perform regular exercise.

The overall objective of this thesis was to investigate the relation between epigenomic and transcriptomic changes following a highly controlled endurance exercise training intervention, and the potential presence of an endurance training-induced skeletal muscle memory. The specific aims of this thesis were to investigate;

1. the baseline transcriptome in human skeletal muscle, *i.e.* the differences within a muscle, between muscles from different legs of the same individual and between male and female muscle.
2. how the global DNA methylation pattern changes in response to long-term endurance training
3. if histone modifications are involved in training adaptation
4. how repeated training, after a period of detraining, affects the transcriptome, *i.e.* if there is any evidence of a skeletal muscle memory of training adaptation
5. the regulation of a specific molecular system, HIF, in highly trained individuals

3 METHODOLOGY

3.1 EXPERIMENTAL PROTOCOLS AND TISSUE SAMPLING

Before each study, the experimental protocol was explained to all subjects and informed consent was obtained. Each study was approved by the Ethics Committee of Karolinska Institutet and conformed to the *Declaration of Helsinki*. Table 1 summarizes the human studies included in this thesis.

	EpiTrain study		Cross-sectional Study (males)		Six-week bicycle study
Number of subjects	Males 12	Females 11	Controls 9	Elite Athletes 12	Males 24
Age (yrs)	28 ± 1	26 ± 1	24 ± 1	22 ± 1	23 ± 1
Height (cm)	181 ± 2	169 ± 2	180 ± 2	181 ± 2	182 ± 2
Weight (kg)	81 ± 5	67 ± 4	72 ± 2	72 ± 2	79 ± 3
Peak VO2 ($ml*kg^{-1}*min^{-1}$)	40 ± 1	38 ± 1	47 ± 2	75 ± 2	48 ± 1

Table 1. Subject characteristics for the different human studies. EpiTrain was the basis for Papers I, II and IV, the cross-sectional study was the basis for Paper V and the six-week study was included in Papers III and V. Data is presented as mean ± SEM.

3.1.1 EpiTrain study (Papers I, II and IV)

Twenty-three young, healthy volunteers (twelve men and eleven women) were included in the study. All individuals completed a questionnaire about exercise habits, disease history, smoking habits, use of contraceptives and other medications. The baseline fitness level was evaluated using the peak oxygen uptake (VO_2-peak) as determined on a cycle ergometer where the load was increased stepwise until exhaustion. Inclusion criteria were < 48 ml kg^{-1} min^{-1} and < 46 ml kg^{-1} min^{-1} for men and women, respectively. Baseline subject characteristics are presented in Table 1 and an outline of the whole study is shown in Figure 3A. The training protocol consisted of 45 min one-legged knee-extension exercise training four times per week for twelve weeks, in total 45 sessions. In this first training period (Period 1), all subjects exercised only one randomized leg. All training was conducted with the load and frequency supervised by an assistant. The load was continuously increased during each session, as well as during the training period.

In humans, skeletal muscle is a relatively unique tissue as it is accessible for sampling in adequate quantities, easily subjected to physiologically relevant interventions, highly plastic and also functionally testable in a reliable and relevant manner. In the EpiTrain study, two biopsies from the *vastus lateralis* muscle of each leg were obtained at rest before and after (24h) each training period (Figure 3A and B). No large blood vessels perforate the lower part

of this outer thigh muscle, making it ideal for sampling. For all studies, the tissue samples were taken under local anesthesia using the percutaneous needle technique (Bergström, 1962), immediately frozen in liquid nitrogen or 2-methylbutane cooled by liquid nitrogen (**Paper III**) and stored at -80°C.

Two one-legged knee-extension performance tests were conducted on a modified bicycle before and after the training period. A one-legged max test was performed using two different protocols; two minutes at a constant load (10W or 20W based on measured VO_2-peak) followed by an increase in load by 3 or 5 watts every 30 seconds until exhaustion. A fifteen minutes one-legged optimal performance test (15-min test) was performed on a different day or at least 6 h after the previous test. After one minute the load was 50% of the achieved maximum at the previous max test. Thereafter the load was kept as high as possible for the subject to be able to keep the pace at 60rpm for the full fifteen minutes. Heart rate was registered every minute, as well as the rate of perceived exertion (RPE, Borg-scale) of the working leg.

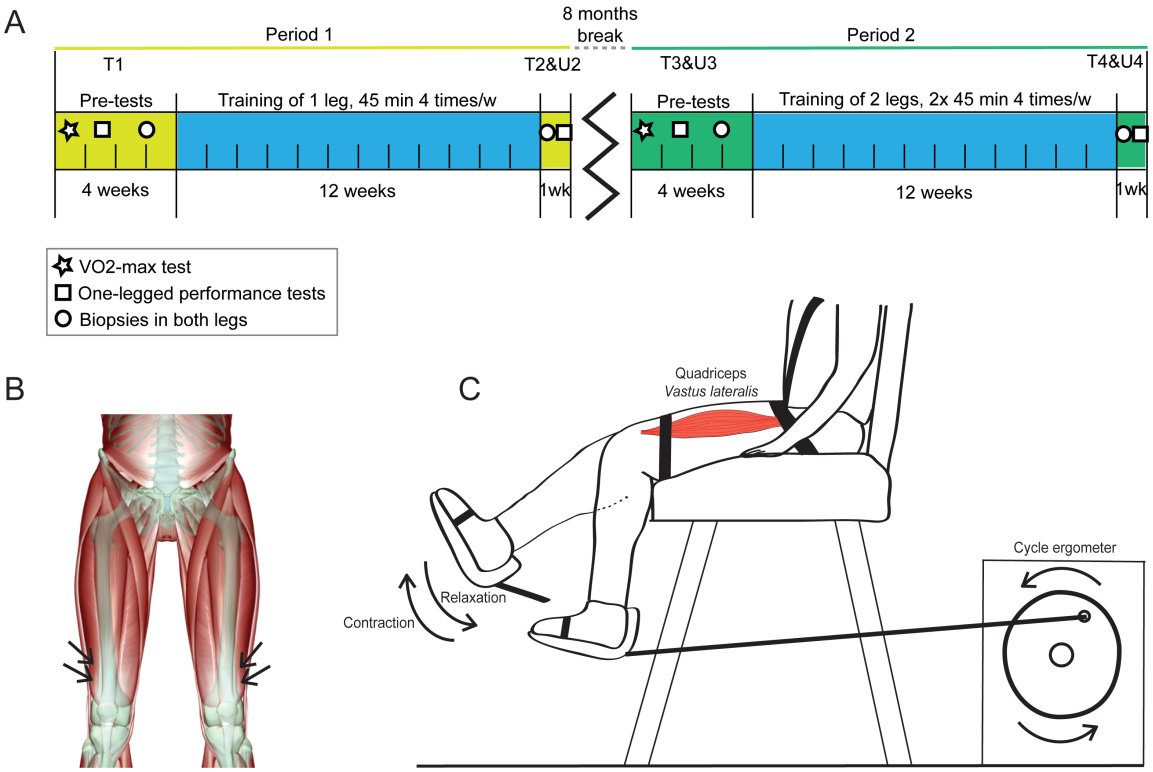

Figure 3. Experimental overview of the EpiTrain study. A) The outline of the study, with performance tests, biopsy time points with the corresponding code (T (for the leg trained in Period 1) or U (for the untrained leg in Period 1) followed by the time point (1-4)). B) Illustration of the biopsy sites. C) Schematic image of the one-legged training regime.

After nine months of detraining, twelve of the subjects started a second training period (Period 2) with the same number of sessions and workload, but now training both legs. Before start of the training period, all subjects performed a VO_2-peak test. The same one-legged performance tests were conducted and skeletal muscle biopsies were obtained from both legs before and after training. The purpose of the one-legged training regime, which is

illustrated in Figure 3C, was to include the other leg as an intraindividual, untrained control. Several environmental factors, like diet and stress can affect DNA methylation and these factors were assumed to affect both legs similarly. In this way, it was possible to isolate the effect of training on the skeletal muscle tissue. The repeated training also made it possible to compare the training response in a previously well-trained with a previously untrained leg.

3.1.2 Cross-sectional study (Paper V)

The cross-sectional study included twelve highly endurance-trained men (elite cyclists and triathletes) and nine moderately active men (controls). Baseline characteristics, including VO_2-peak, are presented in Table 1, and has previously been published (Fischer et al., 2006). Biopsies were obtained at rest from the *vastus lateralis* muscle, frozen in liquid nitrogen and stored at -80°C.

3.1.3 Six-week bicycle study (Papers III and V)

24 young, moderately active men (baseline characteristics are presented in Table 1) performed a supervised six-week endurance training program, four 45-min sessions per week at 70% of their pre-training VO_2-peak. Biopsies were taken from the *vastus lateralis* muscle before and 24 hours after the last training session. The effect on physical fitness was evaluated using the mean improvement in VO_2-peak (13%), data that has been published before (Vollaard et al., 2009). The biopsy procedure used for all human tissue sampling is shown in Figure 4.

Figure 4. Biopsy procedure using a Bergstrom needle. *Image courtesy of Susanna Appel.*

3.1.4 Skeletal muscle primary myoblasts

In **Paper V**, skeletal muscle satellite cells were extracted from resting skeletal muscle biopsies. After initial proliferation of the myoblasts until 80% confluence, the cells were differentiated into myotubes in a low-serum culture medium. The cells were treated with the prolyl hydroxylase inhibitor Dimethyloxallyl glycine (DMOG) to investigate the association between PHD activity and the expression of PDK-1.

3.1.5 C2C12 mouse myoblast cell line

C2C12 is an immortal mouse skeletal muscle cell line (Blau et al., 1985) that originates from satellite cells from the thigh muscle of a mouse from the CH3 strain (Yaffe and Saxel, 1977). It was used in **Paper V** for *in vitro* confirmation of the human data. Once the myoblasts were 80% confluent, the cells were differentiated into myotubes. To study the effect of an altered HIF-1 activity on skeletal muscle metabolism *in vitro*, the cells were treated either with DMOG, to induce HIF-1, or transfected with siRNA (small interfering RNA) against PDK-1, to inhibit translation of the response gene transcript. Transfection was performed overnight using Lipofectamine®. Lactate levels were investigated in the culture medium using standard lactate strips and the cells were subsequently harvested in Trizol® for mRNA analysis and in 0.1 M phosphate buffer with 0.5 % BSA (pH 7.7) for enzyme activity analysis at 6, 48 and 96 hours after removal of the Lipofectamine.

3.2 Enzymatic activity analyses

The experimental procedure was done according to the fluorometric principles of Lowry & Passonneau (Lowry and Passonneau, 1972). For skeletal muscle, a small section of each biopsy was freeze-dried and homogenized in 0.1 M phosphate buffer (pH 7.7) with 0.5% BSA (bovine serum albumin). For the *in vitro* experiments, cells were cultured in six-well plates and scraped off in 250ul of phosphate buffer with 0.5% BSA and 1x PI (protease inhibitor) cocktail per well. The cells were subsequently homogenized with a polytron for ten seconds on ice. For the cell experiments, the homogenate was also freeze-thawed a couple of times to break the cell and mitochondrial membranes.

3.2.1 Citrate synthase

Citrate synthase (CS) is a rate-limiting metabolic enzyme of the citric acid cycle that has been shown to reflect the mitochondrial content of skeletal muscle (Holloszy et al., 1970). A well-established adaptation of skeletal muscle to endurance training is an increase in mitochondrial density (Duscha et al., 2012; Holloszy and Booth, 1976; Holloszy et al., 1970), which enables muscle to produce more aerobic energy. CS activity was therefore used to obtain an objective, biochemical indication of the training response. To measure the activity of CS, the tissue lysates were added to a reagent solution containing NAD^+ and MDH (Malate dehydrogenase). Acetyl-CoA was added to start the enzymatic reaction and the reaction velocity was subsequently registered with a fluorometer (reduction of NAD^+ to NADH). To quantify the activity, the velocity was related to a standard curve computed from known amounts of NADH.

In **Papers III** and **V**, there was a 26% increase in CS activity with six weeks of training (Vollaard et al., 2009). In **Paper V**, the elite athletes in the cross-sectional study had > 2-fold higher CS activity compared to the moderately active controls (Wiik et al., 2005).

3.2.2 3-hydroxyacyl-CoA dehydrogenase

Another mitochondrial enzyme is 3-hydroxyacyl-CoA dehydrogenase, or β-HAD, that catalyzes a step in the β-oxidation of fatty acids. In response to long-term endurance training, there is an increase in the triglyceride content of skeletal muscle and the ability of muscle to utilize these stores through *e.g.* an increase in the lipoprotein lipase activity (Holloszy and Booth, 1976). β-HAD is also induced with endurance training (Ngo et al., 2012; Perry et al., 2007) and was investigated as an additional biochemical marker for training response. The β-HAD activity was measured in a reagent solution (0,5M Imidazol (pH 7), 0.1M EDTA, 5mM NADH diluted in carbonate buffer and 2mM Aceto-Acetyl-CoA) where addition of the tissue lysate started the reaction. Fluorescence of β-HAD oxidation of NADH to NAD^+ was recorded and quantified in relation to a standard curve from known amounts of NADH.

3.3 Transcriptional activity

3.3.1 RNA extraction

Total RNA was isolated from cells and skeletal muscle biopsies using the Trizol® reagent, based on a method by (Chomczynski and Sacchi, 1987). In short, each sample was homogenized with a polytron on ice. Total RNA was precipitated using isopropanol, the final RNA pellet diluted in RNase-free water and stored at -80°C. The concentration and quality was determined using the RNA 6000 Nano chip on the 2100 Bioanalyzer automated electrophoresis system (**Papers I**, **II** and **IV**) and/or absorbance at 260nm and agarose gel (all RNA).

3.3.2 RNA sequencing

RNA sequencing (RNA-seq) uses massive parallel sequencing of complementary DNAs (cDNAs) for genome-wide gene expression profiling. In addition to high-resolution quantification of exon expression, RNA-seq also allows for analysis of splice variants and identification of previously unknown alternative splicing events and novel transcripts (Nagalakshmi et al., 2010). RNA-seq results have been shown to be highly reproducible and well correlated to microarray expression data (Marioni et al., 2008) and the method has a large dynamic range. As the majority of a total RNA sample is rRNA, an enrichment step is necessary. For these studies, two μg of total RNA was used for each sample. Poly-A-tail hybridization to poly-d(T) probes made the analysis mainly include transcripts containing a poly-A-tail, which may also include transcripts marked for degradation (Ozsolak and Milos, 2011). The RNA was subsequently fragmented to improve the sequence coverage and then converted into cDNA. The libraries were barcoded and clustered on a cBot cluster-generation system using an Illumina HiSeq paired-end cluster-generation kit and sequenced as paired-end, 2x100 bp on the Illumina HiSeq 2000. All lanes were spiked with 1% phiX control library. In total, 155 skeletal muscle tissue samples from the EpiTrain study (**Papers I**, **II** and **IV**) underwent sequencing. Efforts have been undertaken to assess the impact of technical limitations for RNA-seq on gene expression analysis, including differences due to processing date, library preparation batches, differences between the Illumina sequencing flow cells and

lanes within each flow cell (Consortium, 2014). Although some residual effects will always be present, these issues have been addressed through batch correction, as described below.

3.3.2.1 Sequence read alignment and analysis

Sequencing of all samples generated over 3 billion paired-end reads in total, a vast challenge analytically. After initial quality control (FastQC), adapter trimming was performed with TrimGalore. The sequencing reads were then aligned to the human genome (version hg19/GRCh37) for identification of their genomic origin. This was performed with the alignment software TopHat2, which is one of the most widely used for RNA-seq data (Kim et al., 2013a). This is an important step as the choice of alignment tool influences the accuracy of the subsequent analysis (Engstrom et al., 2013). The aligned reads were then assembled into transcripts using Cufflinks (Trapnell et al., 2010). To obtain an accurate expression level, it is important to account for sequencing depth, gene length and composition of the RNA population, which is done through normalization (Robinson and Oshlack, 2010). In **Papers I** and **IV**, Fragments Per Kilobase of Exon per Million mapped fragments (FPKM) values were calculated and the aligned reads were also used to count the number of reads per gene using HTSeq (Anders et al., 2014). The gene counts were then used to calculate differential gene expression through DESeq2 (Anders and Huber, 2010) (**Paper I**) or Limma (Smyth, 2005) (**Paper IV**). In **Paper II**, the number of reads per gene were also calculated with HTSeq, lowly expressed genes were filtered out (counts per million (CPM) mapped reads < 1 in at least half of the samples) and normalization factors were calculated using TMM (trimmed mean of M values) to normalize for RNA composition (Robinson and Oshlack, 2010). Differentially expressed genes were analyzed with edgeR (Robinson et al., 2010). DESeq and edgeR are similar in several aspects, for example both use a negative binomial distribution and adopt an adjusted version of Fishers exact test (Rapaport et al., 2013).

Although the ability of RNA-seq to accurately detect relative expression is high, it has to be used with spike-in samples with known concentrations in order to attempt to quantify absolute expression levels (Consortium, 2014), which was not done in the present studies.

3.3.2.2 Novel transcript discovery

In **Paper I**, we investigated the presence of previously unknown transcripts in skeletal muscle. A reference-based assembly of the aligned reads was performed (using Cufflinks, (Trapnell et al., 2010)) without any transcript reference. The resulting 45 assemblies (one per sample) were merged with the Cuffmerge tool within Cufflinks, to which transcript annotation (ENSEMBL v. 71) was supplied. Transcripts that were classified as “unknown intergenic” were extracted, and additional removal of still overlapping transcripts due to slightly different assemblies from different samples was performed. The resulting list of putative novel intergenic transcripts was compared to the BodyMap data set and to a published list of lncRNAs (Hangauer et al., 2013). Some were also defined as protein-coding through comparison to a genome-wide set of known and novel protein-coding loci obtained

through HiRIEF (high-resolution isoelectric focusing) mass spectrometry (Branca et al., 2013).

3.3.3 Quantitative real-time PCR

Two µg of total RNA was used to generate cDNA through reverse transcription. Quantitative real time PCR (qPCR) was then used for quantification of specific transcripts. The analysis was done using TaqMan gene expression assays (Applied Biosystems) (**Papers I** and **V**) or SYBR green reactions on specifically designed primers and probes (see **Paper IV** for details). For SYBR green reactions, the specificity of the primer pairs was investigated using a melting curve. For all qPCR analyses, an endogenous control was used to correct for loading discrepancies and potential variation in efficiency of the reverse transcription. The different controls were; in **Paper I** beta-actin (*ACTB)*, in **Paper IV** 40S ribosomal protein S18 (*RPS18*) and in **Paper V** 18S ribosomal RNA or glyceraldehyde dehydrogenase (*GAPDH*) for human studies and hypoxanthine phosphoribosyltransferase (*Hprt*) for the murine cell line. For each subject, all samples were simultaneously analyzed in one assay run. The expression of each target was then evaluated by the $2^{-\Delta\Delta CT}$ method (Schmittgen et al., 2000).

3.4 DNA methylation analyses (Paper II)

3.4.1 DNA extraction

DNA was extracted with the Gentra Puregene Tissue Kit (Qiagen). In brief, 5-10 mg of tissue was homogenized by hand in lysis buffer with proteinase K, and incubated at 55°C for 1h. After RNase A treatment, protein was precipitated followed by centrifugation at 14 000 x g. DNA was subsequently precipitated using ice-cold isopropanol. DNA was washed in 70% EtOH and mixed with 50ul of DNA hydration solution. After incubation at 65°C for 1h, the concentration and quality of the DNA was analyzed with Nanodrop® and gel electrophoresis.

3.4.2 Luminometric Methylation Assay

LUminometric Methylation Assay, LUMA, measures the global level of methylation at CCGG sequences in the genome. It utilizes the restriction enzymes *MspI*, that cleaves all CCGG sites, and *HpaII*, which is methylation sensitive (Karimi et al., 2006). Sample DNA was mixed with Tango buffer, *EcoRI* enzyme for internal control, and *HpaII* or *MspI* (in separate plates). After 4 hours of incubation at 37°C, annealing buffer was added and the plates were subjected to pyrosequencing. The ratio between *HpaII* and *MspI* was then calculated to investigate unidirectional changes in the DNA methylation level.

3.4.3 450K DNA methylation arrays

Genome-wide DNA methylation profiling was generated with the Infinium HumanMethylation450K BeadChip array (Illumina) on bisulfite-treated DNA. Bisulfite converts cytosine bases to uracil, whilst methylated cytosines are protected. The Illumina array then employs two different chemistries and probe designs, Infinium I and II, to measure

methylation of 485 577 sites at single-base resolution. Infinium I has two probes per locus, one for methylated and one for unmethylated C's, while Infinium II has one probe, with a subsequent base extension with adenine for unmethylated and guanine for methylated sites resulting in differences in fluorescence emission. The array covers 99% of genes, 96% of CpG islands, >80 000 predicted enhancers as well as non-CpG sites (Bibikova et al., 2011). It cannot differentiate between 5-mC and 5-hmC. The samples were run at the BEA core facility, Karolinska Institutet.

3.4.4 Bisulfite pyrosequencing

In order to validate some of the specific findings from the array, bisulfite pyrosequencing was adopted. Primers for this purpose were designed using the PyroMark assay design software (Qiagen) or the BiSearch primer design tool (http://bisearch.enzim.hu). After bisulfite treatment, the DNA was amplified with a specific primer pair, one biotinylated at the 5' end. The PCR product was subsequently mixed with streptavidin-coated sepharose beads to acquire a single-stranded product and mixed with a specific sequencing primer. Pyrosequencing was then performed by sequential injection of nucleotides in a controlled order, where each nucleotide is degraded before addition of the next. Successful incorporation of a nucleotide renders a bioluminometric signal as the resulting release of pyrophosphate (PP_i) provides energy for a luciferase (Tost and Gut, 2007).

3.5 Protein analyses

3.5.1 Immunohistochemistry

In **Paper III**, human skeletal muscle tissue sections (10µm) were cut in a cryostat at -20°C and placed on microscope slides. For each subject, three sections from the before and three from the after training samples were placed together on one slide. The slides were air dried for 5 min in room temperature followed by fixation in 4% PFA for 10 min. After washing, the sections were blocked in 5% horse serum in PBT (Phosphate-buffered saline with 0.1% Tween). Primary antibodies, against a specific histone modification and laminin, were diluted in blocking solution and incubated over night at 4°C. After wash in PBT, all slides were incubated for 1 h dark at room temperature with secondary antibody (anti-rabbit-Red X and anti-goat-FITC). After additional washing, the slides were mounted in VECTASHIELD® with DAPI stain for DNA.

3.5.1.1 Confocal microscopy analysis

Tissue sections were visualized with the Zeiss LSM710 confocal microscope at 40x magnification, using channels for DAPI (blue), FITC (green) and Rhodamine RedX (red). For each subject, three images were taken before and three after training, all from the same glass slide. The images were analyzed using the Volocity® software, where objects were automatically excluded by size (>700µm^3 and <50µm^3) or manually excluded if they did not appear within a nucleus (blue).

3.5.2 Chromatin Immunoprecipitation and sequencing

Chromatin immunoprecipitation (ChIP) analysis was included in **Paper III**. Skeletal muscle tissue was cross-linked with 1% formaldehyde, washed (in PBS with 1mM PMSF and 1% PI cocktail) and then lysed in lysis buffer to isolate nuclear material. The nuclear extracts were clarified by centrifugation and the lysate was diluted five-fold in ChIP lysis buffer, mixed and subsequently sonicated (16x16 pulses, Branson Sonicator). This generated DNA fragments between 100 and 600bp long. For each sample, 100ul of sonicated lysate was used as input control. The samples were then immunoprecipitated with an H3K27me3 antibody (07-449, Merck Millipore, where previous lots have been validated for ChIP experiments (Egelhofer et al., 2011)), and magnetic A/G beads at 4°C overnight. The immune complexes were washed with RIPA buffer and TE buffer with 50mM NaCl. DNA was subsequently eluted with elution buffer and NaCl was added to 0.2M, followed by reverse cross-linking at 65°C overnight. Purification of the DNA was performed with phenol/chloroform/isoamylalcohol.

DNA was prepared with the NEBNext ChIP-Seq Library Prep and sequenced on the Illumina HiSeq 2000 (by the BEA core facility, Karolinska Institutet). Sequencing generated on average 11.4 million reads per sample. Reads were aligned to the human genome using Bowtie 2 (>90% alignment) and peak calling was performed using MACS (Zhang et al., 2008).

3.5.3 Protein extraction and immunoblotting

For **Paper V**, skeletal muscle biopsies were homogenized in the appropriate buffer on ice using glass homogenizers. The concentration was measured using the Bradford Assay and an equal amount of protein from all samples was subsequently boiled in Laemmli loading buffer. The protein was separated on sodium dodecyl sulfate (SDS) polyacrylamide gels and blotted onto Protran nitrocellulose or Polyvinylidene fluoride (PVDF) membranes. The membranes were blocked for 1h with 5 % dry milk or 0.5 % gelatine in tris-buffered saline containing 0.1 % Tween 20 (TBS-T) and then incubated with the appropriate primary antibody at 4°C overnight. After washing, the membranes were incubated with an HRP-linked secondary antibody for 1 h at room temperature. After additional washing, the immune-complex was detected using enhanced chemiluminescence (ECL) or SuperSignal West Femto Sensitivity Substrate. To correct for potential differences in loading, the protein of interest was related to the amount of α-actinin, which was investigated on the same blot.

3.6 Statistics

3.6.1 Student's t-test

In **Paper I**, unpaired Student's *t* test was used to compare CS levels and specific mRNA's (from qPCR validation) between males and females. In **Paper II**, the bisulfite pyrosequencing validation was analyzed with a paired Student's *t* test for before and after training in the trained and untrained legs. In **Paper IV**, the qPCR validation of RNA-seq data was also analyzed with a paired Student's *t* test, while in **Paper V**, unpaired tests were adopted for the cross-sectional study and primary cell experiments while paired tests were

used to analyze protein and mRNA in the longitudinal study. For all analyses, the results were considered significant at $p<0.05$.

3.6.2 ANOVA

A two-way repeated measures ANOVA (Analysis of variance) was adopted for all measurements with repeated sampling and more than one group. That includes the physiological measurements and enzyme activity analyses in **Papers II** and **IV**, and LUMA and hydroxymethylation analyses in **Paper II**. In **Paper V**, results from the *in vitro* experiments with C2C12 cells were analyzed with ANOVA. Specific differences were investigated with the Bonferroni post hoc test and considered significant at $p<0.05$.

3.6.3 Bioinformatics of RNA-seq data

The RNA-seq expression data has been analyzed in two principal ways. For **Papers I** and **IV**, the analyses have been based on multivariate statistical methods. For **Paper II**, where the primary focus was on correlating changes in expression to changes in DNA methylation, the Bioconductor package edgeR was used to identify differentially expressed genes. Basic treatment of the raw count data was very similar, as described under Sequence read alignment and analysis (3.3.2.1).

3.6.3.1 Multivariate data analysis

The multivariate statistical analyses were performed using the software Simca p+13.0.3x64 (Umetrics, Sweden) on normalized gene-level counts. Principal component analysis (PCA) aims to reduce the number of dimensions of a data set into visually comprehensible components that explain the largest amount of variability found in the data. By identifying the linear combinations of the data points in a matrix where the highest variance is found, the data points can be rotated into a more informative co-ordinate system where the axes, *i.e* the components, represent the highest variance. The first principal component thus identifies the dimension where the highest variance can be found in the data, the second component the second highest and so on. Each component is orthogonal to the next. For PCA, no initial categorization of the data was performed, *i.e.* the samples were not grouped in any way. For the majority of comparisons, the data was visualized as the first three principal components with coloring to identify the different samples. The results were also reported as an R_x^2 value, which accounts for the variance explained by the components, also described as the goodness of fit of the model, and a Q_x^2 value, which is the cumulative fraction of the total variation of X (which is the entire gene or isoform expression data set) that can be predicted by the model, *i.e.* the quality of R_x^2. The Q_x^2 value is derived from cross-validation, a procedure that aims to simulate the predictive ability of a PCA model. The principle behind it is to continuously remove part of the data and validate how well the model predicted the removed values in comparison to the actual observed values. The resulting Q^2 is the sum of squares of the difference between the predicted and observed variables (Eriksson et al., 2013). The application of PCA to transcriptome data has been previously described (Kjellqvist et al., 2013; Kurtovic et al., 2011).

Orthogonal Projections to Latent Structures by means of partial least squares discriminant analysis (OPLS-DA) is a multivariate statistical method where the data set can be categorized into groups. Similarly to PCA, OPLS also aims to reduce the number of dimensions of the data, while the discriminant analysis part includes the classification component. The relationship between a data matrix X, gene or isoform expression values in this case, can thereby be investigated in relation to a class Y (right and left legs, males and females or before and after training). What OPLS-DA does is that is separates the systematic variation in X into two different parts, one that is linearly related to Y (referred to as the predictive component) and one that is orthogonal, *i.e.* unrelated, to Y. The purpose of this procedure is to remove potential systematic variation in X that is not related to the investigated groups defined by Y. An OPLS model is also described by the R_x^2 value, together with R_y^2 that accounts for the explained variance of the Y component and Q_y^2, which is the cumulative fraction of the total variation of Y that can be predicted by the model (Eriksson et al., 2013).

If the model was good based on the model parameters, differentially expressed genes and isoforms were identified through the loadings in the OPLS models. Loadings are the cosine of the angle between a variable and the principal component, which therefore indicates how similar the variable is to each respective principal component, *i.e.* how much it contributes to the total variance described by the principal component (Eriksson et al., 2013). If, in theory, a specific variable is in the exact same direction as the first principal component, that variable would explain all the variance observed for the first principal component. The level of significance of each loading was calculated as the absolute value of each loading (ABS (loading)) subtracted by the absolute value of the jack knife confidence interval (Martens et al., 2001), where a positive value indicated that a gene was significant. The jack knife confidence interval is a cross-validation technique where the average for the whole set of samples is calculated and subsequently there is an iterative process where one sample is dropped from the set to obtain partial estimates. The differences between the partial estimates and the estimate for the whole set are then used to estimate standard errors and calculate confidence intervals. This technique has also been used previously for differential expression analysis (Kjellqvist et al., 2013; Kurtovic et al., 2011). In **Paper I**, a significance level of 0.05 was used, while for **Paper IV** the significance level was set to 0.01.

Prior to the multivariate analyses all data was subject to unit variance scaling (UV-scaling), where each variable was divided by its standard deviation. This equalizes the variance between different variables, which is important since both PCA and OPLS aim to identify the maximum variability in a data set. The data was also mean-centered, where the average for each variable was subtracted from the data, resulting in a mean of 0 for all variables. In the original co-ordinate system, the average for all variables will therefore be centered at the origin.

3.6.3.2 Differential gene expression using DESeq2, edgeR and Limma

Differential expression was also evaluated with the univariate methods DESeq2, edgeR and Limma. DESeq2 and edgeR (used in **Papers I** and **II** respectively) are very similar methods.

They both evaluate differential gene expression based on generalized linear models, use count data as input, and utilize a negative binomial model to account for over-dispersion of the count data (the sample variance exceeds the sample mean) (Anders and Huber, 2010; Robinson et al., 2010). See (Soneson and Delorenzi, 2013) for a comparison of these and other methods for analysis of differential expression of RNA-seq data.

In **Paper I**, DESeq2 was used to investigate differential expression within a muscle and between the right and left leg of an individual. In **Paper II**, the time point (T2 vs T1), library preparation and the individual were included as covariates and for T2 vs T1 the log2FoldChange and false discovery rate (FDR) were reported for each gene. A multidimensional scaling plot was obtained using the top 1000 genes (displaying the largest biological variation between samples) where the distance between each pair of samples (T1 and T2 for each individual) represent the biological coefficient of variation from edgeR. In **Paper IV**, Limma was used to investigate potential differential expression with training for the list of putative novel transcripts presented in Paper I, as well as to validate the significant differentially expressed genes identified with OPLS. Putative novel isoforms were investigated using a new tool called Ballgown, because of its ability to perform differential expression analysis on transcriptome assemblies (Frazee et al., 2015).

3.6.4 Bioinformatics of DNA methylation data

The pre-processing and normalization of DNA methylation data was carefully evaluated in a separate paper, not included in this thesis (Marabita et al., 2013). The data processing in **Paper II** included color adjustment (because of different fluorescent colors used on the array) and quantile normalization on pooled signal intensities from methylated and unmethylated probes to reduce the between sample variation due to position on the actual slide. BMIQ (beta mixture quantile dilation) is a correction method that aims to eliminate probe-type bias (Marabita et al., 2013), which was performed on the calculated β-values ($\beta=I_M/(I_U+I_M+\alpha)$, where I_M and I_U represent fluorescence intensities for the methylated and unmethylated probes, respectively, and α is a constant. β-values (range between 0 and 1) basically reflect percentage methylation of a specific locus. Before differential methylation analyses, we filtered out 65 SNP-associated (single nucleotide polymorphism) probes, probes on chromosomes X and Y (because the study included both males and females) and probes with a detection p-value > 0.01 in over 5% of the samples (detection p-value represents the confidence with which the probe is positively detected above a negative control background level). Batch effect correction was performed using ComBat. The subsequent analyses were performed on M-values ($M=\log_2((I_M+\alpha)/(I_U+\alpha))$), partly because the data distribution is more homoscedastic (approximately equal standard deviations across the data set), which most statistical methods assume.

Differentially methylated positions (DMPs) were identified using Limma (Smyth, 2005), with the group (T2 vs T1) and the subject as covariates. DMPs were considered significant if FDR<0.05. Standard Illumina annotation was used to annotate the DMPs (relation to genes, relation to CpG islands, enhancers etc.) in relation to non-DMPs and the entire array probe

set. To investigate the DMPs further, we used publically available NIH roadmap epigenomic data for skeletal muscle to identify H3K4me1, H3K4me3 and H3K27ac peaks using MACS (Zhang et al., 2008). A list of putative promoters (H3K4me3) (Guenther et al., 2007) enhancers (H3K4me1) and active enhancers (H3K27ac ∩ H3K4me1) (Creyghton et al., 2010) was obtained and the relative fractions of DMPs, non-DMPs and the entire array probe set located within those elements were calculated. In addition, chromatin segmentation tracks for an HSMM (human skeletal muscle myoblast) cell line were downloaded from the UCSC genome browser and the relative fraction falling in each segment category was calculated for with DMPs and the entire array position; a fold difference was then obtained (DMPs vs. array).

3.6.5 Relation between DNA methylation and transcriptome data

For DNA methylation data, the Illumina annotation for genes associated to analyzed cytosines on the 450K array was used to acquire a methylation associated gene list. Together with the gene expression data, a Spearman correlation was calculated for the normalized CPM value and the corresponding M-values. The calculation was done for all methylation/gene pairs and for DMPs/DEGs (Differentially Expressed Genes) only. The differential gene expression data was also used to construct a transcriptional network based on a mutual information algorithm developed in ARACNE (Algorithm for the Reconstruction of Accurate Cellular Network) (Margolin et al., 2006). This algorithm aims to reverse engineer gene regulatory networks and identify transcriptional interactions. Mutual information between differentially expressed genes was considered significant at $P<10^{-14}$ and the final consensus networks were generated with bootstrap runs and a significance level of $P<10^{-10}$. The networks were visualized using the Cytoscape software.

3.6.6 Functional categorization and network analyses

To obtain a functional categorization of the differentially expressed or methylated high-throughput data, several different approaches were taken. In **Paper I**, the entire skeletal muscle transcriptome (expressed at an average FPKM>1) was functionally annotated using the Panther database (Thomas et al., 2003), while the DAVID (Database for Annotation, Visualization and Integrated Discovery) gene ontology (GO) annotation tool (Dennis et al., 2003) was used for DEGs between males and females. DAVID results were reported for functional categories that had a fold enrichment >1.5 and a FDR<0.05. The presented p-values represent the probability of finding that subset of genes from a certain GO term in a random set of genes drawn from the complete background gene list. The Ingenuity pathway analysis software was also used to investigate functional enrichment of expression differences between males and females, including potential upstream regulators. In **Paper II**, GO analysis of the gene expression data was performed using GOseq (Young et al., 2010) and the resulting enriched KEGG pathways were visualized with pathview (Luo and Brouwer, 2013). For DNA methylation data, functional gene categories associated with DMPs were investigated using GREAT. Potential enrichment of known transcriptional motifs in a region of ± 100bp around DMPs was investigated using HOMER. In **Paper IV**, ontology of

differential gene and isoform expression was performed with DAVID and Ingenuity. For novel transcripts, conserved regions were investigated using the UCSC genome browser and presence of open reading frames and protein-coding motifs were analyzed with the ORF Finder Sequence Manipulation Suits and Motif Search tool, respectively. For details, please see each respective paper.

4 RESULTS AND DISCUSSION

4.1 THE BASELINE SKELETAL MUSCLE TRANSCRIPTOME

For most skeletal muscle studies, repeated sampling is necessary. Samples taken from different legs or different parts of the same leg are often assumed to be similar at baseline, despite little scientific evidence. Although alterations in gene expression levels in different conditions and in response to interventions have been studied extensively in skeletal muscle, there is no systematic analysis of the baseline skeletal muscle transcriptome. **Paper I** aimed to investigate this with specific regard to sex differences, alternative splicing and tissue homogeneity.

4.1.1 Tissue homogeneity

From RNA sequencing of 48 biopsies from nine men and nine women, 12 659 unique known transcripts were found to be expressed in skeletal muscle, with 10 419 representing protein-coding genes. Many mitochondrial and structural genes were among the most highly expressed. The number of expressed genes was similar to what has been observed in other human tissues, as well as skeletal muscle, with RNA-seq (Wang et al., 2008). A comparison between 28 different human tissues shows that skeletal muscle is most similar to heart muscle and adipose tissue and the number of genes that are considered uniquely elevated in skeletal muscle is <400, mainly structural components involved in contraction. Similar to heart muscle, a majority of the transcribed genes in skeletal muscle are, however, involved in energy turnover (Lindskog et al., 2015). At the isoform level, just over 23 000 different transcripts were transcribed above the threshold (FPKM>1), out of which 15 455 were classified as protein-coding (**Paper I**). Alternative splicing occurs in virtually all multi-exon genes and appears to be highly tissue-specific (Wang et al., 2008). The average number of isoforms per gene was 2 in this study, but there were 113 genes that expressed 10 different isoforms or more where *NDRG2* was the most versatile with 38 expressed isoforms. The tissue homogeneity in these repeated samples was found to be very high. Multivariate statistics, with unsupervised PCA and supervised OPLS-DA, was adopted to compare global expression differences within legs (*i.e.* with two biopsies from the same leg) and between legs, but no systematic differences were found. This was confirmed by linear model statistics as well. When interpreting the results, it is important to consider the individual variability that has been observed for isoform expression (Wang et al., 2008). The reported numbers represent averages across all individuals and it is unlikely that all individuals expressed all the investigated isoforms. A closer look at all expressed isoforms from a few genes, however, showed that the most abundant isoforms were high among most subjects, while the lowly expressed ones were more variable. This could also reflect a technical limitation as most software, including Cufflinks, perform worse in estimating expression of lowly expressed genes and isoforms (Kanitz et al., 2015).

4.1.2 Sex differences

A supervised OPLS-DA analysis comparing male and female human skeletal muscle showed clear differences both at gene- (>3000) and isoform (>5000) level (Figure 5A-B), regardless of whether the sex chromosomes were excluded or not. In females, enriched transcripts belonged to genes coding for mitochondrial factors for cellular respiration and fatty acid oxidation, while those enriched in males were cytoplasmic and involved in protein catabolism (Figure 5C). Some difference in gene expression between the sexes was expected, as this has been investigated previously with microarray (Maher et al., 2009; Welle et al., 2008). Maher et al. compared skeletal muscle from 12 men and 12 women with a global microarray approach and found 66 genes with a significant fold difference >1.2 (Maher et al., 2009). Fifty percent of those were found among our differentially expressed genes. Using oligonucleotide microarrays to compare muscle between 15 men and 15 women (age 20-75), Welle et al. (Welle et al., 2008) found approximately the same number of genes to be differentially expressed as in our study, but only 25% of those were the present in both studies. Age differences between the included subject populations, as well as methodological and statistical differences can likely explain some of the discrepancies.

Fiber type composition is one physiological difference between male and female skeletal muscle that, in part, can explain the differential gene expression. In males, glycolytic type II fibers are more abundant, while females have a higher percentage area of more oxidative type I fibers (Staron et al., 2000). This was reflected in our transcriptome data, where female muscle was highly enriched for genes involved in mitochondrial function while many glycolytic genes (*e.g. LDHA*, *GAPDH*, *ALDOA* and *PFKM*) were more highly expressed in male muscle. The more oxidative phenotype of women could also reflect a small difference in average fitness level, and although there were no significant differences between the groups, the average CS activity was somewhat higher in women. There was no difference in VO_2-peak. Type I fibers also have a higher capillary density, and female muscle was enriched for endothelial markers as well as several angiogenic factors, including *VEGFA*.

ACTN3 is a structural protein in the sarcomere that is only expressed in type II fibers. A SNP (R577X) in the ACTN3 gene results in a premature stop codon that leads to a dysfunctional protein, *i.e.* lack of ACTN3. An XX genotype for this SNP in humans has, for example, been associated to reduced response to sprint exercise training (Norman et al., 2014). The expression of ACTN3 was higher in males, likely due to the higher percentage area of type II fibers. However, it is also important to consider the potential difference in the genotype prevalence among these subjects. A small sample size increases the risk for a skewed genotype distribution, and this could influence observed expression differences. This particular SNP has been analyzed in these subjects and no one has the XX genotype, however it is not possible to exclude for other genes.

In order to further investigate the reason behind the sex differences in skeletal muscle gene expression, we also used Ingenuity to predict upstream regulators of the enriched genes. In females, the most significant predictor was P53 (*TP53*). Other predictors were the PPARs

(Peroxisome proliferator-activated receptors), regulators of for example lipid metabolism (Ehrenborg and Krook, 2009), and ß-estradiol. Hormonal differences were expected, but there was no predicted regulation by male sex hormones. Two factors involved in translation were predicted as upstream regulators of genes enriched in male skeletal muscle. The substantial differences between male and female skeletal muscle may explain differences in physiological phenotype and possibly also disease susceptibility between the sexes.

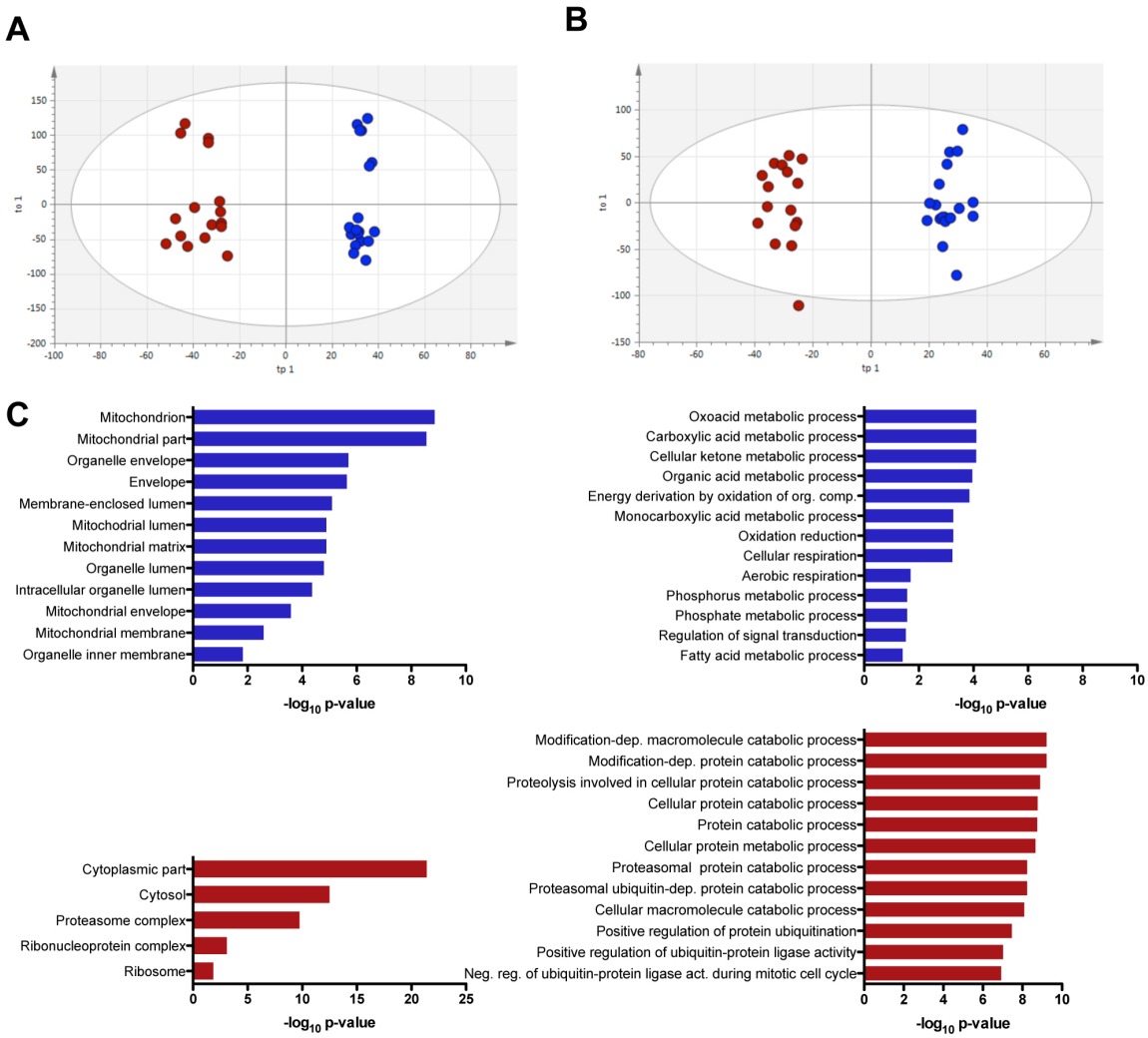

Figure 5. Sex differences in the human skeletal muscle transcriptome. A) OPLS-DA score plots of isoform and B) gene level expression data between males (*red*) and females (*blue*). Data is analyzed in 34 samples in total, two from each individual. C) Ontology analysis for cellular component (*left*) and biological process (*right*) for genes enriched in females (*blue*) and males (*red*), respectively.

4.1.3 Novel transcripts

The high number of sequenced samples generated over 830 million paired-end reads in total. The whole data set was used for a reference-based *de novo* assembly for identification of previously unannotated loci, which resulted in 2430 novel transcripts. The majority were either associated with lncRNAs (17%), (from a study by (Hangauer et al., 2013)), found in other RNA-seq data sets (78%) or conserved in mammals (99%), indicating that they are transcripts of functional importance. Five transcripts co-localized in the genome with newly

identified peptides (Branca et al., 2013), which strongly indicates that they are protein-coding genes. It is important to note that all these transcripts were unannotated in ENSEMBL, which means that some of them are likely included in for example RefSeq, and some might have been described before, but not annotated.

In summary, **Paper I** provides a deep and extensive baseline reference for the human skeletal muscle transcriptome, both with regards to alternative splicing events, novel transcript expression and sex differences in functional ontology. The high tissue homogeneity observed is of great importance for studying interventions, which is the basis for the rest of the papers in this thesis.

4.2 THE TRAINING-INDUCED TRANSCRIPTOME

The training-induced transcriptome was investigated in the EpiTrain study and is reported in **Papers II** and **IV**. Three months of one-legged endurance training (Period 1) induced significant physiological changes in the trained leg, measured by two different one-legged performance tests (Figure 6, top). A small performance improvement was also detected in the untrained leg. A more objective measure of the skeletal muscle adaptation was obtained by investigating changes in citrate synthase (CS) and 3-hydroxyacyl-CoA dehydrogenase (β-HAD) activity, which increased by 50% and 20%, respectively (Figure 6, bottom). There was no change in the untrained leg. This confirms that the training stimulus induced a specific and significant adaptation.

4.2.1 Gene expression changes

In **Paper II**, linear model analysis of the RNA sequencing data showed that 4076 genes significantly changed in response to training. Fifty-four percent of the regulated genes increased in expression, which is relatively low compared to other studies showing a much larger majority of genes increasing in expression (Timmons et al., 2010).

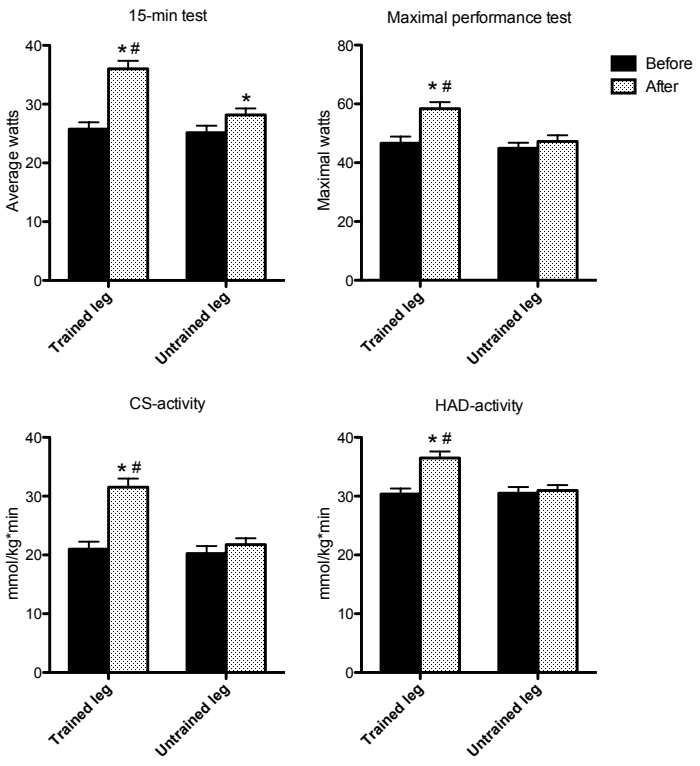

Figure 6. Physiological and biochemical adaptation. Data from Period 1 of the EpiTrain study. Data is presented as mean ± SEM. * indicates significance between before and after training and # between legs, $P<0.05$.

Gene ontology and pathway analysis of the differentially expressed genes (DEGs) revealed many processes known to be associated with training adaptation, for example cellular respiration and blood vessel development (Figure 7). Enriched pathways included for example oxidative phosphorylation and ECM-receptor interaction. To identify transcriptional interactions, the data was also visualized as a gene regulatory network based on mutual information from the gene expression data, as implemented in ARACNE (Margolin et al., 2006). The network can be viewed in full online in the freely available supplementary material of **Paper II** (supplementary figure S7).

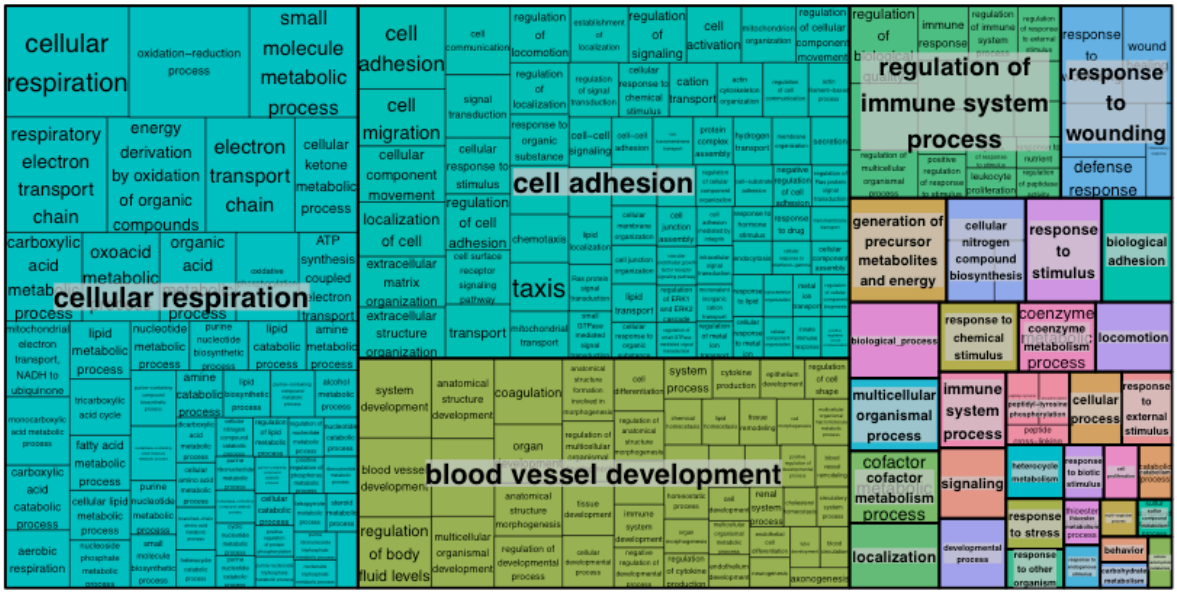

Figure 7. Molecular function ontology. Gene ontology for molecular function of the differentially expressed genes in response to training in Period 1 of the EpiTrain study. Significant categories are visualized using TreeMaps.

In **Paper IV**, multivariate statistics was applied to the RNA-seq data and confirmed the results from **Paper II** with a clear difference between the samples before and after training. 2394 genes were found to be differentially expressed, at a significance level of 0.01 (resulting in lower number of genes compared to **Paper II**). The untrained leg was also included in the analysis, showing that the training response of the trained leg was very specific. However, there were genes that changed also in the untrained leg, out of which approximately 500 also changed in the trained leg. This can for example be due to the element of time (three months between the biopsies) or through vascular or other systemic effects induced by training in the trained leg. Skeletal muscle tissue releases myokines into the circulation (Pedersen and Febbraio, 2012) that could potentially induce effects also in the contralateral leg, as well as other parts of the body. The observed changes in the untrained leg highlights the importance of including a relevant control when conducting exercise training intervention studies.

Of the 2394 DEGs in **Paper IV**, 75% were also found in **Paper II**. These multivariate and univariate methods complement each other, although there are evident discrepancies that are inevitable for different statistical approaches. Multivariate statistics is highly appropriate for

this type of data that is multivariate in nature. It is unpaired, however, which is a disadvantage considering the increased power of a paired analysis that is more appropriate in this experimental set-up with repeated sampling from the same individual. We are currently exploring a potential paired multivariate approach for this data. EdgeR is univariate, but allows for efficient pairing of the samples and is commonly used, which simplifies comparison to other studies.

4.2.2 Isoform expression changes

In **Paper IV**, the training-induced transcriptome changes were also investigated at isoform level. With training, 3404 isoforms, distributed across 2624 genes, were differentially expressed. For a majority of the genes, only one isoform was thus differentially expressed, while 2% of the genes differentially expressed 4-8 isoforms. The *NDRG2* gene, with 18 differentially expressed isoforms, displayed 17 isoforms decreasing and one isoform increasing in expression. *NDRG2* has been shown to be a PGC1α/ERRα transcriptional target in C2C12 myotubes, influencing protein turnover and the regulation of genes involved in muscle contraction and function (Foletta et al., 2013). From the *SPARC* gene (Secreted protein, acidic, cystein-rich, or osteonectin), five isoforms significantly increased in expression. This protein has recently been identified as a myokine that is induced by exercise and suppresses tumorigenesis in the colon (Aoi et al., 2013). Gene ontology analysis showed that the differentially expressed isoforms were mainly associated to cellular respiration and ATP synthesis coupled electron transport, and enriched pathways included oxidative phosphorylation and the TCA cycle. These results were thus very similar to the analysis performed at gene level. Many differentially expressed isoforms were highly interesting for endurance training adaptation and its health effects.

Fifty-four genes differentially expressed multiple isoforms that changed in opposite directions. This information further highlights the complexity of gene regulation and the importance of recognizing this before drawing conclusions solely from gene-based data. Examples of these genes include the mitochondrial ATP synthase *ATP5G1*, the *FUS* gene (Fused In Sarcoma) that encodes an RNA-binding protein involved in pre-mRNA processing (Vance et al., 2009), and Pyruvate Kinase Muscle (*PKM*) (Figure 8). A potential reason for differential regulation of isoforms from the same gene is that the isoforms have different functions and therefore contrasting roles in muscle adaptation to training. It could, however, also be due to a change in cellular composition of the muscle, which is known to occur with endurance training (Gustafsson, 2011) and isoform expression shows cell-type specificity (Wang et al., 2008).

The analysis at isoform level is less robust compared to the analysis at gene level because it is more difficult to correctly map reads to specific isoforms compared to genes (Consortium, 2014). Therefore, we utilized the second training period in the EpiTrain study to validate the results from the first period and identify the most training-responsive isoforms. The power of Period 2 was much smaller, as only twelve subjects finalized the entire study. However, we found 153 isoforms that were differentially expressed with training in the leg trained in

Period 1 and in both legs trained in Period 2. A Venn diagram showing the specific numbers for all trained legs as well as a table of all 153 isoforms are included in **Paper IV**. Many isoforms expectedly belonged to genes that are known to be induced by training, for example collagens and genes encoding proteins for the mitochondrial oxidative phosphorylation process.

Three months of endurance training also altered the expression of 752 previously unannotated splice variants of already known genes. The differences range from a single base to unannotated exons, and further analysis of these is necessary.

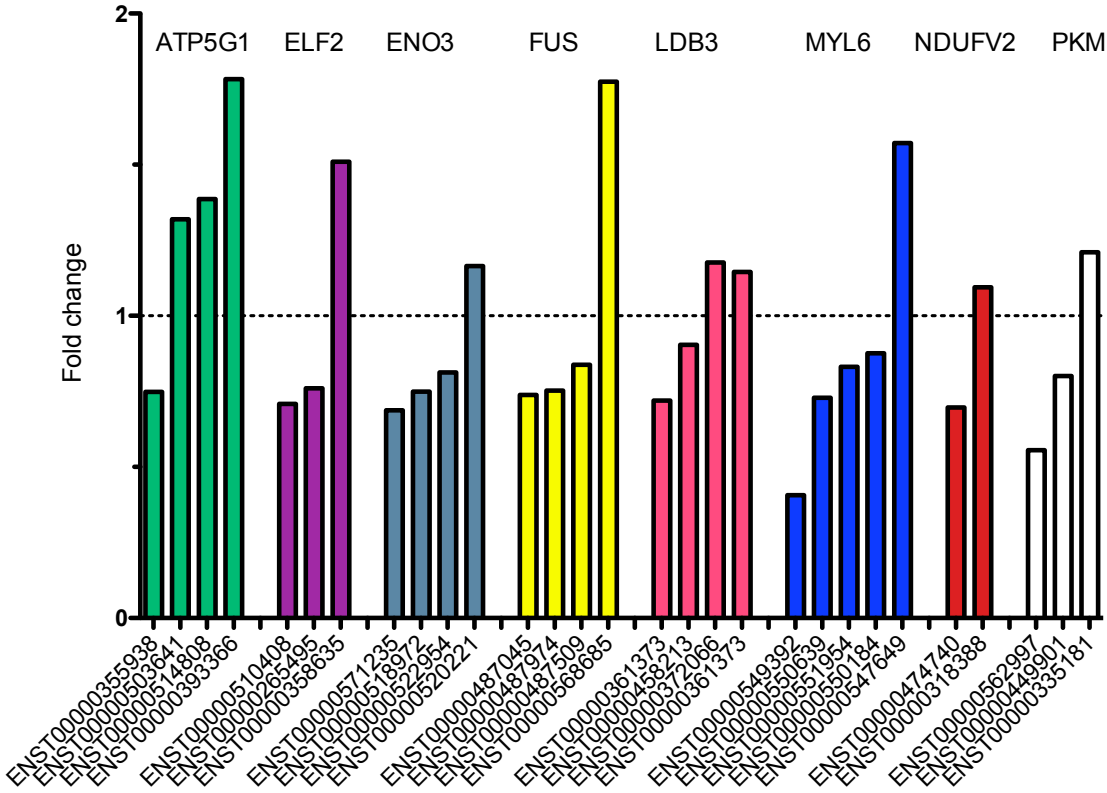

Figure 8. Differentially expressed isoforms. Examples of isoforms that were all significantly differentially expressed with three months endurance training and where isoforms from the same gene changed in different directions. The bars represent the mean fold changes. All displayed isoforms are significant ($P<0.01$).

4.2.3 Novel transcripts

In order to investigate if endurance training could alter expression of transcripts in unannotated parts of the genome, we performed a differential expression analysis (using Limma) of the 2400 novel transcripts identified in **Paper I**. Initially, we found 35 transcripts that were differentially expressed and not annotated in Ensembl v.71. Manual inspection of all 35 transcripts showed that one transcript overlapped an already known gene (*PECAM1*) that was annotated in RefSeq. This gene has previously been shown to be induced by training (Roudier et al., 2013), which was confirmed by our data. Out of the remaining 34 novel

transcripts, nine were upregulated and 25 were downregulated with training. The expression level was low in general, ranging from only four to 178 reads before training. All 34 transcripts contained predicted open reading frames (ORFs) and protein-coding motifs (ranging from two to eleven), suggesting that they have protein-coding potential. A majority of the assembled transcripts also overlapped conserved regions in chimp, rhesus, mouse, rat or dog, indicating that they are functionally interesting. Two transcripts were also identified as enhancer RNAs when compared to a paper by (Arner et al., 2015).

4.3 EPIGENETIC CHANGES WITH ENDURANCE TRAINING

Epigenetics is a very important mechanism in the regulation of gene expression, and there are very few studies that have investigated the impact of epigenetic changes on adaptation to endurance training. The aim of **Papers II** and **III** was therefore to investigate the role of DNA methylation (**Paper II**) and histone modifications (**Paper III**) in skeletal muscle adaptation to long-term endurance training.

4.3.1 Training-induced changes of the methylome

In **Paper II**, the effect of training on the skeletal muscle methylome was investigated using LUMA and Illumina 450K arrays on DNA isolated from skeletal muscle biopsies taken before and after Period 1 of the EpiTrain study. The array specifically investigates the level of methylation at 486 000 sites throughout the genome.

With training, 4919 sites (FDR<0.05) were found to be differentially methylated. The absolute changes were small in general, which was partly expected from this environmental stimulus where observed changes often are small and susceptible to large individual variation in comparison to for example cancer (Feil and Fraga, 2011). Clustering analysis of the data from autosomal chromosomes identified training and gender as the main determinants of variability between the samples. The corresponding analysis was performed on gene expression data, with very similar results. The localization of the differentially methylated positions (DMPs) in relation to the distribution on the array showed that there was a relative enrichment outside of CpG islands/shores/shelves. In relation to genes, the DMPs were primarily present in gene bodies and intergenic regions, and to a lesser degree in promoters. Further analysis based on the Illumina annotation showed significant enrichment of DMPs in regulatory enhancer regions. This was additionally confirmed using independent data from an HSMM cell line. That differential methylation mainly occurred at regulatory enhancer regions is very novel from an exercise biology perspective. However, DNA methylation has previously been shown to be most dynamically regulated in enhancers and other regulatory regions in humans (Ziller et al., 2013). Enhancer methylation has also shown tissue-specific patterns (Xie et al., 2013; Ziller et al., 2013), which is important to consider in the context of studying differential methylation in a mixed tissue sample.

To investigate the potential functional relevance of the DMPs, an ontology analysis of the genes associated to the DMPs was performed. The DMPs were grouped as increasing or decreasing with training. Genes in the vicinity of increasing DMPs were associated to skeletal

muscle processes, highly relevant to endurance training adaptation, *e.g.* contractility, collagen binding and energy metabolism. Decreasing DMPs showed enrichment for developmental processes, transcriptional regulation and inflammation (Figure 9), very different from the increasing DMPs. This indicates that the observed changes were part of a controlled adaptation process rather than a random effect across the genome.

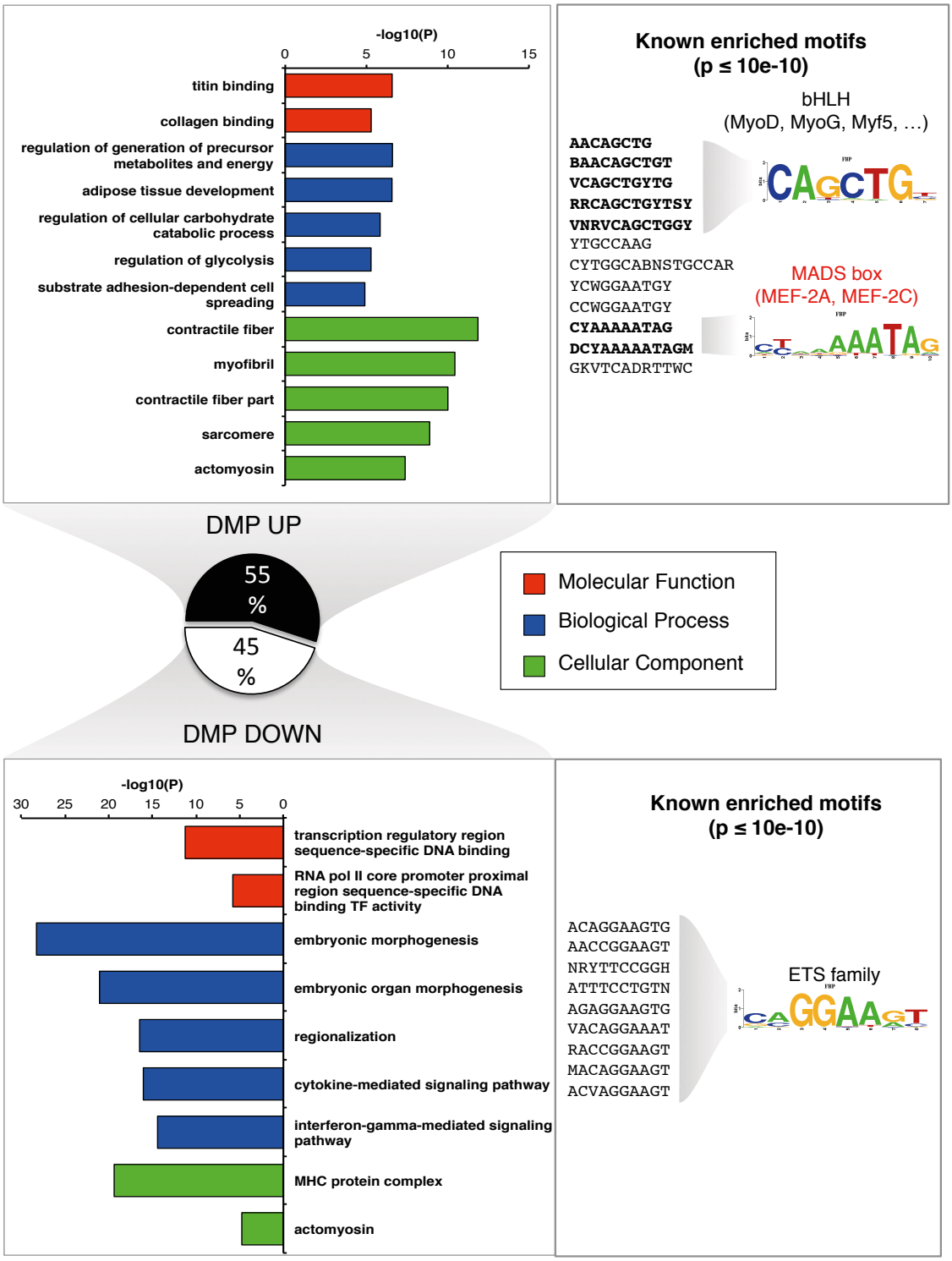

Figure 9. Methylation ontology and motif enrichment. GREAT analysis of genes associated to DMPs increasing (UP) and decreasing (DOWN) in methylation. Enriched motifs were retrieved using a 200bp window surrounding each DMP.

The high degree of DMPs in enhancer regions prompted for additional analyses of the potential presence of known transcriptional motifs in close proximity of the DMPs. For increasing DMPs, motifs for the skeletal muscle-related Myogenic Regulatory Factors (MRFs) and Myocyte Enhancer Factors (MEFs) were significantly enriched. For decreasing DMPs, motifs for the ETS family of transcription factors were enriched (Figure 9). Enhancer methylation has been shown to be very dynamic and more closely associated to gene expression changes compared to promoter methylation (Aran et al., 2013). MRFs (including Myf5, MyoD, Myogenin and Mrf4) are important transcription factors for development of the muscle cell lineage (Gundersen, 2011). They are also highly interesting for endurance training adaptation by regulating expression of PGC1α (Amat et al., 2009) and through a possible role in promotion of a more oxidative muscle. MyoD has been shown to promote slow-to-fast twitch fiber transformation in rats, while myogenin has been implicated in regulation of several mitochondrial enzymes (Gundersen, 2011). Our results were confirmed in a recent endurance training study on mice, where binding sites for MyoD and myogenin were significantly enriched around differentially methylated sites (Kanzleiter et al., 2015). The ETS family of transcription factors has diverse functions, and several members regulate exercise-responsive genes. One example is GABPA (or NRF-2, nuclear respiratory factor 2), which controls gene expression of many mitochondrial genes important for cellular respiration (Fredriksson et al., 2008). Other members of the family control expression of angiogenic factors, and could thereby contribute to the increased capillary density that occurs with endurance training.

To validate the array data, seven selected sites were analyzed using bisulfite pyrosequencing. Six sites were in accordance with the array data. There was no change observed in the untrained leg for any of the sites. We also performed an analysis of unidirectional changes in CpG methylation of global CCGG sites with LUMA. In accordance with the observed array data of approximately similar number of sites increasing and decreasing in methylation, we observed no global unidirectional change with training, nor any difference between the trained and untrained leg. The level of methylation was instead remarkably similar between samples. Global reduction of CCGG methylation with LUMA has previously been observed with an acute bout of endurance exercise in humans (Barres et al., 2012).

Several different approaches were employed to couple changes in DNA methylation with transcription, in order to evaluate potential functional effects. A Pearson correlation of DMPs and DEGs showed that there was both positive and negative correlation, which was related to the position of the DMP. Negative correlation was primarily present in promoters/5'UTR/1st exon while positive correlation was more prominent in gene bodies, in accordance with known effects of methylation on transcription. To link the change in methylation to the change in transcription, we also produced a starburst plot where this is illustrated (Figure 10). Indeed, two-thirds of the DMP/DEG pairs showed an inverse correlation. The methylation data was also laid on top of the transcriptional gene network produced with ARACNE. This data showed that the gene regulatory domain that mainly increased in expression showed mostly decreasing levels of methylation, while the opposite was true for the domain that

showed mostly decreases in expression. Several interesting examples of genes were found among the altered DMP/DEG pairs, many related to the known training-induced changes in structural components, metabolic enzymes and transcriptional regulation. Although the DMPs were not directly correlated to gene expression at isoform level, > 3000 isoforms were found to be differentially expressed with training in the EpiTrain study. DNA methylation changes are highly interesting from this perspective as 22% of alternative exons are regulated by DNA methylation and splice sites are highly methylated in comparison to adjacent introns (Lev Maor et al., 2015).

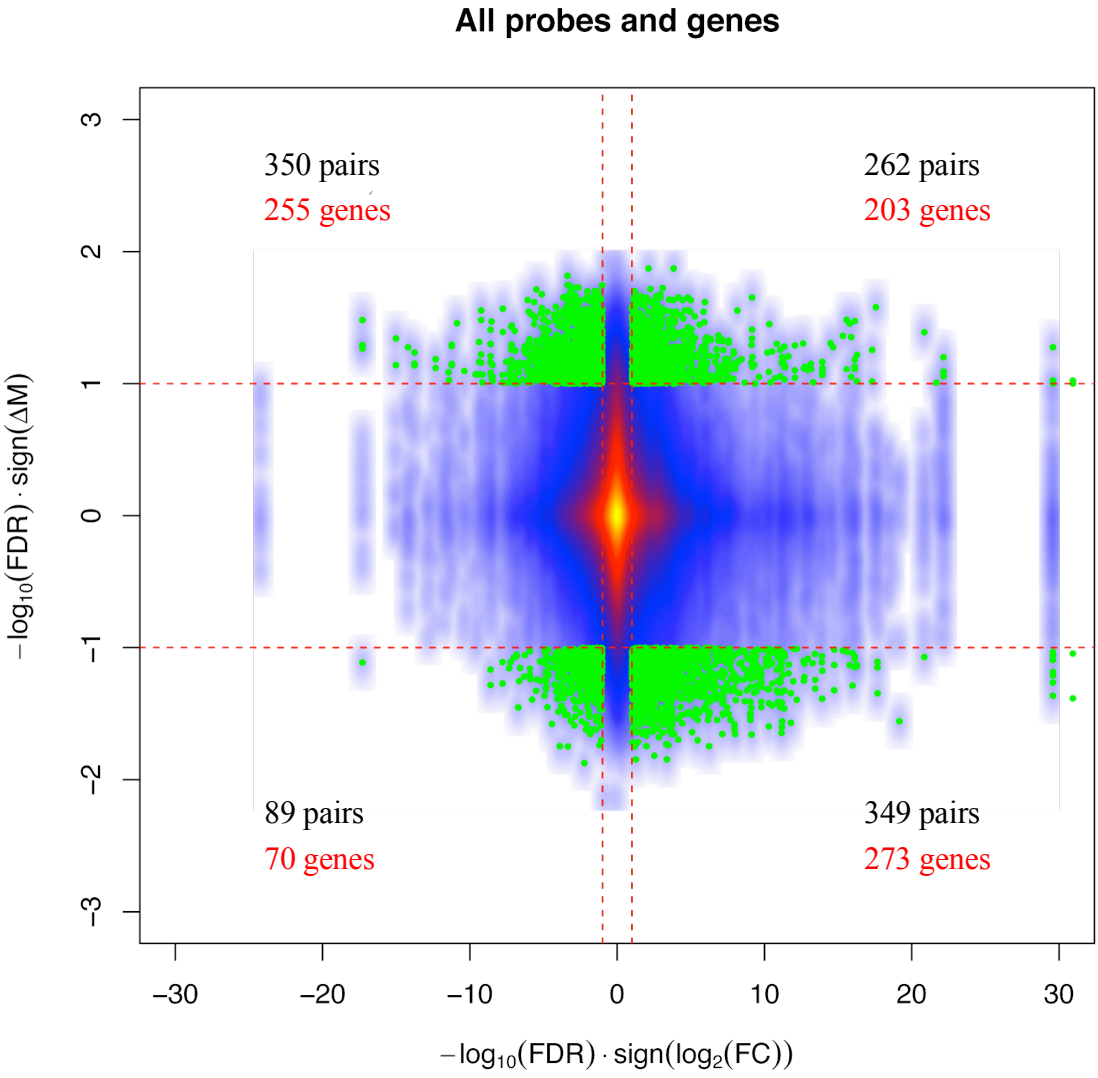

Figure 10. Relationship between methylation and expression changes. A starburst plot illustrating the relationship between gene expression changes (x-axis) and DNA methylation changes (y-axis). Green dots correspond to significant gene expression/methylation pairs.

One other study has investigated DNA methylation changes in human skeletal muscle with exercise training, using an immunoprecipitation-based method (MeDIP) and microarrays (Nitert et al., 2012). Methodological and statistical differences render very different results, but both their and our (**Paper II**) studies show reduced methylation of *RUNX1* and *COL4A1* for example. *THADA* was highlighted by Nitert et al., who observed a reduced methylation in the promoter region, although no change in expression was seen. In contrast, we saw an

increase in methylation of two CpG sites in the gene body, and no change in the promoter region or in gene expression of *THADA*. In total, Nitert et al. found 46 DMP/DEG pairs changing in opposite direction (Nitert et al., 2012), the corresponding figure in our data was 210. The effect of endurance training on DNA methylation has also been studied in human adipose tissue. Six months of endurance training altered almost 18 000 CpG sites in the adipose tissue genome (Rönn et al., 2013). A striking difference in comparison to our findings in skeletal muscle is that over 90% of the DMPs in adipose tissue increased in methylation, while the distribution was 55% increasing DMPs in our study.

The mechanisms behind the observed training-induced changes in DNA methylation remain to be investigated. Altered metabolism is one hypothesis, as changes in metabolites from the citric acid cycle, for example α-ketoglutarate, are known substrates for enzymes catalyzing methylation reactions (Lu and Thompson, 2012). Another potential mechanism is interaction between epigenetic DNMT and TET enzymes and exercise induced transcriptional regulators, like AMPK, NRFs and HIF-1. In mice, DNA-binding factors have been shown to establish lowly methylated active enhancer regions (Stadler et al., 2011) and an increased DNA methylation at regulatory regions has been shown to be due to vacating transcription factors and a subsequent passive introduction of methylated cytosines (Thurman et al., 2012). It is also possible that something in the exercise stimulus can directly activate DNMTs and/or TETs, for example changes in redox status or inflammation. Oxidative stress has been hypothesized to induce TET activity through increased production of α-ketoglutarate via a SIRT-dependent mechanism (Chia et al., 2011). The expression of DNMT1 itself increases with training in this study, as well as in the six-week training study analyzed in **Paper III**.

4.3.2 Global changes of the histone modification H3K27me3

Post-translational modifications of histones are important gene regulatory epigenetic mechanisms that were investigated in **Paper III**. At the global level, six weeks of endurance training induced no change in H3ac, H3K4me3, H3K9ac, H3K9me3 or H3K36me3. However, there was a strong tendency for a reduction of the inhibitory modification H3K27me3 (FC=0.87, p=0.058, Figure 11). Previously, an acute bout of endurance exercise has been shown to increase the global level of H3K36ac in humans (McGee et al., 2009). The study of overall levels of different histone modifications with immunohistochemistry has many limitations and should be interpreted with caution. The size of a skeletal muscle nucleus is, for example, large in relation to thickness of the skeletal muscle section. Skeletal muscle nuclei are approximately 23 μm^2 in a cross-section (Vassilopoulos et al., 1976) and normally have a convoluted shape, although this varies for example with degree of contraction (Franke and Schinko, 1969). The section can vary in thickness and fibers are not perfect cross-sections, which will affect the measureable nuclear volume. This method can at best be considered semi-quantitative. It was, however, used to obtain an indication of global changes for further investigation of gene-specific changes using chromatin immunoprecipitation followed by sequencing, which was performed for H3K27me3. The analysis is currently ongoing and the resulting data will be correlated to gene expression

changes that have already been published for the included subjects (Keller et al., 2011; Timmons et al., 2010).

Trimethylation of H3K27 is a repressive histone mark that has been associated to silent promoters (Barski et al., 2007). Reduced levels of this inhibitory modification in gene promoters and/or enhancers may thus stimulate gene activation in response to training. H3K27me3 is demethylated by the two demethylases JMJD3 (*KDM6B*) and UTX (*KDM6A*) (Swigut and Wysocka, 2007), while the Polycomb-Group proteins (PcG) methylates H3K27 (Grossniklaus and Paro, 2014). Interestingly, several lysine specific demethylases (KDMs) were differentially expressed with training in the EpiTrain study, indicating that training does affect protein methylation levels. Transcription of other histone-modifying enzymes, for example several HDACs and both lysine and arginine methyltransferases was also altered with three months of endurance training. Metabolic perturbations that occur with exercise is important for training adaptation (Egan and Zierath, 2013). Similarly to enzymes regulating DNA methylation, histone lysine demethylases are also α-ketoglutarate-dependent (Tsukada et al., 2006) and can be inhibited by two other metabolites from the citric acid cycle; succinate and fumarate (Xiao et al., 2012).

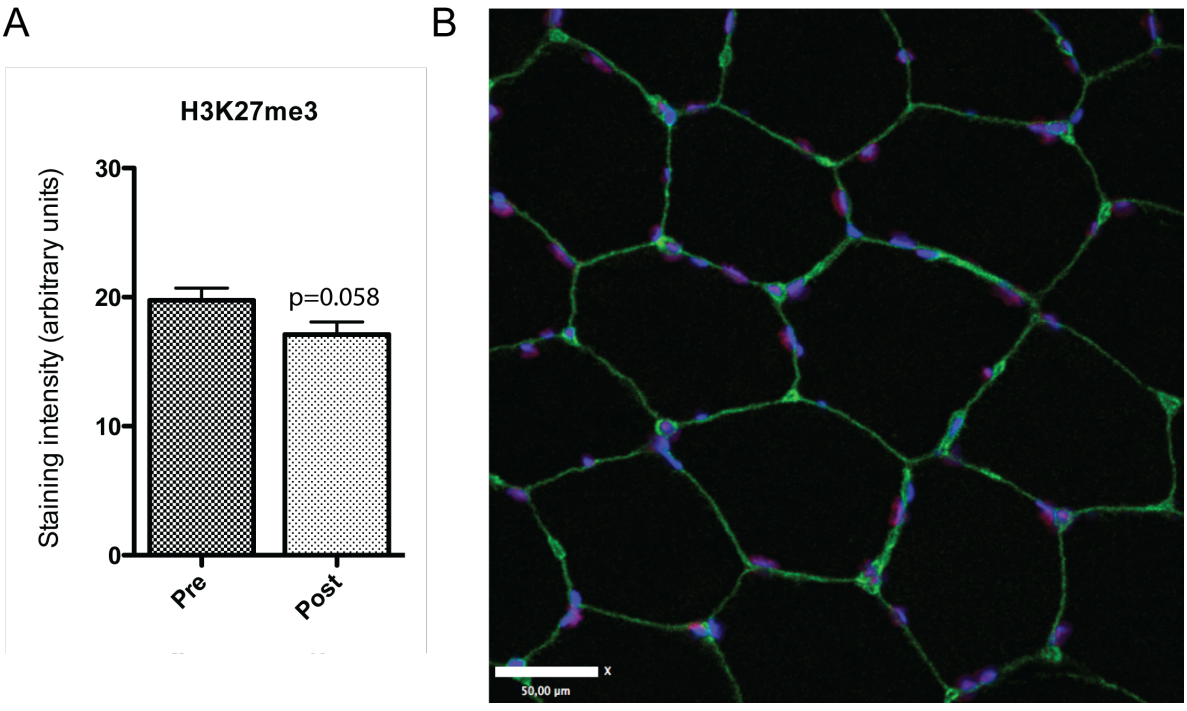

Figure 11. Immunohistochemistry of H3K27me3. A) Semi-quantitative representation of the staining for H3K27me3 before (pre) and after (post) training. Data is presented as mean ± SEM. B) Example of a two-dimensional visualization of the confocal image of a skeletal muscle cross-section with the nuclei in blue, the histone modification in red and basal lamina in green.

The immunohistochemical analysis of histone modifications also included a three-dimensional volume measurement of the nucleus, *i.e.* the DAPI stained DNA. The primary objective was to relate the intensity of the immunohistochemical stainings for the different histone modifications to the nuclear volume. However, an analysis of only the volume before and after six weeks of training showed that the nuclear volume increased by 10%. Despite the

above-mentioned limitations that applies also to this measurement, the difference in volume was consistent for the repeated measurements represented by the separate sections stained for the different histone modifications. To investigate if this could possibly be due to the presence of polyploid skeletal muscle nuclei, a flow cytometry based method was adopted. However, there was no sign of polyploidy. Nuclear volume is tightly controlled and highly correlated to the cell size (Huber and Gerace, 2007), thus an increase in the cytoplasmic volume might promote an increase in nuclear size. Another possible explanation could be an increased protein content in the nuclei, due to for example an increased transcriptional activity, something that remains to be investigated.

The complexity of the overall chromatin structure approaches infinite as the chromosomes show structural stochasticity in the three-dimensional nuclear space and that the structural organization may affect gene activation (Nagano et al., 2013).

4.4 SKELETAL MUSCLE TRANSCRIPTIONAL MEMORY

The purpose of the repeated training period after detraining in the EpiTrain study was to investigate the potential presence of a skeletal muscle transcriptional memory in response to endurance training (**Paper IV**). Epigenetic mechanisms are highly interesting for muscle memory since they have the potential to mark genomic regions for future activation or suppression. Currently, only the transcriptome data is available for analysis of endurance-induced skeletal muscle memory.

4.4.1 Baseline skeletal muscle memory

We first investigated if there were any remaining effects from the previous training period in the previously trained leg before start of Period 2 (T3). Comparing before training (T1) with T3, there were some differences that could be due to time (one year between biopsies), that the subjects were slightly less fit based on VO_2-peak (mean 40 before Period 1 vs 37.5 ml/kg/min before Period 2) or remaining effects of training. Next, we therefore compared T3 with U3 (*i.e.* the untrained leg before start of Period 2), and found no significant differences with the multivariate methods (Q_y^2 was negative) or with paired Limma. The power of this analysis was low, considering that only twelve subjects continued in Period 2. As there were no significant difference between the skeletal muscle performance and enzymatic measures in T3 and U3, we found no indication of a baseline skeletal muscle endurance-induced memory at the physiological or transcriptome level.

4.4.2 Repeated response to training

If a possible memory exists at the epigenetic level, this could potentially influence the transcriptional response to a second training period. During the first training session in Period 2, the rate of perceived exertion (Borg-scale measurement of effort) was smaller in the previously trained leg (Figure 12A) compared to the untrained leg. This could potentially be a neuromuscular or other neurophysiological “memory” effect. It is impossible to blind the training, why this may introduce a bias. Unfortunately, the acute exercise response at the

transcriptome level is not possible to investigate in this study since there was no acute exercise biopsy in association to either Period 1 or 2. The transcriptional response to Period 2 was investigated for both the previously trained (T3 vs T4) and untrained (U3 vs U4) leg. At the global level, the degree of transcriptome differences was similar between the two legs (Figure 12B-C), with 852 genes differentially expressed in T3 vs T4 and 928 genes in U3 vs U4. However, there were specific differences between the responses, as only 180 DEGs were common between the two legs. Ingenuity analysis of the training-responsive genes showed that the main pathways induced were oxidative phosphorylation and mitochondria for both legs. In the previously untrained leg, EIF2 and mTOR signaling were downregulated, which was not the case for the previously trained leg. Yet, these differences in response were not significant enough to introduce global differences at the end of Period 2 (T4 vs U4), as no separation was seen in a PCA and the model Q_y^2 of the OPLS was negative.

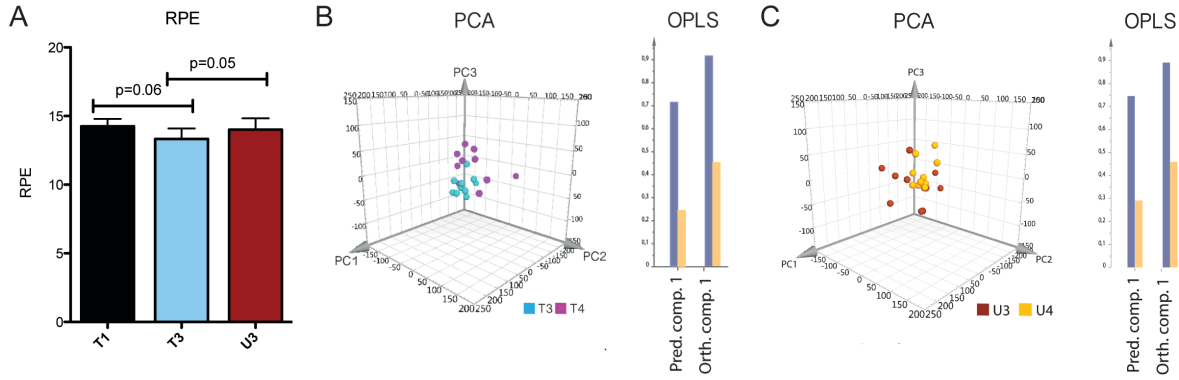

Figure 12. Response to repeated training. A) The rate of perceived exertion (RPE) in the exercising leg during the first training session in Period 1 (T1) and Period 2 in the previously trained (T3) and untrained (U3) legs. Data is presented as mean ± SEM. The presented p-values are based on Student's *t*-test. B-C) PCA and OPLS analyses of the training response of the transcriptome to Period 2 of the EpiTrain study in the previously trained (B) and untrained (C) leg.

In a separate, univariate approach we also specifically analyzed the DEGs in Period 1 (using Limma, in total 2203 genes) to see if those genes also changed in Period 2 and if there were any differences between the legs. Limma identified 465 DEGs in T3 vs T4 and 398 DEGs in U3 vs U4, with 150 genes in common between the two legs. Although the number of common DEGs is still low, the correlation between the fold changes between Periods 1 and 2 was very high (Pearson's r=0.96 for both legs). This shows that the training response still was very consistent, although differences definitely exist. Whether this is due to natural biological variability or some type of memory mechanism is not possible to determine with this experimental setup, but the current evidence does not support the presence of a differential response due to a skeletal muscle memory effect.

4.5 THE HYPOXIA INDUCIBLE FACTOR 1 SYSTEM IN TRAINING ADAPTATION

At the protein level, which is functionally most important, transcription factors are essential transcriptional regulators. Cellular responses to hypoxia are to a large degree conferred by the

transcription factor HIF-1 (Stroka et al., 2001), which is activated in response to an acute bout of exercise (Ameln et al., 2005). HIF-1 was early on suggested to be an interesting factor from an endurance training adaptation perspective (Hoppeler and Vogt, 2001) as several studies have shown an increased adaptation when training is performed with restricted blood flow or at high altitude (Hoppeler and Vogt, 2001; Kaijser et al., 1990; Sundberg et al., 1993). However, based on observations indicating that HIF-1 effects on skeletal muscle metabolism are in some ways opposite to those of long-term endurance training, together with an attenuation of HIF-1 regulated genes with an acute exercise bout after training, we hypothesized that training attenuates HIF-1 activity by inducing some of its negative regulators (**Paper V**).

The results showed that PHD2 was the most abundant prolyl hydroxylase in human skeletal muscle, and together with PHD3, it was higher in elite athletes compared to moderately active individuals for mRNA (Figure 13). Protein levels of PHD2 were also higher in the elite group. FIH and SIRT6, two other negative regulators, were also higher in elite athletes at both mRNA (Figure 13) and protein levels, while expression of the HIF-response gene PDK-1 was lower. A longitudinal study with six weeks of endurance training also showed similar results with training. The HIF-1 inhibition was thus present at several different regulatory levels, with the prolyl hydroxylases marking HIF for proteasomal degradation (Ivan et al., 2001; Jaakkola et al., 2001), FIH inhibiting binding to the coactivator CBP/P300 (Mahon et al., 2001) and SIRT6 that deacetylates H3K9 on glycolytic response genes of HIF (Zhong et al., 2010).

The effect of HIF on performance has been studied in humans suffering from the autosomal recessive disorder Chuvash polycythemia. These patients carry a mutation in the VHL gene, leading to a general stabilization and thus elevated activity of HIF-1. In response to exercise, the patients experience an altered skeletal muscle metabolism with a reduced muscle pH and increased serum lactate accumulation, as well as a reduced maximal exercise performance.

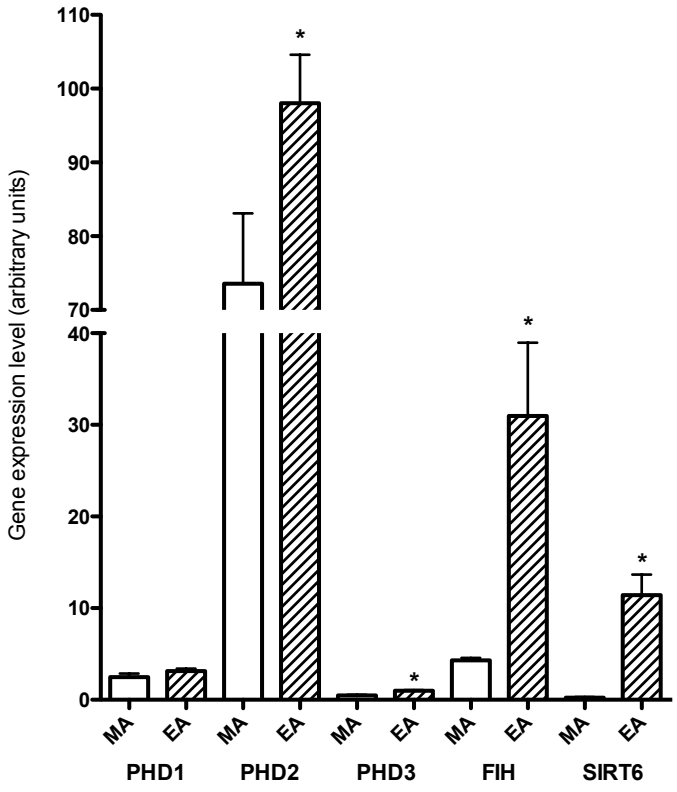

Figure 13. Higher negative regulators of HIF in elite athletes. mRNA expression in moderately active individuals (MA) compared to elite athletes (EA). Data is presented as mean ± SEM. * indicates significance at $P<0.05$.

Transcriptional analyses of skeletal muscle showed enhanced levels of the HIF-1 targets PFK-1 and PDK-1, indicating high glycolytic activity (Formenti et al., 2010). In humans, there are also different variants of the *HIF1A* gene. One SNP (P582S) changes the amino acid sequence, where presence of a T-allele-associated HIF-1 protein has shown a reduced degradation and thus increased activation of HIF-1 (Tanimoto et al., 2003). Twenty-four weeks of endurance training induces lower changes of VO_2-peak in elderly humans with the T-allele compared to those with the CC genotype (Prior et al., 2003). The opposite has, however, been observed in young women where the T-allele was associated to greater gains in VO_2-peak (McPhee et al., 2011), although that study was shorter (six weeks).

The functional role of a protein is otherwise more easily studied in KO animals. Mice with a muscle-specific deletion of HIF-1α actually show higher endurance compared to their normal littermates, and higher activities of CS and ß-HAD, concordant with a better aerobic capacity. The improved mitochondrial function in the KO mice was suggested to be due to activation of the pyruvate dehydrogenase complex (PDHc) through its HIF-1 inducible inhibitor PDK-1 (Mason et al., 2004; Mason et al., 2007a). HIF-1 regulates mitochondrial oxygen consumption during hypoxia through increased expression of PDK-1 (Kim et al., 2006), which then inhibits PDHc and provides an active repression of mitochondrial function to reduce the cellular oxygen consumption. Thus, a lower HIF-1 activity seems to be beneficial for endurance performance, and in humans, skeletal muscle appears to adapt to long-term endurance training by inducing the negative regulators of the HIF system.

To mechanistically establish a link between an altered prolyl hydroxylase activity and PDK-1 mRNA levels in muscle, human primary myoblasts were treated with the prolyl hydroxylase inhibitor DMOG. This resulted in a high induction of PDK-1 in the cells, which would be expected from an increased HIF activity when the negative regulator was inhibited. To study the opposite effect of a reduced HIF-1 activity on metabolism, we silenced the response gene PDK-1 in C2C12 cells. This resulted in an increase in CS activity, while treatment with DMOG reduced CS and concurrently also increased the lactate production (as measured in the cell media), indicating a higher glycolytic activity. The reason for not measuring HIF-1 directly in the human subjects is that the protein is continuously degraded and only stabilized in response to acute hypoxia or exercise. In biopsies taken at rest, as in these studies, it is therefore not possible to measure HIF.

The hypothesis of a moderation of the HIF response in a trained muscle is supported by recent data from a rat study where acute exercise decreased the nuclear accumulation and DNA binding activity of HIF-1α, as well as the expression of VEGF and eNOS after training (Rodriguez-Miguelez et al., 2015). Tibetans, who live at high altitude where they have been exposed to chronic hypoxia for generations, have a very high frequency of a variant of PHD2 with a higher affinity for oxygen (Lorenzo et al., 2014). This would prevent HIF-1α stabilization also under mild hypoxia, likely resulting in a similar effect as that expected from an increased expression of the HIF negative regulators.

The studied factors are not the only factors regulating HIF-1 activity. The N(Alpha)-Acetyltransferase 10 (NAA10), for example, acetylates a lysine in HIF-1, which results in an accelerated degradation during normoxia (Jeong et al., 2002). In relation to epigenetics, regulation of histone modifications is highly associated to HIF-1. HDAC7 binds directly to HIF-1α (Kato et al., 2004), as does HDACs 4 and 5 that inhibit binding to FIH, which results in an increased transactivation capacity of HIF (Seo et al., 2009). Cell experiments have demonstrated that hypoxia increases global histone methylation. HIF-1 has been suggested to compensate for this by inducing expression of several known histone demethylases (*e.g.* JARID1B, JMJD1A, JMJD2B, and JMJD2C) to maintain histone methylation homeostasis (Pollard et al., 2008; Xia et al., 2009). In addition, HIF-1 regulation of the histone demethylase JMJD1A has been shown to be important for tumor growth (Krieg et al., 2010). However, the effect of HIF-1 on chromatin regulation in skeletal muscle is unknown as are the effects of a possible attenuation of HIF with long-term training on chromatin regulation. These issues are highly interesting and remain to be investigated.

In addition to its effects on angiogenesis and metabolism, regulation of glucose transport in skeletal muscle has also been attributed to HIF-1α activity (Sakagami et al., 2014). Hyperglycaemia, which is an important feature of type II diabetes, inhibits hypoxia-induced stabilization of HIF-1 (Catrina et al., 2004), suggesting that this protein is also highly relevant in glucose control.

4.6 GENERAL DISCUSSION

Skeletal muscle is a heterogeneous tissue. Whole tissue samples include skeletal muscle fibers and satellite cells, as well as other cell types, for example endothelial cells, fibroblasts and immune cells. The transcriptome and methylome analyzed in this thesis represent a mixed human tissue, not only skeletal muscle cells (myocytes). The analyzed samples also represent a snap shot of the skeletal muscle biology, taken 24h after the last training bout for the intervention studies or at rest for the cross-sectional study and before the interventions. Gene transcription is known to occur stochastically between individual cells, and this occurs for example for metabolic gene networks where cells stochastically activate or deactivate metabolic pathways, perhaps to be prepared for nutrient arrival (Raj and van Oudenaarden, 2008). The stochasticity of gene activity will inevitably affect the results, although the number of subjects and high number of cells in a tissue sample aim to reduce the impact of stochastic effects.

Gene expression data was the basis for the majority of the studies in this thesis. Quantification of gene expression changes, regardless of method, is not synonymous with a concurrent functional change in skeletal muscle. Specific mRNA levels do not always correlate with protein levels, although there is a high overall correlation (Vogel and Marcotte, 2012), and post-translational modifications of proteins are also highly important for the activity of individual factors.

A major limitation in studying training-induced skeletal muscle responses in humans is the large intra- and interindividual variability that exists in comparison to model organisms. Genetic differences, lifestyle choices, psychological factors and medical status may all affect the investigated phenomena. This is something that needs to be considered before any general conclusions can be drawn from the data. Also, human studies are limited in their descriptive nature and need to be complemented with animal models and cell systems to establish direct mechanistic relationships. Nevertheless, there are evident advantages with studying humans. For animal studies and *in vitro* experiments, results are commonly extrapolated to the human situation, which is sometimes not correct. If humans are the intended species of investigation, we should study humans whenever possible from a methodological and ethical perspective.

5 CONCLUSIONS

The main findings from this thesis were;

- The baseline transcriptome in human skeletal muscle was remarkably homogenous both within a muscle and in the corresponding contralateral leg of an individual. Furthermore, the transcriptome difference between male and female skeletal muscle was very high.
- That 23 000 isoforms were expressed in skeletal muscle at baseline together with almost 2500 novel transcripts, out of which at least five were protein-coding, and how these isoforms and novel transcripts changed in expression with three months endurance training.
- Three months endurance training induced changes in DNA methylation across the human skeletal muscle genome that were associated to functionally relevant transcriptional changes. Many changes occured in regulatory enhancer regions.
- Trimethylation of histone 3 lysine 27 was likely reduced at a global level with endurance training, showing that histone modifications are involved in training adaptation.
- There was no clear evidence of an endurance-induced skeletal muscle memory at the transcriptome level after nine months of detraining or in response to a repeated training period.
- The HIF-1 response was attenuated in response to long-term endurance training due to activation of its negative regulators.

Collectively, the results of the present thesis show that endurance exercise training can induce associated changes in the epigenome and transcriptome of human skeletal muscle. The data provides an in-depth analysis of the human skeletal muscle transcriptome at baseline, including sex-specific differences with potential implications for function, and how gene expression changes in response to repeated endurance training periods. With the present analyses, there was no detectable muscle memory of previous training at the transcriptome level. In all, the results contribute to a better understanding of the molecular pathways involved in physiological adaptation to endurance training and can provide a basis for how training prevents disease development and different dysfunctions.

6 FUTURE PERSPECTIVES

We are born to move and it has even been argued that, in order to maintain health, we need a certain amount of physical activity to reach a high enough expression level of certain genes that evolved during a time when we were required to be much more physically active than today (Booth et al., 2002). The mechanisms behind training adaptation and how genes respond to environmental lifestyle interventions have been the focus of intensive research for many years, but despite substantial progress, they are still not completely understood. The mechanisms of exercise will most certainly continue to be of great scientific and public interest. The studies in this thesis have contributed to a part of the picture, but more studies are obviously needed.

The DNA methylation and histone modification responses to endurance training shown in this thesis can hopefully serve as one starting point for future studies on the mechanisms behind these training-induced epigenetic changes. Future studies should focus on identifying the specific factors responsible for inducing changes in DNA methylation and histone modifications. It is not understood whether changes in transcription drive methylation changes or vice versa. Therefore, the studies herein need to be complemented with mechanistic studies to functionally link changes in epigenetic modifications and concurrent transcriptional regulation in training adaptation.

Transcriptional activity and epigenetic patterns are cell-type specific and most likely also cell-specific. Future studies would therefore benefit from a single-cell approach to elucidate the cell-to-cell variability in expression and its epigenetic regulation. From an exercise physiology perspective, it would be highly interesting to compare the transcriptome and epigenome between the different fiber types and how these are affected by different training modalities. From a skeletal muscle biology perspective, an investigation of the epigenetic pattern in individual skeletal muscle nuclei from the same fiber would also provide valuable information on the natural variability in human biology. The inter- and intraindividual response variability poses a great challenge scientifically, but is also one of the most fascinating and important elements of studying human physiology.

Exercise genomics has been of great interest for many years (Loos et al., 2015), but exercise epigenomics is still in its infancy. There are some studies on miRNAs, very few on DNA methylation and studies of histone modifications are sparse. 5-hydroxymethylcytosine (5-hmC) is considered to be the sixth base of DNA and has only been mentioned briefly in this thesis. However, mechanistic studies regarding its role in gene regulation have emerged in recent years, for example through the presence of TET enzymes in enhancer regions (Stadler et al., 2011). 5-hmC is likely involved also in endurance training adaptation considering its role in gene activation.

In sports, an optimal physical performance is of interest and many hope that epigenetics will be able to bridge the gap between the genotype-phenotype interaction (Ehlert et al., 2013).

Although most predictions of athletic maximal performance potential based on genetic and epigenetic information will likely never reach significance due to enormous biological complexity, the impact of the genome and epigenome on performance and performance change is presumably substantial. Current and future studies on epigenetic changes with training will also be of interest for performance improvement purposes, and may therefore be relevant in development of anti-doping strategies. This could also be of interest for the negative regulators of HIF, investigated in this thesis.

Epigenetic memory mechanisms have been studied in relation to inheritance, *i.e.* that epigenetic marks acquired from environmental exposures can be passed on to the next generation, enabling multigenerational influences on inheritance and phenotype. This is a relatively common phenomenon in plants (Reik, 2007) and likely also occurs in humans (Pembrey et al., 2014). It would be intriguing if the effects of exercise could be conveyed to the next generation, and a recent study actually shows that three months of sprint interval training changes the methylation pattern of sperm cells in humans (Denham et al., 2015). Future studies will most likely develop these initial findings.

Endurance exercise effects on skeletal muscle are fundamental for the health benefits of training, which includes prevention of many of our most common chronic diseases, including obesity, type II diabetes and cardiovascular disease. The prevalence of these diseases is increasing at a scary pace, which in combination with a growing and aging population is a huge challenge for society. Exercise training prevents disease, and can also be a cheap and easily accessible treatment (or part thereof) for numerous different diseases, including metabolic diseases, cardiovascular disease and even cancer. In this respect, exercise is a medicine for the future.

7 POPULÄRVETENSKAPLIG SAMMANFATTNING

Gener och deras aktivitet

Människans kropp består av tusentals miljarder enskilda celler som samverkar för att vi ska fungera som en organism. Grunden för människan som livsform ligger i arvsmassans DNA. Dess specifika sammansättning bestäms då ägget och spermien smälter samman med hälften av arvsmassan från mamma och hälften från pappa. Från denna embryonala stamcell utvecklas en mängd olika celler och celltyper som bygger upp en människa. Alla celler innehåller samma genetiska material, samma DNA, men har ändå olika funktion och utseende. Det beror på skillnader i hur de ca 20 000 gener som finns DNA:t används. Generna uttrycks (kopieras) till en molekyl som kallas mRNA som sedan kan översättas till ett funktionellt protein i cellerna. Något förenklat kan man säga att varje gen kodar för ett enskilt protein, och det är proteinerna som står för majoriteten av kroppens olika funktioner. Våra proteiner ger t.ex. struktur, producerar rörelse och fungerar som biologiska katalysatorer och markörer. Genom att olika gener uttrycks i olika celler så gör det att celler har varierande funktion. Det gör också att gener kan svara på förändringar i vår miljö genom att öka eller minska sitt uttryck och därmed produktion av proteiner. En av de mekanismer som styr hur mycket gener uttrycks kallas för epigenetik, vilket betyder ovanför genetiken. Epigenetiken påverkar hur lättillgängliga olika gener är för uttryck och därmed hur mycket protein som kan produceras från varje gen. Förutom våra gener så uttrycks också andra delar av arvsmassan, som inte kodar för proteiner men ändå bildar RNA. Dessa kan till stor del påverka hur mycket generna uttrycks och därmed hur mycket funktionella proteiner som bildas.

Muskler och träning

Muskel är den av kroppens vävnader som är specialiserad på rörelse. Skelettmuskler är ansvariga för alla våra viljemässigt styrda rörelser, vilket bl.a. inkluderar allt kroppsspråk, att andas, springa äta och tala. En frisk människa består till ca 40-45% av skelettmuskulatur, vilket innebär att det är vårt viktmässigt största organ. Skelettmusklerna är också ett av kroppens mest anpassningsbara organ. Genom träning anpassar sig musklerna till den typ av aktivitet som utförs. I denna avhandling ligger fokus på uthållighetsträning, som leder till att musklernas förmåga att utvinna energi med hjälp av syrgas förbättras och att vi kan arbeta längre och hårdare. Musklerna får t.ex. fler små blodkärl som ökar transporten av syre, de får fler mitokondrier som är muskelcellernas energiproducenter, och de ökar sin inlagring av näringsämnen som kolhydrater och fettsyror. Tillsammans med den anpassning som sker i andra organ förbättras kroppens funktion och hälsa avsevärt, vi minskar risken att utveckla en mängd olika sjukdomar och till och med risken för förtidig död. Syftet med den här avhandlingen var att studera vilka förändringar i geners uttryck som är viktiga för anpassning till träning och om det finns någon form av muskelminne av tidigare träning som kan påverka hur vi svarar på upprepad träning.

Muskelbiopsier

En stor fördel med att studera just muskler är att den är tillgänglig för provtagning i människor. Alla studier som ligger till grund för denna avhandling inkluderar muskelbiopsier före och efter uthållighetsträning, alternativt från vältränade jämfört med normalt tränade, friska frivilliga försökspersoner. Muskelbiopsier är små bitar, på ca 100 mg (vilket är mindre än en tiondels tusendel av lårets sträckarmuskels vikt), som tas ut med hjälp av en nål. Huden lokalbedövas först och därefter läggs ett litet snitt ner till muskelns yttre hinna, den så kallade muskelfascian. Nålen förs därefter ner genom snittet och in i muskeln. Genom att ett undertryck skapas utifrån med hjälp av en sug gör det att en liten bit av muskeln sugs in i nålen och kan kapas av. Alla biopsier för dessa studier har tagits från den yttre lårmuskeln, *m. Vastus lateralis.* Med hjälp av modern teknik och avancerad statistik har sedan nästan allt RNA i muskelbiopsierna kunnat analyseras för att studera hur detta förändras med träning. Även epigenetiska förändringar har kunnat analyseras på en mängd olika ställen i musklernas arvsmassa.

Resultat från studierna

Resultaten från dessa studier visar att epigenetiken, på såväl DNA:t som på de proteiner (histoner) som DNA:t finns upprullat på, påverkas av tre månaders, respektive sex veckors träning hos människor. Förändringar på DNA kallas DNA-metylering, vilket innebär att en metylgrupp (som består av en kolatom och tre väteatomer) sätts på eller tas bort från DNA. Tre månaders träning förändrade DNA-metylering på nästan 5000 olika platser i försökspersonernas DNA och många av dessa förändringar kunde kopplas till skillnader i olika geners uttryck. Det visade sig också att många av förändringarna skedde i delar av DNA som är specialiserat för att reglera geners uttryck, så kallade enhancers. Processer som påverkades av förändring i DNA-metylering med träning var relaterade till muskelns förmåga att dra ihop sig, inflammation, samt reglering av geners kopieringhastighet. Preliminära resultat tyder också på att histonerna kan påverkas av träning så att DNA sitter hårdare eller lösare, vilket är viktigt för hur mycket en gen kan uttryckas. En gen som sitter lösare bunden är lättare tillgänglig för uttryck och därmed för potentiell produktion av dess protein. Ett specifikt protein som undersöktes extra noga var HIF som står för hypoxi-inducerad faktor. Tidigare studier har visat att HIF aktiveras vid träning hos otränade och stimulerar då sannolikt till bildandet av nya blodkärl och större mängd syretransporterande röda blodkroppar. Mycket tyder dock på att HIF hindrar mitokondriernas aktivitet, vilket inte är bra för anpassning till en längre tids träning. Själva HIF-proteinet är svårt att mäta, men de proteiner som reglerar HIF visade sig vara mycket högre hos vältränade personer jämfört med normaltränade och dessutom ökade proteinerna med sex veckors träning. De reglerande proteinerna hindrar olika delar av HIFs normala verkan. Det gör att resultaten tyder på att HIFs verkan dämpas med träning, troligtvis för att tillåta en större syreberoende (aerob) energiproduktion.

En gen kan uttryckas till lite olika varianter, så kallade isoformer. Dessa studier visar för första gången hur en stor mängd olika isoformer förändrades med träning. Det fanns till och

med gener med isoformer som reglerades i olika riktning, d.v.s. isoformer från samma gen som uttrycktes mer respektive mindre med träning. Vi hittade även 34 tidigare okända uttryck av DNA som förändrades med träning. Funktionen för dessa är dock okänd, men mycket tyder på att några av dem kan koda för funktionella proteiner.

Ett viktigt syfte med avhandlingen var också att studera om det fanns någon form av muskelminne som kvarstod efter ett längre träningsuppehåll och i så fall kunde påverka hur anpassningen till en andra träningsperiod såg ut. I den träningsstudie som bestod av tre månaders träning tränades först bara det ena benet och efter nio månaders uppehåll tränades båda benen i ytterligare tre månader, vilket innebar att ett ben tidigare var vältränat och ett ben tidigare otränat. Vi kunde dock inte hitta några tydliga tecken på att uthållighetsträning kunde ge kvarvarande effekter. Även om det fanns vissa skillnader i hur ett tidigare vältränat ben respektive otränat ben svarade på en andra träningsperiod kunde vi inte tillskriva detta till ett muskelminne.

En av studierna visade också att det grundläggande genuttrycket i skelettmuskler var väldigt lika mellan muskelprover tagna från en och samma muskel, samt från vänster jämfört med höger ben. Däremot var skillnaderna mellan män och kvinnor stora, där kvinnor verkar ha mer uthålliga muskler och män ha större omsättning av proteiner.

Framtida användningsområden

Resultaten från dessa studier bidrar till ny grundläggande kunskap om vad som händer med geners uttryck i skelettmuskler vid uthållighetsträning och hur det grundläggande genuttrycket ser ut hos unga män och kvinnor. Resultaten kan i förlängningen bidra till utvecklingen av individualiserade träningsprogram för att optimera träningseffekter i olika personer. De kan också vara värdefulla för utvecklingen av framtida behandlingar mot t.ex. metabola sjukdomar och för att på konstgjord väg kunna få fram effekter liknande de som fås av träning hos personer som av olika anledningar inte kan utföra fysisk aktivitet, t.ex. på grund av rörelsehinder.

8 ACKNOWLEDGEMENTS

First, I would like to thank all the fantastic **volunteers** who have been generous and brave enough to provide their time and muscle for these studies. Without you, none of this would have been possible. You are simply the best.

I wish to express my sincere gratitude to everyone that has contributed to this thesis. Especially, I wish to thank;

My main supervisor, **Carl Johan Sundberg** for providing an inspiring and creative work environment. Your enthusiasm and generosity is unbeatable. I will always be grateful for the faith you have put in me, and the confidence that has given me over these years. Thank you so much for the great experience of doing a PhD in your group.

My fantastic co-supervisors:
Tomas Ekström. Your dedication and commitment to science is inspiring. Thank you for all the great input and for always being there for meetings, often the same day that I've asked. Although I know you would rather be sailing some times ☺
Helene Rundqvist. You are the most hard-working scientist I know. Your enthusiasm and love for science has motivated me so many times. I have mostly left our many meetings excited about all the extra work hours ahead. I will always appreciate your many great ideas and opinions, whether it has been over a cell culture flask, coffee or red wine. Without you, I would have given up a long time ago.
Helene Fischer. For your great willingness to help and for always taking the time to talk about everything, from single results to the whole picture. I am so grateful that you have been present through both success and failure all these years, without ever giving up on me.

All my co-authors;
especially **Francesco Marabita** and **David Gomez-Cabrero** for introducing me to the power of R and an endless array of bioinformatic methods. Thank you for patiently answering all my questions, even though they sometimes have been the same again and again. You have brought this thesis to another level. I want to thank **Mikael Huss** for your genuine interest in our data, for always being positive towards new ideas and believing that anything is possible. It has been a true pleasure working with such an excellent scientist. To **Thomas Gustafsson**, for always fighting for what is best for science and for your genuine interest and willingness to help. To **Stefania Giacomello**, for a great collaboration with many fruitful and interesting discussions. Different perspectives are always good ☺ To **Sanela Kjellqvist**, your intellect and drive is something I really look up to. I hope we will work together also in the future. To **Beata Werne Solnestam** for your great efforts, for always finding time and not least for baby talk over fun lunches. To **Randall Johnson**, for inviting me to your lab at UCSD. To **Anita Öst** for a great collaboration in Linköping. For valuable contributions to these papers; **Jesper Tegnér**, **Joakim Lundeberg**, **Lorenz Poellinger**, **Amarjit Saini**, **Gert Auer**, **Juan Castro** and **Mattias Alenius.**

Great colleagues and friends;
especially my sister in arms, **Susanna Appel**. Thank you so much for making it fun to go to work every day. For great scientific discussions, for being there for good and bad, for great laughs, and not least for making my wardrobe as well as my personality more colourful. I hope we will always be friends; to **Eva-Karin Gidlund**, the most gorgeous woman in FyFa. Your commitment to science and teaching is extraordinary and I really appreciate everything we have done together, at work and outside; to **Mia Ydfors**, for being a great friend by always speaking your mind, for intriguing statistical discussions and for being my constant computer support. I hope we will work together sometime in the future! To **Charlotte Schönbeck**, the best and most caring runner and "alltiallo" I know ☺ Thank you for keeping the whole lab organized despite my sometimes confused wishes; to **Anna Strömberg**, for endless intellectual and less intellectual discussions, for keeping me sane when things go crazy and for always being up for a glass of red wine and some gossip. You excel at everything you do; to **Sara Mijwel** for bringing another perspective on exercise science to the group, I hope to get to know you more during the spring ☺; to **Jessica Norrbom**, for being the boss. I think you are the organizational glue that keeps the group together; to the newcomers **Maike Winters**, **Mark Chapman** and **Stefan Reitzner**.

Past and present colleagues at FyFa;
the previous corridor pals **Lars Karlsson** and **Ellinor Kenne** for sharing your vast experience and for all the fun times! **Ferdinand von Walden**, it was great sharing lab with you! To **Anne van der Does** for organizing OLM and parties, and for all the contagious loud laughs. FyFa is a bit too quiet now; the rest of the Lindbom gang; **Joel Rasmusson** and **Erik Lundeberg**; the other people at FyFa who have understood that muscle is the best tissue: **Håkan Westerblad**, **Johanna Lanner**, **Niklas Ivarsson** and **Rodrigo Fernandez Gonzalo**; the fantastic senior scientists; **Lennart Lindbom**, **Bo Rydqvist**, **Mats Rundgren**, **Lars Gustafsson**, **Dag Linnarsson** and **Jan Lännergren** for countless interesting lunch conversations; to **Stefan Eriksson** for providing a great work environment for everyone; the administrative stars **Sofia Pettersson** and **Peter Wolf** (I think I won the bet ☺); the financial administration, especially **Camilla Fors-Holmberg** and **Freddie Hellström**; **Marie-Louise Sundin** for all your help with courses; **Robert Frithiof**, **Iréne Lund**, **Olivia Frånberg**, **Björn Schilström**, **Åsa Konradsson-Geuken**, **Gunnar Schulte**, **Jacomijn Dijksterhuis**, **Michaela Kilander**, **Magdalena Kegel**, **Lena Norrbrand**, **Micke Elm**, **Daniel Andersson**.

Past and present members of the "muscle lab" at Labmed; **Eva Jansson** for running the whole lab, your vast knowledge in exercise physiology is amazing; **Håkan Rundqvist** for great pep talk at the gym; **Seher Alam** for cell culture advice; **Erik Rullman**, **Tommy Lundberg**, **Mats Lilja**, **Ted Österlund**, **Karl Olsson**, **Barbara Norman**, **Mona Esbjörnsson**, **Anna Wiik**, **Maria Westerståhl**, **Helena Wallin**, **Elisabeth Berg**, **Majken Lidén**, **Andreas Montelius**.

All my wonderful friends – I hope you know who you are and what you mean to me.

To all the members of the **Klintberg**, **Stendahl** and **Lindholm** families for great support over the years, especially **Andreas** for making me believe that everything is possible and always inspiring me to dare to do new things; **Niklas** for caring so much about everyone and making me appreciate the small things in life. I think it's great that you're also pursuing a career in medicine! **Mamma**, for all our interesting and challenging discussions that help me argue for my case and **Pappa,** for all the fun trips, homework help and driving me everywhere. The list is endless ☺ Thank you both for never doubting me and always being there, no matter what has happened. You have no idea how much it means to my family and me. Thank you for your unconditional love and support. None of this would have been possible without you. **Mormor**, for encouraging me to pursue academic studies and always do my best. I still miss you so much.

Peter. For all your love and encouraging support throughout the years, for always knowing how to improve my mood when things were bad and for the wonderful life we share ❤

Valter och **Filip**. För att ni visar vad som verkligen är viktigt i livet. För allt bus och skoj. För all kärlek. Ni är det bästa i mitt liv och det finns inga ord som kan beskriva hur mycket jag älskar er ❤

The studies included in this thesis were supported by grants from the Swedish National Center for Research in Sports, the Karolinska Institutet PhD research funds (KID), Karolinska Institutet Physiology and Pharmacology funds, the Wallenberg Advanced Bioinformatics Infrastructure and Magn. Bergvall Foundation.

9 REFERENCES

Allis, D., Jenuwein, T., and Reinberg, D. (2007). Overview and Concepts. In Epigenetics, D. Allis, T. Jenuwein, and D. Reinberg, eds. (New York: Cold Spring Harbor Laboratory Press), pp. 23-61.

Amat, R., Planavila, A., Chen, S.L., Iglesias, R., Giralt, M., and Villarroya, F. (2009). SIRT1 controls the transcription of the peroxisome proliferator-activated receptor-gamma Co-activator-1alpha (PGC-1alpha) gene in skeletal muscle through the PGC-1alpha autoregulatory loop and interaction with MyoD. J Biol Chem *284*, 21872-21880.

Ameln, H., Gustafsson, T., Sundberg, C.J., Okamoto, K., Jansson, E., Poellinger, L., and Makino, Y. (2005). Physiological activation of hypoxia inducible factor-1 in human skeletal muscle. Faseb J *19*, 1009-1011.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics.

Aoi, W., Naito, Y., Takagi, T., Tanimura, Y., Takanami, Y., Kawai, Y., Sakuma, K., Hang, L.P., Mizushima, K., Hirai, Y.*, et al.* (2013). A novel myokine, secreted protein acidic and rich in cysteine (SPARC), suppresses colon tumorigenesis via regular exercise. Gut *62*, 882-889.

Aran, D., Sabato, S., and Hellman, A. (2013). DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. Genome Biol *14*, R21.

Arner, E., Daub, C.O., Vitting-Seerup, K., Andersson, R., Lilje, B., Drablos, F., Lennartsson, A., Ronnerblad, M., Hrydziuszko, O., Vitezic, M.*, et al.* (2015). Gene regulation. Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. Science *347*, 1010-1014.

Ayoubi, T.A., and Van De Ven, W.J. (1996). Regulation of gene expression by alternative promoters. Faseb J *10*, 453-460.

Baar, K. (2014). Nutrition and the adaptation to endurance training. Sports medicine *44 Suppl 1*, S5-12.

Ball, M.P., Li, J.B., Gao, Y., Lee, J.H., LeProust, E.M., Park, I.H., Xie, B., Daley, G.Q., and Church, G.M. (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol *27*, 361-368.

Barres, R., Yan, J., Egan, B., Treebak, J.T., Rasmussen, M., Fritz, T., Caidahl, K., Krook, A., O'Gorman, D.J., and Zierath, J.R. (2012). Acute exercise remodels promoter methylation in human skeletal muscle. Cell metabolism *15*, 405-411.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bennett, D.A., Yu, L., Yang, J., Srivastava, G.P., Aubin, C., and De Jager, P.L. (2015). Epigenomics of Alzheimer's disease. Transl Res *165*, 200-220.

Bergström, J. (1962). Muscle electrolytes in man. Scand J Clin Lab Invest *14*, 1–110.

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L.*, et al.* (2011). High density DNA methylation array with single CpG site resolution. Genomics *98*, 288-295.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev *16*, 6-21.

Bird, A. (2007). Perceptions of epigenetics. Nature *447*, 396-398.

Bishop, D.J., Granata, C., and Eynon, N. (2014). Can we optimise the exercise training prescription to maximise improvements in mitochondria function and content? Biochim Biophys Acta *1840*, 1266-1275.

Blau, H.M., Pavlath, G.K., Hardeman, E.C., Chiu, C.P., Silberstein, L., Webster, S.G., Miller, S.C., and Webster, C. (1985). Plasticity of the differentiated state. Science *230*, 758-766.

Blauuw, B., Shiaffino, S., and Reggiani, C. (2013). Mechanisms modulating skeletal muscle phenotype. Comprehensive Physiology *3*, 1645-1687.

Booth, F.W., Chakravarthy, M.V., and Spangenburg, E.E. (2002). Exercise and gene expression: physiological regulation of the human genome through physical activity. J Physiol *543*, 399-411.

Booth, F.W., Laye, M.J., and Spangenburg, E.E. (2010). Gold standards for scientists who are conducting animal-based exercise studies. J Appl Physiol (1985) *108*, 219-221.

Booth, F.W., Roberts, C.K., and Laye, M.J. (2012). Lack of exercise is a major cause of chronic diseases. Compr Physiol *2*, 1143-1211.

Booth, F.W., and Thomason, D.B. (1991). Molecular and cellular adaptation of muscle in response to exercise: perspectives of various models. Physiol Rev *71*, 541-585.

Branca, R.M., Orre, L.M., Johansson, H.J., Granholm, V., Huss, M., Perez-Bercoff, A., Forshed, J., Kall, L., and Lehtio, J. (2013). HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics. Nature methods.

Bruusgaard, J.C., Johansen, I.B., Egner, I.M., Rana, Z.A., and Gundersen, K. (2010). Myonuclei acquired by overload exercise precede hypertrophy and are not lost on detraining. Proc Natl Acad Sci U S A *107*, 15111-15116.

Catoire, M., Mensink, M., Boekschoten, M.V., Hangelbroek, R., Muller, M., Schrauwen, P., and Kersten, S. (2012). Pronounced effects of acute endurance exercise on gene expression in resting and exercising human skeletal muscle. PLoS One *7*, e51066.

Catrina, S.B., Okamoto, K., Pereira, T., Brismar, K., and Poellinger, L. (2004). Hyperglycemia regulates hypoxia-inducible factor-1alpha protein stability and function. Diabetes *53*, 3226-3232.

Chia, N., Wang, L., Lu, X., Senut, M.C., Brenner, C., and Ruden, D.M. (2011). Hypothesis: environmental regulation of 5-hydroxymethylcytosine by oxidative stress. Epigenetics *6*, 853-856.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem *162*, 156-159.

Clamp, M., Fry, B., Kamal, M., Xie, X., Cuff, J., Lin, M.F., Kellis, M., Lindblad-Toh, K., and Lander, E.S. (2007). Distinguishing protein-coding and noncoding genes in the human genome. Proc Natl Acad Sci U S A *104*, 19428-19433.

Coffey, V.G., and Hawley, J.A. (2007). The molecular bases of training adaptation. Sports medicine *37*, 737-763.

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Consortium, S.M.-I. (2014). A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. Nat Biotechnol *32*, 903-914.

Crane, J.D., MacNeil, L.G., Lally, J.S., Ford, R.J., Bujak, A.L., Brar, I.K., Kemp, B.E., Raha, S., Steinberg, G.R., and Tarnopolsky, M.A. (2015). Exercise-stimulated interleukin-15 is controlled by AMPK and regulates skin metabolism and aging. Aging Cell *14*, 625-634.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A*., et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A *107*, 21931-21936.

Crick, F. (1970). Central dogma of molecular biology. Nature *227*, 561-563.

Dahl, C., Gronbaek, K., and Guldberg, P. (2011). Advances in DNA methylation: 5-hydroxymethylcytosine revisited. Clin Chim Acta *412*, 831-836.

De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.L., and Natoli, G. (2010). A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol *8*, e1000384.

Denham, J., O'Brien, B.J., Harvey, J.T., and Charchar, F.J. (2015). Genome-wide sperm DNA methylation changes after 3 months of exercise training in humans. Epigenomics, 1-15.

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol *4*, P3.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F*., et al.* (2012). Landscape of transcription in human cells. Nature *489*, 101-108.

Duscha, B.D., Annex, B.H., Johnson, J.L., Huffman, K., Houmard, J., and Kraus, W.E. (2012). Exercise dose response in muscle. Int J Sports Med *33*, 218-223.

Egan, B., and Zierath, J.R. (2013). Exercise metabolism and the molecular regulation of skeletal muscle adaptation. Cell Metab *17*, 162-184.

Egelhofer, T.A., Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwierz, P., Alekseyenko, A.A., Cheung, M.S., Day, D.S., Gadel, S., Gorchakov, A.A*., et al.* (2011). An assessment of histone-modification antibody quality. Nat Struct Mol Biol *18*, 91-93.

Ehlert, T., Simon, P., and Moser, D.A. (2013). Epigenetics in sports. Sports medicine *43*, 93-110.

Ehrenborg, E., and Krook, A. (2009). Regulation of skeletal muscle physiology and metabolism by peroxisome proliferator-activated receptor delta. Pharmacol Rev *61*, 373-393.

Ellison, G.M., Waring, C.D., Vicinanza, C., and Torella, D. (2012). Physiological cardiac remodelling in response to endurance exercise training: cellular and molecular mechanisms. Heart *98*, 5-10.

Engstrom, P.G., Steijger, T., Sipos, B., Grant, G.R., Kahles, A., Ratsch, G., Goldman, N., Hubbard, T.J., Harrow, J., Guigo, R*., et al.* (2013). Systematic evaluation of spliced alignment programs for RNA-seq data. Nature methods *10*, 1185-1191.

Epstein, A.C., Gleadle, J.M., McNeill, L.A., Hewitson, K.S., O'Rourke, J., Mole, D.R., Mukherji, M., Metzen, E., Wilson, M.I., Dhanda, A*., et al.* (2001). C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. Cell *107*, 43-54.

Eriksson, L., Byrne, T., Johansson, E., Trygg, J., and Vikström, C. (2013). Multi- and megavariate data analysis - Basic principles and applications, Third edn (Umetrics).

Feil, R., and Fraga, M.F. (2011). Epigenetics and the environment: emerging patterns and implications. Nat Rev Genet *13*, 97-109.

Feinberg, A.P., and Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature *301*, 89-92.

Fischer, H., Gustafsson, T., Sundberg, C.J., Norrbom, J., Ekman, M., Johansson, O., and Jansson, E. (2006). Fatty acid binding protein 4 in human skeletal muscle. Biochem Biophys Res Commun *346*, 125-130.

Foletta, V.C., Brown, E.L., Cho, Y., Snow, R.J., Kralli, A., and Russell, A.P. (2013). Ndrg2 is a PGC-1alpha/ERRalpha target gene that controls protein synthesis and expression of contractile-type genes in C2C12 myotubes. Biochim Biophys Acta *1833*, 3112-3123.

Formenti, F., Constantin-Teodosiu, D., Emmanuel, Y., Cheeseman, J., Dorrington, K.L., Edwards, L.M., Humphreys, S.M., Lappin, T.R., McMullin, M.F., McNamara, C.J.*, et al.* (2010). Regulation of human metabolism by hypoxia-inducible factor. Proc Natl Acad Sci U S A *107*, 12722-12727.

Fox, K.R. (1999). The influence of physical activity on mental well-being. Public Health Nutr *2*, 411-418.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J.*, et al.* (2005). Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A *102*, 10604-10609.

Franke, W.W., and Schinko, W. (1969). Nuclear shape in muscle cells. J Cell Biol *42*, 326-331.

Frazee, A.C., Pertea, G., Jaffe, A.E., Langmead, B., Salzberg, S.L., and Leek, J.T. (2015). Ballgown bridges the gap between transcriptome assembly and expression analysis. Nat Biotechnol *33*, 243-246.

Fredriksson, K., Tjader, I., Keller, P., Petrovic, N., Ahlman, B., Scheele, C., Wernerman, J., Timmons, J.A., and Rooyackers, O. (2008). Dysregulation of mitochondrial dynamics and the muscle transcriptome in ICU patients suffering from sepsis induced multiple organ failure. PLoS One *3*, e3686.

Gidlund, E.K., Ydfors, M., Appel, S., Rundqvist, H., Sundberg, C.J., and Norrbom, J.M. (2015). Rapidly elevated levels of PGC-1alpha-b protein in human skeletal muscle after exercise: exploring regulatory factors in a randomized controlled trial. J Appl Physiol (1985), jap 01000 02014.

Green, H.J. (2000). Adaptations in the muscle cell to training: role of the Na+-K+-Atpase. Can J Appl Physiol *25*, 204-216.

Grossniklaus, U., and Paro, R. (2014). Transcriptional silencing by polycomb-group proteins. Cold Spring Harb Perspect Biol *6*, a019331.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77-88.

Gundersen, K. (2011). Excitation-transcription coupling in skeletal muscle: the molecular pathways of exercise. Biol Rev Camb Philos Soc *86*, 564-600.

Gustafsson, T. (2011). Vascular remodelling in human skeletal muscle. Biochem Soc Trans *39*, 1628-1632.

Gustafsson, T., Puntschart, A., Kaijser, L., Jansson, E., and Sundberg, C.J. (1999). Exercise-induced expression of angiogenesis-related transcription and growth factors in human skeletal muscle. Am J Physiol *276*, H679-685.

Hangauer, M.J., Vaughn, I.W., and McManus, M.T. (2013). Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. PLoS Genet *9*, e1003569.

Haskell, W.L., Lee, I.M., Pate, R.R., Powell, K.E., Blair, S.N., Franklin, B.A., Macera, C.A., Heath, G.W., Thompson, P.D., and Bauman, A. (2007). Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc *39*, 1423-1434.

Hawley, J.A. (2002). Adaptations of skeletal muscle to prolonged, intense endurance training. Clin Exp Pharmacol Physiol *29*, 218-222.

Henriksson, J. (1995). Influence of exercise on insulin sensitivity. J Cardiovasc Risk *2*, 303-309.

Hirst, M., and Marra, M.A. (2009). Epigenetics and human disease. Int J Biochem Cell Biol *41*, 136-146.

Holloszy, J.O., and Booth, F.W. (1976). Biochemical adaptations to endurance exercise in muscle. Annual review of physiology *38*, 273-291.

Holloszy, J.O., Oscai, L.B., Don, I.J., and Mole, P.A. (1970). Mitochondrial citric acid cycle and related enzymes: adaptive response to exercise. Biochem Biophys Res Commun *40*, 1368-1373.

Hoppeler, H., and Vogt, M. (2001). Muscle tissue adaptations to hypoxia. J Exp Biol *204*, 3133-3139.

Huber, M.D., and Gerace, L. (2007). The size-wise nucleus: nuclear volume control in eukaryotes. J Cell Biol *179*, 583-584.

International Human Genome Sequencing, C. (2004). Finishing the euchromatic sequence of the human genome. Nature *431*, 931-945.

Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M.*, et al.* (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet *41*, 178-186.

Issa, J.P. (2014). Aging and epigenetic drift: a vicious cycle. J Clin Invest *124*, 24-29.

Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J.M., Lane, W.S., and Kaelin, W.G., Jr. (2001). HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. Science *292*, 464-468.

Jaakkola, P., Mole, D.R., Tian, Y.M., Wilson, M.I., Gielbert, J., Gaskell, S.J., von Kriegsheim, A., Hebestreit, H.F., Mukherji, M., Schofield, C.J.*, et al.* (2001). Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. Science *292*, 468-472.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet *33 Suppl*, 245-254.

Jeong, J.W., Bae, M.K., Ahn, M.Y., Kim, S.H., Sohn, T.K., Bae, M.H., Yoo, M.A., Song, E.J., Lee, K.J., and Kim, K.W. (2002). Regulation and destabilization of HIF-1alpha by ARD1-mediated acetylation. Cell *111*, 709-720.

Jjingo, D., Conley, A.B., Yi, S.V., Lunyak, V.V., and Jordan, I.K. (2012). On the presence and role of human gene-body DNA methylation. Oncotarget *3*, 462-474.

Kaijser, L., Sundberg, C.J., Eiken, O., Nygren, A., Esbjornsson, M., Sylven, C., and Jansson, E. (1990). Muscle oxidative capacity and work performance after training under local leg ischemia. J Appl Physiol (1985) *69*, 785-787.

Kanitz, A., Gypas, F., Gruber, A.J., Gruber, A.R., Martin, G., and Zavolan, M. (2015). Comparative assessment of methods for the computational inference of transcript isoform abundance from RNA-seq data. Genome Biol *16*, 150.

Kanzleiter, T., Jahnert, M., Schulze, G., Selbig, J., Hallahan, N., Schwenk, R.W., and Schurmann, A. (2015). Exercise training alters DNA methylation patterns in genes related to muscle growth and differentiation in mice. Am J Physiol Endocrinol Metab *308*, E912-920.

Karimi, M., Johansson, S., Stach, D., Corcoran, M., Grander, D., Schalling, M., Bakalkin, G., Lyko, F., Larsson, C., and Ekstrom, T.J. (2006). LUMA (LUminometric Methylation Assay)--a high throughput method to the analysis of genomic DNA methylation. Exp Cell Res *312*, 1989-1995.

Kato, H., Tamamizu-Kato, S., and Shibasaki, F. (2004). Histone deacetylase 7 associates with hypoxia-inducible factor 1alpha and increases transcriptional activity. J Biol Chem *279*, 41966-41974.

Keller, P., Vollaard, N.B., Gustafsson, T., Gallagher, I.J., Sundberg, C.J., Rankinen, T., Britton, S.L., Bouchard, C., Koch, L.G., and Timmons, J.A. (2011). A transcriptional map of the impact of endurance exercise training on skeletal muscle phenotype. J Appl Physiol (1985) *110*, 46-59.

Khoury, G., and Gruss, P. (1983). Enhancer elements. Cell *33*, 313-314.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013a). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Kim, G.H., Ryan, J.J., and Archer, S.L. (2013b). The role of redox signaling in epigenetics and cardiovascular disease. Antioxid Redox Signal *18*, 1920-1936.

Kim, J.W., Tchernyshyov, I., Semenza, G.L., and Dang, C.V. (2006). HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia. Cell Metab *3*, 177-185.

Kim, M.S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R., Madugundu, A.K., Kelkar, D.S., Isserlin, R., Jain, S.*, et al.* (2014). A draft map of the human proteome. Nature *509*, 575-581.

Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S.*, et al.* (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182-187.

Kimura, H. (2013). Histone modifications for human epigenome analysis. J Hum Genet *58*, 439-445.

Kirchner, H., Osler, M.E., Krook, A., and Zierath, J.R. (2013). Epigenetic flexibility in metabolic regulation: disease cause and prevention? Trends Cell Biol *23*, 203-209.

Kjellqvist, S., Maleki, S., Olsson, T., Chwastyniak, M., Branca, R.M., Lehtio, J., Pinet, F., Franco-Cereceda, A., and Eriksson, P. (2013). A combined proteomic and transcriptomic approach shows diverging molecular mechanisms in thoracic aortic aneurysm development in patients with tricuspid- and bicuspid aortic valve. Mol Cell Proteomics *12*, 407-425.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kraniou, Y., Cameron-Smith, D., Misso, M., Collier, G., and Hargreaves, M. (2000). Effects of exercise on GLUT-4 and glycogenin gene expression in human skeletal muscle. J Appl Physiol (1985) *88*, 794-796.

Krieg, A.J., Rankin, E.B., Chan, D., Razorenova, O., Fernandez, S., and Giaccia, A.J. (2010). Regulation of the histone demethylase JMJD1A by hypoxia-inducible factor 1 alpha enhances hypoxic gene expression and tumor growth. Mol Cell Biol *30*, 344-353.

Kulis, M., Heath, S., Bibikova, M., Queiros, A.C., Navarro, A., Clot, G., Martinez-Trillos, A., Castellano, G., Brun-Heath, I., Pinyol, M.*, et al.* (2012). Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat Genet *44*, 1236-1242.

Kurtovic, S., Paloschi, V., Folkersen, L., Gottfries, J., Franco-Cereceda, A., and Eriksson, P. (2011). Diverging alternative splicing fingerprints in the transforming growth factor-beta signaling pathway identified in thoracic aortic aneurysms. Mol Med *17*, 665-675.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lando, D., Peet, D.J., Gorman, J.J., Whelan, D.A., Whitelaw, M.L., and Bruick, R.K. (2002). FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor. Genes Dev *16*, 1466-1471.

Lennartsson, A., and Ekwall, K. (2009). Histone modification patterns and epigenetic codes. Biochim Biophys Acta.

Lev Maor, G., Yearim, A., and Ast, G. (2015). The alternative role of DNA methylation in splicing regulation. Trends Genet *31*, 274-280.

Lindskog, C., Linne, J., Fagerberg, L., Hallstrom, B.M., Sundberg, C.J., Lindholm, M., Huss, M., Kampf, C., Choi, H., Liem, D.A.*, et al.* (2015). The human cardiac and skeletal muscle proteomes defined by transcriptomics and antibody-based profiling. BMC Genomics *16*, 475.

Little, J.P., Safdar, A., Cermak, N., Tarnopolsky, M.A., and Gibala, M.J. (2010). Acute endurance exercise increases the nuclear abundance of PGC-1alpha in trained human skeletal muscle. Am J Physiol Regul Integr Comp Physiol *298*, R912-917.

Liu, Y., Aryee, M.J., Padyukov, L., Fallin, M.D., Hesselberg, E., Runarsson, A., Reinius, L., Acevedo, N., Taub, M., Ronninger, M.*, et al.* (2013). Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nat Biotechnol *31*, 142-147.

Loos, R.J., Hagberg, J.M., Perusse, L., Roth, S.M., Sarzynski, M.A., Wolfarth, B., Rankinen, T., and Bouchard, C. (2015). Advances in exercise, fitness, and performance genomics in 2014. Med Sci Sports Exerc *47*, 1105-1112.

Lorenzo, F.R., Huff, C., Myllymaki, M., Olenchock, B., Swierczek, S., Tashi, T., Gordeuk, V., Wuren, T., Ri-Li, G., McClain, D.A.*, et al.* (2014). A genetic mechanism for Tibetan high-altitude adaptation. Nat Genet *46*, 951-956.

Lowry, O.H., and Passonneau, J.V. (1972). A flexible system of enzymatic analysis (New York: Academic Press).

Lu, C., and Thompson, C.B. (2012). Metabolic regulation of epigenetics. Cell Metab *16*, 9-17.

Lundby, C., Gassmann, M., and Pilegaard, H. (2006). Regular endurance training reduces the exercise induced HIF-1alpha and HIF-2alpha mRNA expression in human skeletal muscle in normoxic conditions. Eur J Appl Physiol *96*, 363-369.

Luo, W., and Brouwer, C. (2013). Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics *29*, 1830-1831.

Luttropp, K., Nordfors, L., Ekstrom, T.J., and Lind, L. (2013). Physical activity is associated with decreased global DNA methylation in Swedish older individuals. Scand J Clin Lab Invest *73*, 184-185.

Maher, A.C., Fu, M.H., Isfort, R.J., Varbanov, A.R., Qu, X.A., and Tarnopolsky, M.A. (2009). Sex differences in global mRNA content of human skeletal muscle. PLoS One *4*, e6335.

Mahon, P.C., Hirota, K., and Semenza, G.L. (2001). FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity. Genes Dev *15*, 2675-2686.

Mahoney, D.J., Parise, G., Melov, S., Safdar, A., and Tarnopolsky, M.A. (2005). Analysis of global mRNA expression in human skeletal muscle during recovery from endurance exercise. Faseb J *19*, 1498-1500.

Marabita, F., Almgren, M., Lindholm, M.E., Ruhrmann, S., Fagerstrom-Billai, F., Jagodic, M., Sundberg, C.J., Ekstrom, T.J., Teschendorff, A.E., Tegner, J.*, et al.* (2013). An evaluation of analysis pipelines for DNA methylation profiling using the Illumina HumanMethylation450 BeadChip platform. Epigenetics *8*, 333-346.

Margolin, A.A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Dalla Favera, R., and Califano, A. (2006). ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics *7 Suppl 1*, S7.

Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M., and Gilad, Y. (2008). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res *18*, 1509-1517.

Martens, H., Hoy, M., Westad, F., Folkenberg, B., and Martens, M. (2001). Analysis of designed experiments by stabilised PLS Regression and jack-knifing. . Chemometrics and Intelligent Laboratory Systems, 151-170.

Mason, S.D., Howlett, R.A., Kim, M.J., Olfert, I.M., Hogan, M.C., McNulty, W., Hickey, R.P., Wagner, P.D., Kahn, C.R., Giordano, F.J.*, et al.* (2004). Loss of skeletal muscle HIF-1alpha results in altered exercise endurance. PLoS Biol *2*, e288.

Mason, S.D., Rundqvist, H., Papandreou, I., Duh, R., McNulty, W.J., Howlett, R.A., Olfert, I.M., Sundberg, C.J., Denko, N.C., Poellinger, L.*, et al.* (2007a). HIF-1{alpha} in endurance training: suppression of oxidative metabolism

10.1152/ajpregu.00335.2007. Am J Physiol Regul Integr Comp Physiol *293*, R2059-2069.

Mason, S.D., Rundqvist, H., Papandreou, I., Duh, R., McNulty, W.J., Howlett, R.A., Olfert, I.M., Sundberg, C.J., Denko, N.C., Poellinger, L.*, et al.* (2007b). HIF-1alpha in endurance training: suppression of oxidative metabolism. Am J Physiol Regul Integr Comp Physiol *293*, R2059-2069.

McGee, S.L., Fairlie, E., Garnham, A.P., and Hargreaves, M. (2009). Exercise-induced histone modifications in human skeletal muscle. J Physiol *587*, 5951-5958.

McGee, S.L., and Hargreaves, M. (2011). Histone modifications and exercise adaptations. J Appl Physiol (1985) *110*, 258-263.

McPhee, J.S., Perez-Schindler, J., Degens, H., Tomlinson, D., Hennis, P., Baar, K., and Williams, A.G. (2011). HIF1A P582S gene association with endurance training responses in young women. Eur J Appl Physiol *111*, 2339-2347.

Michishita, E., McCord, R.A., Berber, E., Kioi, M., Padilla-Nash, H., Damian, M., Cheung, P., Kusumoto, R., Kawahara, T.L., Barrett, J.C.*, et al.* (2008). SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. Nature *452*, 492-496.

Munzel, M., Globisch, D., and Carell, T. (2011). 5-Hydroxymethylcytosine, the sixth base of the genome. Angew Chem Int Ed Engl *50*, 6460-6468.

Nagalakshmi, U., Waern, K., and Snyder, M. (2010). RNA-Seq: a method for comprehensive transcriptome analysis. Curr Protoc Mol Biol *Chapter 4*, Unit 4 11 11-13.

Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E.D., Tanay, A., and Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature *502*, 59-64.

Neufer, P.D., Bamman, M.M., Muoio, D.M., Bouchard, C., Cooper, D.M., Goodpaster, B.H., Booth, F.W., Kohrt, W.M., Gerszten, R.E., Mattson, M.P.*, et al.* (2015). Understanding the Cellular and Molecular Mechanisms of Physical Activity-Induced Health Benefits. Cell Metab.

Ng, R.K., and Gurdon, J.B. (2008). Epigenetic inheritance of cell differentiation status. Cell Cycle *7*, 1173-1177.

Ngo, K.T., Denis, C., Saafi, M.A., Feasson, L., and Verney, J. (2012). Endurance but not resistance training increases intra-myocellular lipid content and beta-hydroxyacyl coenzyme A dehydrogenase activity in active elderly men. Acta Physiol (Oxf) *205*, 133-144.

Nitert, M.D., Dayeh, T., Volkov, P., Elgzyri, T., Hall, E., Nilsson, E., Yang, B.T., Lang, S., Parikh, H., Wessman, Y.*, et al.* (2012). Impact of an exercise intervention on DNA methylation in skeletal muscle from first-degree relatives of patients with type 2 diabetes. Diabetes *61*, 3322-3332.

Norman, B., Esbjornsson, M., Rundqvist, H., Osterlund, T., Glenmark, B., and Jansson, E. (2014). ACTN3 genotype and modulation of skeletal muscle response to exercise in human subjects. J Appl Physiol (1985) *116*, 1197-1203.

Norrbom, J., Wallman, S.E., Gustafsson, T., Rundqvist, H., Jansson, E., and Sundberg, C.J. (2010). Training response of mitochondrial transcription factors in human skeletal muscle. Acta physiologica *198*, 71-79.

Olesen, J., Kiilerich, K., and Pilegaard, H. (2010). PGC-1alpha-mediated adaptations in skeletal muscle. Pflugers Arch *460*, 153-162.

Ookawara, T., Suzuk, K., Haga, S., Ha, S., Chung, K.S., Toshinai, K., Hamaoka, T., Katsumura, T., Takemasa, T., Mizuno, M., *et al.* (2002). Transcription regulation of gene expression in human skeletal muscle in response to endurance training. Res Commun Mol Pathol Pharmacol *111*, 41-54.

Ozsolak, F., and Milos, P.M. (2011). RNA sequencing: advances, challenges and opportunities. Nat Rev Genet *12*, 87-98.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet *40*, 1413-1415.

Papandreou, I., Cairns, R.A., Fontana, L., Lim, A.L., and Denko, N.C. (2006). HIF-1 mediates adaptation to hypoxia by actively downregulating mitochondrial oxygen consumption. Cell Metab *3*, 187-197.

Pedersen, B.K., and Febbraio, M.A. (2012). Muscles, exercise and obesity: skeletal muscle as a secretory organ. Nat Rev Endocrinol *8*, 457-465.

Pembrey, M., Saffery, R., Bygren, L.O., Network in Epigenetic, E., and Network in Epigenetic, E. (2014). Human transgenerational responses to early-life experience: potential impact on development, health and biomedical research. J Med Genet *51*, 563-572.

Pennisi, E. (2012). Genomics. ENCODE project writes eulogy for junk DNA. Science *337*, 1159, 1161.

Perry, C.G., Talanian, J.L., Heigenhauser, G.J., and Spriet, L.L. (2007). The effects of training in hyperoxia vs. normoxia on skeletal muscle enzyme activities and exercise performance. J Appl Physiol (1985) *102*, 1022-1027.

Pilegaard, H., Saltin, B., and Neufer, P.D. (2003). Exercise induces transient transcriptional activation of the PGC-1alpha gene in human skeletal muscle. J Physiol *546*, 851-858.

Pollard, P.J., Loenarz, C., Mole, D.R., McDonough, M.A., Gleadle, J.M., Schofield, C.J., and Ratcliffe, P.J. (2008). Regulation of Jumonji-domain-containing histone demethylases by hypoxia-inducible factor (HIF)-1alpha. Biochem J *416*, 387-394.

Prior, S.J., Hagberg, J.M., Phares, D.A., Brown, M.D., Fairfull, L., Ferrell, R.E., and Roth, S.M. (2003). Sequence variation in hypoxia-inducible factor 1alpha (HIF1A): association with maximal oxygen consumption. Physiol Genomics *15*, 20-26.

Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating mRNA processing with transcription. Cell *108*, 501-512.

Raj, A., and van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. Cell *135*, 216-226.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc Natl Acad Sci U S A *97*, 5237-5242.

Ramskold, D., Wang, E.T., Burge, C.B., and Sandberg, R. (2009). An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. PLoS Comput Biol *5*, e1000598.

Rapaport, F., Khanin, R., Liang, Y., Pirun, M., Krek, A., Zumbo, P., Mason, C.E., Socci, N.D., and Betel, D. (2013). Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Biol *14*, R95.

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. Nature *447*, 425-432.

Robertson, K.D. (2005). DNA methylation and human disease. Nat Rev Genet *6*, 597-610.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol *11*, R25.

Rodriguez-Miguelez, P., Lima-Cabello, E., Martinez-Florez, S., Almar, M., Cuevas, M.J., and Gonzalez-Gallego, J. (2015). Hypoxia-inducible factor-1 modulates the expression of vascular endothelial growth factor and endothelial nitric oxide synthase induced by eccentric exercise. Journal of applied physiology *118*, 1075-1083.

Rönn, T., Volkov, P., Davegardh, C., Dayeh, T., Hall, E., Olsson, A.H., Nilsson, E., Tornberg, A., Dekker Nitert, M., Eriksson, K.F*., et al.* (2013). A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue. PLoS Genet *9*, e1003572.

Roudier, E., Aiken, J., Slopack, D., Gouzi, F., Mercier, J., Haas, T.L., Gustafsson, T., Hayot, M., and Birot, O. (2013). Novel perspective: exercise training stimulus triggers the expression of the oncoprotein human double minute-2 in human skeletal muscle. Physiological reports *1*, e00028.

Rullman, E., Rundqvist, H., Wagsater, D., Fischer, H., Eriksson, P., Sundberg, C.J., Jansson, E., and Gustafsson, T. (2007). A single bout of exercise activates matrix metalloproteinase in human skeletal muscle. J Appl Physiol *102*, 2346-2351.

Russell, A.P., Feilchenfeldt, J., Schreiber, S., Praz, M., Crettenand, A., Gobelet, C., Meier, C.A., Bell, D.R., Kralli, A., Giacobino, J.P*., et al.* (2003). Endurance training in humans leads to fiber type-specific increases in levels of peroxisome proliferator-activated receptor-gamma coactivator-1 and peroxisome proliferator-activated receptor-alpha in skeletal muscle. Diabetes *52*, 2874-2881.

Sakagami, H., Makino, Y., Mizumoto, K., Isoe, T., Takeda, Y., Watanabe, J., Fujita, Y., Takiyama, Y., Abiko, A., and Haneda, M. (2014). Loss of HIF-1alpha impairs GLUT4 translocation and glucose uptake by the skeletal muscle cells. American journal of physiology Endocrinology and metabolism *306*, E1065-1076.

Saltin, B., and Gollnick, P.D. (2011). Skeletal Muscle Adaptability: Significance for Metabolism and Performance, First published in print 1983 edn (Comprehensive Physiology).

Sandoval, J., Heyn, H., Moran, S., Serra-Musach, J., Pujana, M.A., Bibikova, M., and Esteller, M. (2011). Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics *6*, 692-702.

Schmittgen, T.D., Zakrajsek, B.A., Mills, A.G., Gorn, V., Singer, M.J., and Reed, M.W. (2000). Quantitative reverse transcription-polymerase chain reaction to study mRNA decay: comparison of endpoint and real-time methods. Anal Biochem *285*, 194-204.

Schwer, B., and Verdin, E. (2008). Conserved metabolic regulatory functions of sirtuins. Cell Metab *7*, 104-112.

Semenza, G.L. (1999). Regulation of mammalian O2 homeostasis by hypoxia-inducible factor 1. Annu Rev Cell Dev Biol *15*, 551-578.

Seo, H.W., Kim, E.J., Na, H., and Lee, M.O. (2009). Transcriptional activation of hypoxia-inducible factor-1alpha by HDAC4 and HDAC5 involves differential recruitment of p300 and FIH-1. FEBS Lett *583*, 55-60.

Short, K.R., Vittone, J.L., Bigelow, M.L., Proctor, D.N., Coenen-Schimke, J.M., Rys, P., and Nair, K.S. (2005). Changes in myosin heavy chain mRNA and protein expression in human skeletal muscle with age and endurance exercise training. J Appl Physiol (1985) *99*, 95-102.

Smyth, G.K. (2005). Limma: linear models for microarray data. In: Bioinformatics and Computational Biology Solutions using R and Bioconductor. R Gentleman, V Carey, S Dudoit, R Irizarry, W Huber (eds) *Springer, New York*, 397-420.

Soneson, C., and Delorenzi, M. (2013). A comparison of methods for differential expression analysis of RNA-seq data. BMC Bioinformatics *14*, 91.

Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Scholer, A., van Nimwegen, E., Wirbelauer, C., Oakeley, E.J., Gaidatzis, D.*, et al.* (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature *480*, 490-495.

Staron, R.S., Hagerman, F.C., Hikida, R.S., Murray, T.F., Hostler, D.P., Crill, M.T., Ragg, K.E., and Toma, K. (2000). Fiber type composition of the vastus lateralis muscle of young men and women. J Histochem Cytochem *48*, 623-629.

Stroka, D.M., Burkhardt, T., Desbaillets, I., Wenger, R.H., Neil, D.A., Bauer, C., Gassmann, M., and Candinas, D. (2001). HIF-1 is expressed in normoxic tissue and displays an organ-specific regulation under systemic hypoxia. Faseb J *15*, 2445-2453.

Sundberg, C.J., Eiken, O., Nygren, A., and Kaijser, L. (1993). Effects of ischaemic training on local aerobic muscle performance in man. Acta Physiol Scand *148*, 13-19.

Swigut, T., and Wysocka, J. (2007). H3K27 demethylases, at long last. Cell *131*, 29-32.

Tanimoto, K., Yoshiga, K., Eguchi, H., Kaneyasu, M., Ukon, K., Kumazaki, T., Oue, N., Yasui, W., Imai, K., Nakachi, K.*, et al.* (2003). Hypoxia-inducible factor-1alpha polymorphisms associated with enhanced transactivation capacity, implying clinical significance. Carcinogenesis *24*, 1779-1783.

Thomas, P.D., Kejariwal, A., Campbell, M.J., Mi, H., Diemer, K., Guo, N., Ladunga, I., Ulitsky-Lazareva, B., Muruganujan, A., Rabkin, S.*, et al.* (2003). PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. Nucleic Acids Res *31*, 334-341.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B.*, et al.* (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75-82.

Timmons, J.A., Knudsen, S., Rankinen, T., Koch, L.G., Sarzynski, M.A., Jensen, T., Keller, P., Scheele, C., Vollaard, N.B., Nielsen, S.*, et al.* (2010). Using molecular classification to predict gains in maximal aerobic capacity following endurance exercise training in humans. J Appl Physiol.

Timmons, J.A., Larsson, O., Jansson, E., Fischer, H., Gustafsson, T., Greenhaff, P.L., Ridden, J., Rachman, J., Peyrard-Janvid, M., Wahlestedt, C.*, et al.* (2005). Human muscle gene expression responses to endurance training provide a novel perspective on Duchenne muscular dystrophy. Faseb J *19*, 750-760.

Tost, J., and Gut, I.G. (2007). DNA methylation analysis by pyrosequencing. Nat Protoc *2*, 2265-2275.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

van Praag, H. (2009). Exercise and the brain: something to chew on. Trends Neurosci *32*, 283-290.

Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P.*, et al.* (2009). Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science *323*, 1208-1211.

Vassilopoulos, D., Lumb, E.M., and Emery, A.E. (1976). Muscle nuclear size in neuromuscular disease. J Neurol Neurosurg Psychiatry *39*, 159-162.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A.*, et al.* (2001). The sequence of the human genome. Science *291*, 1304-1351.

Vogel, C., and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat Rev Genet *13*, 227-232.

Voisin, S., Eynon, N., Yan, X., and Bishop, D.J. (2015). Exercise training and DNA methylation in humans. Acta Physiol (Oxf) *213*, 39-59.

Vollaard, N.B., Constantin-Teodosiu, D., Fredriksson, K., Rooyackers, O., Jansson, E., Greenhaff, P.L., Timmons, J.A., and Sundberg, C.J. (2009). Systematic analysis of adaptations in aerobic capacity and submaximal energy metabolism provides a unique insight into determinants of human aerobic performance. J Appl Physiol *106*, 1479-1486.

Waddington, C. (1942). The Epigenotype. Endeavour, 18-20.

Wadley, G.D., Tunstall, R.J., Sanigorski, A., Collier, G.R., Hargreaves, M., and Cameron-Smith, D. (2001). Differential effects of exercise on insulin-signaling gene expression in human skeletal muscle. J Appl Physiol (1985) *90*, 436-440.

Wallace, M.A., Hock, M.B., Hazen, B.C., Kralli, A., Snow, R.J., and Russell, A.P. (2011). Striated muscle activator of Rho signalling (STARS) is a PGC-1alpha/oestrogen-related receptor-alpha target gene and is upregulated in human skeletal muscle after endurance exercise. J Physiol *589*, 2027-2039.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470-476.

Wang, G.L., and Semenza, G.L. (1993). General involvement of hypoxia-inducible factor 1 in transcriptional response to hypoxia. Proc Natl Acad Sci U S A *90*, 4304-4308.

Welle, S., Tawil, R., and Thornton, C.A. (2008). Sex-related differences in gene expression in human skeletal muscle. PLoS One *3*, e1385.

White, A.J., Sandler, D.P., Bolick, S.C., Xu, Z., Taylor, J.A., and DeRoo, L.A. (2013). Recreational and household physical activity at different time points and DNA global methylation. European journal of cancer *49*, 2199-2206.

White, A.T., and Schenk, S. (2012). NAD(+)/NADH and skeletal muscle mitochondrial adaptations to exercise. Am J Physiol Endocrinol Metab *303*, E308-321.

Whyte, J.J., and Laughlin, M.H. (2010). The effects of acute and chronic exercise on the vasculature. Acta Physiol (Oxf) *199*, 441-450.

Widegren, U., Wretman, C., Lionikas, A., Hedin, G., and Henriksson, J. (2000). Influence of exercise intensity on ERK/MAP kinase signalling in human skeletal muscle. Pflugers Arch *441*, 317-322.

Wiik, A., Gustafsson, T., Esbjornsson, M., Johansson, O., Ekman, M., Sundberg, C.J., and Jansson, E. (2005). Expression of oestrogen receptor alpha and beta is higher in skeletal muscle of highly endurance-trained than of moderately active men. Acta Physiol Scand *184*, 105-112.

Windhorst, U., and Mommaerts, W.F.H.M. (1996). Physiology of Skeletal Muscle in: *Comprehensive Human Physiology - from cellular mechanisms to integration*, Vol 1 (Heidelberg: Springer).

Xia, X., Lemieux, M.E., Li, W., Carroll, J.S., Brown, M., Liu, X.S., and Kung, A.L. (2009). Integrative analysis of HIF binding and transactivation reveals its role in maintaining histone methylation homeostasis. Proc Natl Acad Sci U S A *106*, 4260-4265.

Xiao, M., Yang, H., Xu, W., Ma, S., Lin, H., Zhu, H., Liu, L., Liu, Y., Yang, C., Xu, Y.*, et al.* (2012). Inhibition of alpha-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. Genes Dev *26*, 1326-1338.

Xie, M., Hong, C., Zhang, B., Lowdon, R.F., Xing, X., Li, D., Zhou, X., Lee, H.J., Maire, C.L., Ligon, K.L.*, et al.* (2013). DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. Nat Genet *45*, 836-841.

Yaffe, D., and Saxel, O. (1977). Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. Nature *270*, 725-727.

Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol *11*, R14.

Yu, J.G., Bonnerud, P., Eriksson, A., Stal, P.S., Tegner, Y., and Malm, C. (2014). Effects of long term supplementation of anabolic androgen steroids on human skeletal muscle. PLoS One *9*, e105330.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhong, L., D'Urso, A., Toiber, D., Sebastian, C., Henry, R.E., Vadysirisack, D.D., Guimaraes, A., Marinelli, B., Wikstrom, J.D., Nir, T.*, et al.* (2010). The histone deacetylase Sirt6 regulates glucose homeostasis via Hif1alpha. Cell *140*, 280-293.

Ziller, M.J., Gu, H., Muller, F., Donaghey, J., Tsai, L.T., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E.*, et al.* (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature *500*, 477-481.