# Proteomic profiling of skeletal muscle mitochondrial adaptations to exercise training through meta-analysis methods

Mr. Siddharth Pruthi

Master of Research

Submitted in fulfilment of the requirements of the degree of Master of Research

Victoria University, Australia

Institute for Health & Sport (iHeS)

July 2023

**Student Declaration**

"I, Siddharth Pruthi, declare that the Master of Research thesis entitled "Proteomic profiling of skeletal muscle mitochondrial adaptations to exercise training through meta-analysis methods" is no more than 50,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work".

"I have conducted my research in alignment with the Australian Code for the Responsible Conduct of Research and Victoria University's Higher Degree by Research Policy and Procedures.

Signature

Date:

17/07/2023

**Acknowledgements**:

This research has been close to two years of lots of days (mostly nights) of fighting with *R*, reading, writing, and re-writing (and scouring documentation to find the one parameter that makes the plot look juust right…). All that being said, and as tough as I find finally letting go of my companion of the last two years, I am quite proud of being able to finish what I started.

I would first like to thank my supervisors Prof. David Bishop and Dr. Nikiesha Caruana for their immense support and patience throughout. Their enthusiasm and passion for this project helped keep me motivated over the course of this long journey. I have been truly fortunate to have them as my mentors.

I would also like to thank Dr. Nicolas Pillon who helped me get familiar with meta-analysis techniques and pipelines and was very supportive and generous with his time. As this project has been built on secondary data, I am grateful to all who timely responded to our requests for their data. This (literally) would not have been possible without them.

Finally, I would like to thank my family. My mom and dad have always been smilingly and unconditionally supportive of all that I have ever done, and I owe them everything. I would also like to thank my girlfriend who has always been a calming and loving presence as I went through all the trials, tribulations, and anxieties that accompany writing a thesis.

While I would have loved to study and learn with the entire Bishop lab group on-campus in sunny Melbourne, being able to stay academically engaged and pursue my interest from home, staying with my family, allowed for invaluable experiences and memories. Had I travelled away as I was planning to when this opportunity came along, we would have never adopted the three cutest little (now big) kittens who are the centre of our universes. And I would not have had a chance to

say goodbye to my grandfather and perform his funeral rites. I am grateful for all that this opportunity offered me in my personal and academic life and I hope I have done it justice. As far as Melbourne goes, maybe it still is in my future!

## Table of Contents

## Figure Legends

## Abstract

Many studies have analysed the skeletal muscle mitochondrial proteome in relation to exercise training. These studies have identified adaptations within metabolic pathways and other pathways governing mitochondria form and function. With the adoption of high-throughput proteomics within the field, the depth of mitochondrial proteome coverage has increased many-fold within the last decade. Information on protein abundance changes following exercise training now exists for hundreds of mitochondrial proteins across multiple studies. Despite this greater data availability, no research has yet sought to compare, contrast, and aggregate the findings from these studies. Therefore, this research aims at a comprehensive analysis of adaptations of the skeletal muscle mitochondria proteome to exercise training using all available high-throughput proteomics data.

Eight independent datasets were shortlisted to be included in the meta-analysis study. For each dataset, raw protein intensities were extracted for each sample from the respective *proteinGroups* file. Uniform filtration, imputation, and normalisation, if required, were applied to raw protein intensities and differential expression analysis was performed using a linear-modelling approach in *DEqMS*, a package built on top of limma and which accounts for peptide-identification information when adjusting the protein-wise variance estimator. A Random-Effects meta-analysis was performed on these aggregated differential expression results using the package *metafor*. Independently of these analyses, the effects of mitochondria-normalisation (*in silico* scaling of raw protein intensities to account for overall change in mitochondria protein abundance) on differential expression results was also explored.

A total of 778 proteins were quantified across the eight datasets and the meta-analysis revealed 200 ($K^1 \geq 5$) differentially expressed proteins at $FDR < 0.05$. The capacity of the increased power of the meta-analysis design was also highlighted as approximately 20% of the significant proteins at the meta-analysis level were not significant in any individual dataset. The findings of the meta-analysis revealed an enrichment for terms related to the Electron Transport Chain, multiple metabolic pathways, and mitochondria biogenesis. While not demonstrating significant enrichment, multiple central proteins within pathways of ROS detoxification (*SOD2*, *GSTK1, PRDX3*), protein import and assembly (*TIMM44, HSPA9, GRPEL1*), and small molecule transport (*SLC25A3*, *MTCH2*) were identified as significantly upregulated. Exploratory analysis of heterogeneity scores ($I^2$ %) of meta-analysis results suggest stronger effects of study variables (exercise-training protocol) on adaptations to oxidative metabolism proteins compared to the rest. Analysis of mitochondria-normalised data revealed different enrichment profiles depending upon exercise exposure of the associated study, suggesting the potential of this technique to help investigate the time-course of mitochondria proteome adaptations to exercise training independent of overall change in mitochondria content.

This research aims to be the basis of a resource for the wider scientific community that can be used to generate novel hypotheses, select targets for further experimental validation, and investigate exercise-training induced adaptations at the level of the mitochondria protein network.

[1] Refers to the number of independent studies in which a particular protein was quantified.

# 1 CHAPTER ONE

## Review of Literature

## Exercise Training and the Skeletal Muscle Mitochondria Proteome: Existing evidence and future directions.

## 1.1 Skeletal Muscle and Mitochondria

Mitochondria are commonly described as the energy-producing "powerhouses" of the cell but play fundamental roles in many other biological functions and cellular signalling pathways. One of their most critical roles is as the hub of cellular metabolism – producing ATP through the process of oxidative phosphorylation [7, 8]. However, independent of energy provision, mitochondria have also been shown to play a role in cell cycle progression, the apoptosis cascade, and as a site for protein complex assembly [9].

Altered mitochondrial characteristics have been associated with a host of metabolic conditions, such as insulin resistance, the onset of obesity, cachexia, and cancer [10, 11]. These characteristics include changes in mitochondrial volume/content, morphology (e.g., cristae density, as well as mitochondrial shape and interconnectedness), and respiration (including the capacity to oxidise different substrates). Factors such as physical inactivity, aging, and hereditary predisposition to Type 2 Diabetes (T2D) have been associated with altered mitochondrial characteristics (e.g., decreased mitochondrial volume, reduced oxidative capacity and metabolic flexibility in substrate utilisation, and a disconnected mitochondrial morphology) [10, 12, 13]. Their central role in cellular functions and associations with disease, and the complexity of its regulation - as will be highlighted, make skeletal muscle mitochondria an important target for scientific investigation.

### *1.1.1 Skeletal Muscle*

Skeletal muscles are highly specialised tissues with primary roles in locomotion, the maintenance of posture, and metabolism. Skeletal muscles are composed of individual muscle fibres - multinucleated single cells with a diverse range of properties that allow the same muscle

to be used for a broad repertoire of functional tasks. Repeated co-contraction of the myofibrils (composed of protein filaments actin, myosin, and titin) of muscle fibres, which rely on electrochemical depolarisation-induced release of $Ca^{2+}$ (Calcium ions) to allow actin-myosin binding followed by ATP (Adenosine Triphosphate) consuming repolarisation, cause the movement of muscles [14]. In humans, the skeletal muscles make up approximately 40% of the total mass and are comprised of 50-75% of the total body proteins [15].

Classification of muscle fibres based on their Myosin Heavy Chain (MHC) isoforms is now a well-accepted method and has broadly identified three major fibre types in humans – Type I, Type IIa, and Type IIx [16]. This characterisation is not absolute and only signifies the predominant MHC type within a muscle fibre; hybrid-type muscle fibres (Type I/IIa or Type IIa/IIx) are often observed [17]. Improvements in techniques of fibre-type identification and isolation have allowed the investigation of fibre-type specific adaptations to ageing [18] and exercise training [19] [20]. Divergent responses have been reported in the transcriptome and proteome between fibre types and when compared to whole-muscle samples, as well in response to exercise training [19], particularly with respect to metabolic proteins, which highlights new directions and hypotheses for future research.

Beyond its mechanical functions, of particular interest to physiologists is the contribution of skeletal muscle to whole-body metabolism, carbohydrate and fat storage, and as an amino acid reservoir [21]. Impaired plasma glucose uptake by the muscle, and increased intramyocellular accumulation of lipid species, are the hallmarks of the metabolic syndrome of diabetes and are central to the development of insulin resistance and Type 2 Diabetes [22] [12]. A healthy quantity of muscle mass also contributes to improved energy balance through expending greater energy in the muscle protein turnover process [23] [24] [25].

Structurally, functionally, and metabolically, the skeletal muscle is a highly adaptive tissue. For instance, with age, both the cross-sectional area and force production/strength of skeletal muscle reduces by 1 to 2 per cent per year [26]. Metabolically, the elderly have been shown to simultaneously demonstrate a higher resting metabolic rate but lower energy availability [27]. On the other hand, physical activity and nutritional interventions have been shown to increase muscle mass [28], efficiency in ATP production, skeletal muscle antioxidant capacity [29], and strength/power generation [30]; these adaptations all help to mitigate the effects of aging.

### *1.1.2 Skeletal Muscle Mitochondria*

The skeletal muscle is a highly metabolically active tissue and therefore is highly reliant on its mitochondria for sustained production of energy (ATP). Indeed, alongside cardiac (heart) muscle, skeletal muscle has been shown to have the highest mtDNA/nDNA (mitochondrial DNA/nuclear DNA) ratio compared to other organs [31] - hinting at the importance of this organelle for healthy muscle fibre functioning.

#### 1.1.2.1 Structure

The mitochondria are tubular-shaped, double-membraned, organelles. The mitochondria are known to house approximately 1300 proteins [32], which are predominantly localised to one of four ‘compartments’ - the matrix, the phospholipid inner membrane (IM) and outer membrane (OM), and the inter-membranal space (IMS) [1] [33] (Figure 1.1). The matrix space enclosed within the inner membrane contains the enzymes of the Tricarboxylic

Acid (TCA) Cycle and the beta-oxidation pathway and is the main site for those reactions. The outer phospholipid membrane is porous and is freely traversed by ions and small uncharged molecules through the action of pore-forming membrane proteins (porins), whereas larger

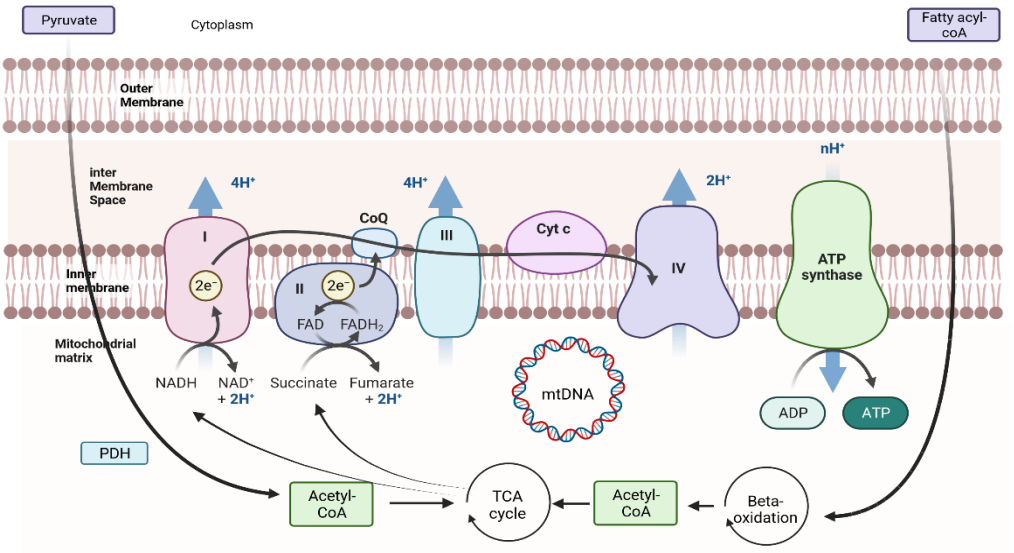


Figure 1:1: Overview of different mitochondrial compartments and the constituent major enzymes and proteins. Adapted "Electron Transport Chain Template" from Biorender

molecules, such as proteins, require membrane translocases. The inner membrane does not allow for diffusion of any ions and molecules but requires the function of membrane transport proteins, each of which are specific to the target and destination site. Due to this selectivity, an electrochemical membrane potential of approximately 180 mV builds up across the inner membrane and is the driving force behind the process of oxidative phosphorylation [34].

The inner membrane cristae hold the five mitochondrial respiratory chain complexes (CI, CII, CIII, CIV, and CV), along with the two main electron carriers - ubiquinone (CoQ), which is membrane-bound and functions to transfer electrons from CI and CII to CIII, and cytochrome c, which exists in the cristae lumen and transfers electrons between CIII and CIV (the main site of biological energy production) [35]. The cristate are dynamic projections of the inner membrane, which are thought to provide greater surface area to support increased respiratory capacity as evidenced by the distinctive cristae morphology of highly energy-

dependant cells such as skeletal and cardiac muscles [36]. Cristae morphologies are also thought to adapt within cells to control localised membrane potential and to promote higher energy production through higher-functioning ETC complexes [37]. Relatedly, the maintenance of these folded structures is integral to respiration and overall mitochondrial and cellular health. This function is performed by the subunits (currently includes seven) of the 'mitochondrial contact site and cristae organising system' (MICOS) complex [36].

Electron microscopy investigations [38] and immunoprecipitation analysis [39] have demonstrated the organisation of respiratory chain complexes into supramolecular structures called *supercomplexes*, instead of just independent entities embedded in the inner membrane [40]. Mainly, CI, CIII and CIV, the three proton-translocating complexes, have been found to associate in higher-order assemblies that have also been termed the respirasome (Figure 1.2) [38] [41], whereas CV is rarely observed as part of a supercomplex [42]. As with other features

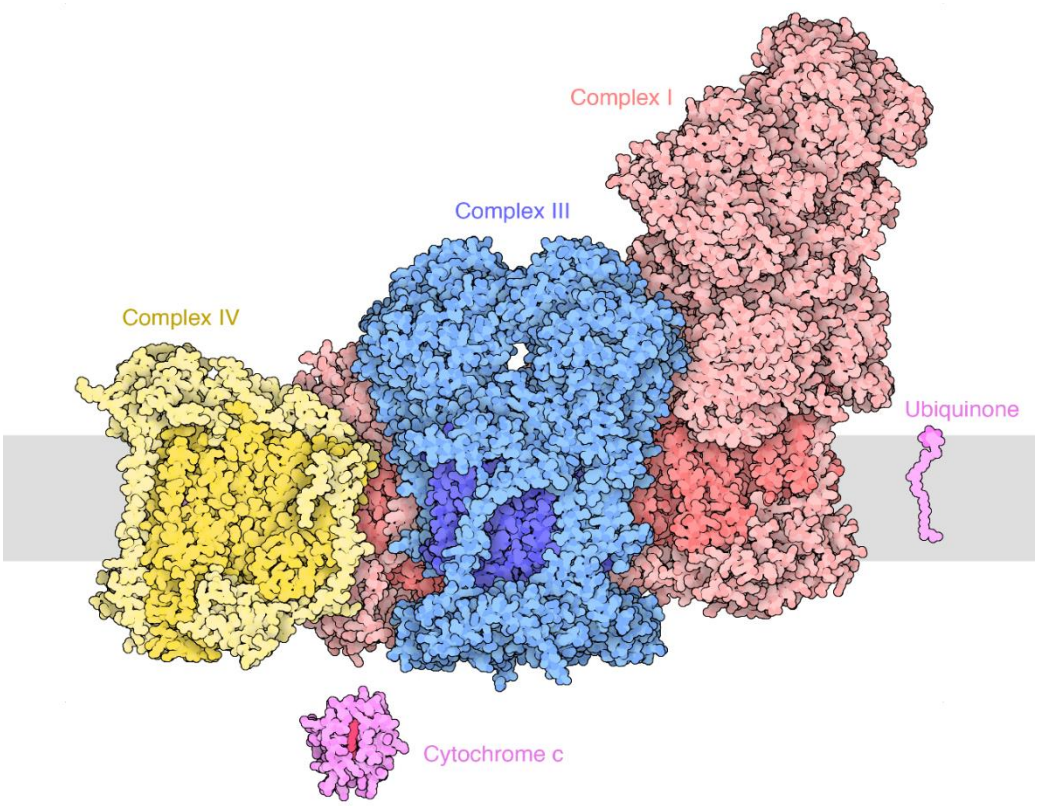


Figure 1:2: Representation of a CI-CIII-CIV supercomplex (respirasome). From https://pdb101.rcsb.org/motm/273s

of the mitochondria, it is proposed that assemblies of supercomplexes of various stoichiometries are dynamic and respond to oxidative stress and stimuli [40]. The inner boundary membrane also holds other key carrier/import protein complexes, such as the ATP/ADP translocase [34]. The intermembrane space, importantly, is spanned by intermediate formations of super-complexes of the TOM (Translocase of the Outer Membrane) and TIM (Translocase of the Inner Membrane) protein complexes, increasing the efficiency of protein-trafficking [43].

The mitochondria have their own set of genetic material (mtDNA), which exists within the matrix in the form of supramolecular assemblies called nucleoids - each of which holds one copy of the mtDNA and is held together in this structure by the action of TFAM (Mitochondrial Transcription Factor) [44]. Mitochondrial Ribosome units, which are formed of approximately 80 interconnected proteins (mitochondrial ribosome proteins; MRPS), also exist within the matrix space attached to the inner membrane boundary [45, 46].

#### 1.1.2.1.1 Network Morphology

Mitochondria are generally classified into two types based on their location within the muscle fibre, with the subsarcolemmal (SS) mitochondria being in proximity to the sarcolemma, allowing for reduced diffusion distances for oxygen transport from muscle capillaries, while the other mitochondrial population is distributed in the intermyofibrillar space (IMF), where they play a greater role in energy provision to the contractile proteins [15]. Despite the common representations of mitochondria as singular oblong structures, mitochondria exist as interconnected networks in skeletal muscle. They are highly dynamic

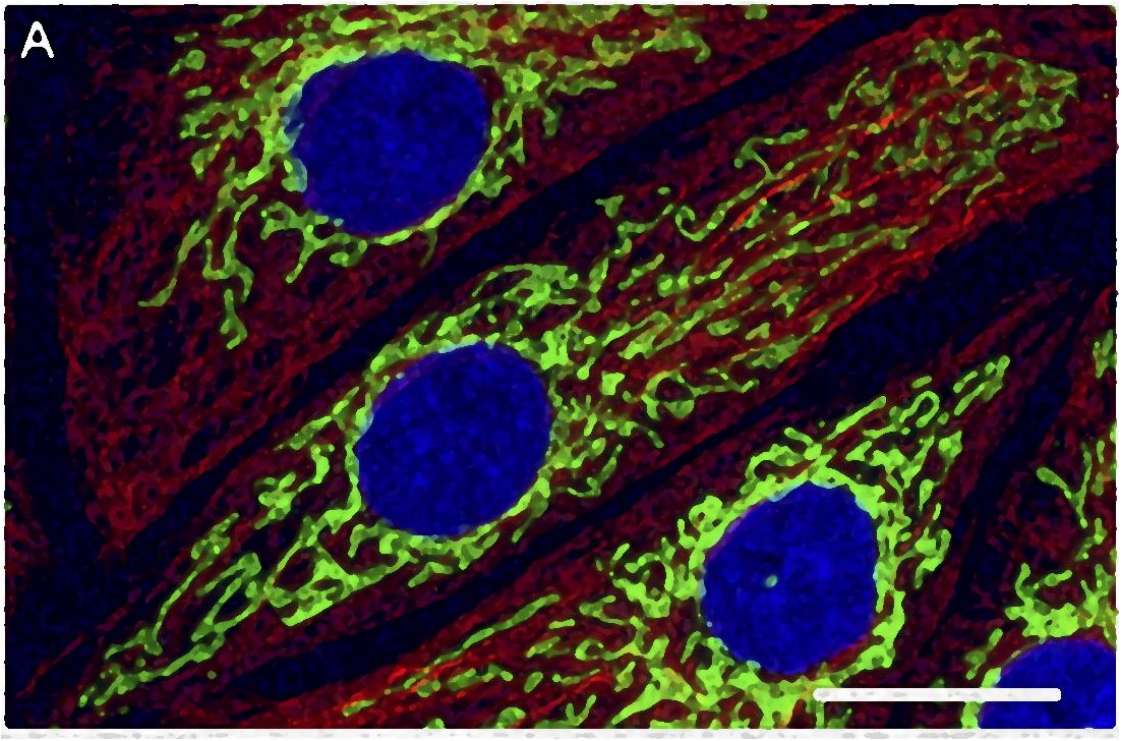


Figure 1:3 Fluorescence microscopy imaging of mitochondria (green) from mammalian endothelial cell lines demonstrating the interconnected nature of the mitochondria reticulum. The red filamentous structures in the background represent cytoskeletal tracks through which these organelles move in response to changing metabolic demands and energy states of the cell. From [3]

and move along cytoskeletal tracks in response to the changing metabolic demands of the muscle cell [47] (see Figure 1:3). Mitochondria in different subcellular localisations are connected by a membrane-like structure into a reticulum network that allows for the transport of oxygen and metabolic substrates within the organelles and plays an important role in the distribution of the inner membrane potential [48, 49]. Recent evidence through EM (Electron Microscopy) imaging also demonstrates the presence of nanotunnels within IMF mitochondria fractions in human skeletal muscle mitochondria networks. It is hypothesised that these double-membraned protrusions allow for 'mitochondrial content exchange' (transfer of small proteins, RNA molecules, and mtDNA), especially within physically constrained organelles, such as the IMF mitochondria that are tethered to myofibrils through the action of Desmin [50], that are unable to respond to adapt to changing metabolic/energy demands through morphological changes (e.g. fusion) [51, 52].

The adaptation of mitochondrial morphology is controlled by the processes of fission, fusion, and movement activities of the organelles [53]. The maintenance of the mitochondrial network requires a constant balance between the antagonistic processes of mitochondrial fusion and fission, collectively referred to as mitochondrial dynamics [12]. These adaptations fine-tune other processes, such as calcium homeostasis, ATP generation, Reactive Oxygen Species (ROS) production, and oxygen sensing in response to mitochondrial quality and the cellular environment. For instance, mitochondrial fusion and the expansion of the interconnected mitochondrial reticulum promote the capacity for oxidative phosphorylation (energy production) [54] through diluting the effects of defective mtDNA and promoting mtDNA inheritance and complementation (sharing mitochondrial components between organelles) [55, 56]. In contrast, mitochondrial fission facilitates the greater endoplasmic reticulum connectivity that plays a role in initiating cellular apoptosis and mitophagy [57]. Some of the key proteins implicated in the regulation of mitochondria dynamics are the dynamin-related GTPases, such as dynamin-related protein 1 (Drp1; alias: DNM1L), Mitofusin (MFN) 1/2 (regulates OMM fusion), OPA1 (regulates IMM fusion), Mitochondrial fission 1 protein (FIS1), and Mitochondrial fission factor (MFF) [58].

### *1.1.3 Mitochondria Formation*

The mitochondrial genome contains 37 genes, which encode for 22 tRNAs (transfer TNA), 2 rRNAs (ribosomal RNAs), and 13 proteins that are all part of the respiratory complexes and integral to the process of oxidative phosphorylation (OXPHOS) [59] [60]. However, most of the proteins found in the mitochondria (> 1300) [1] are nuclear-encoded, also referred to as NuGEMPS (nuclear genes encoding mitochondria proteins), and are

transcribed in the nucleus before being translated in the cytoplasm and then relying on protein translocation complexes and chaperone proteins for import, stabilisation, and targeted delivery into the mitochondria. Most cytosolic pre-proteins are imported into the mitochondria through the outer membrane by the TOM (translocase of the outer membrane) complex [61]. Each of the four intra-mitochondrial compartments, namely the matrix, IM, OM, and IMS, have their own respective translocase complexes that subsequently transport proteins to their destinations based on their signalling information. The small TIM (Translocase of the inner membrane) chaperones of the intermembrane space allow for the movement of the pre-proteins between the mitochondrial membranes. TIM 22 and TIM23 are the two main complexes for delivery of pre-proteins to their sites on the inner membrane and into the matrix, respectively [61]. The Mitochondrial import complex (MIM) on the outer membrane promotes the import and insertion of several outer-membrane-specific proteins [62]. The process of protein import through the inner membrane is ATP-dependent [63]. The mitochondrial import and assembly machinery (MIA) handles the import and assembly of inter-membranal space proteins. The matrix-bound proteins are further processed and folded via the actions of mitochondrial processing peptidase (MPP) and matrix chaperones and chaperonins - principally HSP60/HSP10 and HSP70 (Heat Shock Proteins) [64-66]. Defects in mitochondria protein import and assembly mechanisms have been shown to be linked with ageing, increased oxidative stress, and metabolic disorders, and have further been shown to be positively associated with chronic muscle stimulation [67, 68].

### *1.1.4 Regulation of Mitochondrial Content*

Mitochondrial content[2] in the skeletal muscle tissue is governed by the opposing forces of mitochondrial biogenesis (the synthesis of new components of the mitochondria reticulum) and mitophagy (the removal of organelles marked for degradation) [53]. Cellular energy-sensing proteins, such as AMPK (AMP-activated protein kinase) and CaMK ($Ca^{2+}$/calmodulin-dependent protein kinase) [69], are activated in response to the changing metabolic demands induced by environmental and physiological perturbations. These kinases induce signalling cascades that cause activation of chromatin deacetylation and transcription factors, which finally result in upregulation of mitochondrial biogenesis. The key group of kinase-activated transcription factors include PGC1 (Peroxisome proliferator-activated receptor-gamma coactivator) proteins, MEF2 (myocyte enhancer factor-2), and HDACs (Histone deacetylases) [5] [70]. Increasing evidence suggests that the PGC1 proteins, particularly PGC1-alpha, are the master regulators for mitochondrial biogenesis through the co-regulation of Nuclear Respiratory Factors (NRF1, NRF2), which upregulate the transcription of NuGEMPs, and their combined co-regulation with ERRs (oestrogen-related receptors) of TFAM – the final effector of mtDNA transcription and replication [69, 71]. Another key group are the sirtuin (SIRT) proteins that are activated in response to flux in $NAD^+$/NADH ratios, another marker of cellular energy demand, and 13eacylated the PGC1-alpha protein [10].

[2] Within this thesis, mitochondria content is defined as the total amount of mitochondria in skeletal muscle as measured via either morphological (mitochondrial volume density using TEM microscopy for example) or biochemical (measurement of representative biomarkers; see for example [69]) measures, unless defined otherwise.

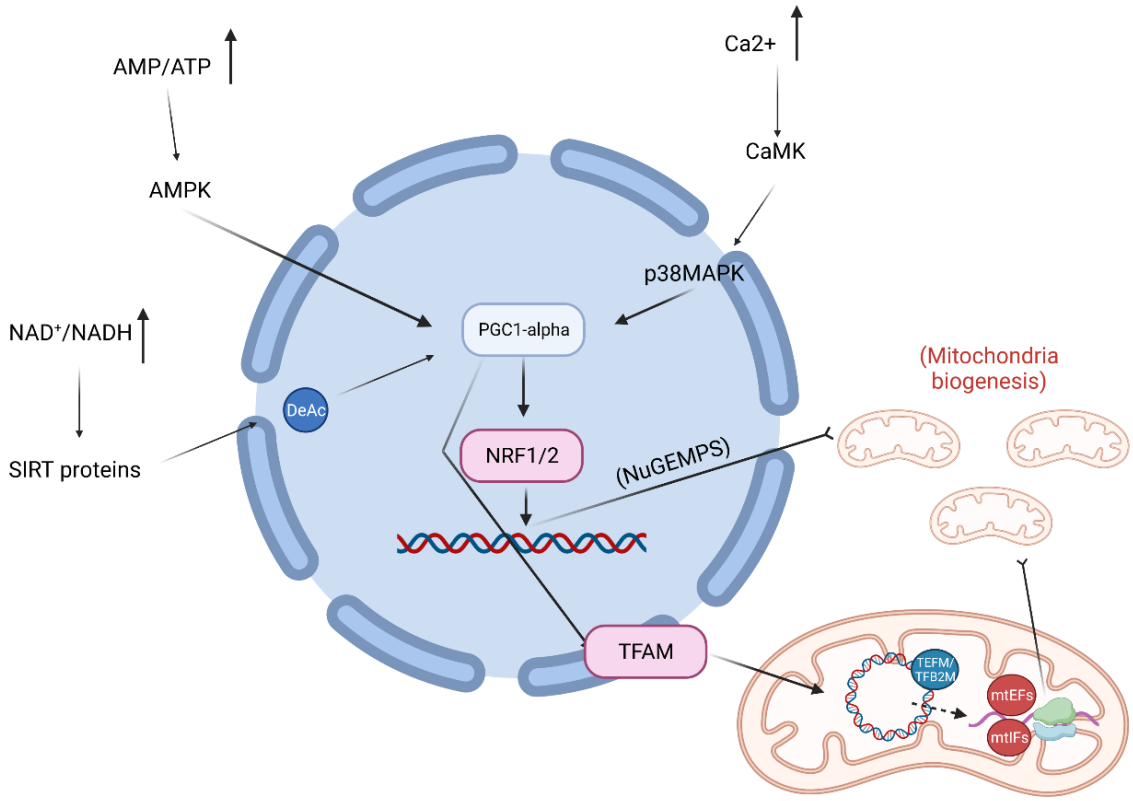


Figure 1:4 Overview of key proteins and their roles in the mitochondria biogenesis pathway. Abbreviations are defined in the text. Created in Biorender.com

Transcriptional control of mitochondrial genes is also performed by key nuclear-encoded regulatory proteins such as transcription factors (TFB2M, TFB1M) and elongation factors (TEFM) [72]. Further, nuclear-encoded translation regulators such as initiation factors (mtIF1, mtIF2), mitochondrial elongation factors (mtEFTu, mtEFTs, and mtEFG1), and mitochondrial release factor (mtRF1L) control the process of translation. Additionally, translational activators, such as the Translational Activator of Cytochrome C Oxidase I (TACO1), bind to the mitochondrial RNA to regulate the rate of protein translation [73].

A decrease in the inner membrane potential and elevated ROS are some of the key signalling events for the initiation of the mitophagy process through the PINK1/Parkin cascade, the most well-known mitochondrial degradation pathway [10, 58]. These signals cause the accumulation of the PTEN-induced kinase 1 (PINK1) protein on the OM, which

phosphorylates several downstream targets with the ultimate objective of recruiting the E3 ubiquitin ligase protein Parkin to the mitochondria. This protein ubiquitinates several other OM substrates, such as VDAC1 (Voltage Dependant Anion Channel 1), that subsequently bind to the P62/SQSTM1 adaptor protein, initiating the process of autophagosome formation and its transportation to the lysozyme [10] [74]. Other key relevant proteins, such as BCL2 interacting protein 3 (BNIP3), B-cell lymphoma 2-like protein 13 (BCL2L13), and Nip3-like protein X (NIX), are also implicated in this process through their direct interactions with the LC3 protein (microtubule-associated protein 1 light chain 3) – a binding molecule on the autophagosome membrane, in pathways independent of PINK1/Parkin [58] [74]. Finally, the balance between activities of fusion and fission proteins, which are described in more detail in the following section, also play a critical role in regulating mitophagy activity [75].

### *1.1.5 Regulation of Mitochondria Function*

Mitochondrial content and respiratory function have been intimately linked. For instance, mitochondria content (mitochondrial volume density measured with TEM) in human skeletal muscle has been shown to be moderately, yet significantly, correlated with mass-specific mitochondrial respiration (picomoles of oxygen consumption per second per milligram; r=0.64) and maximal oxygen consumption ($VO_{2max}$ measured as L/min/kg; r=0.59) [76]. Relatedly, increases in mitochondrial respiration in response to high-intensity interval training (HIIT) are no longer apparent when normalised for the increase in mitochondria content (as measured by CIV activity), suggesting that adaptions to skeletal muscle respiratory capacity can often be largely attributable to an increase in mitochondria content and not qualitative changes in mitochondria respiratory function [76]. Nonetheless, studies in both

animals [77] and humans [78] have sometimes reported an increase in functional markers, such as mass-specific maximal respiration and ADP sensitivity, independent of changes in mitochondria content in response to exercise training. Physiological variables, such as glucose homeostasis and insulin signalling, have also been found to be reduced in fusion-protein-deficient mice despite the absence of changes in mtDNA content and relative abundance of ETC subunits [79]. Cross-sectional analysis of data comparing individuals across different exercise/training statuses demonstrates relatively greater adaptations in mass-specific mitochondria respiration than markers of mitochondria content [80] – further suggesting qualitative adaptations in mitochondrial respiratory function. Finally, recent work from the Bishop lab has highlighted that changes in individual mitochondrial proteins, which are not stoichiometrically linked to the overall increase in mitochondrial content, may contribute to changes in mitochondrial respiratory function with training [81].

A major, unresolved topic in the exercise-physiology field, therefore, concerns the underlying mitochondrial adaptations that contribute to improvements in respiratory function /capacity beyond the well-characterised adaptations in ETC complex subunits and the overall increase in mitochondria content. In disease and dysfunctional states, such as cardiovascular and metabolic disorders, fragmentation of the mitochondria network and reduced fusion capacity has been observed [82]. These morphological changes have further been linked to reduced OXPHOS function (measured as decreased mitochondrial respiration), elevated ROS production, insulin signalling, and mtDNA depletion through gene-specific knockout models in mice [82-85]. Thus, mitochondria-associated respiratory and signalling functions are negatively affected by a more fragmented mitochondrial network – even in the absence of changes in mitochondrial content [84].

It has also been suggested that the assembly/formation of *supercomplexes* is dynamically responsive to cellular metabolic state and influences maximal respiratory capacity

[86, 87]. The biogenesis and structural organisation of ETC proteins along the mitochondria inner membrane, regulated through complex control of phospholipids, is integral to the formation and maintenance of these supercomplexes [88]. However, there is also contrasting evidence to suggest that enhanced mitochondrial bioenergetics observed with greater supercomplex formation may simply be attributable to a higher overall mitochondrial content [81, 89].

Changes in many different mitochondrial pathways may also contribute to altered respiratory function in the absence of changes in mitochondrial content. One of the key aspects of mitochondria quality control is the management of intra-mitochondrial ROS levels. High levels of ROS cause oxidative stress on the cell and can be severely damaging for cell constituents, reducing DNA stability, and affecting protein structure. Accumulation of ROS within the mitochondria can also cause the opening of mitochondrial channels leading to the release of ATP, $Ca^{2+}$, and cytochrome c, which can lead to cellular apoptosis [90]. Low levels of ROS production are an inevitable by-product of mitochondrial respiration and predominantly occur at sites on respiratory complex I and III through partial oxidation of molecular $O_2$. Elevated ROS production is substantially linked to mitochondrial metabolic states [91], with it being highest during conditions of low oxygen consumption through oxidative phosphorylation (low ATP demand) and high inner membrane potential as complexes are in their reduced states [92]. Phenotypically, consistently elevated levels of ROS have been shown to cause fragmentation of the mitochondrial reticulum and upregulate expression of fission and mitophagy genes [93, 94].

Prominent defences against ROS accumulation are proteins that make up the antioxidant system, converting oxygen anions to less reactive states (such as Hydrogen Peroxide), and the mitochondrial uncoupling mechanism that lowers the membrane potential and discourages the formation of ROS species [91, 92, 95]. Key enzymatic proteins involved

in the process of dismutation and scavenging of $O_2$ include Superoxide dismutase (SOD) and Glutathione (GSH), Glutathione peroxidase (GPX), Thioredoxin (TRX) and catalase [93]. Uncoupling of ETC flux with ATP production is performed mainly through the UCP (Uncoupling Protein) family of proteins and the ANT (Adenine Nucleotide Translocase) proteins through affecting ADP transport back into the matrix and promoting proton leakage back across the membrane [13, 92]. A decline in enzymatic proteins and impaired ADP sensitivity has also been associated with reduced oxidative capacity and ATP synthesis in ageing and in insulin-resistant populations [96, 97].

Another important aspect of mitochondrial quality control is the modulation of the endoplasmic reticulum (ER)-mitochondrial contacts, also known as mitochondria-associated ER membranes (MAMs). These sites have a vital role to play in $Ca^{2+}$ trafficking, lipid transport, and autophagy, among others. Beyond the fusion proteins discussed before, other crucial proteins such as VDAC, VAPP (Vesicle Associated Membrane Protein B), PACS-2 (Phosphofurin Acidic Cluster Sorting Protein 2), and RMDN (Regulator of microtubule dynamics) proteins act as tethers, physically connecting the outer mitochondrial membrane and the endoplasmic reticulum [98, 99]. The localised MAM molecular chaperone Glucose Regulated Protein (Grp75, alias: HSPA9), a cytosolic heat shock protein, allows $Ca^{2+}$ transport between VDAC and the ER calcium-release channel and is key to the maintenance of calcium homeostasis, as evidence suggests that deficiency of this protein causes a decrease in ER-mitochondria $Ca^{2+}$ exchange [100] [99]. The maintenance of the mitochondrial membranes requires phospholipid synthesis in the ER and transport to the mitochondria through the action of MAM localised phospholipid and glycosphingolipid-synthesising enzymes, which highlights another integral function of MAM structures [101]. Evidence from the skeletal muscle of mice and humans with T2D suggests that disruption of these mitochondria-ER

interactions contributes to reduced insulin sensitivity and thus the development of insulin resistance [102].

Proteostasis, the regulation and degradation of misfolded and damaged proteins, is another important factor in the maintenance of mitochondrial health [103]. Skeletal muscle samples from aged human and animal populations demonstrate a higher accumulation of damaged proteins in the mitochondria, and it has been suggested that this loss of proteostasis may have a causal role to play in the functional decline associated with age [104]. Further, protein turnover data from model organisms also demonstrate that new protein synthesis does not decline with age, and, therefore, pathways regulating protein assembly and degradation may be affected [104]. The protein quality control mechanism is mainly composed of two functional groups – chaperones and proteases. Chaperones, such as the mtHSPs (mitochondrial Heat Shock Proteins), fold and assemble imported proteins and re-fold damaged proteins. Proteases, specific to each compartment of the mitochondria, digest the excess damaged proteins [105]. This quality control mechanism is regulated by the mitochondrial Unfolded Protein Response (mtUPR), a retrograde pathway that upregulates the nuclear transcription of chaperone and protease genes in response to proteotoxic stress [106]. Elevated ROS accumulation, elevated protein damage, and mito-nuclear imbalance, where there are disproportionate volumes of mitochondrial and nuclear-encoded protein subunits, are some of the causes of proteotoxic stress [105]. Chaperone and protease gene and protein expression has been shown to be strongly correlated in human and mouse populations, further demonstrating co-regulation of the UPR response [107].

## 1.2 Skeletal Muscle Mitochondria and Exercise Training

As has been highlighted so far, the maintenance of healthy mitochondria content and function relies on many interconnected pathways. The following sections will discuss the evidence around exercise-dependent regulation of biomarkers of mitochondria content and some key functional pathways.

### *1.2.1 Adaptations to markers of Mitochondria content and Biogenesis in Animals*

Interest in the study of mitochondria adaptations to an exercise/training stimulus arose from the findings of greater respiratory enzyme content in the skeletal muscle of animals demonstrating higher degrees of physical activity [108] [109]. To understand the factors regulating these adaptations, Holloszy, in 1967, studied changes in mitochondrial protein content in rats following 12 weeks of endurance training (120 min of continuous exercise, with 12 sprints at 42 m per min, each lasting 30 seconds, interspersed at intervals through the workout, 5 days per week for 12 weeks) [110]. A significant increase in respiratory enzyme activities, cytochrome c content, and total protein content, of the mitochondria fraction was widely reported for the first time–- establishing upregulation of mitochondrial biogenesis in response to exercise. Greater coupling between the ETC and ATP-synthase (i.e., greater ATP-producing capacity from the OXPHOS mechanism) was also observed [110]. This was not the very first finding of its kind, however, as in 1939 Chepinoga demonstrated an increase in mitochondrial respiration (measured as rate of oxygen consumption) and SDH (Succinate Dehydrogenase) activity following electrical stimulation of rabbit skeletal muscles over 14-25 days [111]. Failures to detect such significant changes in similar previous trials [112, 113] had

largely been attributed to insufficient duration and intensity of the exercise exposure [110], highlighting the important role of exercise-training variables for mitochondrial adaptations.

Following these major findings, multiple other studies have now validated these results using the protein content of CIV subunits and cytochrome c, ETC Complex activity, and transmission electron microscopy-derived mitochondria density, as proxies for mitochondria content [114] [115] [116]. Affecters of mitochondrial biogenesis, such as PGC1, NRF, TFAM, and SIRT proteins, have expectedly also been found to be upregulated with exercise training [5] (citing others). The broad consensus within the area is that PGC1-α is a central node in exercise-induced mitochondrial biogenesis response–- due to findings of increased mitochondrial gene expression, increased oxidative metabolism/mitochondrial respiration, and increased mitochondrial mass following its specific overexpression in model organisms [117]. However, contrasting evidence to the centrality of this transcriptional activator does exist, especially as PGC1 knockout organisms have not widely been shown to demonstrate a blunted adaptive response to exercise compared to wild type; this suggests that there may be other compensatory mechanisms regulating mitochondrial biogenesis [114] [117].

### *1.2.2 Adaptations to markers of Mitochondria content and Biogenesis in Humans*

Following findings of exercise-induced mitochondria biogenesis in mice, similar studies were designed to observe these adaptations in humans. The *Bergstrom* needle technique for biopsies, developed in the 1960s, allowed for rapid and less traumatic sampling over previous methods and accelerated skeletal muscle research in humans [118]. Numerous studies have confirmed increases in markers of mitochondria content (e.g., ETC complex proteins [119, 120] and elevated citrate synthase (CS) activity [121, 122]) following endurance exercise

training protocols. As with animal studies, exercise training has been shown to induce increases in the content of proteins associated with mitochondrial biogenesis, such as PGC1-alpha and p53 (another regulator of mitochondrial transcription machinery) [123], NRF and TFAM [124], and sirtuins (SIRT) [125]; these adaptations highlight potential energy-sensing and calcium dependant mechanisms [126] [127] through which mitochondrial protein content is increased through exercise training [69]. Exercise training has also been found to directly upregulate protein translation through its effects on increasing mitochondria ribosomal protein (MRP) gene expression and protein abundance [128, 129] [81]. The magnitude of change in mitochondrial content and the degree of activation of genes in the mitochondria-biogenesis pathways has been shown to vary with training volume, intensity, and modality [8], and participant health characteristics [130].

### *1.2.3 Adaptations to specific Mitochondria Proteins*

Since the development of the western blot technique (immunoblotting) in the early 1980s, an antibody-based method for the relative quantification of targeted proteins, this technique has increasingly been utilised in the field of exercise physiology to determine adaptations to specific skeletal muscle proteins, including mitochondrial proteins [131]. Previously, quantification of mitochondrial adaptations was limited to the total protein content of mitochondria-enriched fractions determined through methods such as the Bradford assay [131]. With the application of this technique, researchers have investigated the regulation of specific proteins within different aspects of mitochondria function, as shall be discussed below.

#### 1.2.3.1 Adaptations to markers of Mitochondria Function in Animals

The content of representative protein subunits of all five ETC complexes and cytochrome c has been shown to be upregulated following exercise training in mice, with the Complex IV subunits being the most studied [132] [133] [134]. Beyond proteins directly involved in OXPHOS, multiple other functional aspects have been found to be regulated with exercise. TCA cycle proteins, such as GADPH, PDH, and PDH kinase (PDK4), are upregulated through exercise, adapting to increased flux through the TCA cycle to increase electron supply to the ETC and thus improve respiratory capacity [132, 133, 135].

Both mitochondria fusion and fission processes appear to be regulated through exercise training in mice, with findings of increased abundance of MFN and FIS proteins following training, which could indicate increased capacity to incorporate new elements into the mitochondria reticulum and to turnover damaged mitochondria [136, 137]. Expression of Heat Shock Proteins (HSPs), TOM proteins, and chaperonins (Cpn10), integral components of the Protein Import Machinery (PIM), is increased through exercise-induced stress [138] [139]. On the other hand, increased mitophagy capacity has also been demonstrated in rodents through elevated levels of the BNIP3 protein, a key marker of this process, further suggesting adaptation of increased mitochondria turnover capacity following exercise training [140]. Despite the observed increase in mitophagy proteins, however, available evidence, mostly from animal studies, does not always demonstrate an increase in basal mitophagy flux. Decreased mitophagy flux has been demonstrated in rat skeletal muscle following nine days of chronic contractile activity through nerve stimulation [141]. Similarly, basal mitophagy flux is decreased in rats following 6 weeks of endurance training despite increased abundance and localisation of Parkin protein [142]. Related to this, HIIT in mice has been shown to improve

mito-nuclear imbalance and unfolded protein response (UPRmt) gene expression and protein content, affecting mitochondrial proteostasis and function [143]. Finally, Uncoupling protein abundance, specifically UCP3 (alias: SLC25A9), has also been found to increase in response to graded acute exercise exposure (between 30 to 200 minutes at approx. 70% $VO_{2peak}$), hypothesised to protect against elevated mitochondrial ROS production induced by repeated contractile activity [144].

#### 1.2.3.2 Adaptations to markers of Mitochondria function in Humans

Many, but not all, of the immunoblot results from animal studies that have investigated mitochondria protein adaptations to exercise have been replicated in human trials. Owing to the ease of designing exercise interventions, and altering modalities and intensities, there is greater breadth and depth of evidence available in exercise research in human studies, as will be highlighted below.

As observed in rodents, the relative protein abundance of all five respiratory chain is responsive to exercise training in humans [120] [86] [145] [146]. However, not all complexes have been found to be similarly regulated and have instead demonstrated variability in their magnitude of adaptation. For instance, it was reported that Complex I is the most upregulated and does not show stoichiometry to the other protein complexes, which are positively correlated with each other, in response to a 4-month, moderate-intensity, exercise intervention in sedentary older men and women Greggio, Jha [86]. Similarly, the Complex I subunit, followed by Complex IV subunits, demonstrated higher fold changes than the other Complex markers following a 2-month progressive endurance training program [146]. Interestingly, no significant changes in baseline or acute (1 and 4 hours following the exercise session) gene

expression of the Complex markers were observed. Contrastingly, significant upregulation of all but the Complex I subunit has been observed in response to Sprint Interval Training (SIT) and Medium Intensity Continuous Training (MICT) for 12 weeks in healthy young men [147]. Moreover, the SIT protocol elucidated equal, if not greater, changes in the complex subunits despite a five-fold lower exercise volume.

Electron Transport Chain complex subunits have also been shown to respond differently and uniquely to exercise volume/exposure in healthy young men. Four weeks (3 times/week) of normal-volume HIIT (NVT) did not cause an increase in the subunit protein abundance of any complex subunits. However, subsequent exposure to the same exercise modality for three weeks (2/day) (HVT), which was nearly 4 times the exercise volume of the NVT block, induced an increase in protein abundance of all but the CIII subunit compared to post-NVT. Only when accounting for the NVT phase, and comparing post-HVT against baseline, was an increase in CIII subunit protein abundance observed. Further, the subsequent reduction in training volume also only downregulated subunits of CI, CII, and CIV, perhaps suggesting exercise-volume dependant regulation of the different complexes [120]. Similarly, four weeks of either SIT or HIIT in healthy young men did not lead to an increase in abundance of any of the ETC complex subunits [123],

The relative protein abundance of metabolic enzyme proteins, such as β-HAD (3-Hydroxyacyl Coenzyme A Dehydrogenase), SDH, and CS, have also been shown to be upregulated in response to training [148]. Contrastingly, isocitrate dehydrogenase (IDH), another TCA cycle enzyme, has been shown not to increase with exercise volumes sufficient to induce mitochondria biogenesis; this suggests differential regulation of TCA cycle components with exercise [119]. Fatty-acid Translocase protein content (FAT/CD36) in isolated mitochondria fractions and whole-muscle CPTI (Carnitine Palmitoyl transferase I) activity, a key enzyme in mitochondrial acyl-CoA import, has been found to increase in

response to 2 and 6 weeks of HIIT training [149, 150]; this highlights potential mechanisms that may contribute to the elevated basal fat oxidation rate observed in trained individuals [151].

As observed in animal studies, increases in fission (FIS1 and DRP1) and fusion (MFN1, MFN2 and OPA1) proteins have been reported following some training interventions [148, 152]. Stratifying gene expression and protein content data by time course and intensity of exercise training intervention, it was hypothesised that fission proteins increase early in the exercise intervention to improve mitochondria quality through remodelling and removing defective sites and adapt to increased energy demand followed later by pro-fusion dynamics to enlarge the mitochondrial reticulum [153]. This would explain the increased abundance of both fusion and fission proteins following training. Chronic exercise training has generally been shown to promote pro-fusion mitochondria dynamics in human skeletal muscle. An increase in fusion-to-fission protein content ratios has been observed following 12 weeks of continuous endurance training in older men [154]. Lifelong exercise-trained individuals demonstrate markedly increased protein content of fusion proteins MFN1 and MFN2, decreased protein content of fission protein DRP1, and increased phosphorylation at $Ser^{637}$, which inhibits DRP1 activity and translocation [155] [74]. Furthermore, SIT, HIT, and MICT protocols have all been shown to affect an increase in mitochondria dynamics-related proteins [153] [156]. There, however, appears to be a threshold effect of exercise volume/exposure as low-volume SIT-based protocols have failed to demonstrate changes in gene expressions of MFN2 and DRP1 - perhaps suggesting greater dependence on exercise volume than intensity [157].

One of the key avenues through which exercise impacts mitochondrial health is through attenuation and management of ROS accumulation, which has been shown to be implicated in pathologies such as insulin resistance [22] [158]. Exercise training has been shown to improve enzymatic handling of ROS molecules through increased expression of SOD2 protein [159]

[160]. Further, life-long physical activity has been shown to elevate circulating levels of GSH, another enzymatic antioxidant, in older individuals compared to sedentary age-matched controls. The production of ROS molecules is also more tightly regulated in exercise-trained individuals through elevated UCP and ATN2 protein content, which dissipates membrane potential and thus reduces the propensity for ROS formation [161].

Turnover of damaged and dysfunctional mitochondria is also found to be upregulated in exercise-trained individuals. Exercise training interventions have been shown to induce increases in the capacity for mitophagy through upregulating protein expression of the LC3-1/BNIP3[162] pathway and, more recently, through upregulation of B-cell lymphoma 2-like protein 13 (BCL2L13) [74]. These mechanisms are implicated in mitochondria docking to the LC3 protein, the initiator of autophagosome formation, and are known to be able to function independently of the well-known PINK/Parkin pathway [163]. It has also been demonstrated that gene expression of BNIP3 and LC3, pro-fission DRP1 protein content, and autophagy flux, is upregulated in Parkin knockout animal skeletal muscle; this further highlights the independence of these pathways and suggests compensatory autophagy mechanisms in conditions of restricted mitophagy [164]. Further, habitual exercise training has also been demonstrated to increase the protein content of the PINK1 and associated proteins (Parkin and VDAC1) compared to controls [58] [74]. Similar to observations in animal studies, as discussed previously, LC3II and p62 protein contents, two markers of autophagy flux, in mitochondrial fractions are not found to increase despite an overall change in mitophagy-implicated proteins following 8 weeks of endurance training in humans [162]. The developing hypothesis is that exercise training results in overall improved mitochondria quality due to accumulated changes following individual exercise sessions, therefore reducing the need for elevated basal mitophagy flux in the trained state, while maintaining the capacity to further respond to greater energetic stress through increased levels of mitophagy proteins [165]. The

precise regulation of mitophagy response to exercise in humans is not fully resolved yet and requires a broader profiling of related markers in participants that differ in health and training status.

With regard to protein import and mitochondria assembly machinery, extensive evidence of exercise-induced regulation is lacking in humans [166], though the upstream mechanistic basis for this adaptation has been explained [167] and demonstrated in animal studies as highlighted previously. Increased expression of mitochondrial chaperone proteins Hsp60 (*alias*: HSPD1) and GRP75 (*alias*: HSPA9) has been demonstrated in response to combined endurance and resistance exercise over two years in middle-aged, sedentary adults [168]. Similarly, an increase in chaperone mtHSP70 protein content, but not of TOM20 - a sub-unit of the outer membrane translocase complex (TOM), has been found following 14 weeks of moderate-intensity endurance training in middle-aged, healthy adults [169]. The logical

hypothesis is that to match the increased mitochondria biogenesis flux with exercise, concomitant increases in the import and assembly complexes should be observed.

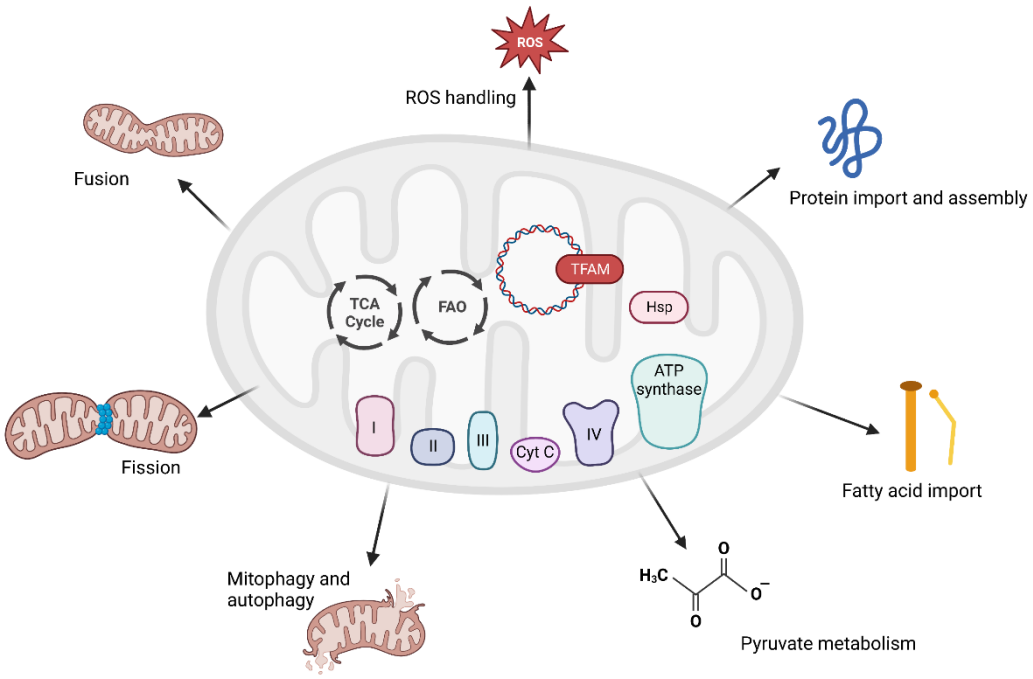


Figure 1:5 Overview of mitochondrial component and pathway proteins that have been shown to be regulated with exercise training according to western blot evidence from either human or animal studies. Created in BioRender.com

### *1.2.4 Limitations of Immunoblotting*

As demonstrated above, many mitochondria proteins are altered with exercise training (Figure 1:5), with evidence for some more extensive than others. However, the proteins discussed in the previous sections only account for a small fraction of the mitochondrial proteome (< 5%) [170], and only reflect the mitochondrial proteins that have been most studied using immunoblotting methods. Despite the obvious utility of western blot methods, they do possess limitations. One key limitation is that the target of analysis must be determined *a priori*. Therefore, this approach is limited to mitochondrial proteins for which commercial antibodies

exist or allow proprietary development and does not lend itself to exploratory data acquisition and analysis. Further, these methods only allow for relative quantification between conditions/controls, are not suited for absolute abundance analysis, and antibodies often demonstrate limited specificity and sensitivity [171]. Also, the accuracy of quantification between experimental conditions is prone to be affected by errors in sample loading, antibody concentration, and saturation of the linear dynamic range, among others [172, 173]. Aside from the fact that commercial antibodies exist for only a fraction of known mitochondrial proteins, if an attempt was made to measure mitochondrial protein abundance at the scale possible with high-throughput techniques, the costs would be unreasonably high.

Another limitation is that antibody-based methods of protein identification often do not allow for valid estimates of mitochondria content/volume. Normalising protein abundances against valid markers of mitochondrial content can allow insight into mitochondrial adaptations that are independent of changes in mitochondrial content [8]. Various markers of mitochondria content, such as mtDNA content, Complex I-V content/activity, and CS activity, have previously been used to indirectly assess changes to mitochondrial content in physiological studies, which include studying the effects of exercise training [174] [175]. While CS and Complex I activity has shown the strongest correlation ($r=0.77$; $p < 0.02$) for mitochondria content, mtDNA concentration and Complex I protein content were found to be poor markers for mitochondria content. [174]. Furthermore, a limited utility of these biomarker-based methods to measure the relative change in mitochondria content in contexts of physiological interventions, especially exercise, has been argued [175] and demonstrated [176]. The adoption of high throughput mass spectrometry-based data collection has allowed for the development of more robust *in silico* normalisation techniques that incorporate a greater breadth of the data [81, 177].

To circumvent many of these limitations discussed, as the subsequent sections will highlight, high-throughput methods of protein analysis are now being employed within exercise physiology to provide greater breadth and resolution to exercise-mediated regulation of skeletal muscle mitochondrial proteins [178].

## 1.3 Proteomics

### *1.3.1 Overview of Proteomics Technology*

Originally, the term proteomics was used to describe the method of two-dimensional gel electrophoresis (2-DE) separation, manual digestion of selected protein spots, and subsequent identification of digested spots using automated protein sequencers. While this method allows for greater resolution of the proteome than the western blot technique, it suffers from limited accuracy of quantification as reliability determining relative intensities for hundreds of spots is problematic even for the sophisticated image analyses software packages used [179]. Advances in mass spectrometry technology in the 1990s allowed for the greater sensitivity required for the analysis of gel-separated proteins, combined with the development of computational resources to correlate mass spectrometry output with sequence databases, which led to mass spectrometry being used for protein identification in tandem with separation techniques [179] [180]. Another major advancement that allowed for this fundamental shift in protein analysis was the development of the two ionisation methods, namely MALDI (Matrix Assisted Laser Desorption/Ionisation) and ESI (Electrospray Ionisation), which allowed for the generation of peptide/protein ions without significant analyte fragmentation [181]. To overcome limitations of 2-DE, including limited sample carrying capacity, limited detection

sensitivity of low-abundant proteins, co-migration of proteins to the same spot, and, most importantly, a lack of amenability to automated and high-throughput analysis, microcapillary RP-LC (Reversed Phase)/HP-LC (High-Pressure Liquid Chromatography) separation protocols were introduced.

### *1.3.2 Workings of high-throughput Mass Spectrometry-based proteomics*

A typical proteomics experiment can be broken down into roughly five stages. First, the proteins to be analysed are isolated from the sample of interest through methods of cellular lysis and fractionation [182]. Next, the complex protein mixture obtained from the sample is typically digested into smaller peptides using proteases, mostly trypsin, and are introduced into the μLC column (micro–Liquid Chromatography Column) for separation based on their physical attributes such as mass, size, or hydrophobicity [180]. Trypsin is generally chosen as the primary protease as it allows cleavage at the carboxyl terminal of the lysine and arginine residues resulting in positively charged peptides that are beneficial for MS analysis [183]. Complementary proteases are often used in conjunction with trypsin to target different amino acid residues and obtain more complex peptide mixtures. For example, Asp-N and Glu-C, as the name suggests, target cleavage at positively charged amino acid residues of Aspartate and Glutamate, respectively [183]. Peptide digestion is then undertaken, as MS of whole proteins is generally less sensitive than peptide MS and intact protein masses are insufficient to make accurate identifications [182]. Further, peptides are also more easily separable by chromatography methods, ionise better, and fragment in a more predictable manner. Experimental set-ups using peptides as the unit of analysis are termed bottom-up proteomic approaches [184].

Following injection into the μLC columns, depending on their characteristics, peptides elute at different times, and this is termed the retention time of those peptides [185]. Multiple different peptides may co-elute together at a given time, and this means that retention time is not sufficient alone to allow for peptide identification. In contrast to gel-based proteomics, continuous separation via HPLC/RPLC can be coupled with continuous ionisation sources, giving rise to automated LC/MS setups [186]. Following elution, the peptides need to be ionised. This is a crucial step for mass spectrometry analysis as peptides in solution or solid-state need to be converted into small, nebulised ions in the gas phase before being subject to the mass analyser [186]. Within MS instrumentation, in general, ion sources are coupled with a mass analyser to detect m/z (mass-to-charge) ratios of eluted ions and a detector to measure the intensity at each m/z value.

The fundamental principle behind quantification and detection in proteomics is as follows - ionised peptides, after eluting at a given time, are measured for their m/z ratio and intensity. In most protocols, computationally prioritised peptides are then selected further for an MS-II (MS/MS) scan in the adjacent instrument for sequence identification of the peptide following fragmentation into product ions [182]. This method is also referred to as tandem MS. Quantification of the peptide is measured either via the peak intensity of the MS-I scan or the area under the curve of the retention time at a given m/z. The former method has been shown to correlate linearly with protein concentration ($r^2$>0.9) in samples of horse myoglobin digests [187].

Following peptide quantification and sequencing, protein-level inference and quantification from peptide data is the final step in the workflow. First, peptide sequence identification is carried out through a database search of obtained MS2 spectra against existing computationally digested theoretical spectra and these comparisons are assigned a probability score. For each identified peptide, the highest scoring match is chosen as the amino acid

sequence for the peptide [188]. Generally, peptide identifications are then filtered after applying an FDR correction to the peptide identification scores to control for false-positive hits against the database [189]. The shortlisted peptides are then used to infer the presence of proteins, also known as the protein inference problem [190]. For each peptide, proteins are gathered to explain the presence of that peptide in the sample. Some peptides are only unique to a specific protein, whereas others may be associated with multiple protein groups. Where proteins are indistinguishable from each other based on peptide identifications (i.e., the same combination of peptides is explained by multiple proteins), those proteins are collapsed into a single entry termed as a protein group (See Figure 1:6) [4].

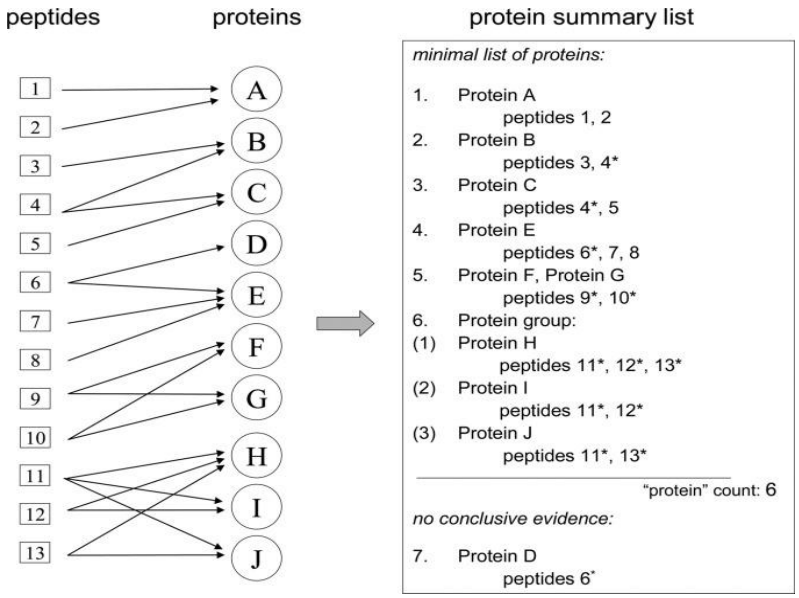


Figure 1:6 Overview of the protein-inference process from peptide-level information. As Protein H, Protein J, and Protein I were identified by the same minimal set of peptides, they are collapsed into one protein group. From [4]

Common algorithmic underpinnings of the peptide-inference problem include greedy searches/parsimonious models (finding the *minimum* set of proteins that explains all peptide identifications), optimistic searches (finding the *largest* set of proteins that explains all peptide identifications), and statistical approaches (modelling peptide matches to a protein as a

probability distribution and generating the likelihood of obtaining *n* number of matches) [190]. Finally, peptide quantification is aggregated in either an additive (either unique or all MS1 peptide intensities are summed together) or a reference manner (an average or median of peptide intensities of a particular protein) to produce protein-level quantification information [191]. Developments in protein-inference algorithms are constantly evolving, and results can differ heavily based on underlying parameters and method of choice [190]. Therefore, it may be beneficial to re-analyse existing mass-spectra files with updated and uniform approaches in the future to verify and improve existing findings. Figure 1:7 summarises the major steps within the proteomics workflow discussed so far.

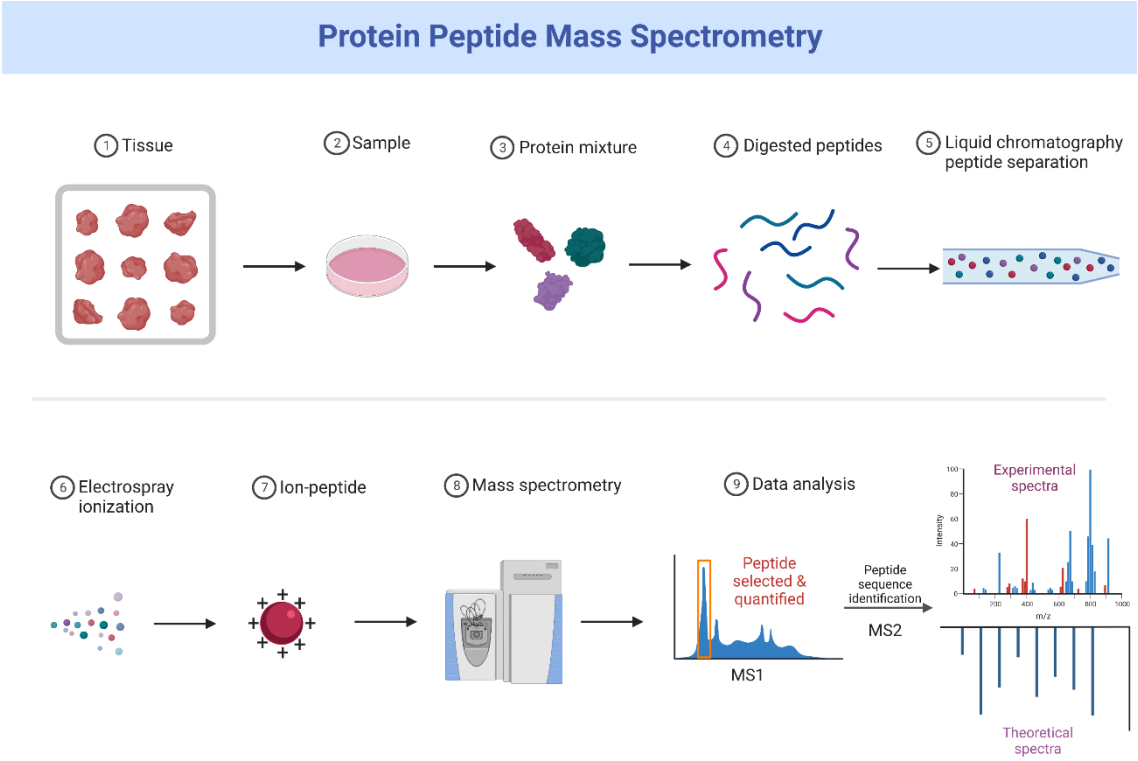


Figure 1:7 Overview of major steps in the mass spectrometry workflow. Adapted from the 'Protein Peptide Mass Spectrometry' template in Biorender.com

There are broadly three types of mass analysers that are most used in biological applications, in either a standalone or tandem manner; these are the ion-trap, quadrupole, and

TOF (Time of Flight) analysers [184]. The triple quadrupole analyser usually consists of three sections - Q1, where the initial MS1 scan across the range of m/z values takes place and, subsequently, all ions of a particular m/z are filtered using time-varying electric fields; Q2 - this is a collision cell where the selected ions from Q1 are dissociated through CID (collision-induced dissociation) into fragment ions; Q3 - where the fragment ions are profiled for an MS2 scan [192] [182]. Ion trap analysers work on a similar principle as the quadrupole. Eluted ions are first 'trapped' in a small volume of space before being subject to sequential, varying, high-frequency voltage fields that allow for specific m/z ions to have unstable oscillations and propel towards the detector [193]. In conjunction with other mass analysers, filtered ions from the ion trap can be directed to another scan following fragmentation. Time-of-Flight (TOF) analysers, as the name suggests, deduce the m/z ratio of an ion through the time taken for the kinetically accelerated particle to reach the detector at the other end of a tube of specified length [194]. At a fixed kinetic energy supply, ions with different m/z are accelerated to a different velocity, which is inversely proportional to their masses; thus, the smaller ions reach the detector first [195]. The TOF analyser is often paired with another TOF analyser (TOF-TOF setup) separated by a quadrupole collision cell, to allow for fragmentation and direction of the product ions, for producing MS-II scans of selected ions [182].

Generally, MALDI is paired with Time-of-flight analysers, whereas the ESI technique has been utilised alongside the triple quadrupole and ion trap analysers [181]. Briefly, in MALDI, the peptide/protein mixture is co-crystalised with an appropriate organic compound (the matrix) with a strong optical absorption capacity. The analyte-matrix crystals are then irradiated with laser pulses, which induce desorption of the crystals into the vapour phase, and single proton transfer from the matrix to the analyte occurs [195]. The ESI method, on the other hand, ionises analytes directly out of the solute phase through the application of a high voltage (see Figure 1:8) (2-6 kV) [186], which causes the creation of small, charged droplets, followed

by exposure to high-temperature gas causing evaporation of the solvent leading to a high build-up of charge at the droplet surface, eventually causing desorption of ions into the gas phase along with the attached solute species [196]. This method is more easily coupled with chromatographic and electrophoretic separation techniques, as it is liquid-based [182]. While the use of MALDI with automated LC may require additional intermediate steps, it also allows for the rate of collection of MS/MS data to be decoupled from the rate of chromatographic separation, thus leading to greater resolution in the data [197].

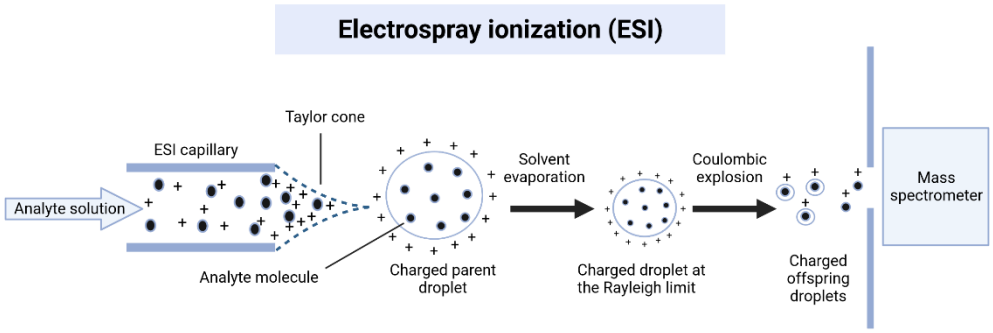


Figure 1:8 Overview of the principle of Electrospray ionisation. Adapted "Electrospray ionisation (ESI)" template from BioRender.com

Generally, TOF analysers demonstrate higher resolution and mass accuracy compared to ion-trap analysers. The latter, on the other hand, allow for greater sensitivity and faster data acquisition but demonstrate limited resolution, accuracy, and ion-trapping capacity [194]. To optimise data acquisition, different analysers are often paired in configurations such as Q-Q-TOF and Q-Q-LIT (MS-I scan and precursor selection in the quadrupole followed by MS-II scan in the second mass-analyser) among others [198]. The most recent development in this area of technology has been the Orbitrap analyser, which outperforms the other commonly used analysers both in terms of accuracy and resolution while also addressing some of the limitations such as the small dynamic range of TOF analysers and the inter-ion interactions within the ion-trap analysers [199]. The Orbitrap works on the principle of current image detection, largely similar to the ion cyclotron resonance (ICT) analyser, where eluted ions are forced into a

harmonic circular spiral through application of a perpendicular voltage. The motion of the ions produces an electric current that is detected by the outer electrodes and this scan is deduced into its constituent frequencies using the Fourier Transform. Under the voltage and kinetic conditions, each m/z has a distinct frequency, and, therefore, a distinct spectrum, and the amplitude of the waveform corresponds to the intensity of the spectra [200]. The Orbitrap is generally paired with either a linear triple quadrupole ion trap, which filters and fragments precursor ions for an MSII scan (e.g., as in the *LTQ-Orbitrap Velos*), or with a c-trap linked to a high-energy collision dissociation cell (HCD) (e.g., as in the *Exactive* benchtop), which allows for preservation of low-mass ions and precursor independent MS-II scan, improving accuracy and resolution of mass measurements [201]. When the input ion source is continuous, as when automated chromatography and ionisation techniques are used, the Orbitrap needs to be paired with a trapping device to control the flow of ions into the instrument in a pulsed manner in the form of *coherently oscillating* ion packets, a pre-requisite for image current detection [202].

In summary, multiple combinations and choices of ionisers and mass spectrometry analysers are used within high-throughput proteomics analysis – each with its own merits and limitations with respect to the ease of set-up, accuracy, and resolution, among other factors. However, the fundamentals of data acquisition across these methods are the same. Enzymatically digested proteins are first separated through the chromatography method, followed by quantification, fragmentation, and identification of peptide sequences. Finally, peptide-level information is computationally reconstructed to infer parent protein identity and quantity. The following section will cover the progression and use of mass spectrometry-based data collection within studies of human exercise physiology.

#### 1.3.2.1 Technical Variability in Mass Spectrometry-based Proteomics

Despite the power and promise of proteomics technologies, variability in methodology, instrumentation, and choice of computational pipelines for downstream analyses, in addition to the factors identified above, ultimately affects the outcomes and reproducibility of research results. This section will briefly summarise the effects of some of these key factors.

Broadly, relative quantification techniques can be classified into two categories: label-based and label-free methods. Label-based methods, the most common of which are iTRAQ (isobaric tags for relative and absolute quantification) and TMTs (tandem mass tags), use chemical reagents made up of different combinations of stable heavy isotopes to tag different samples. All samples are then run simultaneously through the mass spectrometer and quantification is calculated as the ratio of intensities of the reporter ions in the MS2 scan [203]. The main advantage of such methods is that it minimises inter-condition variation introduced during the LC-MS process [204]. Label-free methods, as the name suggests, rely on comparisons of features (i.e., peptide intensities or spectra counts) between different LC-MS runs to infer differences in samples [205]. These methods require fewer steps in sample preparation and do not incur the additional costs of isotopic labels and additional instrumentation time, which can prove to be expensive, especially in larger experiments [206]. However, due to the inherent variability between multiple LC-MS runs, the same peptides may not be quantified and identified in all samples, thus affecting accuracy and increasing the propensity of having missing data [204, 205]. Due to these differences in procedures, the overlap in protein coverage may be limited and relative quantification for all common proteins may not always show strong agreement [206].

Another key distinction to consider is the difference between Data-Dependant (DDA) and Data-Independent Acquisition (DIA) methods. Briefly, in DDA methods, a select number

(usually 10 or 15) of the most abundant peptide peaks (precursor ions), usually called 'top N peaks', from narrow mass-to-charge (m/z) ranges in the MS1 scan are selected for identification [207]. The DIA methods instead aim to identify all quantified peptides (precursor ions) from the MS1 scan [208]. DDA is generally less resource intensive and computationally simpler to run and allows for more sensitive quantification. However, proteome coverage is limited, as only approximately 16% of all identifiable peptides are generally selected for MS/MS scan using standard DDA protocols [208]. The DIA protocols, on the other hand, allow for deeper coverage of the proteome and peptide identification is more reproducible; however, quantification is less sensitive due to reduced acquisition time per data point [209]. Due to the fragmentation of multiple peptides following an MS1 scan, attributing fragment spectra peaks to the correct precursor ion is challenging with DIA [210]. To overcome this, the DIA method requires the building of spectra libraries of all possible peptides in a sample annotated with their MS2 scan, m/z ratio, and elution time [210]. Spectra libraries are built using DDA on multiple fractions of the sample of interest, or samples similar to the type of the sample of interest, and peptide identification is performed through comparison to theoretical spectra using search engines [211]. The results of these multiple runs are combined to build a consensus spectra library, and this is used as the reference for peptide identification following an MS2 scan [211].

Another source of variation that might affect reproducibility or agreement between studies is the differences in computational platforms for peptide identification and quantification, and steps in data pre-processing [212, 213], such as normalisation and missing data imputation, which will be discussed in detail in the following chapter. Differences in techniques of sample preparation and cellular fractionation affect the robustness of protein quantification and may contribute to inter-study variability - especially in the analysis of skeletal muscle tissue, which has a high propensity to get contaminated with biomolecules from

other cell types [205]. Contamination of tissue extractions further exacerbates the problem of detection of low-abundant proteins as skeletal muscle inherently demonstrates a high dynamic range in protein abundance, especially if not accounted for in sample fractionation or choice of platform and data acquisition [214].

Due to these myriad factors, and more, that have the potential to contribute to intra-study variability, the need to re-analyse existing datasets has been recognised [214]. The assumption is that the application of uniform methods, wherever possible in the post-data-acquisition pipeline, should help contextualise, validate, and identify spurious findings if any – which is one of the aims of this project.

### *1.3.3 Application of High-throughput Proteomics in Exercise Physiology*

Consistent with the wider proteomics field, the field of exercise physiology has also been undergoing a shift toward the adoption of high-throughput MS technologies -primarily motivated by the prospect of a much wider proteome coverage. Over approximately the last decade, the number of studies published with MS-based proteomics within exercise physiology per year has doubled compared to before 2011[215]. The following section will review the landmark studies within this area, starting with the very first application of MS-based data acquisition in a human exercise trial.

Holloway et. al., in 2009, claiming to be the first study to use mass spectrometry-based proteomics in conjunction with 2D/1D gel electrophoresis to study skeletal muscle protein responses to exercise training in humans, identified 20 differentially expressed gel spots following 6 weeks of interval training in five untrained young men [216]. They also noted that 2-D separation does not allow for the requisite resolution required to resolve a high number of

mitochondrial proteins in the presence of high-abundance contractile proteins and glycolytic enzymes, necessitating the use of MS for quantification and identification. An iTRAQ-based MS analysis on 1-D separated gel spots, on the other hand, detected differential expression of Complex V subunits α and β, Complex II subunit SDH-A, and Mitochondrial Trifunctional Protein subunit-α (MTPA, alias: HADHA); this result further highlights the functional relevance of MS analysis in the context of investigating training-induced changes in mitochondrial proteins.

To resolve a greater number of mitochondria proteins and achieve greater sensitivity in gel separation, 2D *Difference* Gel Electrophoresis (DIGE) was employed in conjunction with LC/MS on mitochondria-enriched fractions of *vastus lateralis* samples following 2-weeks of MICT endurance training [217]. This allowed for the first mitochondria-specific proteome profile of adaptations to exercise training and identified 18 differentially expressed mitochondria proteins. Beyond the previously observed increases in proteins of the OXPHOS complexes [216], increased abundance of proteins in pathways of the TCA cycle, the Malate Aspartate shuttle, ROS handling (Mn-SOD), creatine resynthesis (mi-CK; Creatine Kinase), and adenine nucleotide homeostasis (AK3; Adenylate Kinase 3) were observed. Further, novel upregulation in the components of the protein synthesis machinery, such as translation elongation factors (EF-Tu) and the isomerase ERp57 (Disulfide isomerase A3), was reported. Western blot results, performed to validate a subset of the identified proteome, which included ATP-β, Mn-SOD (Manganese Superoxide Dismutase) and MDH (Malate Dehydrogenase), demonstrated similar DE patterns as reported by the 2D-DIGE and supported the proteomic results.

According to the literature search, Hussey et.al (2013) appear to be the first to use gel-free quantitative proteomics to investigate exercise-training-induced adaptations in human skeletal muscle (54 ± 4 year-old insulin resistant population) following 4 weeks of MICT [130].

They reported a general trend for higher myocellular adaptations related to oxidative metabolism, with an increase in the abundance of proteins of the TCA cycle, the β-oxidation pathway (SUCLA2, FH, IDH2), ETC (Complex II and II subunits) processes, and a decrease in cytosolic glycolytic proteins. An increase in proteins of the Malate Aspartate shuttle complex (GOT1, GOT2: mitochondrial) was also observed, suggesting enhanced NADH delivery capacity to the ETC. Overall, eight mitochondrial proteins were reported to have significantly increased abundance following exercise training, of which GOT2 (Aspartate aminotransferase) appears to be a novel adaptation identified.

Following this, Robinson et.al (2017) appear to be the first to have utilised gel-free MS proteomics to investigate exercise training adaptations in healthy young adults Robinson, Dasari [128]. These authors performed a comprehensive multi-omics analysis (methylome, transcriptome, and proteome) of HIIT-induced adaptations in both young (mean age 24 years) and old (mean age 70 years) adults separately, and quantified changes in the abundance of over 150 mitochondrial proteins following 12 weeks of training. Most notably, a far greater number of mitochondrial proteins (169), including multiple Complex I-V subunits, TCA cycle proteins, and MA shuttle proteins (GOT2, MDH2), were found to be significantly upregulated in the older cohort compared to the young group (25), perhaps indicating at a lack of requisite exercise intensity to induce stronger adaptations in the young cohort. Further, there was a lack of overlap in the genes upregulated at the mRNA and protein level. Relatedly, ribosomal proteins and proteins within translation pathways were upregulated alongside decreased post-translational markers of protein oxidative damage, suggesting significant exercise-related proteome regulation at the post-transcriptional level. Most of these significant findings, however, were only observed in the older cohort and were not replicated in the young group, warranting future investigation.

Similarly, Hostrup et.al (2018) utilised gel-free proteomics methods to analyse the effects of four weeks of high-intensity training in nine healthy young men (mean age = 24 y), identifying over 450 skeletal muscle proteins (of which 70 were mitochondrial) [218]. The depth achieved in this study appears to be lower than in Robinson et.al (2017), despite the use of similar mass analysers (LTQ-Orbitrap) and data acquisition methods. This might be attributable to the filtering out of proteins (~ 300) that were not annotated with being expressed in the human skeletal muscle. Of the identified proteins, 26 were found to be differentially upregulated following HIT (of which 19 were mitochondrial). Consistent with previous results, oxidative phosphorylation, glycogen metabolism, and Complex I proteins were enriched among the differentially abundant proteins. STOML2 (Stomatin-like protein 2) - implicated in positive cardiolipin regulation and mtDNA replication, CHCHD3 (Coiled-Coil-Helix-Coiled-Coil-Helix Domain Containing 3 protein) - an inner-membrane protein integral for the maintenance of cristae structure [219], and Prohibitin (PHB), appear to be novel exercise adaptions identified in humans.

Granata et. al (2021) have performed the most comprehensive analysis of mitochondria adaptations to training in humans so far, profiling approximately 500 mitochondrial proteins [81]. Utilising a longitudinal experimental design, participants completed a 6-week, high-intensity exercise intervention on a cycle ergometer (with varying weekly blocks of exercise volume) and muscle biopsies were extracted at multiple time points (including at the start of the intervention to establish baseline levels for each participant) to uncover patterns of regulation relative to exercise exposure. Samples were fractionated using differential centrifugation to enrich for mitochondrial proteins and allow for a deep and high-confidence proteome profile. The mitochondrial enrichment also allowed for one of the most detailed observations to date of the adaptations in OXPHOS complex subunits (78 of 92 [220] quantified); previously, most studies have only been able to rely on the differential abundance

of representative proteins to approximate the regulation of the complexes. A novel contribution of this research was to disentangle changes in individual mitochondrial proteins from the overall increase in mitochondrial protein abundance, by applying an *in silico* normalisation method. Post normalisation, a distinct and co-ordinated relative downregulation/deprioritisation of OXPHOS subunits was observed, which is contrary to the well-accepted hypotheses of exercise response backed up previous unscaled mitochondria data (for example, see [221], [218]) [128]. Further, proteins of the TCA cycle, fatty-acid oxidation enzymes, protein import and transport (chaperones), enzymatic ROS scavengers, and other functional groups were found to be upregulated post-normalisation following the first block of training. Measures of mitochondria biogenesis, such as citrate synthase activity and summed mitochondria protein intensity values, were found to be elevated following each block of training and appeared to follow a dose-response. This demonstrated for the first time that certain functional clusters, most importantly the OXPHOS subunits, are non-stoichiometrically regulated relative to changes in overall mitochondria content. This research also demonstrates the comprehensive utilisation of statistical enrichment analysis, clustering methods, and visualisation tools in exercise proteomic studies.

The first application of DIA-based proteomics in exercise-training studies in humans, which allows for deeper proteome coverage (as described in section 1.3.2.1), profiled the response of over 3000 proteins following five weeks of HIIT in young men [222]. Highlighting the capacity of increased depth obtained, more than 30 unique mitochondrial proteins were quantified compared to some of the other data discussed here (see section 3.1.2). Alongside the proteome, the acetylome was also investigated; this analysis revealed increased abundance and novel post-translational modifications, which affect protein activity, stabilisation, and localisation, for TCA cycle enzymes and ETC complex subunits (particularly CV). Further demonstrating the application of multiple-omics analysis, another study conducted a same-

subject transcriptomic and proteomic analysis of skeletal muscle responses to two months of endurance training on seven young, untrained men, and reported the upregulation of 250 skeletal muscle proteins Makhnovskii, Zgoda [223]. Using methods of correlation analysis (between transcriptomic and proteomic data), in conjunction with algorithms to identify the enrichment of translational regulatory motifs, the authors were able to generate hypotheses regarding the mechanisms that regulate protein abundance depending on their function and cellular location. These two studies highlight the evolution of *omics* analysis in the field of exercise physiology and how data stratified across multiple layers of regulation can potentially help uncover the molecular events underlying adaptations to exercise training (e.g., changes in gene expression, protein abundance, and post-translational modifications etc).

Fibre-type specific adaptations to exercise training in human skeletal muscle have also recently been the subject of massspectrometry-based proteomics investigations [19] [20]. Deshmukh et.al (2021) profiled over 4000 proteins in Type I and Type IIa/x single fibres isolated from whole-muscle samples following 12 weeks of continuous cycling training. Most notably, highly specific responses were observed between fibre types; approximately only 32 proteins (of which 19 were mitochondrial) were differentially regulated across both Type I and Type II fibres, compared to approximately 200 and 150 proteins that were significant within each fibre type, respectively. Regarding mitochondrial proteins, transcription factor TFAM, translation-related proteins, and mtDNA encoded ETC subunits, all demonstrated significant increases in expression irrespective of fibre type; this highlights a common mitochondria biogenesis response across the fibre types. Expectedly, a greater number of mitochondrial proteins demonstrated increased expression in Type I compared to Type II fibres in response to endurance exercise - including proteins of the import complexes (TOM and TIM) and calcium-uptake proteins (MICU1 and MICU2). Finally, little overlap was observed between differential expression profiles of whole-muscle samples compared to either fibre type. It was

hypothesised that this discrepancy could have been due to opposing changes in specific fibres that may have diminished the overall effect observed in whole-lysate samples, in addition to the potential contribution of batch-effects and other technical variability confounders. Pooling together the increasing amounts of data that are starting to become available will allow for more robust findings of exercise-dependant regulation in whole-muscle tissues and enable stronger comparisons and verifications against the 'newer' fibre-type specific evidence.

Cross-sectional studies have also previously utilised high-throughput mass spectrometry methods to study the distinctions in skeletal muscle proteome profiles in populations with differing levels of exercise exposure. The skeletal muscle proteome profile of five endurance-trained (>5 h/wk for > 5 years) young men has been compared to an aged-matched sample from the sedentary population [224]. These authors expectedly reported upregulation of subunits of all five OXPHOS complexes, TCA enzymes, and mitochondrial carriers (SLC25 family), suggesting that the adaptations from short-term training interventions are retained with regular exercise. They reported 25 mitochondrial proteins as differentially abundant between the groups. Similarly, to decipher the protective effects of exercise on age-related loss of muscle mass and strength (sarcopenia), the proteome profiles of 60 individuals (20-87 years old) with varying physical activity levels were compared and over 4000 proteins were quantified [225]. Approximately 250 of the 600 mitochondrial proteins identified demonstrated differential abundance with physical activity exposure independent of age. Significant positive correlations with physical activity were observed in multiple subunits of each ETC complex, Complex I assembly factors, regulators of mitochondrial biogenesis (mainly SIRT proteins), mitochondrial solute carrier proteins (SLC family), TCA cycle proteins, and other functional classes.

Since 2011, the number of proteomic studies within the area of exercise physiology per year has almost doubled [215]. As observed, the depth obtained through MS-based techniques

far exceeds the number of mitochondrial proteins that have been investigated in the context of exercising training via the western blot (immunoblot) technique (~ 70 different proteins) [170]; this highlights the scope and potential utility of proteomic techniques to identify novel mitochondrial adaptations to exercise training. Given the multifaceted nature of mitochondrial regulation, and the diversity of mitochondrial proteins altered through exercise, as demonstrated above, proteomics technology holds the potential to create an intricate map of functionally relevant mitochondrial adaptations to exercise training and cellular response to stress more broadly.

## 1.4 Meta-Analysis in Exercise Physiology

Given the increased interest and utility in using high-throughput mass spectrometry methods within exercise physiology research, a large-scale meta-analysis of exercise-training response data, which is the main objective of this research, has now become possible. Aspects of technical variability, as mentioned in the previous section, along with inherent biological variability, differences in sample size and makeup, and characteristics of the exercise interventions, are some of the major contributors to the overall variability, and sometimes inconsistencies, ultimately observed in effect-sizes reported in individual studies. Aggregating existing results together using meta-analytic techniques may help mitigate the contribution of these non-systematic factors and approximate the true effect of exercise training on the identified proteins. This section summarises the utility and application of the meta-analysis technique within the area of exercise physiology.

The meta-analysis study design is a standard tool in biomedical research. It serves to not only combine existing data in an objective manner, but also confers higher statistical power

in the analysis due to the greater volume of data [226]. This type of design may find significant differences in a dependent variable in response to a given intervention, even when no single study has reported differences at the level of significance. The growth in available data and uniformity in data collection techniques has made this study design feasible in exercise physiology, allowing for a more robust understanding of exercise-mediated adaptations and explanations of contrasting findings.

Transcriptomic responses to exercise training in humans have recently been the topic of large-scale, meta-analysis studies. A cumulative sample size of 739 individuals from 43 studies was included in a random-effects, meta-analysis design that demonstrated differential regulation of 114 genes in skeletal muscle following exercise training [227]. Interestingly, genes for contractile proteins and ECM (extra-cellular matrix) modelling proteins, among others, demonstrated stronger enrichment than mitochondrial genes. Following a similar approach, another study had previously identified the gene NR4A3 as an important regulator of the exercise response and validated this finding through *in vitro* experiments [228], demonstrating the capability of pooling data to uncover new hypotheses. Pathways related to ETC complexes and mitochondrial biogenesis were found to be significantly upregulated with exercise training and downregulated in conditions of inactivity. Interestingly, however, there was little intersection between the gene sets significantly regulated in conditions of exercise training and inactivity despite the overlap in functional enrichment. These two studies have been accompanied by the development of interactive online tools and websites sharing detailed analysis of their data (metamex.eu and extrameta.org) to aid fellow scientists to develop experimental questions and targets to analyse.

Petriz et al. (2017) have previously conducted a narrative review of studies analysing the skeletal muscle proteome using both gel-based and mass spectrometry-based methods [229]. However, only the direction of change and not the magnitude was reported. Furthermore,

human and animal studies were not segregated. Srisawat et al., in 2017, attempted the only existing meta-analysis investigating the proteomic response to exercise in humans [230]. However, a lack of amenable data also limited the sample size in the analysis and proteome coverage was extremely limited (only 4 mitochondria proteins were profiled). Further, effect sizes for only the proteins demonstrating significant changes were reported.

As demonstrated, a meta-analysis study holds great potential to advance the knowledge and understanding of mechanisms of physiological regulation in response to exercise stress. While gene expression data has been investigated in detail, no such large-scale study has yet been performed on proteomic data. Since the latest meta-analysis on proteomics data [230], at least five more studies, utilising gel-free MS in this context, have been published, with each profiling hundreds of mitochondrial proteins at once. Greater access and uniformity of raw data, paired with the upward trend in the adoption of this technology [178, 215], now allows for and warrants the development of an informative analysis and database mapping the regulation of the skeletal muscle mitochondrial proteome network.

## 1.5 Relationship between gene and protein regulation

A far greater number of omics-based studies have analysed the transcriptomic response than the proteomic response to exercise, due to the greater feasibility and ease of applying transcriptomic techniques [231], and have been the subject of meta-analyses (as discussed in section 1.4). Nearly 4000 genes have been reported to have a variable response to aerobic exercise [227, 228]. The interest in gene-expression data is based on the hypothesis that the cumulative effects of transient increases in mRNA following single exercise sessions will subsequently lead to an increase in steady-state protein abundance [5] (Figure 1.9). However,

inferring changes or levels of protein abundance from mRNA data is not straightforward, in human tissue or cultured cell lines, and remains a subject of debate due to the lack of strong correlations between mRNA and protein results [232, 233]. In the context of exercise training, it has been demonstrated that increases in muscle proteins were greater than changes in mRNA, and the relationship was particularly weak for mitochondrial and ribosomal proteins Robinson, Dasari [128]. Similarly, another study demonstrated an overrepresentation of mitochondria proteins in a group of genes having increased baseline protein abundance without significant changes in mRNA in response to two months of exercise training Makhnovskii, Zgoda [223]. An increase in the relative abundance of respiratory complex proteins following exercise has been reported without a corresponding transcriptional activation of the encoding genes following a single session of the same exercise Popov, Lysenko [146]. Further, genes upregulated following a single session of exercise do not always demonstrate increased protein abundance after an exercise training intervention [125, 223]. A portion of this observed discordance may be attributed to variability in factors such as post-exercise biopsy times and training intensity, modality, and duration, particularly when comparing results between studies [152, 178, 234].

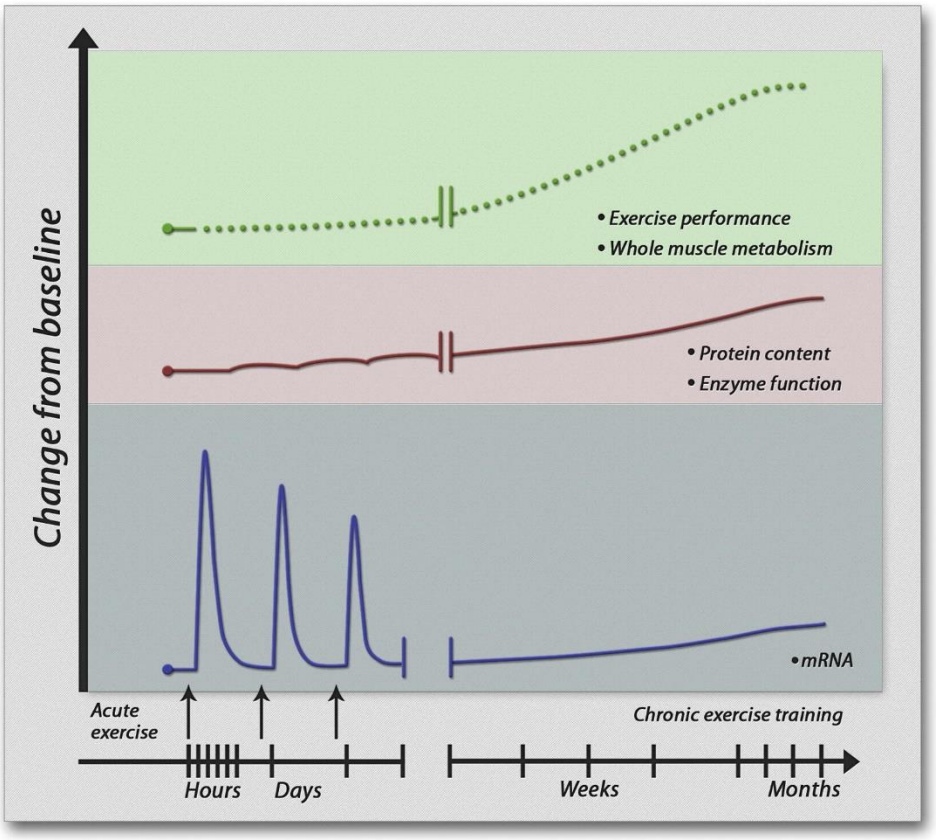


Figure 1:9 Relationship between gene expression and protein abundance regulation with acute exercise and chronic exercise training. From [5]

Qualitatively comparing transcriptomic results with proteomic results may help contextualise, stratify, and validate the findings generated by gene-expression data so far and uncover dynamics of protein regulation. A novel and extensive meta-analysis of exercise training proteomics data can be a valuable tool in this regard, which is another motivation behind the conception of this project.

## 1.6 Conclusions of literature review

The profiling of exercise-induced adaptations in skeletal muscle and skeletal muscle mitochondria has evolved drastically in the last decade. As this section has highlighted, exercise training has been shown to affect protein abundances across multiple key mitochondrial metabolic and regulatory pathways. While western blot and other traditional targeted wet-lab approaches have been able to demonstrate adaptations to certain biomarker proteins across various mitochondrial pathways, the adoption of high-throughput proteomics has allowed for and resulted in a detailed map of protein responses at a far greater depth than was available before. This has also led to the identification of many 'novel' adaptations that were not previously known. The greater and constantly growing availability of protein abundance data within exercise physiology has now created the opportunity for aggregating data from different sources and creating robust results - thus mitigating the effects of technical variability and random noise. Re-analysing existing data and comparing findings between previous exercise-training studies has also been identified as a gap in the academic field [214]. Comparing, contrasting, and summarising high volume of exercise-training proteomics data has the potential to uncover novel protein adaptations in response to metabolic stress and help generate new hypotheses for future targeted experimental investigations.

**Aims:**

- Collect and re-analyse gel-free LC-MS proteomics data from existing literature and identify adaptations to the skeletal muscle mitochondria proteome in response to exercise training in humans.
- Compare the differential expression analysis results between studies using enrichment analysis tools and discuss qualitative differences.
- Perform a meta-analysis by aggregating the results of individual differential expression analyses and present the findings in the context of existing literature.

**Contribution to knowledge:**

A comprehensive and integrated analysis of available data analysing muscle mitochondrial proteomic adaptations to exercise training has not yet been conducted, as the literature review has highlighted. This research and its by-products will therefore aim to provide an important resource that researchers can use to generate new hypotheses and devise targets for experimental validation.

**Statement of Significance:**

While there is significant evidence for the positive effects of exercise training on mitochondrial content and respiratory function, there are still long-standing questions about the effects of physiological stress on the underlying mechanisms and pathways governing mitochondria form, function, and quality control. With the greater and deeper mitochondria proteome coverage available now, it is possible to obtain a more comprehensive understanding of the underlying processes governing mitochondria adaptation to exercise and cellular stress

more broadly. Successfully identifying central adaptive responses implicated in changes to mitochondria structure and function could be of great significance to the understanding, treatment, and management of muscular disorders and age-related degeneration linked to mitochondrial characteristics [235].].

# 2 CHAPTER TWO

## Methodology and Methods

## 2.1 Study design

### *2.1.1 Collection of Protein Datasets*

Existing proteomics datasets investigating the response to endurance-type exercise training in human skeletal muscle were first shortlisted. Databases, such as Scopus, Web of Science, and Google Scholar, were mined using a combination of keywords that included '*human, muscle, exercise, training, and proteomics'*. The PRIDE, ProteomeXchange, and BioArchive (*BioRxv*) databases were also searched for additional proteomics studies.

Only datasets utilising high-throughput MS-based approaches for protein quantification were included in the meta-analysis design to maintain uniformity of downstream data processing. All the LC-MS datasets identified approximately 10-fold more proteins than the most resolved on-gel proteomic study, thus limiting any improvement in statistical power by including these results. Furthermore, western blot-based studies demonstrate an approximately 100-fold lower depth than most LC-MS datasets and hence were excluded from the meta-analysis design. Also, due to the fundamental differences in the principles of quantification [237] and the reported effect sizes (for example: [238]), aggregating fold changes across LC-MS and western blot results was not deemed appropriate. Excluded data were nonetheless tracked to be summarised separately and qualitatively compared with the results of the meta-analysis.

Where available, raw protein groups (non-imputed and non-normalised/logged) files for each shortlisted MS dataset were acquired. If raw protein-groups files were not available, attempts were made to contact the authors and solicit the data. If the acquisition of raw files was unsuccessful, supplementary files for each publication were scoured for processed protein quantification data or differential expression results. Beyond the datasets that are part of the meta-analysis, additional candidate datasets were also identified for inclusion in the study-

Hussey et.al (2013) [130], Robinson et.al (2017) [128], Hostrup et.al (2018) [218], and Jacques et.al (2022) [239]. No proteomics files were able to be obtained for Robinson et.al (2017) and Hussey et.al (2013) and no supplementary files with comprehensive differential expression results were found either. Similarly, for Jacques et.al (2022), supplementary differential expression data was only available for significant results and protein identification was described with only Gene Names and not Uniprot IDs. Raw files for the [218] samples were found on PRIDE under the accession PXD005480; however, no accompanying labelling information was found signifying the grouping of the samples. In all these cases, attempts to contact the authors for further information were unsuccessful. The overview of the study design and methodology is presented in Figure 2:1.

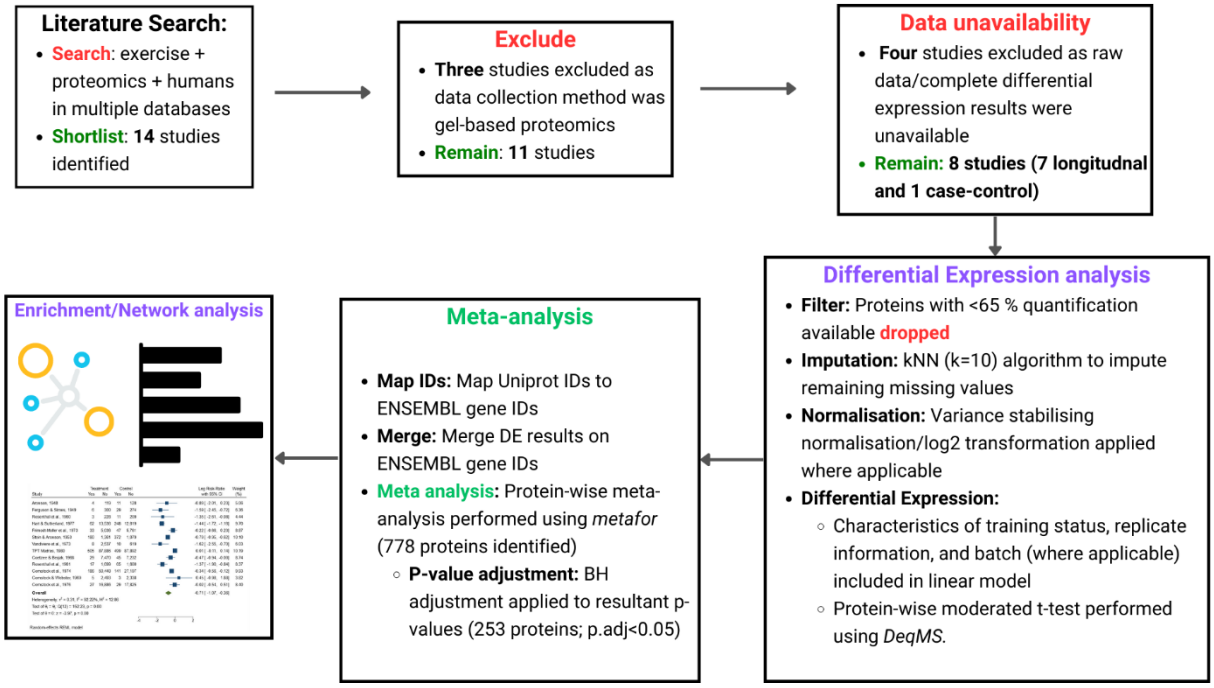


Figure 2:1 Overview of the study design and methodology

Study and participant characteristics data were collected along the categories of duration and modality of exercise intervention, number of participants (and covariates such as age, health, and sex, where reported), muscle sampled, method of protein quantification and identification, and biopsy times. Variables related to training intensity and duration were used to calculate exercise exposure/volume in arbitrary units (as described in section 2.1.1.1). Critical information for all datasets included in the meta-analysis are summarised in Figure 2:2. For a detailed summary of these studies, see the Appendix file 'Data characteristics and results.xls' (*sheet:* Study_characteristics).

| Protocol | Repository ID | Author | DOI | Reference label |
|---|---|---|---|---|
| High Intensity Interval Training (Normal Volume) | PXD026219 | Granata, 2021 | https://doi.org/10.1038/s41467-021-27153-3 | HVT_Granata |
| High Intensity Interval Training (High Volume) | PXD026219 | Granata,2021 | https://doi.org/10.1038/s41467-021-27153-3 | NVT_Granata |
| High Intensity Interval Training | PXD023084 | Hostrup, 2021 | https://doi.org/10.7554/eLife.69802 | Hostrup_2021 |
| High Intensity Interval Training | NA | Makhnovskii, 2020 | https://doi.org/10.1038/s41598-020-60578-2 | Popov_2020 |
| Continuous Endurance | PXD012824 | Deshmukh, 2021 | https://doi.org/10.1038/s41467-020-20556-8 | Deshmukh_2021 |
| Continuous Endurance | NA | Botella, 2022 | https://doi.org/10.1101/2022.10.23.512956 | Botella_MICT |
| Continuous Endurance (Case-control comparison) | NA | Schild, 2015 | https://doi.org/10.1016/j.jprot.2015.03.028 | Schild_2015 |
| Sprint Interval | NA | Botella, 2022 | https://doi.org/10.1101/2022.10.23.512956 | Botella_SIT |

Figure 2:2 Overview of studies included in the meta-analysis. The Reference label column describes the tag used to refer to these datasets within this study.

#### 2.1.1.1 Calculating exercise volume/exposure

The exercise volume/exposure associated with each dataset included in this study was calculated using measures of duration (minutes of exercise) and average intensity (%$VO_{2max}$) for each exercise session, and the total number of exercise sessions completed within the intervention. Session exercise intensity in the associated publications of the included datasets was reported as a percentage of either the *lactate threshold (LT)*/*maximal lactate steady state* (MLSS), the maximal power output achieved in an incremental exercise test ($W_{max}$), or the maximal heart rate ($HR_{max}$) maintained during each exercise session. All measures of intensity were converted to %$VO_{2max}$ to calculate exercise intensity uniformly across the different datasets. Measures of LT/MLSS were converted to %$W_{max}$ using approximations detailed in Jamnick et. al (2018) [240], which was then converted to an equivalent value of %$VO_{2max}$ (for example, 85% $W_{max}$ ≈ 85% $VO_{2max}$). The %$HR_{max}$ values were converted to approximations of %$VO_{2max}$ using the *Swain* equation, as reported in Swain et. al (1994) [241]. Total exercise duration (in hours) was calculated by multiplying the average exercise session duration with the total number of exercise sessions across the intervention. Finally, exercise intensity and duration measures were multiplied to obtain a total exercise volume expressed as %$VO_{2max}$.hour (%$VO_{2max}$.h). No exercise volume measure was calculated for Scihld_2015 as it is based on a case-control comparison (untrained v. endurance-trained for > 5 years) and not a longitudinal study-design. The calculated exercise volume measures for each included dataset in this study is reported in the Appendix file 'Data characteristics and results.xls' (*sheet:* Study_characteristics).

### *2.1.2 Processing of Protein Data*

Of the datasets shortlisted for downstream quantitative analysis, only one utilised a label-based (iTRAQ) technique (Popov_2020[223]). The remaining studies utilised LFQ techniques in either data-dependant or data-independent mode. For more detailed information on the data acquisition methods for each study, see the Appendix file 'Data characteristics and results.xls' (*sheet:* Study_characteristics). A sample of the pipeline used for pre-processing and differential expression analysis of proteomics data in this study can be found in the Appendix folder 'Code'.

#### 2.1.2.1 Data Pre-processing and cleaning

Raw Protein Groups files were obtained in either the .xlsx or. tsv format. Files were read into the R workspace using either base R (for .tsv files) or the *readxl* package *(for .xlsx files; [242]).* Where reported, entries with *"Potential Contaminant", "only.identified.by.site",* and *"is.reverse"* characterisations were filtered out from the data. Where applicable, zeros in the intensity columns were converted to NA values. Only the data columns with the Protein Accessions/Protein Group Majority Protein IDs, the Razor +Unique Peptide Count for each sample, and the normalised protein groups intensities (where applicable) were retained
Each row was assigned a unique character value as a *rowname* to keep track of the entries during data manipulation and to avoid duplication of rows. Finally, all intensity values were extracted into an only numeric matrix, which was the basis of all downstream analysis.

#### 2.1.2.2 Imputation

Missingness of data can be broadly categorised into two categories: Missing (Completely) at Random (MAR/MCAR) and Missing Not at Random (MNAR). Briefly, in the context of proteomics, MAR/MCAR means that a missing value is not related to the nature of the peptide/protein and further that the missingness affects the entire dataset uniformly. MNAR, on the other hand means that a feature of the peptide/protein contributes to its tendency to be missing [243]. Within LC-MS proteomics, particularly DDA acquisition, MNAR typically refers to abundance-dependant missingness where peptides/proteins whose abundances are close to the instrument limit of detection demonstrate a higher degree of missing values [244] [243]. To test the nature of missingness in the DDA datasets used in the quantitative analysis, data from *Granata_2021* (mitochondria-enriched muscle fraction) and *Deshmukh_2021* (whole-muscle lysate) was used to compare average protein intensity against the degree of missingness (see Figure 2:4 and Figure 2:3 respectively). Firstly, it is evident that proteins with high abundance values relative to the data generally do not show any degree of missingness. Further, beyond zero/no missing values, the tendency of missingness does not appear to be dependent on the average intensity and is abundance independent. Generally, missingness in LC-MS data is a mixture of MAR and MNAR. For example, using tests of statistical significance, it was found that only 43% of missing values could be linked to average peptide abundance in samples of human plasma obtained through LC-MS acquisition [245]. Indeed, this finding was also confirmed for the data included in this study. For the *Deshmukh_2021* and *Granata_2021* datasets, the *model.selector* functionality from the package *imputeLCMD* [246] was applied; this function identifies rows (proteins) as either MAR/MNAR based on the empirical relationship between average protein intensities and missingness. While MNAR tendencies were observed for high degrees of missingness (missing

in > 80-85% of the data), the rest of the data was almost exclusively categorised as MAR by the model. Further, after the filtering threshold of missingness was applied (as mentioned in 2.1.2.1), all the remaining data was exclusively categorised as MAR. In summary, the MS data used within this study appears to demonstrate a combination of both types of missingness, with a tendency toward MAR, and thus warrants a tested imputation approach that is valid for such mixed/heavily MAR data.

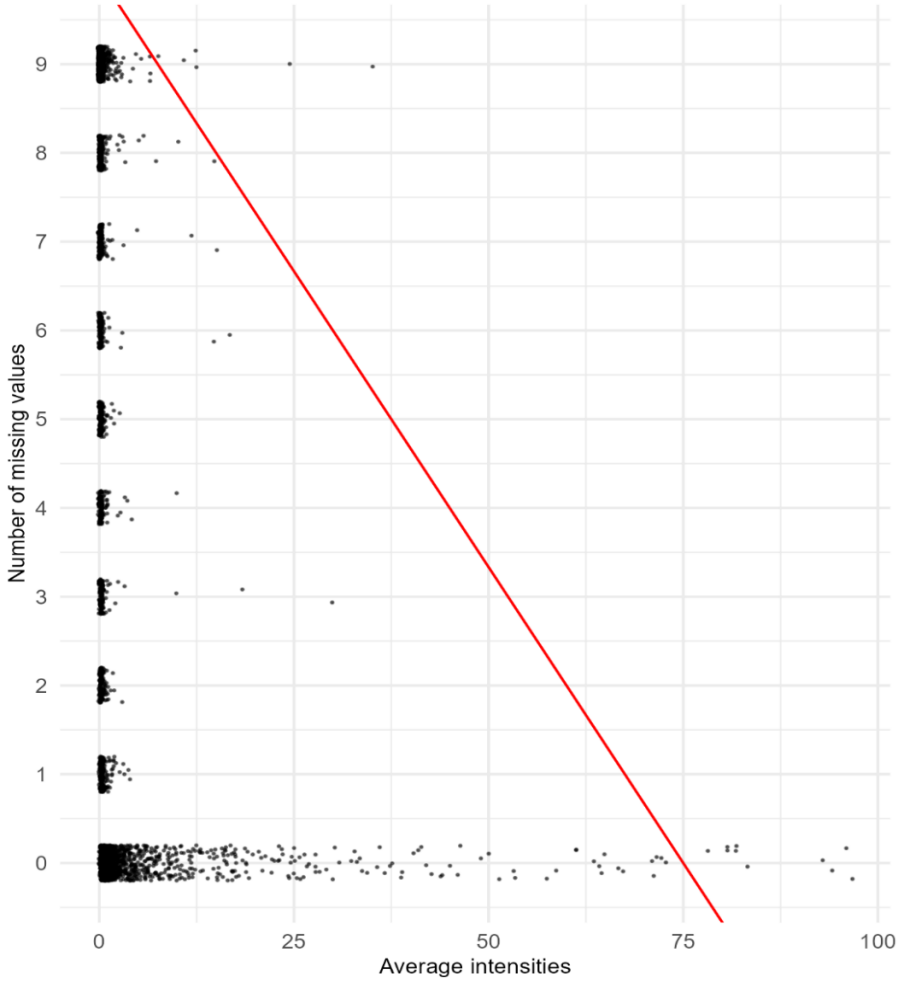


Figure 2:3 Comparison of missingness in the raw data against scaled average intensities. As can be observed, beyond zero missing values, the degree of missingness does not correlate with average intensities (scaled). The red line is an approximate trend line of the association that should have been observed in the case of intensity-dependent missingness. The y-axis has been limited to better represent the trend. Data from Deshmukh 2021

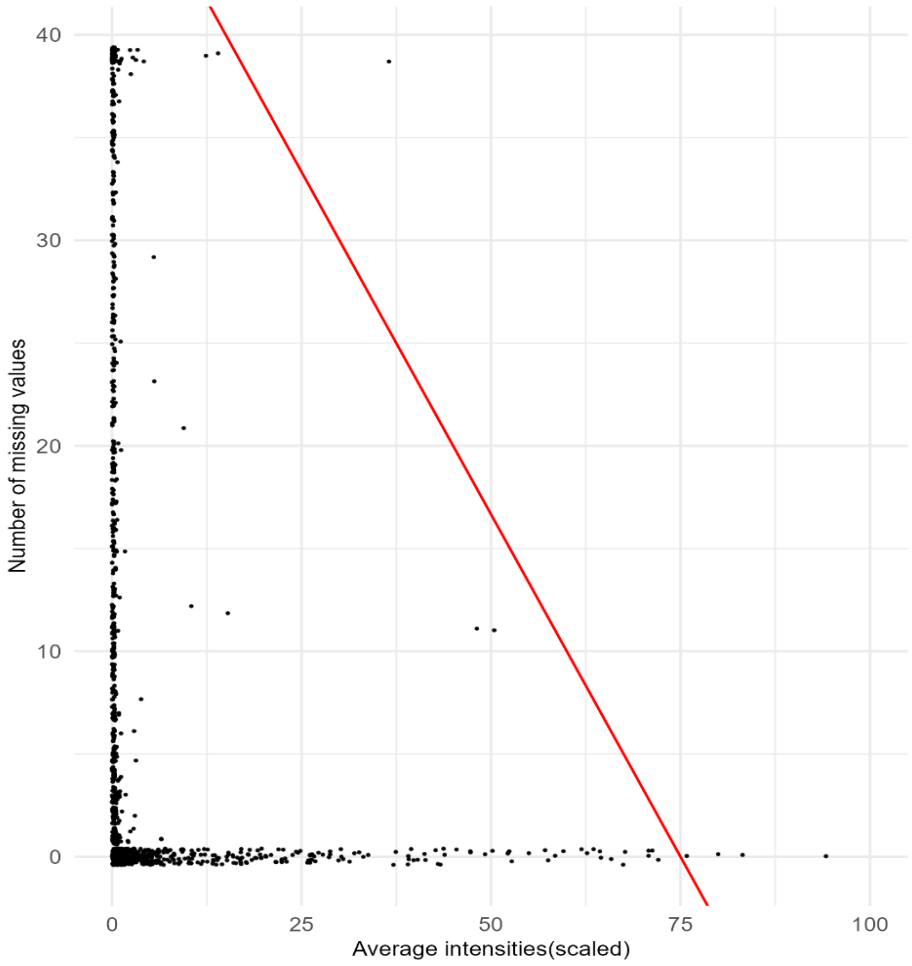


Figure 2:4 Comparison of missingness in the raw data against scaled average intensities as in Figure 2.3. Data from Granata_2021

There are multiple MAR/MCAR-appropriate imputation methods previously applied to proteomics data that use local-similarity-based approaches (estimating missing values based on other proteins with similar intensity profiles), such as Least Squares Regression (LLS) and k-Nearest Neighbours (kNN), or global approaches (iterative, corrective imputation on 'lossy' approximations of data) such as Singular Value Decomposition (SVD) and Probabilistic Principal Component Analysis (PPCA) [243, 245, 247]. For data evaluated to be heavily MNAR, or to contain intensity-dependant missingness, left-censored techniques, such as random draws from left-censored normal distribution (ND) and quantile-regression, and single-

value methods, such as Deterministic Minimum Imputation (MinDet) or Lowest of Detection (LOD), are applicable [243] [247]. However, these methods appear to only be applicable in conditions where there is a high degree of left censoring; i.e., low (left tail) values across the protein abundance distribution generally demonstrate missingness [245]. For example, the accuracy of ND imputation is negatively affected by higher rates of MAR in the data [247]. Furthermore, local-similarity-based approaches (kNN, LSS, RF) were shown to outperform the accuracy of left-censored-devoted methods (ND, LOD) even with MNAR rates as high as 80%. It has also been advocated that in the absence of determining the nature of missing values, MAR/MCAR-devoted methods outperform MNAR methods [243]. In summary, given the estimated hybrid nature of missingness in the data for this study and the robustness of these techniques on real [245] and simulated data [247], local approaches were preferred over left-censored appropriate methods. Global approaches, on the other hand, have been found to slightly outperform similarity-based approaches [245]; however, this is not always observed [247] and they are also computationally more complex to run. The accuracy of the local-similarity-based imputation is affected more by the proportion of MNAR values [247] and the proportion of total missing values [243] than the method of imputation applied. Therefore, proteins with a high degree of missingness were filtered out (as discussed in 2.1.1) before applying imputation.

Within local-similarity-based approaches, Random Forest (RF) imputation appears to outperform the other methods tested, such as LLS, LSA (Least Squares Adaptive), and kNN for LC-MS proteomics [247] and metabolomics data [248]. RF imputation was attempted with the package *missForest* [249]; however, this method failed to converge consistently for all datasets. It is suspected that because this imputation is iterative (until convergence), personal computer resources were not sufficient to reliably complete the process, especially given the size of some of the data. The LLS imputation method has also been shown to perform well

and is an appropriate choice; however, it has been shown to perform unexpectedly poorly, and much worse than kNN, in accurately predicting protein ratios [247]. The LSA, EM (Expectation Maximisation), and kNN imputations have been shown to perform comparably and the latter is even outperformed by the other two in certain cases [245]. However, this study [245] examined the effects of peptide-level and not protein-level imputation, unlike in Jin et. al (2021) [247]; therefore, limited applicable inferences can be drawn from their findings. Moreover, no further review or applications of LSA or EM for the imputation of LC-MS data was found. The kNN-imputation, on the other hand, has been widely applied and well-reviewed in the context of MS data [245] [247] [248] [250]. With MS metabolomics data, kNN imputation has even been shown to outperform global approaches, such as PPCA and Multivariate Imputation [250]. Given the feasibility of application, and evidence supporting the strength of performance, kNN was chosen as the preferred method for imputing missing data values in this study. RF imputation may have been the better choice, as supported by the literature; however, due to the limitations discussed, it could not be used in this study.

After filtering out rows with missing values greater than the set threshold (35%), the rest of the missing values were filled with imputation. A *kNN* algorithm with $k$=10 was applied to the raw intensity data matrix using the *impute* package from Bioconductor [251]. No scaling or transformation to intensity values was applied prior to the imputation. It has been suggested that imputation at the peptide level is a more accurate and robust approach compared to imputation at the protein level [245]. However, a limitation of this study is that for most datasets included in this investigation the raw MS spectra files or peptide-level quantification data were not available and imputation could not be performed at the peptide level.

#### 2.1.2.3 Normalisation

Measures of intra-group variance and sensitivity of differential expression analysis within proteomics have been shown to improve with the application of normalisation to protein intensity data compared to simply the $Log_2$ transformation of the intensities [252]. Primarily, this has been attributed to the attenuation of the inherent bias in the data either due to random noise or the accumulation of systematic errors through the experimentation and the sample preparation process [253]. Of the common methods utilised, such as median normalisation, regression-based techniques, PCA-based techniques, and others, the Variance Stabilising Normalisation (VSN) has been shown to be a highly suitable method [252, 254] [253]. The use of VSN in proteomics has been adapted from its applications on microarray gene expression data [254]. This method was chosen as it provides an improvement on $log_2$ transformations, and it can transform the data such that the variance is approximately independent of the mean [255]. As can be observed in Figure 2:5, which depicts the mean-variance relationship of raw intensity data obtained from my research group and provides a representative example of the data in this study, the VSN markedly reduces the standard deviation associated with each intensity value on the x-axis and further disentangles the mean from the variance (an almost straight line compared to a clear downward trend for the log transformation). Further, as observed in the Principal Component Analysis (PCA) figures below, applying VSN normalisation (see Figure 2:7) over log transformation (see Figure 2:6) may also help better

stratify the data compared to log transformations; this strengthens the downstream differential expression analysis.

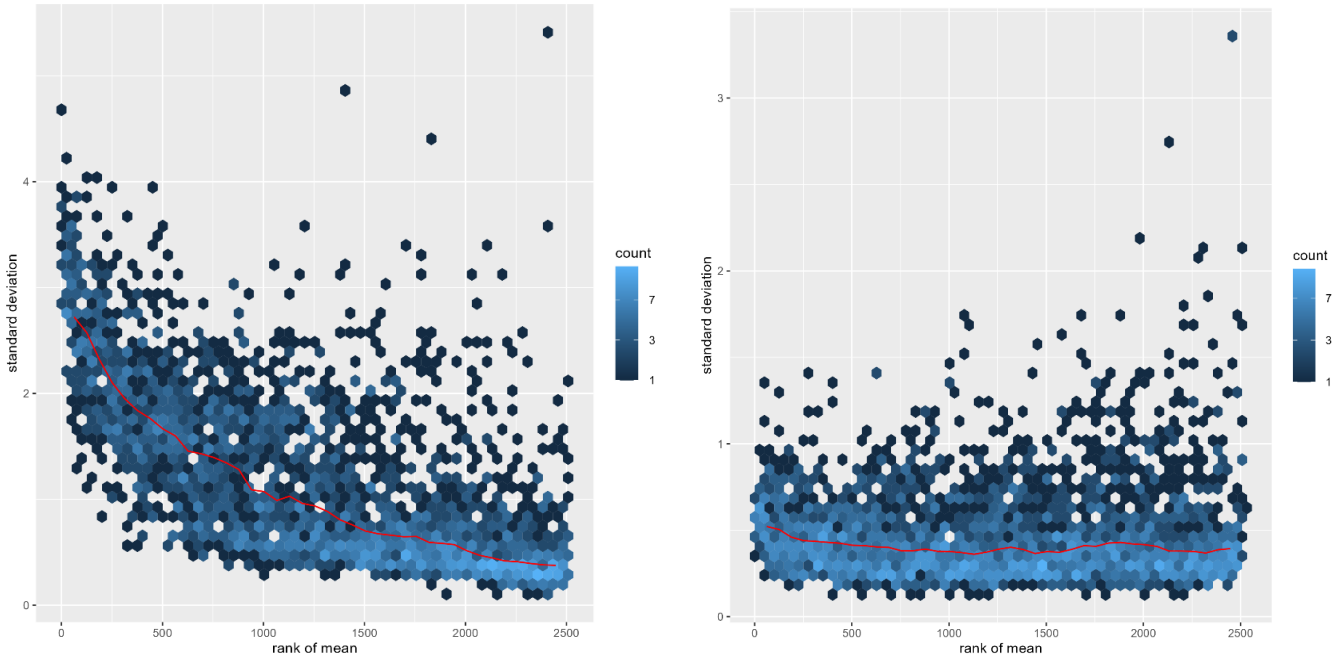


Figure 2:5 Mean vs. standard deviation plot of log-transformation (left) and VSN-normalisation (right) applied on unpublished DIA data from *Botella_2022*

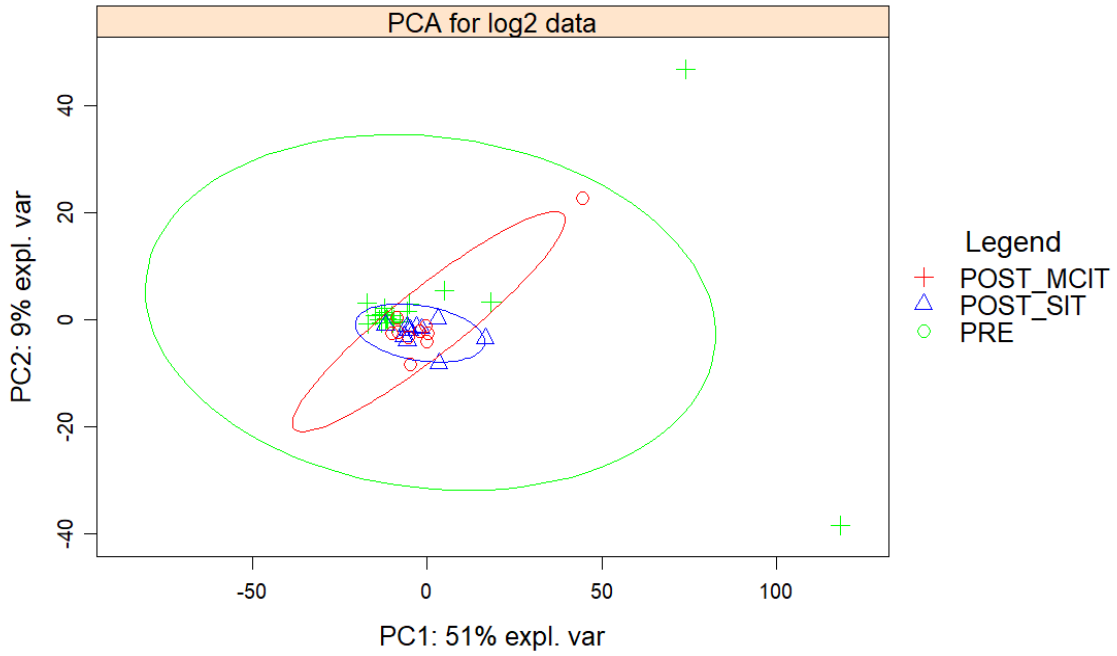


Figure 2:6 PCA plot of log-transformed mitochondrial protein quantification data. Extracted from unpublished DIA data from the *Botella_2022* dataset.

Per the rationale discussed above, following pre-processing and imputation, variance stabilising normalisation was applied to the unpaired raw intensity data using the *normalizeVSN* function from the package *limma* from *Bioconductor* [256]. Certain studies/datasets (Deshmukh_2021, Hostrup_2021, and Schild_2015) did not appear to improve their mean-variance relationship when comparing $\log_2$ vs. VSN transformation. Further, for these data, no marked differences were observed in the PCA/MDS plot's ability to separate the groups either (e.g.: when comparing Pre v. Post data; see Figure 2:8 and 2:9). Therefore, only the log2 transformation was applied for these data.

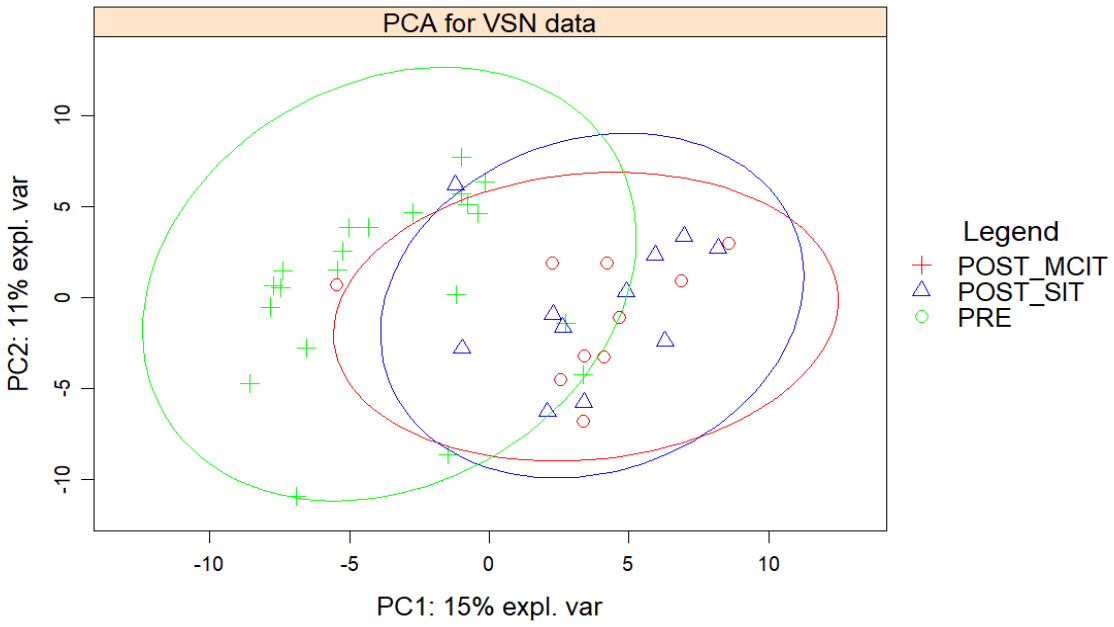


Figure 2:7 PCA plot of VSN-transformed mitochondrial protein quantification data. Extracted from unpublished DIA data from the *Botella_2022* dataset.

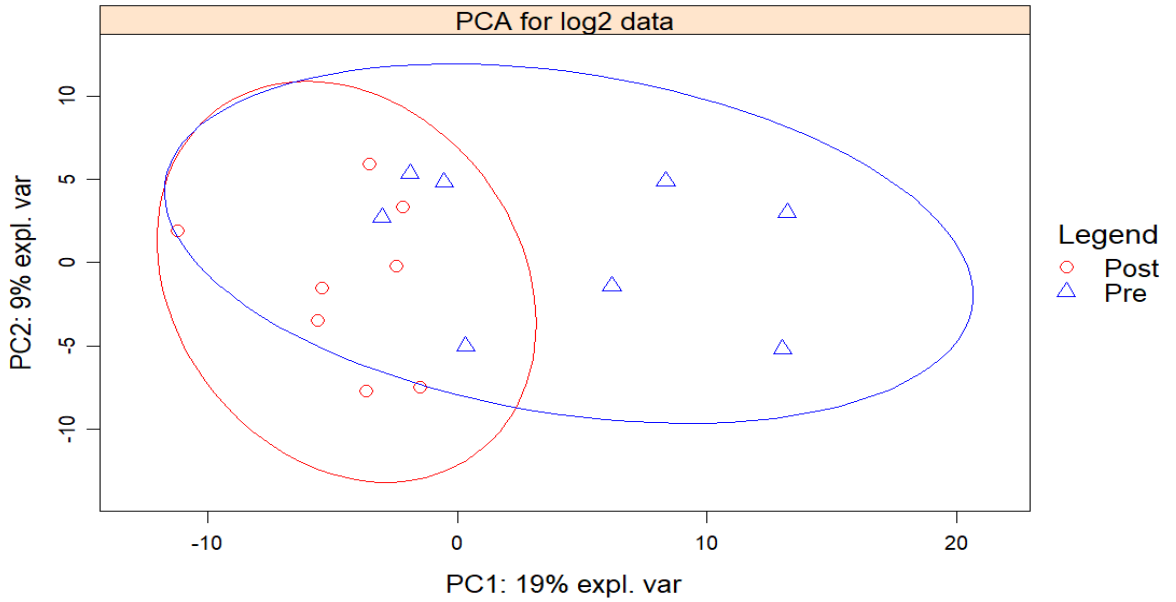


Figure 2:9 PCA plot of log transformed mitochondrial proteome quantification data. Extracted from the *Hostrup_2021 dataset.*

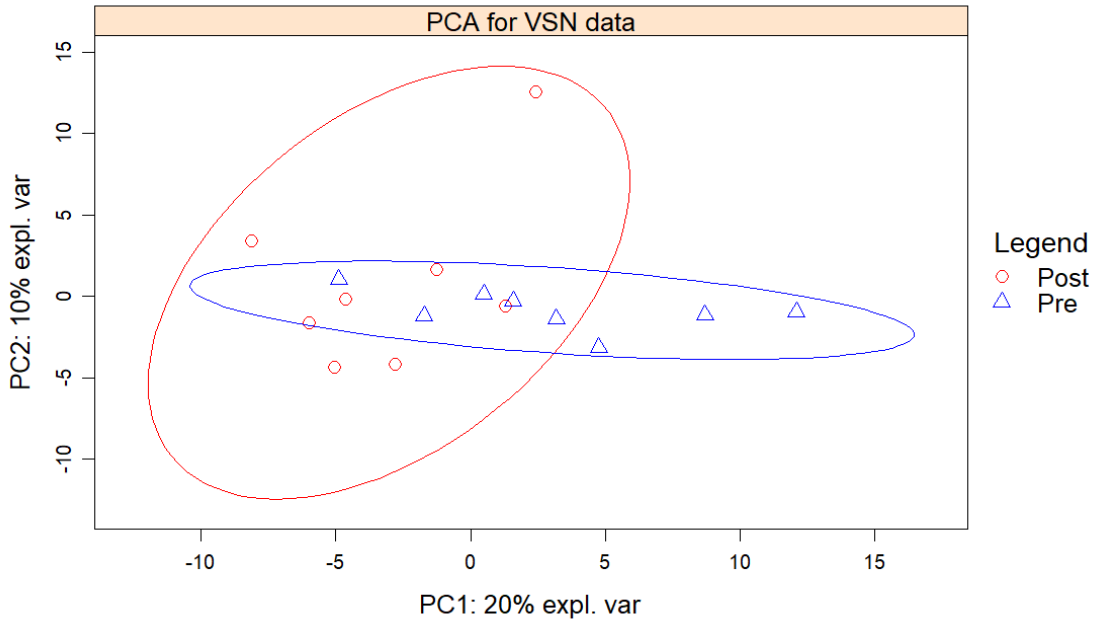


Figure 2:8 PCA plot of VSN-transformed mitochondrial proteome quantification data. Extracted from the *Hostrup_2021dataset.*

### 2.1.2.4 Mitochondria-specific Normalisation

To account for changes in overall mitochondria content and to discern non-stoichiometric adaptations to the mitochondria proteome, scaling through the *normalizeVSN* function was applied on the mitochondria protein intensities after sub-setting them out from the whole dataset as described in [81]. The VSN function works through estimating column-specific (sample-specific) parameters and applying unique scaling and offset to each of the intensity columns followed by a generalised log transformation ($glog_2$) [255]. Only retaining the mitochondria proteins allows the estimation of the scaling factors that account for mitochondria proteome abundance in any sample/column and to calibrate accordingly. The efficacy of this method can be confirmed visually by examining the volcano plots of

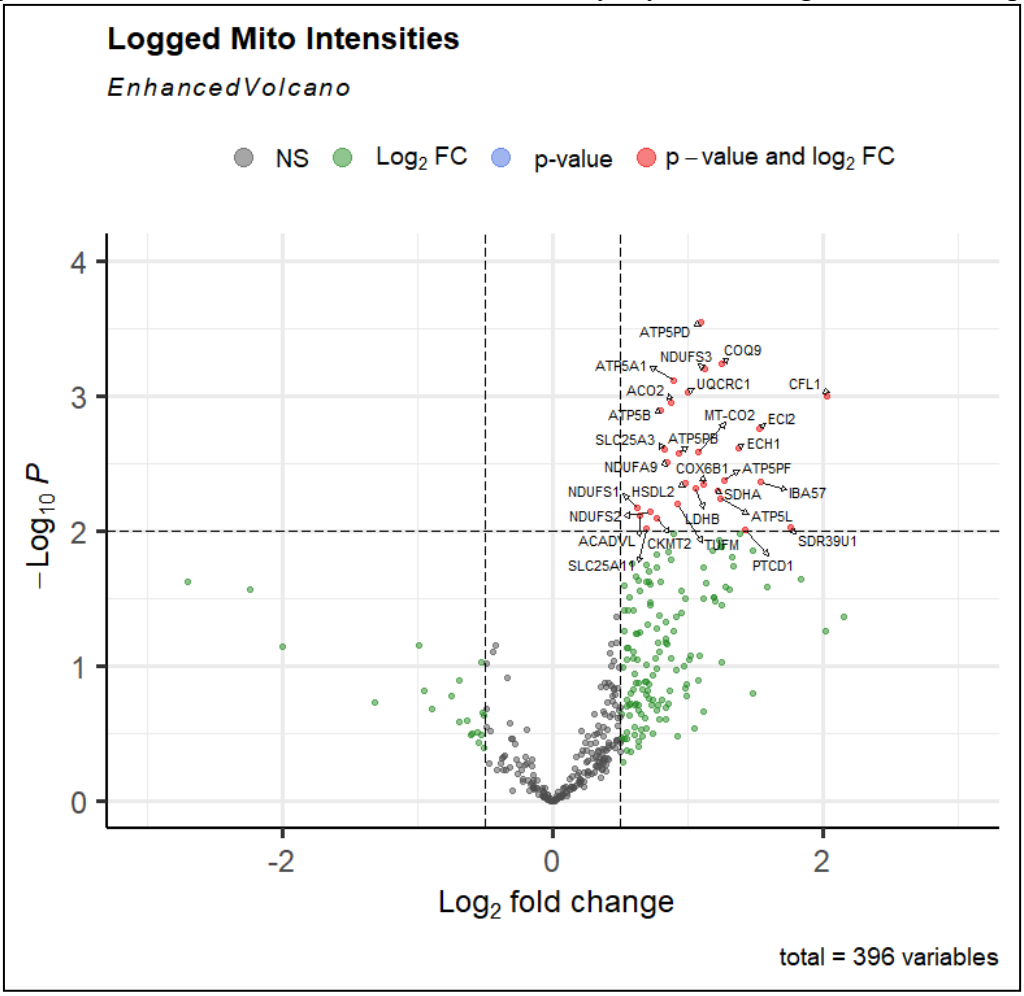


Figure 2:10 Volcano plot of differential expression of logged mitochondrial intensities. Data from *Schild_2015*.

mitochondria-normalised against the non-normalised data from Schild_2015. It is observed that fold changes are uniformly centred around the 0 on the X-axis in the mitochondria-normalised plot (Figure 2:11), whereas the fold changes are positively skewed in the non-normalised plot signifying the overall change in mitochondria protein abundance/content (Figure 2:10).

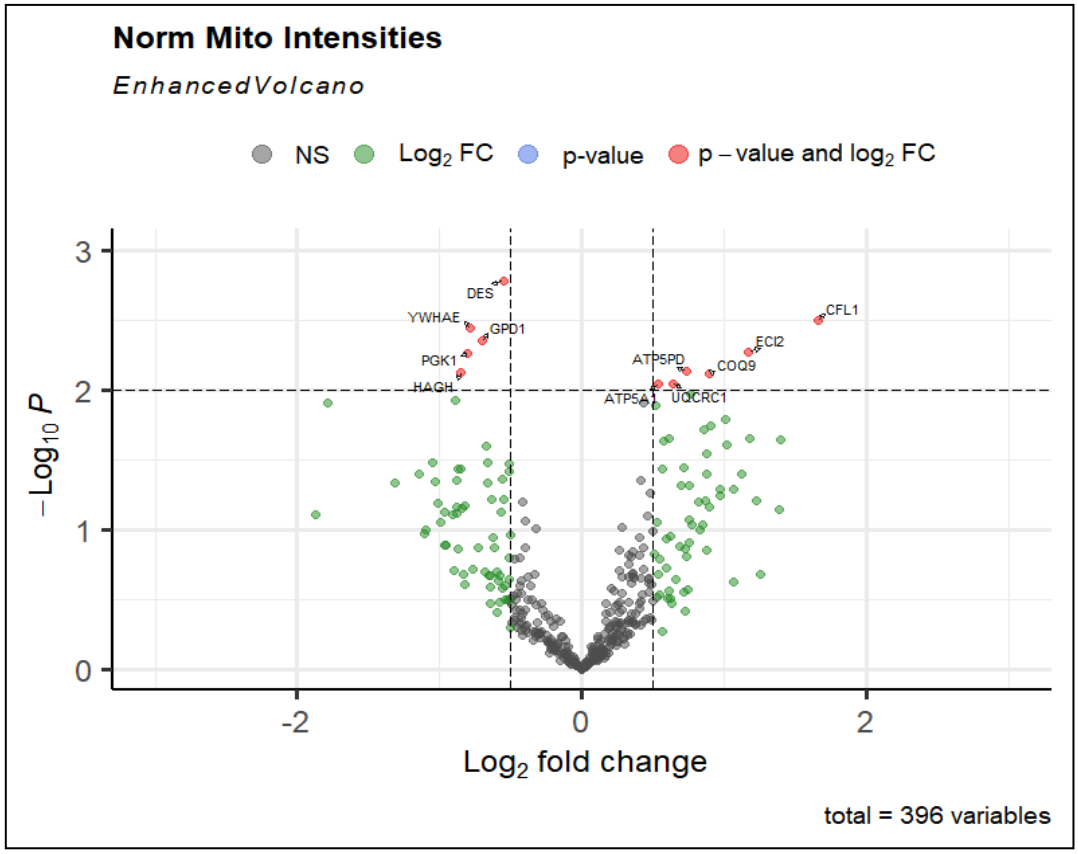


Figure 2:11 Volcano plot of differential expression of mitochondria-normalised intensities. Data from Schild et.al (2015).

### 2.1.2.5 Determination of mitochondrial proteins

The Integrated Mitochondrial Protein Index (IMPI) was used as the basis for the identification and shortlisting of mitochondrial proteins [32]. Genes found as being localised to the mitochondria in either MitoCarta 3.0 or IMPI (labelled as *verified)*, where localisation

has been assessed through GFP (Green Fluorescent Protein) tagging and microscopy in isolated mitochondria, were first shortlisted as genes of interest (1492 genes) [32]. Further, genes with evidence of either localisation through APEX (Proximity dependent labelling) tagging (labelled as *associated)* or that have been reported to affect mitochondria function or morphology (labelled as *ancillary)* were also shortlisted. From these associated and ancillary genes, only those with an SVM (Support Vector Machine) score of >= 0.9 on a trained classifier algorithm were ultimately selected (14 genes). A total of 1506 unique mitochondrial genes [257] were identified.

This list of identified Gene Name/ Symbol was fed into the Uniprot ID mapping service (https://www.uniprot.org/id-mapping; From database setting=Gene Name; To database setting= UniProtKB; Taxonomy setting: 9606 (Human)). A comprehensive list of approximately 6500 unique UniProtKB IDs was generated. Often a protein/protein group has multiple accession IDs (one primary and multiple secondary) associated with it, due to entries being merged and demerged (see: https://www.uniprot.org/help/accession_numbers). Further, each gene may have multiple protein isoforms associated with it, each with its own unique accession numbers. Hence, the number of Uniprot IDs shortlisted was approximately four times the number of unique gene names. All Uniprot IDs in the cleaned protein-groups file were searched against this list to create a matrix of mitochondria-only intensities. Where multiple Uniprot IDs existed for any particular row in the protein-groups file, the rows were expanded

and only the first distinct Uniprot ID with a matching entry in the curated list was selected. A total of 906 unique Uniprot IDs were found across the eight datasets.

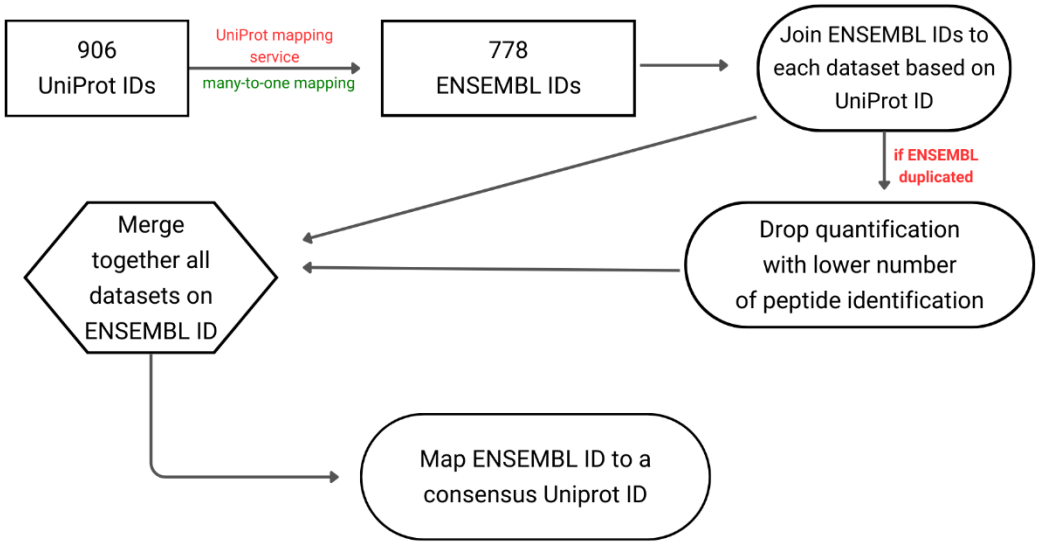


Figure 2:12 Overview of UniProt ID to ENSEMBL ID mapping

These Uniprot IDs were mapped to ENSEMBL Gene IDs to merge downstream differential expression analysis results and perform the meta-analysis. ENSEMBL Gene IDs offered a higher degree of intersection between datasets compared with Uniprot IDs, mainly due to the issue of multiplicity of Uniprot IDs mentioned above. Uniprot IDs in each dataset mapped one-to-one with ENSEMBL IDs, except for minor duplications in Schild_2015 and Deshmukh_2021 (one ENSEMBL ID appeared twice), and in Hostrup_2021 (two ENSEMBL IDs appeared twice). With these duplications, the row with the lower peptide count (number of peptides used for quantification) was dropped. A total of 778 unique ENSEMBL IDs were observed across the eight datasets (see Figure 2:12).

## 2.2 Differential Expression Analysis

### *2.2.1 Use of DEqMS*

Following imputation and normalisation (where applicable), differential expression analysis of both mitochondria scaled/normalised (as described above) and unscaled data was performed primarily using the *DEqMS* package [258]. This package was built as a supplement to the *empiricalBayes* method in *limma,* which has previously been used to analyse LC-MS data (for example, see: [213] [259] [81] [20]), to tailor it better for applications in proteomics. Briefly, the *empiricalBayes* works through performing a 'moderated' t-test which, unlike a two-sample t-test, is not reliant on observed variances. Instead, it updates the posterior variance estimators for each gene based on a universal prior calculated from the input data and the observed variance for each gene using a closed-form approach. This 'shrinking/squeezing' of the variance allows for more robust inferences, particularly when the sample size is low, reduces false positive rates for genes with small variances, and improves power for genes with large variances [256].

When applied to proteomics, however, *limma* has an inherent limitation as it is unable to account for the difference in quantitative variation for proteins identified by varying number of peptides or peptide spectrum matches (PSMs) [258]. In its default setting, *limma* estimates a singular prior for all the genes in the dataset. As can be observed in the data from Granata_2021 (Figure 2:13) and unpublished DIA data from the Bishop lab (*Botella_2022*) (Figure 2:14), this heavily underestimates the variance of proteins with a low peptide count. Further, the estimated prior variance does not track the underlying observed data. On the other hand, the DEqMS is able to affect the underestimation and overestimation of variances,

respectively, towards either end of the peptide count/PSM distribution. According to [258], DeqMS was able to identify approximately 70% more differentially expressed proteins, all whose evidence of regulation was supported by RNA-seq and miRNA target prediction, than *limma*, in microRNA targeted U1810 cell-lines. This can be attributed to *limma's* tendency, which can also be observed in Figure 2:15, to overestimate variance for proteins identified with higher PSMs/peptide counts.

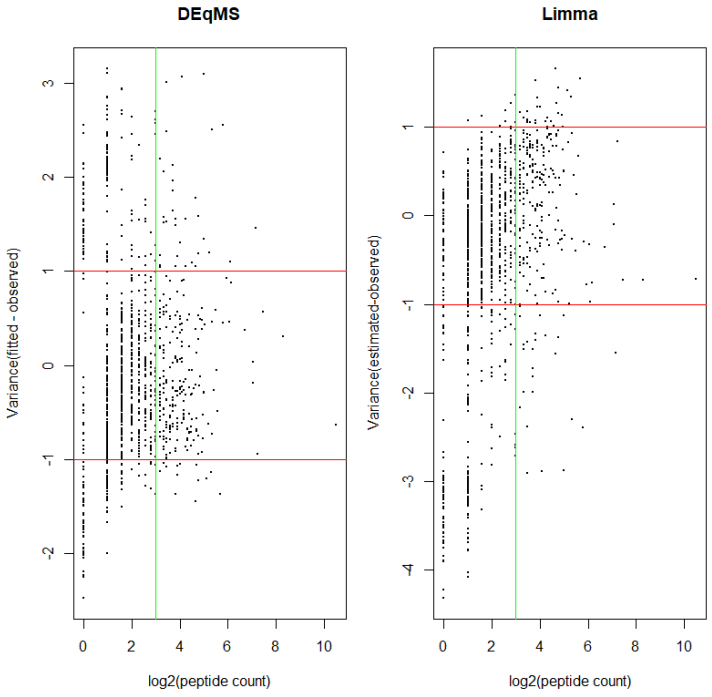


Figure 2:14 Comparison of accuracy of estimated variance of VSN-transformed protein expression values between limma and DEqMS. Data from Granata et.al (2021).

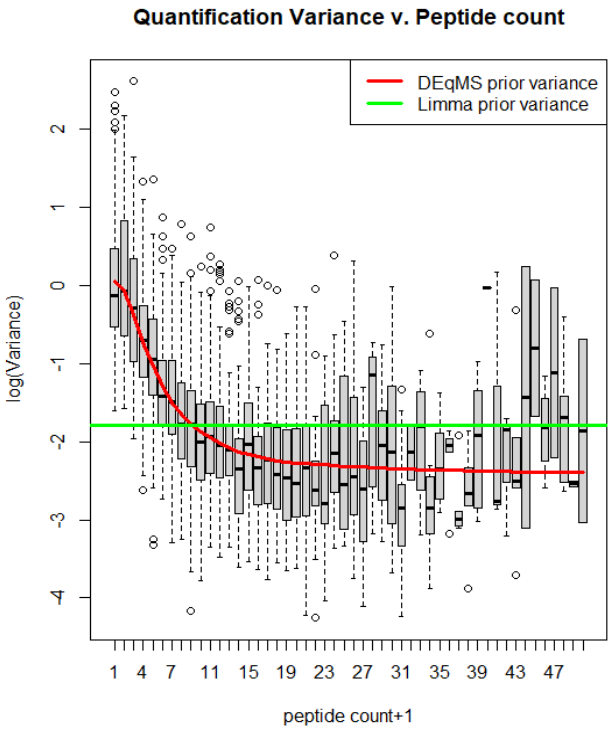


Figure 2:13 Observed and estimated prior variance for limma and DEqMS on transformed protein expression values. Boxplots demonstrate the spread of observed variance at different degrees of peptide identification. The lines demonstrate the trend of estimated variance over different degrees of peptide identification. Data from Botella_2022 (unpublished).

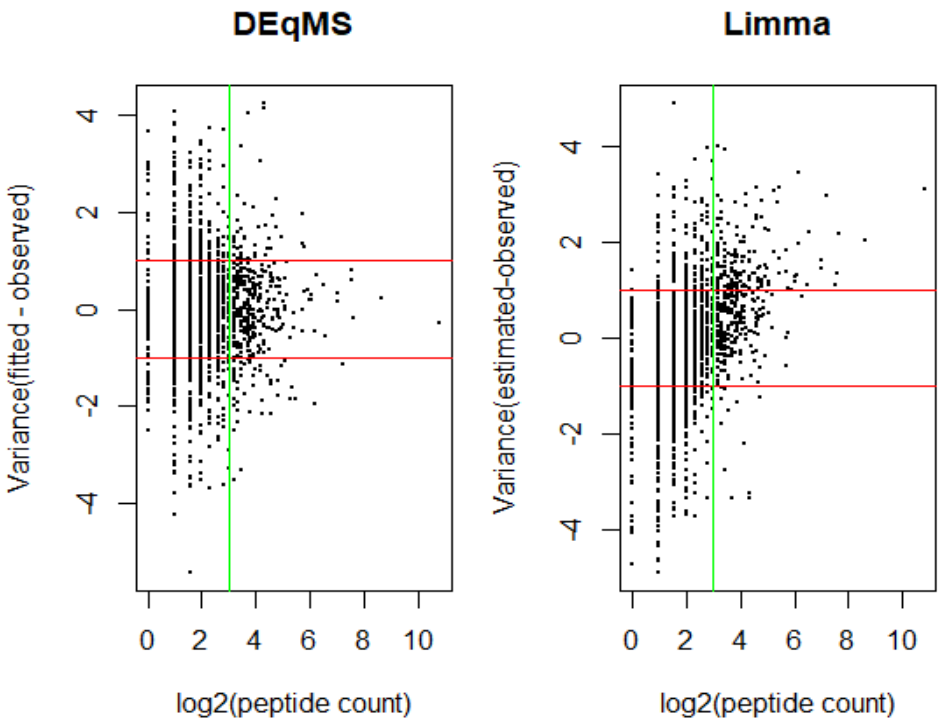


Figure 2:15 Comparison of the accuracy of estimated variance of transformed protein expression values between limma and DEqMS. Notice the greater concentration of points in the upper right quadrant of the limma plot, demonstrating the tendency of limma to overestimate the variance for high peptide count/PSM proteins. Data from Deshmukh_2021.

As discussed above, the peptide count/PSM associated with each Protein/Protein group identification was used as an input to quantify the degree and confidence of the differential expression. The *DEqMS* documentation recommends using the minimum of the sum of the 'Razor peptides' and 'Unique peptides' associated with each sample to correct the bias of variance estimate. Where 'Razor +Unique' peptide count was not reported, 'Razor peptides' and 'Unique peptides' for each sample were summed and the *rowmin* (row minimum) was selected. In the case of DIA data generated from *Spectronaut* [260], these measures of peptide attribution are not available and, therefore, the 'Number of Stripped Sequences used for Quantification' or 'Number of Precursors Used for Quantification' was used as these metrics equivalently provide peptide identification information in DIA data.

### *2.2.2 Implementation of Differential Expression Analysis*

The *model.matrix* function from the *stats* module in R [261] was used to build the design matrix for the differential expression analysis. Samples were initially split on groups with respect to their sampling time; i.e., Pre-training, Post-training etc. For studies where there were multiple timepoints/training volumes at which samples were extracted, the samples were split into multiple groups (e.g., see Granata_2021 *[81]*). For each replicate in the sample, a factor-level variable was created and included in the design matrix. For data where 'batch' information was available, a factor-level variable was created to account for batch effects. Factors of age and sex were not included in the model, largely as with most data included in this study there was a high degree of homogeneity along those factors. Further, for most if not all of the data available, sample-specific information along the variables of age and sex was not reported. Finally, no intercept term was included in the linear model.

For non-mitochondria-normalised data, differential expression analysis was run on the complete available data, followed by the extraction of mitochondria-specific proteins as described above. This allowed for a wider range of peptide level information to be available to the *DEqMS* model and, thus, to better model the prior variance for each protein.

The *lmfit* function from *limma* was used to estimate a linear model, for each protein, of the log/*vsn* transformed expression matrix along the factors included in the design matrix. The *contrast.fit* function was used to generate the $\log_2$ fold change (*logFC)* values for each protein. This was used in conjunction with the *ebayes* function to calculate the p-values for moderated t-tests as described before. The output object from the *ebayes* was supplemented with the 'peptide count data' (calculated as described above) for each protein and was subjected to the *spectraCounteBayes* function from *DEqMS* to calculate the 'count-adjusted p-values' for each

protein. For the purposes of the meta-analysis, the protein-wise standard error of the mean (SEM) was calculated using the *unscaled standard deviation* for each of the contrasts of interest and the estimated posterior variance of each protein as calculated by *DEqMS.*

A *Benjamini-Hochberg* (BH) correction was applied to the 'count-adjusted p-values' to correct for multiple hypothesis testing [262] using the *p.adjust* function from the *stats* module. For the purposes of downstream enrichment, visualisation and comparative analysis, proteins with a 'count adjusted p-value' of less than or equal to 0.01 were considered to be differentially expressed. False-discovery-rate (FDR) adjusted p-values were not used to shortlist differentially expressed proteins from each dataset for the purposes of comparative enrichment analysis. This is primarily because the proteome coverage, and thus the number of independent comparisons, within each dataset was considerably different. As the purpose of this component of the study was to analyse the degree of consensus/divergence in regulation at the level of pathways and individual proteins, using raw p-values allowed the comparisons between data to be unaffected by the depth of each dataset. To control the false positive rate, the significance threshold of 0.01 was used instead of the more generally applied 0.05 threshold. Finally, the use of unadjusted p-values has been previously argued to be justified in scenarios such as this where the increased type II error rate with p-value adjustments may limit the full potential of exploratory comparative analysis [263] [264].

## 2.3 Application of Meta-analysis Techniques

The *Metafor* package [265] in R was used to compute the meta-analysis results[3] through the *rma* function. The Mean and Standard Error of the fold change for each protein, calculated

[3] The script used for running and building the meta-analysis can be found in the Appendix in the folder 'Code'.

as described above, along with the sample size (n) for each study, was used as the input to the model. A Random-Effects meta-analysis was used, instead of a Fixed-Effects model, as the underlying assumption is that the true effect of the treatment (exercise intervention) differs between the studies included due to the inherent differences in the modality and duration of the treatment applied, along with the random sampling and technical variability because of the different sample extraction and proteomic measurement techniques used [266]. The 'Restricted Maximum Likelihood Estimator' (REML) was used to estimate the variance of the distribution of the true effect sizes/as the estimator of between-study heterogeneity. Other widely used methods for between-study variance estimations include the DerSimonian-Laird (DL) estimator, which has been shown to perform poorly with smaller sample sizes, and the Sidik-Jonkman estimator (SJ) [267], which is regarded as highly conservative compared to other methods [268]. The REML is the most commonly used method for continuous type data and is the default setting for the Random-Effects function in *metafor*; adjusted p-values for summary effects were reported alongside the effect size and confidence intervals. $I^2$ percentage values and prediction intervals for each protein, mainly due to their wide acceptability and ease of interpretability, were reported as a quantification of inter-study heterogeneity [268]. The *metaviz* package was used to create visualisations in the form of forest plots using the *vizforest* function [269].

## 2.4 Enrichment analysis and visualisation

The *clusterProfiler* [270] and *ReactomePA [271]* R packages were used for the purposes of overrepresentation analysis and Gene Set Enrichment (GSE) analysis. Enrichment analysis and GSE analysis were generally performed for GO:BP terms and Reactome Pathway

terms. The background set/universe for all enrichment analysis, where applicable, was set as the approximately 1500 mitochondrial proteins shortlisted as described earlier in this chapter. The FDR threshold was set to 0.05 for the GO and Reactome enrichment analysis. The FDR threshold for the GSEA was set at 0.25 as recommended by Yu et.al (2019)[272] and Subramanian (2005) [273]. Minimum Gene Set Size (MinGS) and Maximum Gene Set Size (MaxGS), where applicable, were amended from their default values to 20 and 200, respectively, in order to capture more specific ontology and pathway terms and to reduce the comparisons for higher-order terms, which are generally less informative. All Uniprot IDs were converted to *Entrez* IDs for the purposes of enrichment analysis using the *mapIds* function from the R package *AnnotationDbi* [274]. The database *org.Hs.eg.db* [274] was used for mapping between different gene and protein IDs.

All volcano plots were generated using the R package *EnhancedVolcano* [275]. The results of all enrichment and GSE analysis were visualised using the R package *enrichplot* [270] and *pathview* [276]. Network analysis and visualisation were performed using Cytoscape [277]and the associated application *StringApp* [278].

# 3 CHAPTER THREE

## Results and Discussion

### 3.1 Summary of the pooled results

A total of 778 proteins were identified across the eight datasets included in this study. Of this set of proteins, 666 were identified in *MitoCarta 3.0* as being localised to the mitochondria. According to IMPI, the rest of the identified proteins were verified to be localised to the mitochondria, except for six proteins that were only predicted as mitochondrial. The DIA-based datasets expectedly demonstrated the greatest depth, with *Botella_2022* and *Hostrup_2021* profiling over 600 mitochondrial proteins each. The data from Popov_2021 [223] had considerably lower depth (approximately 180 proteins profiled) than all the other included datasets. It is hypothesised that this may be because data acquisition was run in multiple batches, and not all of the intersection of the batched data met the threshold set for missingness in this study (i.e., rows with less than 65% of data available were filtered out). Finally, the depth of the mitochondrial proteome obtained in each study also appears to be affected by the recency of the study. For instance, approximately 400 mitochondrial proteins were ultimately extracted from the data of Schild_2015 compared to 465 mitochondrial proteins from the Deshmukh_2021 data and approximately 550 mitochondrial proteins from the Granata_2021 data. This observation is consistent with the trend of technological advances in mass spectrometry instruments, improvements in sample extraction and preservation methods, and the increasing sophistication in the use of this technique in the area of exercise physiology [279] [178, 214]. Lastly, the depth of the mitochondrial proteome identification reported here may differ slightly from what is reported in the respective publication for the following reasons. First, only the mitochondrial fraction, as opposed to the whole muscle proteome, has been the subject of this research and, therefore, only mitochondrial protein quantifications have been reported. Secondly, the published studies may have used a different

threshold for missingness compared to the common method used in this research to filter out data with missing values in all of the included datasets.

### *3.1.1 Localisation and basal expression of identified mitochondria proteins*

Sub-mitochondrial localisation of the identified mitochondrial proteome was investigated to examine the representation of different mitochondria components. As can be observed from Table 3:1 and Figure 3:2, most of the identified mitochondria-annotated proteins were reported in the MitoCarta 3.0 and approximately 70% of the identified proteome was annotated as being localised to the Inner Membrane (MIM) and the matrix. The major difference between the distribution of the identified mitochondria proteome compared to the unidentified set was that a larger proportion of the latter was not annotated in *MitoCarta 3.0* (see Table 3:1). The reported total peak intensity of proteins in skeletal muscle tissue from *Mitocarta 3.0* was also used to analyse whether there was an association between basal protein expression and identification in the data. Approximately 80% of the *MitoCarta 3.0* annotated proteins in the unidentified protein set showed missing spectral intensity values for skeletal muscle tissue compared to only 27% in the identified proteome. In other words, it is likely that the lack of identification for most of the unidentified mitochondrial proteins could be related to their low basal expression in skeletal muscle tissue. Finally, the association between basal skeletal muscle expression and degree of identification was analysed. As can be observed from Figure 3:1, when there is a greater degree of identification (x-axis indicates the number of studies in which a protein was identified) fewer proteins have missing expression scores (denoted by 0) and quantified proteins trend toward having higher intensity values.

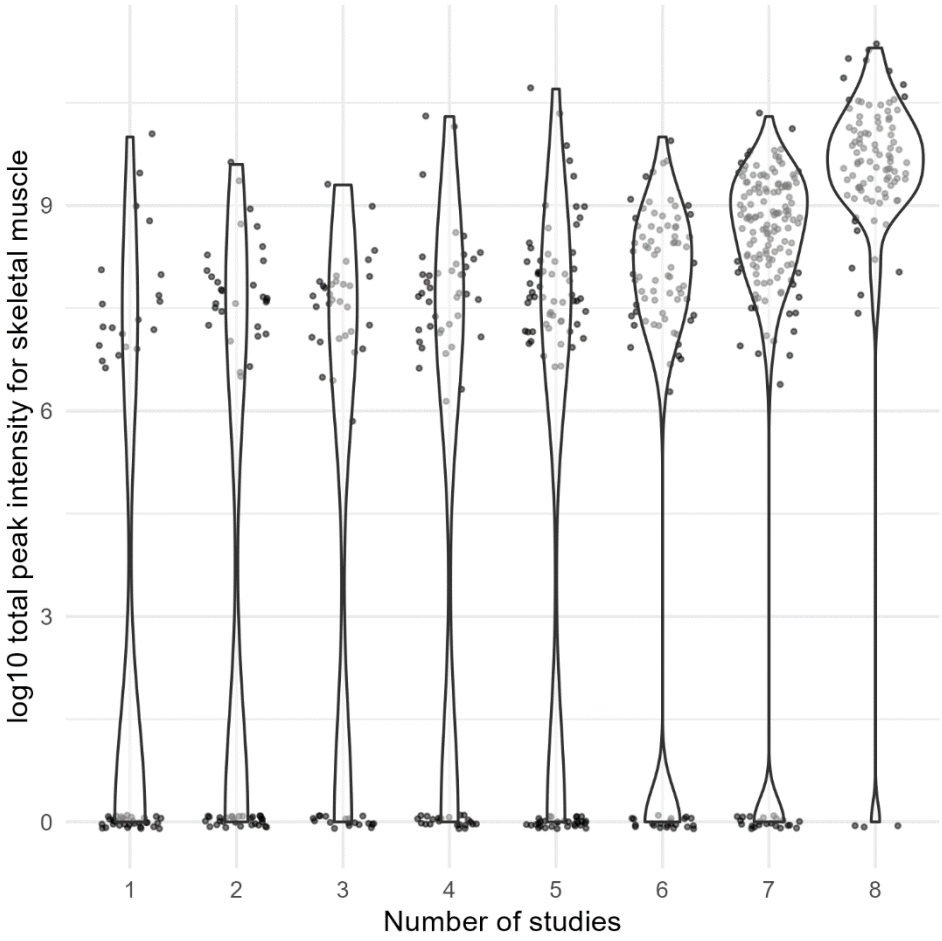


Figure 3:1 Comparison of basal skeletal muscle expression of proteins against the number of datasets (studies) in which a protein was quantified. Zero values denote missing skeletal expression quantification in MitoCarta 3.0[1]

| Localisation | Number identified | Percent identified per category | Percent of identified Proteins | Number unidentified | Percent of unidentified Proteins |
|---|---|---|---|---|---|
| IMS | 32 | 62.7 | 4.1 | 19 | 2.6 |
| Matrix | 314 | 59.3 | 40.4 | 215 | 29.3 |

| | | | | | |
|---|---|---|---|---|---|
| Membrane | 12 | 35.3 | 1.5 | 22 | 3.0 |
| MIM | 228 | 63.7 | 29.0 | 130 | 17.7 |
| MOM | 51 | 46.3 | 6.6 | 59 | 8.0 |
| unknown | 28 | 51.0 | 3.6 | 27 | 3.7 |
| Not in *MitoCarta 3.0* | 112 | 30.0 | 14.4 | 261 | 35.5 |

Table 3:1 Distribution of identified and unidentified mitochondria proteins across sub-mitochondrial localisations. Abbreviations: IMS= Inter-membrane space, MOM=Outer Membrane, MIM=Inner Membrane

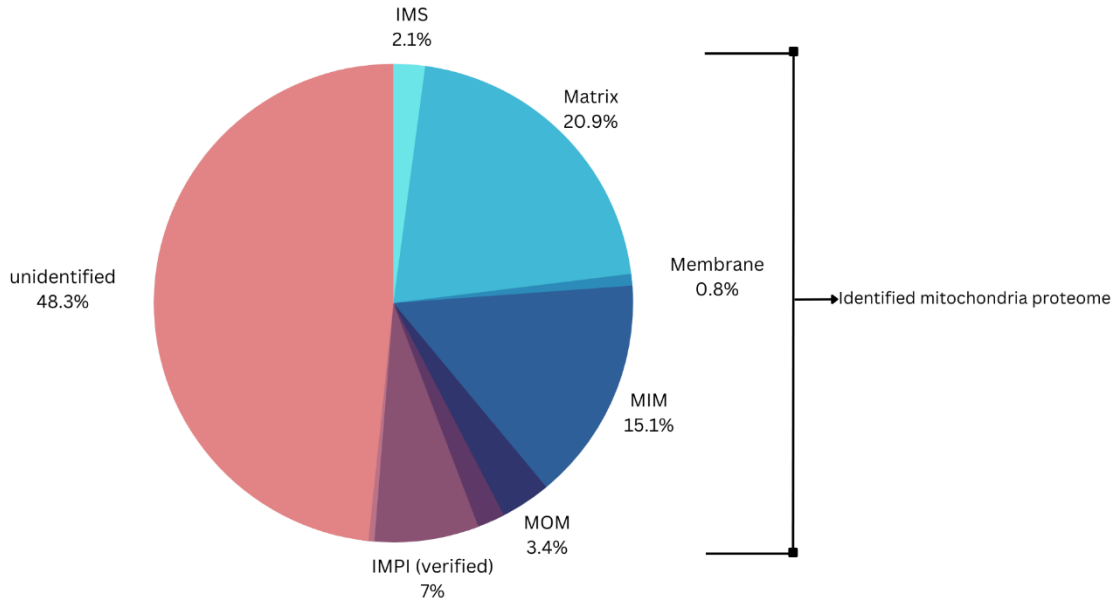


Figure 3:2 Distribution of identified mitochondria proteins across sub-mitochondrial localisations, alongside the percentage of unidentified mitochondria proteins. Abbreviations: IMS= Inter-membrane space, MOM=Outer Membrane, MIM=Inner Membrane

### *3.1.2 Intersection of datasets*

Of the 778 quantified proteins identified across the eight datasets, 113 proteins (~ 15%) were identified and quantified in all eight datasets and the cumulative participant count for these proteins was 66. Figure 3:3 provides comprehensive information on the intersections of

protein identifications between the datasets. Excluding the *Popov_2020* dataset, which demonstrated considerably lower mitochondrial proteome depth than any other data included, 173 common proteins were quantified across the other seven datasets. Of all the datasets, the data from Hostrup*_2021* contained the highest number of uniquely identified mitochondrial proteins with 58; this may be attributable to its DIA data acquisition protocol. The other DIA-based dataset, *Botella* _2022, uniquely identified 35 mitochondrial proteins. The union of the DIA-based datasets (i.e., *Hostrup_2021*, *Botella_SIT*, and *Botella_MICT*,) identified 127 proteins over the union of the DDA datasets (i.e., *NVT_Granat*a, *HVT_Granata*, *Deshmukh_2021*, *Schild_2015* and *Popov_2020)* (see Figure 3:4). A GO:CC analysis was performed on this set of unique DIA-only proteins to detect whether there was a localisation-driven pattern to the increased detection observed with DIA data acquisition. However, no significant enrichment of any terms was observed. Further, the distribution of these DIA-only proteins across sub-mitochondrial compartments did not differ substantially from the distribution of the entire quantified mitochondrial proteome. A stark difference was, however, observed in the basal skeletal muscle protein expression profiles (tissue-specific spectra peak-intensity obtained from MitoCarta 3.0) of the DIA-only proteins compared to entire quantified mitochondrial proteome. Approximately 60% of the *MitoCarta* annotated DIA-only protein set demonstrated missing spectra values in skeletal muscle tissue analysis compared to only 27 % for the whole quantified proteome. This suggests that the increased depth and unique identifications from the DIA datasets is due to its profiling of lower abundance proteins. The capacity of DIA-based data acquisition methods to be less biased toward higher abundance proteins and thus acquire a deeper proteome coverage has been demonstrated previously, as discussed in the literature review, and it is evident in these data as well.

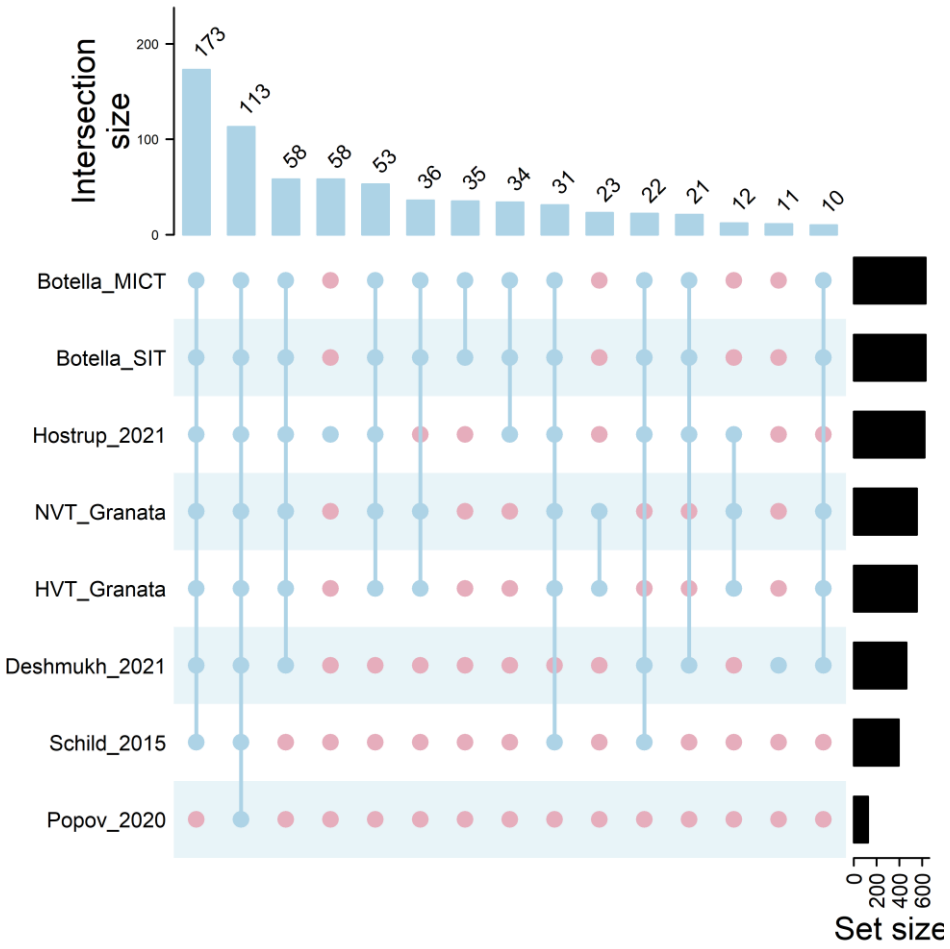


Figure 3:3 Intersection of protein identifications across datasets. The set size denotes the number of proteins within each dataset.

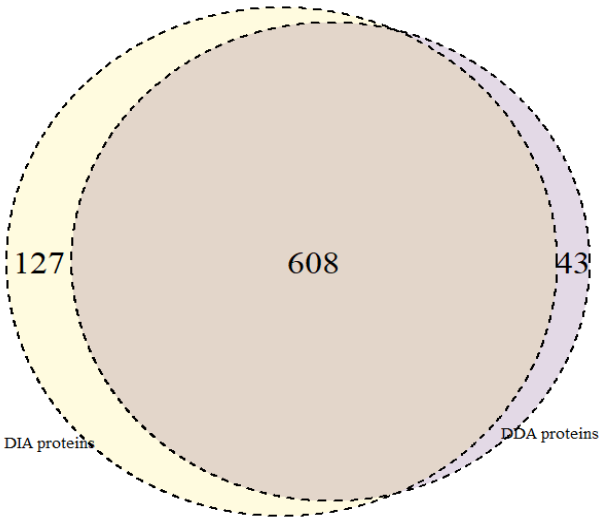


Figure 3:4 A *Venn* diagram demonstrating the overlap between Data Dependent Acquisition and Data Independent Acquisition identified protein sets

A GO:CC enrichment analysis was performed on the mitochondrial proteins commonly identified in seven or more datasets[4] ($K \geq 7$) (289) to determine whether any sub-mitochondrial locations were overrepresented in these datasets (Figure 3:5). Relevant enrichment of only inner membrane-related terms was found despite the identification of all four mitochondrial compartments. Further, except for CII, terms related to all other individual respiratory complexes were enriched. The finding of enrichment of inner membrane localisations in the set of the most common proteins is supported by the high relative basal abundance of respiratory chain complex proteins in the mitochondria [280].

Relatedly, the 289 high-intersection proteins were also analysed for their representation of major mitochondrial functional groups. Within the ETC complex proteins, CI, CII, and CIII demonstrated the highest depth with approximately 75% of each complex's subunits represented in the set. Complexes IV and V, on the other hand, had only approximately 50%

[4] The degree of identification (number of datasets quantified in) for proteins is abbreviated as *K* within this thesis.

of their subunits represented. Further, CIV and CV also demonstrated the highest proportions of non-quantified protein subunits. Similarly, approximately 75% of the TCA cycle enzymes were present in this set compared to 50% of the fatty acid oxidation related proteins and 30 % of the Protein Import and Mitochondria Dynamics related proteins (Figure 3:6 and 3:7).

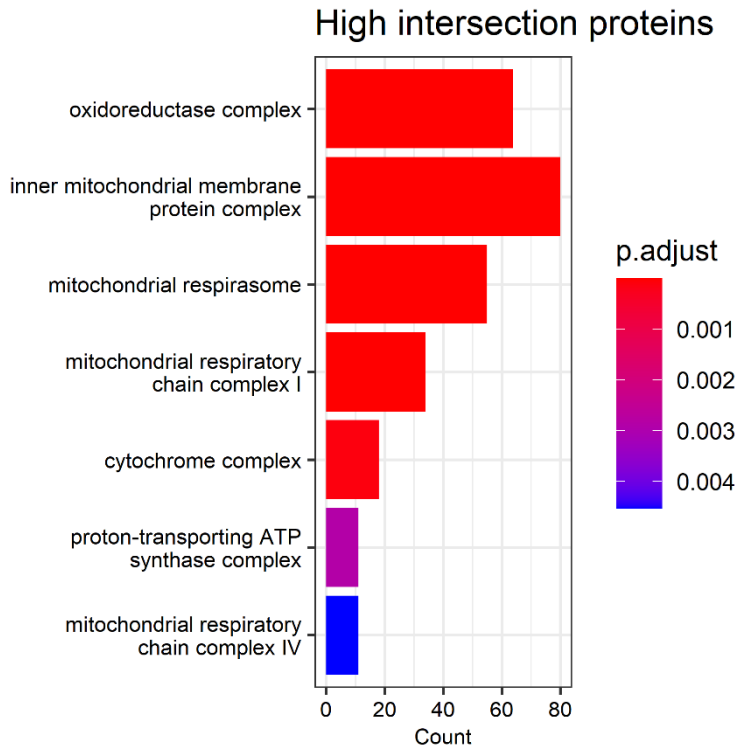


Figure 3:5 Results of a GO:CC (Gene Ontology-Cellular Component) enrichment of commonly identified proteins between the datasets.

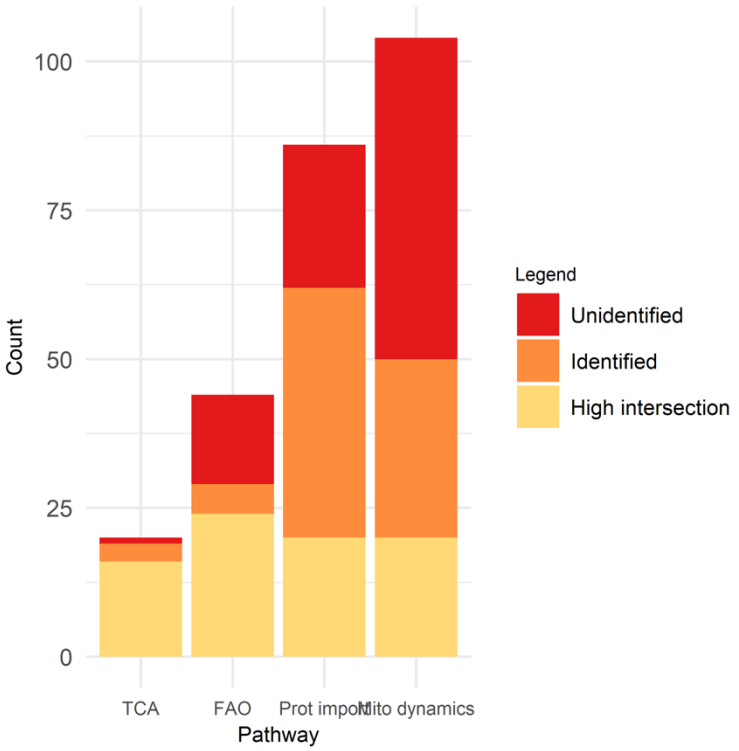


Figure 3:7 Distribution of other pathway proteins according to identification status. *Unidentified* refers to the proteins not observed in the data. *Identified* refers to proteins identified only once within the datasets. High Intersection refers to proteins with K ≥ 7. Abbreviations- TCA: TCA cycle enzymes; FAO: Fatty Acid Oxidation proteins; Prot. import: Protein Import pathway proteins; Mito dynamics: proteins affecting fusion/fission.

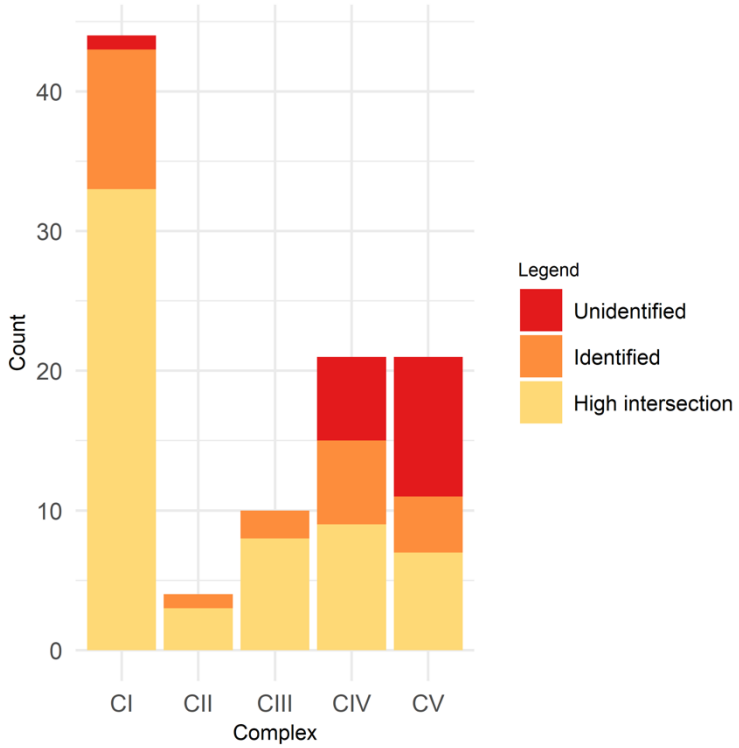


Figure 3:6 Distribution of Electron Transport Chain complex proteins according to identification status as described in Figure 3.6.

### *3.1.3 Pooled results for differentially expressed mitochondria proteins following training*

Differentially expressed proteins (DEPs) were classified as those demonstrating peptide-identification adjusted p-values (sca.p.value[5]) ≤ 0.01. In total, 365 mitochondrial proteins were identified as differentially expressed following training in at least one of the

[5] This refers to the p-values as calculated by the package DEqMS when performing the differential expression analysis

datasets[6]. Functional analysis identified approximately 70% of the CI and CII proteins, 50% of the CIII, CIV and CV proteins, 85% of TCA cycle enzymes, 50% of the fatty acid oxidation related proteins, and 25-30% of Protein Import and mitochondria Dynamics annotated proteins (Figure 3:8). Relatedly, a GO: BP enrichment analysis of this set of proteins yielded terms related to the ETC process, TCA cycle, fatty acid oxidation and amino acid metabolism. This finding also demonstrates that, possibly in response to the increased energy demand, exercise training can upregulate protein expression across all major metabolic pathways (Figure 3:9).

[6] Differential expression results for all proteins for each dataset is reported in the Appendix file 'Data characteristics and results.xls'

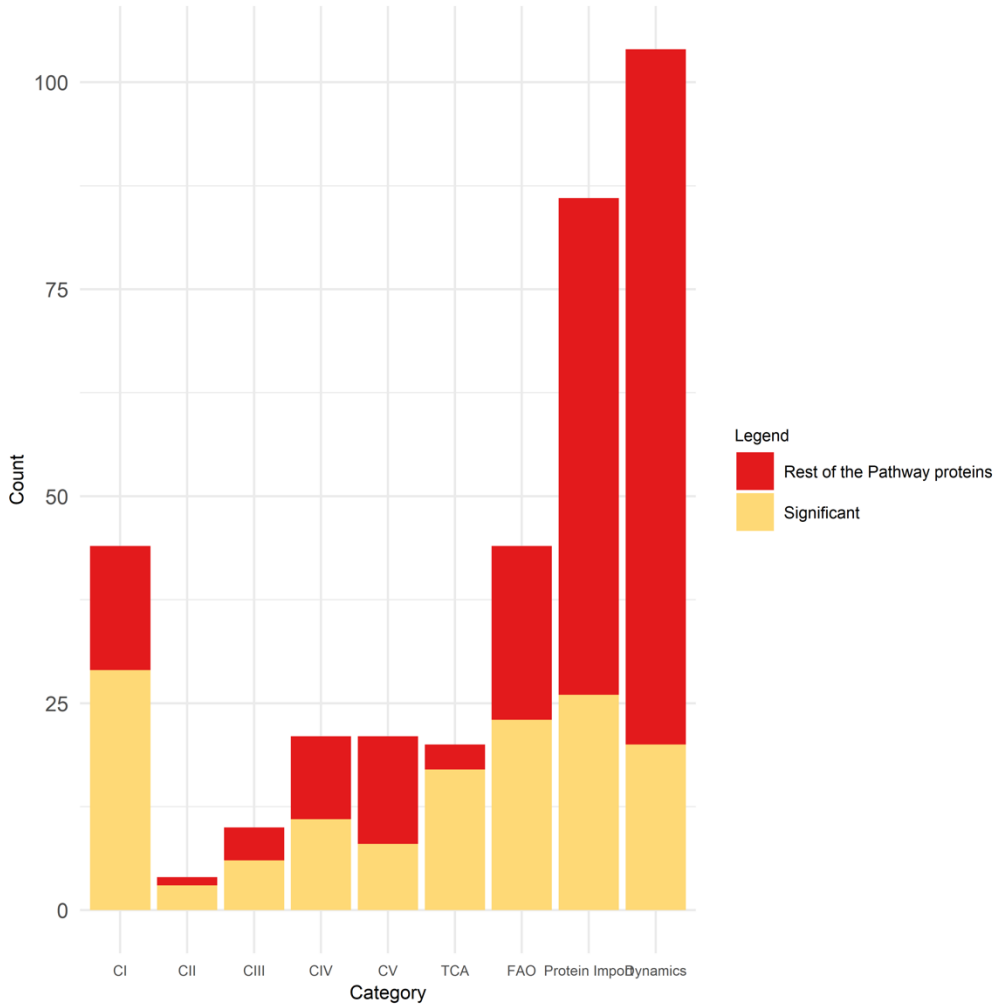


Figure 3:8 Overview of distribution of the differentially expressed proteins following training within key mitochondrial functional pathways. *Significant* refers to proteins found significant in at least one of the datasets. Abbreviations as in Figure 3.6

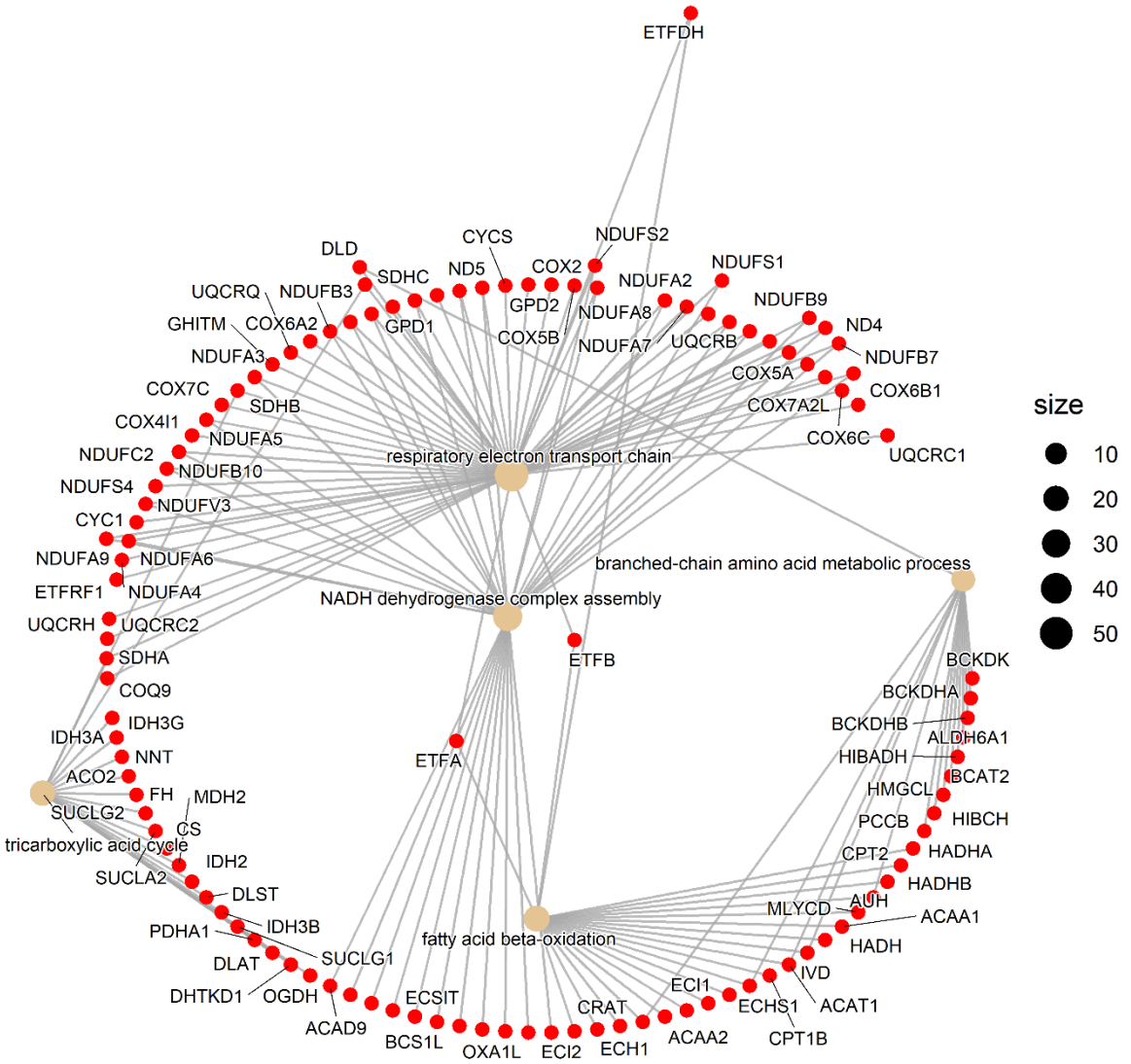


Figure 3:9 Visualisation of the GO:BP (Gene Ontology- Biological Process) analysis on proteins identified as significant in at least one dataset

The highest number of upregulated DEPs (with a sca.p.value < 0.01) were observed in the two *Granata_2021* datasets, followed by the *Hostrup_2021* and *Botella_SIT* datasets (see Figure 3:10). Very few DEPs were observed in the *Popov_2020* dataset, and this may be attributed to the small number of proteins identified and the strength of the batch effects observed. Differential expression profiles of data depend on many factors, which may have contributed to the observed differences in the number of significant DEPs observed between the datasets. For the *Granata* datasets, MS data was collected from isolated mitochondrial

fractions of muscle lysate samples, and this could have improved the accuracy of protein quantification due to the absence of high-abundance contractile proteins [81]. The sample size in this study (10) was also the highest among those included in this research and was twice as high as the sample sizes for the *Deshmukh_2021, Schild_2015,* and *Popov_2020* datasets; *this* would have conferred greater statistical power. These technical and study-design variables may have had a stronger effect on the number of DEPs identified than exercise training-related variables, especially as some of the other studies demonstrated higher overall exercise volume. The *Schild_2015* study, despite having the highest overall exercise exposure of all included studies, identified one of the lowest numbers of DEPs. As this is the oldest dataset included in this study, technical limitations of mass analysers prevalent in 2015 may have significantly affected the differential expression profile. Further, the nascency in the adoption of gel-free mass spectrometry-based proteomics in the exercise physiology field in 2015 may have also been a factor as sample preparation techniques and LC-MS/MS assemblies have evolved in sophistication [215]. Finally, as this study followed a cross-sectional design, biological variability between replicates may have strongly affected intra-group variance and thus diminished the significance of the true effect of exercise training.

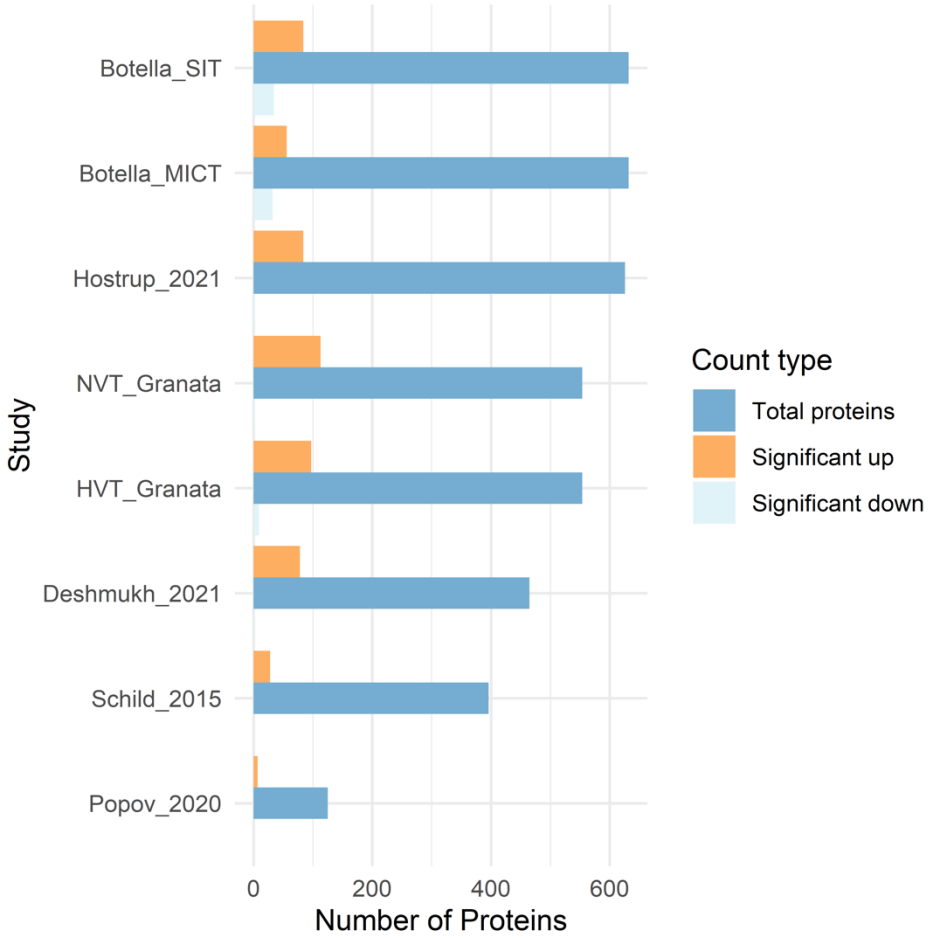


Figure 3:10 Distribution of the number of differentially expressed proteins across the included datasets.

For the *Popov_2020* data, the lowest number of DEPs were observed in both absolute (7) and proportional terms (5% of the identified mitochondria proteins). The results for this data in the associated publication, on the other hand, reported significant changes in abundance for 250 proteins (of which > 100 were mitochondrial); this is very different to the findings of the current analysis [223]. The published study utilised a paired Wilcoxin Signed-Ranks test, as compared to the moderated t-tests in the pipeline for the current analysis and reported median instead of mean $\log_2$ fold changes. The strength of the batch effects[7] within the data from this

[7] Due to the strength of batch effects observed, the adjustment of batch effects for Popov_2020 was attempted using the *removeBatchEffect* function from *limma* on the VSN-transformed intensities before performing differential expression analysis

study is evident in an exploratory analysis of the VSN-normalised data (as shown in Figure 3:11) and it is suspected that this was the major contributing factor affecting the differential expression results observed. The published results may have been less affected by the batch effects, however, as the Wilcoxin Signed-Ranks test only requires pairwise differences as input [281]. According to the sample-level information, all samples within a pair were included in the same batch and this would have mitigated the influence of batch-effects. Given that the pipeline utilised has yielded reproducible results for all the other data included in the current analysis, and in the interest of maintaining uniformity of inputs for the meta-analysis design, the same parametric statistical approach as described in the methods was applied to this data. The effects of exercise modality and volume on differential expression results will be discussed in a subsequent section. However, even when excluding the *Popov_2020* and *Schild_2015* results, it is unclear whether the number of significant changes observed is related to variability in mass spectrometry protocols, instrumentation, sample preparation, sample size, or differences in exercise training variables.

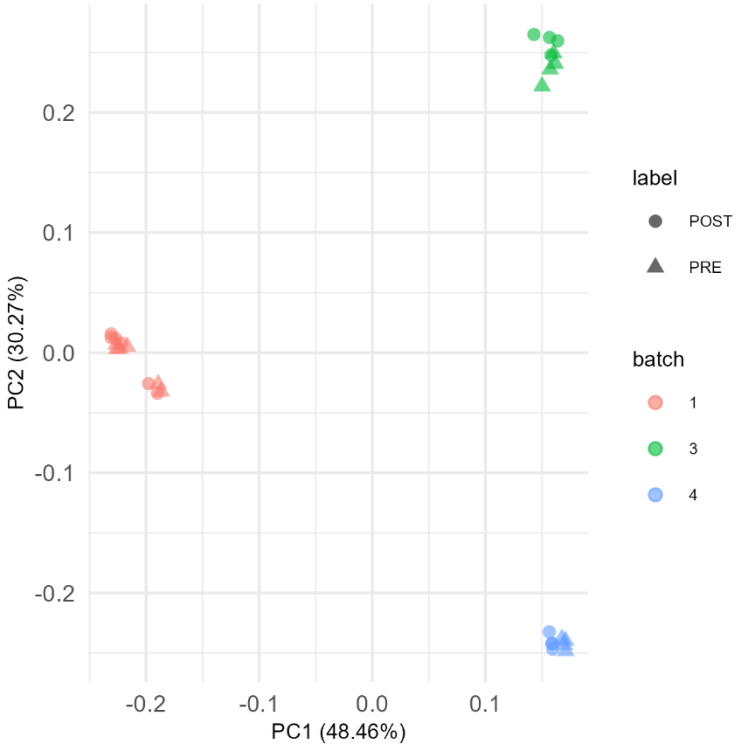


Figure 3:11 PCA plot of VSN-transformed data from Popov_2020, which demonstrates the strength of the batch effects

### *3.1.4 Intersection of the Differentially Expressed Mitochondrial Protein Results*

A comprehensive intersection analysis was performed on the proteins identified as differentially expressed (henceforth referred to as DEPs) in each of the datasets (see Figure 3:12). While many proteins demonstrated differential expression across multiple datasets, it was surprising that no single protein was found to be differentially expressed in all datasets. It is unclear why this was observed; however, it is hypothesised that this was not driven by differences in exercise-related variables but by the technical variability between the data. This is reinforced by the finding that despite a relatively low degree of intersection of common proteins, multiple common pathways were found to be enriched between the differential expression results of these datasets (discussed later).

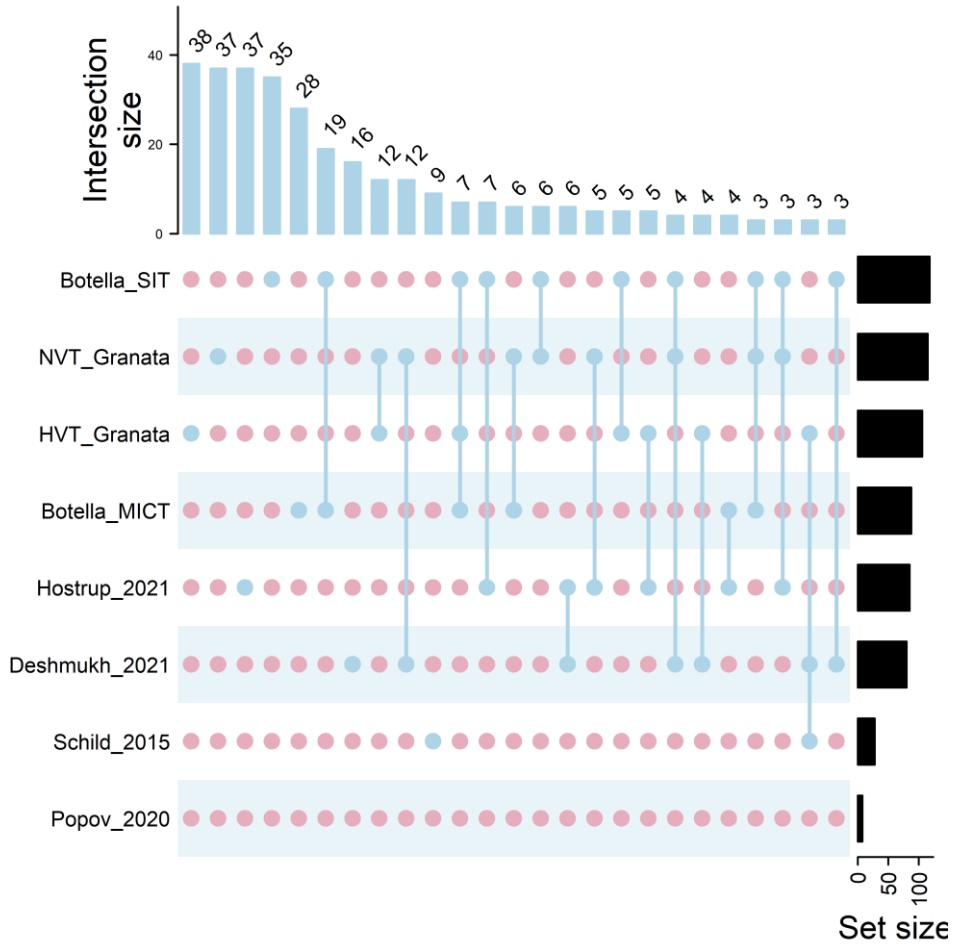


Figure 3:12 Overview of intersection degrees between significant proteins identified within each dataset

The highest degree of unique intersections was observed in the two *Botella_2022* datasets, with 19 unique proteins identified as differentially expressed despite following distinctly different training protocols. No significant enrichment of any GO terms was observed in this intersection set. The *Granata* datasets cumulatively identified 87 unique DEPs not observed in the other data. No significant GO:BP enrichment for any GO terms was observed in this set; however, there were many proteins within pathways of mitochondrial translation (16) and the ETC (8). The data with the deepest proteome profile, *Hostrup_2021*, similarly identified 37 unique DEPs; these included eight Complex I and three Complex V subunits. Running a GO:CC enrichment on this set revealed an enrichment of terms related to the ETC

complexes (see Figure 3:13), perhaps suggesting that the DIA acquisition method allowed for greater resolution in discerning the fold changes of this group of proteins.

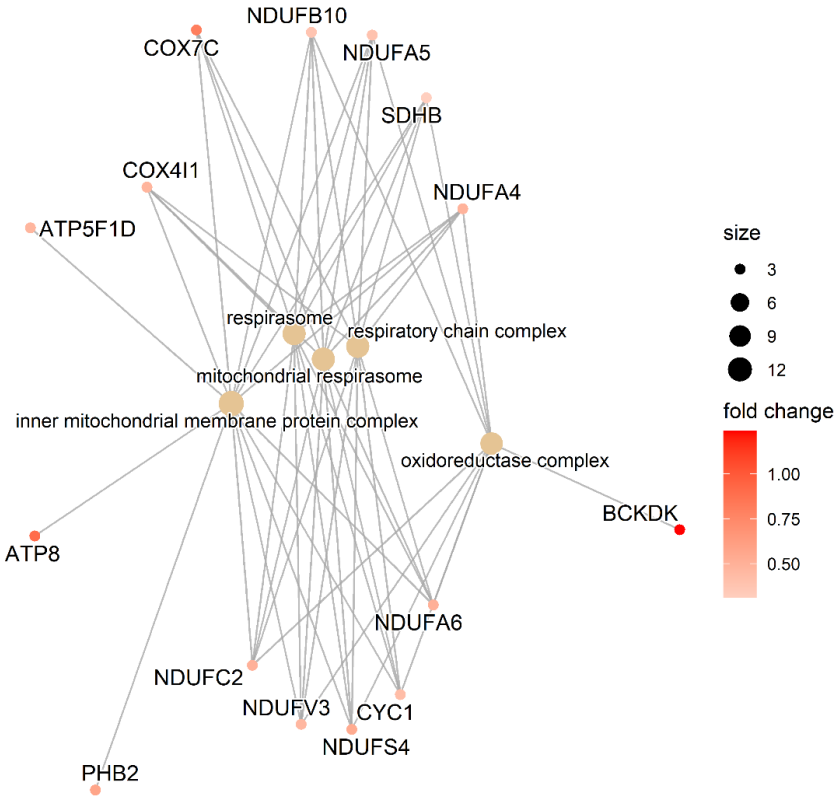


Figure 3:13 Visualisation of Reactome enrichment analysis on unique significant proteins identified from Hostrup_2021.

There were 164 mitochondrial DEPs in more than one dataset ($K \geq 2$), and 62 DEPs in three or more datasets ($K \geq 3$). A Reactome enrichment analysis on DEPs with $K \geq 3$, which include multiple Complex I subunits, Complex III and IV subunits, and four Citric Acid cycle enzymes, revealed an overrepresentation of terms related to Complex I biogenesis, the TCA cycle, and Electron Transport Chain (see Figure 3:14). Interestingly, the terms related to fatty acid oxidation and amino acid metabolism that were seen on the larger set of proteins (DEPs with K=1) were lost, and, instead, mitochondria protein import and localisation terms were observed to be enriched.

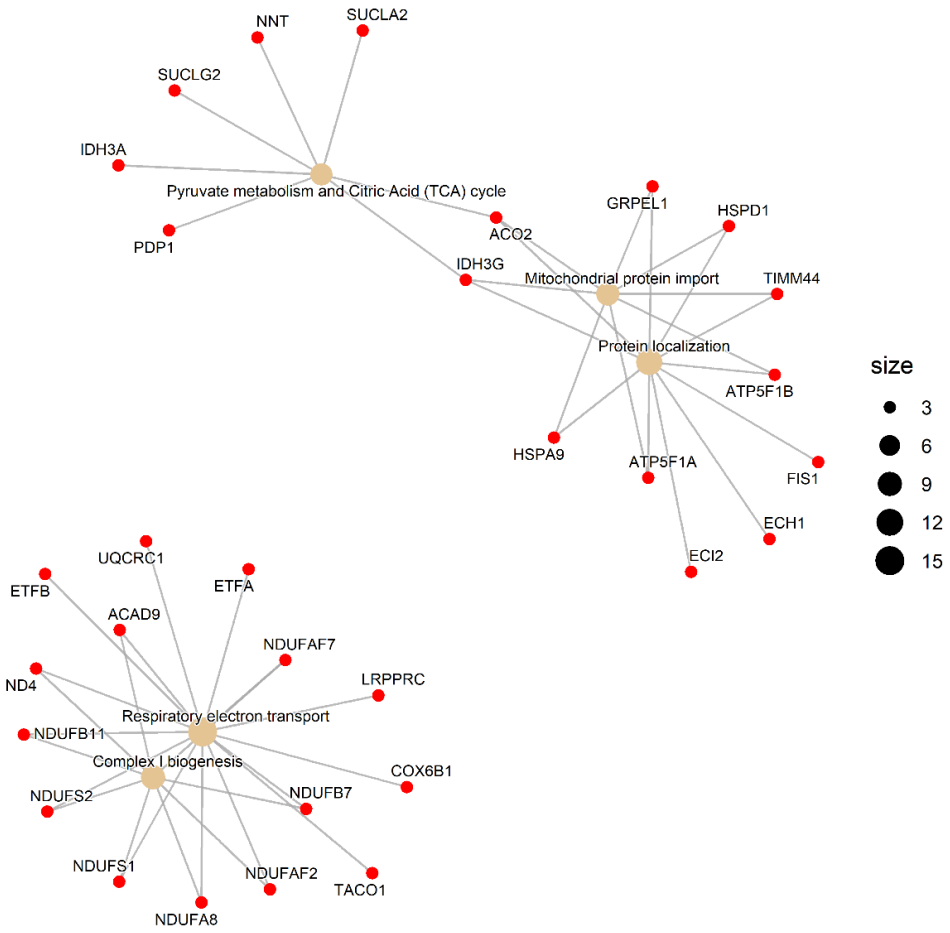


Figure 3:14 Visualisation of Reactome enrichment analysis of proteins significant in three or more datasets.

The Botella_SIT data, which was the only study to employ a sprint interval training (SIT) protocol, identified 35 unique DEPs. Among the set of unique upregulated SIT proteins (22) were two members of the Pyruvate Dehydrogenase Complex (PDC), PDHA1 (Pyruvate Dehydrogenase E1 Subunit Alpha 1; P08559) and DLAT (Dihydrolipoyl transacetylase subunit E2; P10515), another pyruvate converting enzyme LDHD (Lactate Dehyrdogenase D; Q86WU2), and two enzymes implicated in acetyl-coA synthesis COASY (Coenzyme A Synthase; Q13057) and ACSS1 (acetyl-CoA synthetase short chain; Q9NUB1). Most of these proteins are of functional significance to pyruvate metabolism either through aerobic (pyruvate

conversion to acetyl-CoA) or anaerobic pathways (pyruvate conversion to lactate through lactate dehydrogenase). Given that these proteins were quantified in multiple datasets, it is interesting that only the SIT-based protocol demonstrated adaptations central to pyruvate metabolism, suggesting specific effects of this training modality. Indeed, PDHA1 protein content [282] and activity [283] have previously been found to be upregulated in response to six and two weeks of SIT, respectively. Increased glycolytic capacity [284] and lactate transport capacity [285] have also been reported previously following SIT, suggesting a shift toward carbohydrate metabolism with this type of training and supporting the current finding of unique mitochondria proteins adapting to meet the increased pyruvate supply.

Similarly, the Schild_2015 data, which is based on a cohort design as opposed to the longitudinal design of all the other data, identified nine unique DEPs between endurance-trained and untrained samples; this is approximately 30% of all DEPs in the data. Four CV subunits (ATP5PB, ATP5PF, ATP5L, ATP5PD) and the CII subunit SDHA (Succinate Dehydrogenase A) – the only ETC complex protein also part of the TCA cycle, were observed among this set. Overall, six different CV subunit proteins were found to be significantly upregulated in the *Schild_2015 dataset*, the highest among all the included data. As the Schild_2015 data was the only study to investigate the effects of long-term exercise training, this result may suggest an effect of exercise volume on adaptations to CV protein abundance. Relatedly, SDHA, the second entry point for electrons into the ETC, also appears to demonstrate a dependence on exercise volume (see Figure 3:15). Indeed, all four CII subunits from *Schild_2015* show large $\log_2$-fold changes ($> 0.8$) compared to the other data and approach significance (sca.p.value $<0.03$). As no strong evidence for CII regulation is observed in any other data, it may be hypothesised that CII protein abundance increases more robustly in response to continuous high energy demand (long-term endurance training) to supplement electron supply from CI into the ETC.

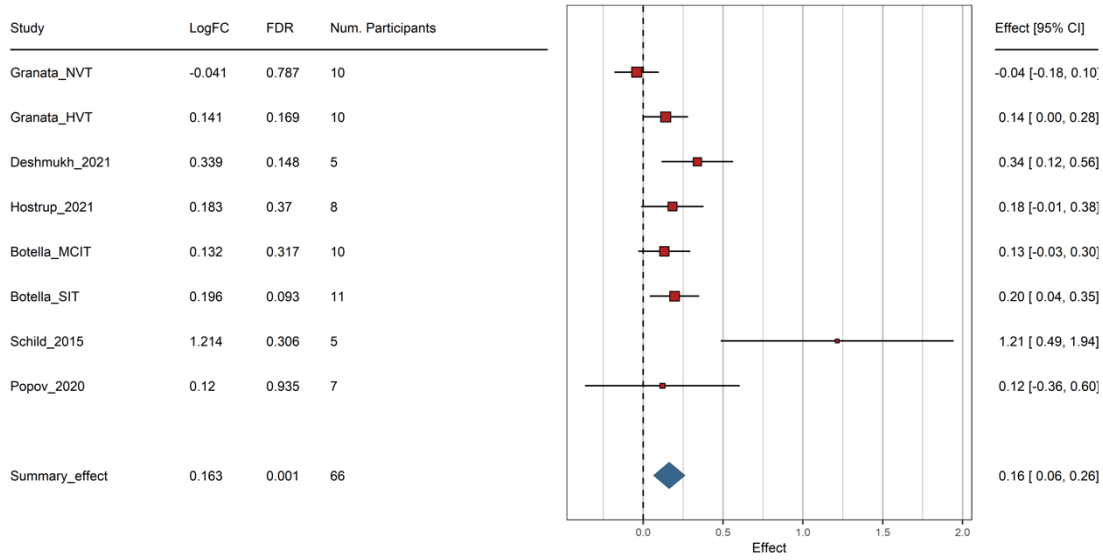


Figure 3:15 A forest plot demonstrating fold change of SDHA protein in all eight datasets included in this study. Particularly notice the large fold-change with Schild_2015. This may be related to the high exercise volume underlying the comparisons within the study.

The highest degree of intersection for any differentially expressed mitochondrial protein was six studies, and three proteins were identified in this set. This included TIMM44 (Translocase of Inner Mitochondrial Membrane 44; O43615) (Figure 3:16), a mitochondrial inner membrane translocase, ECI2 (Enoyl-CoA delta isomerase 2; O75521), a key enzyme for beta-oxidation of unsaturated fatty acids (Figure 3:17), and SLC25A3 (alias: PiC ;Solute Carrier Family 25 member 3; Q00325), the only importer of phosphate ions across the inner membrane and thus integral to the final step of oxidative phosphorylation [286, 287] (Figure 3:18).

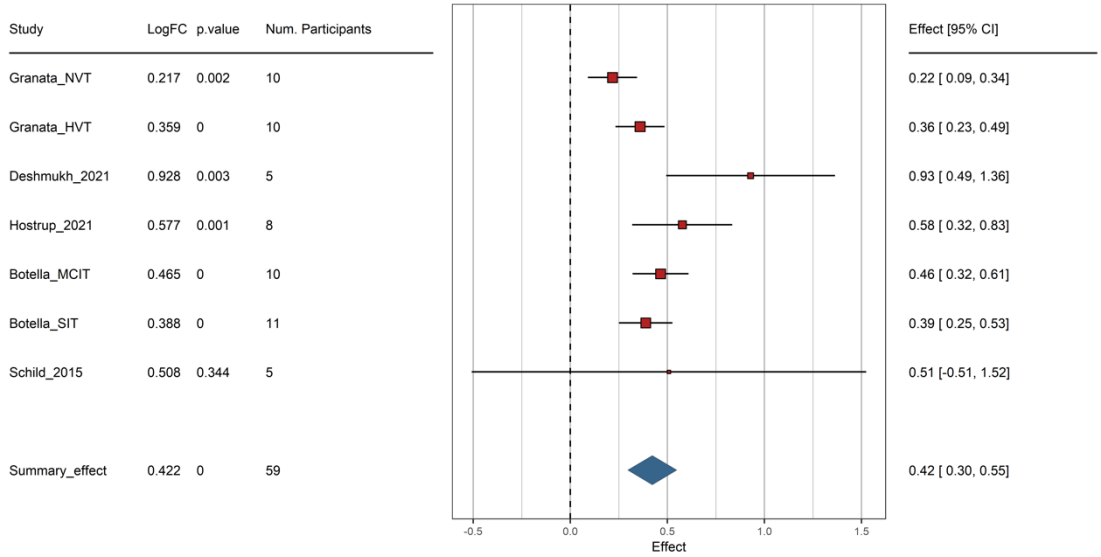


Figure 3:16 A Forest plot for *TIMM44* showcasing the high degree of differential expression demonstrated by this protein

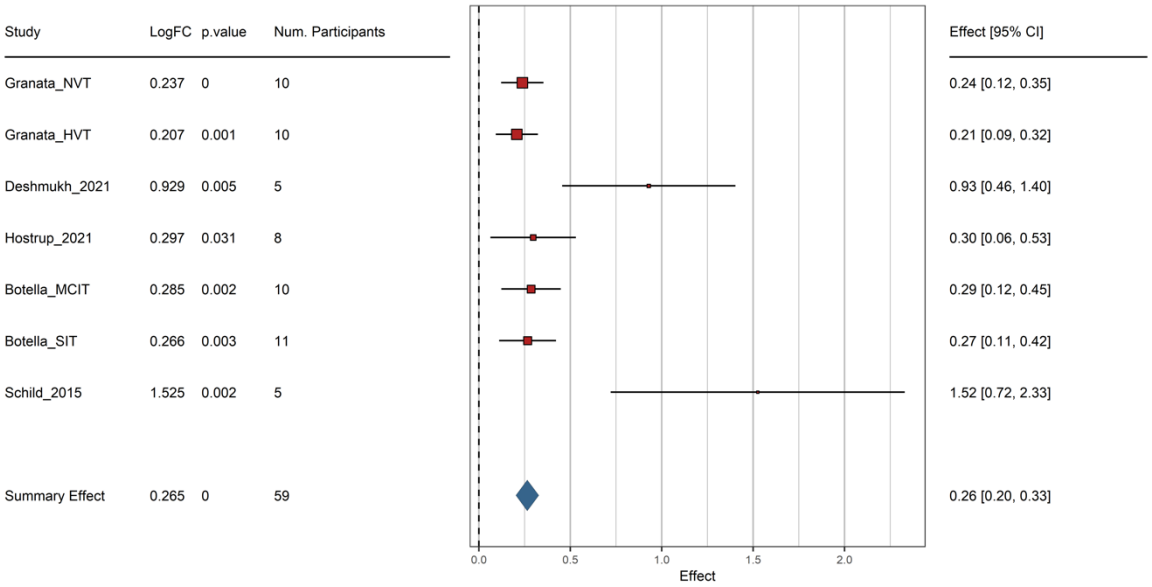


Figure 3:17 A Forest plot for *ECI2* showcasing the high degree of differential expression demonstrated by this protein

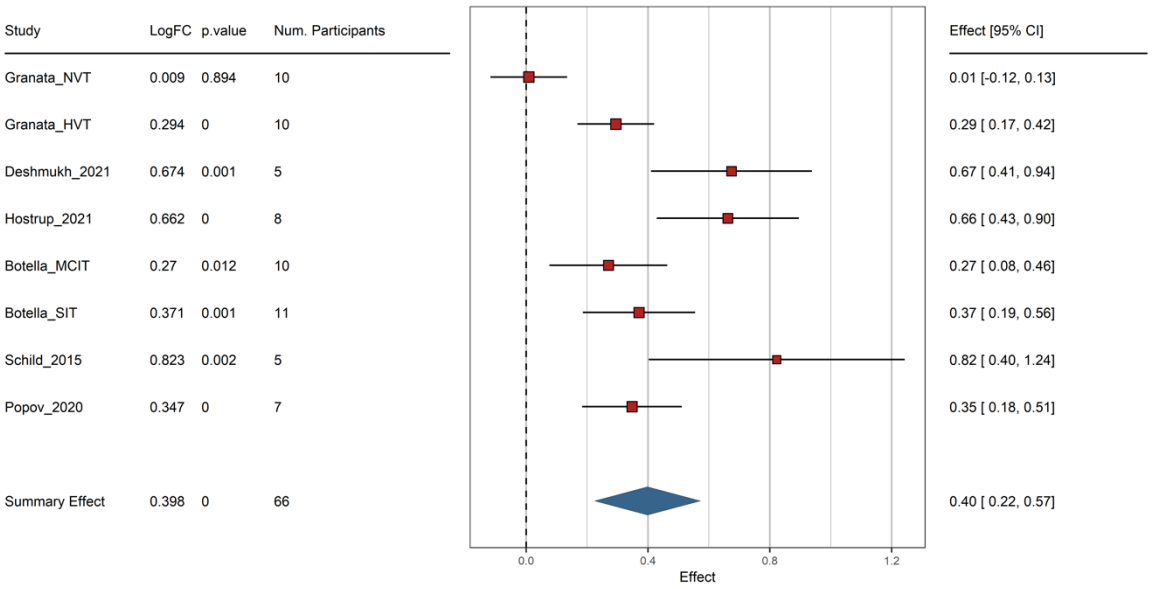


Figure 3:18 A Forest plot for *SLC25A3* showcasing the high degree of differential expression demonstrated by this protein

Furthermore, six proteins were identified as differentially expressed in five different datasets. These proteins included ACAD9 (Acyl-CoA Dehydrogenase; Q9H845) and ECH1 (Enoyl-CoA Hydratase 1; Q13011), both of which are implicated in the fatty-acid β-oxidation pathway. The ACAD family of proteins are crucial as they catalyse the rate-limiting step in the β-oxidation pathway [288]. The rate-limiting TCA cycle enzyme Isocitrate dehydrogenase (IDH3A; P50213) was also observed in this set, as were the outer membrane Sorting and Assembly complex (SAM) proteins GRPEL1 (GrpE Like 1; Q9HAV7) and HSPA9 (Heat Shock Protein member 9; P38646), and the antioxidant enzyme PRDX3 (Peroxiredoxin 3; P30048). Of these proteins, IDH3A, HSPA9, and GRPEL1 have been previously investigated in the context of exercise training, and this shall be discussed in the subsequent sections alongside their meta-analysis results. Finally, 16, 37, and 102 proteins were identified as differentially expressed in four, three, and two different datasets, respectively. The remaining 201 proteins were only identified as differentially expressed in only one of the eight datasets.

Figure 3:19 highlights a subset of the common proteins identified as differentially expressed in multiple datasets.

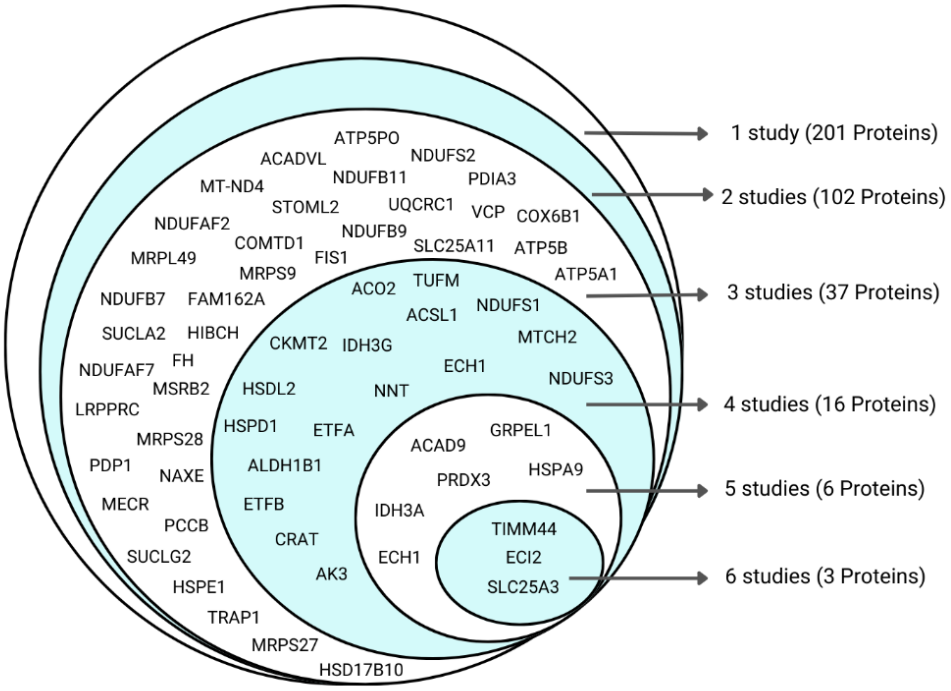


Figure 3:19 A Venn diagram highlighting the common proteins identified as differentially expressed across multiple datasets.

### *3.1.5 Enrichment Analysis of individual datasets*

First, the enrichment of the upregulated proteins in each dataset, shortlisted as described in the Methods, were compared against each other. The Figure 3:20 demonstrates the GO enrichment plot of upregulated proteins from all datasets included in the current analysis. The *Popov_2020* results were excluded from this analysis as very few DEPs were observed. No significant enrichment was observed for any set of downregulated proteins. Expectedly, terms related to the ETC (oxidative phosphorylation and cellular respiration) were observed to be enriched across the greatest number of datasets (5). Terms related to the TCA cycle and fatty-acid oxidation were also consistently observed. As supported by the results of the publication

[81], upregulation of ETC-related proteins was only observed following high-volume training along with a concomitant *deprioritisation* of the TCA cycle enzymes. On the other hand, the enrichment of fatty acid oxidation proteins was observed following both NVT and HVT, which may suggest different effects of training volume on these two pathways.

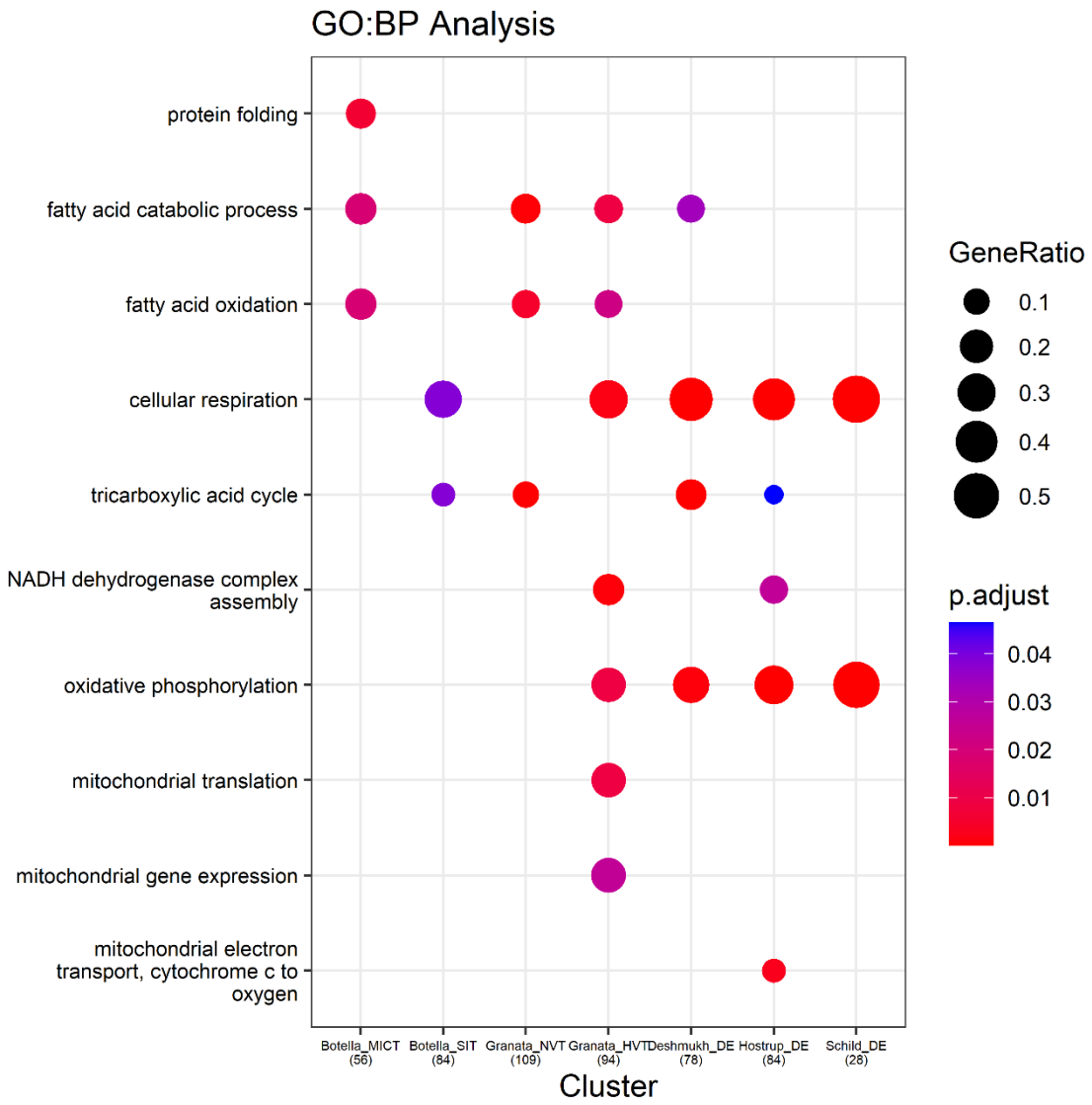


Figure 3:20 Visualising comparative Gene Ontology enrichment profiles for significantly upregulated proteins within each dataset.

Only the *HVT_Granata* dataset demonstrates an enrichment of mitochondrial protein translation and transcription machinery. This may be related to the combination of intensity

and volume of the exercise protocol used and not just volume (i.e,, total exercise exposure), as higher exercise volumes were observed in other datasets (e.g., Deshmukh_2021 and Schild_2015). Within this protocol, a high volume of HIIT exercise was concentrated within a short duration (training twice a day, 7 days a week, for 3 weeks). Existing translation capacity in these previously untrained men may have been insufficient to maintain the required rate of translation/mitochondria biogenesis to meet the rapidly elevating respiratory demand and thus necessitating increases in mitochondria ribosome subunit protein content. With Reactome enrichment analysis, both Schild_2015 and Deshmukh_2021, which have the highest exercise exposures among the datasets, were uniquely enriched for terms related to mitochondria biogenesis and cristae formation, which may signify the dependence of these adaptive responses on exercise volume/exposure (Figure 3:21)

The SIT protocol, despite being low in exercise volume, demonstrated an enrichment of proteins related to respiration. This is supported by previous findings of improved skeletal muscle oxidative capacity (measured as changes in CS activity) and increase in CIV subunit protein content following sprint interval training in healthy young men [282]. Finally, it is unclear why the *Botella_MICT* results do not show any enrichment for OXPHOS-related or TCA cycle-related terms as the exercise volume should not have been the limiting factor (datasets with lower overall exercise volume e.g., Hostrup_2021, NVT_Granata demonstrate enrichment for OXPHOS or TCA cycle-related terms). As *Botella_MICT* demonstrated the

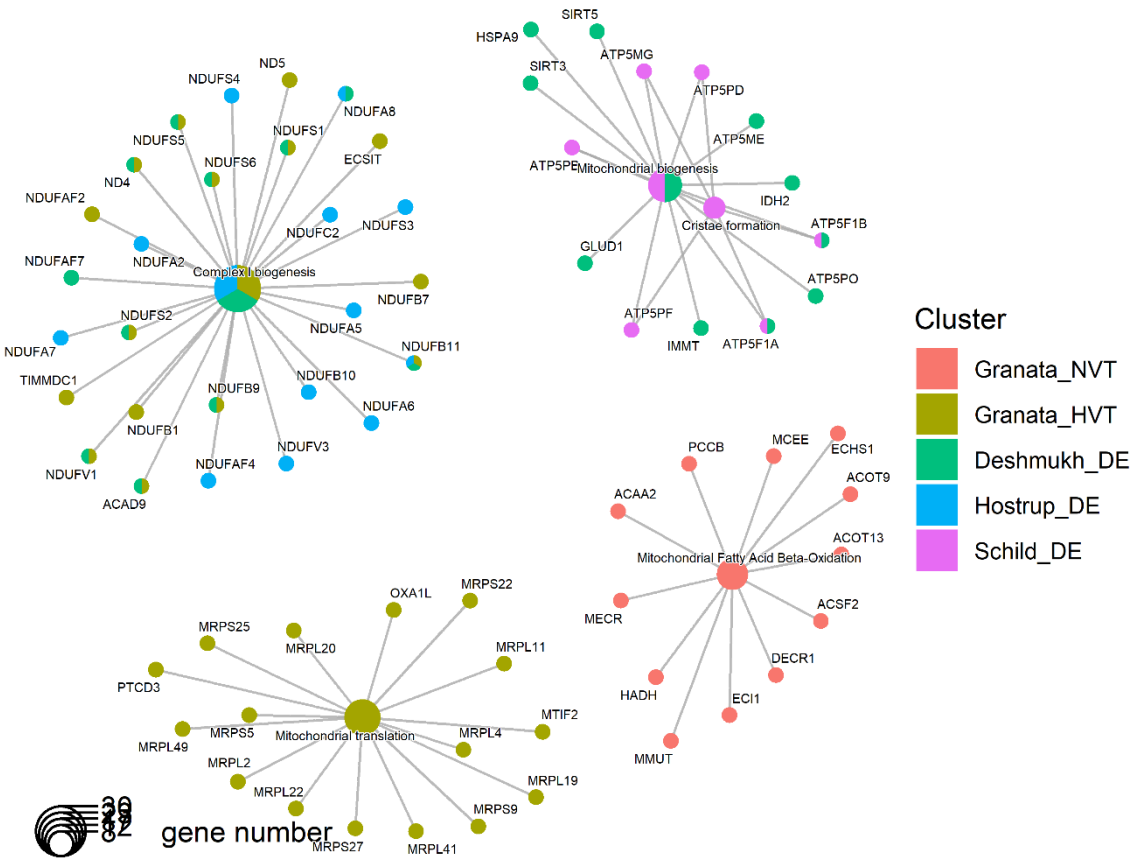


Figure 3:21 Visualising a subset of comparative Reactome enrichment profiles for significant upregulated proteins within each dataset

lowest exercise intensity (approximately 50%$VO_2$max) of all included datasets, it is suspected that this may have blunted the adaptive response within metabolic pathways and components thus contributing to the lack of enrichment.

A Gene Set Enrichment analysis (GSEA) on the individual results revealed similar enrichment profiles as the ontology analysis described above. The major additional finding was that mitochondria translation related pathways were positively enriched in *HVT_Granata,* Deshmukh*_2021*, and *Hostrup_2021 datasets*. This suggests there may be a tendency towards an adaption to increase capacity for mitochondria protein synthesis following exercise training of sufficient volume. However, the changes in relevant proteins may not always be large or consistent enough to be recognised by pathway analysis of significant proteins only.

### 3.2 Effect of Mitochondria Normalisation

As described in the Methods section, an *in silico* normalisation was applied on the mitochondrial subset of the quantified proteome for each dataset in this study demonstrating an increase in overall mitochondria protein content. This section will describe the results of the differential expression analysis performed on mitochondria-normalised data and briefly discuss the exercise-training-induced adaptations to the mitochondria proteome corrected for estimated changes in overall mitochondria content. It is hypothesised that this method will help provide greater resolution to study proteome changes and uncover qualitative adaptations that are prioritised in response to exercise training stimulus.

As can be observed in Table 3:2, normalisation affects the number of proteins detected as differentially expressed within the datasets considered[8]. The result of disentangling changes

[8] The Popov_2020 dataset was excluded from this analysis as it did not return any significant results after applying the normalisation.

in protein abundance from the overall change in mitochondrial protein abundance influenced by exercise training (exercise-induced mitochondria biogenesis response), which was the purpose of the normalisation, reflects mitochondrial proteins that are non-stoichiometrically regulated. It therefore follows that fewer proteins are detected as differentially expressed following normalisation compared to non-normalised results, as most proteins are expected to follow a stoichiometric response.

| Data | Non-Normalised sig up | Non-Normalised sig down | Normalised significant up | Normalised significant down | Normalised total significant | %Sig up after norm |
|---|---|---|---|---|---|---|
| Botella_MICT | 56 | 32 | 48 | 47 | 95 | 85.7 |
| Botella_SIT | 84 | 34 | 71 | 54 | 125 | 84.5 |
| NVT_Granata | 113 | 2 | 67 | 41 | 108 | 59.3 |
| HVT_Granata | 97 | 9 | 29 | 34 | 63 | 29.9 |
| Deshmukh_2021 | 78 | 1 | 31 | 29 | 60 | 39.7 |
| Hostrup_2021 | 84 | 2 | 27 | 23 | 50 | 32.1 |
| Schild_2015 | 28 | 0 | 6 | 5 | 11 | 21.4 |

Table 3:2 Distribution of the number of significant proteins within each dataset before and after the application of mitochondria-normalisation.

Enrichment analyses were performed on the differential expression results of the mitochondria-normalised data to ascertain whether there might be functional significance to these observed non-stoichiometric changes. A de-prioritisation (non-stoichiometric decrease) of ETC complex proteins was observed in the *NVT_Granata* dataset, as evidenced by the Reactome enrichment of *Respiratory Complex* and *Cristae Formation* related terms in the

normalised-downregulated set of proteins; this is consistent with the published findings [81], which identified reduced intensities of ETC complex proteins in the mitochondria-normalised protein results following Normal Volume Training (see Figure 3:22). Terms related to the TCA cycle were enriched for the set of normalised-upregulated proteins, further supporting the conclusions that early adaptions in response to exercise training may be focused on increasing electron flow to the OXPHOS complexes instead of increasing overall OXPHOS capacity [81]. The *NVT_Granata* dataset, however, is the only data within this study that examined adaptations following only two weeks of training and thus it was not possible to validate this finding with other data and strengthen these conclusions. A similar analysis of the *Hostrup_2021* dataset [222] was performed and Reactome terms related to the TCA cycle were only found to be enriched for the non-normalised data, whereas the OXPHOS related terms were enriched for the normalised-upregulated protein set (Figure 3:23). Alongside these, terms related to 'Protein import' and 'Protein localisation' were also enriched exclusively in the normalised-upregulated set of proteins. The *HVT_Granata* dataset also demonstrated enrichment of the 'Protein Import' term for the normalised-upregulated protein set. However, no other non-stoichiometric adaptations in the TCA, OXPHOS, or other functional pathways were observed. Finally, analysis of normalised data from the *Deshmukh_2021* [19] (see Figure

3:24) yielded no significant enrichment and not enough proteins were significant in the normalised results from *Schild_2015* to perform any meaningful enrichment analysis.

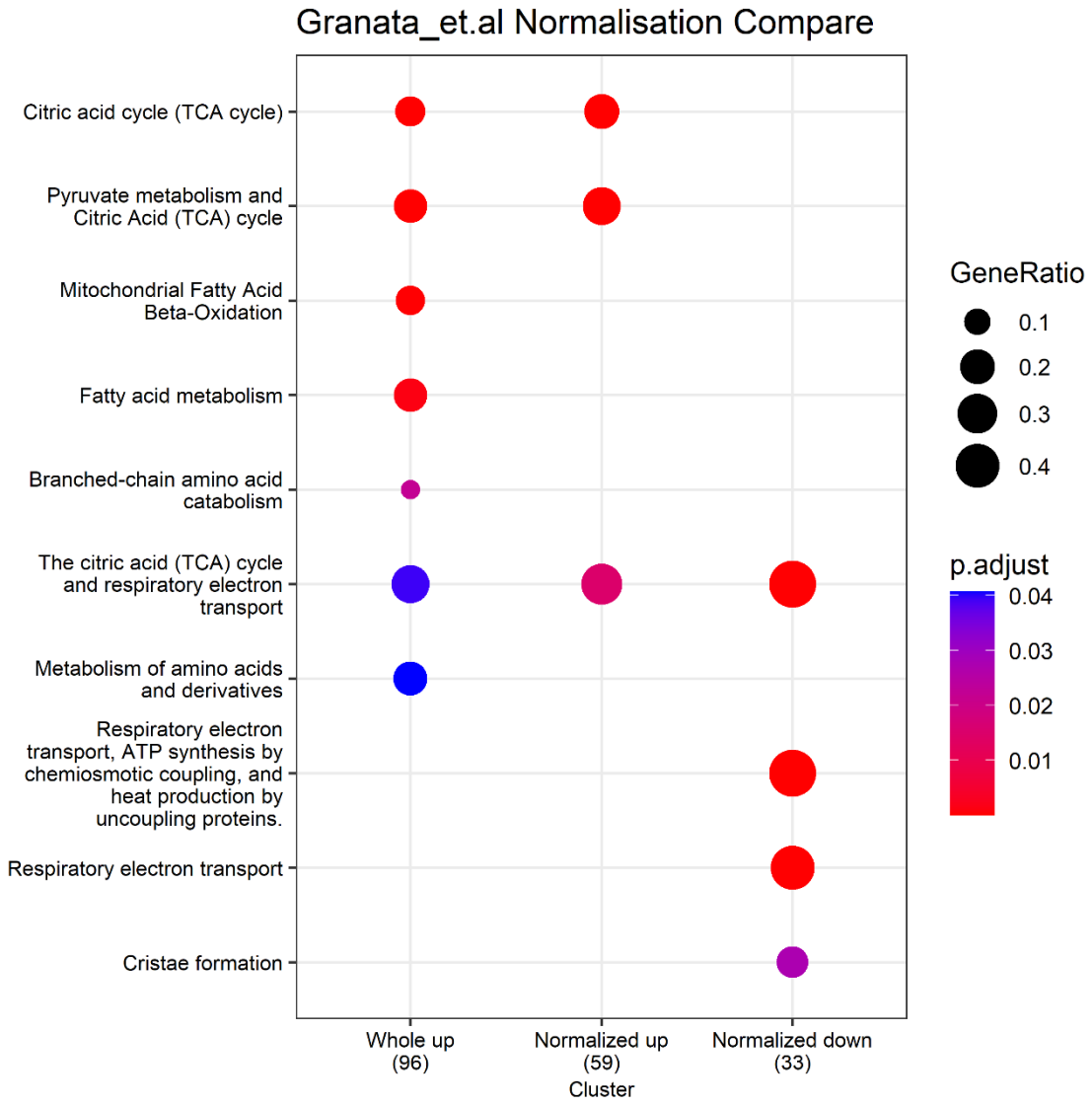


Figure 3:22 Comparing Gene Ontology enrichment profiles between non mitochondria normalised (whole) and mitochondria normalised results for *NVT_Granata*

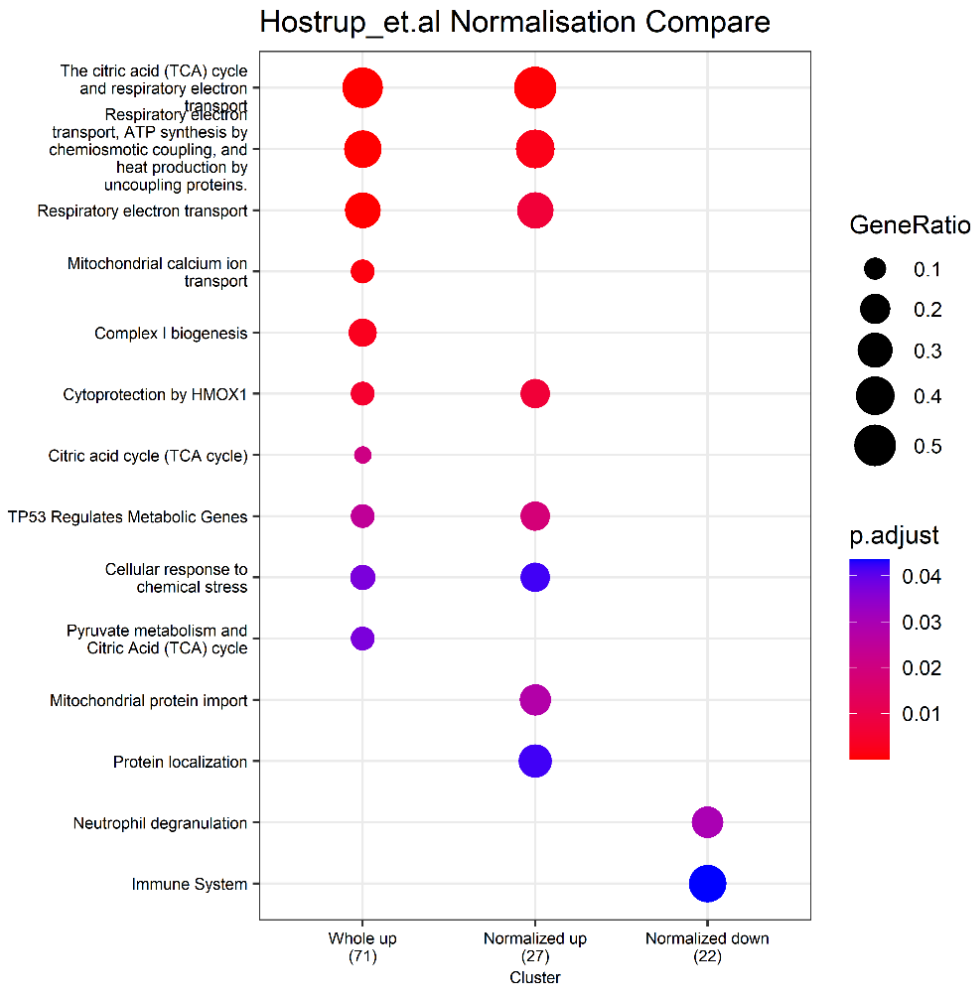


Figure 3:23 Comparing GO enrichment profiles between non-mitochondria normalised (whole) and mitochondria normalised results for *Hostrup_2021*

Based on this exploratory analysis there appears to be a trend in the time-course of non-stoichiometric proteome changes, with different functional protein adaptations prioritised at varying levels of exercise exposure. As observed with *Granata_NVT*, following two weeks of (Normal Volume) training, adaptations are focused on increasing TCA cycle capacity, whereas OXPHOS proteins are largely unchanged and are even proportionally lower than pre-exercise levels (when expressed relative to the overall increase in mitochondrial proteins). The *Hostrup_2021* dataset provides a representation of adaptations after five weeks of high-intensity interval training where a significant increase in non-stoichiometric abundance is observed for OXPHOS subunits. Coupled with the significant increases in import-related

proteins, which was also observed with *HVT_Granata,* this suggests that the organelle is adapting to build new functional units and expand the mitochondria reticulum (biogenesis) during this phase of the training-induced response. After prolonged exposure to exercise, such as with 12 weeks of continuous endurance training (as represented in *Deshmukh_2021)* or with over five years of endurance training (as represented in *Schild_2015*), overall mitochondria content is observed to increase to meet the physiological demand for adaptations and the non-stoichiometry in functional pathways apparent in the early exercise response is no longer apparent.

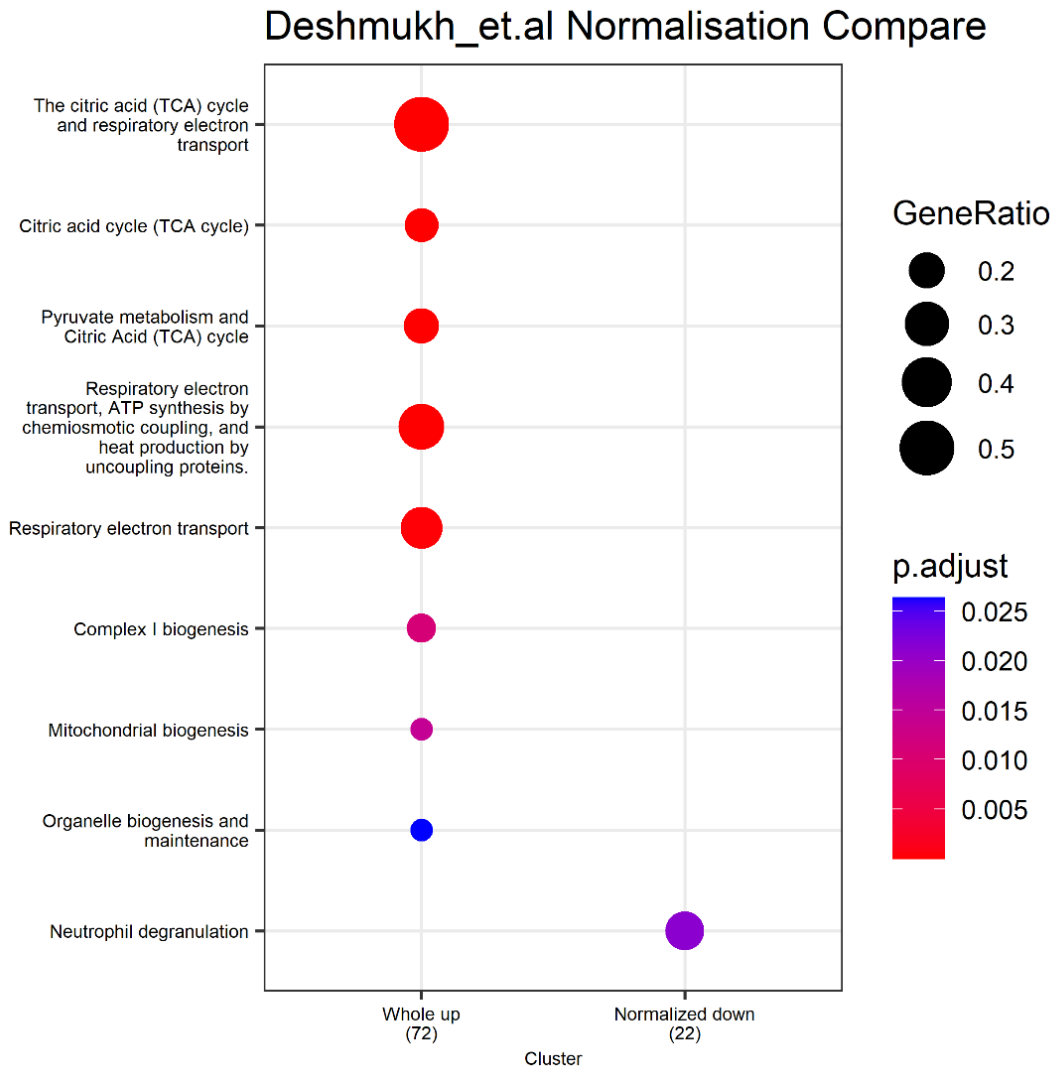


Figure 3:24 Comparing Gene Ontology enrichment profiles between non-mitochondria normalised (whole) and mitochondria normalised results for *Deshmukh_2021*

### 3.3 Mitochondria Protein Enrichment Analysis

Changes in the total mitochondrial protein abundance following exercise was analysed by comparing imputed raw protein intensities in the Pre and Post conditions for each dataset included in this meta-analysis. Mitochondria Protein Enrichment (MPE) was also calculated as the mitochondrial fraction of total protein intensities for each sample within each dataset. One-sided unpaired t-tests for total mitochondrial protein intensities and MPE was performed for each dataset to approximate regulation of mitochondria biogenesis.

Significantly higher total mitochondrial protein intensities post exercise were observed in the NVT_Granata ($p<0.01$), HVT_Granata ($p<0.001$), Deshmukh_2021($p<0.05$), and Schild_2015 ($p<0.05$) datasets. Similarly, significantly higher post-exercise MPE values were observed in the NVT_Granata ($p<0.01$), HVT_Granata ($p<0.01$), Deshmukh_2021 ($p<0.01$), and Schild_2015 ($p<0.05$) data. There was a trend for increased MPE post-exercise in the Botella_MICT ($p=0.1$), Hostrup_2021 ($p=0.09$), and Popov_2020 ($p=0.08$) datasets; however, these differences did not meet the significance threshold.

According to the published results of *Hostrup_2021,* a significant increase in summed overall mitochondrial protein intensities was observed ($p<0.01$) along with a significant increase in CS activity, thus supporting the claim for increased mitochondria content [222]. It is unclear why this finding was not reproduced in this study, but it is hypothesised that it may be due to differences in filtration and imputation of missing values and how mitochondria proteins were annotated. The Botella_SIT protocol had one of the lowest exercise volumes of all the included studies, which may help explain the absence of an overall increase in MPE or total mitochondria protein intensities (consistent with the absence of a significant increase in overall mitochondria protein content in either fibre type along with unchanged CS activity following SIT; [20]). Increases in markers of mitochondria content (CS activity, CIV subunit

protein content, and PGC-1α protein content), however, have previously been observed following 6 weeks of SIT [282]. Further investigations may be necessary to develop a consensus around the regulation of mitochondria content with SIT, using both wet-lab and *in silico* methods of protein quantification. The Botella_MICT data, on the other hand, demonstrated evidence for increased overall mitochondria content in both fibre-types and increased CS activity following MICT, according to the publication [20]. The discrepancy between the observed and published results may be due to the differences in samples (whole-muscle lysate versus single-fibre) and the method of data acquisition (DIA v. DDA). The lack of change in overall mitochondria protein content in Popov_2020 may be attributed to the technical limitations of the data (strength of batch effects) and the relatively shallow proteome coverage in the processed data. Changes in biomarkers of mitochondria content were not discussed in the associated publication and, therefore, can't be compared with the *in silico* findings. However, it is likely that there was an underlying increase in mitochondria content as the exercise volume (8 weeks of alternating HIIT and continuous endurance training) should not be limiting when compared to the other studies (see, for example, [81], [289]).

## 3.4 Meta-analysis Results:

### *3.4.1 Overview of Meta-analysis results:*

All identified proteins (as described in Section 3.1) were included in the meta-analysis. The design and parameters of the meta-analysis design have been described in detail in the Methods section.

A total of 253 of the 778 proteins were found to be differentially expressed at the FDR threshold of less than 0.05 (see Figure 3:27). A total of 493 proteins were quantified in five or more studies (K ≥ 5), of which 200 were significant at the FDR threshold of less than 0.05. A total of 113 proteins were quantified in all eight studies (K=8), of which 80 were significant at the FDR threshold of less than 0.05 (Figure 3:25).

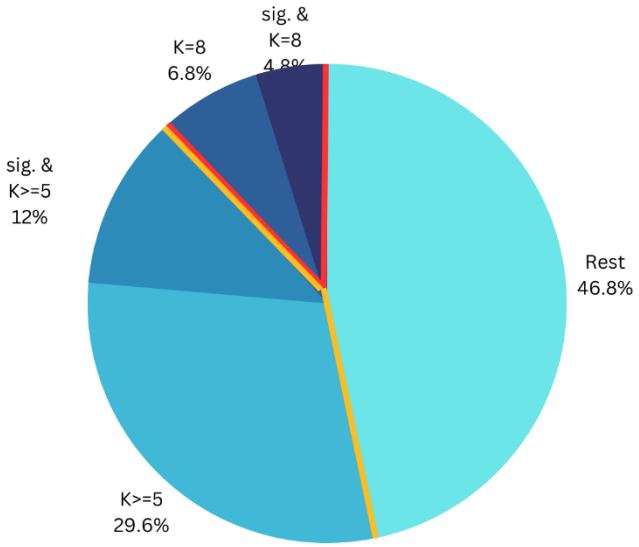


Figure 3:25 The distribution (percentage of entire quantified proteome) of meta-analysis results according to degree of identification (*K* refers to the number of datasets the protein was quantified in) and significance (*sig.* denotes that the protein is significant at the meta-analysis level)*Rest* refers to the remaining proportion of the quantified proteome.

The average absolute $\log_2$ fold change of proteins across the included datasets was found to be approximately 0.25, translating to a 1.2x fold change. When only considering significantly differentially expressed proteins, the average absolute $\log_2$ fold change was found to be 0.35, translating to a 1.27x fold change. As a stringent data filtration threshold and conservative imputation technique were used within the data analysis pipelines, this may have contributed to the relatively small overall fold changes compared to some of the results from existing proteomics literature (See for example Deshmukh et.al, 2021[19]; Hostrup et.al, 2021[222]).The distribution of $\log_2$ fold changes of proteins with $K \geq 5$ is shown in Figure 3:26. Of the 274 significant mitochondrial DEPs, 228 demonstrated a positive aggregate $\log_2$ fold change and 25 demonstrated a negative aggregate $\log_2$ fold change (See Figure 3:27).

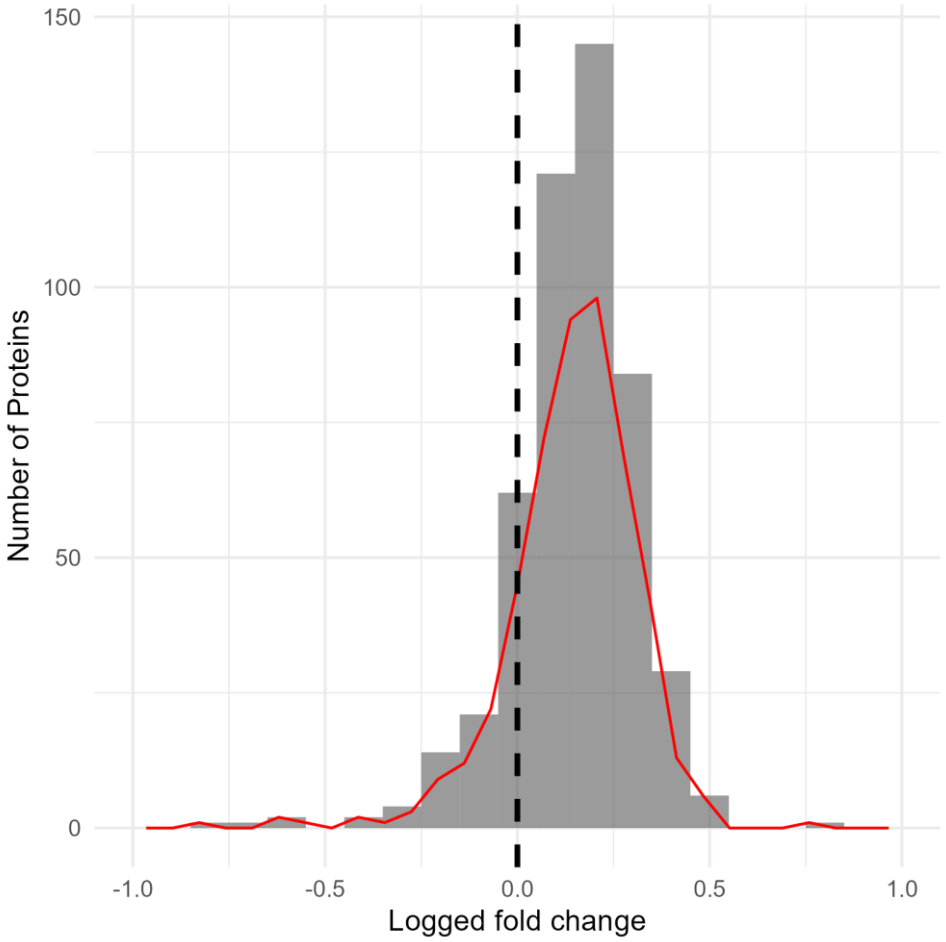


Figure 3:26 Histogram and distribution of aggregate log fold changes of proteins with $K \geq 5$.

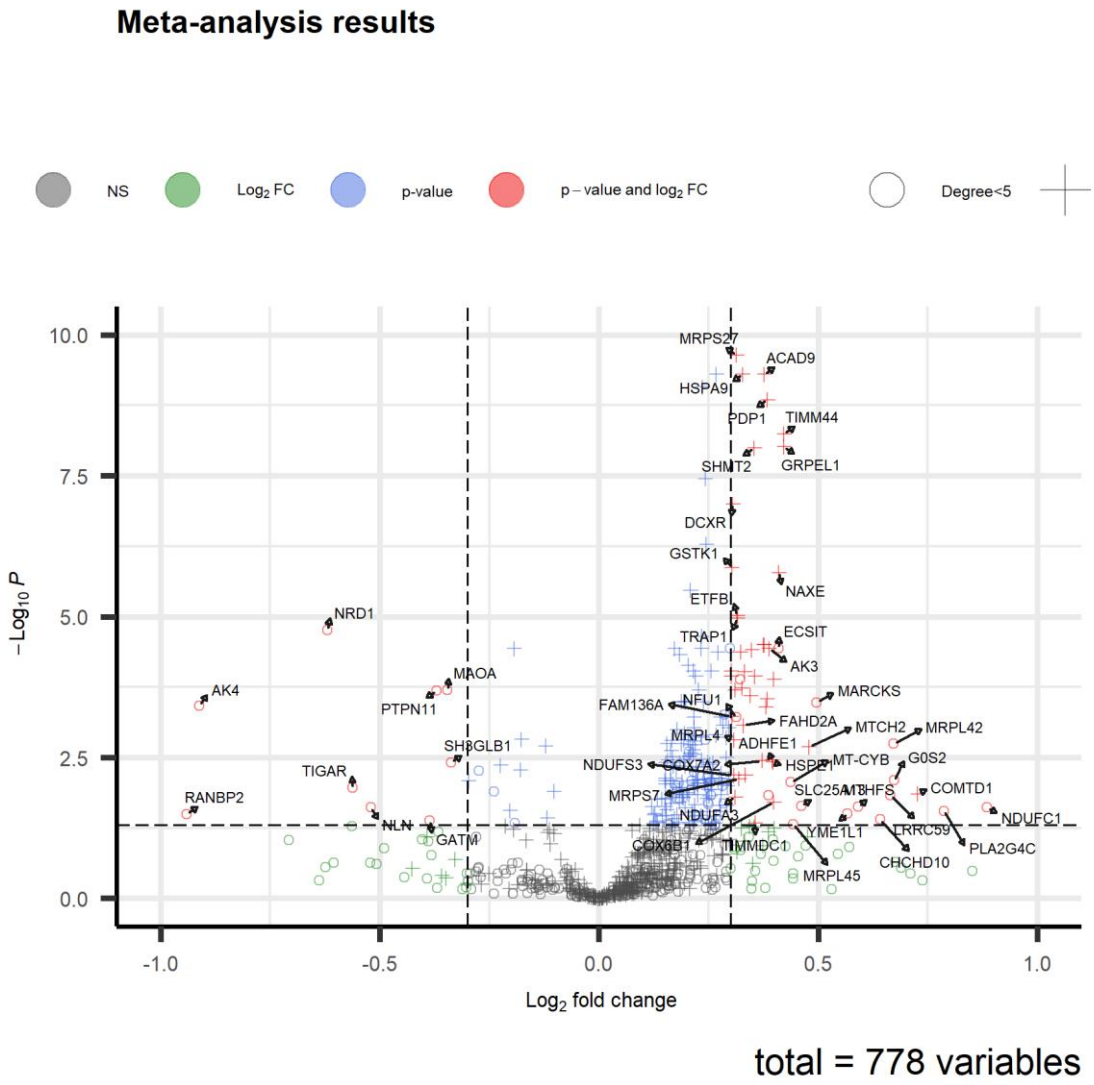


Figure 3:27 Volcano plot of meta-analysis results. Points denoted as + signify K ≥ 5. The horizontal cut-off denoted the FDR threshold of 0.05. The vertical cutoffs denote log2 fold change of 0.3.

Measures of heterogeneity found in the aggregate fold change observed for each protein were reported as $I^2$ percentage values. The mean $I^2$ percentage value for proteins identified in five or more datasets was approximately 44%, suggesting an overall low-to-medium level of heterogeneity in the meta-analysis results. With an increasing number of studies in which proteins were identified, the distribution of heterogeneity scores shifted to the right (i.e., was

higher), as observed in Figure 3:28. This suggests there is inherent variation in mitochondrial protein responses to exercise training interventions, which depends on a variety of study-specific factors, and this becomes more apparent to the meta-analysis model when more studies are inputted into the model. The heterogeneity score for any protein also appears to be affected by uniformity in the degree and direction of its response to the exercise interventions. This shall be discussed in more detail for a subset of the results in the following section.

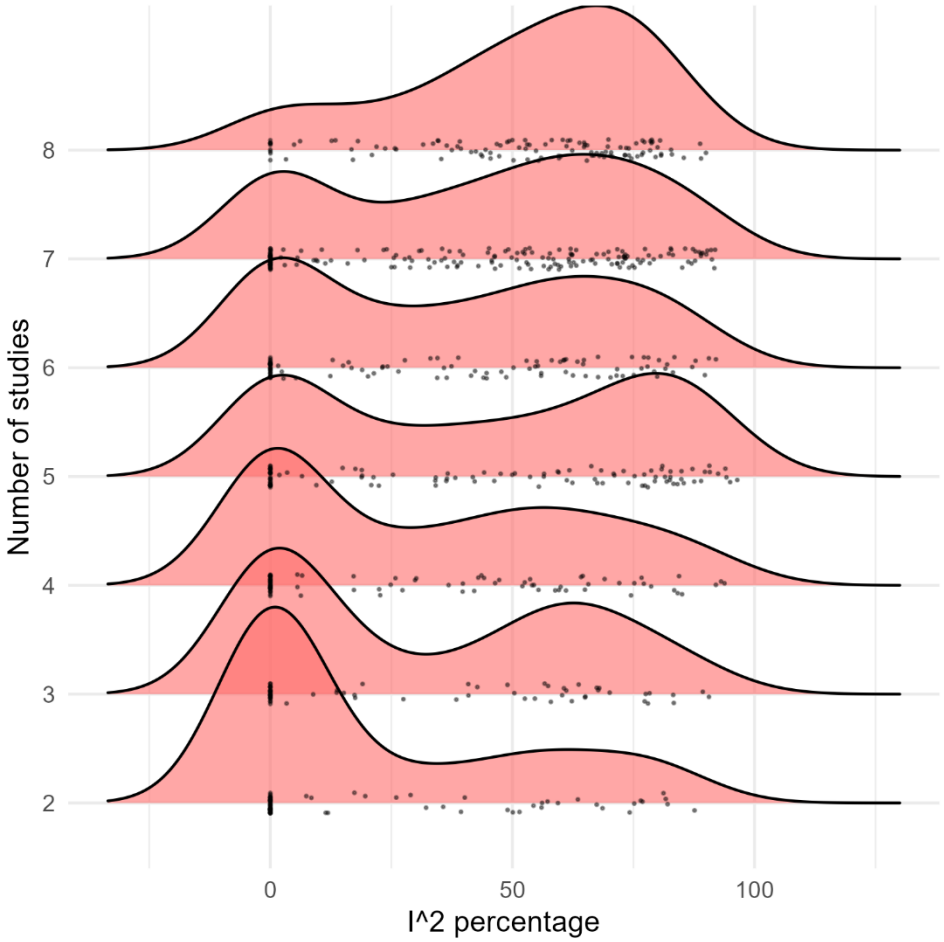


Figure 3:28 Distribution of heterogeneity scores for proteins stratified by the number of studies in which the proteins were identified. Each point is a single protein. The x-axis denotes the $I^2$percentage of the meta-analysis result of the protein and the y-axis denotes the number of studies in which the protein was quantified.

### *3.4.2 Analysis of heterogeneity*

Given the distribution of heterogeneity values displayed in Figure 3:29, which describes the distribution of $I^2$ values for proteins identified in five or more studies ($K \geq 5$), we can observe that many results are estimated to have extremely low heterogeneity scores. Approximately 60 proteins were estimated to have a heterogeneity score of zero and over 100 proteins demonstrated a heterogeneity score under 5%. Thresholds of 25%, 50%, and 75% have generally been used to categorise low, medium, and high levels of heterogeneity in meta-analyses [6]. With these meta-analysis results, however, different thresholds were chosen due to the right skew of the distribution of $I^2$ percentage values. Low heterogeneity proteins were classified as demonstrating $I^2 < 10\%$, which accounted for approximately 20% of the proteins considered. Similarly, high heterogeneity proteins were classified as demonstrating $I^2 > 70\%$, which is roughly the 75th percentile of the $I^2$ percentage distribution.

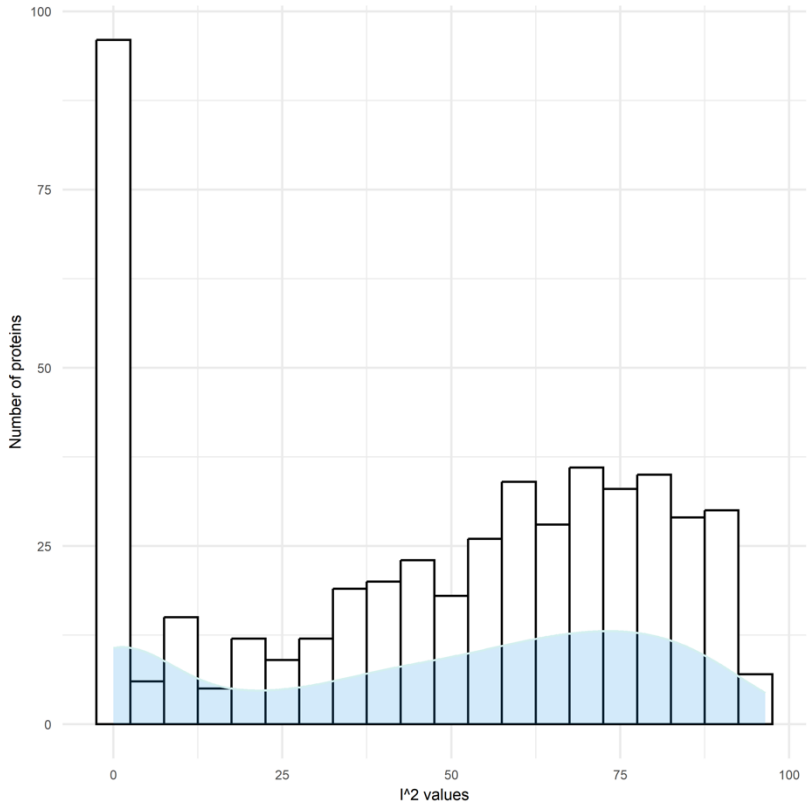


Figure 3:29 Histogram and density plot of $I^2$ % values for proteins with $K \geq 5$.

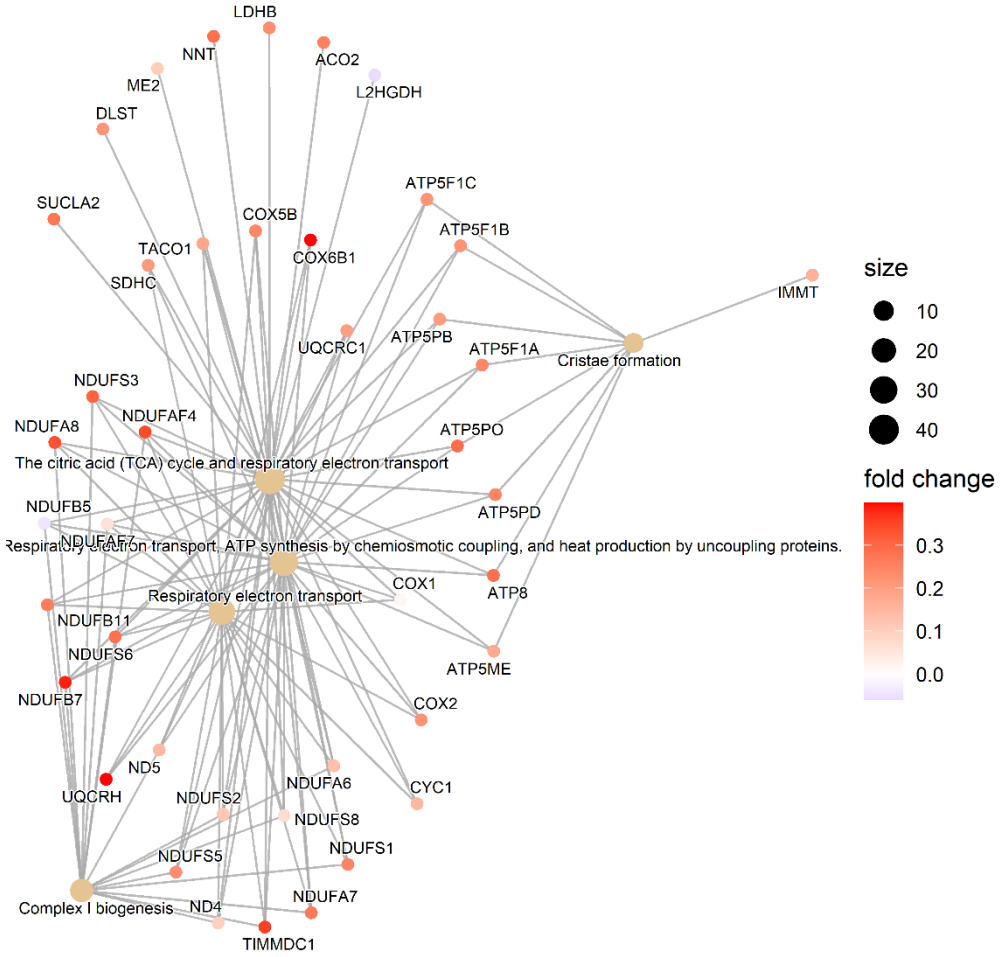


Figure 3:30 Visualisation of a subset of Reactome enrichment results for high heterogeneity set of proteins. The fold change included is the aggregate fold change in the meta-analysis results.

A Reactome and GO enrichment analysis was performed on the subsets of proteins with a K>=5 and with high and low heterogeneity scores. The high heterogeneity set of proteins demonstrated an enrichment of terms related to OXPHOS activity, Complex I assembly, and respirasome formation (Figure 3:30). On the other hand, the low heterogeneity set of proteins (Figure 3:31) demonstrated an enrichment of terms related to mitochondria membrane organisation and protein import. These differences in enriched terms might suggest that exercise training variables (study-specific characteristics) may have a stronger effect on the regulation of ETC complex proteins compared to mitochondria protein import and assembly

mechanisms. This finding is in keeping with the findings of dependence of mitochondria respiratory capacity adaptations on exercise training volume [120]. However, caution must be applied when making inferences about the extent of heterogeneity from $I^2$ percentage values, especially when confidence intervals around the $I^2$ statistic are wide and the number of studies is relatively low [6] [290].

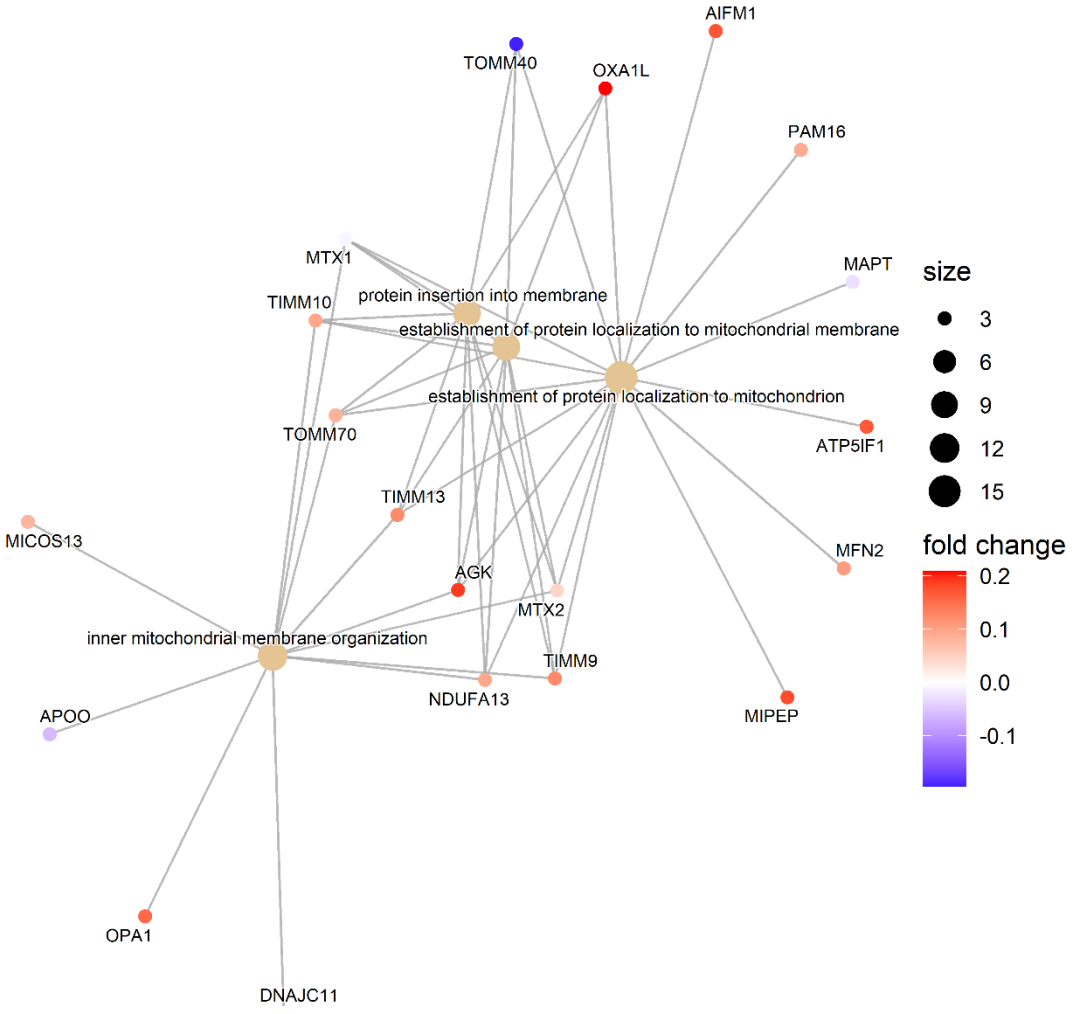


Figure 3:31 Visualisation of a subset of Reactome enrichment results for low heterogeneity set of proteins. The fold change included is the aggregate fold change in the meta-analysis results.

The plot in Figure 3:33 shows the forest plot for the gene CrAT (Carnitine Acyltransferase; P43155). According to its $I^2$ percentage value (0), this protein demonstrates extremely low variance. As can be observed, the spread of the summary effect is narrow, whereas the study-level variances are comparably wider; therefore, the model attributes all the variance in the observed effect to sampling/random error. The opposite is the case for the protein visualised in Figure 3:32 (MTCH2, $I^2 = 87\%$). The variance in the observed effect is higher than most of the study-level variances. Therefore, the model attributes a much greater proportion of the observed variance to heterogeneity in effect rather than sampling/random error. This interpretation is also supported by the mathematical design of the $I^2$ statistic, which is monotonically dependent upon the Q statistic (i.e., the higher the Q statistic the larger the $I^2$ percentage) [291, 292]. The Q statistic is a measure of the dispersion of each study's effect size relative to the meta-analysis effect size weighted by the variance of the study [291]. In other words, it is dependent on the relative degree of intra-study variance, interpreted by the model as a measure of sampling error, and variance of the effect-size, interpreted as a measure of true heterogeneity. This is also demonstrated in the visual illustrations.

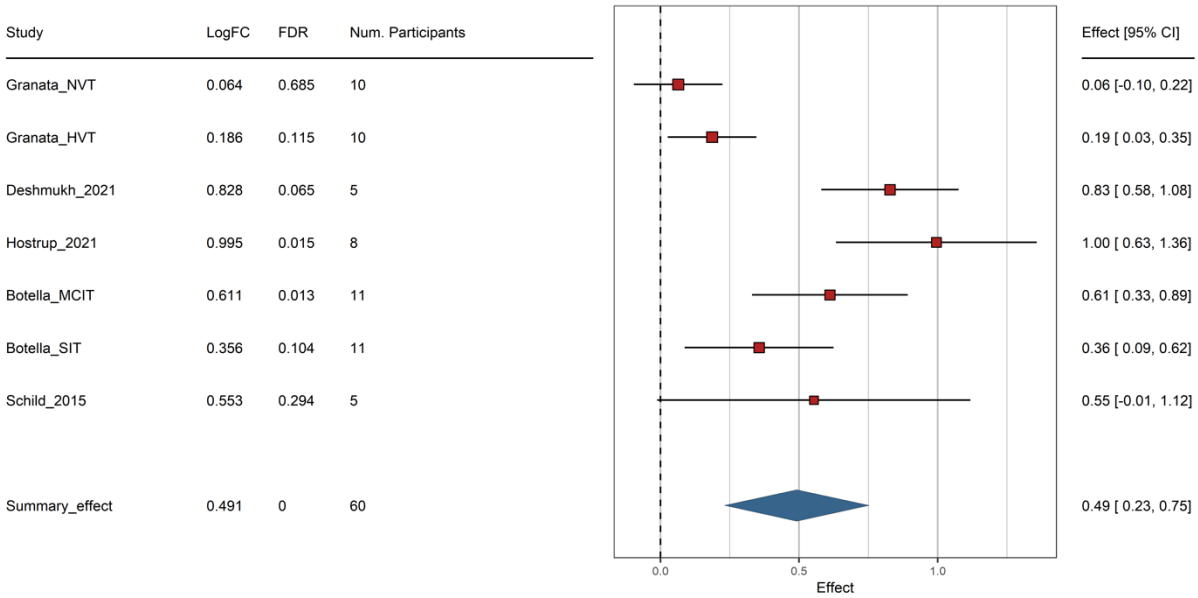


Figure 3:32 Forest plot for the MTCH2 protein demonstrating high proportion of heterogeneity in the overall effect size.

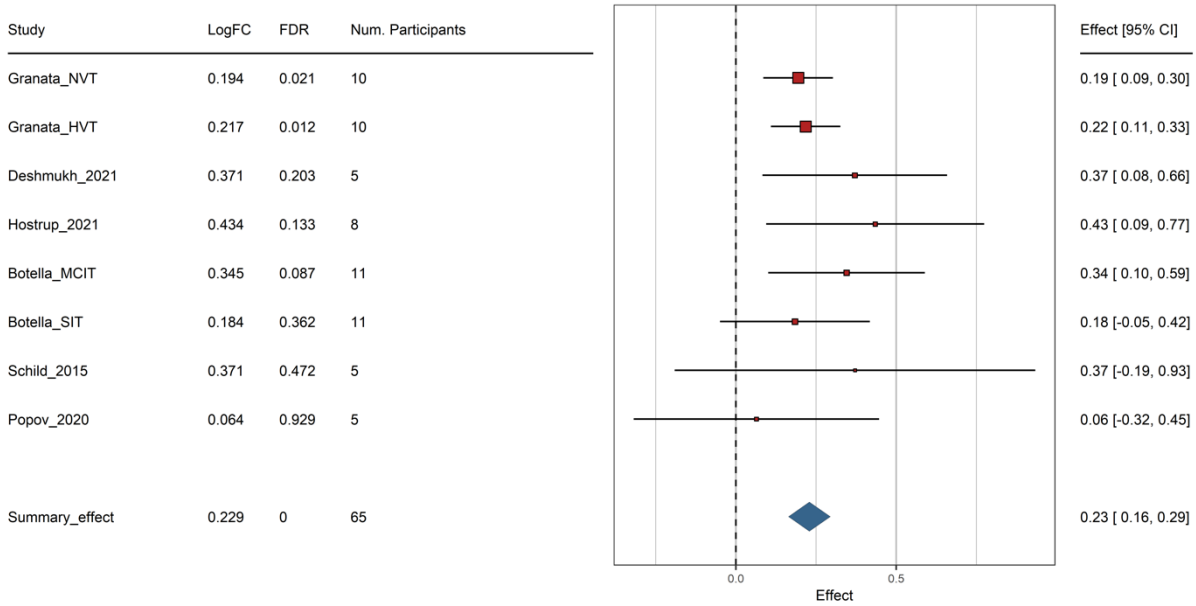


Figure 3:33 Forest plot for the CrAT protein demonstrating low proportion of heterogeneity in the overall effect size.

The Figures 3:34 and 3:35 demonstrate the behaviour of a subset of the 'high' and 'low' heterogeneity set of proteins, respectively, against each study in which they were identified. The x-axis is the studies ranked ordinally in ascending order of total exercise volume. There appears to be a trend for the fold changes of the high heterogeneity protein set to be associated with total training volume. The fold changes tend to be larger with greater training volume (see especially Deshmukh_2021 and Schild_2015). Similarly, the low heterogeneity protein set appears to demonstrate a more consistent and uniform relationship with increasing training volume/across different exercise-training characteristics.

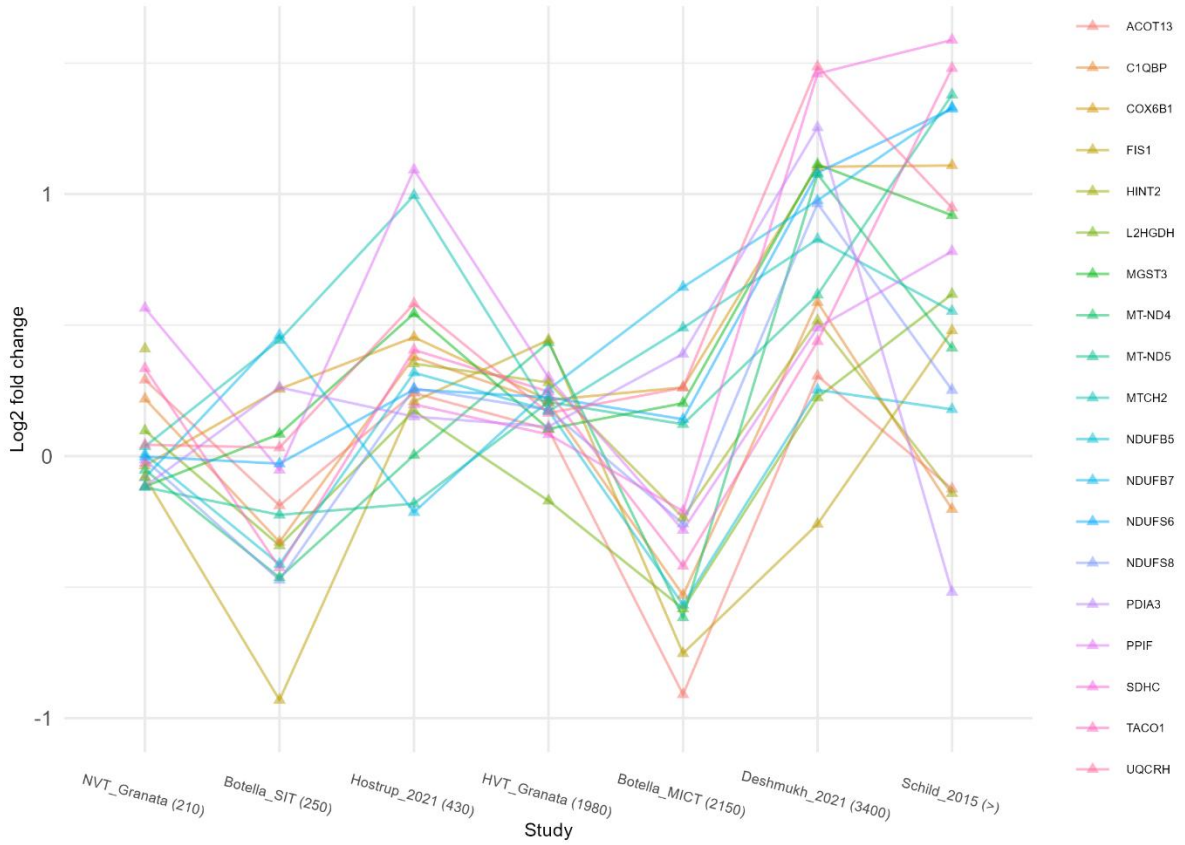


Figure 3:34 Plot demonstrating average log fold change of a subset of high heterogeneity proteins across the different studies. The X-axis is ordinally ranked and the numbers denote approximations of exercise volume calculated as described in the Methods section. * > denotes exercise volume higher than all others. The Y-axis limits between this plot and 3:34 have been kept the same to highlight the difference between the plots.

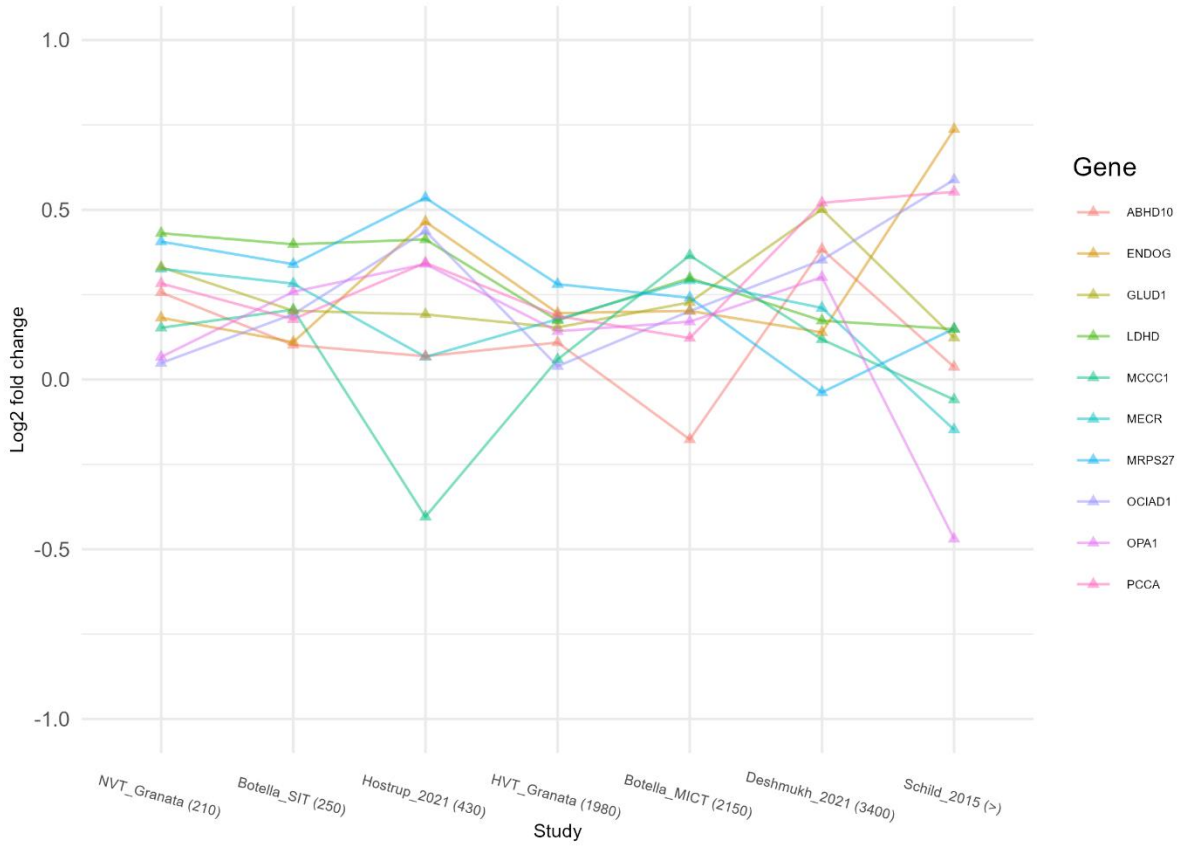


Figure 3:35 Plot demonstrating average log fold change of a subset of high heterogeneity proteins across the different studies. The X-axis is ordinally ranked and the numbers denote approximations of exercise volume calculated as described in the Methods section. * > denotes exercise volume higher than all others. The Y-axis limits between this plot and 3:34 have been kept the same to highlight the difference between the plots.

It has been demonstrated that $I^2$% statistics can be significantly biased toward underestimation by almost 30% in conditions of high true heterogeneity - especially in small meta-analyses designs [290]. Further, the confidence intervals of estimated $I^2$ % values are often so wide that they cross into thresholds of other heterogeneity 'categories', leaving researchers and readers prone to spurious conclusions [6]. The Figure 3:36 demonstrates an example of the spread of confidence intervals of $I^2$ % values from Cochrane and genetic risk-factor meta-analyses adapted from [6] (left) and the spread of confidence intervals of $I^2$ values in the results of the current meta-analysis (right). As can be observed, the trend is in keeping with similar observations of existing meta-analyses. The upper confidence bound for several proteins with an $I^2$ % point estimate of zero is greater than 80% and, therefore, no concrete

inferences can be made about the absence or low degree of heterogeneity within these observations. On the other hand, proteins with a high $I^2$ % point estimate demonstrate narrow confidence intervals and generally have a significant Q statistic after adjusting for multiple hypothesis correction.

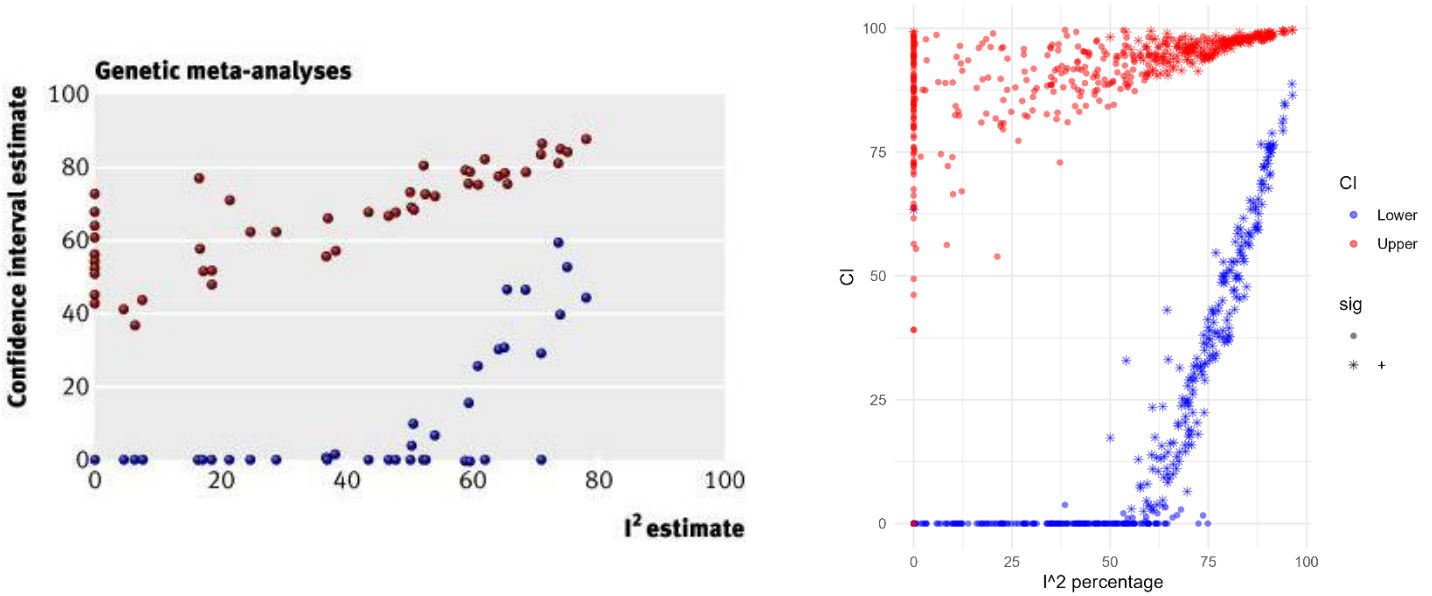


Figure 3:36 Comparison of $I^2$ values against the confidence interval associated with those heterogeneity scores. The plot on the left demonstrates the spread of heterogeneity scores and confidence intervals from a published meta-analysis [6] and the plot on the right are the results from the current study where each point represents the upper/lower CI bound for a particular protein. The significance label denotes the significance of the Q statistic associated with each $I^2$ value.

In summary, stronger claims can be made about presence of inter-study heterogeneity in effect-size for a certain subset of proteins than about the absence of heterogeneity in seemingly homogenous results. However, despite the trend in fold changes observed in a subset of 'high' heterogeneity proteins (Figure 3:34), no conclusive assertions can be made about exercise training volume/intensity-dependant effects due to the limitations in the accurate measurement of heterogeneity as discussed above and the current low number of studies in this meta-analysis. The current exploratory observations are nonetheless intriguing, and they should

be further investigated in the future if these trends persist and/or stronger patterns are observed with the addition of more data.

### *3.4.3 Meta-analysis of biological results:*

Of the significant proteins identified, the protein COMTD1 (Catechol-o-Methyltransferase; Q86VU5; K=5)[9] demonstrated the highest degree of aggregate fold change (0.72 ± 0.24) (Figure 3:37). According to MitoCarta 3.0 [1], this protein is localised to the mitochondria membrane and is predicted to be involved in catecholamine metabolism. This protein has previously been identified in the context of skeletal muscle mitochondria dysfunction but has not been discussed in the context of exercise response in humans [293]. The next highest $\log_2$ fold changes were observed for the proteins MTCH2 (Q9Y6C9; K=7) (0.48 ± 0.13) and TIMM44 (O43615; K=7) (0.42 ± 0.06). The MTCH2 (Mitochondrial Carrier 2) protein is localised to the mitochondria membrane and is a key facilitator of pro-apoptotic signalling and thus plays a key role in the maintenance of mitochondrial dynamics [294]. The TIMM44 protein is integral to the inner membrane translocase mechanism that allows causes the movement of peptides from the IMM into the mitochondrial matrix. Other notable proteins include AK-3 (Adenylate kinase-3; Q9UIJ7; K=7) (0.39 ± 0.08), integral to the replenishment of adenine nucleotides within the matrix, GRPEL1 (GrpE Like 1; Q9HAV7; K=6) (0.42 ± 0.06), another essential component of the inner membrane protein import complex, and NAXE

[9] Results of the meta-analysis are reported in the following format: Gene Symbol (Gene Name; Uniprot ID; *K* value) (aggregate $\log_2$ fold change ± SEM). SEM stands for the standard error of the mean. If Gene Names and Uniprot IDs have been noted before, only the Gene Symbol is used. *K* values are only reported sometimes and mostly in cases where K<5.

(NADPH Epimerase; Q8NCW5; K=7) (0.41 ± 0.07), a quality control mechanism that catalyses the conversion of NAD(P)H back from the hydrated form NAD(P)HX.

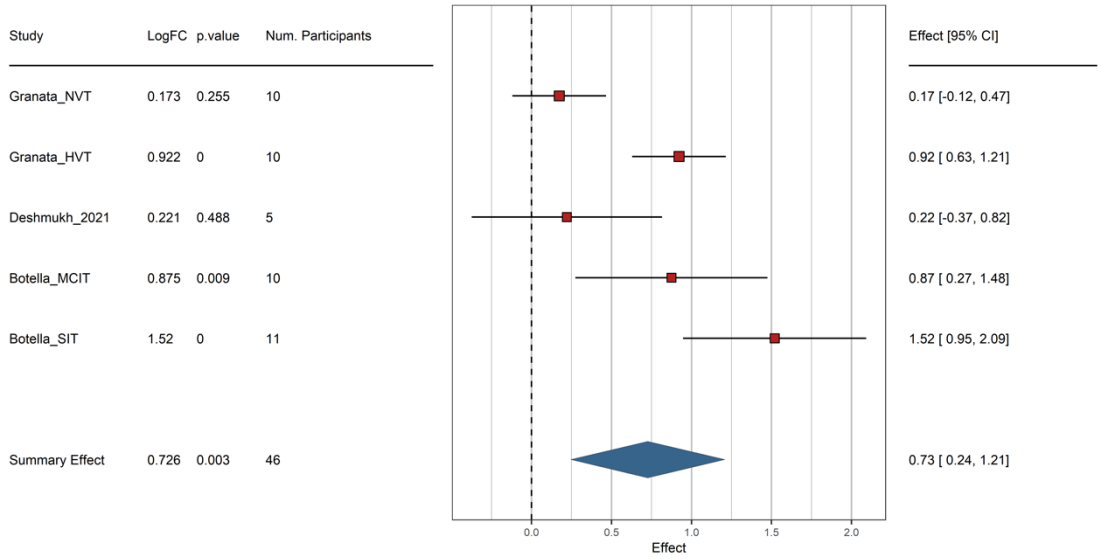


Figure 3:37 Forest plot for the meta-analysis results of the COMTD1 protein.

Of the identified proteins, the lowest p values were demonstrated for ECI2 (K=7; q[10] < 1e-12) (0.26 ± 0.03), CRAT (K=8; q < 1e-10) (0.22 ± 0.03), HSD17B10 (Hydroxysteroid 17-Beta Dehydrogenase; Q99714; K=8) (0.27 ± 0.04; q < 1e-10), and ACAD9 (0.38 ± 0.05; q < 1e-10; K=6), which are all implicated in the fatty-acid oxidation pathway. Other notable proteins include HSPA9 (mtHSP70) (0.33 ± 0.05; q < 1e-10; K=8) (Figure 3:38) - an essential component for matrix-localised protein import, DCXR (Dicarbonyl xylulose reductase; Q7Z4W1; K=7; q < 1e-7) (0.31 ± 0.05) - an NADPH-dependant oxidoreductase of sugars that has been shown to be downregulated in conditions of T2D [130]. Another key protein within this criterion was GLUD1 (Glutamate Dehydrogenase 1; P00367; K=7) (0.25 ± 0.04; q < 1e-08), which is activated by SIRT3, an exercise-responsive deacetylase that is implicated in ATP-synthesis and gluconeogenesis through amino acid metabolism [295] [296]. Further, GLUD1

[10] q values refer to the Benjamini-Hochberg adjusted p-values of the meta-analysis results

may also have a role in mediating mTORC1 (mechanistic Target of Rapamycin) activation, a critical regulator of muscle anabolism, through its effects on amino acid oxidation [297] [298]. Finally, the CS enzyme protein (O75390; K=8; q <1 e-10) (Figure 3:39), a standard biomarker for mitochondria content [174], which has been extensively shown to be regulated with exercise, also demonstrated highly consistent evidence for regulation within the data.

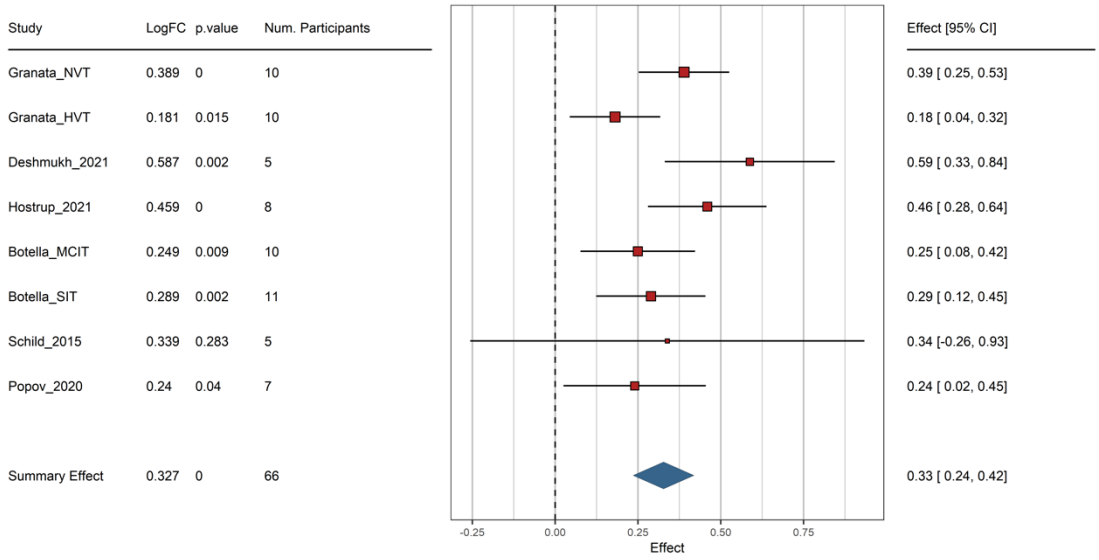


Figure 3:38 Forest plot for the meta-analysis results of the HSPA9 protein.

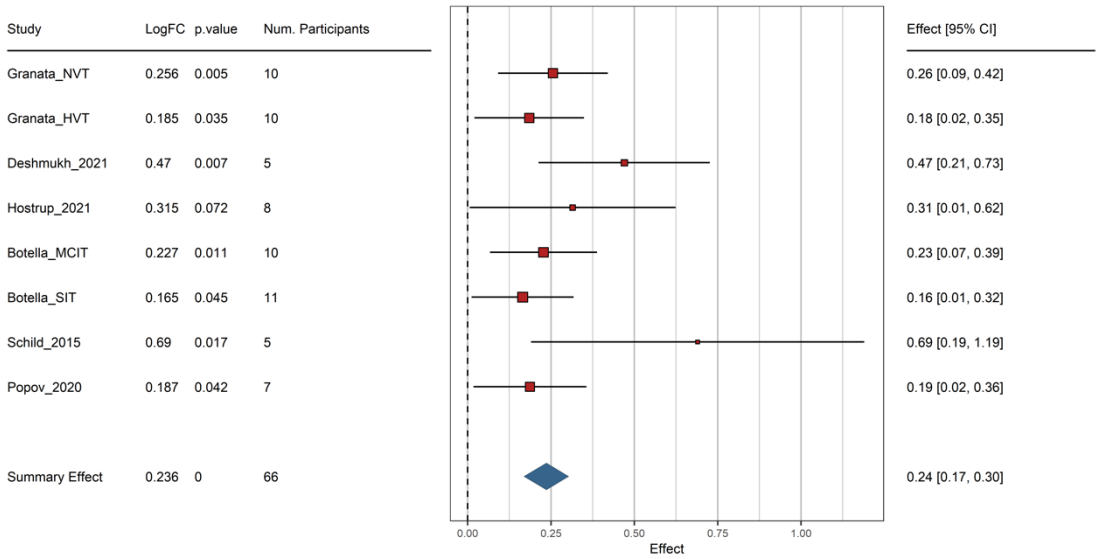


Figure 3:39 Forest plot for the meta-analysis results of the CS protein.

The following sections will describe the results and findings of the meta-analysis with respect to key metabolic and functional mitochondrial pathways shortlisted using MitoCarta 3.0 [1].

### 3.4.3.1 Regulation of fatty acid oxidation proteins

Of the 44 annotated proteins within the fatty acid oxidation pathway, 29 were identified at least one of the datasets. Of these, 23 were found to be differentially expressed and upregulated in the meta-analysis results.

The first mitochondria-related step in the fatty acid oxidation pathway is the import of activated fatty acids (fatty acyl-co-A) across the outer membrane into the mitochondria matrix [299]. This step is regulated by the Carnitine palmitoyltransferase (CPT) enzymes, CPTIB (the prevalent isoform in human muscle) and CPTII, both of which were quantified multiple times

in the meta-analysis dataset. CPTIB (Q92523), which regulates the rate-limiting step of formation of acyl-carnitine and thus fatty acid uptake into the mitochondria, was found to trend toward significance (0.25 ± 0.11; $q < 0.06$) and CPTII (P23786) (0.15 ± 0.04), which catalyses the formation of fatty acyl-CoA back from acyl-carnitine, was significantly increased as well.

The first step in the beta-oxidation of fatty acyl-CoA is 'oxidation', producing one FADH molecule and enoyl-CoA, which is primarily controlled by the activities of the ACAD (acyl-CoA dehydrogenase) family [299]. Several ACAD family proteins were identified in multiple datasets in the meta-analysis results. This included significantly upregulated ACAD family proteins, ACADS (P16219) (0.19 ± 0.05) & ACDSB (P45954) (0.22 ± 0.07), ACAD9 (0.37 ± 0.05), and ACADVL (P49748) (0.22 ± 0.04), which act on small, long, and very long acyl-CoA molecules, respectively. The next step in the process is 'hydration', the addition of a hydroxyl group to the acyl-CoA molecule, which is primarily regulated by the action of the Enoyl-CoA hydratase enzyme (ECH) family. Both identified ECH enzyme proteins, ECH1 (Q13011) (0.35 ± 0.07) and ECHS1 (P30084) (0.17 ± 0.07), were found to be significantly increased. Next is another oxidation step, which utilises the previously added hydroxyl group to yield an NADH molecule through the activity of hydroxyl acyl-CoA dehydrogenase enzymes. Of this family of proteins, mitochondria trifunctional proteins subunits HADHA (P40939) (0.16 ± 0.04) and HADHB (P55084) (0.16 ± 0.04), and HSD17B10 (0.27 ± 0.03), were found to be significantly upregulated in the meta-analysis results. The final step in the process is thiolysis, which converts fatty acyl-CoA molecules to acetyl CoA– the entry point into the TCA cycle. The mitochondrial enzyme protein ACAA2 (P42765) (0.23 ± 0.04) is a key catalyser of this step and was also found to be significantly upregulated.

Other significant proteins within the fatty oxidation pathway include ACSL1 (Acyl-CoA Synthetase Long Chain; P33121) (0.34 ± 0.08) and ACSS1 (Acyl CoA Synthetase Short Chain; Q9NUB1) (0.21 ± 0.08), catalysers of ATP dependant conversion of long chain fatty

acids to acyl-CoA molecules [300], ECI (Enoyl-CoA delta isomerase) proteins I (P42126) (0.19 ± 0.05) and II (O75521) (0.26 ± 0.03) and DECR1 (Dienoyl-CoA reductase 1; Q16698) (0.27 ± 0.06), integral enzymes for degradation of unsaturated fatty acid-CoA intermediates, and CrAT (Carnitine acetyltransferase; P43115) (0.23 ± 0.06), a modulator of acetyl-coA/acetylcarnitine intracellular pool [301].

Increases in the abundance/activity of proteins CPT1 and HAD have been previously reported in humans following exercise training using traditional wet-lab techniques (i.e., western blots and activity assays) [302]. Upregulation of ACS gene expression, specifically ACSL1, has also been demonstrated in skeletal muscle samples of overweight men following 8 weeks of MICT [303]. Very few of the proteins identified and found to be differentially abundant in the meta-analysis results have previously been studied using traditional wet-lab techniques [302]. Training-induced adaptations to the fatty acid oxidation pathway have previously been identified via changes in the respiration of isolated mitochondria [304, 305], measures of substrate utilisation during exercise bouts following a training intervention [306, 307], and increased myocellular capacity for fatty acid uptake [149]. However, no extensive evidence existed toward the mitochondrial proteome adaptions that contribute to the changes in physiological measures of fatty acid oxidation capacity. Aggregation of proteomic results in this meta-analysis comprehensively demonstrates the upregulation of proteins involved in all major steps of the beta-oxidation pathway in response to exercise training.

### 3.4.3.2 Regulation of TCA cycle proteins

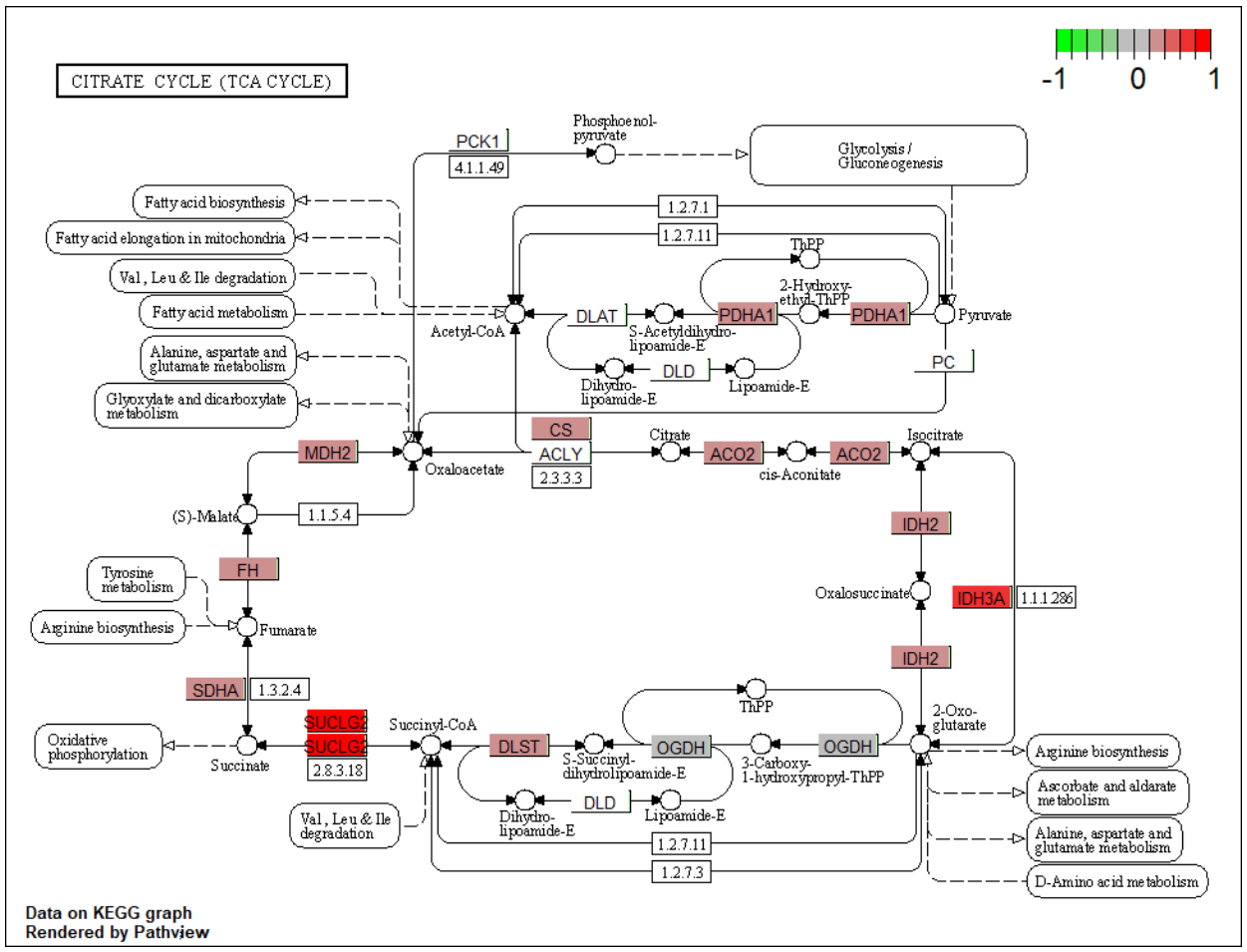


Figure 3:40 Overview of the adaptations to the TCA cycle enzyme proteins highlighting the significant results from the meta-analysis. Red highlights reflect positive $\log_2$ fold changes, and green highlights reflect negative $\log_2$ fold changes.

Of the 20 TCA related enzymes [1], all except OGDHL (Oxoglutarate Dehydrogenase L) were identified in at least one of the datasets. These proteins demonstrated a high degree of identification, as 17 of these proteins had a $K \geq 5$. Sixteen of these proteins were found to be significantly increased, and none were observed to be downregulated. The mean $\log_2$ fold change of these significant proteins was observed to be 0.25 ($\pm$ 0.06).

As can be observed from Figure 3:40, enzyme proteins involved in all reactions of the TCA cycle were found to be significantly upregulated with exercise training. While this finding

makes intuitive sense, it is important to note that many of these enzyme proteins were found to be differentially regulated in only two or three of the individual datasets; this provides an example of the increased statistical power due to the meta-analysis. Catalytic steps by the enzymes PDH (Pyruvate Dehydrogenase), CS, and IDH are known to be rate-limiting [308] [309] [310] and were all found to be significantly increased in the meta-analysis results.

The Pyruvate Dehydrogenase Complex (PDC) is made up of the components E1 (PDHA1 & PDHB), E2 (DLAT), and E3 (DLD) [311]. Of these, only the subunits of component E1, alpha (PDHA1; P08559) (0.16 ± 0.04) and beta (PDHB; P11177) (0.15 ± 0.05), were found to be upregulated. The E2 component (Dihydrolipoyl transacetylase; P10515) (0.16 ± 0.07) protein also demonstrated a trend towards significance (q = 0.055). Components E1 and E2 completely catalyse the production of acetyl-CoA from pyruvate, whereas the E3 component (Dihydrolipoamide Dehydrogenase) is known to moonlight across several other metabolic pathways [312]. The protein PDHX (E3-binding protein; O00330) (0.2 ± 0.05), which anchors the E2 and E3 components together [313], was also found to be upregulated. Further, the rate of activity of the PDC complex is mainly regulated by phosphorylation at sites on the E1-alpha subunit by the action of phosphatases (PDP) and kinases (PDK) [314] [315]. Of these, PDK2 (Q15119) (0.26 ± 0.1) and PDP1 (Q9P0J1) (0.38 ± 0.05) were found to be differentially expressed. Both these regulatory proteins, along with PDHA1, have been observed to be significantly upregulated in endurance-trained compared to sedentary individuals in a population of healthy young men [316], which corroborates our findings. Similarly, the PDHA1 and PDK2 protein has also been demonstrated to increase in abundance, along with a trend for increases in PDHB and PDHX (E3-binding protein) protein abundance, following eight weeks of continuous endurance training in young men [317]. Interestingly, the PDK4 (Q16654) protein has been reported to be unresponsive to exercise training [317] and long-term exercise [316], and this was also observed in the individual and meta-analysis results.

In summary, the pyruvate metabolising components of the PDC complex and associated regulatory kinases and phosphatases strongly appear to be regulated with exercise training.

The CS enzyme catalyses the first step in the TCA cycle, and it has been extensively studied in the context of exercise; CS activity is often used as a biomarker to measure changes in mitochondria volume/content [174]. While CS protein abundance was expectedly upregulated in the meta-analysis results, its regulation within each dataset is interesting. While strong evidence for regulation of CS protein abundance was observed in each study, it was only observed to be significant in *NVT_Granata* ($q<0.05$) and *Deshmukh_2021* ($p < 0.01$, $q < 0.06$), and trended towards significance in *Schild_2015* ($p < 0.02$). These were also the only datasets, except *HVT_Granata*, that also demonstrated significant increase in total mitochondrial protein intensity and mitochondrial protein enrichment. Given that the exercise protocol in *HVT_Granata* was performed following two weeks of training, a further large increase in CS abundance may have been blunted despite an overall increase in total mitochondria protein intensities. The strong evidence for CS abundance regulation, which is a commonly accepted adaptation to exercise training [318], from both the overall meta-analysis results ($q < 1e^{-10}$) and the study-specific observations, helps validate the findings of the meta-analysis results and the use of the bioinformatics pipeline utilised.

Beyond the proteins discussed above, sparse western blot evidence exists for training-induced changes in the protein abundance of other enzymes in humans. Measures of maximal enzyme activity have generally formed the basis of evidence for TCA cycle adaptations following exercise [319]. The results of this meta-analysis now demonstrate that a coordinated upregulation of enzyme proteins across all major reactions underlies the increased TCA cycle capacity observed with exercise training.

Of the TCA cycle-associated proteins, SLC25A11 (Solute Carrier Family 25 Member 11; Q02978) (0.19 ± 0.06), which is part of the Malate Aspartate Shuttle (MAS) complex and has been shown to be upregulated with increased PGC1-alpha expression in mice [320], was differentially expressed. Similarly, ABHD11 (Abhydrolase Domain Containing 11; Q8NFV4) (0.29 ±0.09), a regulator of 2-oxoglutarate metabolism within the TCA cycle that has been shown to be severely downregulated in SM mitochondria in conditions of obesity [293], also demonstrated significant upregulation.

#### 3.4.3.3 Regulation of OXPHOS (ETC) complexes

- **Complex I:**

All the 44 Complex I (CI) subunits [321], except MT-ND4L, were identified and quantified in at least one of the datasets in the current study. Thirty-five CI subunit proteins were identified in five or more datasets, of which 21 were found to be differentially upregulated (Figure 3:41[11]). Twenty-two CI subunits, which were identified in more than one dataset, were found to be differentially upregulated. All 14 core subunits were identified more than once, and all, except MT-ND3 and MTND-6, were identified in five or more datasets ($K \geq 5$). The mean log-fold change of significant CI proteins was observed to be 0.27. No significantly downregulated proteins were observed for any of the CI subunits.

Assembly factor proteins AIFM_1 (Apoptosis -inducing factor 1; O95831) (0.18 ± 0.08), ACAD9 (Q9H845) (0.37 ± 0.12), ECSIT (ECSIT Signaling Integrator; Q9BQ95; K=3) (0.4 ± 0.08), and CI-associated chaperone protein TIMMDC1 (Translocase of Inner

[11] The colour of the nodes signify the fold changes on a spectrum. Red colours denote positive log2 fold changes and the blue colours denote negative log2 fold changes. For context, the highest and lowest aggregate log2 fold changes for each plot are noted in the caption to help contextualise each plot.

Mitochondrial Membrane Domain Containing 1; Q9NPL8) (0.36 ± 0.15) also demonstrated significant upregulation [322] [323]. See Figure 3:41 for a comprehensive summary of the regulation of Complex I proteins.

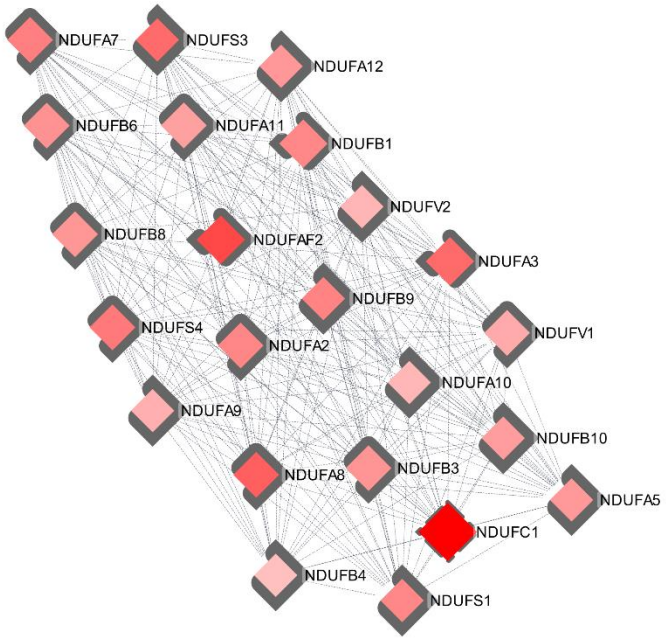


Figure 3:41 Representation of significant Complex I subunits from the meta-analysis results. The width of the borders serves to indicate the number of quantifications for each protein - the thicker the borders, the greater the number of studies in which it was quantified. Red colour represents positive $log_2$ fold change. Highest $log_2$ foldchange: 0.89; Lowest $log_2$ fold change :0.06

- Complex II

All four Complex II (CII) subunits were identified in more than one dataset, with SDH-A (succinate dehydrogenase complex flavoprotein-subunit A), SDH-B, and SDH-C being identified in seven or more datasets. Of these, only SDHA (0.16 ± 0.04) and SDHB (0.17 ± 0.05) were found to be significantly upregulated (Figure 3:42). Two assembly factors (SDHAF4 and SDHAF2) were also quantified in the data, but only in the *Hostrup_2021* dataset and neither was found to be significant.

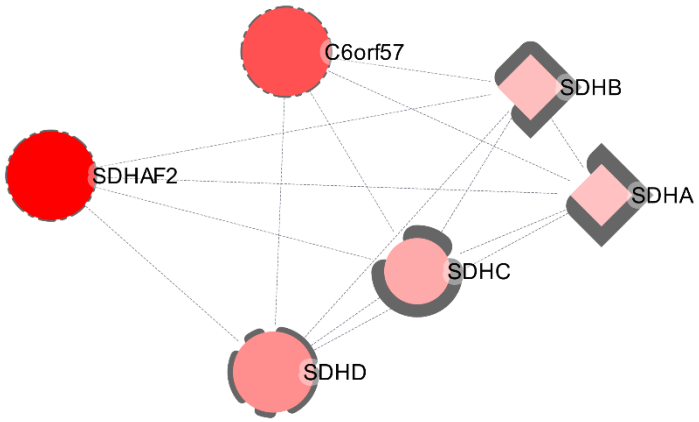


Figure 3:42 Representation of all identified CII subunits and associated proteins from the meta-analysis results. The width of the borders serves to indicate the number of quantifications for each protein - the thicker the borders, the greater the number of studies in which it was quantified. Rhomboid nodes signify significant proteins at the FDR threshold. Red colour represents positive $\log_2$ fold change. Highest $\log_2$ fold change: 0.7; Lowest $\log_2$ fold change: 0.16

- Complex III

All ten Complex III (CIII) subunits were identified in more than one dataset. All subunits except UQCR11 (Ubiquinol-Cytochrome C Reductase-Subunit 11) and MT-CYB (Cytochrome b) were identified in seven or more datasets, of which UQCRQ (Subunit VII), UQCRB (Subunit VI), UQCRC2 (Core Protein 2), and UQCRFS1 (Iron-Sulfur Polypeptide 1) were found to be upregulated (Figure 3:43). The CIII assembly factor LYRM7 protein (LYR Motif Containing 7, Q5U5XO) [324], which works as a UQCRFS1 chaperone, was also found

to be upregulated (0.18 ± 0.06). The CIII subunit cytochrome b (P00156; 0.43 ± 0.13; K=3) was the only mitochondrially encoded subunit that demonstrated significant upregulation in the meta-analysis results. The average fold change of differentially regulated CIII subunits was approximately 0.24 (± 0.03).

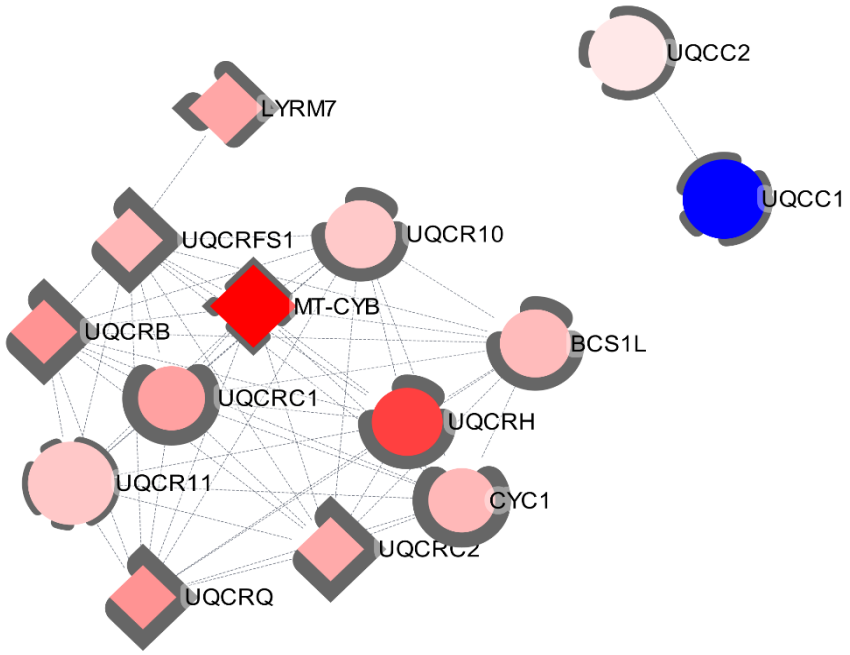


Figure 3:43 Representation of all identified CIII subunits and associated proteins from the meta-analysis results. The width of the borders serves to indicate the number of quantifications for each protein - the thicker the borders, the greater the number of studies in which it was quantified. Rhomboid nodes signify significant proteins at the FDR threshold. Red colour s represents positive $\log_2$ fold change; Blue colours represent negative $\log_2$ fold change.Highest $\log_2$ fold change: 0.44; Lowest $\log_2$ fold change: -0.14

- Complex IV

Of the 21 subunit proteins of Complex IV (CIV) [1], 15 were identified in more than one study within the meta-analysis dataset. COX (Cytochrome c Oxidase) subunits 4I2, 6A1, 6B2, 7B, 7B2, and 8C were not identified in any of the datasets in the meta-analysis. COX subunits 6C, 5B, 4I1, 7A2, and 6B1, all of which were quantified in five or more studies, were found to be significantly upregulated in the meta-analysis results (Figure 3:44). The MT-CO2

(alias: COX2) stabilising protein COA6 (Cytochrome c Oxidase Assembly Factor 6; Q5JTJ3) (0.15 ± 0.05) was also found to be significantly upregulated (Figure 3:44) While many other assembly factors, such as COA1 COA3, COA4, and COA5, were identified, none of them were significant in the meta-analysis results [323]. The average fold change of CIV subunits was found to be 0.28 (± 0.07).

- Complex V

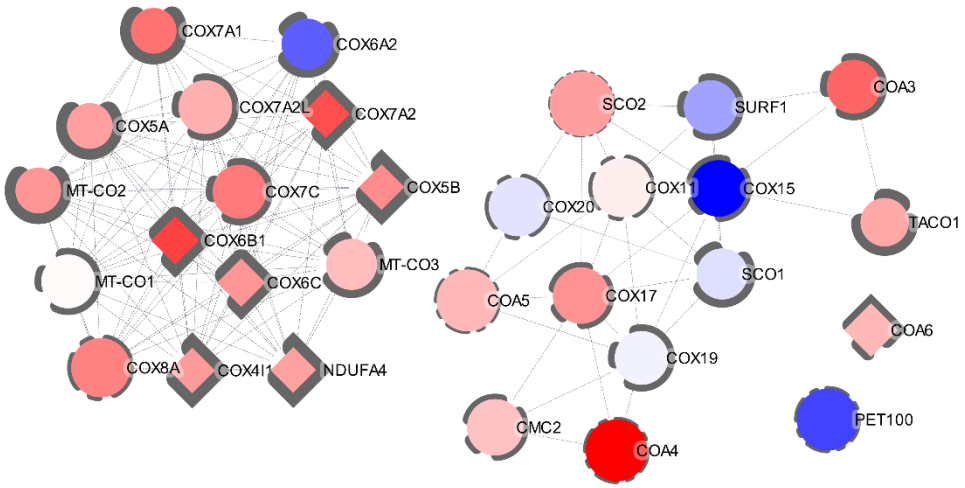


Figure 3:44 Representation of all identified CIV subunits and associated proteins from the meta-analysis results. The width of the borders serves to indicate the number of quantifications for each protein- the thicker the borders, the greater the number of studies in which it was quantified. Rhomboid nodes signify significant proteins at the FDR threshold. Red colours represents positive $log_2$ fold change; Blue colours represent negative $log_2$ fold change. Highest $log_2$ fold change: 0.74; Lowest $log_2$fold change: -0.4

Of the 21 subunits proteins of Complex V (CV) [1], ten were identified in more than one study within the dataset. Of the identified proteins, subunits PD, PB, F1C, PO, ME and DMAC2L (Distal Membrane Arm Assembly Complex) were found to be significantly upregulated (Figure 3:45). Assembly factors ATPAF1, ATPAF2, and TMEM70 were identified in the results; however, only ATPAF1 (ATP synthase mitochondrial F1 complex assembly factor 1; Q5TC12) (0.3 ± 0.06) was significant. The mean fold change demonstrated by significant CV proteins was 0.24 (± 0.03).

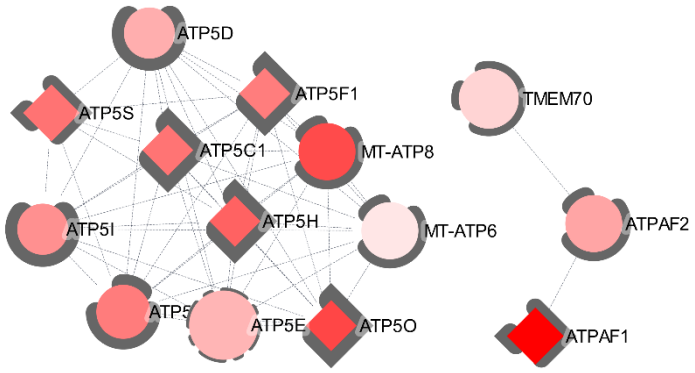


Figure 3:45 Representation of all identified CV subunits and associated proteins from the meta-analysis results. The width of the borders serves to indicate the number of qualifications for each protein - the thicker the borders, the greater the number of studies in which it was quantified. Rhomboid nodes signify significant proteins at the FDR threshold. Red colour s represents positive $log_2$ fold change; Blue colours represent negative $log_2$ fold change. Highest $log_2$ fold change: 0.3; Lowest $log_2$ fold change: 0.04

### 3.4.3.4 Regulation of Mitochondrial Dynamics

As discussed before, mitochondrial dynamics refers to the interaction between the processes of mitochondrial fusion, fission, and organisation, and exercise interventions have previously been shown to affect all these aspects of mitochondria function.

With respect to fusion proteins, OPA1 (O60313) (0.17 ± 0.09) and MTCH2 (0.47 ± 0.13) were identified as significantly upregulated and were quantified in six and seven different datasets, respectively. Mitofusin protein 1 (MFN1) was not quantified in any of the datasets; however, MFN 2 appeared in six different datasets but did not reach the overall significance threshold. Seven of the 15 proteins classified as 'Fission' proteins, including FIS1 (quantified in seven datasets), were observed in the meta-analysis results; however, none of them were found to be significantly upregulated.

Cellular apoptosis-implicated proteins AIFM1 (Apoptosis Inducing Factor Mitochondria 1; O95831), BNIP3, ENDOG (Endonuclease G; Q14249) (0.22 ± 0.06) and

GHITM (Growth Hormone Inducible Transmembrane Protein; Q9H3K2) (0.26 ± 0.07) were found to be differentially abundant in the meta-analysis results. The AIFM1 protein is also thought to be involved in OXPHOS as an NADH-oxidoreductase [325] and an essential component of the OXPHOS assembly and maintenance [326] mechanisms; this may help to better explain the evidence of increased abundance with exercise training as observed in the results. The ENDOG protein has been shown to be necessary for mtDNA replication and for cleavage of oxidatively damaged (through ROS) mtDNA fragments through investigations on knockout human cell lines [327]. The GHITM protein is implicated in the negative regulation of apoptosis as its increased expression has been shown to be linked with delayed cytochrome c release and maintenance of mitochondria cristae integrity and organisation; these processes promote cell survival [328]. Despite being annotated within apoptotic pathways, the regulation of these specific proteins within this pathway signifies promotion of mitochondria biogenesis and maintenance of mitochondria components.

All but one (CHCHD6) of the MICOS complex proteins were identified in at least one of the datasets. Of these, APOOL (Apolipoprotein O Like protein; Q6UXV4) (0.16 ± 0.05) and CHCHD3 (Q9NX63) (0.19 ± 0.05) were found to be differentially expressed in the meta-analysis results. The APOOL protein has been previously shown to bind specifically only to the phospholipid cardiolipin and, therefore, has been thought to be implicated in cardiolipin transport from the inner membrane to the cristae membrane especially at sites of super-complex formation and Complex V assembly [329] [330]. The CHCHD3 protein has been demonstrated to be integral to the maintenance of crista integrity via knockout analysis that results in fragmented mitochondria morphology, impaired fusion, and severely reduced oxygen consumption [219]. Immunoprecipitation studies have also shown this protein to be bound with SAMM50 (Sorting and Assembly Machinery Component 50) [219] [331], which was also differentially expressed in the meta-analysis results (0.19 ± 0.04). Alongside maintaining

cristae stability, the SAMM50 protein plays a crucial role in assembling and integrating β-barrel proteins. This has been demonstrated by knockout analysis, which results in significant reductions in the protein levels of CHCHD3, IMMT, TOMM subunits 20 and 40, and VDAC [332] [333]. Figure 3:46 summarises these findings from the meta-analysis.

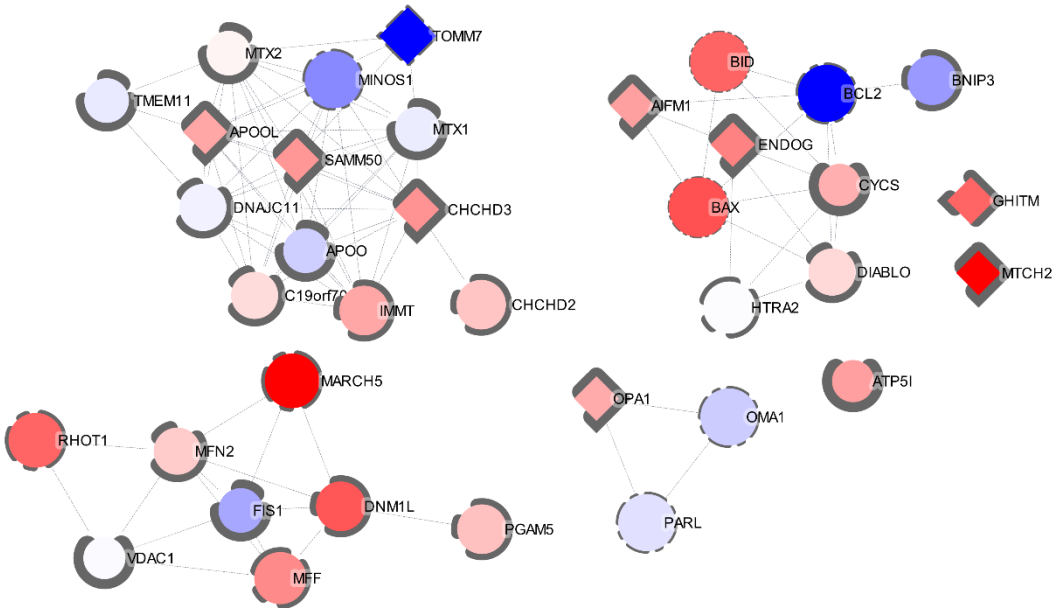


Figure 3:46 Summary of regulation of fusion/fission/mitophagy annotated proteins according to the meta-analysis results. The width of the borders serves to indicate the number of qualifications for each protein - the thicker the borders, the greater the number of studies in which it was quantified. Rhomboid nodes signify significant proteins at the FDR threshold. Non-significant nodes without protein-protein interactions with other nodes in the annotation are not shown. Red colour s represents positive $\log_2$ fold change; Blue colours represent negative $\log_2$ fold change. Highest $\log_2$ fold change: 0.69; Lowest $\log_2$ fold change: -1.3

### 3.4.3.5 Detoxification and ROS handling

Forty of the 51 genes annotated as ‘Detoxification’ were detected in the results of the meta-analysis, of which 12 were found to be significantly differentially expressed.

ROS scavenging antioxidant enzymes SOD2 (P04179) (0.17 ± 0.05), PRDX3 (Peroxiredoxin 3; P30048) (0.3 ± 0.07), PRDX5 (P30044) (0.2 ± 0.08), and PRDX6 (P30041) (-0.1 ± 0.03), were observed to be differentially expressed in the meta-analysis results. The protein abundance of SOD2, a converter of free radicals into hydrogen peroxide, has previously been shown to be positively regulated in response to exercise training in healthy young men [159], as well as in physically active older populations compared to age-matched sedentary controls [334], through western blot quantification; these findings support the results of the current meta-analysis. The PRDX family of proteins are modulators of cellular hydrogen peroxide ($H_2O_2$) levels in human cells, and it is estimated that > 90% of peroxide in the mitochondria is reduced by the action of these proteins [335]. Oxidation of PRDX proteins is also thought to play a role in cellular signalling pathways that modulate activities of transcription factors STAT3 and p38, which are linked downstream to muscle anabolism and mitochondria biogenesis, respectively [336]. The only previous evidence of exercise training on PRDX protein abundance is from a 3-month (2 times a week) moderate-intensity mixed training protocol on older diabetic men, which demonstrated an increase in PRDX5, SOD2 and GPX1 (Glutathione peroxidase) protein levels. ABHD10 (Abhyrdolase Domain Containing 10; Q9NUJ1) (0.14 ± 0.03), a mitochondria localised positive regulator of PRDX5 antioxidant activity [337], was also found to be upregulated in the meta-analysis results.

The TXN proteins function by reducing other antioxidant enzymes, including PRDXs, and thus replenish the supply of these peroxide-reducing enzymes. Of the TXN protein family, TXN2 is localised exclusively to the mitochondria [338] and was found to approach significance in the meta-analysis results (0.26 ± 0.12, q=0.09). TXNR (Theoredoxin reductase) proteins are, in turn, responsible for reducing the oxidised TXN proteins and thus maintaining their continuous supply. Both the cytosolic (TXNRD1) and mitochondrial (TXNRD2) isoforms of this protein were identified in the dataset; however, none of these were found to be

differentially expressed. Training-related upregulation of the TXN family of proteins, specifically TXNRD2, has been demonstrated via western blotting in rat skeletal muscle [339]; however, no evidence from human studies has been reported.

Reactive aldehyde species are produced as a result of lipid peroxidation and the oxidative degradation of lipid membranes by radical species, which are both accelerated in conditions of oxidative physiological stress such as during bouts of endurance exercise [340]. The Aldehyde Dehydrogenase (ALDH) family of proteins are some of the key detoxifying enzymes for aldehydes in the human body, of which at least five are known to localise to the mitochondria [341]. Five members of the ALDH family, including ALDH2, were identified in the meta-analysis results and only ALDH1B1 (Uniprot ID: P30837; 0.39 ± 0.08) was found to be significantly differentially expressed. Of this family of proteins, ALDH2 protein overexpression has been reported in mice following eight months of HIIT and MICT training; however, no such evidence for ALDH proteins has been studied or reported in humans. Another aldehyde reducing protein, AKR7A2 (Aldo-keto reductase family 7, O43488; 0.16 ± 0.04) was also found to be differentially expressed in the meta-analysis results.

Glutathione (GSH) is a non-enzymatic molecule that directly eliminates free radicals, as well as acting upon the intermediate $H_2O_2$ molecules [342]. Glutathione Reductase (GSR) and Glutathione peroxidase (GPx) are catalytic enzymes that are involved in the reduction and oxidation (by free radicals and intermediate species) of GSH, respectively, thus maintaining redox balance in the cell [343]. GPx1 and GPx4, the mitochondria-localised isoforms of GPx [344], and GSR, were all identified in multiple datasets within the meta-analysis; however, these proteins were not found to be significantly differentially expressed in any of the datasets. Glutathione-related enzymes HAGH (Hydroxyacylglutathione Hydrolase; Q16775; -0.19 ± 0.04), which helps prevent accumulation of toxic by-products of glycolysis [345], and GSTK1 (Gluathione-S-transferase kappa 1; 0.3 ±0.05), a catalyser of glutathione conjugation (i.e.,

addition of GSH to ROS molecules) [342] and the only mitochondria-specific GST family member [346], were found to be differentially expressed. See Figure 3:47 for a comprehensive summary of significant anti-oxidant and detoxification proteins from the meta-analysis results.

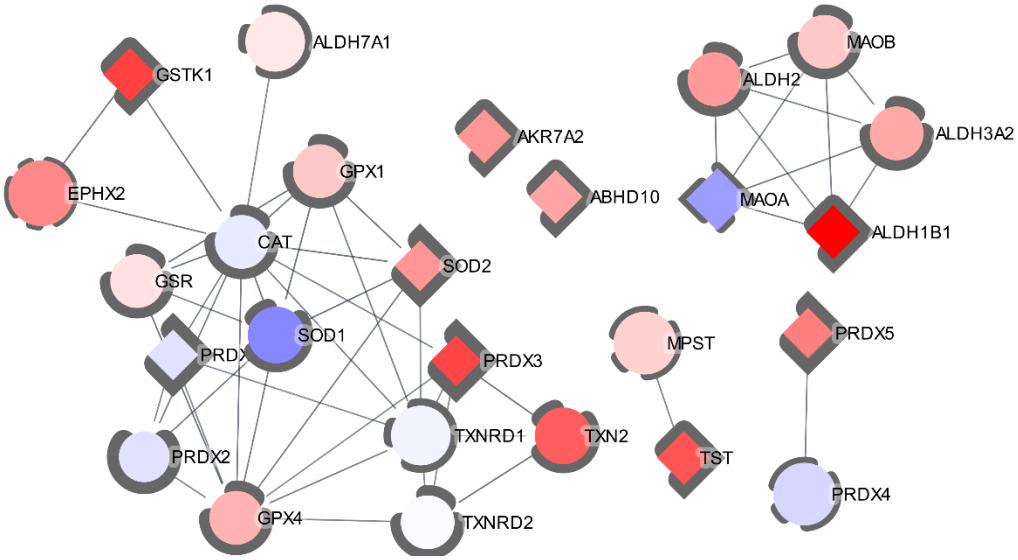


Figure 3:47 Summary of regulation of enzymatic ROS dismutating proteins according to the meta-analysis results. The borders' width indicates the number of quantifications for each protein - the thicker the borders, the greater the number of studies in which it was quantified. Rhomboid nodes signify significant proteins at the FDR threshold. Non-significant nodes with sparse protein-protein interactions with other nodes in the annotation are not shown. Red colour s represents positive $log_2$ fold change; Blue colours represent negative $log_2$ fold change. Highest $log_2$ fold change: 0.39 ; Lowest $log_2$ fold change: -0.43

### 3.4.3.6 Regulation of protein import, sorting, and homeostasis mechanisms

Sixty-two of the 86 proteins annotated to the protein import, sorting, and homeostasis pathways were observed in the meta-analysis results. The Figure 3:48 summarises the major components of the mitochondria protein import pathway.

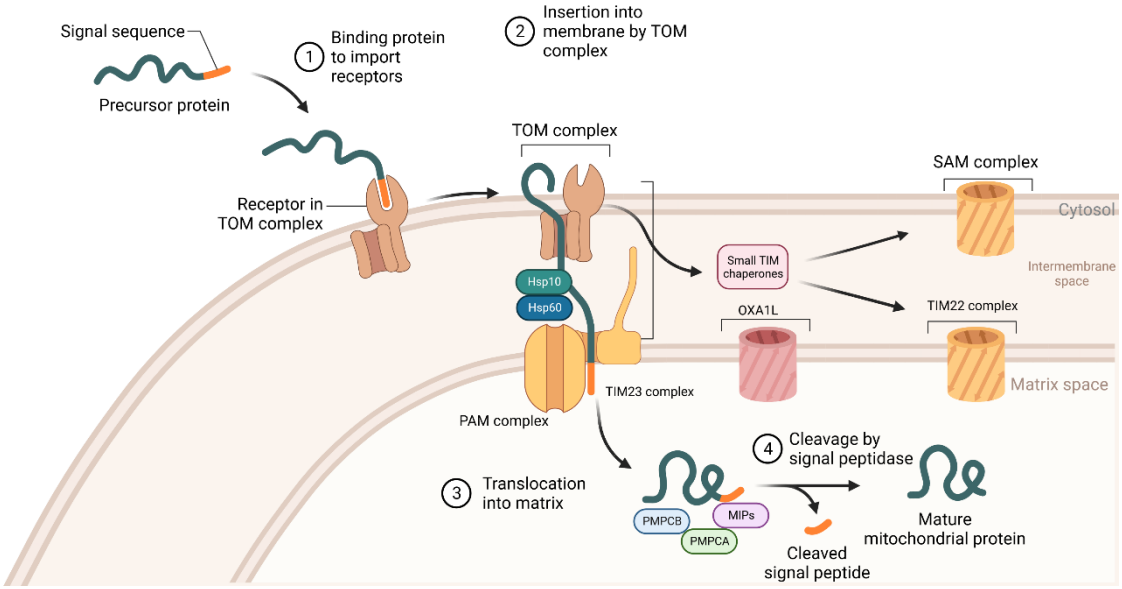


Figure 3:48 Summary of mitochondria protein import mechanisms. Adapted from [2] and "Protein Import into the Mitochondria" from Biorender

Of the TOM complex proteins, TOMM 22, TOMM40, TOMM40L, and TOMM70 were identified in the meta-analysis results, of which TOMM 22 (Q9NS69) (0.25 ± 0.1; K=6) was differentially expressed. Interestingly, TOMM22 was not found to be differentially expressed in any of the individual datasets. The TOMM22 protein, along with TOMM20 and TOMM70 proteins, are known to be receptor proteins that accept cytosolic precursor proteins and transport them to the OM import pore. The TOMM20 protein, a receptor protein like TOMM22, has previously been shown to be upregulated with contractile activity in the skeletal muscle of animals [347] and humans [169]. In rat skeletal muscles following chronic contractile activity, upregulation of TOMM20, along with chaperone proteins HSP60 and HSP70, has also been shown to coincide with an accelerated rate of import of MDH and TFAM proteins, both of which were significantly upregulated in this meta-analysis [348].

All three proteins annotated as part of the Sorting and Assembly Machinery (SAM) complex, which is the primary assembler and inserter of beta-barrel proteins into the

mitochondria OM and plays a crucial role in OM biogenesis, were found in the meta-analysis results [2]. Of these, SAMM50 (Sorting And Assembly Machinery Component 50; Q9Y512) (0.18 ± 0.05), the primary subunit, was differentially expressed. No previous investigation for this protein was found in the context of exercise training in animal or human studies, and it was also not found to be differentially expressed in any single dataset within the meta-analysis. However, this protein, along with other components of the SAM complex, has been found to be negatively regulated with ageing in mice skeletal muscle and is thought to therefore contribute to the disruption in mitochondria-endoplasmic reticulum contacts observed in aged skeletal muscle [349].

Proteins localised to the IMS undergo folding and modifications through the 'mitochondrial IMS assembly' (MIA) pathway, which allow these proteins to remain 'entrapped' within the IMS [350]. Multiple subunits of different ETC complexes, assembly factors, and MICOS complex proteins, are known to be the target of this import pathway [351]. All three annotated proteins (CHDH4, GFER, AIFM1) of the MIA pathway were identified in the meta-analysis results, and of these AIFM1 was significantly upregulated after exercise training (0.17 ± 0.03). The protein GFER (alias: ERV1, P55789) (-1.01 ± 0.36) was also found to be differentially expressed; however, this was only observed in one dataset (*Hostrup_2021*). The AIFM1 protein has been shown to be crucial to Complex I biogenesis and the 'disulphide relay' process (the functional process underlying the MIA pathway) through knockout experiments in mice that demonstrate impaired insertion of CI subunit NDUFS5 in complex intermediaries, reduced total abundance of multiple CI subunits, and reduced import and oxidation of MIA substrates despite unchanged abundances of CHCHD4 (alias: MIA40) [352]. Given the centrality of this protein ultimately to CI biogenesis and assembly, it is relevant that this protein demonstrates strong evidence for regulation with exercise training alongside the observed increases in CI subunit proteins in the meta-analysis results.

Transport of proteins through the mitochondrial IM and into the matrix is mediated mainly by the action of the TIM23 complex, whereas the TIM22 complex functions to import carrier proteins localised to the IM [353] [2]. All but one of the subunits of the TIMM23 complex (TIMM17A) were identified in the meta-analysis results; however, none of them were differentially expressed. TIMM17B is the predominantly expressed isoform in mammalian skeletal muscle [354], which explains the lack of detection of the alternate isoform A in the results. Of the TIMM22 complex proteins, TIMM29 and acyl glycerol kinase (AGK, Q53H12) were observed in the results, but only AGK was found to be differentially expressed ($0.19 \pm 0.06$). The core component of the TIMM22 complex, TIMM22, was not detected in any of the datasets. Alongside being a subunit of the TIMM22 complex, confirmed by previous immunoprecipitation analysis [355], the AGK protein has been shown to play a role as a lipid kinase ultimately involved in the synthesis process of cardiolipin [356], which may also explain its increased abundance in response to exercise to support the increased mitochondria biogenesis. Sparse western blot evidence exists for exercise-induced regulation in the protein abundance of skeletal muscle TIM subunits through either animal or human investigations. Only TIMM23 complex subunits TIMM17 and TIMM23 have been previously shown to be significantly upregulated following artificial chronic contractile stimulation [357] and exercise training [358] in rats and mice, respectively.

Associated with the TIM23 complex is the presequence translocase-associated motor (PAM), which is essential to the ATP-dependant process of preprotein import into the matrix [359]. Component subunits of this complex, DNAJC19 (alias: TIMM13, DnaJ Homolog Subfamily C Member 19), GRPEL1, PAM16 (alias: TIMM16, Presequence Translocase Associated Motor 16), and TIMM44, were identified in the meta-analysis results. Of these, GREPL1 and TIMM44 were found to be differentially expressed. Interestingly, these proteins also demonstrate two of the highest $\log_2$ fold changes within the DEPs. Hydrolysis of ATP

molecules to drive the import process across the IM is performed by another PAM-associated component protein HSPA9 (alias: mtHsp70; P38646), which was also observed to be differentially expressed (0.33 ± 0.04; K=8). The protein GRPEL1 acts as a 'nucleotide exchange factor' supporting the function of HSPA9 by controlling the binding and release of HSPA9 substrate proteins [360] [361]. As discussed previously, TIMM44 plays a crucial role in the binding of HSAP9 to the TIM23 complex and is also associated in a complex with HSPA9, as confirmed by immunoprecipitation [362] [363]. The TIMM44 protein is also implicated in the mitochondrial biogenesis process by mediating the binding of mtHSP70 with the TIM complex (TIM44/mtHSP70 complex), which allows the conformational change necessary for the TFAM protein, a key regulator of mtDNA transcription, to attach to the chaperone protein. This is a necessary step for transporting the TFAM protein into the mitochondrial matrix where it interacts with the mtDNA to initiate transcription [364] [365]. The functional proximity of GRPEL1, HSPPA9, and TIMM44 with each other may help to explain the similar high fold changes for these three proteins (all > 90 percentile of proteins and with K ≥ 5). Of the PAM and associated subunits, HSPA9 (mtHSP70) has been well-studied previously shown via western blot quantification to increase in relative abundance in response to endurance training in human skeletal muscle through [169] [366]. Similarly, GRPEL1 protein abundance has also been found to be upregulated in response to contractile activity in striated muscle from rat diaphragm samples [367].

The inner membrane localised protein OXA1L (Oxidase Assembly 1-Like) is essential to the insertion and assembly of mitochondria-encoded subunits of the OXPHOS complexes [368]. This protein has also been shown to be physically associated to mitoribosome MRPL45, which is bound to the ribosome exit tunnel - the site from where polypeptide chains emerge post translation [369] [370] [2]. Both OXA1L (Q15070) (0.21 ± 0.06, K=6) and MRPL45 (Q9BRJ2) (0.44 ± 0.18; K=1) were found to be differentially expressed in the results; however,

MRPL45 quantification was only derived from the *Hostrup_2021* dataset. Despite its functional importance, OXA1L protein content does not appear to have been previously studied in the context of exercise training in either human or animal skeletal muscle. Further, it was only identified as differentially expressed in one dataset (*HVT_Granata*) of the six in which it was quantified. This finding again highlights the power of this meta-analysis design which, despite the limited number of studies, has identified another protein that was previously not identified as changing in response to exercise training. Depending on the availability of antibodies, this protein could be a target of western blot-based validation, and further proteomic studies in the future.

Small TIMM chaperone proteins localised to the IMS function to protect hydrophobic precursor proteins in transit within the IMS to either the SAM complex or the TIMM22 complex for protein insertion [2]. Of the TIM chaperone-annotated proteins, TIMM8A/B, TIMM9, TIMM10, and TIMM13 were identified in the meta-analysis results. No differentially expressed proteins from this set were found in the meta-analysis results; however, Subunit 13 (Q9Y5L4) (0.12 ± 0.06; K=6) was found to approach significance (q =0.09). The protein content of TIMM13 has previously been shown to be downregulated in conditions of chronic disuse in mice skeletal muscle [371]. Alongside the TIMM proteins, HSP proteins HSPE1 (*alias*: HSP10) and HSPD1 (*alias*: HSP60), which exist together in a complex in the IMS and assist in protein folding [65] [372], are also classed as chaperone proteins. Approximately half of the matrix localised proteins are estimated to interact with the HSP60/HSP10 chaperonin complex [65]. Both HSPE1 (P61604) (0.4 ± 0.11; K=7) and HSPD1 (P10809) (0.29 ± 0.06; K=8) were found to be differentially expressed in the meta-analysis results. HSEPE1 (HSP60) has previously been the subject of multiple exercise-focused investigations in both animal and human studies [373]. Elevated basal HSP60 protein abundance has been demonstrated via western blot quantification in endurance-trained participants compared to

sedentary/recreationally active controls in young men [374] [375]. Despite this evidence, and the finding of significance at the meta-analysis level, neither HSPE1 (HSP60) or HSPD1 (HSP10) were found to be differentially expressed in the *Schild_2015* dataset - a case-control comparison of endurance-trained subjects and non-trained participants. No previous studies appear to have demonstrated or investigated the effects of exercise training on the protein abundance of the HSP60/HSP10 complex. However, both HSPE1 and HSPD1 were observed to be upregulated in three individual datasets included in this study as well as in the meta-analysis results, which provide comprehensive evidence of regulation of this complex in response to endurance exercise training.

The final subset of proteins to be discussed under this category belong to the functional mechanism of regulation of proteostasis (protein homeostasis). Key differentially expressed proteins under this annotation included STOML2 (Stomatin Like 2, Q9UJZ1; K=6) (0.24 ± 0.04) and the peptidases, LAP3 (Leucine Aminopeptidase, P28838; K=8) (0.24 ± 0.06), AFG3L2 (ATPase Family Gene 3-Like 2, Q9Y4W6; K=7) (0.21 ± 0.05), LONP1 (Lon Peptidase 1; P36776; K=8) (0.2 ± 0.05), and MIPEP (Mitochondrial Inner Membrane Peptidase; Q99797; K=5) (0.18 ± 0.07). The protein STOML2 is involved in many functions. It is known to associate in a complex (*SPY*) with and modulate the substrate-specific activities of proteases PARL (Presenilins-associated Rhomboid-Like) and YME1L (YME1 Like 1 ATPase). This complex promotes mitochondria fusion and protects against apoptosis by inhibiting the proteolysis of pro-fusion protein OPA1 but facilitating the degradation of pro-mitophagy protein PINK1 [376]. Interestingly, YME1L (Q5T8D1) (0.57 ± 0.21) was also observed to be differentially expressed in the results; however, it was only quantified in one dataset (*Hostrup_2021*). Loss of function analysis of YME1L in C2C12 myotubes has been shown to upregulate MuRF1 (Muscle Ring-Finger Protein-1; alias: TRIM63) expression and thus directly contributes to pro-atrophy signalling. This protein has been shown to play a role

in controlling the accumulation of non-assembled respiratory chain subunits and in degradation of oxidatively damaged membrane proteins [377]. Exercise-induced stress is known to upregulate pathways of biogenesis of respiratory complex subunits and cause an increase in ROS production within skeletal muscle [378]. This may lead to accumulation of non-assembled respiratory complex proteins, especially as it is observed that not all complex subunits respond to training stimuli or respond to the same degree (see section on ETC complex proteins; or see [81] for example), and increased oxidative damage of membrane proteins [378], respectively. This may provide support for explaining the positive regulation of YME1L with exercise training.

Independent of its function in the *SPY* complex, STOML2 is also known to interact with the prohibitin complex that consists of the subunits PHB (Prohibtin 1; P35232) (0.21 ± 0.05; K=8) and PHB2 (Prohibitin 2; Q99623) (0.17 ± 0.06; K=8), both of which were found to be significantly upregulated in the meta-analysis results, to upregulate cardiolipin biosynthesis [379]. The PHB complex has been thought to be implicated in multiple mitochondrial processes, including attenuation of fission and apoptosis signalling and the maintenance of mitochondria morphology (evidenced by ultrastructure studies in PHB2 deficient cells) through regulation of OPA1-mediated mechanisms [380] [381], regulating proteolysis of unassembled OXPHOS proteins [382], promotion of mtDNA replication through stabilising DNA structure, promoting organisation of IM—bound nucleoids, affecting TFAM binding[383], and increasing respiratory complex chain assembly [384] and activity [385]. Further support for this complex's vital role in multiple facets of mitochondria regulation is its high basal expression in high-ATP demand mitochondria-reliant cell types, such as skeletal and cardiac muscle, brown adipocytes, and renal tubules [386] [387]. A significant effect of exercise training (HIIT) on prohibitin protein abundance has been previously demonstrated via proteomics in a dataset independent from this study [218], which supports this finding.

Interestingly, both PHB subunits are found to be interdependent in multiple organisms, including humans, such that one is degraded in the absence of the other [388]. Relatedly, both PHB1 and PHB2 are found to co-precipitate together and are not detected as free proteins [389] [390], supporting our finding that both prohibitin subunits were differentially expressed with similar aggregate fold changes. See Figure 3:49 for the comprehensive summary of significant protein import, assembly, and proteostasis proteins from the meta-analysis results.

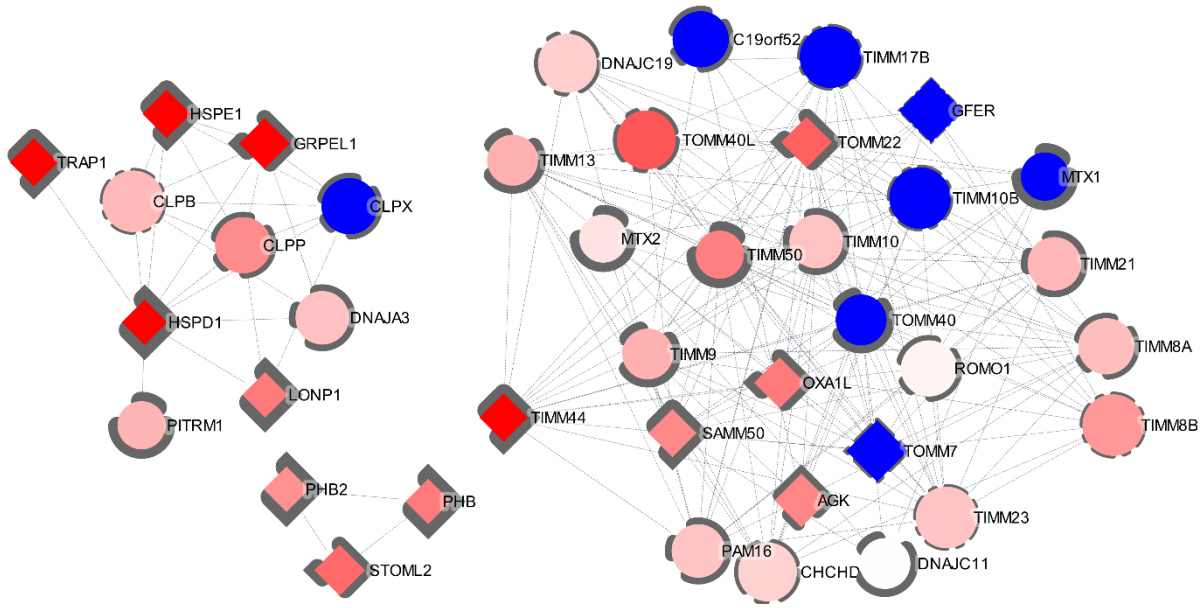


Figure 3:49 Summary of regulation of protein-import mechanism proteins according to the meta-analysis results. The width of the borders serves to indicate the number of qualifications for each protein- the thicker the borders, the greater the number of studies it was quantified in. Rhomboid nodes signify significant proteins at the FDR threshold. Non-significant nodes with sparse protein-protein interactions with other nodes in the annotation are not shown. Highest log2 fold change: 0.42. All negative log2 fold changes are shown as blue (no gradient). White colour denotes log2fold change ~0.

### 3.4.3.7 Regulation of electron carrier proteins and small molecule transport proteins

Of the 15 proteins annotated as ‘electron carriers’, 11 were quantified in the meta-analysis results. Of these, four were identified as differentially expressed.

Electron Transport Flavoprotein (ETF) subunits alpha and beta, encoded by genes ETFA (P13804) (0.32 ± 0.06) and ETFB (P38117) (0.06) respectively, constitute the ETF enzyme [391] and both subunit proteins were found to be upregulated in the meta-analysis results. The protein ETF: QO (electron transfer flavoprotein-ubiquinone oxidoreductase; alias: ETFDH; Q16134) (0.19 ± 0.07), which links the ETF complex to the ETC through facilitating electron transfer to coenzyme-Q (CoQ) and ultimately supplies electrons to CIII, was also observed to be upregulated in the meta-analysis results. The ETF complex has been shown to accept electrons from enzymes of fatty-acid oxidation and amino acid catabolism [391]; this links multiple substrate-oxidation pathways with the ETC and may help explain the consistent response in the component proteins following exercise training. Relatedly, both subunits (ETFA and ETFB) were observed to be significantly upregulated in four individual datasets. Finally, mitochondrial cytochrome-b (MT-CYB) , which is involved in the redox cycle of coenzyme Q (also referred to as the Q-cycle) [35], was also observed to be significantly upregulated. See Figure 3:50 for the comprehensive summary of significant electron transfer proteins from the meta-analysis results.

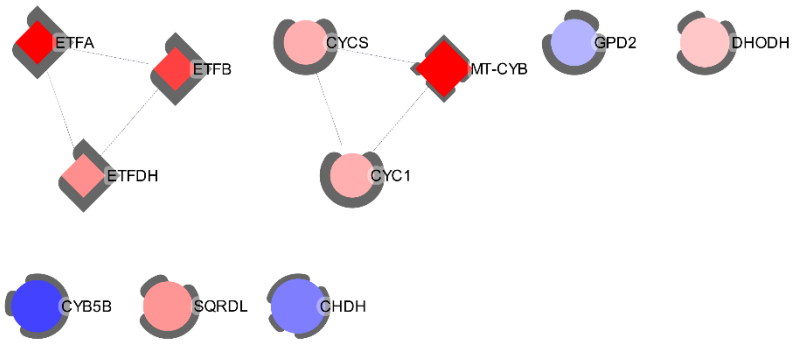


Figure 3:50 Summary of regulation of electron transfer proteins according to the meta-analysis results. The width of the borders serves to indicate the number of qualifications for each protein- the thicker the borders, the greater the number of studies it was quantified in. Rhomboid nodes signify significant proteins at the FDR threshold. Red colour s represents positive $log_2$ fold change; Blue colours represent negative $log_2$ fold change. Highest $log_2$ fold change: 0.44; Lowest $log_2$ fold change: -0.21

Of the annotated small molecule carrier proteins, 29 were quantified in the meta-analysis results and ten were identified as differentially expressed.

Members of the mitochondrial solute carrier family (SLC25), which are involved in the transport across the inner membrane of amino acids, nucleotides, fatty-acids, and carboxylates, among others [287], predominantly featured in the set of differentially expressed proteins. This includes the proteins SLC25A4 (P12235) (0.25 ± 0.08) and SLC25A5 (P05141) (0.23 ± 0.06), also referred to as adenine nucleotide translocases (ANT), which are primary ATP/ADP transporters in muscle cells and thus integral for energy production, SLC25A20 (alias: CAC/CACT; O43772) (0.25 ± 0.08), which is the carnitine/acylcarnitine transporter component of the carnitine shuttle system exchanging acyl-carnitine and free carnitine across the inner membrane and thus facilitating fatty-acid oxidation[392], member proteins of the malate-aspartate shuttle complex SLCA2511 (Q02978) (0.19 ± 0.06), SLC25A12 (O75746) (0.12 ± 0.05), and SCL25A13 (Q9UJS0; K=1) (0.46 ± 0.16) [393], and SLC25A3 (alias: PiC;

Q00325) (0.39 ± 0.08), the only importer of phosphate ions across the inner membrane [286, 287]. The final member of the SCLA25 family to be discussed is the protein sideroflexin 1 (SFXN1, alias: SLC56A1; Q9H9B4) (0.25 ± 0.08; K=4), an inner membrane multipass protein implicated as the only known importer of cytosolic serine into the mitochondria [394]. The mitochondrial concentration of serine is important for supporting the process of one-carbon metabolism, the generation of reactive one-carbon species that is carried out in parallel in the cytosol on folate as a substrate, particularly in replicating cells. This supply of one-carbon species ultimately contributes mainly to the process of nucleotide biosynthesis. Supporting the finding of increased capacity for serine import is the observed increase in MTHFS (5,10-methenyltetrahydrofolate synthetase; P49914; K=2) (0.59 ± 0.21), the only other quantified component of the one-carbon pathway, which is involved in the downstream conversion of serine following import [394].

Another significant protein under this annotation was the previously mentioned MTCH2 (alias: SLC25A50), which, despite being a member of the SLC25A family, does not have known substrates or transporter activity and is also localised to the outer membrane in contrast to other proteins of the family [395]. This protein has been well-studied and implicated in the regulation of multiple pathways [395]. Specifically, within skeletal muscle, loss-of-function of this protein in mice models has been shown to increase mitochondrial mass, mitochondrial area, protein content of OXPHOS complex subunits, and mitochondrial respiration [396]. Due to this evidence (see also [397]), this protein has been referred to as a repressor of mitochondrial mass and function, which is highly conflicting with its consistent increased expression observed following exercise training (found to be significant in four independent datasets in this study and also reported to be significant in a published study [224]). Conversely, loss-of-function analysis in K562 cell lines has demonstrated that MTCH2 is integral for outer membrane insertion of α-helical proteins, particularly multipass proteins

[395], and this might better explain the functional relevance of the relatively large increase observed with exercise training as in the meta-analysis results. No previous western blot investigations have analysed the response of this protein to exercise training in either animals or humans. Given the strength of evidence for its regulation as demonstrated in this study, and the drastic positive functional impact of its deficiency as seen in animal models [396], further investigations to validate its exercise training-mediated response and elucidate its concomitant functional significance in human skeletal muscle is warranted. See Figure 3:51 for a comprehensive summary of significant small molecule carrier transport proteins from the meta-analysis results.

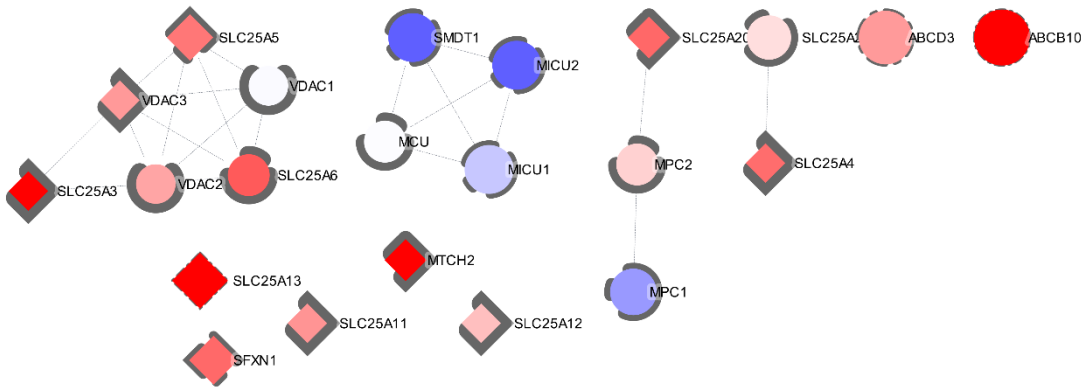


Figure 3:51 Summary of regulation of small-molecule carrier proteins according to the meta-analysis results. The borders' width indicates the number of quantifications for each protein - the thicker the borders, the greater the number of studies in which it was quantified. Rhomboid nodes signify significant proteins at the FDR threshold. Red colour s represents positive $\log_2$ fold change; Blue colours represent negative $\log_2$ fold change. Highest $\log_2$ fold change: 0.48; Lowest $\log_2$2fold change: -0.28

### 3.4.3.8 Regulation of translation and transcription proteins

Of the proteins annotated to be involved in the translation of mitochondria, 98 were quantified in the meta-analysis results and 30 were identified as differentially expressed.

Mitochondrial ribosome component proteins (MRPs) featured heavily in this set of annotated proteins, as 60 distinct components were quantified - of which more than a third were found to be differentially expressed. The mammalian mitochondria ribosome consists of the small 28S subunit, which contains 30 proteins, and the large 39S subunit, which contains 48 proteins. Of the large subunit proteins (MRPLs) and small subunit proteins (MRPSs), 38 and 22 distinct components were quantified in the results respectively. Despite the wide profiling of ribosomal proteins, depth was limited as only 25 (of which six of each subunit were significantly differentially expressed) components were quantified in five or more datasets ($K \geq 5$).

Small mitoribosome subunit MRPS27 (Q92552; $0.32 \pm 0.04$, $K=7$) was one of the upregulated proteins that demonstrated the lowest q values in the meta-analysis results ($q < 3e-10$). This protein is also part of the Pentatricopeptide repeat domain protein family (PPR), which are RNA-binding proteins that have an essential role to play in specifically controlling mitochondrial gene expression in humans through post-transcriptional control of RNA [398] [399]. All seven members of the PPR family are localised to the mitochondria, as they only modulate organelle DNA expression [398]. Some of the other proteins in this family include LRPPRC (Leucine Rich Pentatricopeptide Repeat Containing; P42704), PTCD1 (Pentatricopeptide Repeat Domain 1; O75127), and PTCD3 (Q96EY7), all of which were identified in the meta-analysis results [399]. Despite being part of a family of transcriptional regulators, its effect on global mitochondrial protein expression is mostly thought to be due to its effect on translation. Its central role in the translation of mitochondrial genes is highlighted by the findings of a knockout analysis in human cell lines, which indicate that the abundance of mitochondrial-encoded subunits across all ETC complexes is severely depleted in conditions of MPPS27 absence. This was observed alongside a lack of an effect of MRPS27 knockout on overall abundance and organisation of the small ribosomal subunit [398]; this suggests that this

protein may be necessary to the process of mitochondrial protein translation, which also helps explain the strong evidence for its upregulation in the meta-analysis results. Another significantly upregulated protein with known associations outside the mitoribosome is MRPS36 (P82909; 0.19 ± 0.04), which has recently been reported as being the fourth component of the 2-oxoglutarate dehydrogenase enzyme complex (OGDHC) that is involved in the TCA cycle [400]. This may help explain its regulation with exercise training compared to other mitoribosome subunits.

Of the regulated MRPL components, only MRPL12 (P52815; 0.27 ± 0.1, K=6) demonstrated functional diversity or any specific relevant findings in the literature. This protein has previously been shown to play a role in stimulating transcription activity by binding to the RNA polymerase protein *in vitro* and is a potential coordinative link between transcription and translation regulation [401]. Indeed, it is the only MRP protein that is also annotated as a transcription regulatory protein in Mitocarta 3.0 [1]. Furthermore, regulation of MRPL12 gene expression is controlled by the well-studied transcription factor p53, which is a key activator of the exercise-responsive mitochondrial biogenesis signalling pathway [402, 403] and provides more context for its observed increase in the meta-analysis results.

Beside the MRP component proteins, a few other translation regulatory proteins were identified as significant in the meta-analysis results. The aforementioned PPR family proteins PTCD3 (0.2 ± 0.03) and LRPPRC (0.18 ± 0.03) both demonstrated strong evidence for exercise-training dependent regulation ($q < 5e\text{-}5$). The LRPPRC protein is implicated in multiple functions across nuclear and mitochondrial encoded transcripts, including export of nuclear mRNAs to the cytosol, interaction with PGC1-α to upregulate mtDNA gene expression, and contributing to the maintenance of mtRNA stability (i.e., protecting against degradation) through a complex with the stem-loop-interacting RNA binding protein (SLIRP; Q9GZT3)

(0.24 ± 0.1) [404-406]. The SLIRP protein has been shown to coprecipitate with LRPPRC [407] [406], confirming their association in a complex, which may help explain its concomitant upregulation (q = 0.051) in the meta-analysis results.

Finally, translation initiation and elongation factors, mtIF2 (Translation Initiation Factor IF-2, P46199; 0.32 ± 0.07, K=3) and TUFM (Elongation Factor Tu, P49411; 0.26 ± 0.07), respectively, were also observed to be upregulated in the meta-analysis results. The translation activator of cytochrome oxidase I (TACO1, Q9BSH4) protein, while not significant in the meta-analysis, demonstrated significant upregulation in three different datasets (and approached significance at $p < 0.02$ for Schild_2015). All three of these proteins are necessary components of the mtDNA translation process and have been specifically investigated in the context of exercise training in animals [408] [409] but not humans. Endurance training for four weeks was shown to cause an increased abundance of TUFM and TACO1 proteins but not of mtIF2 [409]. Interestingly, no effect of exercise training on the abundances of ribosomal subunits was observed either [409]. The protein content of mtIF2, on the other hand, has been shown to increase in abundance following two weeks of resistance training [408]. However, this change in mtIF2 content was not accompanied by increased abundance of content markers of ETC complexes IV and V; this suggests an absence of a mitochondria biogenesis response despite the increase in mtIF2 protein expression[408].

The only transcription regulatory protein identified in the meta-analysis results was TFAM (Q00059; 0.12 ± 0.04, K=5), and it was found to be significantly upregulated. The upregulation of TFAM gene expression has been canonically used as a marker for assessing mitochondria biogenesis [410], and endurance exercise studies have also demonstrated significant increases in TFAM protein content following training [411] [124] [120]. Despite this evidence for increased TFAM protein abundance from western blot quantification, there

was only a small overall effect size and TFAM was not observed to be differentially expressed in any of the individual datasets. Even in the reported results from Deshmukh et.al (2021) [19] (TFAM was excluded from *Deshmukh_2021* due to not meeting the threshold of missingness), which followed a 12-week continuous endurance training protocol, regulation of TFAM protein abundance did not reach significance. It is unclear why a larger and more consistent effect on TFAM protein abundance was not observed in the data especially considering measures of mitochondria biogenesis (increased overall protein intensity, mitochondria protein enrichment, CS activity as reported in the associated literature) were found significant in multiple datasets.

### *3.4.4 Enrichment of Meta-analysis results and effects of Exercise volume*

Pathway enrichment analysis of the 200 differentially expressed proteins with $K \geq 5$ from the meta-analysis results was performed on Reactome and all GO category terms.

Reactome enrichment yielded significant terms related to the TCA cycle, OXPHOS, fatty acid β-oxidation, mitochondrial biogenesis, and cristae formation, among others. The Figure 3:52 shows a comprehensive profile of the proteins belonging within these pathways, along with their fold changes. Similarly, GO enrichment was also significant for the TCA cycle, lipid oxidation, OXPHOS related terms specifically for CI, CIV, and CV, and respiratory complex chain assembly. The Figure 3:53 demonstrates a subset of these enriched GO terms. A Reactome Gene Set Enrichment Analysis (GSEA) was also performed on the list of meta-analysis analysis results for the 493 proteins with K >= 5 (ordered by effect size from high to low). Beyond the already identified Reactome terms, ‘Mitochondrial Translation’ was found to be significantly enriched with a positive normalised enrichment score.

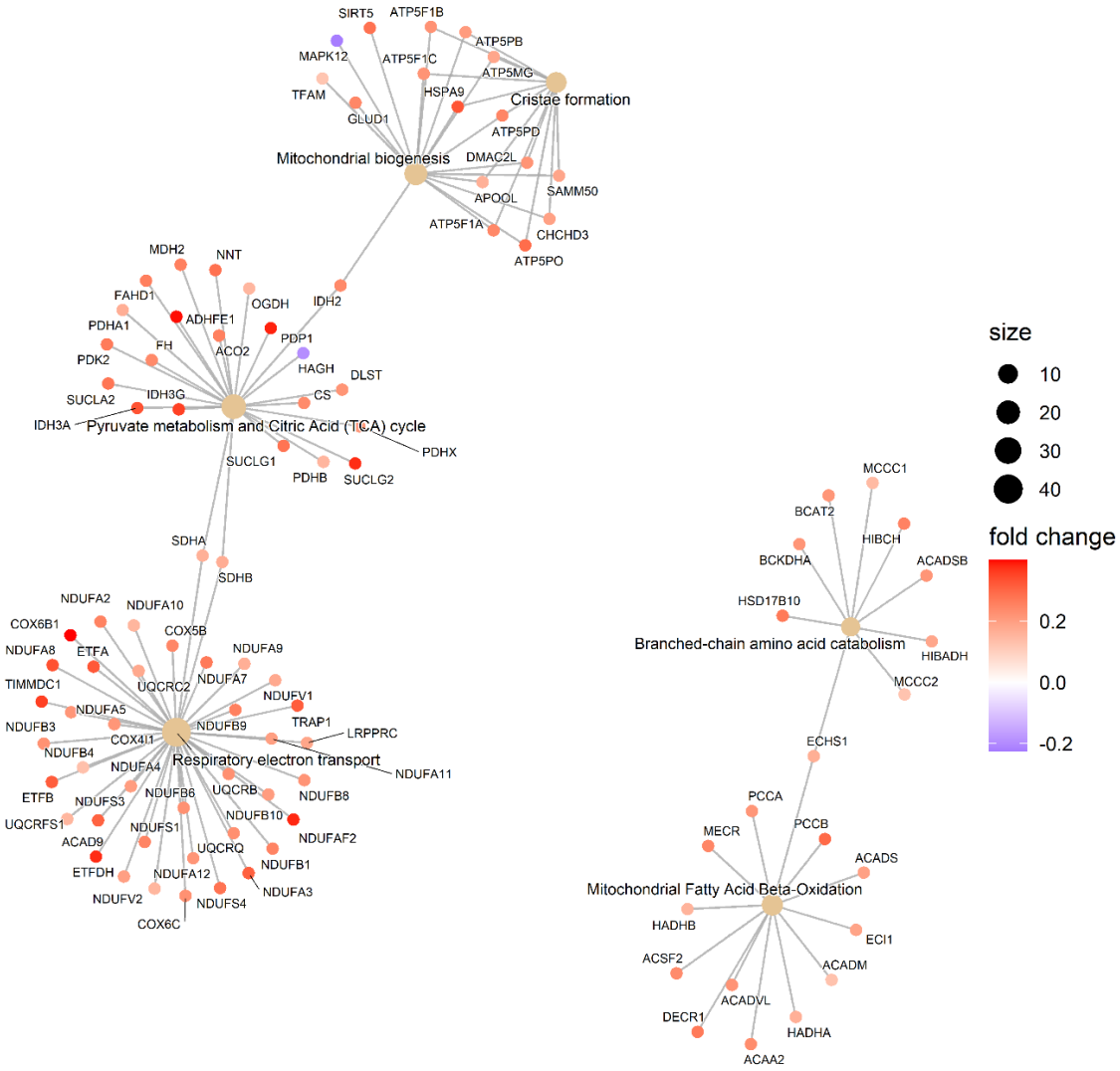


Figure 3:52 Visualisation of enriched Reactome terms and associated proteins from the significant results of the meta-analysis.

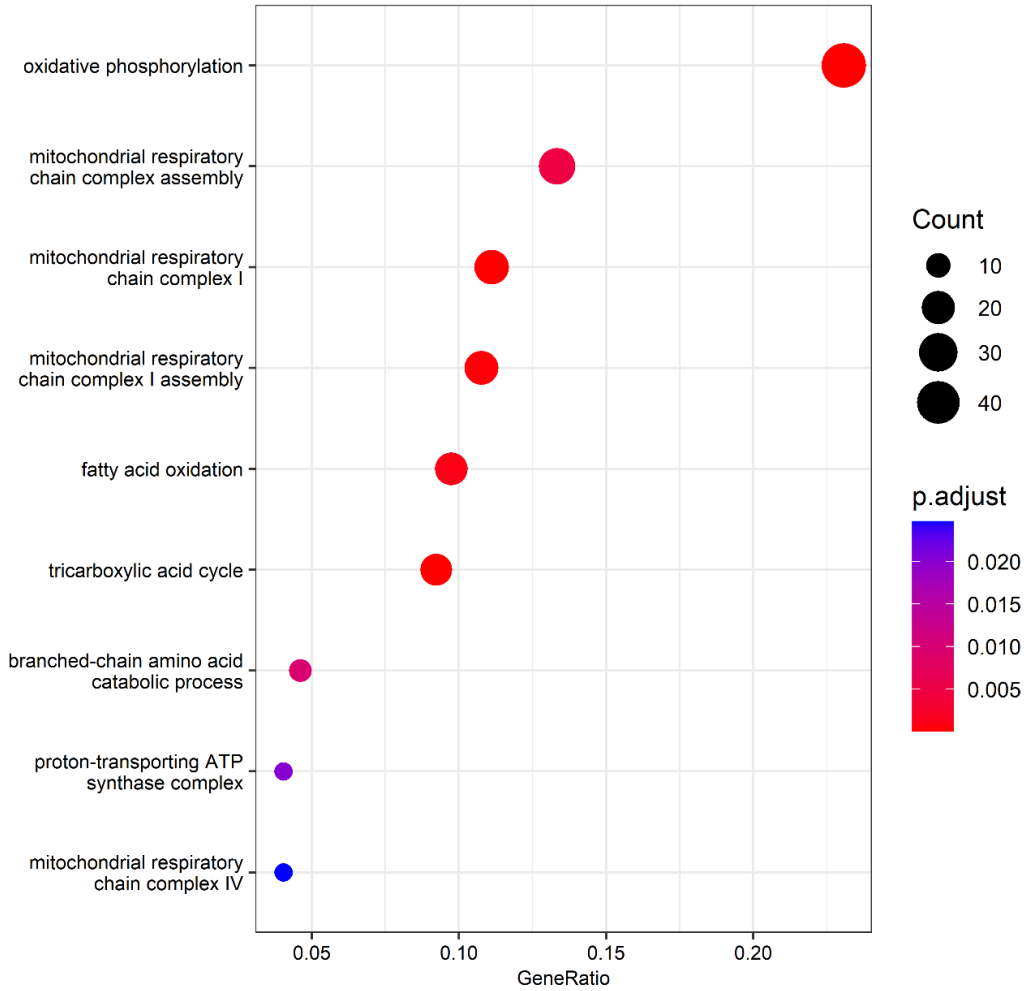


Figure 3:53 Summary of enriched GO terms from significant results of the meta-analysis.

No distinct metabolic pathway or functional terms that hadn't previously been identified in at least one of the individual datasets were found enriched in the meta-analysis results. However, the increased power of the meta-analysis design is still apparent in the results. For instance, Reactome terms related to mitochondria biogenesis, and fatty acid β-oxidation were only enriched in one or two of the eight independent datasets. Due to the limited power of differential expression analysis of individual exercise studies, a proportion of this subset of proteins may not have been detected as differentially expressed in individual datasets despite the underlying fold changes. For example, 41 of the 200 (approximately 20%) significant proteins (with $K \geq 5$) did not demonstrate significant differential expression in any single

dataset. Aggregating data using a meta-analysis design was able to demonstrate training-dependent regulation of these pathways with greater statistical robustness than may have been possible or available before.

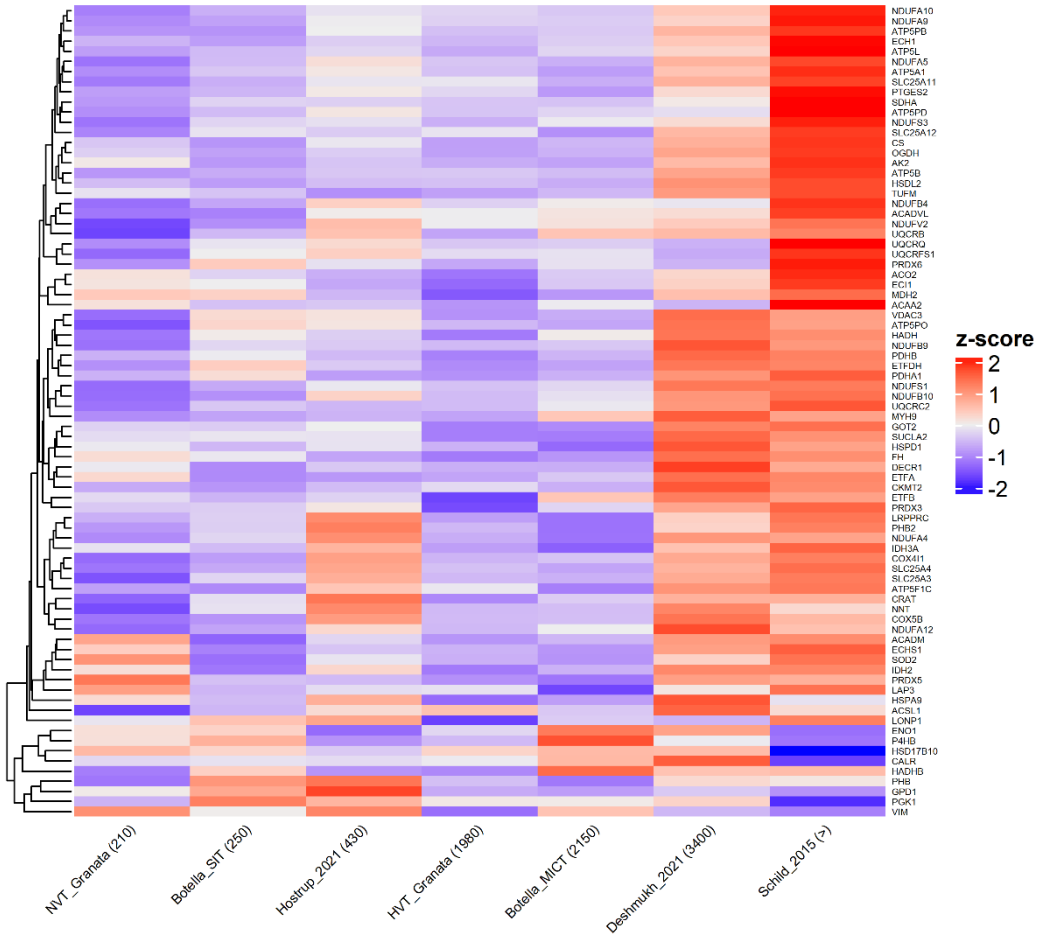


Figure 3:54 A heatmap of z-scored log2 fold changes for significant meta-analysis results with $K \geq 7$. The columns are ordered based on ordinal ranking of exercise exposure/volume which is denoted for each study.

The trend of individual fold change across the different datasets for meta-analysis significant proteins was also briefly explored. As can be observed from the Figure 3:54[12], there is expectedly a trend toward larger $\log_2$ fold changes with increasing exercise volume. The two largest datasets in terms of underlying exercise volume/exposure consistently demonstrate the

[12] Popov_2020 was excluded from this analysis of exercise volume on study-wise log2 fold changes.

largest $log_2$ fold changes. While *Botella_MICT* and *HVT_Granata* are both based on training protocols with larger overall exercise volume compared to *Hostrup_2021*, they did not consistently demonstrate larger fold changes compared to the latter. As noted before, it is suspected that while exercise volume for *Botella_MICT* was relatively high, the average exercise intensity was the lowest (~50% $VO_{2max}$) among all the included studies which may have contributed to blunted the effect-size of exercise adaptations. However, contribution of technical variability/error cannot be ruled out. The exercise volume in *HVT_Granata* was approximately 4 times that of *Hostrup_2021* and the exercise intensity and modality was comparable, yet the exercise response is seen to be similar/slightly lower for *HVT_Granta.* It is suspected that due to the concentrated nature of training in *HVT_Granata* (training twice a day, 7 days a week, for 3 weeks), the proteomic adaptations may not have been fully realised at the time of the biopsy sampling following the end of training. Observed fold changes may have been higher if the exercise exposure was spread over a longer duration, or if biopsies were sampled later, in order to allow the increased protein synthesis and accumulation, in response to exercise training, to manifest completely. As before, however, this is not a conclusive explanation of the observed trend and the contribution of factors of technical variability/error in potentially supressing effect-sizes cannot be ruled out.

# 4 CHAPTER FOUR

## Conclusions and Future Direction

The overall aim of this research was to compare the differential expression results of existing exercise-training-derived MS proteomics data in humans and perform the first-of-its-kind meta-analysis of this data. Tools of enrichment analysis and network analysis were applied to the differential expression and meta-analysis results to capture potential functional and metabolic implications of the changes. Finally, the meta-analysis results for proteins in key mitochondria pathways were qualitatively discussed with respect to existing cellular or physiological evidence.

## 4.1 Summary of key findings

### *4.1.1 Comparison of differential expression results across the included datasets*

- A total of 778 unique mitochondrial proteins were identified across the eight datasets included in this study. Of these, 113 proteins were commonly identified across all datasets. The DIA-based datasets achieved a deeper proteome coverage compared to the DDA-based datasets, and this was observed to be related, in part, to the former's capacity to capture more low-abundance proteins.
- A total of 365 proteins were identified as differentially expressed in at least one of the datasets. An enrichment analysis on this set of proteins identified the key pathways of OXPHOS, TCA cycle, fatty acid oxidation, and amino acid metabolism to be significant.
- A low degree of overlap was observed among the set of proteins identified as differentially expressed across the individual datasets, and no single protein was identified as differentially expressed in all eight datasets. The highest degree of intersection for common differentially expressed proteins was six datasets, and this was limited to only three proteins (TIMM44, ECI2, SLC25A3).
- Despite the low overlap for common, differentially-expressed, individual proteins, comparing the enrichment profiles of the different datasets revealed a high degree of overlap for terms related to the OXPHOS complexes, the TCA cycle, and fatty acid oxidation. The *HVT_Granata* dataset uniquely identified terms related to mitochondria translation and gene expression.
- The effects of *in silico* mitochondria normalisation, as described in [81], on differential expression profiles of individual datasets was also analysed and reported. A pattern in

stoichiometry of proteome adaptation with respect to exercise exposure was observed, with the early adaptive response focused on increasing TCA cycle protein abundance and a *deprioritisation* of the OXPHOS system (stoichiometrically reduced abundance). With further exercise exposure, the OXPHOS system proteins were then observed to be upregulated/prioritised and the non-stoichiometric adaptations to the TCA cycle observed previously were lost. Finally, with prolonged exercise exposure, stoichiometry with respect to mitochondria proteins in all observed functional classes is regained.

### *4.1.2 Meta-analysis Results*

- A total of 254 proteins were found to be differentially expressed at the FDR threshold of 0.05. Of these, 228 were observed to be upregulated and 25 were observed to be downregulated. Of the 254 significant proteins, 200 were quantified in five or more datasets ($K \geq 5$).
- Multiple subunits from each OXPHOS complex were found to be differentially expressed.
    - Of the 43 identified Complex I subunits, 22 were found to be significant, with a mean $\log_2$ fold change of 0.26.
    - All four CII subunits were identified in the results; however, only SDHA and SDHB, the only two ETC proteins also involved in the TCA cycle, were found to be significant with a mean $\log_2$ fold change of 0.17.
    - All ten CIII subunits proteins were identified in the meta-analysis results, and five were found to be significantly upregulated with a mean log fold-change of 0.24. All three ETF complex subunit proteins that supply

electrons to CIII through Coenzyme Q (ETFA, ETFB, and ETFDH), were also observed to be upregulated.

    - 15 CIV subunits were identified in the meta-analysis results, and six were found to be significantly upregulated with a mean $\log_2$ fold change of 0.28.
    - Five of the 11 identified CV subunits were found to be significantly upregulated with a mean fold-change of 0.24.

- Major enzymes involved in all intermediate TCA cycle reactions, starting from the conversion of pyruvate until the formation of oxaloacetate, were significantly upregulated.
- Mitochondrial proteins involved in all major steps of the beta-oxidation pathway, which includes fatty acid import through the *CPT* proteins, oxidation through the *ACAD* proteins, hydration (*HADH* and *HSD17B10*), and thiolysis (*ACAA2*), were found to be upregulated.
- Multiple proteins implicated in regulation of mitochondrial dynamics were also observed to be differentially regulated in response to exercise training. Two key fusion-related proteins (OPA1 and MTCH2) were observed to be significantly upregulated; however, no significant changes in fission-related proteins were observed. Mitochondrial proteins within cellular apoptosis pathways (*AIFM*, *GHITM*, *ENDOG*) were also found to be upregulated; however, these proteins have also been shown to be essential to other functional pathways of mitochondria biogenesis according to the literature. Finally, protein components of the MICOS complex (*APOOL*, *CHCHD3*) were also upregulated.
- Enzymatic ROS mutating antioxidant proteins (*SOD2*, PRDX family members III and V, and reactive aldehyde species detoxifying enzyme *ALDH1B1* and *AKR7A2*)

were found to be upregulated. Non-enzymatic protein GSTK1, the only mitochondrial glutathione conjugating protein, was also observed to be upregulated.

- Subunit proteins from all components of the import and assembly pathway were found to be upregulated. This includes the import through the OM by the TOM complex (*TOMM20*), the assembly of OM localised proteins through the SAM complex (*SAMM50*), the modification and 'trapping' of IMS localised proteins (*AIFM1*), inner membrane insertion through the TIMM22 complex (*AGK*), and matrix import through the PAM complex (*GREPL1*, *TIMM44, HSPA9*). Other key findings include upregulation of OXA1L, integral to the assembly of mtDNA-encoded OXPHOS proteins, and both subunits of the chaperone complex (*HSPE1* and *HSPD1)*.
- Key proteostasis proteins (*STOML2*, *PHB1* and *PHB2*; the prohibitin complex), which have all been shown to interact with each other to promote pro-fusion mitochondria dynamics and biogenesis, were also found to be upregulated.
- Multiple members of the SLC25 protein family implicated in ADP/ATP shuttling (SLC25A5), carnitine/acyl-carnitine transport (SLC25A20), the malate-aspartate shuttle complex (SLC25A11, SLC25A12, SLCA25A13), and phosphate ion import (SLC25A3) were found to be differentially expressed. The SLC family protein MTCH2 demonstrated one of the highest fold changes and multiple findings of significance across individual datasets. However, the functional relevance of this adaptation is unclear and warrants further investigation.
- Sixty mitoribosome protein subunits were quantified in the data, of which 23 were found to be differentially expressed. However, the depth of coverage was limited as only 25 protein subunits were observed with a $K \geq 5$. Other translation-related differentially expressed proteins included the PPR family proteins LRPPRC and

PTCD3, and the initiation and elongation factors mtIF2 and TUFM respectively. The well-known transcription activator TFAM was the only transcription-related protein observed to differentially expressed. However, it only demonstrated a mild effect size and was not found to be significant in any individual dataset, which contrasts with the consistent evidence from western blot studies for its increased relative abundance following exercise training.

- Reactome and GO enrichment analysis of significant findings from proteins with a $K \geq 5$ expectedly revealed enrichment of terms related to the OXPHOS complexes and their assembly, the TCA cycle, fatty acid oxidation, amino acid metabolism, and mitochondria biogenesis. Of the OXPHOS complexes, only CI, CIV and CV demonstrated specific enrichment.

## 4.2 Conclusion and Future Directions

High-throughput mass spectrometry is increasingly gaining popularity as the most prevalent data acquisition method to analyse protein-level changes in the field of exercise physiology [215] [214]. In anticipation of, and preparation for, the growing amounts of exercise training-related data that is starting to be available, this study aimed to develop uniform pipelines to re-analyse and aggregate secondary MS data. Further, there has been a demand in the academic community for revisiting existing proteomics datasets to compare, contrast, and validate previous findings [214]. Addressing this need, this study is the first and only attempt, so far, at a meta-analysis of gel-free MS-derived proteomics data. The findings of this research

hope to elucidate the agreement and divergence within existing data and help confer greater statistical power on findings of novel exercise training-regulated proteins.

The differential expression analysis and meta-analysis results expectedly revealed comprehensive regulation of oxidative metabolic pathways. While most existing research has relied on quantification of representative subunits or component proteins to signify metabolic adaptation, aggregation of wide proteome-profiling data demonstrates adaptations to protein abundance at all major steps in the fatty acid oxidation, TCA cycle, and OXPHOS pathways. Beyond metabolically relevant proteins, the meta-analysis identified key proteins within functions of antioxidant defence, protein import, mitochondria biogenesis and fusion, and small molecule transport, as well as other pathways. The functional centrality and significance of most of the identified proteins within their respective pathways was supported by findings in the literature from cell lines or animal models. Very few of the meta-analysis-identified proteins had been previously investigated in the context of exercise training in either human or animal models using traditional protein quantification methods, which highlights the utility of MS data acquisition and the application of meta-analytic methods. While this analysis may be limited in statistical power due to the relatively small number of studies available, especially compared to the existing transcriptomic meta-analysis work in the field [228] [227], numerous 'novel' proteins demonstrated strong and consistent evidence for exercise training-dependant regulation. Therefore, this research can be a resource for the scientific community to mine and can help generate hypotheses for targeted validation studies.

As more data becomes publicly available in the future, this analysis can and should be supplemented with more input to strengthen and refine existing findings. Relatedly, as the sample size grows, these findings can form the basis of an interactive web application, akin to *extrameta* [227] or *metamex* [228], allowing researchers worldwide to query their protein of interest. For such analysis to improve further and truly benefit from the expected growth in MS

proteomics investigations, it is imperative that higher, more uniform standards are followed in the sharing and maintenance of raw data files. The lack of shared well-annotated raw data, or any raw data at all, was one of the main reasons MS spectra files were not re-analysed in this study and more published studies could not be included in the meta-analysis. Applying uniform processing steps on raw spectra files might have provided greater agreement among the differential expression results of the included data and this should be investigated in the future. Finally, as this meta-analysis provides a degree of consensus around the adaptation of mitochondrial proteome with exercise training, it may be worthwhile for future research to contextualise these findings against the results of gene expression analysis and characterise the dynamics of protein regulation.

## References


1. Rath, S., et al., *MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations.* Nucleic Acids Res, 2021. **49**(D1): p. D1541-d1547.
2. Eaglesfield, R. and K. Tokatlidis, *Targeting and insertion of membrane proteins in mitochondria.* Frontiers in Cell and Developmental Biology, 2021. **9**: p. 803205.
3. Stagge, F., *Untersuchung der Heterogenität submitochondrialer Proteinverteilungen mit hochauflösender Mikroskopie*. 2014, Georg-August-Universität Göttingen.
4. Nesvizhskii, A.I. and R. Aebersold, *Interpretation of shotgun proteomic data.* Molecular & cellular proteomics, 2005. **4**(10): p. 1419-1440.
5. Egan, B. and J.R. Zierath, *Exercise metabolism and the molecular regulation of skeletal muscle adaptation.* Cell Metab, 2013. **17**(2): p. 162-84.
6. Ioannidis, J.P., N.A. Patsopoulos, and E. Evangelou, *Uncertainty in heterogeneity estimates in meta-analyses.* Bmj, 2007. **335**(7626): p. 914-6.
7. Spinelli, J.B. and M.C. Haigis, *The multifaceted contributions of mitochondria to cellular metabolism.* Nature cell biology, 2018. **20**(7): p. 745-754.
8. Granata, C., N.A. Jamnick, and D.J. Bishop, *Training-Induced Changes in Mitochondrial Content and Respiratory Function in Human Skeletal Muscle.* Sports Med, 2018. **48**(8): p. 1809-1828.
9. McBride, H.M., M. Neuspiel, and S.J.C.b. Wasiak, *Mitochondria: more than just a powerhouse.* 2006. **16**(14): p. R551-R560.
10. Oliveira, A.N., et al., *Exercise Is Muscle Mitochondrial Medicine.* Exerc Sport Sci Rev, 2021. **49**(2): p. 67-76.
11. Johannsen, D.L. and E. Ravussin, *The role of mitochondria in health and disease.* Current Opinion in Pharmacology, 2009. **9**(6): p. 780-786.
12. Hesselink, M.K., V. Schrauwen-Hinderling, and P. Schrauwen, *Skeletal muscle mitochondria as a target to prevent or treat type 2 diabetes mellitus.* Nat Rev Endocrinol, 2016. **12**(11): p. 633-645.
13. Frangos, S.M., D.J. Bishop, and G.P. Holloway, *Revisiting the contribution of mitochondrial biology to the pathophysiology of skeletal muscle insulin resistance.* Biochem J, 2021. **478**(21): p. 3809-3826.
14. Calderón, J.C., P. Bolaños, and C. Caputo, *The excitation-contraction coupling mechanism in skeletal muscle.* Biophysical reviews, 2014. **6**(1): p. 133-160.
15. Frontera, W.R. and J. Ochala, *Skeletal Muscle: A Brief Review of Structure and Function.* Calcified Tissue International, 2015. **96**(3): p. 183-195.
16. Schiaffino, S. and C. Reggiani, *Fiber types in mammalian skeletal muscles.* Physiological reviews, 2011. **91**(4): p. 1447-1531.
17. Biral, D., et al., *Myosin heavy chain composition of single fibres from normal human muscle.* Biochemical Journal, 1988. **250**(1): p. 307.
18. Murgia, M., et al., *Single muscle fiber proteomics reveals fiber-type-specific features of human muscle aging.* Cell reports, 2017. **19**(11): p. 2396-2409.
19. Deshmukh, A.S., et al., *Deep muscle-proteomic analysis of freeze-dried human muscle biopsies reveals fiber type-specific adaptations to exercise training.* Nat Commun, 2021. **12**(1): p. 304.
20. Reisman, E., et al., *Single-fibre proteomics reveals a greater fibre-type specific mitochondrial remodelling with moderate than high-intensity training.* bioRxiv, 2022: p. 2022.10. 23.512956.
21. Wilkinson, D.J., M. Piasecki, and P.J. Atherton, *The age-related loss of skeletal muscle mass and function: Measurement and physiology of muscle fibre atrophy and muscle fibre loss in humans.* Ageing research reviews, 2018. **47**: p. 123-132.
22. Di Meo, S., S. Iossa, and P. Venditti, *Skeletal muscle insulin resistance: role of mitochondria and other ROS sources.* Journal of Endocrinology, 2017. **233**(1): p. R15-R42.

23. Schutz, Y., *Protein turnover, ureagenesis and gluconeogenesis.* International Journal for Vitamin and Nutrition Research, 2011. **81**(2): p. 101.
24. Joanisse, S., et al., *Understanding the effects of nutrition and post-exercise nutrition on skeletal muscle protein turnover: Insights from stable isotope studies.* Clinical Nutrition Open Science, 2021. **36**: p. 56-77.
25. Sparti, A., et al., *Relationship between resting metabolic rate and the composition of the fat-free mass.* Metabolism, 1997. **46**(10): p. 1225-1230.
26. McArdle, A., A. Vasilaki, and M. Jackson, *Exercise and skeletal muscle ageing: cellular and molecular mechanisms.* Ageing research reviews, 2002. **1**(1): p. 79-93.
27. Rezuş, E., et al., *Inactivity and Skeletal Muscle Metabolism: A Vicious Cycle in Old Age.* International Journal of Molecular Sciences, 2020. **21**(2): p. 592.
28. Hanach, N.I., F. McCullough, and A. Avery, *The Impact of Dairy Protein Intake on Muscle Mass, Muscle Strength, and Physical Performance in Middle-Aged to Older Adults with or without Existing Sarcopenia: A Systematic Review and Meta-Analysis.* Advances in Nutrition, 2019. **10**(1): p. 59-69.
29. Ringholm, S., et al., *Impact of Aging and Lifelong Exercise Training on Mitochondrial Function and Network Connectivity in Human Skeletal Muscle.* The Journals of Gerontology: Series A, 2022. **78**(3): p. 373-383.
30. Hunter, G.R., J.P. McCarthy, and M.M. Bamman, *Effects of resistance training on older adults.* Sports medicine, 2004. **34**: p. 329-348.
31. Fernández-Vizarra, E., et al., *Tissue-specific differences in mitochondrial activity and biogenesis.* Mitochondrion, 2011. **11**(1): p. 207-213.
32. Smith, A.G., et al., *A Curated Collection of Human Mitochondrial Proteins—the Integrated Mitochondrial Protein Index (IMPI).* Available at SSRN 4042282.
33. Schenkel, L.C. and M. Bakovic, *Formation and regulation of mitochondrial membranes.* Int J Cell Biol, 2014. **2014**: p. 709828.
34. Kühlbrandt, W., *Structure and function of mitochondrial membrane protein complexes.* BMC Biology, 2015. **13**(1): p. 89.
35. Ahmad, M., A. Wolberg, and C.I. Kahwaji, *Biochemistry, electron transport chain.* 2018.
36. Eramo, M.J., et al., *The 'mitochondrial contact site and cristae organising system' (MICOS) in health and human disease.* The Journal of Biochemistry, 2019. **167**(3): p. 243-255.
37. Wolf, D.M., et al., *Individual cristae within the same mitochondrion display different membrane potentials and are functionally independent.* The EMBO Journal, 2019. **38**(22): p. e101056.
38. Letts, J.A., K. Fiedorczuk, and L.A. Sazanov, *The architecture of respiratory supercomplexes.* Nature, 2016. **537**(7622): p. 644-648.
39. Schägger, H. and K. Pfeiffer, *Supercomplexes in the respiratory chains of yeast and mammalian mitochondria.* The EMBO journal, 2000. **19**(8): p. 1777-1783.
40. Acin-Perez, R. and J.A. Enriquez, *The function of the respiratory supercomplexes: The plasticity model.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2014. **1837**(4): p. 444-450.
41. Genova, M.L. and G. Lenaz, *Functional role of mitochondrial respiratory supercomplexes.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2014. **1837**(4): p. 427-443.
42. Vartak, R., C.A.-M. Porras, and Y. Bai, *Respiratory supercomplexes: structure, function and assembly.* Protein & Cell, 2013. **4**(8): p. 582-590.
43. Gold, V.A.M., et al., *Visualizing active membrane protein complexes by electron cryotomography.* Nature Communications, 2014. **5**(1): p. 4129.
44. Kukat, C., et al., *Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA.* Proceedings of the National Academy of Sciences, 2011. **108**(33): p. 13534-13539.
45. O'Brien, T.W., *Properties of Human Mitochondrial Ribosomes.* IUBMB Life, 2003. **55**(9): p. 505-513.
46. Amunts, A., et al., *The structure of the human mitochondrial ribosome.* Science, 2015. **348**(6230): p. 95-98.

47. Bo, H., Y. Zhang, and L.L. Ji, *Redefining the role of mitochondria in exercise: a dynamic remodeling.* Annals of the New York Academy of Sciences, 2010. **1201**(1): p. 121-128.
48. Glancy, B., et al., *Mitochondrial reticulum for cellular energy distribution in muscle.* Nature, 2015. **523**(7562): p. 617-620.
49. Kirkwood, S.P., E.A. Munn, and G.A. Brooks, *Mitochondrial reticulum in limb skeletal muscle.* American Journal of Physiology-Cell Physiology, 1986. **251**(3): p. C395-C402.
50. Milner, D.J., et al., *Desmin Cytoskeleton Linked to Muscle Mitochondrial Distribution and Respiratory Function.* Journal of Cell Biology, 2000. **150**(6): p. 1283-1298.
51. Vincent, A.E., et al., *Mitochondrial nanotunnels.* Trends in cell biology, 2017. **27**(11): p. 787-799.
52. Vincent, A.E., et al., *Quantitative 3D mapping of the human skeletal muscle mitochondrial network.* Cell reports, 2019. **26**(4): p. 996-1009. e4.
53. Hood, D.A., et al., *Maintenance of Skeletal Muscle Mitochondria in Health, Exercise, and Aging.* Annu Rev Physiol, 2019. **81**: p. 19-41.
54. Gao, S. and J. Hu, *Mitochondrial fusion: the machineries in and out.* Trends in Cell Biology, 2021. **31**(1): p. 62-74.
55. Youle, R.J. and A.M. van der Bliek, *Mitochondrial Fission, Fusion, and Stress.* Science, 2012. **337**(6098): p. 1062-1065.
56. Westermann, B., *Bioenergetic role of mitochondrial fusion and fission.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2012. **1817**(10): p. 1833-1838.
57. Archer, S.L., *Mitochondrial dynamics—mitochondrial fission and fusion in human diseases.* New England Journal of Medicine, 2013. **369**(23): p. 2236-2251.
58. Tanaka, T., et al., *Mitochondrial dynamics in exercise physiology.* Pflugers Arch, 2020. **472**(2): p. 137-153.
59. Chen, X.J. and R.A. Butow, *The organization and inheritance of the mitochondrial genome.* Nature Reviews Genetics, 2005. **6**(11): p. 815-825.
60. Taanman, J.-W., *The mitochondrial genome: structure, transcription, translation and replication.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1999. **1410**(2): p. 103-123.
61. Wiedemann, N. and N. Pfanner, *Mitochondrial Machineries for Protein Import and Assembly.* Annual Review of Biochemistry, 2017. **86**(1): p. 685-714.
62. Becker, T., et al., *The mitochondrial import protein Mim1 promotes biogenesis of multispanning outer membrane proteins.* Journal of Cell Biology, 2011. **194**(3): p. 387-395.
63. Moro, F., et al., *Mitochondrial protein import: molecular basis of the ATP-dependent interaction of MtHsp70 with Tim44.* Journal of Biological Chemistry, 2002. **277**(9): p. 6874-6880.
64. MacKenzie, J.A. and R.M. Payne, *Mitochondrial protein import and human health and disease.* Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2007. **1772**(5): p. 509-523.
65. Bie, A.S., et al., *An inventory of interactors of the human HSP60/HSP10 chaperonin in the mitochondrial matrix space.* Cell Stress Chaperones, 2020. **25**(3): p. 407-416.
66. Bahr, T., et al., *Mitochondrial chaperones in human health and disease.* Free Radic Biol Med, 2022. **179**: p. 363-374.
67. Szczesny, B., et al., *Age-dependent deficiency in import of mitochondrial DNA glycosylases required for repair of oxidatively damaged bases.* Proceedings of the National Academy of Sciences, 2003. **100**(19): p. 10670-10675.
68. Wright, G., et al., *Oxidative stress inhibits the mitochondrial import of preproteins and leads to their degradation.* Experimental cell research, 2001. **263**(1): p. 107-117.
69. Jornayvaz, F.R. and G.I. Shulman, *Regulation of mitochondrial biogenesis.* Essays Biochem, 2010. **47**: p. 69-84.
70. Adhihetty, P.J., et al., *Plasticity of skeletal muscle mitochondria in response to contractile activity.* Experimental physiology, 2003. **88**(1): p. 99-107.
71. Popov, L.D., *Mitochondrial biogenesis: An update.* J Cell Mol Med, 2020. **24**(9): p. 4892-4899.

72. D'Souza, A.R. and M. Minczuk, *Mitochondrial transcription and translation: overview.* Essays Biochem, 2018. **62**(3): p. 309-320.
73. Yokokawa, T., et al., *Exercise-induced mitochondrial biogenesis coincides with the expression of mitochondrial translation factors in murine skeletal muscle.* Physiological Reports, 2018. **6**(20): p. e13893.
74. Arribat, Y., et al., *Distinct patterns of skeletal muscle mitochondria fusion, fission and mitophagy upon duration of exercise training.* Acta Physiologica, 2019. **225**(2): p. e13179.
75. Russell, A.P., et al., *Skeletal muscle mitochondria: a major player in exercise, health and disease.* Biochim Biophys Acta, 2014. **1840**(4): p. 1276-84.
76. Lundby, C. and R.A. Jacobs, *Adaptations of skeletal muscle mitochondria to exercise training.* Experimental Physiology, 2016. **101**(1): p. 17-22.
77. Zoll, J., et al., *Quantitative and qualitative adaptation of skeletal muscle mitochondria to increased physical activity.* Journal of Cellular Physiology, 2003. **194**(2): p. 186-193.
78. Pesta, D., et al., *Similar qualitative and quantitative changes of mitochondrial respiration following strength and endurance training in normoxia and hypoxia in sedentary humans.* American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 2011. **301**(4): p. R1078-R1087.
79. Sebastián, D., et al., *Mitofusin 2 (Mfn2) links mitochondrial and endoplasmic reticulum function with insulin signaling and is essential for normal glucose homeostasis.* Proceedings of the National Academy of Sciences, 2012. **109**(14): p. 5523-5528.
80. Bishop, D.J., C. Granata, and N. Eynon, *Can we optimise the exercise training prescription to maximise improvements in mitochondria function and content?* Biochimica et Biophysica Acta (BBA) - General Subjects, 2014. **1840**(4): p. 1266-1275.
81. Granata, C., et al., *High-intensity training induces non-stoichiometric changes in the mitochondrial proteome of human skeletal muscle without reorganisation of respiratory chain content.* Nature communications, 2021: p. 2021.02.19.431993.
82. Fealy, C.E., et al., *Skeletal muscle mitochondrial network dynamics in metabolic disorders and aging.* Trends in Molecular Medicine, 2021. **27**(11): p. 1033-1044.
83. Wai, T. and T. Langer, *Mitochondrial Dynamics and Metabolic Regulation.* Trends in Endocrinology & Metabolism, 2016. **27**(2): p. 105-117.
84. Iqbal, S. and D.A. Hood, *Oxidative stress-induced mitochondrial fragmentation and movement in skeletal muscle myoblasts.* American Journal of Physiology-Cell Physiology, 2014. **306**(12): p. C1176-C1183.
85. Bach, D., et al., *Mitofusin-2 determines mitochondrial network architecture and mitochondrial metabolism: a novel regulatory mechanism altered in obesity.* Journal of Biological Chemistry, 2003. **278**(19): p. 17190-17197.
86. Greggio, C., et al., *Enhanced Respiratory Chain Supercomplex Formation in Response to Exercise in Human Skeletal Muscle.* Cell Metabolism, 2017. **25**(2): p. 301-311.
87. den Brave, F. and T. Becker, *Supercomplex formation boosts respiration.* EMBO reports, 2020. **21**(12): p. e51830.
88. Nunnari, J. and A. Suomalainen, *Mitochondria: In Sickness and in Health.* Cell, 2012. **148**(6): p. 1145-1159.
89. Hirst, J., *Open questions: respiratory chain supercomplexes—why are they there and what do they do?* BMC biology, 2018. **16**(1): p. 1-4.
90. Scherz-Shouval, R. and Z. Elazar, *ROS, mitochondria and the regulation of autophagy.* Trends Cell Biol, 2007. **17**(9): p. 422-7.
91. Starkov, A.A., *The role of mitochondria in reactive oxygen species metabolism and signaling.* Annals of the New York Academy of Sciences, 2008. **1147**(1): p. 37-52.
92. Echtay, K.S., *Mitochondrial uncoupling proteins--what is their physiological role?* Free Radic Biol Med, 2007. **43**(10): p. 1351-71.
93. Trewin, A.J., B.J. Berry, and A.P. Wojtovich, *Exercise and Mitochondrial Dynamics: Keeping in Shape with ROS and AMPK.* Antioxidants (Basel), 2018. **7**(1).
94. Xiao, B., et al., *Reactive oxygen species trigger Parkin/PINK1 pathway–dependent mitophagy by inducing mitochondrial recruitment of Parkin.* Journal of Biological Chemistry, 2017. **292**(40): p. 16697-16708.

95. McArdle, A., et al., *Role of mitochondrial superoxide dismutase in contraction-induced generation of reactive oxygen species in skeletal muscle extracellular space.* American Journal of Physiology-Cell Physiology, 2004. **286**(5): p. C1152-C1158.
96. Chistiakov, D.A., et al., *Mitochondrial aging and age-related dysfunction of mitochondria.* BioMed research international, 2014. **2014**.
97. Schrauwen, P. and M.K.C. Hesselink, *Oxidative Capacity, Lipotoxicity, and Mitochondrial Damage in Type 2 Diabetes.* Diabetes, 2004. **53**(6): p. 1412-1417.
98. Csordás, G.r., et al., *Structural and functional features and significance of the physical linkage between ER and mitochondria.* The Journal of cell biology, 2006. **174**(7): p. 915-921.
99. Sun, Y. and S. Ding, *ER–Mitochondria Contacts and Insulin Resistance Modulation through Exercise Intervention.* International Journal of Molecular Sciences, 2020. **21**(24): p. 9587.
100. Helle, S.C.J., et al., *Organization and function of membrane contact sites.* Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 2013. **1833**(11): p. 2526-2541.
101. Vance, J.E., *Molecular and cell biology of phosphatidylserine and phosphatidylethanolamine metabolism.* Progress in nucleic acid research and molecular biology, 2003. **75**: p. 69-111.
102. Arruda, A.P., et al., *Chronic enrichment of hepatic endoplasmic reticulum–mitochondria contact leads to mitochondrial dysfunction in obesity.* Nature medicine, 2014. **20**(12): p. 1427-1435.
103. Vilchez, D., I. Saez, and A. Dillin, *The role of protein clearance mechanisms in organismal ageing and age-related diseases.* Nature communications, 2014. **5**(1): p. 5659.
104. Drake, J.C. and Z. Yan, *Mitophagy in maintaining skeletal muscle mitochondrial proteostasis and metabolic health with ageing.* The Journal of Physiology, 2017. **595**(20): p. 6391-6399.
105. Jovaisaite, V. and J. Auwerx, *The mitochondrial unfolded protein response—synchronizing genomes.* Current Opinion in Cell Biology, 2015. **33**: p. 74-81.
106. Zhu, L., et al., *Mitochondrial unfolded protein response: An emerging pathway in human diseases.* Free Radical Biology and Medicine, 2021. **163**: p. 125-134.
107. Wu, Y., et al., *Multilayered Genetic and Omics Dissection of Mitochondrial Activity in a Mouse Reference Population.* Cell, 2014. **158**(6): p. 1415-1430.
108. Lawrie, R., *The activity of the cytochrome system in muscle and its relation to myoglobin.* Biochemical Journal, 1953. **55**(2): p. 298.
109. Paul, M.H. and E. Sperling, *Cyclophorase System XXIII. Correlation of Cyclophorase Activity and Mitochondrial Density in Striated Muscle.* Proceedings of the society for experimental biology and medicine, 1952. **79**(3): p. 352-354.
110. Holloszy, J.O., *Biochemical Adaptations in Muscle: EFFECTS OF EXERCISE ON MITOCHONDRIAL OXYGEN UPTAKE AND RESPIRATORY ENZYME ACTIVITY IN SKELETAL MUSCLE.* Journal of Biological Chemistry, 1967. **242**(9): p. 2278-2282.
111. Botella, J., E.S. Motanova, and D.J. Bishop, *Muscle contraction and mitochondrial biogenesis – A brief historical reappraisal.* Acta Physiologica, 2022. **235**(1): p. e13813.
112. Gould, M. and W. Rawlinson, *I. The effect of swimming on the levels of lactic dehydrogenase, malic dehydrogenase and phosphorylase in muscles of 8, 11-and 15-week-old rats.* Biochem. J., 1959. **73**: p. 41-44.
113. Hearn, G.R. and W. Wainio, *Succinic dehydrogenase activity of the heart and skeletal muscle of exercised rats.* American Journal of Physiology-Legacy Content, 1956. **185**(2): p. 348-350.
114. Rowe, G.C., et al., *PGC-1α is dispensable for exercise-induced mitochondrial biogenesis in skeletal muscle.* PloS one, 2012. **7**(7): p. e41817.
115. Geng, T., et al., *PGC-1α plays a functional role in exercise-induced mitochondrial biogenesis and angiogenesis but not fiber-type transformation in mouse skeletal muscle.* American Journal of Physiology-Cell Physiology, 2010. **298**(3): p. C572-C579.
116. Dudley, G.A., W.M. Abraham, and R.L. Terjung, *Influence of exercise intensity and duration on biochemical adaptations in skeletal muscle.* Journal of Applied Physiology, 1982. **53**(4): p. 844-850.
117. Villena, J.A., *New insights into PGC-1 coactivators: redefining their role in the regulation of mitochondrial function and beyond.* The FEBS Journal, 2015. **282**(4): p. 647-672.

118. Bergström, J., *Percutaneous needle biopsy of skeletal muscle in physiological and clinical research.* Scandinavian journal of clinical and laboratory investigation, 1975. **35**(7): p. 609-616.
119. Fiorenza, M., et al., *High-intensity exercise training enhances mitochondrial oxidative phosphorylation efficiency in a temperature-dependent manner in human skeletal muscle: implications for exercise performance.* The FASEB Journal, 2019. **33**(8): p. 8976-8989.
120. Granata, C., et al., *Mitochondrial adaptations to high-volume exercise training are rapidly reversed after a reduction in training volume in human skeletal muscle.* The FASEB Journal, 2016. **30**(10): p. 3413-3423.
121. Murias, J.M., et al., *Adaptations in Capillarization and Citrate Synthase Activity in Response to Endurance Training in Older and Young Men.* The Journals of Gerontology: Series A, 2011. **66A**(9): p. 957-964.
122. Bengtsson, J., et al., *Mitochondrial transcription factor A and respiratory complex IV increase in response to exercise training in humans.* Pflügers Archiv, 2001. **443**(1): p. 61-66.
123. Granata, C., et al., *Training intensity modulates changes in PGC-1alpha and p53 protein content and mitochondrial respiration, but not markers of mitochondrial content in human skeletal muscle.* FASEB J, 2016. **30**(2): p. 959-70.
124. Islam, H., et al., *The impact of acute and chronic exercise on Nrf2 expression in relation to markers of mitochondrial biogenesis in human skeletal muscle.* European Journal of Applied Physiology, 2020. **120**(1): p. 149-160.
125. Edgett, B.A., et al., *The effect of acute and chronic sprint-interval training on LRP130, SIRT3, and PGC-1α expression in human skeletal muscle.* Physiological Reports, 2016. **4**(17): p. e12879.
126. Combes, A., et al., *Exercise-induced metabolic fluctuations influence AMPK, p38-MAPK and Ca MKII phosphorylation in human skeletal muscle.* Physiological reports, 2015. **3**(9): p. e12462.
127. Hawley, J.A. and J.O. Holloszy, *Exercise: it's the real thing!* Nutrition Reviews, 2009. **67**(3): p. 172-178.
128. Robinson, M.M., et al., *Enhanced Protein Translation Underlies Improved Metabolic and Physical Adaptations to Different Exercise Training Modes in Young and Old Humans.* Cell Metab, 2017. **25**(3): p. 581-592.
129. Merry, T.L. and M. Ristow, *Mitohormesis in exercise training.* Free Radical Biology and Medicine, 2016. **98**: p. 123-130.
130. Hussey, S.E., et al., *Effect of exercise on the skeletal muscle proteome in patients with type 2 diabetes.* Med Sci Sports Exerc, 2013. **45**(6): p. 1069-76.
131. Wackerhage, H., *Molecular exercise physiology: an introduction*. 2014: Routledge.
132. Tanner, C.B., et al., *Mitochondrial and performance adaptations to exercise training in mice lacking skeletal muscle LKB1.* American Journal of Physiology-Endocrinology and Metabolism, 2013. **305**(8): p. E1018-E1029.
133. Dent, J.R., et al., *Muscle-specific knockout of general control of amino acid synthesis 5 (GCN5) does not enhance basal or endurance exercise-induced mitochondrial adaptation.* Molecular Metabolism, 2017. **6**(12): p. 1574-1584.
134. Chavanelle, V., et al., *Effects of high-intensity interval training and moderate-intensity continuous training on glycaemic control and skeletal muscle mitochondrial function in db/db mice.* Scientific Reports, 2017. **7**(1): p. 204.
135. Wright, D.C., et al., *Calcium induces increases in peroxisome proliferator-activated receptor γ coactivator-1α and mitochondrial biogenesis by a pathway leading to p38 mitogen-activated protein kinase activation.* Journal of Biological Chemistry, 2007. **282**(26): p. 18793-18799.
136. Moore, T.M., et al., *The impact of exercise on mitochondrial dynamics and the role of Drp1 in exercise performance and training adaptations in skeletal muscle.* Molecular metabolism, 2019. **21**: p. 51-67.
137. Yan, Z., V.A. Lira, and N.P. Greene, *Exercise training-induced regulation of mitochondrial quality.* Exercise and sport sciences reviews, 2012. **40**(3): p. 159-164.

138. Samelman, T., L. Shiry, and D. Cameron, *Endurance training increases the expression of mitochondrial and nuclear encoded cytochrome c oxidase subunits and heat shock proteins in rat skeletal muscle.* European journal of applied physiology, 2000. **83**(1): p. 22-27.
139. Ljubicic, V., et al., *Transcriptional and post-transcriptional regulation of mitochondrial biogenesis in skeletal muscle: Effects of exercise and aging.* Biochimica et Biophysica Acta (BBA) - General Subjects, 2010. **1800**(3): p. 223-234.
140. Lira, V.A., et al., *Autophagy is required for exercise training-induced skeletal muscle adaptation and improvement of physical performance.* The FASEB Journal, 2013. **27**(10): p. 4184.
141. Carter, H.N., et al., *Autophagy and mitophagy flux in young and aged skeletal muscle following chronic contractile activity.* The Journal of physiology, 2018. **596**(16): p. 3567-3584.
142. Chen, C.C.W., A.T. Erlich, and D.A. Hood, *Role of Parkin and endurance training on mitochondrial turnover in skeletal muscle.* Skelet Muscle, 2018. **8**(1): p. 10.
143. Cordeiro, A.V., et al., *High-intensity exercise training induces mitonuclear imbalance and activates the mitochondrial unfolded protein response in the skeletal muscle of aged mice.* GeroScience, 2021. **43**(3): p. 1513-1518.
144. Jiang, N., et al., *Upregulation of uncoupling protein-3 in skeletal muscle during exercise: a potential antioxidant function.* Free radical biology & medicine, 2008. **46**(2): p. 138-145.
145. MacNeil, L.G., et al., *The order of exercise during concurrent training for rehabilitation does not alter acute genetic expression, mitochondrial enzyme activity or improvements in muscle function.* PloS one, 2014. **9**(10): p. e109189.
146. Popov, D.V., et al., *Effect of aerobic training on baseline expression of signaling and respiratory proteins in human skeletal muscle.* Physiol Rep, 2018. **6**(17): p. e13868.
147. Gillen, J.B., et al., *Twelve Weeks of Sprint Interval Training Improves Indices of Cardiometabolic Health Similar to Traditional Endurance Training despite a Five-Fold Lower Exercise Volume and Time Commitment.* PLOS ONE, 2016. **11**(4): p. e0154075.
148. Konopka, A.R., et al., *Markers of Human Skeletal Muscle Mitochondrial Biogenesis and Quality Control: Effects of Age and Aerobic Exercise Training.* The Journals of Gerontology: Series A, 2013. **69**(4): p. 371-378.
149. Talanian, J.L., et al., *Exercise training increases sarcolemmal and mitochondrial fatty acid transport proteins in human skeletal muscle.* American Journal of Physiology-Endocrinology and Metabolism, 2010. **299**(2): p. E180-E188.
150. Perry, C.G., et al., *High-intensity aerobic interval training increases fat and carbohydrate metabolic capacities in human skeletal muscle.* Applied Physiology, Nutrition, and Metabolism, 2008. **33**(6): p. 1112-1123.
151. Blaak, E.E. and W.H. Saris, *Substrate oxidation, obesity and exercise training.* Best Pract Res Clin Endocrinol Metab, 2002. **16**(4): p. 667-78.
152. Perry, C.G., et al., *Repeated transient mRNA bursts precede increases in transcriptional and mitochondrial proteins during training in human skeletal muscle.* J Physiol, 2010. **588**(Pt 23): p. 4795-810.
153. Casuso, R.A. and J.R. Huertas, *The emerging role of skeletal muscle mitochondrial dynamics in exercise and ageing.* Ageing Res Rev, 2020. **58**: p. 101025.
154. Axelrod, C.L., et al., *Exercise training remodels human skeletal muscle mitochondrial fission and fusion machinery towards a pro-elongation phenotype.* Acta Physiologica, 2019. **225**(4): p. e13216.
155. Chang, C.-R. and C. Blackstone, *Cyclic AMP-dependent protein kinase phosphorylation of Drp1 regulates its GTPase activity and mitochondrial morphology.* Journal of Biological Chemistry, 2007. **282**(30): p. 21583-21587.
156. Huertas, J.R., et al., *Human muscular mitochondrial fusion in athletes during exercise.* The FASEB Journal, 2019. **33**(11): p. 12087-12098.
157. Fiorenza, M., et al., *Metabolic stress-dependent regulation of the mitochondrial biogenic molecular response to high-intensity exercise in human skeletal muscle.* The Journal of physiology, 2018. **596**(14): p. 2823-2840.

158. Ayer, A., et al., *The role of mitochondrial reactive oxygen species in insulin resistance.* Free Radic Biol Med, 2021.
159. Morrison, D., et al., *Vitamin C and E supplementation prevents some of the cellular adaptations to endurance-training in humans.* Free Radical Biology and Medicine, 2015. **89**: p. 852-862.
160. Ludzki, A., et al., *Rapid repression of ADP transport by palmitoyl-CoA is attenuated by exercise training in humans: a potential mechanism to decrease oxidative stress and improve skeletal muscle insulin signaling.* Diabetes, 2015. **64**(8): p. 2769-2779.
161. Fritz, T., et al., *Low-intensity exercise increases skeletal muscle protein expression of PPARδ and UCP3 in type 2 diabetic patients.* Diabetes/Metabolism Research and Reviews, 2006. **22**(6): p. 492-498.
162. Brandt, N., et al., *Exercise and exercise training-induced increase in autophagy markers in human skeletal muscle.* Physiological Reports, 2018. **6**(7): p. e13651.
163. Villa, E., S. Marchetti, and J.-E. Ricci, *No Parkin zone: mitophagy without Parkin.* Trends in cell biology, 2018. **28**(11): p. 882-895.
164. Gouspillou, G., et al., *Protective role of Parkin in skeletal muscle contractile and mitochondrial function.* The Journal of physiology, 2018. **596**(13): p. 2565-2579.
165. Guan, Y., J.C. Drake, and Z. Yan, *Exercise-Induced Mitophagy in Skeletal Muscle and Heart.* Exercise and sport sciences reviews, 2019. **47**(3): p. 151-156.
166. Islam, H., et al., *Exercise-Induced Mitochondrial Biogenesis: Molecular Regulation, Impact of Training, and Influence on Exercise Performance*, in *The Routledge Handbook on Biochemistry of Exercise*. 2020, Routledge. p. 143-161.
167. Hood, D.A., *Mechanisms of exercise-induced mitochondrial biogenesis in skeletal muscleThis paper is one of a selection of papers published in this Special Issue, entitled 14th International Biochemistry of Exercise Conference – Muscles as Molecular and Metabolic Machines, and has undergone the Journal's usual peer review process.* Applied Physiology, Nutrition, and Metabolism, 2009. **34**(3): p. 465-472.
168. Venojärvi, M., et al., *Exercise training with dietary counselling increases mitochondrial chaperone expression in middle-aged subjects with impaired glucose tolerance.* BMC Endocrine Disorders, 2008. **8**(1): p. 3.
169. Adhihetty, P.J., et al., *The effect of training on the expression of mitochondrial biogenesis- and apoptosis-related proteins in skeletal muscle of patients with mtDNA defects.* American Journal of Physiology-Endocrinology and Metabolism, 2007. **293**(3): p. E672-E680.
170. Reisman, E., *ANALYSIS OF CHANGES IN*

*MITOCHONDRIAL PROTEINS IN*

*SINGLE MUSCLE FIBRES WITH*

*DIFFERENT TYPES OF TRAINING*, in *Institute for Health and Sport (IHeS)*. 2020, Victoria University: Melbourne, Australia.
171. Ghosh, R., J.E. Gilda, and A.V. Gomes, *The necessity of and strategies for improving confidence in the accuracy of western blots.* Expert Review of Proteomics, 2014. **11**(5): p. 549-560.
172. Koller, A. and H. Wätzig, *Precision and variance components in quantitative gel electrophoresis.* Electrophoresis, 2005. **26**(12): p. 2470-5.
173. Mollica, J.P., et al., *Are genuine changes in protein expression being overlooked? Reassessing Western blotting.* Analytical biochemistry, 2009. **386**(2): p. 270-275.
174. Larsen, S., et al., *Biomarkers of mitochondrial content in skeletal muscle of healthy young human subjects.* The Journal of physiology, 2012. **590**(14): p. 3349-3360.
175. Groennebaek, T., et al., *Utilization of biomarkers as predictors of skeletal muscle mitochondrial content after physiological intervention and in clinical settings.* American Journal of Physiology-Endocrinology and Metabolism, 2020. **318**(6): p. E886-E889.
176. Meinild Lundby, A.-K., et al., *Exercise training increases skeletal muscle mitochondrial volume density by enlargement of existing mitochondria and not de novo biogenesis.* Acta Physiologica, 2018. **222**(1): p. e12905.

177. Gonzalez-Franquesa, A., et al., *Mass-spectrometry-based proteomics reveals mitochondrial supercomplexome plasticity.* Cell Reports, 2021. **35**(8): p. 109180.
178. Hoffman, N.J., *Omics and Exercise: Global Approaches for Mapping Exercise Biological Networks.* Cold Spring Harb Perspect Med, 2017. **7**(10).
179. Patterson, S.D., *Proteomics: evolution of the technology.* Biotechniques, 2003. **35**(3): p. 440-444.
180. Roe, M.R. and T.J. Griffin, *Gel-free mass spectrometry-based high throughput proteomics: Tools for studying biological response of proteins and proteomes.* Proteomics, 2006. **6**(17): p. 4678-4687.
181. Aebersold, R. and D.R. Goodlett, *Mass Spectrometry in Proteomics.* Chemical Reviews, 2001. **101**(2): p. 269-296.
182. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.
183. Dau, T., G. Bartolomucci, and J. Rappsilber, *Proteomics Using Protease Alternatives to Trypsin Benefits from Sequential Digestion with Trypsin.* Anal Chem, 2020. **92**(14): p. 9523-9527.
184. Dupree, E.J., et al., *A critical review of bottom-up proteomics: The good, the bad, and the future of this field.* Proteomes, 2020. **8**(3): p. 14.
185. Podwojski, K., et al., *Retention time alignment algorithms for LC/MS data must consider non-linear shifts.* Bioinformatics, 2009. **25**(6): p. 758-764.
186. Yates, J.R., C.I. Ruse, and A. Nakorchevsky, *Proteomics by mass spectrometry: approaches, advances, and applications.* Annual review of biomedical engineering, 2009. **11**(1): p. 49-79.
187. Zhu, W., J.W. Smith, and C.-M. Huang, *Mass Spectrometry-Based Label-Free Quantitative Proteomics.* Journal of Biomedicine and Biotechnology, 2010. **2010**: p. 840518.
188. Eng, J.K., et al., *A face in the crowd: recognizing peptides through database search.* Molecular & Cellular Proteomics, 2011. **10**(11).
189. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.* Nature Biotechnology, 2008. **26**(12): p. 1367-1372.
190. Huang, T., et al., *Protein inference: a review.* Briefings in Bioinformatics, 2012. **13**(5): p. 586-614.
191. Matzke, M.M., et al., *A comparative analysis of computational approaches to relative protein quantification using peptide peak intensities in label-free LC-MS proteomics experiments.* Proteomics, 2013. **13**(3-4): p. 493-503.
192. Johnson, J.V., et al., *Tandem-in-space and tandem-in-time mass spectrometry: triple quadrupoles and quadrupole ion traps.* Analytical Chemistry, 1990. **62**(20): p. 2162-2172.
193. Shimadzu. *Introduction to mass analyzers*. Available from: https://www.shimadzu.com/an/service-support/technical-support/analysis-basics/fundamental/mass_analyzers.html#:~:text=Quadrupole%20MS%20systems%20separate%20and,unstable%20oscillations%20from%20the%20system.
194. Domon, B. and R. Aebersold, *Mass spectrometry and protein analysis.* science, 2006. **312**(5771): p. 212-217.
195. Greco, V., et al., *Applications of MALDI-TOF mass spectrometry in clinical proteomics.* Expert Review of Proteomics, 2018. **15**(8): p. 683-696.
196. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.
197. Fernández-Puente, P., et al., *LC-MALDI-TOF/TOF for Shotgun Proteomics*, in *Shotgun Proteomics*. 2014, Springer. p. 27-38.
198. Karpievitch, Y.V., et al., *Liquid Chromatography Mass Spectrometry-Based Proteomics: Biological and Technological Aspects.* Ann Appl Stat, 2010. **4**(4): p. 1797-1823.
199. Scigelova, M., et al., *Fourier transform mass spectrometry.* Molecular & Cellular Proteomics, 2011. **10**(7).
200. Zubarev, R.A. and A. Makarov, *Orbitrap Mass Spectrometry.* Analytical Chemistry, 2013. **85**(11): p. 5288-5296.

201. Bateman, K.P., et al., *Quantitative-qualitative data acquisition using a benchtop Orbitrap mass spectrometer.* Journal of the American Society for Mass Spectrometry, 2009. **20**(8): p. 1441-1450.
202. Hecht, E.S., et al., *Fundamentals and Advances of Orbitrap Mass Spectrometry*, in *Encyclopedia of Analytical Chemistry*. 2019. p. 1-40.
203. Nikolov, M., C. Schmidt, and H. Urlaub, *Quantitative mass spectrometry-based proteomics: an overview.* Quantitative methods in proteomics, 2012: p. 85-100.
204. Bittremieux, W., et al., *Quality control in mass spectrometry-based proteomics.* Mass Spectrometry Reviews, 2018. **37**(5): p. 697-711.
205. Deshmukh, A.S., *Proteomics of Skeletal Muscle: Focus on Insulin Resistance and Exercise Biology.* Proteomes, 2016. **4**(1).
206. Patel, V.J., et al., *A Comparison of Labeling and Label-Free Mass Spectrometry-Based Proteomics Approaches.* Journal of Proteome Research, 2009. **8**(7): p. 3752-3759.
207. Sajic, T., Y. Liu, and R. Aebersold, *Using data-independent, high-resolution mass spectrometry in protein biomarker research: Perspectives and clinical applications.* PROTEOMICS – Clinical Applications, 2015. **9**(3-4): p. 307-321.
208. Michalski, A., J. Cox, and M. Mann, *More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority is Inaccessible to Data-Dependent LC−MS/MS.* Journal of Proteome Research, 2011. **10**(4): p. 1785-1793.
209. Natasha Beeton-Kempen, P.D., *<007- DDA vs DIA_V2.pdf>.* Technology Networks, 2021.
210. Hu, A., W.S. Noble, and A. Wolf-Yadlin, *Technical advances in proteomics: new developments in data-independent acquisition.* F1000Research, 2016. **5**: p. F1000 Faculty Rev-419.
211. Zhang, F., et al., *Data-Independent Acquisition Mass Spectrometry-Based Proteomics and Software Tools: A Glimpse in 2020.* PROTEOMICS, 2020. **20**(17-18): p. 1900276.
212. Elias, J.E., et al., *Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations.* Nature methods, 2005. **2**(9): p. 667-675.
213. van Ooijen, M.P., et al., *Identification of differentially expressed peptides in high-throughput proteomics data.* Brief Bioinform, 2018. **19**(5): p. 971-981.
214. Cervone, D.T., et al., *Mass spectrometry-based proteomics approaches to interrogate skeletal muscle adaptations to exercise.* Scandinavian Journal of Medicine & Science in Sports. **n/a**(n/a).
215. Hesketh, S.J., et al., *The application of proteomics in muscle exercise physiology.* Expert Rev Proteomics, 2020. **17**(11-12): p. 813-825.
216. Holloway, K.V., et al., *Proteomic investigation of changes in human vastus lateralis muscle in response to interval-exercise training.* Proteomics, 2009. **9**(22): p. 5155-74.
217. Egan, B., et al., *2-D DIGE analysis of the mitochondrial proteome from human skeletal muscle reveals time course-dependent remodelling in response to 14 consecutive days of endurance exercise training.* Proteomics, 2011. **11**(8): p. 1413-28.
218. Hostrup, M., et al., *Chronic beta2 -adrenoceptor agonist treatment alters muscle proteome and functional adaptations induced by high intensity training in young men.* J Physiol, 2018. **596**(2): p. 231-252.
219. Darshi, M., et al., *ChChd3, an inner mitochondrial membrane protein, is essential for maintaining crista integrity and mitochondrial function.* J Biol Chem, 2011. **286**(4): p. 2918-32.
220. Koopman, W.J., et al., *OXPHOS mutations and neurodegeneration.* Embo j, 2013. **32**(1): p. 9-29.
221. Egan, B., et al., *Time course analysis reveals gene-specific transcript and protein kinetics of adaptation to short-term aerobic exercise training in human skeletal muscle.* PLoS One, 2013. **8**(9): p. e74098.
222. Hostrup, M., et al., *High-Intensity Interval Training Remodels the Proteome and Acetylome of Human Skeletal Muscle.* 2021.
223. Makhnovskii, P.A., et al., *Regulation of Proteins in Human Skeletal Muscle: The Role of Transcription.* Sci Rep, 2020. **10**(1): p. 3514.

224. Schild, M., et al., *Basal and exercise induced label-free quantitative protein profiling of m. vastus lateralis in trained and untrained individuals.* J Proteomics, 2015. **122**: p. 119-32.
225. Ubaida-Mohien, C., et al., *Physical Activity Associated Proteomics of Skeletal Muscle: Being Physically Active in Daily Life May Protect Skeletal Muscle From Aging.* Front Physiol, 2019. **10**: p. 312.
226. Haidich, A.-B.J.H., *Meta-analysis in medical research.* 2010. **14**(Suppl 1): p. 29.
227. Amar, D., et al., *Time trajectories in the transcriptomic response to exercise-a meta-analysis.* Nature Communications, 2021. **12**(1): p. 1-12.
228. Pillon, N.J., et al., *Transcriptomic profiling of skeletal muscle adaptations to exercise and inactivity.* Nature communications, 2020. **11**(1): p. 1-15.
229. Petriz, B.A., et al., *The Effects of Acute and Chronic Exercise on Skeletal Muscle Proteome.* J Cell Physiol, 2017. **232**(2): p. 257-269.
230. Srisawat, K., et al., *A systematic review and meta-analysis of proteomics literature on the response of human skeletal muscle to obesity/type 2 diabetes mellitus (T2DM) versus exercise training.* Proteomes, 2017. **5**(4): p. 30.
231. Sidoli, S., K. Kulej, and B.A. Garcia, *Why proteomics is not the new genomics and the future of mass spectrometry in cell biology.* Journal of Cell Biology, 2016. **216**(1): p. 21-24.
232. Miller, B.F., A.R. Konopka, and K.L. Hamilton, *The rigorous study of exercise adaptations: why mRNA might not be enough.* J Appl Physiol (1985), 2016. **121**(2): p. 594-6.
233. Fournier, M.L., et al., *Delayed correlation of mRNA and protein expression in rapamycin-treated cells and a role for Ggc1 in cellular sensitivity to rapamycin.* Mol Cell Proteomics, 2010. **9**(2): p. 271-84.
234. McGinley, C. and D.J. Bishop, *Distinct protein and mRNA kinetics of skeletal muscle proton transporters following exercise can influence interpretation of adaptations to training.* Experimental physiology, 2016. **101**(12): p. 1565-1580.
235. Ke, L., et al., *The mitochondrial biogenesis signaling pathway is a potential therapeutic target for myasthenia gravis via energy metabolism.* 2021. **22**(1): p. 1-10.
236. Carey, A.L. and B.A. Kingwell, *Novel pharmacological approaches to combat obesity and insulin resistance: targeting skeletal muscle with 'exercise mimetics'.* Diabetologia, 2009. **52**(10): p. 2015-26.
237. Mehta, D., et al., *The incongruity of validating quantitative proteomics using western blots.* Nature Plants, 2022. **8**(12): p. 1320-1321.
238. Faulkner, S., et al., *A comparison of the bovine uterine and plasma proteome using i TRAQ proteomics.* Proteomics, 2012. **12**(12): p. 2014-2023.
239. Jacques, M., et al., *DNA methylation and proteomics integration uncover dose-dependent group and individual responses to exercise in human skeletal muscle.* bioRxiv, 2022: p. 2022.07.11.499662.
240. Jamnick, N.A., et al., *Manipulating graded exercise test variables affects the validity of the lactate threshold and V˙ O 2 peak.* PloS one, 2018. **13**(7): p. e0199794.
241. Swain, D.P., et al., *Target heart rates for the development of cardiorespiratory fitness.* Medicine and science in sports and exercise, 1994. **26**(1): p. 112-116.
242. Wickham H, B.J., *readxl: Read Excel Files*. 2022.
243. Lazar, C., et al., *Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies.* Journal of Proteome Research, 2016. **15**(4): p. 1116-1125.
244. Karpievitch, Y.V., A.R. Dabney, and R.D. Smith, *Normalization and missing value imputation for label-free LC-MS analysis.* BMC Bioinformatics, 2012. **13**(16): p. S5.
245. Webb-Robertson, B.J., et al., *Review, evaluation, and discussion of the challenges of missing value imputation for mass spectrometry-based label-free global proteomics.* J Proteome Res, 2015. **14**(5): p. 1993-2001.
246. Lazar, C., *imputeLCMD: A collection of methods for left-censored missing data imputation*. 2019.
247. Jin, L., et al., *A comparative study of evaluating missing value imputation methods in label-free proteomics.* Scientific Reports, 2021. **11**(1): p. 1760.

248. Wei, R., et al., *Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data.* Scientific Reports, 2018. **8**(1): p. 663.
249. Stekhoven, D.J. and M.D.J. Stekhoven, *Package 'missForest'.* R package version, 2013. **1**.
250. Hrydziuszko, O. and M.R. Viant, *Missing values in mass spectrometry based metabolomics: an undervalued step in the data processing pipeline.* Metabolomics, 2012. **8**: p. 161-174.
251. Hastie T, T.R., Narasimhan B, Chu G, *impute: impute: Imputation for microarray data. R package version 1.70.0.* 2022.
252. Välikangas, T., T. Suomi, and L.L. Elo, *A systematic evaluation of normalization methods in quantitative label-free proteomics.* Briefings in Bioinformatics, 2016. **19**(1): p. 1-11.
253. Chawade, A., E. Alexandersson, and F. Levander, *Normalyzer: A Tool for Rapid Evaluation of Normalization Methods for Omics Data Sets.* Journal of Proteome Research, 2014. **13**(6): p. 3114-3120.
254. Kultima, K., et al., *Development and evaluation of normalization methods for label-free relative quantification of endogenous peptides.* Molecular & Cellular Proteomics, 2009. **8**(10): p. 2285-2295.
255. Huber, W., et al., *Variance stabilization applied to microarray data calibration and to the quantification of differential expression.* Bioinformatics, 2002. **18**(suppl_1): p. S96-S104.
256. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies.* Nucleic acids research, 2015. **43**(7): p. e47-e47.
257. Howe, K.L., et al., *Ensembl 2021.* Nucleic acids research, 2021. **49**(D1): p. D884-D891.
258. Zhu, Y., et al., *DEqMS: A Method for Accurate Variance Estimation in Differential Protein Expression Analysis.* Mol Cell Proteomics, 2020. **19**(6): p. 1047-1057.
259. Schwämmle, V., I.R. León, and O.N. Jensen, *Assessment and Improvement of Statistical Tools for Comparative Proteomics Analysis of Sparse Data Sets with Few Experimental Replicates.* Journal of Proteome Research, 2013. **12**(9): p. 3874-3883.
260. Bernhardt, O.M., et al., *Spectronaut: a fast and efficient algorithm for MRM-like processing of data independent acquisition (SWATH-MS) data.* Biognosys. ch, 2012.
261. Team, R.C., *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. 2022: Vienna, Austria.
262. Benjamini, Y. and Y. Hochberg, *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.* Journal of the Royal Statistical Society. Series B (Methodological), 1995. **57**(1): p. 289-300.
263. McDonald, J.H., *Handbook of biolological statistics*. 2014: New York•.
264. Rothman, K.J., *No adjustments are needed for multiple comparisons.* Epidemiology, 1990: p. 43-46.
265. Viechtbauer, W.J.J.o.s.s., *Conducting meta-analyses in R with the metafor package.* 2010. **36**(3): p. 1-48.
266. Borenstein, M., et al., *A basic introduction to fixed-effect and random-effects models for meta-analysis.* Research synthesis methods, 2010. **1**(2): p. 97-111.
267. Sidik, K. and J.N. Jonkman, *Simple heterogeneity variance estimation for meta-analysis.* Journal of the Royal Statistical Society: Series C (Applied Statistics), 2005. **54**(2): p. 367-384.
268. Harrer, M., et al., *Doing meta-analysis with R: a hands-on guide*. 2021: Chapman and Hall/CRC.
269. Kossmeier, M., U.S. Tran, and M. Voracek, *Visualizing meta-analytic data with R package metaviz.* R Package Version, 2020. **3**: p. 1.
270. Wu, T., et al., *clusterProfiler 4.0: A universal enrichment tool for interpreting omics data.* The Innovation, 2021. **2**(3): p. 100141.
271. Yu, G. and Q.-Y. He, *ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization.* Molecular BioSystems, 2016. **12**(2): p. 477-479.
272. Yu, G., *Enrichplot: visualization of functional enrichment result.* R package version, 2019. **1**(2).
273. Subramanian, A., et al., *Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles.* Proceedings of the National Academy of Sciences, 2005. **102**(43): p. 15545-15550.

274. Carlson, M., *org.Hs.eg.db: Genome wide annotation for Human. R package version 3.8.2*. 2019.
275. Blighe K, R.S., Lewis M *EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R package version* 2022.
276. Luo, W. and C. Brouwer, *Pathview: an R/Bioconductor package for pathway-based data integration and visualization.* Bioinformatics, 2013. **29**(14): p. 1830-1831.
277. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* 2003. **13**(11): p. 2498-2504.
278. Doncheva, N.T., et al., *Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data.* J Proteome Res, 2019. **18**(2): p. 623-632.
279. Yates III, J.R., *Recent technical advances in proteomics.* F1000Research, 2019. **8**.
280. Pfanner, N., B. Warscheid, and N. Wiedemann, *Mitochondrial proteins: from biogenesis to functional networks.* Nat Rev Mol Cell Biol, 2019. **20**(5): p. 267-284.
281. Woolson, R.F., *Wilcoxon Signed-Rank Test*, in *Wiley Encyclopedia of Clinical Trials*. p. 1-3.
282. Burgomaster, K.A., et al., *Similar metabolic adaptations during exercise after low volume sprint interval and traditional endurance training in humans.* The Journal of Physiology, 2008. **586**(1): p. 151-160.
283. Burgomaster, K.A., G.J. Heigenhauser, and M.J. Gibala, *Effect of short-term sprint interval training on human skeletal muscle carbohydrate metabolism during exercise and time-trial performance.* Journal of applied physiology, 2006. **100**(6): p. 2041-2047.
284. MacDougall, J.D., et al., *Muscle performance and enzymatic adaptations to sprint interval training.* Journal of Applied Physiology, 1998. **84**(6): p. 2138-2142.
285. Juel, C., et al., *Effect of high-intensity intermittent training on lactate and H+ release from human skeletal muscle.* American Journal of Physiology-Endocrinology and Metabolism, 2004. **286**(2): p. E245-E251.
286. Palmieri, F., *The mitochondrial transporter family (SLC25): physiological and pathological implications.* Pflügers Archiv, 2004. **447**(5): p. 689-709.
287. Ruprecht, J.J. and E.R. Kunji, *The SLC25 mitochondrial carrier family: structure and mechanism.* Trends in biochemical sciences, 2020. **45**(3): p. 244-258.
288. Zhang, J., et al., *Cloning and functional characterization of ACAD-9, a novel member of human acyl-CoA dehydrogenase family.* Biochem Biophys Res Commun, 2002. **297**(4): p. 1033-42.
289. Little, J.P., et al., *A practical model of low-volume high-intensity interval training induces mitochondrial biogenesis in human skeletal muscle: potential mechanisms.* The Journal of physiology, 2010. **588**(6): p. 1011-1022.
290. von Hippel, P.T., *The heterogeneity statistic I2 can be biased in small meta-analyses.* BMC Medical Research Methodology, 2015. **15**(1): p. 35.
291. Huedo-Medina, T.B., et al., *Assessing heterogeneity in meta-analysis: Q statistic or I2 index?* Psychol Methods, 2006. **11**(2): p. 193-206.
292. Thorlund, K., et al., *Evolution of heterogeneity (I2) estimates and their 95% confidence intervals in large meta-analyses.* PLoS One, 2012. **7**(7): p. e39471.
293. Kras, K.A., et al., *Obesity modifies the stoichiometry of mitochondrial proteins in a way that is distinct to the subcellular localization of the mitochondria in skeletal muscle.* Metabolism, 2018. **89**: p. 18-26.
294. Zaltsman, Y., et al., *MTCH2/MIMP is a major facilitator of tBID recruitment to mitochondria.* Nature cell biology, 2010. **12**(6): p. 553-562.
295. Zhou, L., et al., *The Role of SIRT3 in Exercise and Aging.* Cells, 2022. **11**(16): p. 2596.
296. He, W., et al., *Mitochondrial sirtuins: regulators of protein acylation and metabolism.* Trends in Endocrinology & Metabolism, 2012. **23**(9): p. 467-476.
297. Lorin, S., et al., *Glutamate dehydrogenase contributes to leucine sensing in the regulation of autophagy.* Autophagy, 2013. **9**(6): p. 850-860.
298. Jaleh, F., *Role of GLUD1 in determining the anabolic response to amino acids in skeletal muscle*. 2020, ETH Zurich.
299. Talley, J.T. and S.S. Mohiuddin, *Biochemistry, fatty acid oxidation*, in *StatPearls [Internet]*. 2022, StatPearls Publishing.

300. Williamson, D.L. and T.C. Rideout, *Is ACSL6 at the crossroads of skeletal muscle lipid synthesis?* J Physiol, 2017. **595**(3): p. 619-620.
301. Davies, M.N., et al., *The acetyl group buffering action of carnitine acetyltransferase offsets macronutrient-induced lysine acetylation of mitochondrial proteins.* Cell reports, 2016. **14**(2): p. 243-254.
302. Fritzen, A.M., A.M. Lundsgaard, and B. Kiens, *Tuning fatty acid oxidation in skeletal muscle with dietary fat and exercise.* Nat Rev Endocrinol, 2020. **16**(12): p. 683-696.
303. Lefai, E., et al., *Exercise training improves fat metabolism independent of total energy expenditure in sedentary overweight men, but does not restore lean metabolic phenotype.* International Journal of Obesity, 2017. **41**(12): p. 1728-1736.
304. Battaglia, G.M., et al., *Effect of exercise training on metabolic flexibility in response to a high-fat diet in obese individuals.* American Journal of Physiology-Endocrinology and Metabolism, 2012. **303**(12): p. E1440-E1445.
305. Bruce, C.R., et al., *Endurance training in obese humans improves glucose tolerance and mitochondrial fatty acid oxidation and alters muscle lipid content.* American Journal of Physiology-Endocrinology and Metabolism, 2006. **291**(1): p. E99-E107.
306. Kiens, B., et al., *Skeletal muscle substrate utilization during submaximal exercise in man: effect of endurance training.* The Journal of physiology, 1993. **469**(1): p. 459-478.
307. Friedlander, A.L., et al., *Effects of exercise intensity and training on lipid metabolism in young women.* American Journal of Physiology-Endocrinology and Metabolism, 1998. **275**(5): p. E853-E863.
308. Gabriel, J.L., P.R. Zervos, and G.W. Plaut, *Activity of purified NAD-specific isocitrate dehydrogenase at modulator and substrate concentrations approximating conditions in mitochondria.* Metabolism, 1986. **35**(7): p. 661-667.
309. Krebs, H.A., *Rate control of the tricarboxylic acid cycle.* Advances in Enzyme Regulation, 1970. **8**: p. 335-353.
310. Heigenhauser, G.J. and M.L. Parolin, *Role of pyruvate dehydrogenase in lactate production in exercising human skeletal muscle.* Adv Exp Med Biol, 1999. **474**: p. 205-18.
311. Patel, M.S. and T.E. Roche, *Molecular biology and biochemistry of pyruvate dehydrogenase complexes1.* The FASEB Journal, 1990. **4**(14): p. 3224-3233.
312. Huberts, D.H. and I.J. van der Klei, *Moonlighting proteins: an intriguing mode of multitasking.* Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2010. **1803**(4): p. 520-525.
313. Harris, R.A., et al., *Dihydrolipoamide dehydrogenase-binding protein of the human pyruvate dehydrogenase complex: DNA-derived amino acid sequence, expression, and reconstitution of the pyruvate dehydrogenase complex.* Journal of Biological Chemistry, 1997. **272**(32): p. 19746-19751.
314. Linn, T.C., F.H. Pettit, and L.J. Reed, *α-Keto acid dehydrogenase complexes, X. Regulation of the activity of the pyruvate dehydrogenase complex from beef kidney mitochondria by phosphorylation and dephosphorylation.* Proceedings of the National Academy of Sciences, 1969. **62**(1): p. 234-241.
315. Roche, T.E., et al., *Distinct regulatory properties of pyruvate dehydrogenase kinase and phosphatase isoforms.* 2001.
316. Gudiksen, A., et al., *Effects of training status on PDH regulation in human skeletal muscle during exercise.* Pflügers Archiv - European Journal of Physiology, 2017. **469**(12): p. 1615-1630.
317. LeBlanc, P.J., et al., *Effects of aerobic training on pyruvate dehydrogenase and pyruvate dehydrogenase kinase in human skeletal muscle.* The Journal of Physiology, 2004. **557**(2): p. 559-570.
318. Vigelsø, A., N.B. Andersen, and F. Dela, *The relationship between skeletal muscle mitochondrial citrate synthase activity and whole body oxygen uptake adaptations in response to exercise training.* International journal of physiology, pathophysiology and pharmacology, 2014. **6**(2): p. 84.
319. Howarth, K.R., et al., *Effect of endurance training on muscle TCA cycle metabolism during exercise in humans.* Journal of Applied Physiology, 2004. **97**(2): p. 579-584.

320. Agudelo, L.Z., et al., *Skeletal muscle PGC-1α1 reroutes kynurenine metabolism to increase energy efficiency and fatigue-resistance.* Nature Communications, 2019. **10**(1): p. 2767.
321. Sharma, L.K., J. Lu, and Y. Bai, *Mitochondrial respiratory complex I: structure, function and implication in human diseases.* Curr Med Chem, 2009. **16**(10): p. 1266-77.
322. Formosa, L.E., et al., *Dissecting the Roles of Mitochondrial Complex I Intermediate Assembly Complex Factors in the Biogenesis of Complex I.* Cell Rep, 2020. **31**(3): p. 107541.
323. Vercellino, I. and L.A. Sazanov, *The assembly, regulation and function of the mitochondrial respiratory chain.* Nature Reviews Molecular Cell Biology, 2022. **23**(2): p. 141-161.
324. Sánchez, E., et al., *LYRM7/MZM1L is a UQCRFS1 chaperone involved in the last steps of mitochondrial Complex III assembly in human cells.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 2013. **1827**(3): p. 285-293.
325. Joza, N., et al., *AIF: Not Just an Apoptosis-Inducing Factor.* Annals of the New York Academy of Sciences, 2009. **1171**(1): p. 2-11.
326. Wischhof, L., et al., *AIFM1 beyond cell death: An overview of this OXPHOS-inducing factor in mitochondrial diseases.* EBioMedicine, 2022. **83**: p. 104231.
327. Wiehe, R.S., et al., *Endonuclease G promotes mitochondrial genome cleavage and replication.* Oncotarget, 2018. **9**(26): p. 18309-18326.
328. Lisak, D.A., et al., *The transmembrane Bax inhibitor motif (TMBIM) containing protein family: Tissue expression, intracellular localization and effects on the ER CA2+-filling state.* Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 2015. **1853**(9): p. 2104-2114.
329. Weber, T.A., et al., *APOOL is a cardiolipin-binding constituent of the Mitofilin/MINOS protein complex determining cristae morphology in mammalian mitochondria.* PLoS One, 2013. **8**(5): p. e63683.
330. Koob, S. and A.S. Reichert, *Novel intracellular functions of apolipoproteins: the ApoO protein family as constituents of the Mitofilin/MINOS complex determines cristae morphology in mitochondria.* Biological chemistry, 2014. **395**(3): p. 285-296.
331. Darshi, M. and S.S. Taylor, *Mitochondrial ChChD3 acts as a Scaffold for Mitofilin, Sam50 and PKA*. 2008, Wiley Online Library.
332. Abudu, Y.P., et al., *SAMM50 acts with p62 in piecemeal basal- and OXPHOS-induced mitophagy of SAM and MICOS components.* Journal of Cell Biology, 2021. **220**(8).
333. Kozjak-Pavlovic, V., et al., *Conserved roles of Sam50 and metaxins in VDAC biogenesis.* EMBO reports, 2007. **8**(6): p. 576-582.
334. Koltai, E., et al., *Master athletes have higher miR-7, SIRT3 and SOD2 expression in skeletal muscle than age-matched sedentary controls.* Redox biology, 2018. **19**: p. 46-51.
335. Karplus, P.A., *A primer on peroxiredoxin biochemistry.* Free Radical Biology and Medicine, 2015. **80**: p. 183-190.
336. Wadley, A.J., S. Aldred, and S.J. Coles, *An unexplored role for Peroxiredoxin in exercise-induced redox signalling?* Redox Biology, 2016. **8**: p. 51-58.
337. Cao, Y., et al., *ABHD10 is an S-depalmitoylase affecting redox homeostasis through peroxiredoxin-5.* Nature Chemical Biology, 2019. **15**(12): p. 1232-1240.
338. Damdimopoulos, A.E., et al., *Human mitochondrial thioredoxin: involvement in mitochondrial membrane potential and cell death.* Journal of Biological Chemistry, 2002. **277**(36): p. 33249-33257.
339. Fisher-Wellman, K.H., et al., *Novel role for thioredoxin reductase-2 in mitochondrial redox adaptations to obesogenic diet and exercise in heart and skeletal muscle.* The Journal of Physiology, 2013. **591**(14): p. 3471-3486.
340. Carvalho, V.H., et al., *Exercise and β-alanine supplementation on carnosine-acrolein adduct in skeletal muscle.* Redox Biology, 2018. **18**: p. 222-228.
341. Chen, C.H., L. Sun, and D. Mochly-Rosen, *Mitochondrial aldehyde dehydrogenase and cardiac diseases.* Cardiovasc Res, 2010. **88**(1): p. 51-7.
342. Marí, M., et al., *Mitochondrial Glutathione: Recent Insights and Role in Disease.* Antioxidants, 2020. **9**(10): p. 909.
343. Baldelli, S., et al., *Glutathione and Nitric Oxide: Key Team Players in Use and Disuse of Skeletal Muscle.* Nutrients, 2019. **11**(10): p. 2318.

344. Powers, S.K., et al., *Exercise Training and Skeletal Muscle Antioxidant Enzymes: An Update.* Antioxidants, 2023. **12**(1): p. 39.
345. Cordell, P.A., et al., *The human hydroxyacylglutathione hydrolase (HAGH) gene encodes both cytosolic and mitochondrial forms of glyoxalase II.* Journal of Biological Chemistry, 2004. **279**(27): p. 28653-28661.
346. Raza, H., *Dual localization of glutathione S-transferase in the cytosol and mitochondria: implications in oxidative stress, toxicity and disease.* The FEBS Journal, 2011. **278**(22): p. 4243-4251.
347. Takahashi, M., et al., *Contractile activity-induced adaptations in the mitochondrial protein import system.* American Journal of Physiology-Cell Physiology, 1998. **274**(5): p. C1380-C1387.
348. Gordon, J.W., et al., *Selected Contribution: Effects of contractile activity on mitochondrial transcription factor A expression in skeletal muscle.* Journal of Applied Physiology, 2001. **90**(1): p. 389-396.
349. Lu, X., et al., *Ultrastructural and proteomic profiling of mitochondria-associated endoplasmic reticulum membranes reveal aging signatures in striated muscle.* Cell Death & Disease, 2022. **13**(4): p. 296.
350. Mordas, A. and K. Tokatlidis, *The MIA pathway: a key regulator of mitochondrial oxidative protein folding and biogenesis.* Acc Chem Res, 2015. **48**(8): p. 2191-9.
351. Modjtahedi, N., et al., *Mitochondrial Proteins Containing Coiled-Coil-Helix-Coiled-Coil-Helix (CHCH) Domains in Health and Disease.* Trends in Biochemical Sciences, 2016. **41**(3): p. 245-260.
352. Salscheider, S.L., et al., *AIFM1 is a component of the mitochondrial disulfide relay that drives complex I assembly through efficient import of NDUFS5.* The EMBO Journal, 2022. **41**(17): p. e110784.
353. Jensen, R.E. and C.D. Dunn, *Protein import into and across the mitochondrial inner membrane: role of the TIM23 and TIM22 translocons.* Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 2002. **1592**(1): p. 25-34.
354. Bauer, M.F., et al., *Genetic and structural characterization of the human mitochondrial inner membrane translocase.* J Mol Biol, 1999. **289**(1): p. 69-82.
355. Vukotic, M., et al., *Acylglycerol kinase mutated in Sengers syndrome is a subunit of the TIM22 protein translocase in mitochondria.* Molecular cell, 2017. **67**(3): p. 471-483. e7.
356. Kang, Y., et al., *Sengers syndrome-associated mitochondrial acylglycerol kinase is a subunit of the human TIM22 protein import complex.* Molecular cell, 2017. **67**(3): p. 457-470. e5.
357. Joseph, A.-M. and D.A. Hood, *Plasticity of TOM complex assembly in skeletal muscle mitochondria in response to chronic contractile activity.* Mitochondrion, 2012. **12**(2): p. 305-312.
358. Koo, J.-H., J.-Y. Cho, and U.-B. Lee, *Treadmill exercise alleviates motor deficits and improves mitochondrial import machinery in an MPTP-induced mouse model of Parkinson's disease.* Experimental Gerontology, 2017. **89**: p. 20-29.
359. Hutu, D.P., et al., *Mitochondrial protein import motor: differential role of Tim44 in the recruitment of Pam17 and J-complex to the presequence translocase.* Mol Biol Cell, 2008. **19**(6): p. 2642-9.
360. Naylor, D.J., et al., *Evidence for the existence of distinct mammalian cytosolic, microsomal, and two mitochondrial GrpE-like proteins, the co-chaperones of specific Hsp70 members.* Journal of Biological Chemistry, 1998. **273**(33): p. 21169-21177.
361. Shin, C.-S., et al., *LONP1 and mtHSP70 cooperate to promote mitochondrial protein folding.* Nature Communications, 2021. **12**(1): p. 265.
362. Schneider, H.-C., et al., *Mitochondrial Hsp70/MIM44 complex facilitates protein import.* Nature, 1994. **371**(6500): p. 768-774.
363. Voos, W., et al., *Differential requirement for the mitochondrial Hsp70-Tim44 complex in unfolding and translocation of preproteins.* The EMBO journal, 1996. **15**(11): p. 2668-2677.
364. Theilen, N.T., G.H. Kunkel, and S.C. Tyagi, *The role of exercise and TFAM in preventing skeletal muscle atrophy.* Journal of cellular physiology, 2017. **232**(9): p. 2348-2358.

365. Hood, D.A., *Invited review: contractile activity-induced mitochondrial biogenesis in skeletal muscle.* Journal of applied physiology, 2001. **90**(3): p. 1137-1157.
366. Gjøvaag, T.F. and H.A. Dahl, *Effect of training and detraining on the expression of heat shock proteins in m. triceps brachii of untrained males and females.* European Journal of Applied Physiology, 2006. **98**(3): p. 310-322.
367. Sollanek, K.J., et al., *Global proteome changes in the rat diaphragm induced by endurance exercise training.* PLoS One, 2017. **12**(1): p. e0171007.
368. Thompson, K., et al., *OXA 1L mutations cause mitochondrial encephalopathy and a combined oxidative phosphorylation defect.* EMBO molecular medicine, 2018. **10**(11): p. e9060.
369. Wruck, F., et al., *The ribosome modulates folding inside the ribosomal exit tunnel.* Communications Biology, 2021. **4**(1): p. 523.
370. Cogliati, S., et al., *Regulation of mitochondrial electron transport chain assembly.* Journal of molecular biology, 2018. **430**(24): p. 4849-4873.
371. Memme, J.M., A.N. Oliveira, and D.A. Hood, *p53 regulates skeletal muscle mitophagy and mitochondrial quality control following denervation-induced muscle disuse.* Journal of Biological Chemistry, 2022. **298**(2).
372. Bross, P. and P. Fernandez-Guerra, *Disease-associated mutations in the HSPD1 gene encoding the large subunit of the mitochondrial HSP60/HSP10 chaperonin complex.* Frontiers in molecular biosciences, 2016. **3**: p. 49.
373. Marino Gammazza, A., et al., *Hsp60 in skeletal muscle fiber biogenesis and homeostasis: From physical exercise to skeletal muscle pathology.* Cells, 2018. **7**(12): p. 224.
374. Morton, J.P., et al., *Trained men display increased basal heat shock protein content of skeletal muscle.* Medicine and science in sports and exercise, 2008. **40**(7): p. 1255-1262.
375. Folkesson, M., et al., *The expression of heat shock protein in human skeletal muscle: effects of muscle fibre phenotype and training background.* Acta Physiologica, 2013. **209**(1): p. 26-33.
376. Wai, T., et al., *The membrane scaffold SLP2 anchors a proteolytic hub in mitochondria containing PARL and the i-AAA protease YME1L.* EMBO reports, 2016. **17**(12): p. 1844-1856.
377. Stiburek, L., et al., *YME1L controls the accumulation of respiratory chain subunits and is required for apoptotic resistance, cristae morphogenesis, and cell proliferation.* Molecular biology of the cell, 2012. **23**(6): p. 1010-1023.
378. Powers, S.K., et al., *Exercise-induced oxidative stress: Friend or foe?* Journal of Sport and Health Science, 2020. **9**(5): p. 415-425.
379. Christie, D.A., et al., *Stomatin-Like Protein 2 Binds Cardiolipin and Regulates Mitochondrial Biogenesis and Function.* MOLECULAR AND CELLULAR BIOLOGY, 2011: p. 3845-3856.
380. Peng, Y.-T., et al., *Multifaceted role of prohibitin in cell survival and apoptosis.* Apoptosis, 2015. **20**: p. 1135-1149.
381. Merkwirth, C., et al., *Prohibitins control cell proliferation and apoptosis by regulating OPA1-dependent cristae morphogenesis in mitochondria.* Genes & development, 2008. **22**(4): p. 476-488.
382. Steglich, G., W. Neupert, and T. Langer, *Prohibitins regulate membrane protein degradation by the m-AAA protease in mitochondria.* Molecular and cellular biology, 1999. **19**(5): p. 3435-3442.
383. Kasashima, K., et al., *Human prohibitin 1 maintains the organization and stability of the mitochondrial nucleoids.* Experimental Cell Research, 2008. **314**(5): p. 988-996.
384. Bourges, I., et al., *Structural organization of mitochondrial human complex I: role of the ND4 and ND5 mitochondria-encoded subunits and interaction with prohibitin.* Biochemical Journal, 2004. **383**(3): p. 491-499.
385. Tsutsumi, T., et al., *Proteomics analysis of mitochondrial proteins reveals overexpression of a mitochondrial protein chaperon, prohibitin, in cells expressing hepatitis C virus core protein.* Hepatology, 2009. **50**(2): p. 378-386.

386. Artal-Sanz, M. and N. Tavernarakis, *Prohibitin and mitochondrial biology.* Trends in Endocrinology & Metabolism, 2009. **20**(8): p. 394-401.
387. Wallace, D.C., *Mitochondrial diseases in man and mouse.* Science, 1999. **283**(5407): p. 1482-1488.
388. Osman, C., C. Merkwirth, and T. Langer, *Prohibitins and the functional compartmentalization of mitochondrial membranes.* Journal of cell science, 2009. **122**(21): p. 3823-3830.
389. Coates, P.J., et al., *Mammalian Prohibitin Proteins Respond to Mitochondrial Stress and Decrease during Cellular Senescence.* Experimental Cell Research, 2001. **265**(2): p. 262-273.
390. Kasashima, K., et al., *Mitochondrial functions and estrogen receptor-dependent nuclear translocation of pleiotropic human prohibitin 2.* Journal of Biological Chemistry, 2006. **281**(47): p. 36401-36410.
391. Henriques, B.J., et al., *Electron transfer flavoprotein and its role in mitochondrial energy metabolism in health and disease.* Gene, 2021. **776**: p. 145407.
392. Tonazzi, A., et al., *The Mitochondrial Carnitine Acyl-carnitine Carrier (SLC25A20): Molecular Mechanisms of Transport, Role in Redox Sensing and Interaction with Drugs.* Biomolecules, 2021. **11**(4): p. 521.
393. Borst, P., *The malate–aspartate shuttle (Borst cycle): How it started and developed into a major metabolic pathway.* IUBMB life, 2020. **72**(11): p. 2241-2259.
394. Kory, N., et al., *SFXN1 is a mitochondrial serine transporter required for one-carbon metabolism.* Science, 2018. **362**(6416).
395. Guna, A., et al., *MTCH2 is a mitochondrial outer membrane protein insertase.* Science, 2022. **378**(6617): p. 317-322.
396. Buzaglo-Azriel, L., et al., *Loss of Muscle MTCH2 Increases Whole-Body Energy Utilization and Protects from Diet-Induced Obesity.* Cell Reports, 2016. **14**(7): p. 1602-1610.
397. Maryanovich, M., et al., *An MTCH2 pathway repressing mitochondria metabolism regulates haematopoietic stem cell fate.* Nature communications, 2015. **6**(1): p. 7901.
398. Davies, S.M., et al., *MRPS27 is a pentatricopeptide repeat domain protein required for the translation of mitochondrially encoded proteins.* FEBS letters, 2012. **586**(20): p. 3555-3561.
399. Manna, S., *An overview of pentatricopeptide repeat proteins and their applications.* Biochimie, 2015. **113**: p. 93-99.
400. Hevler, J.F., et al., *MRPS36 provides a structural link in the eukaryotic 2-oxoglutarate dehydrogenase complex.* Open Biology, 2023. **13**(3): p. 220363.
401. Wang, Z., J. Cotney, and G.S. Shadel, *Human mitochondrial ribosomal protein MRPL12 interacts directly with mitochondrial RNA polymerase to modulate mitochondrial gene expression.* Journal of Biological Chemistry, 2007. **282**(17): p. 12610-12618.
402. Bartlett, J.D., et al., *The emerging role of p53 in exercise metabolism.* Sports medicine, 2014. **44**: p. 303-309.
403. Han, Y., et al., *P53 regulates mitochondrial biogenesis via transcriptionally induction of mitochondrial ribosomal protein L12.* Experimental Cell Research, 2022. **418**(1): p. 113249.
404. Siira, S.J., et al., *LRPPRC-mediated folding of the mitochondrial transcriptome.* Nature communications, 2017. **8**(1): p. 1532.
405. Cui, J., et al., *LRPPRC: a multifunctional protein involved in energy metabolism and human disease.* Frontiers in Physiology, 2019. **10**: p. 595.
406. Ruzzenente, B., et al., *LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs.* The EMBO journal, 2012. **31**(2): p. 443-456.
407. Sasarman, F., et al., *LRPPRC and SLIRP interact in a ribonucleoprotein complex that regulates posttranscriptional gene expression in mitochondria.* Molecular biology of the cell, 2010. **21**(8): p. 1315-1323.
408. Greene, N.P., et al., *Impaired exercise-induced mitochondrial biogenesis in the obese Zucker rat, despite PGC-1α induction, is due to compromised mitochondrial translation elongation.* American Journal of Physiology-Endocrinology and Metabolism, 2014. **306**(5): p. E503-E511.
409. Lee, D.E., et al., *Translational machinery of mitochondrial mRNA is promoted by physical activity in Western diet-induced obese mice.* Acta Physiologica, 2016. **218**(3): p. 167-177.

410. Kozhukhar, N. and M.F. Alexeyev, *Limited predictive value of TFAM in mitochondrial biogenesis.* Mitochondrion, 2019. **49**: p. 156-165.
411. Norrbom, J., et al., *Training response of mitochondrial transcription factors in human skeletal muscle.* Acta physiologica, 2010. **198**(1): p. 71-79.

## Appendix

All appendix files referred to in this thesis can be found in the *DropBox* folder:

https://www.dropbox.com/scl/fo/uece3kkygizz05qsam1g3/h?rlkey=ekhh4fprj3sof9708xbcyc151&dl=0