From the DEPARTMENT OF MOLECULAR MEDICINE AND SURGERY
Karolinska Institutet, Stockholm, Sweden

SYSTEMS BIOLOGY OF MITOCHONDRIAL DYSFUNCTION

Florian A. Schober

Karolinska Institutet

Stockholm 2021

Published by Karolinska Institutet

Printed by Universitetsservice US-AB, 2021

ISBN 978-91-8016-166-4

Cover illustration: The digital mitochondria, a puzzle. ASCII code (outer membrane) and ASCII code at 700 ppm mass tolerance, charge +1, carbamidomethylation (C) as fixed, methylation (KR) and oxidation (M) as variable modifications, 6% missing values (inner membrane).

SYSTEMS BIOLOGY OF MITOCHONDRIAL DYSFUNCTION

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Florian A. Schober

The thesis will be defended in public **online** and at Biomedicum in the **Ragnar Granit lecture hall**, Karolinska Institutet, Solnavägen 9, on **May 7, 2021 at 9:00**.

Principal Supervisor:
Dr. Anna Wredenberg
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Molecular Metabolism

Co-supervisors:
Dr. Christoph Freyer
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Molecular Metabolism

Prof. Anna Wedell
Karolinska Institutet
Department of Molecular Medicine and Surgery
Division of Inborn Errors of Metabolism

Opponent:
Dr. Christian Frezza
Cambridge University
MRC Cancer Unit

Examination Board:
Prof. Maria Ankacrona
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society
Division of Neurogeriatrics

Dr. Bianca Habermann
Aix-Marseille Université
Institut de Biologie du Développement de Marseille
Computational Biology Team

Prof. Roman Zubarev
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Physiological Chemistry I

The scientist does not study nature because it is useful to do so. He studies it because he takes pleasure in it, and he takes pleasure in it because it is beautiful. If nature were not beautiful it would not be worth knowing, and life would not be worth living.

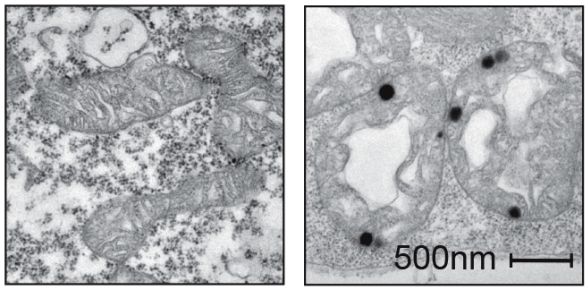
Henri Poincaré, Science and Method (1908)

To my dear family.

POPULAR SCIENCE SUMMARY OF THE THESIS

For their actions, growth and programmed death, human cells need energy that can be stored and distributed *via* the energy-rich molecule ATP. A large proportion of ATP is produced in the cellular organelle mitochondria, which are central platforms for converting metabolites into building blocks for cells. Their malfunction can have detrimental effects on human health. Inborn mitochondrial disease affects at least 1 in 5,000 individuals and causes a wide range of clinical symptoms. However, it is poorly understood how defective mitochondria cause disease on the level of cells and the whole body. A new set of *omics* methods and powerful computers now enable the measurement of thousands of molecules at the same time. In this thesis, I explore the use of *omics* tools in combination with genetic models of human disease in order to capture the systemic biological consequences of mitochondrial malfunction.

I shed light on small protein modifications in mitochondria, which are known to regulate protein function and stability. I map the smallest chemical modification called methylation on mitochondrial proteins, and I pinpoint processes that depend on methylation. By modeling patient-specific fruit flies I demonstrate that a simple dietary increase of the amino acid methionine was beneficial for the diseased flies. Moreover, I develop a method called ”SILAF” to identify novel proteins that are modified by phosphorylation and that are relevant to human disease. Furthermore, I make use of *omics* tools on human neuronal stem cells to find the cause for brain defects in patients with inborn errors of metabolism. Lastly, I monitor how mitochondria signal stress to the rest of the cell through leakage of RNA components in fly models.



Electron microscopy picture of fruit fly mitochondria, healthy (left) and sick (right).

These studies give insight into a novel layer of complexity in the cell, which is far from completely mapped and understood. Importantly, the translational perspective provides therapeutic avenues for patients with mitochondrial malfunction, which is also relevant to diseases with disturbed metabolism during ageing and in cancer.

SVENSK SAMMANFATTNING

Mänskliga celler behöver energi för att växa och genomföra deras specifika funktioner. ATP är en energirik molekyl som behövs i många cellulära reaktioner och som framförallt bildas i mitokondrierna. Dessa cellorganeller är centrala för ämnesomsättningen av molekylära byggstenar och nedsatt mitokondriefunktion kan leda till sjukdomar. Ungefär en av 5000 människor riskerar att utveckla mitokondriell sjukdom, som är en grupp av sjukdomar med en mycket varierad klinisk bild. Detta forskningsarbete behandlar frågan om hur dessa sjukdomar uppstår på cellnivå. En ny generation av metoder som kallas ”omics” kan nu användas för att mäta tusentals molekyler samtidigt. Jag använde dessa metoder tillsammans med genetiska modeller som bananflugor, möss och humana celler för att förstå systemiska konsekvenser av mitokondriella sjukdomar. Studierna visar på ett nytt sätt hur komplex regleringen av cellernas funktion är utöver den som sker på proteinnivå. Dessutom bidrar denna translationella forskning till nya terapikoncept för patienter med medfödd mitokondriell nedsättning och utökar vår förståelse för bristfälliga metabola processer i cancer och i den normala åldrandeprocessen.

DEUTSCHE ZUSAMMENFASSUNG

Menschliche Zellen benötigen Energie, um wachsen und ihre Funktion aufrechterhalten zu können. Das Molekül ATP ist ein energiereicher Zwischenspeicher, der in einer Vielzahl zellulärer Reaktionen verwendet und hauptsächlich in Mitochondrien produziert wird. Diese Zellorganellen sind zentrale Drehscheiben für den Stoffwechsel von zellulären Bausteinen. Funktionieren sie nicht richtig, kann dies zu Erkrankungen führen. Ungefähr einer von 5,000 Menschen ist von einer mitochondrialen Erkrankung betroffen, die sich in einer vielzahl klinischer Symptome manifestieren kann. Wie diese Erkrankungen auf zellulärer Ebene entstehen ist Gegenstand dieser Forschungsarbeit. Eine neue Generation sogenannter ”omics” Methoden erlaubt es, tausende Moleküle gleichzeitig zu vermessen. Ich verwende diese Methoden zusammen mit Fliegen-, Maus- und zellulären Modellen, um die systemischen Auswirkungen schadhafter Mitochondrien zu verstehen. Diese Studien zeigen auf eine neue Art und Weise, wie komplex die Regulierung von Stoffwechselvorgängen abläuft. Außerdem bieten die translationalen Ansätze neue therapeutische Perspektiven für Patienten mit mitochondrialen Erkrankungen und bei Stoffwechselstörungen in Alterungsprozessen und Krebs.

ABSTRACT

The human body consumes vast amounts of metabolites that are transformed into one another, modified to useful building blocks and broken down to harvest their energy. Mitochondria are at the core of this metabolic turnover and oxidative phosphorylation provides most cellular ATP in almost all human tissues. Despite the variability of metabolite and oxygen supply, mitochondria can readily adapt to their cellular niche. This requires a general flexibility in the expression patterns of the roughly 1,150 mitochondrial proteins, and fine tuning of protein actions. Failure to meet the cellular metabolic demand causes a wide range of tissue specific symptoms in human.

This thesis explores the use of high-throughput *omics* techniques in understanding enzymatic remodeling during mitochondrial disease. In the first part, I introduce the reader to the concepts of systems biology, and broadly discuss current methods and tools with a focus on mitochondria. Then, I describe the fine-tuning of mitochondrial function by post-translational modifications, in particular protein methylation and phosphorylation, and how this links to the large metabolic network of the one-carbon cycle. I conclude with primers on mitochondria in development and mitochondrial RNA metabolism. After that, I will present four publications in light of the discussed concepts and methods:

In **study I**, we find that the highly abundant metabolite *S*-adenosylmethionine is indispensable for mitochondrial function. Its cytoplasmic production controls mitochondrial function by regulating iron-sulfur clusters biosynthesis and stability of the electron transport chain complex I, which has implications during ageing and cancer development. We apply a novel labeling method and mass spectrometry-based proteomics to identify 205 high-confidence methylation sites on mitochondrial proteins in fruit flies, and validate several by targeted proteomics in mouse and human.

In **study II**, we describe SILAF, a novel and highly efficient method to label amino acids in the fruit fly proteome. We exploit SILAF to characterize the mitochondrial phosphoproteome in a fly model of mitochondrial disease, and we pinpoint two regulatory phosphorylation sites.

In **study III**, we investigate the role of the scaffold protein SQSTM1/p62 in neuronal development. We find that the protein is required for differentiation of patient-derived neuronal epithelial stem cells, caused by an impaired switch from glycolytic to oxidative metabolism.

In **study IV**, we use various fruit fly models to examine the interactions of proteins in mitochondrial RNA metabolism. Using transcriptomics, we identify leakage of double-stranded RNA into the cytosol when mitochondrial RNA degradation is impaired and we suggest that this contributes to increased susceptibility to infection upon mitochondrial dysfunction.

Our studies take a novel view on mitochondrial dysfunction, and our post-translational modification screens give insight into a novel layer of complexity in the cell. The studies expose the opportunities and challenges of data-driven life science and can serve as a primer towards a digital representation of mitochondrial disease.

LIST OF SCIENTIFIC PAPERS

I. **Schober FA***, Moore D*, Atanassov I, Moedas MF, Clemente P, Végvári Á, El-Fissi N, Filograna R, Bucher A-L, Hinze Y, The M, Hedman E, Chernogubova E, Begzati A, Wibom R, Jain M, Nilsson R, Käll L, Wedell A, Freyer C†, & Wredenberg A† (2021). The One-Carbon Pool Controls Mitochondrial Energy Metabolism via Complex I and Iron-Sulfur Clusters. *Science Advances*, 7: eabf0717.

II. **Schober FA***, Atanassov I*,†, Moore D, Calvo-Garrido J, Moedas MF, Wedell A, Freyer C†, & Wredenberg A† (2021). Stable Isotope Labeling of Amino Acids in Flies (SILAF) Reveals Differential Phosphorylation of Mitochondrial Proteins Upon Loss of OXPHOS Subunits. *Molecular and Cellular Proteomics*, 20: 100065.

III. Calvo-Garrido J*, Maffezzini C*, **Schober FA**, Clemente P, Uhlin E, Kele M, Stranneheim H, Lesko N, Bruhn H, Svenningsson P, Falk A, Wedell A, Freyer C†, & Wredenberg A† (2019). SQSTM1/p62-Directed Metabolic Reprogramming Is Essential for Normal Neurodifferentiation. *Stem Cell Reports*, 12(4): 696–711.

IV. Pajak A*, Laine I*, Clemente P, El-Fissi N, **Schober FA**, Maffezzini C, Calvo-Garrido J, Wibom R, Filograna R, Dhir A, Wedell A, Freyer C†, & Wredenberg A† (2019). Defects of mitochondrial RNA turnover lead to the accumulation of double-stranded RNA in vivo. *PLOS Genetics*, 15(7): e1008240.

* These authors contributed equally.

† Corresponding authors.

SCIENTIFIC PAPERS NOT INCLUDED IN THE THESIS

- Engvall M, Kawasaki A, Carelli V, Wibom R, Bruhn H, Lesko N, **Schober FA**, Wredenberg A, Wedell A, Traisk F† (2021). A novel mutation in the mitochondrial MT-ND5 gene is associated with Leber hereditary optic neuropathy (LHON). *Frontiers in Neurology*, 12: 652590.
- Bruhn H†, Samuelsson K, **Schober FA**, Engvall M, Lesko N, Wibom R, Nennesmo I, Calvo-Garrido J, Press R, Stranneheim H, Freyer C, Wedell A*, & Wredenberg A* (2021). A novel mutation m.10372A>G in MT-ND3 causing sensorimotor axonal polyneuropathy. *Neurology: Genetics*, 7: e566.
- Correia SP*, Moedas MF*, Naess K, Bruhn H, Maffezzini C, Calvo-Garrido J, Lesko N, Wibom R, **Schober FA**, Jemt A, Stranneheim H, Freyer C, Wedell A†, & Wredenberg A† (2021). Severe congenital lactic acidosis and hypertrophic cardiomyopathy caused by an intronic variant in NDUFB7. *Human mutation*: 1-7.
- Cipullo M, Pearce SF, Lopez Sanchez IG, Gopalakrishna S, Krüger A, **Schober FA**, Busch JD, Li X, Wredenberg A, Atanassov I, & Rorbach J† (2021). Human GTPBP5 is involved in the late stage of mitoribosome large subunit assembly. *Nucleic Acids Research*, 49(1): 354–370.
- **Schober FA**, Atanassov I, Freyer C, & Wredenberg A† (2021). Mitochondrial Gene Expression, Methods and Protocols. *Methods in Molecular Biology*, 2192: 75–87.
- Alsina D*, Lytovchenko O*, Schab A, Atanassov I, **Schober FA**, Jiang M, Koolmeister C, Wedell A, Taylor RW, Wredenberg A, & Larsson NG† (2020). FBXL4 deficiency increases mitochondrial removal by autophagy. *EMBO Mol. Med.*, 12(7):e11659.
- Gopalakrishna S*, Pearce SF*, Dinan AM, **Schober FA**, Cipullo M, Spåhr H, Khawaja A, Maffezzini C, Freyer C, Wredenberg A, Atanassov I, Firth AE, & Rorbach J† (2019). C6orf203 is an RNA-binding protein involved in mitochondrial protein synthesis. *Nucleic Acids Research*, 47(17): 9386–9399.
- Jiang S, Koolmeister C, Misic J, Siira S, Kühl I, Ramos ES, Miranda M, Jiang M, Posse V, Lytovchenko O, Atanassov I, **Schober FA**, Wibom R, Hultenby K, Milenkovic D, Gustafsson CM, Filipovska A, & Larsson NG† (2019). TEFM regulates both transcription elongation and RNA processing in mitochondria. *EMBO Reports*, 20(6): e48101.
- Siibak T*, Clemente P*, Bratic A, Bruhn H, Kauppila TES, Macao B, **Schober FA**, Lesko N, Wibom R, Naess K, Nennesmo I, Wedell A, Peter B, Freyer C, Falkenberg M†, & Wredenberg A† (2017). A multi-systemic mitochondrial disorder due to a dominant p.Y955H disease variant in DNA polymerase gamma. *Human Molecular Genetics*, 26(13): 2515–2525.

* These authors contributed equally. † Corresponding author(s).

CONTENTS

ACRONYMS

1C one-carbon

5mC 5-methylcytosine

ANT ADP/ATP translocase

BCKDH branched-chain ketoacid dehydrogenase

bsf bicoid stability factor

cyt *c* cytochrome *c*

Dm *Drosophila melanogaster*

ETF β electron transferring flavoprotein (ETF) β

FDR false discovery rate

GCS glycine cleavage system

GSEA Gene Set Enrichment Analysis

KGDH α-ketoglutarate dehydrogenase

LC-MS/MS liquid chromatography, coupled to tandem mass spectrometry

LDHA lactate dehydrogenase A

LRPPRC leucine-rich pentatricopeptide-repeat containing

m/z mass-to-charge

MAT methionine adenosyltransferase

MEF mouse embryonic fibroblast

mitoSAM mitochondrial *S*-adenosylmethionine

MOAC metal oxide affinity chromatography

MS mass spectrometry

mtDNA mitochondrial DNA

MTPAP mitochondrial poly(A) polymerase

NAD nicotinamide adenine dinucleotide

NESC neuroepithelial stem cells

ORA Over-Representation Analysis

OXPHOS oxidative phosphorylation

PC phosphatidylcholine

PDH pyruvate dehydrogenase complex

PKA cAMP-dependent protein kinase

PNPase polynucleotide phosphorylase

PTMs post-translational modifications

Q ubiquinone

redox reduction/oxidation

RNA-Seq RNA sequencing

ROS reactive oxygen species

SAH *S*-adenosylhomocysteine

SAM *S*-adenosylmethionine

SAMC *S*-adenosylmethionine carrier protein

SILAC stable isotope labeling of amino acids in cell culture

SILAF stable isotope labeling of amino acids in flies

SQSTM1/p62 sequestome 1

TCA tricarboxylic acid

TFAM mitochondrial transcription factor A

THF tetrahydrofolate

TiO_2 titanium dioxide

The world is everything that is the case.

Ludwig Wittgenstein, Tractatus Logico-Philosophicus (1922)

1 INTRODUCTION

1.1 What is systems biology?

A doctoral thesis with *systems biology* in its title is expected to be clear about this term. Provokingly, it cannot be. After the human genome project had been completed in 2001, the emergence of this linguistic compound coincided with a rapid progression of computing power and human imagination for how to use computational resources. This has opened up new possibilities for this young field, and hence has continuously modulated and shaped the definition of the term until today [1].

A unifying aim of systems biology is to understand the larger picture of a living entity as a sum of individual reactions in order to explain its behavior. Thereby, the focus can be a cell, tissue, a whole organism or beyond. To get an understanding of such a system, it is necessary to have knowledge about the individual components and their interactions. Rainer Breitling suggested three aesthetic qualities of systems biology: *diversity*, *simplicity* and *complexity* [2]. The term *diversity* appreciates the elements of a system and recognizes that these are different from one another, for instance the proteins in a cell or the cell types in a tissue. As Darwin proposed, diversity is a product of evolutionary selection, and thus the system itself is blind. Researchers aim at finding patterns and causation in systems that follow a number of *simplistic* working models in order to understand how the entity works. The most striking difference of a systems compared to a molecular biology approach is that systems biology strives towards assembling many simple models into complex networks in order to mimic the behaviour of a living unit, rather than to reduce a complex system to simple building blocks. Following this school of thought, systems biology emerges as a largely data-driven science that marries biology with technological development and computational tools [3].

Cells require vast amounts of energy for homoeostasis and growth, and to fulfil specialized functions like secretion of enzymes or propagation of action potentials. The bioenergetic demands of these processes limit the complexity of prokaryotes [4]. Following the current dogma, the endosymbiotic integration of α-proteobacteria that formed the early mitochondria was a key evolutionary event that allowed the transition from unicellular prokaryotes to multicellular eukaryotes [5]. This thesis work studies the consequences of dysfunctional mitochondria in a multicellular context. The aim is to adapt and apply tools to map the diversity of mitochondria, to generate hypotheses on the etiology of mitochondrial dysfunction, and to build novel functional networks that are useful in understanding human disease.

1.2 The power of eukaryotes – mitochondria

Mitochondria are organelles that form dynamic networks in eukaryotic cells. Under aerobic conditions, they are able to generate large amounts of ATP, which serves as an energetic currency in cellular processes including enzymatic reactions, cytoskeletal remodeling and maintenance of membrane potentials. The α-proteobacterial ancestors of mitochondria [5] lost and gained numerous genes and functions in exchange with the host cell during the course of about 1.5 to 2 billion years [6]. Yet, almost all eukaryotic mitochondria retained some bacterial hallmarks, including a double membrane and an origin-of-replication bearing, coding DNA molecule.

The size, shape and genetic content of mitochondrial DNA (mtDNA) varies substantially across eukaryotic kingdoms [7]. The human mitochondrial matrix harbors circular, double-stranded mtDNA molecules, which are 16.3 kb long and thus small in contrast to the nuclear genome. Each organelle contains multiple copies of mtDNA. Depending on the cell type [8, 9], the copy number ranges from $3x10^3$ per fibroblast [10] to 10^5 in one oocyte [11]. The mitochondrial genome contains information for 13 proteins of the ATP-generating oxidative phosphorylation (OXPHOS) system, and further for two ribosomal RNAs and 22 transfer RNAs that form parts of an intra-mitochondrial protein biosynthesis machinery [12]. However, the genetic information for most of the mitochondrial proteins including the largest part of the OXPHOS system is encoded in the nuclear genome. Most of the 1,136 recognized human mitochondrial proteins [13] are synthesized on cytoplasmic ribosomes and transported as unfolded polypeptide chains across one or both organellar membranes by specialized machineries [14].

The mammalian OXPHOS system consists of five multiprotein complexes that convert a reductive potential stored in electron carriers into ATP (**Figure 1**). The most prominent reduction/oxidation (redox) reactions upstream of OXPHOS occur as part of the tricarboxylic acid (TCA) cycle, which integrates multiple catabolic and anabolic pathways, such as glycolysis, fatty acid degradation and amino acid turnover. The TCA cycle of a non-fasting, non-ischaemic cell reduces both free oxidized nicotinamide adenine dinucleotide (NAD) to NADH and covalently bound flavin adenine dinucleotide (FAD) to $FADH_2$ [15]. These, in turn, reduce iron-sulfur clusters in complex I and complex II, respectively. Driven by a progressive increase of cluster affinity, electrons flow through ubiquinone, complex III and cytochrome *c* (cyt *c*) to complex IV, where the reaction with oxygen yields water. The electric mobility of electrons is used by all OXPHOS complexes but complex II to establish a negative chemiosmotic gradient across the inner mitochondrial membrane. This drives a proton flux through complex V into the mitochondrial matrix that converts kinetic into chemical energy as ATP. Electrons can enter the OXPHOS system not only *via* the TCA cycle, but multiple further redox reactions in the cell are known that either generate NADH or reduce ubiquinone and shuttle electrons directly to complex III.

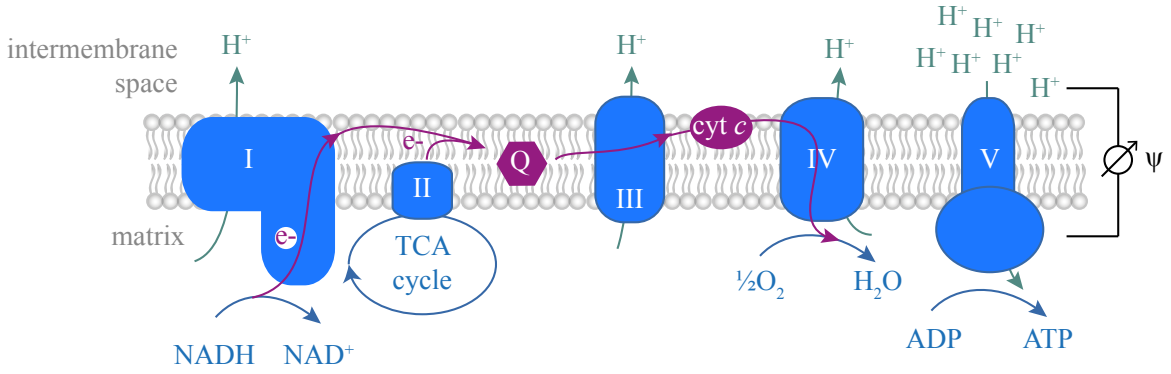


Figure 1: Schematic view of the OXPHOS system. By oxidation of NADH and TCA cycle substrates, electrons enter the electron transport chain in the inner mitochondrial membrane. They are passed along redox centers in complex I to IV, ubiquinone (Q) and cyt *c* onto oxygen. The conformational change of OXPHOS subunits pumps protons (H^+) into the intermembrane space and establishes a concentration gradient and membrane potential Ψ that is harvested at complex V to produce ATP.

1.3 Monogenic mitochondrial disorders

The stoichiometry of OXPHOS complexes is tightly regulated [16] and OXPHOS activity meets the energetic demands of a certain cell type, developmental phase or circadian activity profile. It is thus not surprising that the clinical picture of monogenic mitochondrial disorders, which are caused by either mtDNA or nuclear DNA mutations, is extremely heterogeneous. Any age group can be affected, but symptoms arise predominantly in the first three years of life [17]. The combined prevalence of monogenic mitochondrial diseases is at least one case per 4,300 individuals reported for North East England [18], and rises profoundly in regions with founder mutations and/or intertwined family trees [19]. Any organ can be affected, although symptoms are primarily evident in energy-demanding tissues: brain and optical nerves, heart and skeletal muscle, liver and lung [17]. In certain cases, symptoms can be grouped to syndromes, for instance Leigh syndrome, which predominantly presents in young children with brain lesions and a subsequent decline in cognitive functions [20]. Since the technical and economic advancement of whole-exome and whole-genome sequencing in combination with global efforts for patient report databases, a biochemical defect in OXPHOS function can be linked to a disease-causing gene in more than 60% of all cases [17].

Yet, the diagnosis of mitochondrial diseases remains complex. Beyond monogenic disorders, a growing body of evidence points towards a mitochondrial involvement also in multifactorial diseases like amyotrophic lateral sclerosis [21], Parkinson's [22] or autoimmune diseases [23]. The underlying molecular observations are not limited to dysfunctions of the OXPHOS system, but include the various roles of mitochondria in calcium homeostasis, metabolism, iron-sulfur cluster biosynthesis and apoptosis [24].

2 THE SYSTEMS BIOLOGIST'S TOOLBOX

2.1 *omics* technologies

Data-driven life science has been enabled by methodologies with the suffix *omics*, derived from the word *genome* introduced in 1920 [25], that aim at comprehensively quantifying cellular states in combination with computational data processing. The last two decades have seen a profound decrease in costs, and several *omics* techniques such as whole-genome sequencing are used in clinical practice today [26]. Each method is able to capture one level of biological complexity and multiple *omes* complement each other. A selection of these that are relevant to this thesis work are discussed below.

2.1.1 Transcriptomics

The term *transcriptomics* broadly comprises all methods that aim at quantifying the sum of RNA in a cell, including methods like expressed sequence tags [27] and microarrays. This thesis work only covers massive parallel RNA sequencing (RNA-Seq), which in itself has evolved into approximately 100 distinct methods that enable insight into various aspects of cellular transcriptomes [28]. Most RNA-Seq applications aim at differential expression analysis, as it is used herein, alongside more specialized applications such as isoform-sensitive [29], single-cell [30], spatial [31] and structural RNA-Seq [32].

Three major platforms have emerged to provide various kinds of biological insight: (1) short-read, (2) long-read cDNA and (3) long-read direct RNA-Seq. While direct RNA-Seq is performed on RNA without conversion, short- and long-read sequencing rely on enzymatic cDNA synthesis with a reverse transcriptase. Currently, short-read sequencing is the most robust technique for high-throughput sequencing, and biases in both library preparation and sequencing are well understood [28]. However, owing to its use of short fragment sizes with less than 200 bases, the identification of transcript isoforms can be compromised. This is of particular relevance for assembling reads from very long RNA species including polycistronic mitochondrial transcripts in disease conditions with processing defects. In addition to whole exome and whole genome sequencing [33], transcriptomics is emerging as a valuable tool in diagnosing rare diseases. It is the most informative *omics* technology in cases of cryptic or erroneous splice sites and allelic exclusion [34] when the bioinformatic pipeline allows alignments against whole genome libraries instead of cDNA libraries.

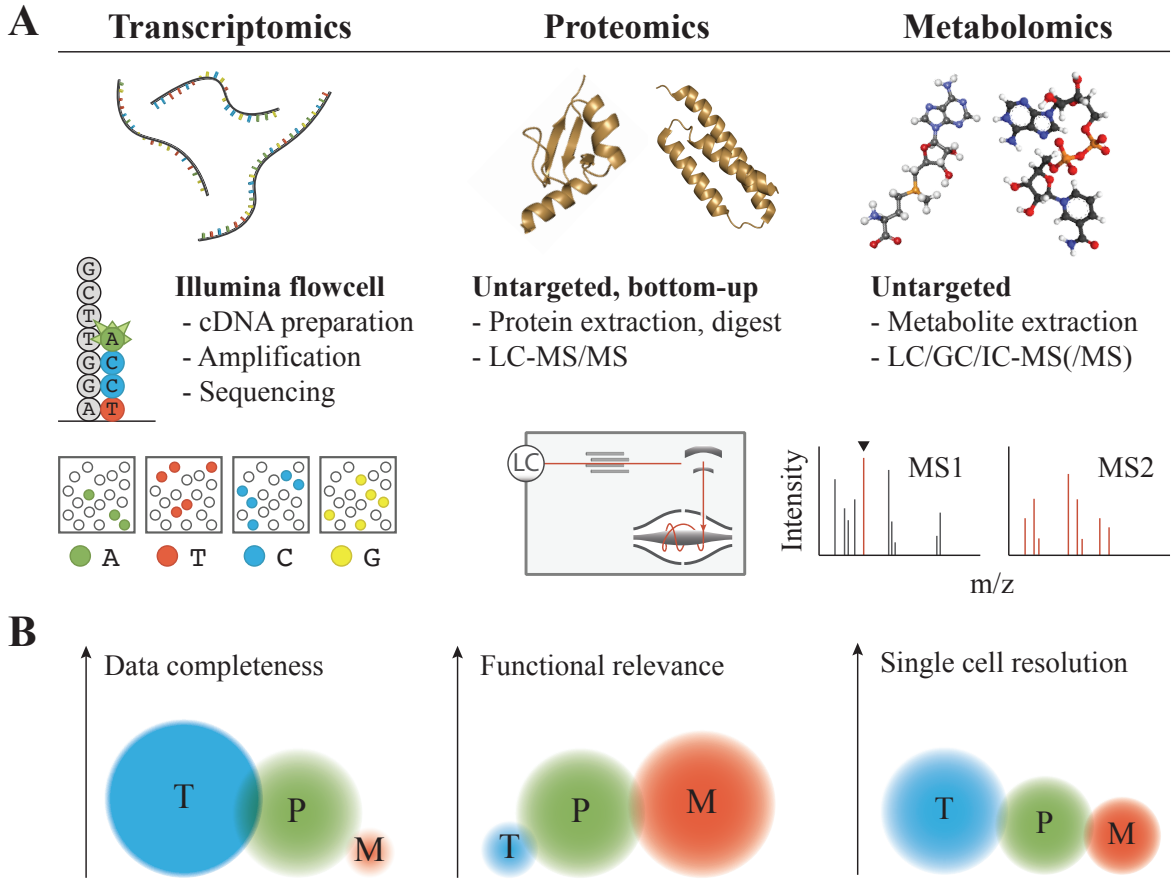


Figure 2: The *omics* trilogy. (A) A short account of methods used in this thesis work. Top to bottom: sample type, technology and key concepts, technical schematic. **(B)** The author's subjective view on key parameters of the methods as they perform today. T: Transcriptomics; P: Proteomics; M: Metabolomics.

2.1.2 Proteomics

While the transcriptome contains the plans for the future functional building blocks of a cell, the proteome reflects the current set of working protein units. Therefore, an accurate and complete proteome quantification is more informative about the state of a cell than its transcriptome. However, the challenge of quantifying proteins is inherently more complex compared to RNA, currently limiting the amount of information that can be obtained from a single proteomics experiment.

About 20,000 human genes give rise to potentially 230,000 different transcripts according to the 2020 ENSEMBL release, including variants and non-coding genes [35]. Of those, 50,000 are protein coding mRNAs [36]. The UniProt database lists 76,000 protein entries at the time of writing [37], but this number grows exponentially when post-translational modifications (PTMs) are taken into account [38]. Although post-transcriptional modification of RNA occurs, post-translational modification of proteins seems to be a much more widespread regulatory mechanism. The unique molecular form of a protein is termed *proteoform* [39], and one [40] to six million [38] proteoforms have been estimated per cell. Determining the molecular identity and quantity of a proteoform requires either an affinity-based method, in

which the specificity of the affine reagent must be known, or a destructive approach that disassembles the molecule in a coordinated manner. Affinity-based methods like antibody-based Western blotting or immunohistochemistry are extremely valuable research tools in molecular biology, and large screening efforts like the Human Protein Atlas combine low-multiplexed antibody-based protein detection with high-throughput microscopy [41]. However, these approaches are limited in their potential to find novel proteoforms.

The advent of proteomics using liquid chromatography, coupled to tandem mass spectrometry (LC-MS/MS) opened up possibilities to quantify thousands of proteins within a short time. LC-MS/MS is a powerful tool to identify novel proteoforms [42]. Yet, the method is constrained by the complexity of cellular samples. The dynamic range of the proteome, that is the quantity range from the least to the most abundant protein, by far exceeds the dynamic range of the transcriptome. While the transcriptome is estimated to range from 0.5 to 50,000 transcripts per cell, arising to five orders of magnitude [43, 44], the cellular proteome spans approximately one copy to ten million protein copies, or an estimate of seven to eight orders of magnitude [44, 45]. Furthermore, while the four building blocks of RNA and cDNA can be reproduced one by one through Watson-Crick base pairing in millions of simultaneous reactions, current methods to quantify proteins cannot be efficiently parallelized yet and the signal of low-abundance proteins cannot be amplified. Therefore, although the most abundant proteins can be quantified by LC-MS/MS within hours, capturing the low abundance range is time consuming and laborious [45]. Despite the methodological constraints, the proteome reflects the state of a cell more accurately than the transcriptome [44] (see also chapter 2.6), and much research is being performed to solve the problem of dynamic ranges.

This thesis work uses discovery- and validation-oriented mass spectrometry (MS)-based proteomics techniques to pinpoint novel cellular proteoforms and to map PTMs. A variety of approaches have been developed for ionization, mass selection and detection, and many are in active development to match specialized applications like single-cell proteomics [46]. By far most MS-based proteomics experiments follow a bottom-up work flow with LC-MS/MS that requires digestion of proteins into peptides [47]. In brief, proteins are extracted from a sample with chaotropic and reducing agents that destroy higher order protein structures. The proteins are digested into peptides using an endoproteinase with a known cleavage pattern. The most commonly used enzyme, trypsin, cuts C-terminally after lysine and arginine and produces a set of peptides with an optimal length range for MS-based protein identification [48]. However, depending on the application, a variety of endoproteinases can be used [49]. For instance, study I further relies on chymotrypsin that increases coverage of hydrophobic amino acid stretches, which are prevalent in mitochondrial membrane proteins. After digest, the produced complex peptide mixture is subjected to LC-MS/MS that combines separation by hydrophobicity on a reversed-phase high pressure liquid chromatography column with ionization, separation by mass-to-charge (m/z) of charged peptides, and selective fragmentation of a subset of abundant peptides. The intensities per m/z beam of both the intact ionized pep-

tide and its fragments, called spectra, are recorded. Peptide sequences can then be assigned by matching measured spectra with *in silico* predicted fragmentation patterns using a library of protein sequences.

As the mass of a post-translational modification is known, the predicted fragmentation patterns can be extended by a possible number of PTMs per peptide. However, this increases the search space dramatically, and potentially 98,000 proteoforms of human histone H4 exist when just considering the 13 most common PTMs [40]. More complex searches impose a computational problem and decrease the sensitivity and specificity of searches at a set false discovery rate (FDR), in part due to similar or identical masses of *de facto* different peptides. Therefore, there is a crucial need for both methodological and computational development, and for accurate libraries of proteoforms in order to limit and speed up spectral searches to biologically relevant hits.

2.1.3 Metabolomics

Metabolites are the downstream products of genes, transcripts and proteins, causing them to be the most dynamic component in an intertwined cellular system [50]. Metabolic networks provide the building blocks and energy for cell maintenance and growth, and shape cellular behavior by modulating signaling pathways [51]. In 1965, the Nobel prize winner Ernst Boris Chain listed ”the study of enzymatic reactions leading to the recognition of metabolic pathways” [52] as one of the ten landmarks in the history of biochemical research. Although Chain did not state that this discovery process was completed, the metabolic maps that are taught in biochemistry classes today mostly date back to the 1960s [53]. Only recently, technical advancements lead to the realisation that the metabolome is far more complex than anticipated. This has caused new excitement to discover novel metabolites and metabolic pathways [53].

A number of methods are available to map and quantify metabolomes, although LC-MS is the predominant technique for discovery metabolomics [53]. As above, the readout is the m/z ratio of molecules after chemical separation, ionization and mass selection. The signals, called features, are then mapped against a library of known metabolites. This is currently the bottleneck of metabolomics, in part because the full content of metabolite libraries is not known, which renders the metabolomics data scientist blindfolded towards the upper limit of the reference database [54]. In addition, while 21 amino acids constitute the human proteome, the structural building blocks of metabolites are combinations of atoms, with an exponentially higher likelihood for identical or very similar masses of two compounds that are actually structurally different. Thus, powerful LC techniques are pivotal to separate two chemically distinct compounds. Indeed, approaches to match computationally predicted with real-world retention times are transforming the field into a treasure box for biological discovery [55].

2.2 Omics beyond steady state

The quantification of transcripts, proteins or metabolites is a snapshot of cellular components with reasonable information when the studied model system is at steady state, for instance immortalized cell lines that have been kept in culture without nutrient restriction. However, once the cells are stimulated by, for instance, changing culture media or through genetic insult, they will adapt to meet the new demands. Capturing these dynamics requires additional dimensions in an experiment. This is of particular interest in search for the primary cause of monogenic disease, in which the molecular role of a dysfunctional protein is yet unclear.

Experimentally, the extra dimension can be added either through consecutive sampling at various time points or by introducing a pseudo-temporal axis through sampling at the same time but from different biological specimens. A powerful tool for the first approach involves stable isotope-labeled tracer molecules and MS to capture proteome and metabolome dynamics. For the popular application stable isotope labeling of amino acids in cell culture (SILAC), cells are grown in defined culture media in which all or selected ”light” amino acids are exchanged for biotechnologically synthesized and chromatographically enriched ”heavy” amino acids [56]. In these, ^{1}H hydrogen, ^{12}C carbon and/or ^{14}N nitrogen atoms are exchanged to non-radioactive ^{2}H deuterium, ^{13}C and ^{15}N isotopes, respectively. After media change, the heavy labels are incorporated into proteins and amino acid-derived compounds. SILAC can be used to quantify protein turnover during metabolic pulse labeling, as the proportion of labeled amino acids in a peptide backbone is dependent on *de novo* synthesis of its protein in the presence of the label [44].

In principle any labeled compound can be used, with the limitation that the light counterpart must be either completely replaceable in the culture medium or the baseline heavy-to-light ratio must be known prior to the experiment. For instance, custom-prepared media with metabolic ^{13}C-labeled precursors were used to deeply label the metabolome of human cancer cell lines and gave insight into the activity of metabolic pathways and the origins of secondary metabolite components [57]. Thus, by measuring the heavy-to-light ratio at various timepoints, the dynamic flow of metabolites, termed flux, in a defined metabolic network can be reconstructed [58].

2.3 Model systems

2.3.1 Cell culture

Homogenous cell populations that are kept at precisely monitored conditions are highly useful models to reduce the complexity of biological problems and to minimize data background noise. Therefore, fewer replicates are required for robust statistics, which decreases analysis time and costs. A wealth of methods has been optimized for cells in culture, including RNA interference, over-expression systems and manipulations on genome level.

Cell culture models can be distinguished by their provenance and possibility to maintain them over longer periods of time. Cell lines originate from tumor tissue or have been immortalized by genetic inhibition of cell cycle blocks [59]. Thus, large quantities of genetically identical cells can be grown within a short time, which allows multiplexing of genetic or drug screens [60]. While cell lines are a widely used and valuable research tool, cellular metabolism has adapted to sustain continuous growth under set culture conditions and data has to be interpreted with great care. By contrast, primary cells are taken from an intact tissue and maintain their potency state. Patient-derived primary cells like fibroblasts or myoblasts are the only ethically acceptable route to test hypotheses that require experimental intervention in a set (epi-)genetic background. Since the discovery that pluripotency can be induced in mouse embryonic and adult fibroblasts by expression of four transcription factors [61], cell models resembling a multitude of tissues can now be evoked in culture from starting material of a different lineage such as fibroblasts. This is of great help in studying diseases that affect the central nervous system [62, 63] including inborn mitochondrial and neurodegenerative disorders [64]. For instance, patient fibroblasts can be reverted to pluripotent neuroepithelial stem cells (NESC) that give rise to mature neurons when a specific differentiation program is induced (see chapter 9.3).

Reprogrammed cells like NESC offer a unique window to study molecular mechanisms of the human brain. However, the major drawback of cell culture models is the lack of tissue complexity. As an example, the metabolism of neurons is tightly linked to surrounding astrocytes that release lactate upon neuronal activity [65]. Therefore, a systems view of cells in culture has to take the simplicity of the model and its inaccuracy in mimicking a tissue setting into account. The highly artificial composition of cell culture media causes adaptations of metabolic pathways so that disease states can be either hidden or exacerbated in *omics* datasets, and is has been shown that components of culture media affect gene essentiality [66]. Furthermore, cells in almost all organs are confronted with much lower oxygen concentration than atmospheric 21% in many tissue culture experiments, and pathways related to mitochondria, iron-sulfur cluster and lipid metabolism are particularly susceptible to inadequate oxygen partial pressures [67]. One solution has been the development of three dimensional cell culture systems resembling human organs. While the reproducibility of these human organoids still suffers from the variety of protocols and standards used in different labs [68], they have the potential to accurately resemble the physiology of organs and complement or even replace animal-based research.

2.3.2 Fruit flies

Since Thomas Hunt Morgan and colleagues established the fruit fly *Drosophila melanogaster (Dm)* as a model system more than one hundred years ago, a wealth of tools have been developed for this working horse of biological research [69, 70]. Numerous hallmark findings have been made especially in the field of developmental biology, which left traces in *Dm* gene

names: *white*, *glass bottom boat* or *toll-like receptors* are gene names, but describe the phenotype of flies deficient for that particular gene. In the field of mitochondrial disorders, the fruit fly has proven to be a powerful *in vivo* model to understand the function of novel disease genes [71, 72] due to short generation times, comparably low maintenance costs and, most importantly, the wealth of genetic tools. Commercially available fly lines allow the knock-down or over-expression of most fly genes, which makes models for genes of interest rapidly available in the laboratory [73].

In 2000, the fly was the second almost completely sequenced multicellular organism [74], and its prospects were therefore promising to become a valuable tool in the era of high-throughput biology. The genome has been fully annotated since 2003 [75] and a fly peptide atlas has been established [76] that forms the basis for larger transcriptomic and proteomic screens. Several developmental proteomes have been established by label-free quantification, of which the latest comprises 15 developmental stages [77]. The most studied PTM in flies is phosphorylation, which has been assessed in embryos [78] and as a differential phosphoproteome in cold exposed flies compared to controls [79]. Differential proteomics in flies, however, has been less applied, in part due to missing tools such as powerful labeling methods. SILAC is the gold-standard for experimental multi-step procedures involving LC-MS/MS and allows accurate PTM quantification [80]. Thereby, organisms are grown on medium that contains non-radioactive, heavy amino acid isotopes, mostly lysine and arginine. SILAC for flies has been established by several research groups [81–83], and all approaches were based on feeding labeled yeast. In practice, these methods are financially and technically difficult [84] and undesirable conversion of amino acids has been shown [83], which depresses the signal of a labeled fraction and compromises the reliability of accurate peptide quantification.

2.3.3 Mouse

Metabolism is highly tissue specific on transcript, protein and metabolite [41, 85, 86] level. Cell culture models can only in part reflect this metabolic diversity and the various clinical symptoms of inborn errors of metabolism. The fruit fly is a valuable compromise between tissue complexity and biological simplicity. However, individual tissues are difficult and laborious to dissect, and tissue organization and function are often fundamentally different to humans. Although ethically more challenging (see chapter 8.7), the mouse allows modeling of human disease in a much more intricate manner. A wealth of genetic tools and well-established research infrastructure make the mouse an obvious choice for many questions regarding tissue-specificity of metabolic dysfunction. Several valuable mouse models of mitochondrial disease have been developed. As an example, knock-out of the mitochondrial transcription factor A (TFAM) causes severe depletion of mtDNA in mice [87]. Tissue specific deletions of TFAM have been used to study the consequences of mitochondrial dysfunction in various organs, including muscle [88], heart [89] and T-cells [90].

When a full-body knock-out of a gene is lethal at embryonic stage, the access to samples is limited and alternative experimental approaches are favorable. Cre-expressing lines are frequently used to knock-out loxP-site flanked genetic segments in a specific tissue, which depends on the chosen Cre-promoter, and these models are often viable. Additionally, mouse embryonic fibroblast (MEF) cells can be obtained from embryos by dissection. After immortalization, MEF cells can be kept in tissue culture with the advantages and disadvantages of cell culture discussed above. Yet, MEF cells can be seen as a valuable and ethically more acceptable model during the discovery phase of a project.

2.4 The dream of the digital cell

For systems biology to be useful in our quest to understand life, its goal must be to represent cells in a digital way so we can make predictions about cellular actions under set circumstances. In a clinical setting, for example, a useful binary model of a cell can pinpoint the genetic mutation that causes mitochondrial dysfunction from whole genome data and it can accurately predict the aberrant metabolic fluxes. This also applies for proteins whose functions are not yet known, and this thesis work has contributed to elucidating the molecular role of several proteins by constructing rudimentary digital representations of cells.

Biological systems are often described as complex, but as Hiroaki Kitano pointed out, the non-linear interactions of multifunctional components lead to a rather coherent behavior [91]. Networks are useful humanized descriptions of these interactions, which are characterized by nodes (components: genes, transcripts, proteins, metabolites, *etc.*) and edges (modulatory interactions between components: transcription factor binding, RNA interference, post-translational modification, protein-protein interactions, *etc.*; **Figure 3**). After several hundred years of developing a strict scientific ethos based on pattern recognition and linear causative relationships, especially within molecular biology, we now face the challenge that the constructed networks from single and multiple *omics* datasets become incomprehensible for the human mind. In other words, the networks are perceived as complex.

Systems biology has the power to give unprecedented and transformative insight into living systems, and a number of tools help us to overcome the limitations of data integration and interpretation [92]. Known protein-protein interactions can help in identifying the primary interaction partners of a protein of unknown function. The popular web-based tool STRING [93] collects and scores publicly available interaction datasets and uses these to find edges between hits of an *omics* screen. For instance, supposedly a gene encoding protein X is knocked out and a proteomics screen identifies that protein Y is lost, then one conclusion can be that X interacts with and stabilizes Y (see study II, in which SLIRP is lost upon knock down of DmLRPPRC1). While highly useful for low-throughput studies of defined model systems, protein-protein interaction networks are more likely to fail when the precise insults to an interaction network are either not known as in the case of unknown disease-causing variants

or multi factorial with environmental contribution. Moreover, incomplete data and varying annotation levels limit this approach further [93].

The interactions on genome level can be computed from gene essentiality screens, which are almost exclusively performed in cell lines. A gene is essential if its gene product is required for cellular survival, and vice versa when its loss compromises cellular fitness [94]. The human Cancer Dependency Map combines genetic screens, drug sensitivity, and other *omics* datasets from various cell types to predict the vulnerability of cells to the loss of gene products [95]. For instance, the Achilles dataset comprises cell viability scores after CRISPR/Cas9-based screening of 18,000 genes in 800 human cancer cell lines [96]. Correlating the viabilities of various genes with each other, or with further *omics* datasets can give insight into relationships between gene products. For instance, Jain *et al.* used gene essentiality data from the Achilles project and correlated it with metabolite levels in the corresponding cell lines, finding a relation between levels of unsaturated lipids and peroxisomal genes [67].

Grouping proteins with known function into functional categories is called pathway analysis and is another popular approach to interpret large datasets. This is appealing because it reduces the complexity of a dataset and has high explanatory potential [97]. Over-Representation Analysis (ORA) comprises the first generation of pathway analysis tools [98]. A list of interesting elements is chosen on the basis of, for instance, *p* values after differential expression analysis, and the list is then statistically evaluated against all detected elements. Each element is annotated with the pathways it contributes to using libraries such as Gene Ontology (GO) resources [99] or the Kyoto Encyclopaedia of Genes and Genomes (KEGG) [100], and over-represented pathways in the interesting list are returned. The main problem is the arbitrary user-based choice of a cutoff value to define the list of interesting elements. The next generation of functional class scoring pathway analyses [97], for instance Gene Set Enrichment Analysis (GSEA), follows a similar functional annotation strategy. GSEA uses ranked lists that are most often based on changes in expression and that contain every detected element instead of just a selection [101] as in ORA. While this approach is more robust towards outliers, it is bias-sensitive when there is no real biological difference. Furthermore, ORA and GSEA suffer from incomplete pathway annotations or assignment of one element to several pathways, which aggravates the biological interpretation. Efforts are being made to standardize library terms [99], and a third generation of pathway topology-based models is on its way that combines pathway analysis with known interactions [97].

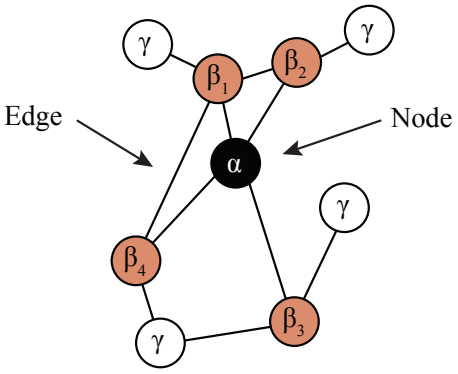


Figure 3: A simple network model. The network is centred on the node α that has four edges to direct direct partners β_i, of which some interact with each other and further γ partners and so forth.

With a growing number of datasets, biologists must pair up with computer scientists to

solve the more and more complex *omics* puzzles. The digital cell is still a dream, but user-friendly machine learning algorithms delivered for instance *via* the TensorFlow interface [102] will contribute greatly in unsupervised network construction to gain biological insight [103]. Importantly, this does not alleviate the need for high-quality molecular follow-up studies by experimental molecular biology .

2.5 Open source data access

Science has become a data intensive enterprise, and it will continue to move into this direction with decreasing costs, more standardized protocols and better accessibility to *omics* core facilities. Moreover, instead of producing new datasets cross-disciplinary studies can greatly profit from re-using and re-purposing already existing data [104]. However, this approach continues to be hampered due to lack of clear best practices in data documentation [105].

The FAIR principles (Findable, Accessible, Interoperable, Reusable) that were formally published in 2016 aim at framing guidelines for the necessary infrastructure and policies [106], and an increasing number of grant agencies and scientific journals require public data deposition [107]. Several platforms including the Gene Expression Omnibus (GEO) of the U.S. National Center for Biotechnology Information (NCBI) and the PRoteomics IDEntifications (PRIDE) Archive hosted by the European Bioinformatics Institute have emerged to support large-scale data deposition and annotation, and to promote the re-usage of datasets. PRIDE Archive stored almost 13,000 datasets at the time of writing predominately from human material [108], and 293 were re-analyzed until September 2018 [109]. Yet, the observation that this thesis provided 3.5% of all fruit fly datasets (5 out of 144 at the time of writing [109]) raises the suspicion that a large body of data is still not made public. With better mining tools, the scientific community will profit from re-analyzing data and it will be exciting to observe how the attitude towards data transparency develops in the near future.

2.6 Nuances of mitochondrial systems biology

Omics data analysis that is focussed on mitochondrial function faces a number of challenges. Firstly, mitochondrial components do not necessarily give insight into a mitochondrial defect when studied in relation to the rest of the cell. Fission and fusion events make the mitochondrial compartment highly dynamic, and the total mitochondrial mass can vary between cells of the same cell type. Thus, a total *omics* dataset might help in observing a generalized loss or accumulation of mitochondrial proteins, but this can reflect an upstream coordination through master regulators such as PGC1α [110] rather than an intrinsic mitochondrial response.

Secondly, the majority of the nuclear transcriptome that encodes mitochondrial proteins does not correlate with protein levels [111]. Consequently, this limits the usability of transcriptomics data in mitochondrial systems biology (**Figure 4**). The reasons for this are unclear. The major part of the approximately 1,150 annotated human mitochondrial proteins

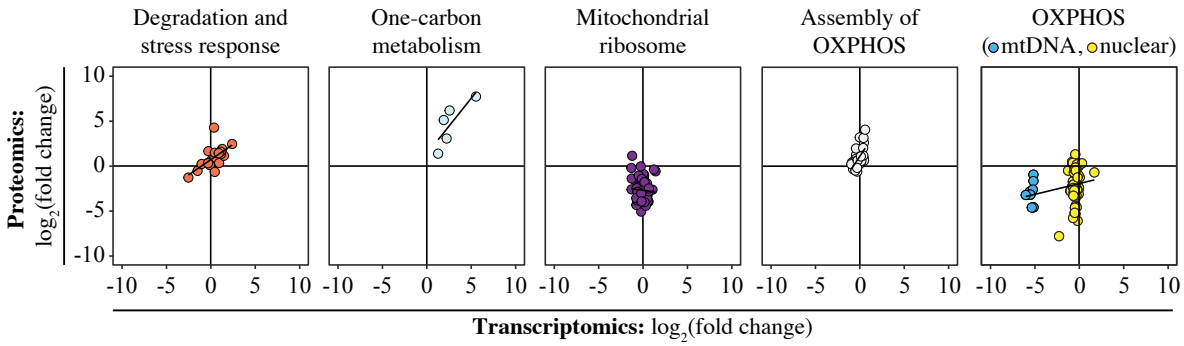


Figure 4: Functional relevance of mitochondrial transcriptomic and proteomic data. Log_2-transformed expression ratios of TFAM knock-out mice against controls, quantified by transcriptomics and proteomics. Data was grouped into functional pathways as indicated. Adapted with permission from Kühl *et al.* [111] under the terms of the Creative Commons Attribution Licence 4.0.

[13] are nuclear encoded, synthesized on cytoplasmic ribosomes and imported into mitochondria *via* a specialized one-way import machinery. There, a considerable fraction assembles into large protein complexes like the mitochondrial ribosome or the OXPHOS system [16]. Thus, the lack of correlation could be a consequence of stabilizing proteins in complexes that prevents their degradation, the absence of a direct translational feedback loop, or the above discussed discrepancy of mitochondrial versus total cell protein mass.

Thirdly, whole cell metabolomics does not retain the subcellular spatial information. Therefore, the organelles are typically enriched by differential centrifugation to quantify mitochondrial matrix metabolite concentrations. This process takes time and metabolites can diffuse through solute carriers along a sharp gradient into the enrichment buffer. One solution has been provided by the rapid antibody-based extraction of mitochondria [112]. However, this requires the expression of a handle protein on outer mitochondrial membrane and the method is currently only available for cultured cells.

Few tools and databases specialise on mitochondria. MitoXplorer uses manually curated lists of mitochondrial proteins including functional annotations. It provides a web-based analysis tool that imports known protein-protein interactions and that subsets user-defined input data into functional categories [113] for manual inspection. MitoCarta was first published in 2008 [114] and has since then been a reference proteome for mitochondrial proteins. The current version 3 [13] assigns hierarchical functional groups and suborganellar localization to mitochondrial proteins. The precise localization of mitochondrial proteins has also been mapped in Saccharomyces cerevisiae [115]. Moreover, the construction of mitochondrial interactive protein networks has been facilitated by recent protein proximity mappings on a global level [116] and with a focus on mitochondrial protein biosynthesis [117].

The comparably small set of proteins and the semi-autonomous characteristics make mitochondria an exciting research object for systems biology. Many challenges on both the experimental and computational level remain, but the integration of more large-scale datasets from genetic models will be useful in the future to construct functional networks of mitochondrial proteins.

3 MITOCHONDRIAL PROTEIN MODIFICATIONS

3.1 General considerations

PTMs are an energy-efficient mechanism to regulate the activity of proteins and diversify the functions of proteomes [118]. With the advent of MS-based proteomics, a large number of proteins was found post-translationally modified. To date, the PhosphoSitePlus inventory version 6.5.9.3 [119] counts 486,000 individual PTMs throughout the cell including mitochondria, and there is increasing evidence for regulatory roles of mitochondrial PTMs [120]. Various types of chemical modifications have been described in mitochondria including phosphorylation, methylation, lipoylation, acetylation, ubiquitination, acylation (succinylation, glutarylation, malonylation), succination, O-GlcNAcylation, and others (**Figure 5**). Some of these are well established while others have been a matter of vivid debate such as acetylation [121] or will still have to face the discussion of their biological significance. In the context of this thesis, only phosphorylation and methylation will be discussed, but several key questions are of general importance in the field and described below.

Is a detected modification real or a technical artefact? Concerns have been raised that sample handling can artificially introduce modification on certain residues. For instance, a larger number of peptides were found methylated specifically after sodium dodecyl sulfate (SDS) gel electrophoresis [122]. Similarly, it has been shown that tryptophan oxidation to kynurenine and N-formyl kyrunenine is a gel electrophoresis-derived artifact [123]. Moreover, phosphorylation is isobaric with sulfation, which is an artifact after silver staining of gels [124]. Gel-free sample preparation for MS counteracts these problems, although Nielsen *et al.* showed that iodoacetamide treatment, which is often used during peptide production, causes artifacts that mimic diglycine signatures of ubiquitination [125]. Thus, the assignment of PTMs to mass spectra needs to be done with caution, and additional evidence using labeling strategies or targeted MS is pivotal for high-confidence mapping of PTMs.

Is a target modified to a biologically relevant extent? Increased sensitivity of mass spectrometers allows the detection of low-abundance proteins and PTMs. An important parameter to assess their biological relevance is the site stoichiometry or occupancy, which is the proportion of a specific site on a given protein that is modified over unmodified [118]. The occupancy value is more informative than fold changes when comparing two biological con-

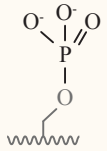
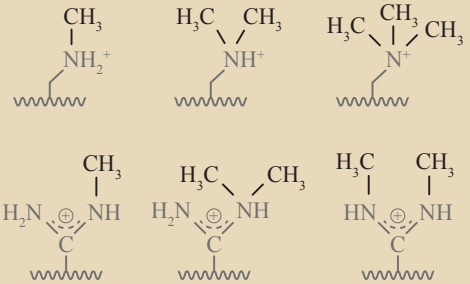
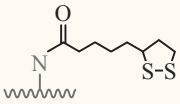
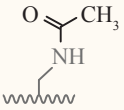
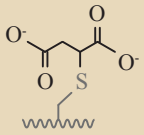
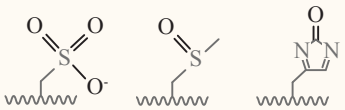
Examples for enzymatic PTMs		
Modification	**Canonical residues**	**Product**
Phosphorylation	Ser, Thr, Tyr	
Methylation	Arg, Lys	
Lipoylation	Lys	
Examples for non-enzymatic PTMs		
Acetylation	Lys	
Succination	Cys	
Oxidation	Cys, Met, Tyr, etc.	

Figure 5: A selection of protein post-translational modifications. Note that several of the modifications can occur on other residues as well. Here, only the most common modified amino acids are listed.

ditions. For instance, a five-fold increase from 1% to 5% would correspond to an increase from 20% to 100% [126]. However, the methodology for absolute quantification of site occupancies has a number of limitations [118] and mere relative ratios between two conditions are easier to obtain. The biological significance of low-abundance occupancy sites is not clear yet [121] and the evaluation currently depends on prior knowledge or biochemical follow-up studies.

Is the modification mediated by enzymes? Nucleophilic or redox-sensitive side chains of amino acids can react with suitable metabolites such as acetyl-CoA without enzymatic catalysis [127]. However, spontaneous reactions do not need to be random, and the specificity of non-enzymatic reactions is an active area of research [127]. On the contrary, phosphorylation, methylation and ubiquitination require kinases, methyltransferases and ubiquitin ligases, respectively, albeit non-enzymatic phosphorylation by inositol pyrophosphates has been reported [128]. From an experimental point, the function of enzymatic PTMs can be studied by both manipulating the target amino acid and by genetic perturbation of the modifying enzyme. The latter approach avoids the problem that exchanging an amino acid to a non-modifiable or constitutively modified analogue only approximates the chemical environment of the actual modification, which is not a definite proof of the biological role of a PTM. For instance, glutamic or aspartic acid are considered mimetics of phosphorylated serine due to their negative charge and small size, while alanine resembles the non-modifiable amino acid [129].

Is the modification reversible? A number of non-enzymatic PTMs are irreversible and signal cellular stress. For instance, succination of cysteine residues is a downstream consequence of fumarate accumulation in the TCA cycle, and it has been shown that mutations in fumarate hydratase lead to elevated succination of proteins required for iron-sulfur-cluster synthesis [130]. Acetylation on lysine, on the contrary, can be removed by NAD^+-dependent SIRT3 in mitochondria [131]. This gave rise to the idea of a direct metabolic feedback loop that omits the need of specific protein relays [121]. Few enzymatic PTMs can be removed through designated proteins, such as phosphatases [132] and deubiquitinases. Demethylases, however, appear to exclusively target histone tails [133], and non-nuclear demethylases have not been identified to date.

Is the target modified inside mitochondria? Few modifying enzymes are targeted to mitochondria such as the methyltransferase NDUFAF7 [134] or the lipoyltransferase LIPT1 [135]. However, most modifications in mitochondria do not have a known corresponding enzyme, which raises the hypothesis that polypeptide chains can be modified prior to mitochondrial import. Indeed, it has been shown that the cytoplasmic enzyme PRMT1 is responsible for multiple asymmetric dimethylated arginine residues on mitochondrial proteins [136]. The extent of pre-import modifications and the underlying molecular mechanisms are still unknown.

Are we done? Several reports suggest that at least 50% of all cellular proteins are phosphorylated at some point [137–139] in mouse, HeLa cell lines and yeast, with similar rates for acetylation [140] and ubiquitination [141]. The rate of false positives as well as false negatives is currently unclear, and Olsen and Mann stated in 2013 that we will not have completed an inventory of PTMs in the near future [42]. The limitations are in part related to the sensitivity of mass spectrometers. In fact, new technical developments have led to ever-increasing sizes of the modified proteome [121]. Additionally, most studies rely on classical peptide preparation approaches using trypsin as an endoproteinase, which renders strongly hydrophobic peptides as in membrane-bound proteins, and short and long tryptic peptides underrepresented [42]. Yet, the biggest challenge remains: To annotate a biological function to the detected sites and modifications.

3.2 Phosphorylation

3.2.1 Biochemical snapshot

Protein phosphorylation was the first post-translational modification to be mapped on glycogen phosphorylase in 1956 by Edwin Krebs and Edmund Fischer [142]. Potentially due to its scientific age and historical role in educating biochemistry students, phosphorylation is well studied and to our knowledge the most abundant PTM. The modification is a covalent addition of a phosphate group to hydroxy groups of side chains of serine (86% of all phosphorylated amino acids), threonine (12%) and tyrosine (2%) [143]. Although protein-protein binding is not energetically more favorable upon phosphorylation [144], the modification increases the interactive capacity of a protein because the dianionic phosphoryl group can participate in hydrogen bonds and salt bridges with other residues such as arginine and lysine [145]. Indeed, modified residues are preferentially located at the binding site of homo- and heterooligomeric complexes [146], and within disordered N- or C-terminal regions [144]. For instance, the aforementioned glycogen phosphorylase is phosphorylated at serine 14, which causes a disorder-to-ordered transition of the N-terminal protein tail, subsequent homotetramerization, the exposition of the active site of the protein, and ultimately the activation of the enzyme to degrade glycogen [147].

3.2.2 Screening and quantification strategies

Phosphorylation is operational in essential cellular signaling pathways, and aberrant stoichiometry or mutations at the phosphorylated site contribute to a wide array of diseases such as cancer [148] and Parkinson's disease [149]. Reassuringly, inhibiting appropriate kinases is a successfully applied therapeutic strategy [150]. The achievable depth of the substoichiometric phosphoproteome is limited by a number of factors that need to be considered during experimental setup. Firstly, although a large fraction of the proteome can be phosphorylated throughout time, only about 1-2% of the entire proteome are believed to be modified at a given

moment [151]. Therefore, if a complex protein sample is subjected to LC-MS/MS, most duty cycles of the mass spectrometer will quantify unmodified peptides and low-intensity modified peptides might not be detected. Secondly, we and others have observed severely reduced signals from phosphorylated over non-phosphorylated synthetic standards due to different ionization properties, although this notion of phosphoproteomics has been challenged [152]. This suggests that phosphorylated peptides are less likely to be detected than their modified counterpeptides.

Enrichment strategies have been developed to enhance the coverage of phosphorylation sites in complex biological samples. Affinity-based enrichments with immobilized metal affinity chromatography (IMAC) or metal oxide affinity chromatography (MOAC) are now widely used, and several pre-enrichment fractionation techniques were optimized including strong-cation exchange (SCX) and high pH reversed phase liquid chromatography (HpH RPLC) [153]. However, since the publication of EasyPhos [154], which is a one-step protocol for enrichment of phosphopeptides with titanium dioxide (TiO_2) as a MOAC material without prior fractionation, TiO_2-based enrichment has become the most widely used method to obtain complementary phosphoproteomes.

A fundamental problem of enrichment strategies in general is the loss of information of site stoichiometry, and the abundance of the corresponding protein has to be quantified in a separate experiment [118]. In a label-free approach, the absolute occupancy of a site can be compared between two biological conditions under the assumption that the abundance of the unmodified peptide changes reciprocally to the modified counterpeptide [155]. However, the requirement for accurate quantification is that the protein levels are unchanged in both conditions, which is in practice not always achievable in a biological context. Furthermore, the site occupancy must range between 10% and 90% so that the less intense peptides can still be reliably quantified [118, 155]. Quantification of site occupancies from SILAC labeled peptides is more robust and takes protein level changes into account [138]. However, it is only reliable when the modification stoichiometry is in fact different between two samples [118]. For this, a user-friendly add-on has been implemented in the classical proteomics data analysis software MaxQuant [156].

3.2.3 Mitochondrial protein phosphorylation

Seminal work on mitochondrial protein phosphorylation has been performed in Saccharomyces cerevisiae. A meta-analysis by Frankovsky *et al.* listed 972 (67.6%) phosphorylated mitochondria-associated proteins and a total of 9,402 phosphorylated sites [132]. This converts to an average of 6.76 sites per protein, which is either slightly more (6.05 in [132]) or considerably less (9.84 in the Fungi Phosphorylation Database [157]) than the total cell site-to-protein ratio. Most sites do not have an annotated function. Growing yeast on either glucose or lactat/galactose medium identified 670 sites in mitochondria and 90 were found differentially phosphorylated [158], suggesting these sites associated with OXPHOS activity. Recently, it has been shown

that mitophagic removal of mitochondrial proteins is selective and depends on the phosphorylation status of the target protein [159]. However, the specificity of this mechanism is unclear.

Prominently, the mitochondrial pyruvate dehydrogenase complex (PDH) is regulated by phosphorylation through the activity of four PDH kinases that inactivate the complex, and two PDH phosphatases that restore enzyme activity [160]. Several kinases and phosphatases with multiple cellular localizations have been found in mitochondria, including the receptor tyrosine kinases epidermal growth factor receptor (EGFR) and ErbB2 (reviewed in [161]). An interesting body of work has been accumulated on the serine/threonine kinase cAMP-dependent protein kinase (PKA). PKA associates with the outer mitochondrial membrane by binding to A-kinase anchoring proteins (AKAP), which can sequester multiple proteins and thus serve as signaling hubs [162]. For instance, by binding to AKAP121, PKA phosphorylates the pro-apoptotic protein BAD and counteracts apoptosis [163]. Moreover, PKA has been localized to the inner mitochondrial membrane and matrix in several reports using immunoelectron microscopy and subcellular fractionation [164, 165]. Silencing of AKAP121 decreases intra-mitochondrial PKA levels [166], however, the mechanism of mitochondrial import remains unknown as the protein does not contain a mitochondrial targeting sequence [162]. Evidence for mitochondrial PKA function stems from PKA-dependent phosphorylation of the complex I subunit NDUFS2 [166] and ATPase Inhibitory Factor 1 (IF1) [167]. Although an intra-mitochondrial cAMP source for PKA has been proposed, it remains unclear whether the vast number of phosphorylations happen inside the mitochondrial membrane boundaries, are added upon import as has been suggested [132], or are non-enzymatic. Due to this conundrum, large-scale PTM screening efforts profit from rigorous biochemical follow-up in order to prove that a particular site is real and of biological importance.

In *Dm*, phosphorylation within the mitochondrial compartment has not been widely studied. A total fly phosphoproteome study found about 500 phosphoproteins and 1,600 phosphopeptides, and changes of PTMs status upon cold exposure associated in particular with microtubular networks [79]. In mammals, mitochondrial proteins related to OXPHOS and metabolism were found differentially phosphorylated in obese mice [168, 169]. In contrast to yeast, the modification was underrepresented in mice [169] and human cell lines [170]. Yet, it is unclear if an intra-mitochondrial feedback loop *via* protein phosphorylation exists and how monogenic mitochondrial disease shapes the modification landscape.

3.3 Methylation

3.3.1 Biochemical snapshot

The protein modification methylation is the addition of one or more CH_3 groups to the amine group of an amino acid side chain, predominantly arginine and lysine. However, non-canonical methylation of histidine, glutamate, glutamine, aspartate, asparagine, cysteine, and protein N- and C-termini were found as well [171]. The primary methyl group donor in cells is *S*-

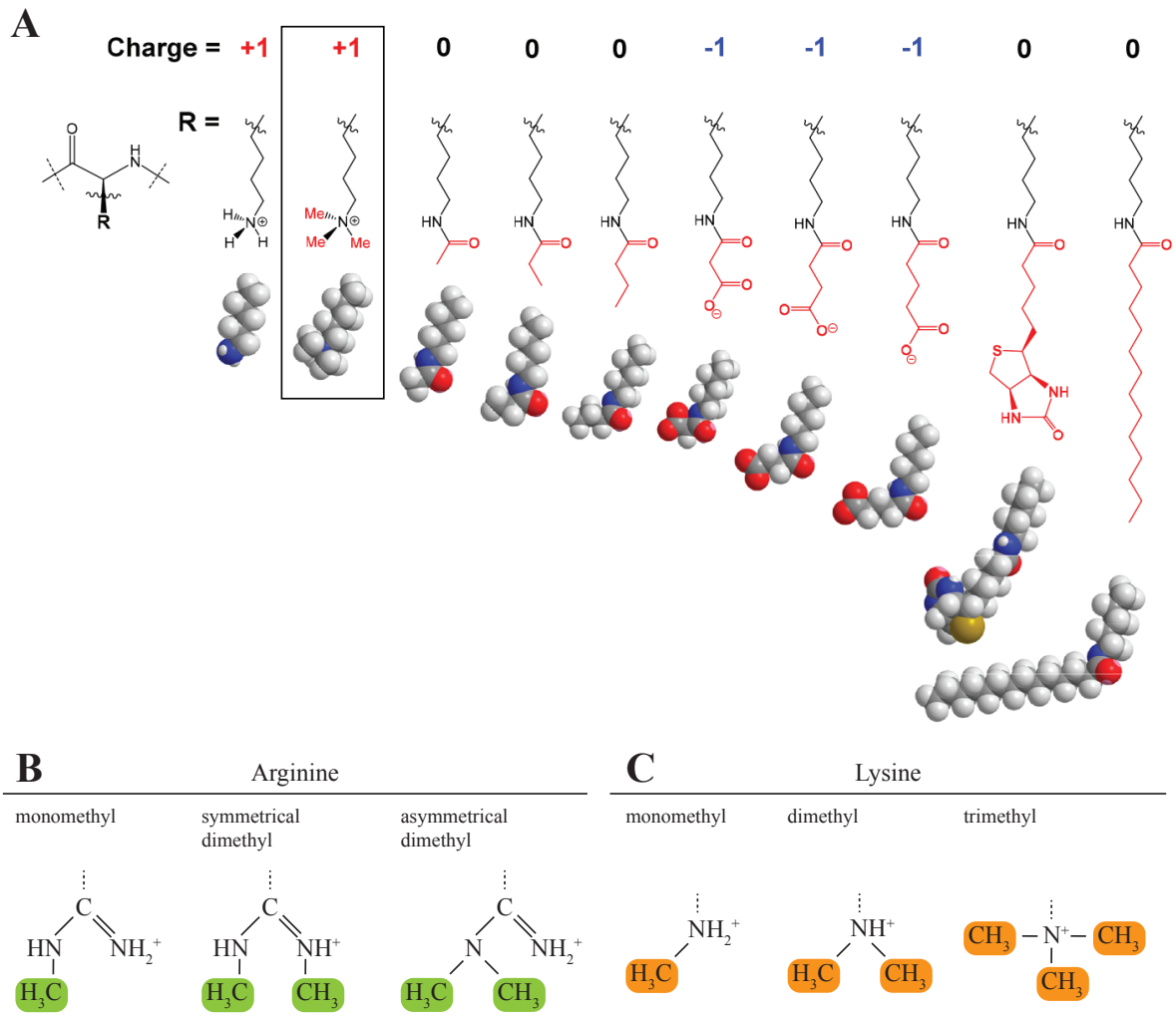


Figure 6: Methylation is a multifaceted and small modification. (A) The formal charge, structure and size of lysine trimethylation (marked with a rectangle) in comparison with other lysine modifications. From left to right: the unmodified lysine residue, trimethylation, acetylation, propionylation, malonylation, butyrylation, succinylation, glutarylation, biotinylation and myristoylation. Reprinted with permission from Minkui Luo [173]. Copyright 2018, American Chemical Society. **(B)** Types of arginine and **(C)** lysine methyl-modifications.

adenosylmethionine (SAM), which is an abundant metabolite second only to ATP [172].

Compared to other PTMs, methylation is a small chemical group that retains the charge of the modified lysine and arginine (**Figure 6A**) [174]. It enlarges the hydrophobic shell around the terminal atoms of the side chain, and thus reduces the amphiphilic character of the unmodified amino acid [173]. On lysine, the ε-N atom can become mono-, di- and trimethylated, which increases the mass of the residue by 14 Da, 28 Da, or 42 Da, respectively. Arginine can be found mono- and dimethylated, and dimethylation can occur both symmetric and asymmetric (**Figure 6B, 6C**). Every consecutive methyl group reduces the hydrogen bonding capacity of the amine group by one. Arginine is a particularly interactive residue that often protrudes from protein surfaces, and its five hydrogen bonding donors enable protein-protein or protein-nucleotide interactions [175]. The naked arginine can interact with two moieties simultaneously through two amine groups, which allows the symmetric and asymmetric dimethyl configurations. Albeit the positive charge is independent of the methylation

state, methylation induces changes in charge distribution as shown by nuclear magnetic resonance (NMR) spectroscopy [174]. However, the precise role of individual modification states is not known and will require additional structural studies.

3.3.2 Screening strategies

Protein methylation is a subtle chemical event that elevates the hydrophobicity and spatial requirements of an amino acid side chain. Yet, as a methyl group adds just 14 Da to a protein, the methylated and unmethylated states are not distinguishable by size selection on Western blotting. Discovery of novel methylation marks with traditional wet-lab techniques is thus a time-consuming process. LC-MS/MS is a powerful technique to detect hundreds to thousands of PTMs at once. In principle, it is required to screen for a 14 Da mass shift from an expected peptide mass to detect a methyl-PTMs. This approach is, however, limited by several obstacles [42]. Firstly, a mass spectrometer can only fully sequence a small number of peptides per time unit, which decreases the likelihood that a medium or low-abundance methylated peptide will be analyzed. Secondly, the methylation occupancy, which is the percentage of a certain residue being modified, is usually low in the case of regulatory PTMs. Thirdly, to confidently localize a methyl group on a certain amino acid, the fragmentation pattern of that peptide has to contain sufficient information, which is not always the case.

Similarly to phosphorylation discussed above, strategies were developed for the enrichment of methylated proteins or peptides over a non-methylated background. Immunoaffinity enrichment involves antibodies that are coupled to agarose or magnetic beads. A set of antibodies has been developed against mono-, symmetrical and asymmetrical dimethylation on arginine and a pan-methyl antibody for lysine [176]. Together with a variety of fractionation methods [177], the mono-methyl arginine antibodies led to the discovery of a large number of methylation sites [178–180]. The false positive rate among those is unknown and likely high [181]. Applying these strategies to lysine methylation has been difficult and tagged peptide domains have been developed that bind mono- and dimethyl lysine without sequence specificity [182]. However, these domains recognize only undigested proteins and not peptides, and a profound background of non-methylated peptides negatively affects discovery depth.

In order to decrease the number of false positive hits and to provide a further *in silico* search strategy, heavy methyl-SILAC was developed. By feeding cells with labeled methionine [183], a 4 Da ”heavy” label is added to every SAM-derived methyl group. The finding of a methyl group depends thus not only on a 14 Da shift of a light methyl group to an unmodified peptide, but also an additional 4 Da heavy to light and 18 Da modified to unmodified shift due to the heavy label. The dedicated software MethylQuant has been developed to screen for mass shifts in complex methyl-SILAC data, which allows a consistent calculation of FDR for methylated peptides across studies [184]. Interestingly, the labeled [$^{13}C^{2}H_{3}$]-SAM is slightly more reactive than its light counterpart [185], which can indicate a minor bias in methyl-SILAC experiments.

In addition to isotopic SAM, unmodified lysine residues can be blocked with formaldehyde or propionic anhydride that introduce distinct masses and furthermore efficiently block trypsin digest at the modified residue. The specific cleavage pattern can thus be used to distinguish the modification state at a certain lysine residue [173]. Furthermore, alkyne-containing SAM analogues can be applied to introduce clickable modifications instead of methyl groups [186]. The modified proteins are then clicked to biotin and enriched with streptavidin-coated beads before digestion and LC-MS/MS [187]. Notably in the mitochondrial field, this approach has been successfully used to detect methyl-histidine on the complex I subunit NDUFB3 [188]. The SAM analogues are powerful in detecting truly methylated proteins in cell culture, yet at the expense of a high false negative rate. It is currently unclear which structures of SAM-binding pockets allow the binding of analogues, and which potential targets will be omitted from labelling in a cell-wide screen [189]. Furthermore, the alkyne instead of a methyl group will alter the local chemistry at the modified residue, and sites that are of structural importance might not be covered well. However, promising solutions are on their way, and SAM analogues are likely to complement existing techniques in the future [189].

4 ONE-CARBON METABOLISM

4.1 Flux of one-carbon units

A unifying early metabolic signature of a mitochondrial insult is an increase of mitochondrial 1C enzymes [24, 111, 190]. 1C metabolism comprises a set of reactions in both cytoplasm and mitochondria to transfer 1C units and modify their oxidation state [191]. The cofactors that carry 1C units within the main pathway are folates, a subsuming term for derivatives of folic acid. Folates cannot be synthesized *de novo* in human cells and must be taken in with a diet. Although natural diets contain almost exclusively 5-methyl-tetrahydrofolate (THF) [192], the precursor folic acid is used as the artificial dietary supplement, which has to be reduced in several NADPH-dependent reactions to be biologically active. These reactions have been shown to be of minor physiological importance in humans and excessive intake of the vitamin retains a large pool of unmetabolized folic acid in the blood that can cause adverse effects [193]. However, folic acid profoundly decreases neural-tube defects in newborns [194] and is indicated as a peri-conceptional dietary supplement [193].

Folates can carry one 1C unit at a time in different oxidation states, most importantly 10-formyl-tetrahydrofolate (THF), 5,10-methylene-THF and 5-methyl-THF (**Figure 7**). Each compound allows a different chemistry and is co-substrate to specific 1C-transfer reactions. Cytoplasmic synthesis of purines is the largest sink of 1C units carried on 10-formyl-THF in proliferating cells [191]. The required enzymes are localized in purinosomes in proximity to mitochondria. These structures have been shown to respond to mitochondrial stress, suggesting a direct link between purine synthesis and mitochondrial function [195]. A minor proportion of 10-formyl-THF [196] is further needed in mitochondria to formylate the methionine tRNA recognition loop [197]. 5,10-methylene-THF is a co-substrate for serine synthesis in both cytosol and mitochondria, whereby the methyl group is transferred onto glycine. Furthermore, 5,10-methylene-THF is required for cytosolic thymidylate biosynthesis. This links back to mtDNA-associated diseases as imbalanced nucleotide pools lead to mtDNA instability [198]. 5,10-methylene-THF can be further reduced to 5-methyl-THF, which provides the methyl group in the re-methylation of homocysteine to methionine, thus fueling the methionine cycle (see section 4.2).

The cytosolic and mitochondrial route of 1C metabolism are linked *via* THF and formate; other folates that carry 1C units cannot pass the mitochondrial membrane and have to be synthesized in both compartments [199]. The primary 1C donor is serine [200], but various other

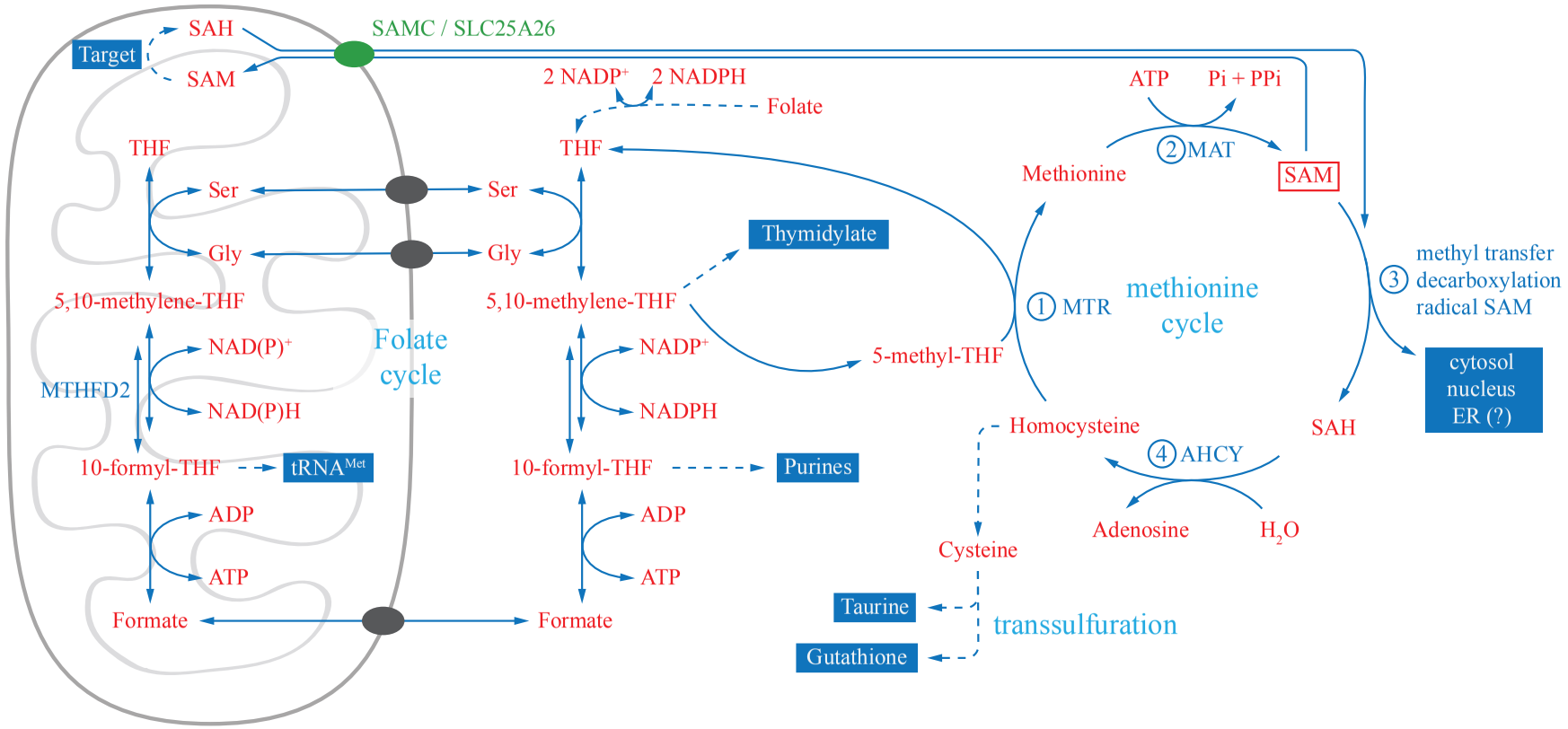


Figure 7: Integration of one-carbon units. The one-carbon cycle comprises a set of reversible reactions in both mitochondria (left) and the cytoplasm (right). In the **folate cycle**, one-carbon units are transferred and their oxidation state is changed. The **methionine cycle** consist of four reactions, in which intially a one-carbon (1C) unit is transferred onto homocysteine, yielding methionine and further SAM. A diverse family of methyltransferases uses SAM in various compartments of the cell, but also SAM decarboxylating and radical SAM reactions are known. Homocysteine can be re-metyhlated to methionine, or is further metabolised in the **transulfuration pathway** to cysteine or the redox-relevant glutathione. MTR, methionine synthase, 5-methyltetrahydrofolate-homocysteine methyltransferase (reaction 1); MAT, methionine adenosyltransferase (reaction 2); AHCY, adenosylhomocysteinase (reaction 4); MTHFD2, methylene-THF dehydrogenase 2.

sources are known including betaine, sarcosine and tryptophan [201]. The two compartmentalized pathways thus form a cycle, known as the folate-mediated 1C cycle. All reactions in this cycle are reversible and their direction depends on the metabolic state of the cell. While serine might be used up in one compartment, it is synthesized in the other one [191]. It has been shown in cancer cells that the mitochondrial route is the predominant one [202] and the transcript for the mitochondrial 1C cycle enzyme MTHFD2 was the most up-regulated among a broad tumor tissue panel [203]. A physiological driving force for a predominant mitochondrial route from THF to formate is potentially the high NAD^+/NADH ratio in mitochondria, thus favoring the oxidation of folates and generation of mitochondrial NADH. It has been proposed that this contributes to redox homeostasis [202]. In agreement with this, the cytosolic route is mostly consuming NADPH and sustains the backflow of THF into mitochondria [204].

4.2 Control of *S*-adenosylmethionine synthesis

5-methyl-THF is required for the re-methylation of homocysteine to methionine by methionine synthase, one of the two known vitamin B_{12}-dependent enzymes. A salvage pathway exists in liver and kidney, whereby betaine-homocysteine methyltransferase (BHMT) abstracts a methyl group from betaine to re-generate methionine [205]. Clearance of homocysteine is essential for human health, and increased levels of plasma homocysteine correlate with cardiovascular diseases [206] and stroke [207]. It has further been found that homocysteine activates NMDA receptors on neurons and immune cells, which stimulate generation of reactive oxygen species (ROS) and cell death [208]. Additionally, the enzyme adenosylhomocysteinase (AHCY) adversely favors the condensation of homocysteine with adenosine to yield *S*-adenosylhomocysteine (SAH), which is a strong inhibitor of SAM-dependent methylation reactions. In support for the arising pathological condition, increased homocysteine levels were linked to a hypo-methylated cellular state [205].

The re-methylation of homocysteine to methionine initiates the methionine cycle, which is formed by four reactions as shown in **Figure 7** [209]. Methionine synthase (reaction 1) links the 1C cycle with the methionine cycle, although the rate-limiting step of the methionine cycle is the subsequent adenylation of methionine (reaction 2). The mammalian genome contains three methionine adenosyltransferase (MAT) genes that can catalyse this reaction, which assemble to three known homo- and dimeric isoforms. MAT I, a tetramer of α_1 subunits, and MAT III, an α_1-dimer, are found primarily in liver tissue [210]. They have a medium ($K_m \approx$ 100 µM) and low ($\approx$ 1 mM) binding affinity to methionine, respectively [211]. Extrahepatic tissue, in contrast, hosts primarily MAT II, which is putatively a heterotrimer of two catalytic α_2 and one regulatory β subunit that enable the highest binding affinity to MAT substrates of all three isoforms ($\approx$ 30µM) [212]. The presence of two isoforms with different K_m in liver points towards a metabolic regulation of MAT gene expression. In fact, a shift in MAT I to

MAT III ratio has been observed in patients with liver cirrhosis [213]. Furthermore, it has been found that the regulatory β subunit of MAT II can bind to $NADP^+$ and thus increases the affinity of the catalytic α_2 subunits [214]. In conjunction with the observation that MAT I/MAT III are inhibited by oxidized glutathione [215], this has sparked the hypothesis that SAM synthesis is sensitive to the redox state of a cell [210].

SAM is, after ATP, the second most common enzymatic cofactor in cells [191]. The concentration of SAM limits the activity of methyltransferases, which stands in contrast to the highly abundant ATP that saturates kinase sites under physiological conditions [216]. This highlights the importance of a tightly regulated MAT reaction. SAM is the primary methyl group donor in cells for about 200 confirmed or predicted methyltransferases [217, 218], which corresponds to 1-2% of the total cellular proteome [171]. Methylation is the addition of a CH_3 group to a target, which includes a multitude of substrates from DNA, RNA, protein and metabolites. Although sometimes considered the universal methyl group donor, some few methylation reactions depend on betaine, choline and folates [219]. Most SAM-dependent reactions (reaction 3) involve the transfer of methyl groups by methyltransferases, but several other reaction mechanisms are known [172]. The methionine moiety of SAM can be decarboxylated by SAM decarboxylase and contribute a propylamine group to polyamine synthesis. The function of polyamines remains elusive, but they have been associated with stress conditions and ageing [220]. SAM can also be reductively cleaved into methionine and an adenosyl radical. The radical is used by a large family of radical SAM enzymes to perform chemically challenging reactions usually by abstraction of target hydrogens [221]. A prominent radical SAM reaction is required during lipoic acid synthesis, which is a prosthetic electron transferring group on PDH, α-ketoglutarate dehydrogenase (KGDH) and others (see 4.4). Radical SAM enzymes are an active field of research with a focus on bacteria; however, I will mostly discuss methylation reactions below.

4.3 Compartmentalization of methylation reactions

Historically, the interest in methylation took off after the discovery that nuclear DNA methylation modifies gene expression [222, 223], which started the field of epigenetics. It is now well established that cytosine can be methylated by DNA methyltransferases (DNMTs) on the 5′-carbon atom yielding 5-methylcytosine (5mC). Although 5mC does not affect Watson-Crick pairing with guanine, the hypermethylation of repetitive CpN elements mostly induces DNA compaction and silencing of gene regions [224]. A layer of complexity is added on top of DNA methylation as tails of histones, the DNA-compacting multimeric proteins, contain numerous PTMs. This includes dynamic methylation on arginine and lysine of various kinds, but also acetylation, phosphorylation, sumoylation, and others [133]. Epigenetic alterations are one of the hallmarks of cancer cells [225], and several reports have proposed a link between cellular SAM levels and cancer invasiveness both *in vitro* and *in vivo* [226–228].

Although much focus has been initially put on the (epi-)genetic side of methylation, an increasing number of marks have been detected throughout the cell on targets with diverse functions. In the cytosol, mRNAs have been found reversibly methylated on adenine and a regulatory role on expression control has been proposed [229]. Furthermore, N^6,2′-O-dimethyladenosine in the 5′-mRNA cap has been described to control transcript stability [230]. Cytosolic ribosomes are heavily methylated both on protein and rRNA components. Ribosomal proteins are methylated mostly on lysine and arginine residues that are both exposed on the ribosomal surface and hidden in the multimer structure [171]. rRNA methylation is found across all domains of life and at least 13 different types of SAM-dependent post-transcriptional modifications on all nucleosides are known with a multitude of mapped methyltransferases [231]. Recent advances in MS techniques led to the detection of 8,030 arginine methylation sites on 3,300 human proteins [180]. These were enriched on components of the protein biosynthesis machinery, but also proteins involved in cell cycle and insulin signaling. Interestingly, these sites have a higher mutation rate compared to unmodified arginine residues across all eukaryotes, but are highly conserved among mammals [180]. The largest fraction of SAM is required during phosphatidylcholine (PC) synthesis at the endoplasmic reticulum [232]. PC is the major lipid component in biological membranes. Deletion of phosphatidylethanolamine methyltransferase, which is the SAM-dependent enzyme in PC synthesis, leads to a 50%-decrease in SAH-plasma levels in mice [233], which highlights the substantial flux of SAM towards synthesis of PC.

Although the primary site of SAM synthesis is in the cytosol, a mitochondrial MATα_1 [234] and nuclear-targeted MAT I/III have been described [235] that might meet the elevated demands of SAM in proliferating cells. Consequently, SAM has to be imported into most organelles that contain SAM-dependent methyltransferases. It has been proposed that the Golgi membrane of *Caenorhabditis elegans* contains a SAM transporter termed SAMT-1. The nematode Golgi apparatus requires SAM to O-methylate fructose and mannose-residues, which are extracellularly exposed on membrane proteins [236]. However, further biochemical evidence for the SAM-transporting activities of SAMT-1 are still missing. The only conclusive SAM transporter was identified in mitochondria [237], which harbor 30% of cellular SAM [238]. SAM carrier, *S*-adenosylmethionine carrier protein (SAMC) or SLC25A26 is a member of the solute carrier family 25, a group of proteins containing 53 nuclear encoded mitochondrial transporters in the inner membrane that shuttle a variety of substrates [239]. SLC25A26 is an antiporter for both SAM and SAH [237], which suggests a mechanism how to resolve the SAH-inhibitory effects on methyltransferases when SAH or homocysteine cannot be broken down close-by. Three patients with mutations in *SLC25A26* are known. Depending on the mutation, symptoms ranged from death shortly after birth, acute episodes of cardiopulmonary failure to progressive muscle weakness [240]. The patient with the most severe symptoms carried a homozygous splice site mutation and likely did not express a functional SAM transporter. Patient fibroblasts showed decreased *de novo* translation and

a reduction of steady-state levels of OXPHOS subunits. Concentrations of SAM-dependent metabolites Q-10 and lipoic acid were depressed, which altogether can account for reduced oxygen consumption in patient cells.

4.4 SAM-dependent modification of mitochondrial components

Known products of SAM-dependent reactions are shown in table **1**. The text highlights conceptually relevant targets.

4.4.1 Proteins

Among the 3,030 proteins that were identified with one or more arginine methylation sites [180], 373 overlapped with Mitocarta [264]. This corresponds to about 25% of all mitochondrial proteins and stands in stark contrast to only one biochemically validated arginine methylation site in the organelle. Five more methylated mitochondrial proteins have been studied, of which four are methylated on lysine and one on glutamine (**Table 1**). The complex I subunit NDUFS2, electron transferring flavoprotein (ETF) β (ETF β), and the complex V subunit ATP synthase subunit c are part of the OXPHOS system. Citrate synthase catalyses the condensation of acetyl-CoA with oxaloacetate, which is a key regulated reaction of the TCA cycle. Mitochondrial translation release factor 1L contributes to translation termination at the mitochondrial ribosome. Finally, the ADP/ATP translocase (ANT) is a major route for ATP into the cytosol in exchange against ADP. All methyl marks are of regulatory or structural importance as shown by knock-out of the respective methyltransferase in human cell lines [134, 242–244, 246], although the directionality of the effect in the case of ANT and ETF β are unclear [241, 244]. In all five cases, the corresponding methyltransferases localizes to mitochondria. For all other proposed methylated proteins in mitochondria it remains unclear if a polypeptide chain can be methylated in the cytoplasm before being imported.

4.4.2 RNA

Mitochondrial tRNAs are heavily methylated both in stem loops and the anticodon [247]. It is generally believed that tRNA-base modifications alter the bonding capabilities and contribute to the stable folding of tRNAs, codon-anticodon recognition and the wobble-characteristics of certain bases [265]. In contrast to tRNAs, the two mitochondrial ribosomal RNAs are less abundantly modified with five methylation sites on 12S and four on 16S [266]. Knock-out of the 12S methyltransferases NSUN4 or TFB1M was embryonic lethal in mice [252, 254] and it has been shown that almost all methylation marks are relevant for either ribosomal assembly or function. Among the eleven mitochondrial mRNAs, 20% of ND5 transcripts were found with a specific 1-methyladenosine site that disrupts base-pairing [259]. However, the functional importance of this observations is not yet clear.

Group	Target	Type & Site	Function	Transferase	Refs.
Protein	Citrate synthase (CS)	tm Lys-395	Enzyme inhibition	METTL12	[241]
	Electron transfer flavoprotein (ETF) β	tm Lys-199+202	Electron transfer modulation	METTL20	[242, 243]
	Complex I subunit NDUFS2	sdm Arg-118	Complex I assembly	NDUFAF7	[134]
	ADP/ATP translocase (ANT)	tm Lys-52	unclear	FAM173A/ANT-KMT	[244]
	ATP synthase subunit c (ATPSc)	tm Lys-43	Complex V assembly	FAM173B/ATPSc-KMT	[245]
	Mt translation release factor 1L (MTRF1L)	N5-mGlu-252	Translation termination	HEMK1	[246]
RNA	mt-tRNA	*diverse*	tRNA folding, codon recognition	*diverse*	[247]
	12S rRNA	m^5U429	unclear, dual function on rRNA and & tRNAs	TRMT2B	[248, 249]
	12S rRNA	m^4C839	Ribosome assembly	METTL15	[250, 251]
	12S rRNA	m^5C841	Ribosome assembly	NSUN4	[252]
	12S rRNA	$dm^6A936+A937$	Ribosome function	TFB1M	[253, 254]
	16S rRNA	m^1A947	Ribosome function	TRMT61B	[255]
	16S rRNA	mG1145	Ribosome, unclear function	MRM1	[256, 257]
	16S rRNA	mU1364	Ribosome assembly	MRM2	[256–258]
	16S rRNA	mG1370	Ribosome assembly	MRM3	[256–258]
	ND5 mRNA	m^1A1374	Translational repression?	TRMT10C	[259]
Metabolites	Lipoic acid	PDH, KGDH, BCKDH, H-protein in GCS	Electron transfer	LIAS	[260, 261]
	Ubiquinone precursors	O-methylation	OXPHOS electron transfer	COQ3	[262]
	Ubiquinone precursor Q	C-methylation	OXPHOS electron transfer	COQ5	[263]

Table 1: Biochemically confirmed SAM-targets in mitochondria with mapped methyltransferase. The "Type & Site" column describes the position of a respective methylation mark within the human protein or human RNA. *m*, methyl; *dm*, dimethyl; *sdm*, symmetrical dimethyl; *tm*, trimethyl; PDH, pyruvate dehydrogenase; KGDH, α-ketoglutarate dehydrogenase; BCKDH, branched-chain ketoacid dehydrogenase; GCS, glycine cleavage system.

4.4.3 Metabolites

Coenzyme Q requires three mitochondrial methylation reactions on the quinone ring for its synthesis [267]. Morover, the E2 subunits of four mitochondrial protein complexes use lipoic acid as a prosthetic group, whose synthesis is SAM dependent: PDH, KGDH, the glycine cleavage system (GCS), and branched-chain ketoacid dehydrogenase (BCKDH). Lipoyl synthase uses a radical SAM mechanism and abstraction of sulfur atoms from iron-sulfur clusters, which then form part of the oxidized lipoic acid ring [268]. In absence of mitochondrial *S*-adenosylmethionine (mitoSAM) in a patient with non-functional SLC25A26, both lipoic acid and Q levels were decreased, although not completely abolished [240]. The multienzyme complexes PDH and KGDH catalyse the NADH-gaining decarboxylation of pyruvate to acetyl-CoA, and α -ketoglutarate to succinyl-CoA, respectively. The H-protein in the GCS is lipoylated and contributes to the catabolism of glycine forming ammonia, carbon dioxide and a one-carbon group that is transferred onto THF [269]. Knock-out of the H-protein in mice was embryonic lethal [270], which could not be rescued with maternal folate feeding. Therefore, the authors suggested that the protein could have additional functions besides its role in the GCS. Lastly, lipoic acid is a prosthetic cofactor on the BCKDH E2 subunit, which catalyzes decarboxylation of downstream catabolic products of valine, leucine, and isoleucine [271]. Defects in the *BCKDHB* gene causes the severe case of maple syrup urine disease type II [272] characterised by encephalopathy and progressive neurodegeneration.

4.4.4 mtDNA

Several studies report the presence of 5mC on mitochondrial DNA [273]. Methylation of DNA without regulating transcriptional activity has been detected in a number of invertebrates [274]. However, there is evidence that mtDNA methylation correlates with certain cellular states. An altered methylation pattern in patients with Down syndrome has been reported [275] and exposure to environmental toxins was shown to alter mtDNA methylation [276]. According to recent data, over-expression of SLC25A26 led to an increased amount of 5mC on mtDNA in a cancer cell line [277]. The catalysing enzyme is under debate. One percent of DNA methyltransferase 1 (DNMT1) isoform 1 has been found to localize to mitochondria [278], but recent studies questioned this finding and suggest DNMT1 isoform 3 instead [279]. The effect size has been small in all reported studies with a maximum of 5% of all cytosines being methylated. The presence of mtDNA methylation has been disputed in several reports. Olova *et al.* showed that a standard library preparation including bisulfite conversion for whole genome methylation screens overestimates the amount of methylated sites [280]. Moreover, an MS-based method failed to detect 5mC on mtDNA [281]. Notably, the methylation signal disappears when scaling up the read depth of sequencing and after decoiling of mtDNA by restriction enzyme treatment [282], rather suggesting a technical artifact and complete absence of mtDNA methylation in humans.

5 MITOCHONDRIAL FUNCTION DURING DEVELOPMENT

Mitochondria are not just passive receivers of cellular signals but are major players development and homoeostasis. Besides their central role in metabolism, mitochondria contribute to calcium signaling, regulate stress responses and apoptosis, synthesise iron-sulfur clusters, and produce the majority of cellular ROS that serve as a pivotal signaling molecule at low doses to impact kinases, growth factors and transcription factors [283, 284]. During embryogenesis, a totipotent oocyte divides and initiates differentiation programs that result in a highly complex multicellular organism in which every niche has characteristic metabolic conditions. These are determined by, for instance, the niche's distance to blood, its oxygen partial pressure and circulatory distance to lungs, and cellular barrier functions between the niche and the primary nutrient supply, which alter metabolite compositions.

Mitochondria must rapidly adapt to meet the constantly changing metabolic demands during development. It is well established that experimental inhibition of OXPHOS prevents stem cell differentiation [285, 286]. For instance, oxidative metabolism is required for lineage commitment during cardiac differentiation of stem cells [287]. Moreover, the above-discussed mitochondrial route of the one-carbon cycle, and the indirect production of acetyl-CoA through TCA-cycle derived citrate, are pivotal for establishing and maintaining the nuclear epigenome in rapidly dividing cells [283].

In addition, a number of findings strengthen the pivotal role of mitochondria during neuronal development. The mammalian brain consists of various cell types that develop during embryonic and postnatal neuro- and gliogenesis [288]. The axonal outgrowth depends on balanced mitochondrial fusion and fission, and loss of the fusion-promoting GTPase mitofusin 2 in dopaminergic midbrain neurons caused a loss of striatal projections, depletion of mitochondrial ubiquinone and severe locomotor dysfunction in mice [289, 290]. Moreover, the young brain relies predominately on glycolysis during synapse formation despite abundant oxygen, termed aerobic glycolysis [291, 292]. Differentiation of single cells in the mouse neocortex induces an early transcriptional program that is characteristic for mitochondrial involvement [293]. Furthermore, an up-regulation of transcripts encoding OXPHOS subunits was observed in adult hippocampal neurogenesis [294]. Specifically, complex V was consistently up-regulated during early lineage progression [295], and inhibition of complex I or complex V abolished cell proliferation *in vitro*. The flux through glycolysis in stem cells

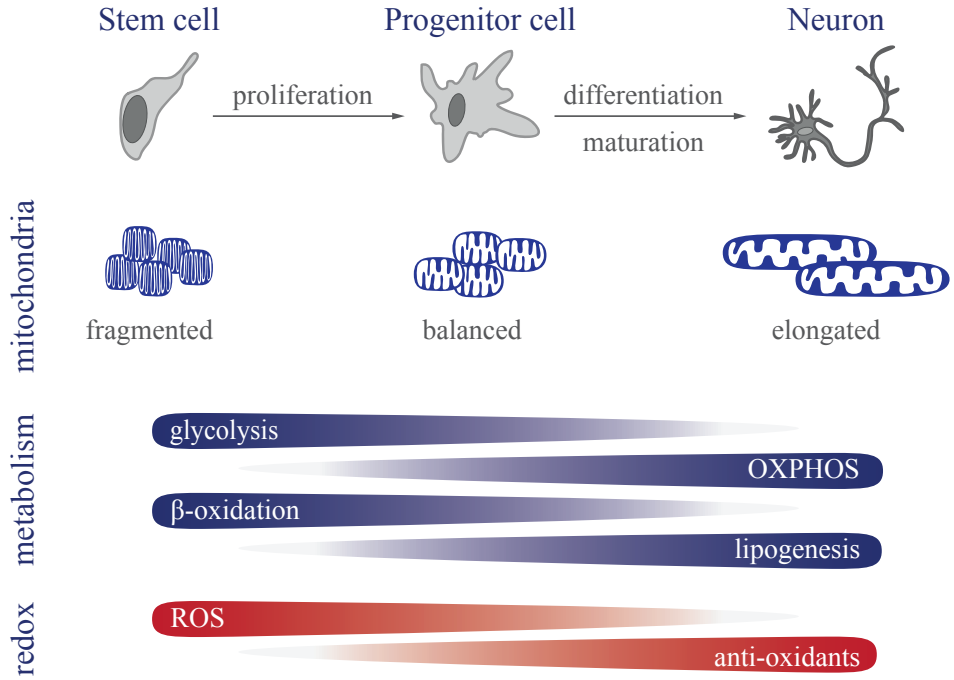


Figure 8: Metabolic changes during neuronal development. Adapted from Maffezzini *et al.* [291] under the terms of the Creative Commons Attribution Licence 4.0 (CC BY).

was found to be regulated at hexokinase 2 and lactate dehydrogenase A (LDHA), and loss of both enzymes characterize the switch to OXPHOS in neuronal progenitor cells *in vitro* [296] (**Figure 8**). Interestingly, the brain regions harboring neuronal stem cell populations during adulthood remain glycolytic [297], suggesting that glycolysis is relevant for maintenance of the stem cell niche in both the developing and adult brain.

Impaired neuronal differentiation is associated with a wide array of disease such as Autism spectrum disorders [298], epilepsy [299], and Alzheimer's disease [300]. Importantly, inborn errors of metabolism are linked to irreversible brain injury [301]. In the context of this thesis, mutations in the gene encoding sequestome 1 (SQSTM1/p62) have been found in exome sequences of nine patients with suspected mitochondrial dysfunction. The patients presented with childhood- or adolescence-onset of neurodegeneration, ataxia and cognitive decline [302]. p62 is an multifunctional scaffold protein that interplays with, among others, KEAP1-Nrf2 signaling during oxidative stress, and LC3-dependent autophagy [303]. In addition, p62 was found in complex with PINK1 to mediate the selective autophagy of mitochondria, termed mitophagy [304], however, this finding is still controversially debated [305]. Eleven more patients with similar symptoms have since been reported [306], yet, the molecular link between defective p62 function and neurodegenerative disease is not clear and subject to study III.

6 MITOCHONDRIAL RNA METABOLISM

Mitochondrial genes are transcribed by the DNA-dependent RNA polymerase POLRMT and auxiliary factors from promoters on both mtDNA strands [307]. This generates polycistronic transcripts that cover almost the entire genome that are processed by endonucleolytic cleavage at tRNA junctions (see **Figure 9** for a comprehensive overview). The cloverleaf folding of tRNAs has been suggested to be essential for recognition of splice sites by the mitochondrial endonucleases, RNase P and RNase Z, which is known as the tRNA punctuation model. However, not all transcripts are flanked by two tRNAs, and the precise mechanisms are not fully understood.

The released transcripts undergo polyadenylation with about 50 adenines by mitochondrial poly(A) polymerase (MTPAP) as part of transcript maturation [308]. In chloroplasts and bacteria, the polyadenylation of transcripts promotes transcript degradation [309, 310], and polyadenylation of cytosolic mRNAs in eukaryotes leads to increased mRNA stability and translational efficiency [71, 311]. In mitochondria, polyadenylation is required to generate functional stop codons of a number of mitochondrial transcripts, but the exact function of the poly(A) tail in mitochondria is still unknown [308]. Patients with mutations in MTPAP presented with spastic ataxia and optic atrophy, indicating an important role for polyadenylation in mitochondria [312].

A number of key players are relevant for mitochondrial RNA turnover, including polynucleotide phosphorylase (PNPase), SUV3 helicase, leucine-rich pentatricopeptide-repeat containing (LRPPRC) protein and SLIRP [307]. In some organisms, polyadenylation is catalysed by PNPase, which displays MTPAP activity *in vitro* and can reversibly catalyse both RNA degradation and 3′ tail extension [313]. A dual localization of PNPase to both matrix and intermembrane space of mitochondria has been proposed, where it may regulate mitochondrial import of nuclear-encoded RNAs into mitochondria [314]. However, the mechanisms of mitochondrial RNA import are still not known [315]. Knock-down of SUV3 is pupal lethal in *Dm* and causes accumulation of unprocessed transcripts, loss of mature tRNAs and compromised OXPHOS activity [316]. This effect was found independent of PNPase, although there is evidence of a degradosome heterocomplex of SUV3 and PNPase [317].

The RNA-binding proteins LRPPRC and its interaction partner SLIRP are required for transcript stability and efficient delivery of RNA to the ribosome [318]. A tissue-specific mouse model with disrupted LRPPRC expression in heart caused mitochondrial cardiomyopathy and presented severely reduced mitochondrial transcript levels and reduced polyadenyla-

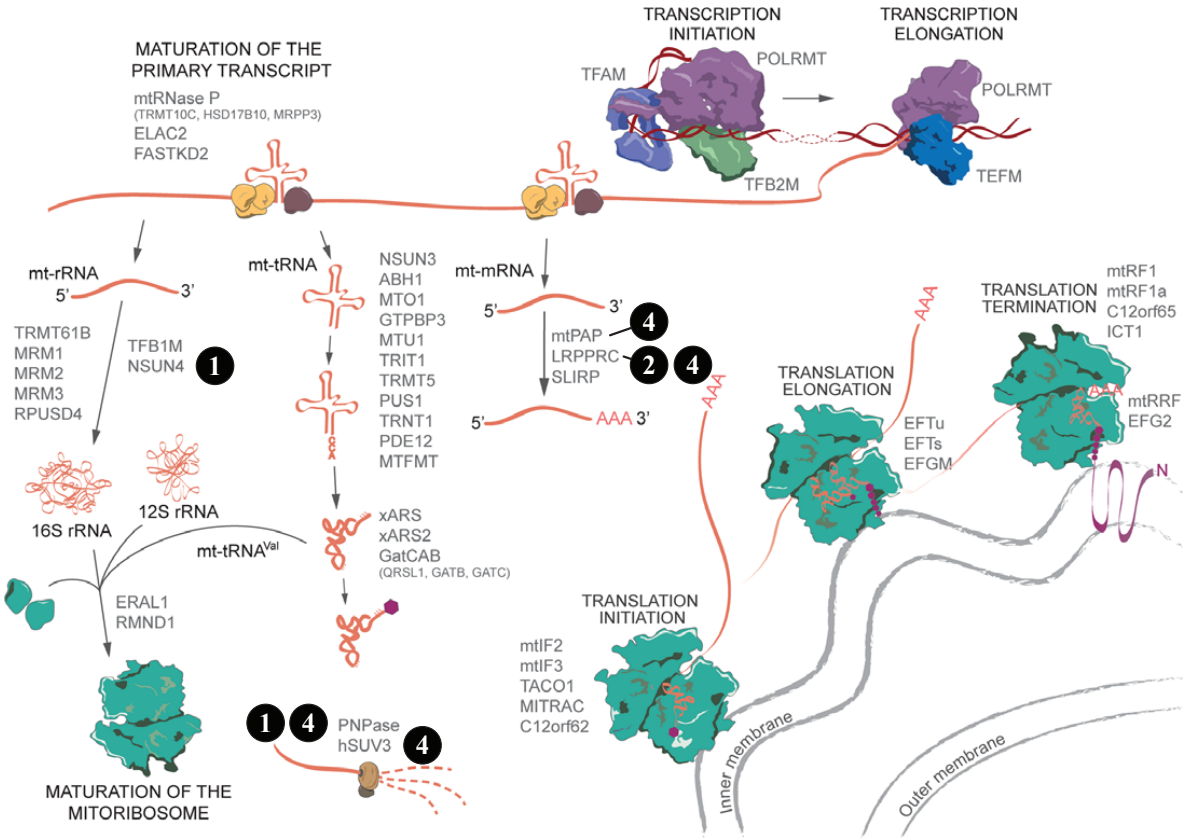


Figure 9: Schematic overview of mitochondrial transcription and translation. Proteins that are relevant to this thesis work are indicated with filled black circles and the study number they relate to. Reprinted from D'Souza & Minczuk [307] under the terms of the Creative Commons Attribution Licence 4.0 (CC BY).

tion [319] without affecting tRNAs and rRNAs. Proteomic analyses of mouse hearts showed loss of complex I, III, IV and V subunits, while complex III and all nuclear-encoded complex II subunits were unaffected [111]. In humans, mutations in LRPPRC cause a French-Canadian type of Leigh syndrome with early-onset of mitochondrial disease [320]. Knock-down of the LRPPRC fly homolog bicoid stability factor (bsf), in study II also called DmLRPPRC1, similarly resulted in loss of mitochondrial mRNAs, polyadenylation and subsequent OXPHOS dysfunction specific to complex I, III and IV [321].

Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits.

Universal Declaration of Human Rights, Article 27 (1948)

7 RESEARCH AIMS

The aim of this thesis work is to adapt and apply tools to map the diversity of mitochondria, to generate hypotheses on the aetiology of mitochondrial dysfunction, and to build novel functional networks that are useful in understanding human disease. In study I and II, I want to shed light on the fine tuning of mitochondrial function by post-translational modifications (PTMs). The overall set of transcripts and proteins is the interest of study III and IV, in which we aim to identify the primary cause of mitochondrial dysfunction and arising secondary consequences by transcriptomics and proteomics.

Methylation is a small protein modification that is essential for epigenetic control of gene expression in the nucleus, and we are also starting to appreciate its widespread relevance for cytosolic protein control. At the onset of **study I**, three patients with mitochondria-specific depletion of the primary methyl group donor SAM had been reported by Kishita *et al.* [240]. The symptoms of the three children ranged from repeated episodes of lactic acidosis, cardiopulmonary complications to death at five days after birth. This highlighted the relevance of mitochondrial methylation in human pathophysiology. However, the molecular disease-causing mechanisms were mostly unclear. The aim of study I is to generate patient-specific fly and mouse models of mitoSAM deficiency by mutating the inner-membrane carrier SAMC, and to characterize the arising defects with molecular and systems biology techniques. In addition, the targets of mitoSAM are not well characterized. I will apply a multi-layered screening strategy to identify methylation sites on mitochondrial proteins (the mitochondrial methylome) in fly, mouse and human. Combining the methylome with model data, we aim at pinpointing relevant sites that are supporting mitochondrial function. SAMC is the only known route for SAM into mitochondria, but considerable residual methylation in a patient with truncated and likely fully dysfunctional SAMC suggested that other transporters might exist. We aim at identifying redundancy in mitoSAM import, and exploit these in the design of future treatment strategies.

The quest to identify the mitochondrial methylome involved SILAC in fruit flies with heavy methionine, for which three similar technical approaches had been published [81–83]. However, we soon realised the technical shortcomings of previous methods with low labeling efficiency, undesirable amino acid conversion and astronomical costs. Therefore, I aim at developing a more suitable and flexible method for SILAC in flies in **study II** based on a fully defined holidic food source [322]. I want to technically validate the method and highlight its power in a difficult-to-solve showcase. Modification of mitochondrial proteins by phos-

phorylation is under-represented across the cell [169, 170] and requires accurate and highly sensitive proteomics approaches during sample preparation and MS data acquisition. In a fly model of mitochondrial dysfunction, I want to identify and quantify mitochondrial protein phosphorylation sites that are sensitive to an OXPHOS-specific primary defect. This study will contribute to higher data robustness that is transferable to study I and, importantly, shall raise attention for the power of proteomics in a marvellous model system in the fly community.

Variants in up to 870 genes are known that are linked to metabolic disease, including mitochondrial dysfunction [33]. In order to identify treatment options, knowledge on the molecular pathomechanism is important, but in many cases not known, especially for novel variants. At the start of **study III**, Haack *et al.* had reported nine patients with three biallelic loss-of-function variants in SQSTM1/p62 that presented with slowly progressive neurodegenerative disease [302]. p62 is an autophagy adaptor protein with a suggested role in mitophagy. We want to elucidate the causative link between neurodegeneration and mitochondria upon loss of p62. In patient-derived NESC, we perform transcriptomics and proteomics to identify dysregulated functional protein networks and metabolic pathways. This is the first study that uses a multi-omics approach in our laboratory, and thus the aim is also to benchmark *omics* techniques for molecular characterisation of mitochondrial dysfunction.

While the role of individual proteins in mitochondrial RNA metabolism has been studied in depth, their functional interplay is not well understood. In **study IV**, we generate individual and combined mutant fly models with affected SUV3, PNPase, DmLRPPRC1/bsf and PAP expression, and characterize their involvement in RNA metabolism and the order of events they act in. Furthermore, by employing antibody-based enrichment and RNA-Seq, we want to explore the accumulation of stress-related RNA intermediates. This will be of relevance in diseases that impair mitochondrial RNA turnover, and contribute to a better understanding of the post-transcriptional events in the organelle.

8 FURTHER MATERIALS AND METHODS

8.1 Generation of fly models

Gene-specific knock-in *Dm* lines were generated by ends-out gene targeting [323], replacing the endogenous locus of CG4743/*SAMC* with a w^+ marker gene and an *attP* cassette (**Figure 10**). The *attP* site was subsequently used to introduce *attB*-site bearing CG4743-constructs with patient or non-sense mutations, or a control rescue construct.

8.2 Mitochondrial SAM import assay

Mitochondrial SAM import in *Dm* was quantified by incubating 50 µg of isolated mitochondria in a suitable buffer containing 1.12 µM radioactively labeled [$^3H_3$] SAM for 20 minutes at 30 °C. Mitochondria were then washed in ice-cold assay buffer supplemented with 1 mM unlabeled SAM, and radioactive decay was quantified on a β-counter.

8.3 Oxygen consumption measurements

Electron transport chain-dependent respiration was quantified in disrupted *Dm* larvae using an Oxygraph-2K respirometer (Oroboros). Samples in respiratory buffer were permeabilizied with digitonin. Complex I-dependent respiration was quantified with pyruvate, glutamate, malate and ADP. Complex I+II-dependent respiration was quantified after addition of succinate, and complex II-dependent oxygen consumption was derived after addition of the complex I inhibitor rotenone. The non-mitochondrial background was determined after addition of antimycin A. Lastly, complex IV activity was measured after addition of tetramethyl-p-phenylenediamine (TMPD) and ascorbate, followed by shutdown of the system with potassium cyanide.

8.4 Bisulfite pyrosequencing

Mitochondrial 12S ribosomal RNA has two methylation marks on cytosines at murine position 909 and 911. An assay has been developed to quantify the site-specific occupancy in fly and mouse by low-throughput sequencing. Purified and DNase-treated RNA was bisulfite-converted, and a segment of 12S was reverse transcribed. A biotinylated primer per species

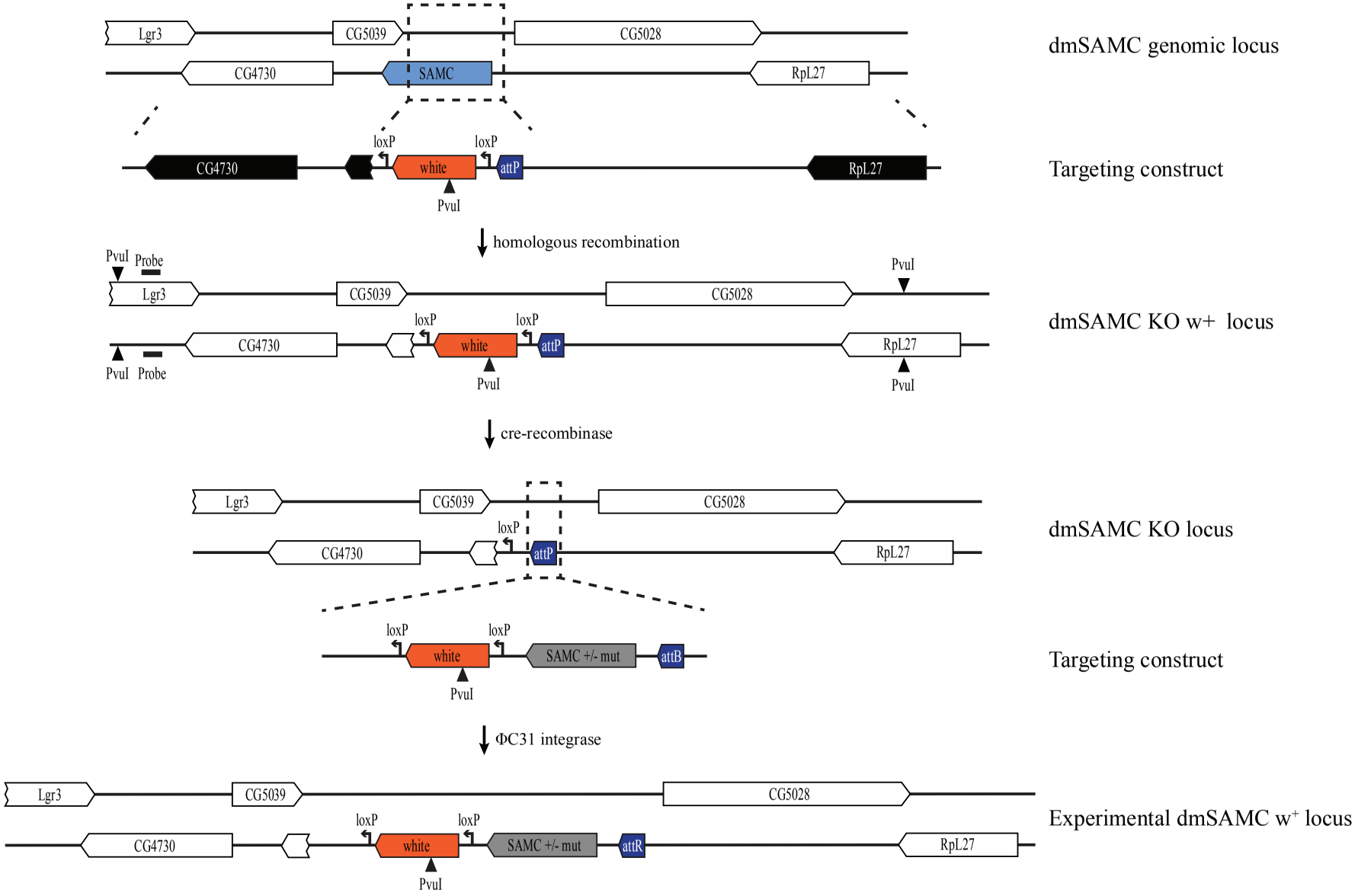


Figure 10: Genomic strategy to obtain mutant SAMC flies. The endogenous CG4743/*SAMC* locus is replaced by an *attP* site, which was subsequently used for introducing designed *SAMC* constructs.

was chosen that annealed next to the two sites of interest, and the sequence was read by pyrosequencing. The occupancy was expressed as modified over total signal at the two sites.

8.5 SILAF and methyl-SILAF

Stable isotope labeling of amino acids in flies (SILAF) was developed and performed as described [324, 325]. In brief, the unlabeled amino acids lysine, serine and methionine of a holidic food source for *Dm* [326, 327] were changed to equimolar amounts of L-lysine-[$^{13}C_6$, $^{15}N_0$], L-serine-[^{13}C,^{15}N,2,2,3-$^{2}H_3$] and L-methionine-[methyl-$^{13}C^{2}H_3$]. Parental flies were kept on the holidic food source for 8 h to lay eggs. Experiments were performed on seven or eight day-old F1 larvae, or on hatched flies. Peptides were prepared by Lys-C or tryptic digest for lysine or methionine labeled proteins, respectively.

8.6 Bioinformatic data analysis

8.6.1 Proteomics

Quality control of raw files was performed with RawMeat (Vast Scientific). Computations were performed on a local machine or on resources provided by Swedish National Infrastructure for Computing (SNIC) through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under project SNIC 2019-8-175 and uppstore2018188. For label-free quantifications, Quandenser was used at standard settings [328] on the cluster. To identify methylation sites in methyl-SILAF data, MaxQuant [156] was set up on the cluster. Light and heavy methyl groups were identified in separate runs, and a confidence score was calculated with [184]. The incorporation rates of labeled amino acids was assessed locally with MaxQuant [156] in a dual-labeling experiment, and the ratio of heavy-to-light peptide intensity for methionine labeling and heavy-to-light protein intensity for lysine labeling was calculated as a proxy for label incorporation rates. The phosphoproteomes on lysine-labeled SILAF samples were quantified on a local machine with MaxQuant [156] in dual label model with phosphorylation as a variable modification. Spectra from tandem mass tag (TMT)-labeled peptides were mapped locally with MaxQuant [156] by manual configuration of reporter ions and batch-dependent correction factors.

All subsequent analysis was performed in R that was run in an R Studio environment or through console scripts. Differential expression proteomics used a limma moderated *t* test [329] and FDR-adjusted *p* values.

8.6.2 RNA-Seq

The 75 base pair (bp) paired-end sequences were quasi-mapped against a cDNA library with Salmon [330] and differential expression was calculated with DESeq2 [331] including *p* value adjustment by independent hypothesis weighing [332]. Double-stranded RNA was quantified

after J2 anti-double stranded RNA antibody-based enrichment following a previously established protocol [333]. 42 bp paired-end sequence reads were mapped against a whole-genome library with Bowtie 2 [334], converted and indexed with Samtools 2 [335] and visualised in IGV viewer [336].

8.7 Ethical considerations

Three of four studies involve a basic molecular characterisation of genetically modified fruit flies with the intention of mimicking the disease status of human patients. In study I, the fly data was subsequently validated in a mouse model. Morally, several layers of complexity arise: (1) terminating the life of a large number of fruit flies, (2) modifying fruit fly genomes, (3) mimicking disorders of alive patients in flies, (4) breeding and terminating the life of mice and (5) using public financial resources to understand the disease mechanism of three known affected individuals instead of funding research on widespread diseases.

Fruit flies are a collection of cells that propagate themselves, react to environmental stimuli including stress, and are able to intermittently keep the fine balance between order and chaos, called life. This description fits to humans, mice, flies and yeast cells alike. Doing research strictly following Kant's categorical imperative would put us into trouble as we would not be allowed to distinguish between any life form and could neither use mice nor cell culture in our work. Yet, our cultural imprinting on moral standards would make us draw a rectangle around flies and yeast cells that we label with "acceptable to kill". It has been argued that the less empathy we feel with an individual, be it human or not, the more acceptable it is to do harm to it. It is scientifically almost impossible to prove or disprove conciousness in flies, which sets them far apart from humans in our perspective. I follow this common perception here, but lacking a truly substantial ethical argument scientists shall take special care not to make the insects suffer unnecessarily and to treat them with respect. On top of this, understanding the nature as a whole is a driving force that sets certain animals apart from others due to natural neuronal constraints. Doing so, the physical environment including wild living flies can benefit from my activities and I think *Homo sapiens* can regard it as their responsibility to protect the planet Earth. That this is blunt theory and not happening in reality is part of another discussion. However, along this line, modifying a fruit fly genome contributes to a general understanding of the world as long as these flies stay contained and do not lead to a shift of the genetic pool of a wild living population.

Following these arguments, it is irrelevant from a fly perspective if the mutation is derived from a patient or not. However, a patient might suffer from the idea that her or his disease is mimicked in fruit flies, which could lead to the perception that the patient is similar to a fly. Yet, it might be a comforting thought for a sick person that researchers invest their life time and public money into finding a molecular disease mechanism. More utilitaristic, the molecular explanation is ultimately a basis for treatment and although we are not able to treat

most monogenic diseases today, my work will help patients on a long run to lead a normal life, that is without permanent medical assistance.

Applying the moral reasoning that the life of a fly can be sacrificed for the good of human kind and beyond, is clearly more difficult to apply to mice with our cultural perception in mind. It is easier to find empathy with a mouse than with a fly. On the positive side, study I uses first fruit flies to identify relevant disease mechanisms and continues with a more targeted strategy to mice. Doing that, we cut down the number of animals. Yet, are we allowed to kill and genetically modify those that remain? As before, it has to be made sure that mice do not suffer unnecessarily from our actions, which requires thorough planning. Additionally to that, I personally do rate the life of human beings higher. I am comforting myself with the thought that 2000 years of natural philosophical history wouldn't have been so controversial if a clear-cut threshold between ”acceptable” and ”unacceptable to kill” would exist.

Lastly, the research we are doing is without doubt costly, however, biomedical research in the last 150 years has majorly contributed to the reduction of human suffering all across the globe. If our moral maxim is to contribute to every humans well-being, responsible biomedical research is fully acceptable. I think this is true irrespective of the number of people the research is targeted at. Our study design originated from the pathologies of three patients. Yet, we think that the outcome has implications for cancer and ageing. The broader outcomes for any research study are hard to predict *a priori*, but understanding the principles of cell biology and a solid ground of knowledge is inevitably required to ultimately cure human diseases.

The sycamore fig cannot ripen unless it is scraped;
but they scrape it with iron claws; the fruits thus scraped ripen in four days.

The first known documentation of a S-adenosylmethionine dependent process
by Theóphrastos (c. 371 – c. 287 BC), Enquiry into Plants, book IV
English translation by Sir Arthur Hort

9 RESULTS AND DISCUSSION

9.1 Paper I: The One-Carbon Pool Controls Mitochondrial Energy Metabolism via Complex I and Iron-Sulfur Clusters

The cytosolic methionine cycle produces the methyl group donor SAM, which is then distributed throughout the cell. In mitochondria, SAM enters the matrix *via* the solute carrier protein SLC25A26/SAMC and is a known cofactor during RNA and protein methylation and metabolite synthesis. However, the severe symptoms of patients with mutations in SAMC suggest that our understanding of mitoSAM-dependent modifications are incomplete.

Therefore, we modeled mitoSAM deficiency by replacing the endogenous SAMC with three patient-specific SAMC variants (*a123v*, *i172g*, *p223l*), a non-sense mutated *null* and a control *rescue* construct in *Dm*. In the fly, the gene CG4743 is the ortholog of human *SLC25A26/SAMC*, and we confirmed the mitochondrial localization of C4743 by immunofluorescence microscopy. All patient mutant and *null* variants caused decreased import of SAM into mitochondria, but only the *p223l* and *null* mutants showed a severe larval-lethal phenotype with developmental arrest at an early stage. We observed loss of the SAM-dependent metabolite lipoic acid. Moreover, markedly reduced ubiquinone-9 levels correlated with a coupled respiratory chain defect of complex I to III and complex II to III, without affecting isolated complex activities. Transcriptomics and proteomics revealed a prominent up-regulation of enzymes related to lipoic acid, ubiquinone-9 and iron-sulfur cluster synthesis.

We confirmed the findings in *Dm* in a mouse model. Homozygous knock-out of SAMC is embryonic lethal, thus we obtained MEF cells from mouse embryos, which we were able to grow in cell culture. Similarly, we observed a complete loss of lipoic acid levels, strongly impaired oxidative phosphorylation, and additionally loss of assembled complex I, III, IV and V. Methylation of two 12S ribosomal RNA methylation marks were completely absent, arguing that no redundant mechanisms for SAM import into mitochondria exist in mouse. Interestingly, the loss of lipoic acid caused stalling of the oxidative TCA cycle and accumulation of α-ketoglutarate. Reassuringly, deep labeling data was compatible with the reductive carboxylation of glutamine-derived α-ketoglutarate as seen by others [337, 338].

We then asked what the protein targets of mitoSAM are. We employed a MS-based methylation screening in *Dm* with methyl-SILAF labeling that was specifically developed for this study. In addition, we increased coverage by multiple protease digestions, antibody-based en-

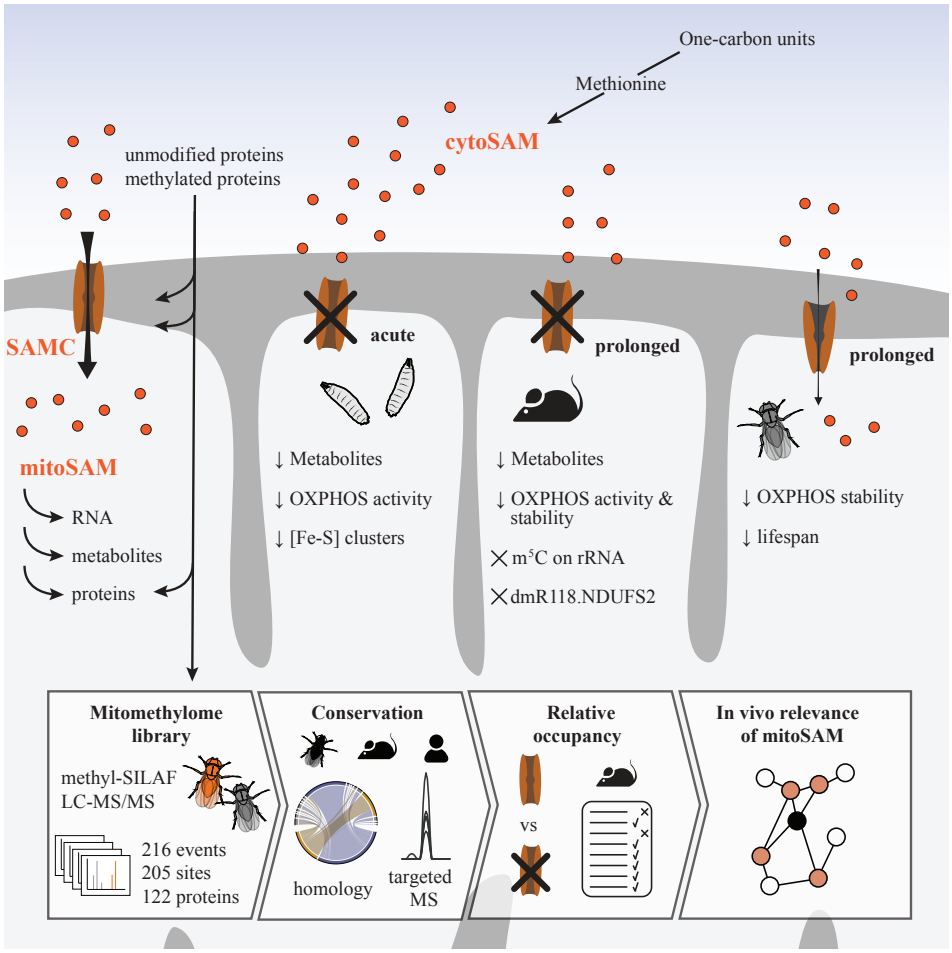


Figure 11: Graphical summary of paper I.

richment of methylated peptides and offline fractionation. We mapped 216 high-confidence methyl groups on 205 mitochondrial proteins, without enrichment of a particular functional category. By medium-scale targeted proteomics, we confirmed 20 and 24 sites in human and mouse samples, respectively. Most of these sites have not been described previously, and several positions have been found mutated in patients with mitochondrial disease including arginine 20 in the complex IV subunit COX6B1 [339] and arginine 72 [340] in the cysteine desulfurase NFS1 (both human protein positions).

We then asked whether methyl groups would be lost in our models of mitoSAM deficiency. While we did not observe any decrease in fly samples, two site-specific signals were lost in MEF cell samples while the majority did not change. This suggests that modification of mitochondrial proteins can happen prior to import and is consistent with previous large-scale data [136, 341]. We combined the fly model *omics* data and the methylome library to pinpoint methyl groups with a potential biological function, and identified a methyl mark on the clinically relevant arginine 72 of NFS1. When we mutated this site in a human cancer cell line, the integration of NFS1 into functional heterocomplexes was impaired. It is currently unclear whether the methylation of NFS1 happens inside of mitochondria, and peptide cleavage patterns limited our experimental approach. When ageing a *i172g* SAMC fly mutant, we

observed decreased lifespan, compensatory up-regulation of serine biosynthesis and a specific loss of complex I subunits. Using data of the Cancer Dependency map, we found a prominent correlation between gene essentiality of SAMC and the complex I subunit NDUFS2, which also lost its methyl mark in SAMC knock-out MEF cells. Yet, we failed to observe a significant reduction in complex I activity, which might be related to the sensitivity of the applied methods.

Finally, we tested the efficiency of a number of metabolites in rescuing larval lethality. We identified that both methionine and SAM could elevate cytosolic SAM levels and in part restore depletion of mitoSAM. Notably, this could partially rescue the larval growth defect, which can be of clinical importance. Furthermore, we could show that SAMC protein levels do not regulate mitoSAM concentration, contradicting reports on increased mitochondrial DNA methylation upon over-expression of SAMC [277]. We thus conclude that the cytosolic production of SAM downstream of the one-carbon cycle directly translates into mitoSAM levels and controls mitochondrial function through complex I and iron-sulfur cluster biosynthesis. Dysregulated one-carbon cycle has been linked to various diseases, most notably tumor development through SAM-dependent epigenetic imbalance [202, 342, 343]. This highlights the relevance of our findings not only in the pathophysiology of mitochondrial disease, but also in cancer and ageing-related disorders.

9.2 Paper II: Stable Isotope Labeling of Amino Acids in Flies (SILAF) Reveals Differential Phosphorylation of Mitochondrial Proteins Upon Loss of OXPHOS Subunits

To map the high-confidence methylome in study I by methyl-SILAC, we were in need for a robust stable isotope-labeling technique in *Dm*. Adaptations of SILAC to *Dm* husbandry practices exist [169, 170] in which labeled yeast is fed to flies. However, the indirect methods suffer from undesirable amino acid conversion and low labeling efficiencies of at most 97.6% just above the threshold of SILAC usability. To this, high prospective costs limited their adaptation by the fly community.

In paper II, we describe the methodological development of a direct labeling technique for fruit flies that is highly efficient, flexible and cost effective. We use a previously described holidic food source [322, 344] that is composed of 47 defined and purified chemicals, including eight essential amino acids. Initially, we followed a classical SILAC approach and exchanged both ”light” lysine-[$^{13}C_0$] and arginine-[$^{15}N_0{}^{13}C_0$] against ”heavy” stable-isotope labeled lysine-[$^{13}C_6$] and arginine-[$^{15}N_4{}^{13}C_6$] with subsequent tryptic digestion. We observed a profound conversion of heavy arginine into proline and thus a loss of MS-signal from heavy peptides and compromised identification rates. Henceforth, we switched to lysine-[$^{13}C_6$]-labeling only with endoproteinase Lys-C digestion. This normalized light and heavy peptide identification and did not perform worse than comparable light tryptic digests. We termed the

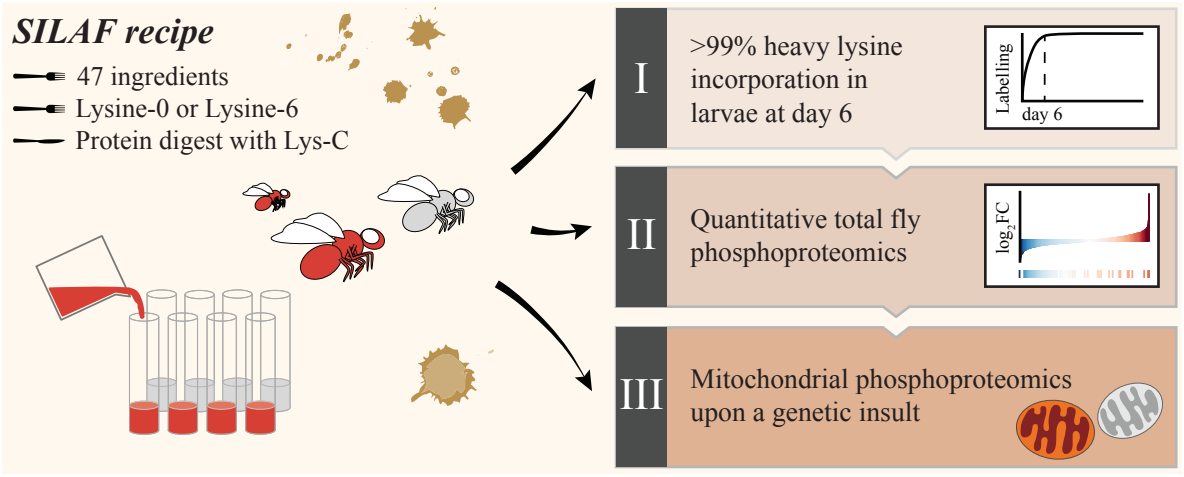


Figure 12: Graphical summary of paper II [324].

method SILAF and found it slightly more reproducible than label-free quantification.

The labeling efficiency reached more than 99.5% at larval stage 2 when larvae were grown from egg laying on labeled food, thus outcompeting all previous methods. Heavy amino acids were also incorporated when unlabeled flies were put onto SILAF food, albeit at a much lower rate reaching at most 45.0% after two labeling weeks. Interestingly, we could use this pulse-SILAF approach to monitor protein turnover in functional protein categories and found that ribosomal components were turned over much faster than OXPHOS subunits in adult flies.

SILAC can reduce variabilities introduced during sample preparation by combining a light and heavy fraction early on in the preparation process. SILAC is therefore especially powerful for analysis pipelines that demand pre-MS offline fractionation. We thus tested SILAF on mixes of male and female flies and performed small-scale peptide fractionation. We observed high correlation between replicates and biologically meaningful differential expression that reproduced previously published yeast-based SILAC experiments [81]. We then asked whether we could quantify phosphorylation site occupancy and chose to compare yeast-food to holidic-food grown flies. SILAF combined with high pH reversed phase liquid chromatography and LC-MS/MS quantified 5,038 proteins and identified 14,008 phosphorylation sites, of which 1,523 occupancies were quantified. We found up-regulation of enzymes related to gluconeogenesis, glycogen degradation and the ribosome, and down-regulation of glycolysis and OXPHOS proteins, suggesting increased energetic usage of amino acids.

Finally, we applied SILAF to a fly model of mitochondrial disease. We knocked down bsf/DmLRPPRC1, which is involved in mitochondrial protein biosynthesis by stabilizing mRNA transcripts. DmLRPPRC1 loss caused an almost exclusive reduction of OXPHOS components, which is consistent with previous molecular data in that model [71]. We mapped the protein phosphorylation landscape in enriched mitochondrial samples, and found that protein phosphorylation sites are under-represented in the mitochondrial compartment compared to the whole cell. This is in agreement with reports in mouse and human [169, 170].

Among the differentially phosphorylated sites upon DmLRPPRC1 knock-down, we con-

firmed previously known regulatory marks and found two novel sites on the complex I subunit NDUFB10 (tyrosine 126) and the complex IV subunit NDUFA4 (serine 66) with a significant stoichiometric increase. Both amino acids are conserved to human and the phospho-groups have been detected in previous large-scale reports, but not linked to a biological function. Tyrosine 126 of NDUFB10 anchors the protein into ND4 in complex I, and absence of a cryogenic electron microscopy trace at the complex I site suggested that only the free, but not the incorporated subunit is phosphorylated. We mutated serine 66 of NDUFA4 to the non-phosphorylatable serine mimetic alanine in a human cancer cell line and observed less association of the mutated protein with intact complex IV on native gel electrophoresis. This argues that phospho-serine 66 might be required for binding of NDUFA4 to complex IV. Of note, it has been proposed that NDUFA4-binding prevents dimerization of complex IV and thus might impact supercomplex formation [345]. However, the initial screening requires additional and extensive molecular follow-up studies to underpin the structural relevance of both phospho-sites.

We conclude that SILAF is a highly powerful technique for label-based fly proteomics with almost complete labeling at a fraction of previous costs. In fact, one milligram of fully labeled complex peptides can be obtained for less than $15. In an additional methods paper, we provide a simple-to-follow recipe with notes [325] and we hope that the fly community uses this opportunity to perform more proteomics studies on fly models.

9.3 Paper III: SQSTM1/p62-Directed Metabolic Reprogramming Is Essential for Normal Neurodifferentiation

Haack *et al.* reported nine individuals with biallelic loss-of-function variants in the autophagy receptor protein SQSTM1/p62 [302]. The suggested role of p62 in removal of damaged mitochondria led us study the link between loss of p62 in patients and compromised mitochondrial function. Due to the strict neurodegenerative clinical symptoms, we obtained NESC from two patient fibroblast lines. Loss of p62 severely compromised differentiation into mature neurons in patient and CRISPR/Cas9 knock-out NESC lines.

Assessment of mitochondrial function showed a mild reduction of complex I activity while ATP production was normal. However, we observed increased sensitivity to the protonophores carbonyl cyanide m-chlorophenyl hydrazone (CCCP) and 2,4-dinitrophenol (DNP). Basal levels of ROS in NESC were normal, but were significantly increased upon induction of differentiation in patient lines alongside reduced levels of transcripts related to glutathione metabolism. Treatment with the antioxidant N-acetyl cysteine (NAC) could in part rescue the defect. Components of the KEAP1-Nrf1 signaling pathway were unchanged in mutant cells, which suggests that p62 acts independent of this relay during oxidative stress. Of note, mitochondrial clearance upon stress was not affected by loss of p62.

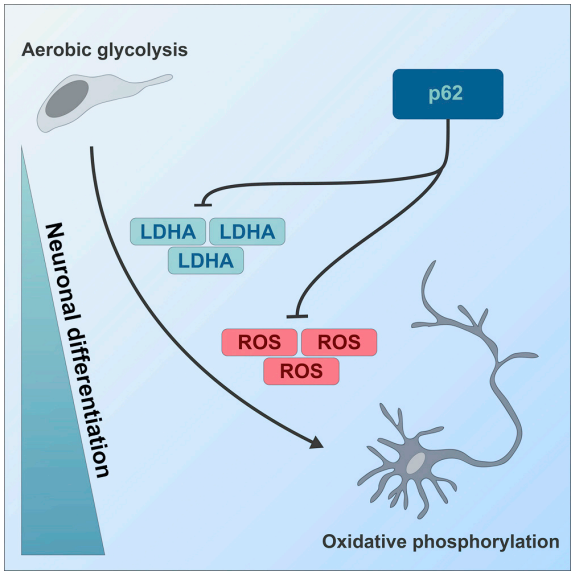


Figure 13: Graphical summary of paper III. Reprinted from Jalvo-Garrido *et al.* [346] under the terms of the Creative Commons Attribution BY-ND Licence 4.0.

We performed transcriptomics and proteomics in resting NESC and at four days after initiation of differentiation. Functional pathway analysis confirmed the onset of neuronal differentiation in controls, but a much less pronounced transcriptional response in knock-out cells. Furthermore, proteins of the OXPHOS system were readily increasing in knock-outs, while this well-known signature of neuronal development was absent in mutant cells. We found a prominent accumulation of LDHA due to loss of p62 at resting state, and disturbed pyruvate metabolism. Reduction of LDHA expression is a hallmark of neuronal stem cell differentiation [296] and the initial abnormally high levels of LDHA in p62 knock-out cells might explain the failure to differentiate, although LDHA was normal in patient fibroblasts.

Our data finds support in a recent report that describes increased lactate levels in cerebrospinal fluid of patients with dysfunctional p62. The same study generated a zebrafish model, in which loss of p62 led to structural abnormalities in hippocampus [306]. Thus, we suggest that progressive neurodegeneration in patients with loss of p62 is caused by a failure to switch from a pro-glycolytic signature to oxidative phosphorylation during differentiation of neuronal stem cells.

9.4 Paper IV: Defects of mitochondrial RNA turnover lead to the accumulation of double-stranded RNA in vivo

Mitochondrial RNA turnover is regulated by a set of enzymes, including PNPase and SUV3 that form the degradosome complex, MTPAP and LRPPRC. How they interact *in vivo* is incompletely understood. In this study, we characterize the function of DmPNPase, and combine this fly model with knock-down, knock-out or over-expression lines of the other factors. CRISPR/Cas9-induced knock-out of the PNPase fly homolog CG11337 caused larval lethality. We found reduction of isolated and coupled OXPHOS activity together with accumulation of mitochondrial mRNAs and loss of specific tRNAs. PNPase and SUV3 acted synergistically on mitochondrial mRNAs, as concomitant knock-down had additive effects on mRNA accumulation. Similarly, knock-out of MTPAP caused loss of polyadenylation and accumulation of mitochondrial mRNAs, which was rescued by simultaneous over-expression of PNPase, suggesting that the

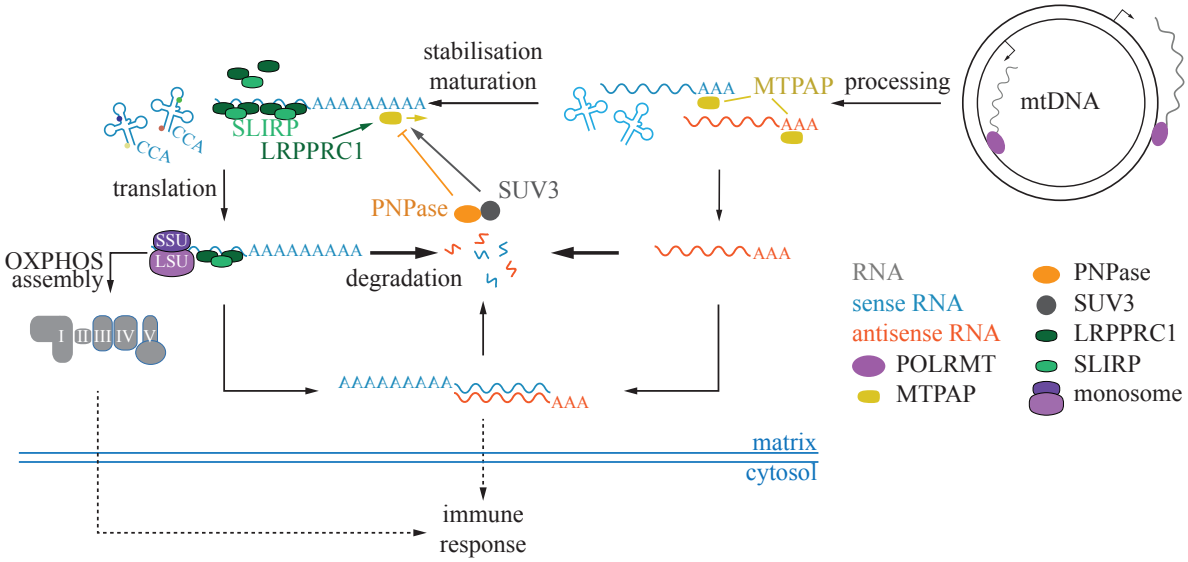


Figure 14: Graphical summary of paper IV. Solving the puzzle of mitochondrial RNA metabolism. Adapted from Pajak *et al.* [347] under the terms of the Creative Commons Attribution Licence 4.0 (CC BY).

polyA-tail on mitochondrial transcripts has neither an exclusively protective role nor signals degradation. Knock-down of LRPPRC reduced mRNA abundance, as has been described in *Dm* before [71], and this was reversed by simultaneous knock-down of either PNPase or SUV3, indicating that LRPPRC is in fact required to protect mitochondrial mRNAs from degradation by the degradosome complex.

Both strands of mtDNA are transcribed from dedicated promoters into polycistronic RNA transcripts that are processed to mature transcripts. This causes the formation of antisense RNA species that are rapidly degraded under physiological conditions [348]. We found accumulation of double-stranded RNA by immunofluorescence microscopy when PNPase or MTPAP was lost, and these transcripts preferentially localized to the cytosol. This was confirmed by antibody-based enrichment of double stranded RNA and subsequent RNA sequencing [333]. This indicates that mitochondrial RNA can be released into the cytosol, and that disruption of mitochondrial RNA metabolism causes accumulation of double-stranded transcripts. In a human cell line, mitochondrial release of double-stranded RNA activates an MDA5-driven antiviral response [333]. Although this pathway is not conserved in flies, we did observe down-regulation of antiviral response genes upon loss of any of the herein studied proteins, suggesting increased sensitivity to infections of these models. Whether this translates into measurable pathological consequences during immune surveillance is still unclear. However, immunological failure upon mitochondrial dysfunction is currently an active area of research [349], and our findings can contribute to the question why some patients with mitochondrial dysfunction are more prone to infection [350].

10 CONCLUSIONS AND PERSPECTIVES

Accurate quantification of thousands of transcripts and proteins within hours - it is difficult for a molecular biologist to dismiss the appeal. In this thesis work, I made large-scale data compatible with molecular findings and developed new methods to unravel the complexity of mitochondrial function in health and disease. This gives an exciting new perspective on the intricate and complex function of this organelle, and offers a roadmap for arising and urgent questions on mitochondrial dysfunction.

We zoomed into post-translational modifications on mitochondrial proteins, and by combining powerful proteomics techniques with genetic models we mapped the mitochondrial protein methylome in the three species fly, mouse and human. While we were able to follow up on just two methylation sites, two hundred more remain to be studied that do not have an annotated biological function. We were surprised to find most proteins methylated outside of the organelle, and it will be highly interesting to understand the molecular machinery that is able to modify polypeptide chains prior to import. We propose that mitochondria are responsive to one-carbon cycle activity and cytosolic SAM production rates, comparable to the epigenome and SAM-dependent metabolite synthesis as shown by others. We can partially rescue the growth defect of a fly with compromised SAM import into mitochondria and the great hope is that this can benefit sick children and adults with SAMC variants. Moreover, with the growing body of literature on the one-carbon cycle in cancer metabolism and ageing, our results can be of wide impact outside of the mitochondrial field.

The fruit fly is an extremely powerful model system, and thousands of lines are publicly available at stock centres. This, in combination with their controlled genetic background makes the fly ideal for *omics*-driven research. Some tools are readily available for this, while others are missing. We developed SILAF, a versatile and efficient labeling technique for proteomic studies on *Drosophila*. While SILAF was highly useful for our PTM screens, we hope that the method excites researchers in the *Drosophila* community to exploit proteomics more. The true power of systems biology can only unfold with a sufficient number of datasets that are made publicly accessible along with assiduous documentation of the acquisition process. To bring fly studies up to speed, gene nomenclatures need to be updated to make large-scale datasets better understandable and comparable to other species. Approximately 75% of human disease causing genes are conserved in *Dm* [351]. However, the dominance of the systematic but vacuous CG nomenclature can make it hard for non-Drosophilists to appreciate the enormous potential of this model system.

I witnessed several times that molecular biologists see the dawn of systems biology as a threat to wet-lab based experiments, and I got asked whether it might be good investment to start learning programming now. I certainly got excited about the idea that we can get a deeper understanding of real biology with digital network models. However, I think molecular biology will continue to be an important backbone of biomedical research. For instance, study IV combines various genetic models to understand causality within the mitochondrial RNA degradation pathway. I encourage the reader to have a detailed look at this beautiful paper that I was very happy to have contributed to. The proposed mechanisms were logically deductible only by using genetics, and currently I cannot see how an *omics* study could have come to the same conclusions.

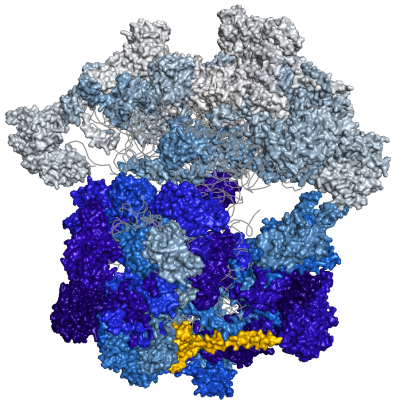
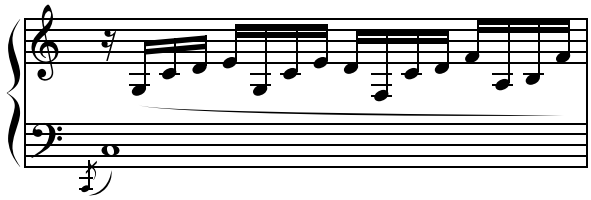


Figure 15: A correlation of fold changes over 100 proteomic datasets, projected onto the mitochondrial ribosome. Dark yellow: MRPL52; shades of blue: correlation rank with MRPL52 levels.

Molecular biology is very powerful when a small set of proteins are under investigation. On a larger systems scale, though, the human mind is limited in understanding complex interactions and networks with thousands of nodes. Systems biology can assist in delineating hypotheses on disease-causing mechanisms, and we will have an increasingly comprehensive understanding of the cell with more datasets. At the Division of Molecular Metabolism at Karolinska Institute, Oleksandr Lytovchenko and I have been collecting more than 100 processed proteomic datasets from mouse models and human cell lines with mitochondrial dysfunction [352]. We are starting to see exciting patterns already with the little data we have (**Figure 15**). The current limit of systems biology is data interpretation and most data points in large-scale studies remain unexplored when they are not of interest for supporting the central hypothesis. Thus, it will require both skilled programmers who make machine learning tools accessible to non-computer scientists, and knowledgeable molecular biologists with biological insight to push these limits of data interpretation. Systems biology is a novel era of bioscience that will be increasingly interdisciplinary, challenging, and fun.



Claude Debussy, Doctor Gradus ad Parnassum (1908)

11 ACKNOWLEDGEMENTS

To **Anna Wredenberg** and **Christoph Freyer**, it has been a privilege to learn from you, to work with you and to go crazy about science. Thank you for trusting in me, a more than solid scientific education, a life time experience and for preparing me for the future. And thank you for being just friends many, many times.

To **Anna Wedell** and **Nils-Göran Larsson**, thank you for your invaluable support, scientific insight and a lot of encouragement. You are role models of scientific leadership.

To **Ilian Atanassov** and **Ákos Végvári**, you have sparked and nurtured my interest in proteomics and systems biology. You have always been there for answering every question, and our interaction with a myriad of Slack messages, e-mails and drop-in sessions reached far beyond the framework of a collaboration.

To **Lukas Käll**, **Roland Nilsson** and **Joanna Rorbach**, you are inspiring personalities, thank you for sharing your insight into science as a profession and passion.

To **Almut Kelber**, your sharp mind helped me see the bigger picture.

To **David Moore**, sometimes we were in need of a couples therapist. Our SAMC team effort caused many different kinds of emotions, but all friction and friendship led to work we can certainly be proud of. Thank you for your kindness and calmness and for getting me excited about the climbing wall.

To **Isabelle** and **Cami**, we share being crazy, each in our own way. Journal clubs at 7.30am, seriously? You brightened up this journey, introduced me to coffee, showed me the upper human limit of hard work and taught me both a long list of very useful swearwords and *konsten att vara otroligt snäll*. Tack och grazie.

To all present and past members of the **Division of Molecular Metabolism**, you make this place a wonderful environment to work in. I wish you all the best! Thank you: **Camilla**, **Sussi** and **Leila**, you really make us unfold our true potential; **Paula**, for being a great teacher and mentor; **Roby**, **Ola**, **Sarah**, my friends and support team in all of life's circumstances,

for giving me courage, lighting fire despite the rain and developing my English; **Marco**, for never complaining about running our samples before dawn; **Javi**, for advise on becoming rich; **Ruth**, **Annika** and **Henrik** for the things you taught me, collaborations and great conversations; **Sasha**, for sharing your excitement about another line of working code; **Mara**, **Diana** and **Leona**, for your idealism and optimism that we can still save the planet, and for never spoiling the weekly cabbage surprise; **Laura**, **Rodolfo** and **David A.** for artworks of cooking, great science and collegiality; **Shan** and **Najla** for some remarkable conversations and cultural insight; **Anas** for never letting me down at fussball and enlightening me with ribosomal stories; **MolMad** including **Jelena**, **Miriam**, **Shree**, **Jing** and **Dima**, for the lightness of being a student, all the laughs and sharing PhD life - good luck with your defenses! **Giovanni**, **Guanbin**, **Yong** and **Duc**, for your infectious happiness. I am proud of the Heidelberg connection and all students, **Tobias**, **Juliane**, **Anna-Lena**, **Frederik** and the 60 youngsters that shared my excitement during five summer schools.

To everyone at **CMMS**, especially Rolf, Erik, Henrik, Helene and Martin for great collaborations and help, and all the CMMS workshops. To our neighbors in quarter 9D of **Molecular Structural Biology** and **MIR** for the chats and helping out many times. To the **Karlsson Hedestam lab** for all the fun, the food, and your quite amazing support.

To the **Studienstiftung des deutschen Volkes** for support since 2011, which allowed me complete academic freedom and gave me invaluable experiences during my studies.

To those who deserve many, many more lines of gratitude than these:

To **Uta**, thank you for sharing and shaping all aspects in life for many years and for being the most honest critic. From the bottom of my heart, all the best to you.

To my closest friends, the **Wuffs**, **Myriam**, **Steffi G.**, **Anna**, **Meike**, **Tilo**, **Carmen** and **Börge**, for miraculously never giving up on my geological perception of time.

An **meine Familie**, der ich alles andere verdanke, die immer für mich da ist und die meine Faszinatinion für Wissenschaft immer unterstützt hat. Sappralott, Zeit wird's, dass ma uns wieder seng!

Till **Steffi**, för hopp och kärlek och framtiden.

REFERENCES

1. Kirschner, M. W. The Meaning of Systems Biology. *Cell* **121,** 503–504 (2005).
2. Breitling, R. What is Systems Biology? *Frontiers in Physiology* **1,** 9 (2010).
3. Tavassoly, I., Goldfarb, J. & Iyengar, R. Systems biology primer: the basic methods and approaches. *Essays in Biochemistry* **62,** 487–500 (2018).
4. Martin, W. F., Garg, S. & Zimorski, V. Endosymbiotic theories for eukaryote origin. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370,** 20140330 (2015).
5. Zachar, I., Szilágyi, A., Számadó, S. & Szathmáry, E. Farming the mitochondrial ancestor as a model of endosymbiotic establishment by natural selection. *Proceedings of the National Academy of Sciences* **115,** 201718707–E1510 (2018).
6. Lane, N. Bioenergetic Constraints on the Evolution of Complex Life. *Cold Spring Harbor Perspectives in Biology* **6,** a015982–a015982 (2014).
7. Ladoukakis, E. D. & Zouros, E. Evolution and inheritance of animal mitochondrial DNA: rules and exceptions. *Journal of Biological Research* **24,** 1476 (2017).
8. Kelly, R. D. W., Mahmud, A., McKenzie, M., Trounce, I. A. & St John, J. C. Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A. *Nucleic Acids Research* **40,** 10124–10138 (2012).
9. Hällberg, B. M. & Larsson, N.-G. Making Proteins in the Powerhouse. *Cell Metabolism* **20,** 226–240 (2014).
10. 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* **108,** 13534–13539 (2011).
11. Chen, X *et al.* Rearranged mitochondrial genomes are present in human oocytes. *American Journal of Human Genetics* **57,** 239–624 (1995).
12. Falkenberg, M., Larsson, N.-G. & Gustafsson, C. M. DNA Replication and Transcription in Mammalian Mitochondria. *Annual Review of Biochemistry* **76,** 679–699 (2007).
13. Rath, S. *et al.* MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. *Nucleic Acids Research* **49,** gkaa1011– (2020).
14. Harbauer, A. B., Zahedi, R. P., Sickmann, A., Pfanner, N. & Meisinger, C. The Protein Import Machinery of Mitochondria—A Regulatory Hub in Metabolism, Stress, and Disease. *Cell Metabolism* **19,** 357–372 (2014).
15. Chouchani, E. T. *et al.* Ischaemic accumulation of succinate controls reperfusion injury through mitochondrial ROS. *Nature* **515,** 431–435 (2014).
16. Letts, J. A. & Sazanov, L. A. Clarifying the supercomplex: the higher-order organization of the mitochondrial electron transport chain. *Nature Structural & Molecular Biology* **24,** 800–808 (2017).

17. Gorman, G. S. *et al.* Mitochondrial diseases. *Nature Reviews Disease Primers* **2,** 16080 (2016).

18. Gorman, G. S. *et al.* Prevalence of nuclear and mitochondrial DNA mutations related to adult mitochondrial disease. *Annals of neurology* **77,** 753–759 (2015).

19. Skladal, D., Halliday, J. & Thorburn, D. R. Minimum birth prevalence of mitochondrial respiratory chain disorders in children. *Brain* **126,** 1905–1912 (2003).

20. Baertling, F. *et al.* A guide to diagnosis and treatment of Leigh syndrome. *J Neurol Neurosurg Psychiatry* **85,** 257–265 (2014).

21. Smith, E. F., Shaw, P. J. & De Vos, K. J. The role of mitochondria in amyotrophic lateral sclerosis. *Neuroscience Letters* (2017).

22. Winklhofer, K. F. & Haass, C. Mitochondrial dysfunction in Parkinson's disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1802,** 29–44 (2010).

23. Lood, C. *et al.* Neutrophil extracellular traps enriched in oxidized mitochondrial DNA are interferogenic and contribute to lupus-like disease. *Nature Medicine* **22,** 146–153 (2016).

24. Nunnari, J. & Suomalainen, A. Mitochondria: in sickness and in health. *Cell* **148,** 1145–1159 (2012).

25. Yadav, S. P. The wholeness in suffix -omics, -omes, and the word om. *Journal of Biomolecular Techniques* **18,** 277 (2007).

26. Karczewski, K. J. & Snyder, M. P. Integrative omics for health and disease. *Nature Reviews Genetics* **19,** 299–310 (2018).

27. Adams, M. *et al.* Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252,** 1651–1656 (1991).

28. Stark, R., Grzelak, M. & Hadfield, J. RNA sequencing: the teenage years. *Nature Reviews Genetics* **20,** 631–656 (2019).

29. Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456,** 470–476 (2008).

30. Halpern, K. B. *et al.* Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542,** 352–356 (2017).

31. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353,** 78–82 (2016).

32. Kertesz, M. *et al.* Genome-wide measurement of RNA secondary structure in yeast. *Nature* **467,** 103–107 (2010).

33. Stranneheim, H. *et al.* Integration of whole genome sequencing into a healthcare setting: high diagnostic rates across multiple clinical entities in 3219 rare disease patients. *Genome Medicine* **13,** 40 (2021).

34. Kremer, L. S., Wortmann, S. B. & Prokisch, H. "Transcriptomics": molecular diagnosis of inborn errors of metabolism via RNA-sequencing. *Journal of Inherited Metabolic Disease* **41,** 525–532 (2018).

35. Yates, A. D. *et al.* Ensembl 2020. *Nucleic Acids Research* **48,** D682–D688 (2019).

36. Piovesan, A. *et al.* Human protein-coding genes and gene feature statistics in 2019. *BMC Research Notes* **12,** 315 (2019).

37. Consortium, T. U. *et al.* UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research* **49,** D480–D489 (2020).

38. Ponomarenko, E. A. *et al.* The Size of the Human Proteome: The Width and Depth. *International Journal of Analytical Chemistry* **2016,** 1–6 (2016).

39. Smith, L. M. *et al.* Proteoform: a single term describing protein complexity. *Nature Methods* **10,** 186–187 (2013).

40. Aebersold, R. *et al.* How many human proteoforms are there? *Nature Chemical Biology* **14,** 206–214 (2018).

41. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* **347,** 1260419 (2015).

42. Olsen, J. V. & Mann, M. Status of Large-scale Analysis of Post-translational Modifications by Mass Spectrometry. *Molecular & Cellular Proteomics* **12,** 3444–3452 (2013).

43. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10,** 57–63 (2009).

44. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473,** 337–342 (2011).

45. Zubarev, R. A. The challenge of the proteome dynamic range and its implications for in-depth proteomics. *PROTEOMICS* **13,** 723–726 (2013).

46. Brunner, A.-D. *et al.* Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *bioRxiv,* 423933 (2020).

47. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537,** 347–355 (2016).

48. Vandermarliere, E., Mueller, M. & Martens, L. Getting intimate with trypsin, the leading protease in proteomics. *Mass Spectrometry Reviews* **32,** 453–465 (2013).

49. Giansanti, P., Tsiatsiani, L., Low, T. Y. & Heck, A. J. R. Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nature Protocols* **11,** 993–1006 (2016).

50. Bar, N. *et al.* A reference map of potential determinants for the human serum metabolome. *Nature* **588,** 135–140 (2020).

51. Lorendeau, D., Christen, S., Rinaldi, G. & Fendt, S. Metabolic control of signalling pathways and metabolic auto-regulation. *Biology of the Cell* **107,** 251–272 (2015).

52. Chain, E. B. Landmarks and Perspectives in Biochemical Research. *British Medical Journal* **1,** 209 (1965).

53. Zamboni, N., Saghatelian, A. & Patti, G. J. Defining the Metabolome: Size, Flux, and Regulation. *Molecular Cell* **58,** 699–706 (2015).

54. Patti, G. J., Yanes, O. & Siuzdak, G. Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology* **13,** 263–269 (2012).

55. Domingo-Almenara, X. *et al.* The METLIN small molecule dataset for machine learning-based retention time prediction. *Nature Communications* **10,** 5811 (2019).

56. Ong, S.-E. *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics* **1,** 376–386 (2002).

57. Grankvist, N. *et al.* Profiling the Metabolism of Human Cells by Deep 13C Labeling. *Cell Chemical Biology* **25,** 1419–1427.e4 (2018).

58. Fernández-García, J., Altea-Manzano, P., Pranzini, E. & Fendt, S.-M. Stable Isotopes for Tracing Mammalian-Cell Metabolism In Vivo. *Trends in Biochemical Sciences* **45,** 185–201 (2020).

59. Kaur, G. & Dufour, J. M. Cell lines. *Spermatogenesis* **2,** 1–5 (2012).

60. Shoemaker, R. H. The NCI60 human tumour cell line anticancer drug screen. *Nature Reviews Cancer* **6,** 813–823 (2006).

61. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126,** 663–676 (2006).

62. Koch, P., Opitz, T., Steinbeck, J. A., Ladewig, J. & Brüstle, O. A rosette-type, self-renewing human ES cell-derived neural stem cell with potential for in vitro instruction and synaptic integration. *Proceedings of the National Academy of Sciences* **106,** 3225–3230 (2009).

63. Falk, A. *et al.* Capture of Neuroepithelial-Like Stem Cells from Pluripotent Stem Cells Provides a Versatile System for In Vitro Production of Human Neurons. *PLoS ONE* **7,** e29597 (2012).

64. Norat, P. *et al.* Mitochondrial dysfunction in neurological disorders: Exploring mitochondrial transplantation. *npj Regenerative Medicine* **5,** 22 (2020).

65. Magistretti, P. J. & Allaman, I. Lactate in the brain: from metabolic end-product to signalling molecule. *Nature Reviews Neuroscience* **19,** 235–249 (2018).

66. Rossiter, N. J. *et al.* CRISPR screens in physiologic medium reveal conditionally essential genes in human cells. *Cell Metabolism* (2021).

67. Jain, I. H. *et al.* Genetic Screen for Cell Fitness in High or Low Oxygen Highlights Mitochondrial and Lipid Metabolism. *Cell* **181,** 716–727.e11 (2020).

68. Kim, J., Koo, B.-K. & Knoblich, J. A. Human organoids: model systems for human biology and medicine. *Nature Reviews Molecular Cell Biology* **21,** 571–584 (2020).

69. Bellen, H. J., Tong, C. & Tsuda, H. 100 years of Drosophila research and its impact on vertebrate neuroscience: a history lesson for the future. *Nature Reviews Neuroscience* **11,** 514–522 (2010).

70. Hales, K. G., Korey, C. A., Larracuente, A. M. & Roberts, D. M. Genetics on the Fly: A Primer on the Drosophila Model System. *Genetics* **201,** 815–842 (2015).

71. Bratic, A. *et al.* The Bicoid Stability Factor Controls Polyadenylation and Expression of Specific Mitochondrial mRNAs in Drosophila melanogaster. *PLOS Genetics* **7,** e1002324 (2011).

72. Clemente, P. *et al.* SUV3 helicase is required for correct processing of mitochondrial transcripts. *Nucleic Acids Research* **43,** 7398–7413 (2015).

73. St Johnston, D. The art and design of genetic screens: Drosophila melanogaster. *Nature Reviews Genetics* **3,** 176–188 (2002).

74. Adams, M. D. *et al.* The genome sequence of Drosophila melanogaster. *Science* **287,** 2185–2195 (2000).

75. Celniker, S. E. *et al.* Finishing a whole-genome shotgun: release 3 of the Drosophila melanogaster euchromatic genome sequence. *Genome Biology* **3,** research0079.1–14 (2002).

76. Brunner, E. *et al.* A high-quality catalog of the Drosophila melanogaster proteome. *Nat Biotech* **25,** 576–583 (2007).

77. Casas-Vila, N. *et al.* The developmental proteome of Drosophila melanogaster. *Genome Research* **27,** 1273–1285 (2017).

78. Zhai, B., Villén, J., Beausoleil, S. A., Mintseris, J. & Gygi, S. P. Phosphoproteome Analysis of Drosophila melanogaster Embryos. *Journal of Proteome Research* **7,** 1675–1682 (2008).

79. Colinet, H., Pineau, C. & Com, E. Large scale phosphoprotein profiling to explore Drosophila cold acclimation regulatory mechanisms. *Scientific Reports* **7,** 1713 (2017).

80. Hogrebe, A. *et al.* Benchmarking common quantification strategies for large-scale phosphoproteomics. *Nature Communications* **9,** 1045 (2018).

81. Sury, M. D., Chen, J.-x. & Selbach, M. The SILAC fly allows for accurate protein quantification in vivo. *Molecular & Cellular Proteomics* **9,** 2173–2183 (2010).

82. Xu, P. *et al.* Stable Isotope Labeling with Amino Acids in Drosophila for Quantifying Proteins and Modifications. *Journal of Proteome Research* **11,** 4403–4412 (2012).

83. Chang, Y.-C. *et al.* Evaluation of Drosophila Metabolic Labeling Strategies for in Vivo Quantitative Proteomic Analyses with Applications to Early Pupa Formation and Amino Acid Starvation. *Journal of Proteome Research* **12,** 2138–2150 (2013).

84. Chung, V. Y. *et al.* Proteomic changes in response to crystal formation in Drosophila Malpighian tubules. *Fly* **10,** 91–100 (2016).

85. Shlomi, T., Cabili, M. N., Herrgård, M. J., Palsson, B. Ø. & Ruppin, E. Network-based prediction of human tissue-specific metabolism. *Nature Biotechnology* **26,** 1003–1010 (2008).

86. Wishart, D. S. *et al.* HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Research* **46,** D608–D617 (2017).

87. Larsson, N.-G. *et al.* Mitochondrial transcription factor A is necessary for mtDNA maintance and embryogenesis in mice. *Nature Genetics* **18,** 231–236 (1998).

88. Wredenberg, A. *et al.* Increased mitochondrial mass in mitochondrial myopathy mice. *Proceedings of the National Academy of Sciences* **99,** 15066–15071 (2002).

89. Hansson, A. *et al.* A switch in metabolism precedes increased mitochondrial biogenesis in respiratory chain-deficient mouse hearts. *Proceedings of the National Academy of Sciences* **101,** 3136–3141 (2004).

90. Desdín-Micó, G. *et al.* T cells with dysfunctional mitochondria induce multimorbidity and premature senescence. *Science* **368,** 1371–1376 (2020).

91. Kitano, H. Computational systems biology. *Nature* **420,** 206–210 (2002).

92. Misra, B. B., Langefeld, C., Olivier, M. & Cox, L. A. Integrated omics: tools, advances and future approaches. *Journal of Molecular Endocrinology* **62,** R21–R45 (2019).

93. Szklarczyk, D. *et al.* STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research* **47,** gky1131 (2018).

94. Bartha, I., Iulio, J. d., Venter, J. C. & Telenti, A. Human gene essentiality. *Nature Reviews Genetics* **19,** 51–62 (2018).

95. Ghandi, M. *et al.* Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569,** 503–508 (2019).

96. Tsherniak, A. *et al.* Defining a Cancer Dependency Map. *Cell* **170,** 564–576.e16 (2017).

97. Khatri, P., Sirota, M. & Butte, A. J. Ten Years of Pathway Analysis: Current Approaches and Outstanding Challenges. *PLoS Computational Biology* **8,** e1002375 (2012).

98. Boyle, E. I. *et al.* GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* **20,** 3710–3715 (2004).

99. Jantzen, S. G., Sutherland, B. J., Minkley, D. R. & Koop, B. F. GO Trimming: Systematically reducing redundancy in large Gene Ontology datasets. *BMC Research Notes* **4,** 267 (2011).

100. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research* **45,** D353–D361 (2017).

101. 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* **102,** 15545–15550 (2005).

102. Abadi, M. *et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems* Software available from tensorflow.org. 2015. `https://www.tensorflow.org/`.

103. Han, H. & Liu, W. The coming era of artificial intelligence in biological data science. *BMC Bioinformatics* **20,** 712 (2019).

104. Vaudel, M. *et al.* Exploring the potential of public proteomics data. *PROTEOMICS* **16,** 214–225 (2016).

105. Vizcaíno, J. A. *et al.* ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nature Biotechnology* **32,** 223–226 (2014).

106. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3,** 160018 (2016).

107. Perez-Riverol, Y. *et al.* Quantifying the impact of public omics data. *Nature Communications* **10,** 3512 (2019).

108. *PRIDE - Proteomics Identification Database* `https://www.ebi.ac.uk/pride/archive/` (2021).

109. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Research* **47,** gky1106– (2018).

110. Austin, S. & St-Pierre, J. PGC1α and mitochondrial metabolism – emerging concepts and relevance in ageing and neurodegenerative disorders. *Journal of Cell Science* **125,** 4963–4971 (2012).

111. Kühl, I. *et al.* Transcriptomic and proteomic landscape of mitochondrial dysfunction reveals secondary coenzyme Q deficiency in mammals. *eLife* **6,** e30952 (2017).

112. Chen, W. W., Freinkman, E., Wang, T., Birsoy, K. & Sabatini, D. M. Absolute Quantification of Matrix Metabolites Reveals the Dynamics of Mitochondrial Metabolism. *Cell* **166,** 1324–1337.e11 (2016).

113. Yim, A. *et al.* mitoXplorer, a visual data mining platform to systematically analyze and visualize mitochondrial expression dynamics and mutations. *Nucleic Acids Research* **48,** 605–632 (2019).

114. Pagliarini, D. J. *et al.* A Mitochondrial Protein Compendium Elucidates Complex I Disease Biology. *Cell* **134,** 112–123 (2008).

115. Vögtle, F.-N. *et al.* Landscape of submitochondrial protein distribution. *Nature Communications* **8,** 290 (2017).

116. Antonicka, H. *et al.* A High-Density Human Mitochondrial Proximity Interaction Network. *Cell Metabolism* **32,** 479–497.e9 (2020).

117. Singh, A. P. *et al.* Molecular Connectivity of Mitochondrial Gene Expression and OXPHOS Biogenesis. *Molecular Cell* **79,** 1051–1065.e10 (2020).

118. Prus, G., Hoegl, A., Weinert, B. T. & Choudhary, C. Analysis and Interpretation of Protein Post-Translational Modification Site Stoichiometry. *Trends in Biochemical Sciences* **44,** 943–960 (2019).

119. Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Research* **43,** D512–D520 (2015).

120. Stram, A. R. & Payne, R. M. Post-translational modifications in mitochondria: protein signaling in the powerhouse. *Cellular and Molecular Life Sciences* **73,** 4063–4073 (2016).

121. Baeza, J., Smallegan, M. J. & Denu, J. M. Mechanisms and Dynamics of Protein Acetylation in Mitochondria. *Trends in Biochemical Sciences* **41,** 231–244 (2016).

122. Jung, S. Y. *et al.* Complications in the Assignment of 14 and 28 Da Mass Shift Detected by Mass Spectrometry as in Vivo Methylation from Endogenous Proteins. *Analytical Chemistry* **80,** 1721–1729 (2008).

123. Perdivara, I., Deterding, L. J., Przybylski, M. & Tomer, K. B. Mass spectrometric identification of oxidative modifications of tryptophan residues in proteins: Chemical artifact or post-translational modification? *Journal of the American Society for Mass Spectrometry* **21,** 1114–1117 (2010).

124. Gharib, M. *et al.* Artifactual Sulfation of Silver-stained Proteins Implications for the Assignment of Phosphorylation and Sulfation Sites. *Molecular & Cellular Proteomics* **8,** 506–518 (2009).

125. Nielsen, M. L. *et al.* Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. *Nature Methods* **5,** 459–460 (2008).

126. Murray, C. I. & Eyk, J. E. V. A Twist on Quantification. *Circulation Research* **111,** 1253–1255 (2012).

127. Harmel, R. & Fiedler, D. Features and regulation of non-enzymatic post-translational modifications. *Nature Chemical Biology* **14,** 244–252 (2018).

128. Saiardi, A., Bhandari, R., Resnick, A. C., Snowman, A. M. & Snyder, S. H. Phosphorylation of Proteins by Inositol Pyrophosphates. *Science* **306,** 2101–2105 (2004).

129. Dephoure, N., Gould, K. L., Gygi, S. P. & Kellogg, D. R. Mapping and analysis of phosphorylation sites: a quick guide for cell biologists. *Molecular Biology of the Cell* **24,** 535–542 (2013).

130. Tyrakis, P. A. *et al.* Fumarate Hydratase Loss Causes Combined Respiratory Chain Defects. *Cell Reports* **21,** 1036–1047 (2017).

131. Lombard, D. B. *et al.* Mammalian Sir2 Homolog SIRT3 Regulates Global Mitochondrial Lysine Acetylation. *Molecular and Cellular Biology* **27,** 8807–8814 (2007).

132. Frankovsky, J., Vozáriková, V., Nosek, J. & Tomáška, L. Mitochondrial protein phosphorylation in yeast revisited. *Mitochondrion* **57,** 148–162 (2021).

133. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Research* **21,** 381–395 (2011).

134. Rhein, V. F., Carroll, J., Ding, S., Fearnley, I. M. & Walker, J. E. NDUFAF7 Methylates Arginine 85 in the NDUFS2 Subunit of Human Complex I. *Journal of Biological Chemistry* **288,** 33016–33026 (2013).

135. Tort, F. *et al.* Mutations in the lipoyltransferase LIPT1 gene cause a fatal disease associated with a specific lipoylation defect of the 2-ketoacid dehydrogenase complexes. *Human Molecular Genetics* **23,** 1907–1915 (2014).

136. Sha, L. *et al.* Asymmetric Arginine Dimethylation Modulates Mitochondrial Energy Metabolism and Homeostasis in Caenorhabditis elegans. *Molecular and Cellular Biology* **37,** e00504–16 (2017).

137. Huttlin, E. L. *et al.* A Tissue-Specific Atlas of Mouse Protein Phosphorylation and Expression. *Cell* **143,** 1174–1189 (2010).

138. Olsen, J. V. *et al.* Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Sci. Signal.* **3,** ra3 (2010).

139. Yachie, N., Saito, R., Sugiyama, N., Tomita, M. & Ishihama, Y. Integrative Features of the Yeast Phosphoproteome and Protein–Protein Interaction Map. *PLoS Computational Biology* **7,** e1001064 (2011).

140. Lundby, A. *et al.* Proteomic Analysis of Lysine Acetylation Sites in Rat Tissues Reveals Organ Specificity and Subcellular Patterns. *Cell Reports* **2,** 419–431 (2012).

141. Mertins, P. *et al.* Integrated proteomic analysis of post-translational modifications by serial enrichment. *Nature Methods* **10,** 634–637 (2013).

142. Krebs, E. G. & Fischer, E. H. The phosphorylase b to a converting enzyme of rabbit skeletal muscle. *Biochimica et Biophysica Acta* **20,** 150–157 (1956).

143. Olsen, J. V. *et al.* Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* **127,** 635–648 (2006).

144. Fong, J. H. & Panchenko, A. R. Intrinsic disorder and protein multibinding in domain, terminal, and linker regions. *Molecular BioSystems* **6,** 1821–1828 (2010).

145. Nishi, H., Shaytan, A. & Panchenko, A. R. Physicochemical mechanisms of protein regulation by phosphorylation. *Frontiers in Genetics* **5,** 270 (2014).

146. Nishi, H., Hashimoto, K. & Panchenko, A. R. Phosphorylation in Protein-Protein Binding: Effect on Stability and Function. *Structure* **19,** 1807–1815 (2011).

147. Johnson, L. N. Glycogen phosphorylase: control by phosphorylation and allosteric effectors. *The FASEB Journal* **6,** 2274–2282 (1992).

148. Radivojac, P. *et al.* Gain and loss of phosphorylation sites in human cancer. *Bioinformatics* **24,** i241–i247 (2008).

149. Canerina-Amaro, A. *et al.* Differential Aggregation and Phosphorylation of Alpha Synuclein in Membrane Compartments Associated With Parkinson Disease. *Frontiers in Neuroscience* **13,** 382 (2019).

150. Müller, S., Chaikuad, A., Gray, N. S. & Knapp, S. The ins and outs of selective kinase inhibitor development. *Nature Chemical Biology* **11,** 818–821 (2015).

151. Li, X.-S., Yuan, B.-F. & Feng, Y.-Q. Recent advances in phosphopeptide enrichment: Strategies and techniques. *Trends in Analytical Chemistry* **78,** 70–83 (2016).

152. Steen, H., Jebanathirajah, J. A., Rush, J., Morrice, N. & Kirschner, M. W. Phosphorylation Analysis by Mass Spectrometry Myths, Facts, and the Consequences for Qualitative and Quantitative Measurements*. *Molecular & Cellular Proteomics* **5,** 172–181 (2006).

153. Yang, C., Zhong, X. & Li, L. Recent advances in enrichment and separation strategies for mass spectrometry-based phosphoproteomics. *ELECTROPHORESIS* **35,** 3418–3429 (2014).

154. Humphrey, S. J., Azimifar, S. B. & Mann, M. High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. *Nature Biotechnology* **33,** 990–995 (2015).

155. Steen, H., Jebanathirajah, J. A., Springer, M. & Kirschner, M. W. Stable isotope-free relative and absolute quantitation of protein phosphorylation stoichiometry by MS. *Proceedings of the National Academy of Sciences* **102,** 3948–3953 (2005).

156. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11,** 2301–2319 (2016).

157. *FPD* `http://bis.zju.edu.cn/FPD/?mod=summary` (2021).

158. Renvoisé, M. *et al.* Quantitative variations of the mitochondrial proteome and phosphoproteome during fermentative and respiratory growth in Saccharomyces cerevisiae. *Journal of Proteomics* **106,** 140–150 (2014).

159. Kolitsida, P. *et al.* Phosphorylation of mitochondrial matrix proteins regulates their selective mitophagic degradation. *Proceedings of the National Academy of Sciences* **116,** 20517–20527 (2019).

160. Kotrasová, V. *et al.* Mitochondrial Kinases and the Role of Mitochondrial Protein Phosphorylation in Health and Disease. *Life* **11,** 82 (2021).

161. Lim, S. *et al.* Regulation of mitochondrial functions by protein phosphorylation and dephosphorylation. *Cell & Bioscience* **6,** 25 (2016).

162. Amer, Y. O. & Hebert-Chatelain, E. Mitochondrial cAMP-PKA signaling: What do we really know? *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1859,** 868–877 (2018).

163. Harada, H. *et al.* Phosphorylation and Inactivation of BAD by Mitochondria-Anchored Protein Kinase A. *Molecular Cell* **3,** 413–422 (1999).

164. Schwoch, G, Trinczek, B & Bode, C. Localization of catalytic and regulatory subunits of cyclic AMP-dependent protein kinases in mitochondria from various rat tissues. *Biochemical Journal* **270,** 181–188 (1990).

165. Technikova-Dobrova, Z., Sardanelli, A. M., Stanca, M. R. & Papa, S. cAMP-dependent protein phosphorylation in mitochondria of bovine heart. *FEBS Letters* **350,** 187–191 (1994).

166. Hebert-Chatelain, E. *et al.* A cannabinoid link between mitochondria and memory. *Nature* **539,** 555–559 (2016).

167. García-Bermúdez, J. *et al.* PKA Phosphorylates the ATPase Inhibitory Factor 1 and Inactivates Its Capacity to Bind and Inhibit the Mitochondrial H+-ATP Synthase. *Cell Reports* **12,** 2143–2155 (2015).

168. Grimsrud, P. A. *et al.* A Quantitative Map of the Liver Mitochondrial Phosphoproteome Reveals Post-translational Control of Ketogenesis. *Cell Metabolism* **16,** 672–683 (2012).

169. Krahmer, N. *et al.* Organellar Proteomics and Phospho-Proteomics Reveal Subcellular Reorganization in Diet-Induced Hepatic Steatosis. *Developmental Cell* **47,** 205–221.e7 (2018).

170. Humphrey, S. J. *et al.* Dynamic Adipocyte Phosphoproteome Reveals that Akt Directly Regulates mTORC2. *Cell Metabolism* **17,** 1009–1020 (2013).

171. Clarke, S. G. Protein methylation at the surface and buried deep: thinking outside the histone box. *Trends in Biochemical Sciences* **38,** 243–252 (2013).

172. Fontecave, M., Atta, M. & Mulliez, E. S-adenosylmethionine: nothing goes to waste. *Trends in Biochemical Sciences* **29,** 243–249 (2004).

173. Luo, M. Chemical and Biochemical Perspectives of Protein Lysine Methylation. *Chemical Reviews* **118,** 6656–6705 (2018).

174. Evich, M., Stroeva, E., Zheng, Y. G. & Germann, M. W. Effect of methylation on the side-chain pKa value of arginine. *Protein science* **25,** 479–486 (2016).

175. Guccione, E. & Richard, S. The regulation, functions and clinical relevance of arginine methylation. *Nature Reviews Molecular Cell Biology* **20,** 642–657 (2019).

176. Guo, A. *et al.* Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Molecular & Cellular Proteomics* **13,** 372–387 (2014).

177. Afjehi-Sadat, L. & Garcia, B. A. Comprehending dynamic protein methylation with mass spectrometry. *Current Opinion in Chemical Biology* **17,** 12–19 (2013).

178. Sylvestersen, K. B., Horn, H., Jungmichel, S., Jensen, L. J. & Nielsen, M. L. Proteomic analysis of arginine methylation sites in human cells reveals dynamic regulation during transcriptional arrest. *Molecular & Cellular Proteomics* **13,** 2072–2088 (2014).

179. Geoghegan, V., Guo, A., Trudgian, D., Thomas, B. & Acuto, O. Comprehensive identification of arginine methylation in primary T cells reveals regulatory roles in cell signalling. *Nature Communications* **6,** 6758 (2015).

180. Larsen, S. C. *et al.* Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells. *Sci. Signal.* **9,** rs9 (2016).

181. Hart-Smith, G., Yagoub, D., Tay, A. P., Pickford, R. & Wilkins, M. R. Large Scale Mass Spectrometry-based Identifications of Enzyme-mediated Protein Methylation Are Subject to High False Discovery Rates. *Molecular & Cellular Proteomics* **15,** 989–1006 (2016).

182. Moore, K. E. *et al.* A general molecular affinity strategy for global detection and proteomic analysis of lysine methylation. *Molecular Cell* **50,** 444–456 (2013).

183. Ong, S.-E., Mittler, G. & Mann, M. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nature Methods* **1,** 119–126 (2004).

184. Tay, A. P., Geoghegan, V., Yagoub, D., Wilkins, M. R. & Hart-Smith, G. MethylQuant: A Tool for Sensitive Validation of Enzyme-Mediated Protein Methylation Sites from Heavy-Methyl SILAC Data. *Journal of Proteome Research* **17,** 359–373 (2017).

185. Linscott, J. A. *et al.* Kinetic isotope effects reveal early transition state of protein lysine methyltransferase SET8. *Proceedings of the National Academy of Sciences* **113,** E8369–E8378 (2016).

186. Bothwell, I. R. *et al.* Se -Adenosyl- l -selenomethionine Cofactor Analogue as a Reporter of Protein Methylation. *Journal of the American Chemical Society* **134,** 14905–14912 (2012).

187. Blum, G., Bothwell, I. R., Islam, K. & Luo, M. Profiling Protein Methylation with Cofactor Analog Containing Terminal Alkyne Functionality. *Current Protocols in Chemical Biology* **5,** 67–88 (2013).

188. Davydova, E. *et al.* The methyltransferase METTL9 mediates pervasive 1-methylhistidine modification in mammalian proteomes. *Nature Communications* **12,** 891 (2021).

189. Deen, J. *et al.* Methyltransferase-Directed Labeling of Biomolecules and its Applications. *Angewandte Chemie International Edition* **56,** 5182–5200 (2017).

190. Bao, X. R. *et al.* Mitochondrial dysfunction remodels one-carbon metabolism in human cells. *eLife* **5,** 1910 (2016).

191. Ducker, G. S. & Rabinowitz, J. D. One-Carbon Metabolism in Health and Disease. *Cell Metabolism* **25,** 27–42 (2017).

192. Wright, A. J. A., Dainty, J. R. & Finglas, P. M. Folic acid metabolism in human subjects revisited: potential implications for proposed mandatory folic acid fortification in the UK. *British Journal of Nutrition* **98,** 667–675 (2007).

193. Choi, J.-H., Yates, Z., Veysey, M., Heo, Y.-R. & Lucock, M. Contemporary Issues Surrounding Folic Acid Fortification Initiatives. *Preventive Nutrition and Food Science* **19,** 247–260 (2014).

194. Smithells, R. W. *et al.* Possible prevention of neural-tube defects by periconceptional vitamin supplementation. *The Lancet* **1,** 339–340 (1980).

195. French, J. B. *et al.* Spatial colocalization and functional link of purinosomes with mitochondria. *Science* **351,** 733–737 (2016).

196. Yang, M. & Vousden, K. H. Serine and one-carbon metabolism in cancer. *Nature Reviews Cancer* **16,** 650 (2016).

197. Moriya, J. *et al.* A Novel Modified Nucleoside Found at the First Position of the Anticodon of Methionine tRNA from Bovine Liver Mitochondria. *Biochemistry* **33,** 2234–2239 (2002).

198. López, L. C. *et al.* Unbalanced deoxynucleotide pools cause mitochondrial DNA instability in thymidine phosphorylase-deficient mice. *Human Molecular Genetics* **18,** 714–722 (2008).

199. Anderson, D. D., Quintero, C. M. & Stover, P. J. Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria. *Proceedings of the National Academy of Sciences* **108,** 15163–15168 (2011).

200. Davis, S. R. *et al.* Tracer-derived total and folate-dependent homocysteine remethylation and synthesis rates in humans indicate that serine is the main one-carbon donor. *American Journal of Physiology-Endocrinology and Metabolism* **286,** E272–E279 (2004).

201. Locasale, J. W. Serine, glycine and one-carbon units: cancer metabolism in full circle. *Nature Reviews Cancer* **13,** 572–583 (2013).

202. Ducker, G. S. *et al.* Reversal of Cytosolic One-Carbon Flux Compensates for Loss of the Mitochondrial Folate Pathway. *Cell Metabolism* **23,** 1140–1153 (2016).

203. Nilsson, R. *et al.* Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. *Nature Communications* **5,** 1131 (2014).

204. Tibbetts, A. S. & Appling, D. R. Compartmentalization of Mammalian Folate-Mediated One-Carbon Metabolism. *Annual Review of Nutrition* **30,** 57–81 (2010).

205. Pajares, M. A. & Perez-Sala, D. Betaine homocysteine S-methyltransferase: just a regulator of homocysteine metabolism? *Cellular and Molecular Life Sciences* **63,** 2792–2803 (2006).

206. Anderson, J. L. *et al.* Plasma homocysteine predicts mortality independently of traditional risk factors and C-reactive protein in patients with angiographically defined coronary artery disease. *Circulation* **102,** 1227–1232 (2000).

207. Yoo, J. H. & Lee, S. C. Elevated levels of plasma homocyst(e)ine and asymmetric dimethylarginine in elderly patients with stroke. *Atherosclerosis* **158,** 425–430 (2001).

208. Boldyrev, A., Bryushkova, E., Mashkina, A. & Vladychenskaya, E. Why is homocysteine toxic for the nervous and immune systems? *Current Aging Science* **6,** 29–36 (2013).

209. Finkelstein, J. D. Methionine metabolism in mammals. *Journal of Nutritional Biochemistry* **1,** 228–237 (1990).

210. Pajares, M. A., Alvarez, L. & Pérez-Sala, D. How are mammalian methionine adenosyltransferases regulated in the liver? A focus on redox stress. *FEBS Letters* **587,** 1711–1716 (2013).

211. Markham, G. D. & Pajares, M. A. Structure-function relationships in methionine adenosyltransferases. *Cellular and Molecular Life Sciences* **66,** 636–648 (2008).

212. Kotb, M & Kredich, N. M. S-Adenosylmethionine synthetase from human lymphocytes. Purification and characterization. *Journal of Biological Chemistry* **260,** 3923–3930 (1985).

213. Cabrero, C, Duce, A. M., Ortiz, P, Alemany, S & Mato, J. M. Specific loss of the high-molecular-weight form of S-adenosyl-L-methionine synthetase in human liver cirrhosis. *Hepatology* **8,** 1530–1534 (1988).

214. González, B. *et al.* NADP+ Binding to the Regulatory Subunit of Methionine Adenosyltransferase II Increases Intersubunit Binding Affinity in the Hetero-Trimer. *PLOS ONE* **7,** e50329 (2012).

215. Pajares, M. A., Durán, C., Corrales, F., Pliego, M. M. & Mato, J. M. Modulation of rat liver S-adenosylmethionine synthetase activity by glutathione. *Journal of Biological Chemistry* **267,** 17587–17605 (1992).

216. Su, X., Wellen, K. E. & Rabinowitz, J. D. Metabolic control of methylation and acetylation. *Current Opinion in Chemical Biology* **30,** 52–60 (2016).

217. Petrossian, T. C. & Clarke, S. G. Uncovering the Human Methyltransferasome. *Molecular & Cellular Proteomics* **10,** M110.000976 (2010).

218. Horning, B. D. *et al.* Chemical Proteomic Profiling of Human Methyltransferases. *Journal of the American Chemical Society* **138,** 13335–13343 (2016).

219. Obeid, R. The Metabolic Burden of Methyl Donor Deficiency with Focus on the Betaine Homocysteine Methyltransferase Pathway. *Nutrients* **5,** 3481–3495 (2013).

220. Miller-Fleming, L., Olin-Sandoval, V., Campbell, K. & Ralser, M. Remaining Mysteries of Molecular Biology: The Role of Polyamines in the Cell. *Journal of Molecular Biology* **427,** 3389–3406 (2015).

221. Bauerle, M. R., Schwalm, E. L. & Booker, S. J. Mechanistic Diversity of Radical S-Adenosylmethionine (SAM)-dependent Methylation. *Journal of Biological Chemistry* **290,** 3995–4002 (2015).

222. Holliday, R & Pugh, J. E. DNA modification mechanisms and gene activity during development. *Science* **187,** 226–232 (1975).

223. Riggs, A. D. X inactivation, differentiation, and DNA methylation. *Cytogenetics and cell genetics* **14,** 9–25 (1975).

224. Jang, H. S., Shin, W. J., Lee, J. E. & Do, J. T. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. *Genes* **8,** 148 (2017).

225. Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome - biological and translational implications. *Nature Reviews Cancer* **11,** 726–734 (2011).

226. Luo, J., Li, Y.-N., Wang, F., Zhang, W.-M. & Geng, X. S-adenosylmethionine inhibits the growth of cancer cells by reversing the hypomethylation status of c-myc and H-ras in human gastric cancer and colon cancer. *International Journal of Biological Sciences* **6,** 784–795 (2010).

227. Wang, Y., Sun, Z. & Szyf, M. S-adenosyl-methionine (SAM) alters the transcriptome and methylome and specifically blocks growth and invasiveness of liver cancer cells. *Oncotarget* **8,** 111866–111881 (2017).

228. Patananan, A. N., Sercel, A. J. & Teitell, M. A. More than a powerplant: the influence of mitochondrial transfer on the epigenome. *Current Opinion in Physiology* **3,** 16–24 (2018).

229. Fu, Y., Dominissini, D., Rechavi, G. & He, C. Gene expression regulation mediated through reversible m^6A RNA methylation. *Nature Reviews Genetics* **15,** 293–306 (2014).

230. Mauer, J. *et al.* Reversible methylation of m6A in the 5' cap controls mRNA stability. *Nature* **541,** 371–375 (2017).

231. Sergiev, P. V., Aleksashin, N. A., Chugunova, A. A., Polikanov, Y. S. & Dontsova, O. A. Structural and evolutionary insights into ribosomal RNA methylation. *Nature chemical biology* **14,** 226–235 (2018).

232. Stead, L. M., Brosnan, J. T., Brosnan, M. E., Vance, D. E. & Jacobs, R. L. Is it time to reevaluate methyl balance in humans? *American Journal of Clinical Nutrition* **83,** 5–10 (2006).

233. Noga, A. A. *et al.* Plasma Homocysteine Is Regulated by Phospholipid Methylation. *Journal of Biological Chemistry* **278,** 5952–5955 (2003).

234. Murray, B. *et al.* Methionine Adenosyltransferase α1 Is Targeted to the Mitochondrial Matrix and Interacts with Cytochrome P450 2E1 to Lower Its Expression. *Hepatology* **70,** 2018–2034 (2019).

235. Reytor, E., Pérez-Miguelsanz, J., Alvarez, L., Pérez-Sala, D. & Pajares, M. A. Conformational signals in the C-terminal domain of methionine adenosyltransferase I/III determine its nucleocytoplasmic distribution. *The FASEB Journal* **23,** 3347–3360 (2009).

236. Wohlschlager, T. *et al.* Methylated glycans as conserved targets of animal and fungal innate defense. *Proceedings of the National Academy of Sciences* **111,** E2787–E2796 (2014).

237. Marobbio, C. M. T., Agrimi, G, Lasorsa, F. M. & Palmieri, F. Identification and functional reconstitution of yeast mitochondrial carrier for S-adenosylmethionine. *The EMBO journal* **22,** 5975–5982 (2003).

238. Farooqui, J. Z., Lee, H. W., Kim, S. & Paik, W. K. Studies on compartmentation of S-adenosyl-L-methionine in Saccharomyces cerevisiae and isolated rat hepatocytes. *Biochimica et Biophysica Acta* **757,** 342–351 (1983).

239. Palmieri, F. Mitochondrial transporters of the SLC25 family and associated diseases: a review. *Journal of Inherited Metabolic Disease* **37,** 565–575 (2014).

240. Kishita, Y. *et al.* Intra-mitochondrial Methylation Deficiency Due to Mutations in SLC25A26. *American Journal of Human Genetics* **97,** 761–768 (2015).

241. Małecki, J. *et al.* Uncovering human METTL12 as a mitochondrial methyltransferase that modulates citrate synthase activity through metabolite-sensitive lysine methylation. *Journal of Biological Chemistry,* jbc.M117.808451 (2017).

242. Rhein, V. F. *et al.* Human METTL20 methylates lysine residues adjacent to the recognition loop of the electron transfer flavoprotein in mitochondria. *Journal of Biological Chemistry* **289,** 24640–24651 (2014).

243. Małecki, J., Ho, A. Y. Y., Moen, A., Dahl, H.-A. & Falnes, P. Ø. Human METTL20 Is a Mitochondrial Lysine Methyltransferase That Targets the β Subunit of Electron Transfer Flavoprotein (ETFβ) and Modulates Its Activity. *Journal of Biological Chemistry* **290,** 423–434 (2015).

244. Małecki, J. M. *et al.* Human FAM173A is a mitochondrial lysine-specific methyltransferase that targets adenine nucleotide translocase and affects mitochondrial respiration. *Journal of Biological Chemistry* **294,** 11654–11664 (2019).

245. Małecki, J. M. *et al.* Lysine methylation by the mitochondrial methyltransferase FAM173B optimizes the function of mitochondrial ATP synthase. *Journal of Biological Chemistry* **294,** 1128–1141 (2019).

246. Ishizawa, T., Nozaki, Y., Ueda, T. & Takeuchi, N. The human mitochondrial translation release factor HMRF1L is methylated in the GGQ motif by the methyltransferase HMPrmC. *Biochemical and Biophysical Research Communications* **373,** 99–103 (2008).

247. Suzuki, T. The expanding world of tRNA modifications and their disease relevance. *Nature Reviews Molecular Cell Biology,* 1–18 (2021).

248. Powell, C. A. & Minczuk, M. TRMT2B is responsible for both tRNA and rRNA m5U-methylation in human mitochondria. *RNA Biology* **17,** 451–462 (2020).

249. Laptev, I. *et al.* Mouse Trmt2B protein is a dual specific mitochondrial metyltransferase responsible for m5U formation in both tRNA and rRNA. *RNA Biology* **17,** 1–10 (2019).

250. Haute, L. V. *et al.* METTL15 introduces N4-methylcytidine into human mitochondrial 12S rRNA and is required for mitoribosome biogenesis. *Nucleic Acids Research* **47,** 10267–10281 (2019).

251. Chen, H. *et al.* The human mitochondrial 12S rRNA m4C methyltransferase METTL15 is required for mitochondrial function. *Journal of Biological Chemistry* **295,** 8505–8513 (2020).

252. Metodiev, M. D. *et al.* NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly. *PLOS Genetics* **10,** e1004110 (2014).

253. Helser, T. L., Davies, J. E. & Dahlberg, J. E. Change in methylation of 16S ribosomal RNA associated with mutation to kasugamycin resistance in Escherichia coli. *Nature New Biology* **233,** 12–14 (1971).

254. Metodiev, M. D. *et al.* Methylation of 12S rRNA is necessary for in vivo stability of the small subunit of the mammalian mitochondrial ribosome. *Cell Metabolism* **9,** 386–397 (2009).

255. Bar-Yaacov, D. *et al.* Mitochondrial 16S rRNA Is Methylated by tRNA Methyltransferase TRMT61B in All Vertebrates. *PLOS Biology* **14,** e1002557 (2016).

256. Lee, K.-W., Okot-Kotber, C., LaComb, J. F. & Bogenhagen, D. F. Mitochondrial Ribosomal RNA (rRNA) Methyltransferase Family Members Are Positioned to Modify Nascent rRNA in Foci near the Mitochondrial DNA Nucleoid. *Journal of Biological Chemistry* **288,** 31386–31399 (2013).

257. Lee, K.-W. & Bogenhagen, D. F. Assignment of 2′-O-Methyltransferases to Modification Sites on the Mammalian Mitochondrial Large Subunit 16 S Ribosomal RNA (rRNA). *Journal of Biological Chemistry* **289,** 24936–24942 (2014).

258. Rorbach, J. *et al.* MRM2 and MRM3 are involved in biogenesis of the large subunit of the mitochondrial ribosome. *Molecular Biology of the Cell* **25,** 2542–2555 (2014).

259. Safra, M. *et al.* The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. *Nature,* 1–20 (2017).

260. Booker, S. J., Cicchillo, R. M. & Grove, T. L. Self-sacrifice in radical S-adenosylmethionine proteins. *Current Opinion in Chemical Biology* **11,** 543–552 (2007).

261. Fujiwara, K, Okamura-Ikeda, K & Motokawa, Y. Chicken liver H-protein, a component of the glycine cleavage system. Amino acid sequence and identification of the N epsilon-lipoyllysine residue. *Journal of Biological Chemistry* **261,** 8836–41 (1986).

262. Poon, W. W. *et al.* Yeast and rat Coq3 and Escherichia coli UbiG polypeptides catalyze both O-methyltransferase steps in coenzyme Q biosynthesis. *Journal of Biological Chemistry* **274,** 21665–21672 (1999).

263. Barkovich, R. J. *et al.* Characterization of the COQ5 gene from Saccharomyces cerevisiae. Evidence for a C-methyltransferase in ubiquinone biosynthesis. *Journal of Biological Chemistry* **272,** 9182–9188 (1997).

264. Calvo, S. E., Clauser, K. R. & Mootha, V. K. MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Research* **44,** 1–7 (2015).

265. Hori, H. Methylated nucleosides in tRNA and tRNA methyltransferases. *Frontiers in Genetics* **5,** 2593 (2014).

266. Sanchez, M. I. G. L., Cipullo, M., Gopalakrishna, S., Khawaja, A. & Rorbach, J. Methylation of Ribosomal RNA: A Mitochondrial Perspective. *Frontiers in Genetics* **11,** 761 (2020).

267. Laredj, L. N., Licitra, F. & Puccio, H. M. The molecular genetics of coenzyme Q biosynthesis in health and disease. *Biochimie* **100,** 78–87 (2014).

268. Cicchillo, R. M. & Booker, S. J. Mechanistic Investigations of Lipoic Acid Biosynthesis in Escherichia coli: Both Sulfur Atoms in Lipoic Acid are Contributed by the Same Lipoyl Synthase Polypeptide. *Journal of the American Chemical Society* **127,** 2860–2861 (2005).

269. Kikuchi, G., Motokawa, Y., Yoshida, T. & Hiraga, K. Glycine cleavage system: reaction mechanism, physiological significance, and hyperglycinemia. *Proceedings of the Japan Academy, Series B* **84,** 246–263 (2008).

270. Leung, K.-Y., Castro, S. C. P. D., Galea, G. L., Copp, A. J. & Greene, N. D. E. Glycine Cleavage System H Protein Is Essential for Embryonic Viability, Implying Additional Function Beyond the Glycine Cleavage System. *Frontiers in Genetics* **12,** 625120 (2021).

271. Solmonson, A. & DeBerardinis, R. J. Lipoic acid metabolism and mitochondrial redox regulation. *Journal of Biological Chemistry* **293,** 7522–7530 (2018).

272. Fisher, C. W. *et al.* A 17-bp insertion and a Phe215 → Cys missense mutation in the dihydrolipoyl transacylase (E2) mRNA from a thiamine-responsive maple syrup urine disease patient WG-34. *Biochemical and Biophysical Research Communications* **174,** 804–809 (1991).

273. Bacalini, M. G. *et al.* The methylation of nuclear and mitochondrial DNA in ageing phenotypes and longevity. *Mechanisms of Ageing and Development* **165,** 156–161 (2017).

274. Roberts, S. B. & Gavery, M. R. Is There a Relationship between DNA Methylation and Phenotypic Plasticity in Invertebrates? *Frontiers in Physiology* **2** (2012).

275. Infantino, V. *et al.* Impairment of methyl cycle affects mitochondrial methyl availability and glutathione level in Down's syndrome. *Molecular Genetics and Metabolism* **102,** 378–382 (2011).

276. Byun, H.-M. *et al.* Effects of airborne pollutants on mitochondrial DNA methylation. *Particle and Fibre Toxicology* **10,** 18 (2013).

277. Menga, A. *et al.* SLC25A26 overexpression impairs cell function via mtDNA hypermethylation and rewiring of methyl metabolism. *The FEBS Journal* **284,** 967–984 (2017).

278. Shock, L. S., Thakkar, P. V., Peterson, E. J., Moran, R. G. & Taylor, S. M. DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. *Proceedings of the National Academy of Sciences* **108,** 3630–3635 (2011).

279. Saini, S. K., Mangalhara, K. C., Prakasam, G. & Bamezai, R. N. K. DNA Methyltransferase1 (DNMT1) Isoform3 methylates mitochondrial genome and modulates its biology. *Scientific Reports* **1,** 1–10 (2017).

280. Olova, N. *et al.* Comparison of whole-genome bisulfite sequencing library preparation strategies identifies sources of biases affecting DNA methylation data. *Genome Biology* **19,** 1827 (2018).

281. Matsuda, S. *et al.* Accurate estimation of 5-methylcytosine in mammalian mitochondrial DNA. *Scientific Reports* **8,** 5801 (2018).

282. Mechta, M., Ingerslev, L. R., Fabre, O., Picard, M. & Barrès, R. Evidence Suggesting Absence of Mitochondrial DNA Methylation. *Frontiers in Genetics* **8,** 1 (2017).

283. Harvey, A. J. Mitochondria in early development: linking the microenvironment, metabolism and the epigenome. *Reproduction* **157,** R159–R179 (2019).

284. Lees, J. G., Gardner, D. K. & Harvey, A. J. Pluripotent Stem Cell Metabolism and Mitochondria: Beyond ATP. *Stem Cells International* **2017,** 1–17 (2017).

285. Mandal, S., Lindgren, A. G., Srivastava, A. S., Clark, A. T. & Banerjee, U. Mitochondrial Function Controls Proliferation and Early Differentiation Potential of Embryonic Stem Cells. *STEM CELLS* **29,** 486–495 (2011).

286. Pereira, S. L. *et al.* Inhibition of Mitochondrial Complex III Blocks Neuronal Differentiation and Maintains Embryonic Stem Cell Pluripotency. *PLoS ONE* **8,** e82095 (2013).

287. Chung, S. *et al.* Mitochondrial oxidative metabolism is required for the cardiac differentiation of stem cells. *Nature Clinical Practice Cardiovascular Medicine* **4,** S60–S67 (2007).

288. Paridaen, J. T. & Huttner, W. B. Neurogenesis during development of the vertebrate central nervous system. *EMBO reports* **15,** 351–364 (2014).

289. Lee, S. *et al.* Mitofusin 2 is necessary for striatal axonal projections of midbrain dopamine neurons. *Human Molecular Genetics* **21,** 4827–4835 (2012).

290. Mourier, A. *et al.* Mitofusin 2 is required to maintain mitochondrial coenzyme Q levelsMFN2 maintains mitochondrial coenzyme Q levels. *The Journal of Cell Biology* **208,** 429–442 (2015).

291. Maffezzini, C., Calvo-Garrido, J., Wredenberg, A. & Freyer, C. Metabolic regulation of neurodifferentiation in the adult brain. *Cellular and Molecular Life Sciences* **77,** 2483–2496 (2020).

292. Goyal, M. S., Hawrylycz, M., Miller, J. A., Snyder, A. Z. & Raichle, M. E. Aerobic Glycolysis in the Human Brain Is Associated with Development and Neotenous Gene Expression. *Cell Metabolism* **19,** 49–57 (2014).

293. Telley, L. *et al.* Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex. *Science* **351,** 1443–1446 (2016).

294. Shin, J. *et al.* Single-Cell RNA-Seq with Waterfall Reveals Molecular Cascades underlying Adult Neurogenesis. *Cell Stem Cell* **17,** 360–372 (2015).

295. Beckervordersandforth, R. *et al.* Role of Mitochondrial Metabolism in the Control of Early Lineage Progression and Aging Phenotypes in Adult Hippocampal Neurogenesis. *Neuron* **93,** 560–573.e6 (2017).

296. Zheng, X. *et al.* Metabolic reprogramming during neuronal differentiation from aerobic glycolysis to neuronal oxidative phosphorylation. *eLife* **5,** e13374 (2016).

297. Bond, A. M., Ming, G.-l. & Song, H. Adult Mammalian Neural Stem Cells and Neurogenesis: Five Decades Later. *Cell Stem Cell* **17,** 385–395 (2015).

298. Schafer, S. T. *et al.* Pathological priming causes developmental gene network heterochronicity in autistic subject-derived neurons. *Nature Neuroscience* **22,** 243–255 (2019).

299. Hattiangady, B., Rao, M. S. & Shetty, A. K. Chronic temporal lobe epilepsy is associated with severely declined dentate neurogenesis in the adult hippocampus. *Neurobiology of Disease* **17,** 473–490 (2004).

300. Moreno-Jiménez, E. P. *et al.* Adult hippocampal neurogenesis is abundant in neurologically healthy subjects and drops sharply in patients with Alzheimer's disease. *Nature Medicine* **25,** 554–560 (2019).

301. Gropman, A. L. Patterns of Brain Injury in Inborn Errors of Metabolism. *Seminars in Pediatric Neurology* **19,** 203–210 (2012).

302. Haack, T. B. *et al.* Absence of the Autophagy Adaptor SQSTM1/p62 Causes Childhood-Onset Neurodegeneration with Ataxia, Dystonia, and Gaze Palsy. *The American Journal of Human Genetics* **99,** 735–743 (2016).

303. Sánchez-Martín, P. & Komatsu, M. p62/SQSTM1 – steering the cell through health and disease. *Journal of Cell Science* **131,** jcs222836 (2018).

304. Geisler, S. *et al.* PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1. *Nature Cell Biology* **12,** 119–131 (2010).

305. Chu, C. T. Mechanisms of selective autophagy and mitophagy: Implications for neurodegenerative diseases. *Neurobiology of Disease* **122,** 23–34 (2018).

306. Muto, V. *et al.* Biallelic SQSTM1 mutations in early-onset, variably progressive neurodegeneration. *Neurology* **91,** e319–e330 (2018).

307. D'Souza, A. R. & Minczuk, M. Mitochondrial transcription and translation: overview. *Essays in Biochemistry* **62,** 309–320 (2018).

308. Bratic, A. *et al.* Mitochondrial Polyadenylation Is a One-Step Process Required for mRNA Integrity and tRNA Maturation. *PLOS Genetics* **12,** e1006028 (2016).

309. Hajnsdorf, E. & Kaberdin, V. R. RNA polyadenylation and its consequences in prokaryotes. *Philosophical Transactions of the Royal Society B: Biological Sciences* **373,** 20180166 (2018).

310. Kudla, J., Hayes, R. & Gruissem, W. Polyadenylation accelerates degradation of chloroplast mRNA. *The EMBO Journal* **15,** 7137–7146 (1996).

311. Villalba, A., Coll, O. & Gebauer, F. Cytoplasmic polyadenylation and translational control. *Current Opinion in Genetics & Development* **21,** 452–457 (2011).

312. Crosby, A. H. *et al.* Defective Mitochondrial mRNA Maturation Is Associated with Spastic Ataxia. *The American Journal of Human Genetics* **87,** 655–660 (2010).

313. Yehudai-Resheff, S., Hirsh, M. & Schuster, G. Polynucleotide Phosphorylase Functions as Both an Exonuclease and a Poly(A) Polymerase in Spinach Chloroplasts. *Molecular and Cellular Biology* **21,** 5408–5416 (2001).

314. Wang, G. *et al.* PNPASE Regulates RNA Import into Mitochondria. *Cell* **142,** 456–467 (2010).

315. Niemann, M. *et al.* tRNAs and proteins use the same import channel for translocation across the mitochondrial outer membrane of trypanosomes. *Proceedings of the National Academy of Sciences* **114,** E7679–E7687 (2017).

316. Clemente, P. *et al.* SUV3 helicase is required for correct processing of mitochondrial transcripts. *Nucleic Acids Research* **43,** 7398–7413 (2015).

317. Borowski, L. S., Dziembowski, A., Hejnowicz, M. S., Stepien, P. P. & Szczesny, R. J. Human mitochondrial RNA decay mediated by PNPase–hSuv3 complex takes place in distinct foci. *Nucleic Acids Research* **41,** 1223–1240 (2013).

318. Siira, S. J. *et al.* LRPPRC-mediated folding of the mitochondrial transcriptome. *Nature Communications* **8,** 1532 (2017).

319. Ruzzenente, B. *et al.* LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs. *The EMBO Journal* **31,** 443–456 (2012).

320. Oláhová, M. *et al.* LRPPRC mutations cause early-onset multisystem mitochondrial disease outside of the French-Canadian population. *Brain* **138,** 3503–3519 (2015).

321. Bratic, A. *et al.* The Bicoid Stability Factor Controls Polyadenylation and Expression of Specific Mitochondrial mRNAs in Drosophila melanogaster. *PLOS Genetics* **7,** e1002324 (2011).

322. Piper, M. D. W. *et al.* A holidic medium for Drosophila melanogaster. *Nature Methods* **11,** 100–105 (2014).

323. Huang, J., Zhou, W., Watson, A. M., Jan, Y.-N. & Hong, Y. Efficient Ends-Out Gene Targeting In Drosophila. *Genetics* **180,** 703–707 (2008).

324. Schober, F. A. *et al.* Stable Isotope Labeling of Amino Acids in Flies (SILAF) Reveals Differential Phosphorylation of Mitochondrial Proteins Upon Loss of OXPHOS Subunits. *Molecular & Cellular Proteomics* **20,** 100065 (2021).

325. Schober, F. A., Atanassov, I., Freyer, C. & Wredenberg, A. Mitochondrial Gene Expression, Methods and Protocols. *Methods in Molecular Biology* **2192,** 75–87 (2020).

326. Piper, M. D. W. *et al.* A holidic medium for Drosophila melanogaster. *Nature Methods* **11,** 100–105 (2014).

327. Piper, M. D. *et al.* Matching Dietary Amino Acid Balance to the In Silico-Translated Exome Optimizes Growth and Reproduction without Cost to Lifespan. *Cell Metabolism* **25,** 610–621 (2017).

328. The, M. & Käll, L. Focus on the spectra that matter by clustering of quantification data in shotgun proteomics. *Nature Communications* **11,** 3234 (2020).

329. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43,** e47–e47 (2015).

330. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* **14,** 417–419 (2017).

331. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15,** 550 (2014).

332. Ignatiadis, N., Klaus, B., Zaugg, J. B. & Huber, W. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. *Nature Methods* **13,** 577–580 (2016).

333. Dhir, A. *et al.* Mitochondrial double-stranded RNA triggers antiviral signalling in humans. *Nature* **560,** 238–242 (2018).

334. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9,** 357–359 (2012).

335. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).

336. Robinson, J. T. *et al.* Integrative genomics viewer. *Nature Biotechnology* **29,** 24–26 (2011).

337. Mullen, A. R. *et al.* Reductive carboxylation supports growth in tumour cells with defective mitochondria. *Nature* **481,** 385–388 (2012).

338. Gaude, E. *et al.* NADH Shuttling Couples Cytosolic Reductive Carboxylation of Glutamine with Glycolysis in Cells with Mitochondrial Dysfunction. *Molecular Cell* **69,** 581–593.e7 (2018).

339. Abdulhag, U. N. *et al.* Mitochondrial complex IV deficiency, caused by mutated COX6B1, is associated with encephalomyopathy, hydrocephalus and cardiomyopathy. *European Journal of Human Genetics* **23,** 159–164 (2015).

340. Farhan, S. M. K. *et al.* Exome sequencing identifies NFS1 deficiency in a novel Fe□S cluster disease, infantile mitochondrial complex II/III deficiency. *Molecular Genetics & Genomic Medicine* **2,** 73–80 (2014).

341. Shishkova, E. *et al.* Global mapping of CARM1 substrates defines enzyme specificity and substrate recognition. *Nature Communications* **8,** 15571 (2017).

342. Avila, M. A. *et al.* Reduced mRNA abundance of the main enzymes involved in methionine metabolism in human liver cirrhosis and hepatocellular carcinoma. *Journal of Hepatology* **33,** 907–914 (2000).

343. Mentch, S. J. *et al.* Histone Methylation Dynamics and Gene Regulation Occur through the Sensing of One-Carbon Metabolism. *Cell Metabolism* **22,** 861–873 (2015).

344. Simpson, S. J. *et al.* Matching Dietary Amino Acid Balance to the In Silico- Translated Exome Optimizes Growth and Reproduction without Cost to Lifespan Article Matching Dietary Amino Acid Balance to the In Silico-Translated Exome Optimizes Growth and Reproduction without Cost to Lifespan. *Cell Metabolism* **25,** 610–621 (2017).

345. Zong, S. *et al.* Structure of the intact 14-subunit human cytochrome c oxidase. *Cell Research* **28,** 1026–1034 (2018).

346. Calvo-Garrido, J. *et al.* SQSTM1/p62-Directed Metabolic Reprogramming Is Essential for Normal Neurodifferentiation. *Stem Cell Reports* **12,** 696–711 (2019).

347. Pajak, A. *et al.* Defects of mitochondrial RNA turnover lead to the accumulation of double-stranded RNA in vivo. *PLOS Genetics* **15,** e1008240 (2019).

348. Szczesny, R. J. *et al.* Human mitochondrial RNA turnover caught in flagranti: involvement of hSuv3p helicase in RNA surveillance. *Nucleic Acids Research* **38,** 279–298 (2010).

349. Buck, M. D., Sowell, R. T., Kaech, S. M. & Pearce, E. L. Metabolic Instruction of Immunity. *Cell* **169,** 570–586 (2017).

350. Kapnick, S. M., Pacheco, S. E. & McGuire, P. J. The emerging role of immune dysfunction in mitochondrial diseases as a paradigm for understanding immunometabolism. *Metabolism* **81,** 97–112 (2018).

351. Cheng, L., Baonza, A. & Grifoni, D. Drosophila Models of Human Disease. *BioMed Research International* **2018,** 1–2 (2018).

352. *MitoDB, database and analysis tool* `https://molmet.ki.se/mass` (2021).