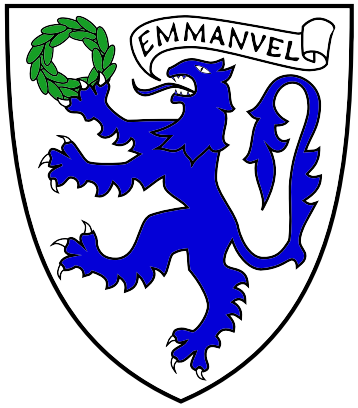
Biogenesis and Function of the Mitochondrial Ribosome



Petra Páleníková

Under the supervision of:

Dr Michal Minczuk

MRC Mitochondrial Biology Unit

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

Emmanuel College September 2021

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification at the University of Cambridge or any other University or similar institution except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the Clinical Medicine and Clinical Veterinary Medicine Degree Committee.

Initial version of ComPrAn R package presented in Chapter 3 of this thesis has been created as part of ITN REMIX secondment with Dr Rick Scavetta (Scavetta Academy). This Chapter is based on the following published work:

> Páleníková, P., Harbour, M.E., Ding, S., Fearnley, I.M., Van Haute, L., Rorbach, J., Scavetta, R., Minczuk, M., and Rebelo-Guiomar, P. "Quantitative density gradient analysis by mass spectrometry (qDGMS) and complexome profiling analysis (ComPrAn) R package for the study of macromolecular complexes." Biochim. Biophys. Acta BBA - Bioenerg. 1862, 148399, 2021

Petra Páleníková
September 2021

PhD Thesis Summary

Biogenesis and Function of the Mitochondrial Ribosome

Petra Páleníková

In this thesis, I present results of my PhD research into the biogenesis and function of the human mitochondrial ribosome.

The mammalian mitochondrial ribosome (mitoribosome) is one of the largest ribonucleoprotein complexes in the cell, with the overall molecular mass of the fully assembled 55S monosome being ~2.7 MDa. It is indispensable for cell survival, as it translates 13 polypeptides encoded in mitochondrial DNA, which are essential for oxidative phosphorylation and therefore for supplying the cell with energy in the form of ATP. The mammalian mitoribosome consists of small 28S and large 39S subunits. It is of dual genetic origin, with all protein components encoded in the nucleus and all RNA components encoded in the mitochondrial DNA. A total of 82 proteins, 2 rRNAs and a structural tRNA comprise one mitoribosomal particle. Proper function and assembly of this molecular machine is reliant on numerous associated factors, such as RNA modifying enzymes and assembly factors. Although recent significant progress has been made in our understanding of how the mitoribosome assembles and its functions in translation, we still lack the knowledge about many factors that are necessary for these processes.

This work aims to provide insight into the function of selected proteins that were predicted to be involved in biogenesis and/or function of mitochondrial ribosome, namely mitochondrial rRNA methyltransferase 1 (MRM1) and GTP binding protein 8 (GTPBP8). The work also presents genomics and proteomics methods for the study of mitochondrial gene expression machinery and of mitoribosome integrity and assembly, respectively.

The first two chapters provide background information for the research performed in the remaining part of the thesis. In the first chapter, I summarise current knowledge of mitochondrial gene expression, with a focus on post-transcriptional RNA modifications and mitochondrial translation. The second chapter describes the materials and methods used.

In the third chapter, I present a computational method for analysis of complexome profiling data from experiments that employ stable isotope labelling by amino acids in cell culture (SILAC). This method is implemented in R and is freely available as the Bioconductor package ComPrAn. It provides tools for analysis of peptide-level data as well as normalisation and clustering tools for protein-level data, dedicated visualisation functions and is accompanied by a graphical user interface. Throughout this thesis ComPrAn has been used for quantitative and qualitative analysis of mitoribosomes in studied cell lines.

In the fourth chapter, I introduce a CRISPR/Cas9-based screening approach designed to target genes with known or predicted function in mitochondrial gene maintenance and expression. I apply this method to identify genes that show genetic interaction with MRM1. The screen identifies MRM2 as a top candidate for genetic interaction with MRM1. This finding is followed up by the generation of a double knockout cell line which shows severe mitochondrial deficiency, with uridine auxotrophy and disruption of assembly of small mitoribosomal subunit being the most striking effects

observed. These findings provide further insight into the role of MRM1 in mitochondria and highlights the complexity of regulation of mitochondrial translation.

The fifth chapter focuses on establishing the role of uncharacterised GTPBP8 protein in the cell. I localised GTPBP8 to mitochondria and studied its function by production of a knockout cell line. GTPBP8 knockout presents a strong oxidative phosphorylation defect due to impaired mitochondrial translation. Quantitative analysis of mitoribosome reveals accumulation of both small and large subunits in the knockout, suggesting that GTPBP8 might play a role in very late assembly of either of the subunits, subunit joining or translation initiation.

Overall, this work improves our understanding of the regulation of mitochondrial translation by characterising two mitochondrial proteins and their role in mitoribosome biogenesis and function. The CRIPSR/Cas9 screening methodology and ComPrAn R package presented here have potential to be used in the study of other proteins, extending the portfolio of methods available for research of mitochondrial function.

Acknowledgements

PhD is a long journey, and it would not be possible to reach the end of it without the continuous support of my family, friends, and colleagues.

I would like to thank my supervisor Dr Michal Minczuk, for guiding me over the past four years on this path towards a PhD and encouraging me to become an independent scientist. I am grateful to all the past and present members of the Mitochondrial Genetics Group, who created a supportive work environment and were happy to help me whenever I encountered difficulties in my research. I would also like to thank the European Commission and the Medical Research Council for funding this research.

Thank you to all my friends from the Emmanuel College, the Czech and Slovak society and beyond. I thoroughly enjoyed my time in Cambridge because of you. Thank you to all the members of the MCR committee throughout the past four years, it was a pleasure being able to contribute to the Emma community together with you. Many thanks to my international "classmates" from the ITN REMIX. I always very much looked forward to our annual meetings.

Finally, I would like to thank my family, for always supporting me and allowing me to pursue my studies far away from home. I want to thank my partner for bearing with all my questions and being there for me over the many stressful times. Although it was sometimes difficult not to be able to spend time with the people that are most important to me, especially over the past year and a half, I always felt you all care about me, and I thoroughly enjoyed the moments we had together.

Ďakujem

Contents

Declaration i
PhD Thesis Summary iii
Acknowledgements v
Contents vii
List of Abbreviations xi
List of Figures xiii
1 Introduction 1
1.1 Mitochondria – origin and function 1
1.1.1 Origin of mitochondria 1
1.1.2 Mitochondria as the energy producing organelle 2
1.1.3 Other functions of mitochondria 3
1.2 Mitochondrial DNA 4
1.2.1 mtDNA replication 6
1.3 Mitochondrial transcription 6
1.4 Post-transcriptional RNA processing 8
1.4.1 mRNA maturation and modifications 9
1.4.2 tRNA maturation and modifications 10
1.4.3 rRNA modifications 11
1.5 Mitochondrial translation 14
1.5.1 Mitochondrial ribosome 14
1.5.2 Mitoribosome assembly 16
1.5.3 Process of mitochondrial translation 18
1.6 Aims of the thesis 21
2 Material and methods 23
2.1 Materials 23
2.1.1 Chemicals and Reagents 23
2.1.2 Solutions 24
2.1.3 Primers 25
2.2 Molecular cloning 26
2.2.1 PCR 26
2.2.2 Sanger sequencing 27
2.2.3 Restriction digestion 27
2.2.4 DNA ligation 27

2.2.5 Gibson Assembly 27
2.2.6 Transformation of chemically competent *E. coli* 28
2.3 SDS-PAGE 28
2.4 Western blotting 28
2.5 Cell culture 29
2.5.1 Maintenance 29
2.5.2 CRISPR/Cas9 mediated gene knock-out 31
2.5.3 Stable transfection and selection 31
2.5.4 Transient DNA transfection and immunocytochemistry imaging 32
2.5.5 Growth rate measurement using Incucyte 32
2.5.6 Production of lentiviral vectors 33
2.5.7 Lentiviral transduction 34
2.6 CRISPR/Cas9 screen 34
2.6.1 Cloning of gRNA plasmid library 34
2.6.1.1 PCR amplification of oligonucleotide pool 34
2.6.1.2 Restriction digestion of plasmid and PCR product and ligation 35
2.6.1.3 Electroporation of bacteria and plasmid extraction 35
2.6.2 Cas9 activity assessment 36
2.6.3 Multiplicity of infection test 37
2.6.4 CRISPR/Cas9 viability screen procedure 38
2.6.5 Sequencing library preparation and next-generation sequencing 38
2.6.6 Analysis of sequencing data 41
2.7 RNA extraction, DNase treatment and cDNA synthesis 43
2.8 RT-qPCR 43
2.9 Mitochondrial isolation 44
2.10 SILAC proteomics analysis 45
2.11 Mass spectrometry analysis 46
2.12 Analysis of mitoribosomes 47
2.12.1 Sucrose gradient fractionation 47
2.12.2 Quantitative density gradient analysis by mass spectrometry (qDGMS) 47
2.13 35S labelling 47
3 Development of Complexome Profiling Analysis (ComPrAn) R package 49
3.1 Introduction 49
3.1.1 Complexome profiling and qDGMS 49
3.1.2 Challenges in analysis of complexome profiling data 51

3.1.3 Data analysis tools for complexome profiling 52
3.2 ComPrAn – design and implementation 54
3.2.1 Installation 55
3.2.2 Data inputs 55
3.2.3 Estimation of protein abundance from peptide abundance 58
3.2.4 Normalisation of protein intensities 59
3.2.5 Hierarchical clustering 60
3.2.6 Graphical user interface 61
3.3 ComPrAn – outputs and visualisations 63
3.4 Conclusions and future development 66
4 Identifying genetic interactions of MRM1 69
4.1 Introduction 69
4.1.1 The role of MRM1 in mitochondria 69
4.1.2 Study of mitochondrial OXPHOS by large-scale genetic screens 71
4.1.3 Use of CRISPR/Cas9 screens for identification of genetic interactions and synthetic lethality 72
4.2 Results 75
4.2.1 Design of gRNA library targeting mitochondrial gene expression machinery 75
4.2.2 Generation of MRM1 KO cells stably expressing Cas9 protein 79
4.2.3 CRISPR/Cas9 screen design 81
4.2.4 CRISPR/Cas9 screen optimisation 84
4.2.5 Analysis of screen results 86
4.2.5.1 Genetic interactions of MRM1 87
4.2.6 MRM2 as a top candidate for MRM1 genetic interaction partner 90
4.2.6.1 Production of MRM1/2 double knockout cell line 90
4.2.6.2 Characterisation of MRM1/2 double knockout cell line 91
4.2.6.3 Complementation of MRM1/2 double knockout 95
4.2.6.4 Mitochondrial proteome in absence of methyltransferases 100
4.2.6.5 Mitochondrial ribosome in MRM1/2 double knockout cell line 109
4.3 Discussion 119
5 The role of GTPBP8 protein in mitochondrial translation 129
5.1 Introduction 129
5.1.1 GTPases and their function in the assembly of mitochondrial ribosome .. 129
5.1.2 Function of homologs of GTPBP8 in other organisms 132

5.2 Results 135
5.2.1 Localisation of GTPBP8 into mitochondria 135
5.2.2 Production of GTPBP8 KO cell line by CRISPR/Cas9 136
5.2.3 Characterisation of cells lacking GTPBP8 138
5.2.4 Complementation of GTPBP8 knockout 142
5.2.5 Mitochondrial ribosome in the absence of GTPBP8 144
5.3 Discussion 153
6 Conclusions 157
7 References 159
Appendix I 180

List of Abbreviations

BFP	Blue fluorescent protein
BioID	Proximity-dependent biotin identification
BN-PAGE	Blue Native polyacrylamide gel electrophoresis
CRISPR	Clustered regularly interspaced short palindromic repeats
cryo-EM	Cryo-electron microscopy
DAPI	4',6-diamidino-2-phenylindole
D-loop	Displacement loop
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DTT	1,4-dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
FACS	Fluorescence-activated cell sorting
FRT	Flippase recognition target
gRNA	Guide RNA
HEK	Human embryonic kidney
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HRP	Horseradish peroxidase
HSP	Heavy strand promoter
ICC	Immunocytochemistry
IMM	Inner mitochondrial membrane
IMS	Intermembrane space
KO	Knockout
LC-MS/MS	Liquid Chromatography with tandem mass spectrometry
LSP	Light strand promoter
MELAS	Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes
MGE	Mitochondrial gene expression
MHC	Major histocompatibility complex

MOI	Multiplicity of infection
MRGs	Mitochondrial RNA granules
mRNA	Messenger RNA
MSH	Mannitol-sucrose-HEPES
mt-	Mitochondrial
mtLSU	Large mitoribosomal subunit
MTS	Mitochondrial targeting sequence
mtSSU	Small mitoribosomal subunit
NCR	Non-coding region
OXPHOS	Oxidative phosphorylation
PBS	Phosphate-buffered saline
qDGSM	Quantitative density gradient analysis by mass spectrometry
RITOLS	Ribonucleotide incorporation throughout the lagging strand
rRNA	Ribosomal RNA
Rt-qPCR	Quantitative reverse transcription PCR
SILAC	Stable isotope labelling by amino acids in cell culture
SOC	Super optimal broth with catabolite repression
TBE	Tris-borate-EDTA
TFAM	Mitochondrial transcription factor A
tRNA	Transfer RNA
WT	Wild type
ZFN	Zinc finger nucleases

List of Figures

Figure 1.1: Schematic representation of mammalian OXPHOS complexes 3
Figure 1.2: Schematic of human mitochondrial DNA 5
Figure 1.3: Nucleolytic processing of primary mtDNA transcript 8
Figure 1.4: Mitochondrial and cytosolic genetic code 11
Figure 1.5: Posttranscriptional modifications present in mitochondrial rRNAs 12
Figure 1.6: Structure of the human mitochondrial ribosome 16
Figure 1.7: Order of assembly of mitoribosomal proteins 17
Figure 1.8: Translation in mammalian mitochondria 20
Figure 2.1: Amplification of gRNAs from genomic DNA 39
Figure 2.2: Fraction of reads that was mapped to the gRNA library 41
Figure 2.3: Gel sections for mass spectrometry analysis 46
Figure 3.1: Schematic of a qDGMS experiment 50
Figure 3.2: Schematic of ComPrAn workflow 54
Figure 3.3: Example of minimal peptide input data 57
Figure 3.4: Example of normalised protein input data 58
Figure 3.5: Selection of representative peptide in ComPrAn 59
Figure 3.6: Normalisation of protein abundance values in ComPrAn 60
Figure 3.7: Introduction page of the ComPrAn graphical user interface 62
Figure 3.8: Conditional access to ComPrAn app sections 62
Figure 3.9: Quantitative analysis of a single protein in ComPrAn 64
Figure 3.10: Visualisation of all peptides detected for a selected protein 64
Figure 3.11: Summary plots available in ComPrAn Shiny app 65
Figure 4.1: Position of the Gm1145 modification in the large mitoribosomal subunit 69
Figure 4.2: Schematic representation of genetic interactions 73
Figure 4.3: Selection of negative control genes 77
Figure 4.4: Selection of positive control genes 78
Figure 4.5: Composition of gRNA library 79
Figure 4.6: Cas9 activity check in HEK293 Cas9 and MRM1 KO Cas9 cells 81
Figure 4.7: Schematic of the CRISPR/Cas9 screen design 82
Figure 4.8: Fraction of infected cells modelled by Poisson distribution 83
Figure 4.9: Titration of viral supernatant 84
Figure 4.10: Evaluation of MOI 84
Figure 4.11: Normalised reads per gRNA in detached and adherent cell populations 85

Figure 4.12: Comparison of normalised reads per gRNA among replicates 86
Figure 4.13: Genetic interactions of MRM1 and other genes of MGE..... 88
Figure 4.14: Production of MRM1/2 double knockout cell line 91
Figure 4.15: Analysis of OXPHOS in double KO cell line 92
Figure 4.16: qPCR analysis of mtDNA copy number and RNA levels in MRM1/2-KO cell line..... 93
Figure 4.17: Growth curve of MRM1/2 KO in glucose and galactose medium. 94
Figure 4.18: Growth curve of MRM1/2 KO in the presence or absence of uridine supplementation 95
Figure 4.19: Western blot analysis of MRM1/2 KO complementation 97
Figure 4.20: Western blot analysis of OXPHOS in MRM1/2 KO after complementation 97
Figure 4.21: Growth curve of MRM1/2 KO after complementation..... 99
Figure 4.22: Sequence alignment of human MRM1 protein and its homologs 100
Figure 4.23: Western blot analysis of MRM1/2 KO complemented with catalytically inactive MRM1 100
Figure 4.24: Number of detected proteins in SILAC mass spectrometry experiments .. 102
Figure 4.25: Fraction of mitochondrial proteins in SILAC proteomics experiments..... 102
Figure 4.26: Changes in mitochondrial proteome in MRM1/2 KO compared to HEK293 cell line..... 104
Figure 4.27: Quantitative analysis of mitochondrial proteome comparing MRM1 KO to HEK293 and MRM2 KO to HEK293..... 106
Figure 4.28: Quantitative comparison of steady state levels of OXPHOS proteins in the absence of mitochondrial methyltransferases..... 107
Figure 4.29: Quantitative comparison of mitochondrial proteome between MRM1/2 KO and MRM2 KO cell lines 108
Figure 4.30: Quantitative comparison of components of mitochondrial ribosome between MRM1/2 KO cell line and HEK293 cell line..... 109
Figure 4.31: Changes in steady state levels of mitochondrial proteins in MRM1 KO, MRM2 KO and MRM1/2 KO cell lines..... 110
Figure 4.32: mtSSU in MRM1/2 KO compared with HEK293 wild type 112
Figure 4.33: mtLSU in MRM1/2KO compared with HEK293 wild type 113
Figure 4.34: Quantitative comparison of mitoribosome between MRM1/2 KO and MRM2 KO by qDGMS 114
Figure 4.35: Migration profile of mtSSU and mtLSU proteins in MRM1/2KO and HEK293 cell lines 115

Figure 4.36: Perturbation of mitoribosomal proteins in mitochondrial lysate 117
Figure 4.37: Evaluation of relationship between protein downregulation and mtSSU and mtLSU assembly sequence 118
Figure 4.38: Mitoribosomal proteins predicted to have genetic interaction with MRM1 121
Figure 4.39: Migration of MRM1 protein in published qDGMS data 124
Figure 5.1: Conserved domains in TRAFAC class of GTPases 130
Figure 5.2: Sequence alignment of GTPBP8 from human and its homologs 133
Figure 5.3: Localisation of GTPBP8 by immunocytochemistry 136
Figure 5.4: Sequence analysis of GTPBP8 knockout clones A5 and B1 137
Figure 5.5: Analysis of GTPBP8 knockout by western blot 138
Figure 5.6: Western blot analysis of OXPHOS complexes in GTPBP8 KO 139
Figure 5.7: mtDNA copy number in GTPBP8 KO cells 139
Figure 5.8: Growth of GTPBP8 KO cells in media with various carbon source 140
Figure 5.9: mRNA levels in GTPBP8 KO 141
Figure 5.10: Metabolic labelling of mitochondrial translation products 141
Figure 5.11: Complementation of GTPBP8 143
Figure 5.12: Effect of GTPBP8 KO complementation on growth rate 143
Figure 5.13: Mitochondrial translation upon complementation 144
Figure 5.14: RT-qPCR analysis of rRNA levels in GTPBP8 KO 145
Figure 5.15: Analysis of steady state levels of mitoribosomal proteins in GTPBP8 KO 146
Figure 5.16: Analysis of mitoribosome in GTPBP8 KO cells 147
Figure 5.17: qDGMS analysis of mitoribosomal small subunit in GTPBP8 KO cells 149
Figure 5.18: qDGMS analysis of mitoribosomal large subunit 150
Figure 5.19: Quantification of mitoribosomal subunits in sucrose gradient 151
Figure 5.20: Evaluation of protein factors associated with translation in GTPBP8 KO cell line 152

1 Introduction

1.1 Mitochondria – origin and function

1.1.1 Origin of mitochondria

Cells are elementary building blocks of all living organisms. We distinguish two basic types of cells – prokaryotic and eukaryotic. Eukaryotic cells are more complex, their genetic information is stored in a membrane-bound nucleus and importantly, they contain organelles such as mitochondria. Mitochondria are separated from the cytosol by a double membrane. Their outer membrane is relatively permeable and allows small molecules to pass between the cytosol and intermembrane space (IMS) by the passive diffusion. The inner mitochondrial membrane (IMM), on the other hand, is a highly impermeable barrier between IMS and mitochondrial matrix. Mitochondria are unique among organelles, as they contain their own genetic information. It is now widely accepted that mitochondria evolved from bacteria, originating in an endosymbiotic event of the bacterium and a protoeukaryote. The coherent endosymbiotic theory of origin of mitochondria was first postulated in 1967 by Lynn Margulis (Sagan, 1967). Mitochondria most likely originated from an alphaproteobacteria from the Rickettsiales order (Wang and Wu, 2015). However, the exact alphaproteobacterial lineage of the original endosymbiont is still under debate. The initial host cell is thought to be related to the uncultivated group of archaea, known as the Asgard superphylum (Zaremba-Niedzwiedzka et al., 2017). This endosymbiotic event which led to the development of present-day mitochondria was a crucial step in the evolution of eukaryotes. Nevertheless, the exact timing of this event is unclear, with theories ranging from denoting the endosymbiosis as the driver of eukaryote evolution to studies dating the endosymbiosis towards the end of eukaryotic evolution (Poole and Gribaldo, 2014; Roger et al., 2017).

For mitochondria to become the organelle we observe today, they had to undergo many transformations from the original endosymbiotic alphaproteobacterium (Gabaldón and Huynen, 2007; Roger et al., 2017). Among others, these include creation of complex protein-import machinery and insertion of small molecule carries into the inner membrane, genome reduction and transfer of genes to the nucleus, origin of organelle division system coordinated with the host cell cycle, or integration of biochemical

pathways (Roger et al., 2017). Only about 10-16% of the present mitochondrial proteome has alphaproteobacterial origin (Gabaldón and Huynen, 2007). It is important to note that mitochondria differ significantly among eukaryotic organisms. For example, there are differences in the functions mitochondria perform based on whether the eukaryote inhabits aerobic or anaerobic environment (Roger et al., 2017). There is even an anaerobic microbial eukaryote, *Monocercomonoides*, that does not have mitochondria (Karnkowska et al., 2016). The absence of mitochondria in this organism is a result of a secondary loss and not a separate evolution event of mitochondria-free eukaryotic organism (Karnkowska et al., 2016). Unless otherwise stated, the mitochondria and their functions described in this thesis refer to the mammalian mitochondria.

1.1.2 Mitochondria as the energy producing organelle

One of the most important functions of the mitochondria in the cell is energy production by the oxidative phosphorylation (OXPHOS). Oxidative phosphorylation refers to the process of creation of ATP by the energy obtained from the transfer of electrons from NADH or $FADH_2$ to oxygen by a series of protein complexes located in the IMM (Figure 1.1). Electrons undergo a series of energetically favourable reactions and energy from these reactions is used to pump protons from mitochondrial matrix into the IMS. These reactions happen on the mitochondrial respiratory chain complexes. There are four respiratory chain complexes: NADH dehydrogenase (Complex I), succinate dehydrogenase (Complex II), ubiquinol:cytochrome *c* oxidoreductase (Complex III) and cytochrome *c* oxidase (Complex IV). Three of these, Complex I, Complex III and Complex IV use the energy of electron transport for pumping of four, two and four protons, respectively, into the IMS. This results into generation of proton motive force (Δp), which is a combination of a proton concentration gradient (ΔpH) and a membrane potential ($\Delta \psi$). In mitochondria the dominant component of Δp is $\Delta \psi$. The proton motive force is utilised by the F_oF_1 ATP synthase (Complex V). The flow of protons from IMS into the matrix via this complex drives the formation of ATP from ADP and inorganic phosphate by the rotary catalysis mechanism (Walker, 2013).

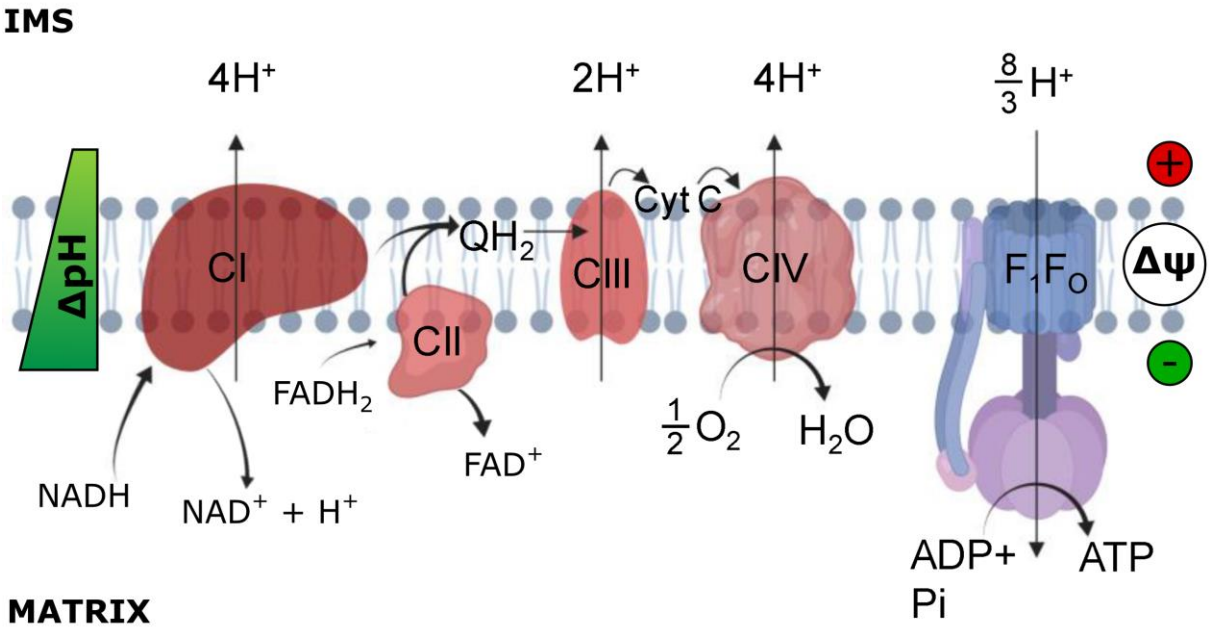


Figure 1.1: Schematic representation of mammalian OXPHOS complexes
Mitochondria produce cellular energy in the form of ATP by the process of oxidative phosphorylation. Oxidation of NADH at Complex I and $FADH_2$ at Complex II leads to a transfer of electrons to ubiquinone and reduces it to ubiquinol (QH2). Ubiquinol is reoxidised to ubiquinone by Complex III, and this complex also reduces cytochrome c. In the final step of electron transport chain, cytochrome c is oxidised by Complex IV and electrons are transferred to the final acceptor, O_2, producing an H_2O molecule. Electron transfer is coupled to pumping of protons across the inner mitochondrial membrane in Complexes I, III and IV. Protons are transported back to the matrix via Complex V, that uses the energy of proton motive force for generation of ATP. Adapted from (Lewis et al., 2019).

1.1.3 Other functions of mitochondria

Although mitochondria are best known for their production of ATP via OXPHOS, they are important for many other cellular functions. These include programmed cell death (Bock and Tait, 2020), Ca^{2+} homeostasis (Giorgi et al., 2018), iron-sulphur cluster assembly (Braymer and Lill, 2017) or heme synthesis (Swenson et al., 2020).

Iron-sulphur (Fe-S) clusters are an important component of several OXPHOS complexes where they function as redox centres for transport of electrons. However they are also important for processes such as genome maintenance (Fuss et al., 2015) or antiviral response (Upadhyay et al., 2014). Fe-S cluster synthesis is a stepwise process that begins in the mitochondrial matrix by synthesising a [2Fe-2S] cluster on a scaffold protein. This cluster is next transported on a transfer protein either to its target proteins or used for a further synthesis of [4Fe-4S] cluster which is then incorporated into proteins (Braymer and Lill, 2017).

Another cofactor involved in electron transfer in Complexes II, III and IV is heme, although participation of Complex II heme in the electron transport is unclear (Kim et al., 2012b). The biosynthetic pathway of heme begins and ends in mitochondrial matrix, with

intermediate steps taking place in the cytosol. First, 5-aminolevulic acid (5-ALA) is synthesised from succinyl-CoA and glycine in mitochondrial matrix. After the transport of 5-ALA into the cytosol by an unknown mechanism it is used for synthesis of coproporphyrinogen III by a series of three reactions. Coproporphyrinogen III is converted into protoporphyrinogen IX in the intermembrane space. The final two steps take place again in mitochondrial matrix where protoporphyrinogen IX is oxidised into protoporphyrin IX by an enzyme PPOX and, finally, insertion of iron is catalysed by a ferrochelatase resulting in the production of heme *b* (Swenson et al., 2020). Heme *b* can be further modified into hemes *a* that are utilised in Complex IV and hemes *c* of Complex III (Kim et al., 2012b). Apart from electron transport in mitochondrial OXPHOS complexes, heme functions as a cofactor in cellular processes including storage and transport of gases (Swenson et al., 2020) and acts as a signalling molecule by binding to transcription factors or kinases (Mense and Zhang, 2006).

1.2 Mitochondrial DNA

Mitochondria are a semiautonomous organelle, which means that they contain their own DNA. Mitochondrial DNA (mtDNA) originates from the bacterial ancestor of the mitochondria; however, it has undergone many changes since the initial endosymbiosis event. Present day eukaryotes contain mtDNA in various shapes and copy numbers. A unicellular phototropic organism, *Chromera velia*, has the smallest known mitochondrial genome in terms of protein-coding capacity with only two genes encoded on heterogenous linear molecules (Flegontov et al., 2015). On the other end of the spectrum are Jakobid protists with more than 60 protein-coding genes (Burger et al., 2013).

Human mtDNA is a circular molecule that encodes 13 protein coding genes, 2 ribosomal RNAs (rRNAs) and 22 transfer RNAs (tRNAs) (Figure 1.2). Many copies of mtDNA molecule are present in each cell. Majority of the molecules have identical sequence, but some can bear mutations leading to mtDNA heteroplasmy. Low levels of heteroplasmy are often present even in healthy individuals, while higher amounts of mutated mtDNA can lead to diseases (Payne et al., 2013).

mtDNA is organised in a ribonucleoprotein complex termed a nucleoid. Mitochondrial nucleoids are uniform in size with a diameter of about 100 nm. Many nucleoids contain a single mtDNA, with the average occupancy of 1.4 mtDNA molecule per nucleoid. The main protein component of the nucleoid is mitochondrial transcription factor A (TFAM), with about 1000 TFAM molecules per mtDNA, corresponding to one TFAM per 16.6 bp of mtDNA (Kukat et al., 2011).

The two strands of the mtDNA are termed "heavy" and "light" based on the higher proportion of guanine nucleotides in the former. This leads to a difference in the molecular weight and separation of the denatured mtDNA into two bands on the cesium chloride density gradient (Corneo et al., 1968). There are different numbers of genes encoded in the two mtDNA strands. Heavy strand encodes 10 mRNAs (translated into 12 proteins), 14 tRNAs and 2 rRNAs whereas only one mRNA and 8 tRNAs are encoded on the light strand. Apart from the different structure, mtDNA differs from the nuclear DNA by the absence of introns. There is also very low amount of non-coding sequence, with only one substantial non-coding region (NCR, Figure 1.2, blue) that contains promoters for transcription of heavy and light strands (HSP and LSP, respectively) as well as heavy strand origin of replication (Ori_H). Replication origin of the light strand (Ori_L) is located approximately 11 kb downstream from the Ori_H in a cluster of tRNA genes. NCR can contain a third DNA strand, termed 7S DNA, which is bound to the light strand and leads to the heavy strand displacement. This stable triple-stranded structure is called a D-loop, and its exact function is currently unknown (Nicholls and Minczuk, 2014; Walberg and Clayton, 1981).

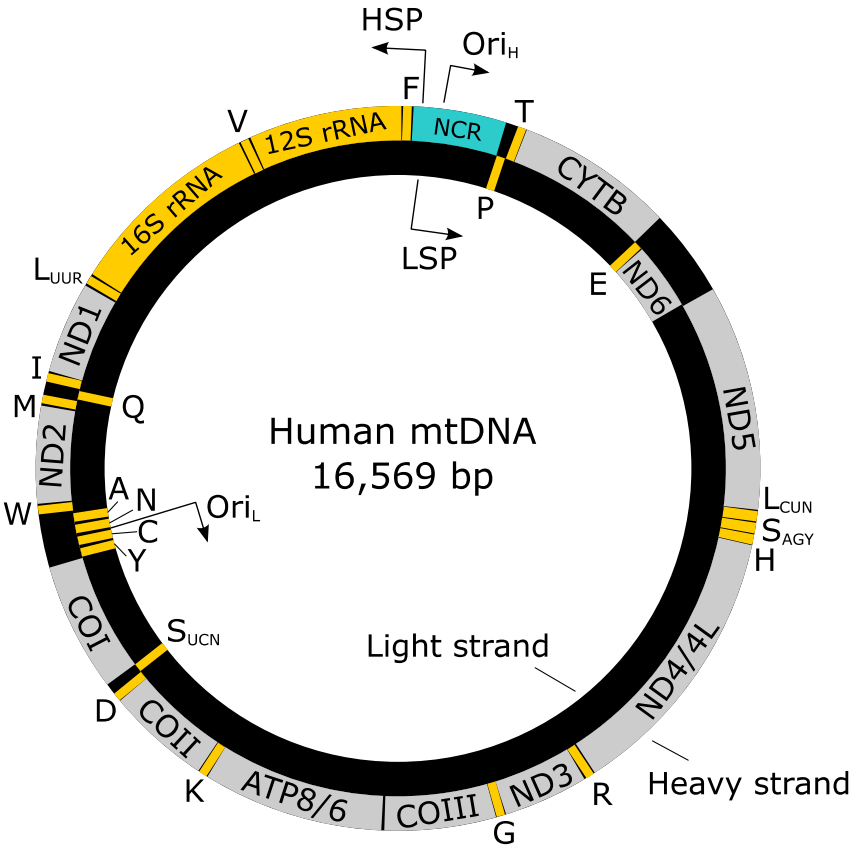


Figure 1.2: Schematic of human mitochondrial DNA
Circular human mitochondrial DNA molecule contains 13 protein coding genes (silver) and 24 non-coding genes (gold). Among the non-coding genes are 2 rRNA (12S and 16S) and 22 tRNAs (labelled with single amino acid names). Regulatory regions of the mtDNA are highlighted (HSP, heavy strand promoter; LSP, light strand promoter; $Ori_{H/L}$, origin of replication heavy/light strand; NCR, non-coding region; genes in the scheme are not to scale).

1.2.1 mtDNA replication

Both replication and gene expression of mtDNA are performed by a specialised expression machinery, distinct from the one responsible for replication and expression of nuclear DNA. *In vitro* experiments identified the minimal components required for replication of mtDNA to be DNA polymerase γ (POLγ), helicase Twinkle, and mitochondrial single-stranded binding protein (mtSSB) (Korhonen et al., 2004). Twinkle is a hexameric helicase that unwinds the mtDNA for POLγ which can synthesise DNA only from a single stranded template. Mammalian POLγ is a heterotrimer composed of one catalytic subunit POLγ A and two subunits of a processivity factor POLγ B (Yakubovskaya et al., 2006). POLγ with Twinkle are sufficient for synthesis of about 2 kb long ssDNA molecules from a dsDNA template. Addition of mtSSB increases the processivity of replication leading to generation of DNA molecules of about the size of the mtDNA (Korhonen et al., 2004).

The mechanism of mtDNA replication is still under debate. Three different models of replication have been proposed: strand displacement model, ribonucleotide incorporation throughout the lagging strand (RITOLS), and strand-coupled DNA replication (SCD). According to the strand displacement model, replication of light and heavy strands initiates from two different origins. First, replication is initiated from Ori_H and the heavy strand is replicated continuously, with mtSSB binding to the displaced ssDNA (Falkenberg, 2018; Fusté et al., 2014). Once the replication passes Ori_L, the exposed ssDNA forms a stem-loop structure allowing POLRMT to initiate RNA synthesis. POLRMT is quickly replaced with POLγ that replicates the light strand (Falkenberg, 2018). Therefore, in the strand displacement model the two mtDNA strands are synthesised asynchronously. In the second RITOLS model, replication is as well proposed to start from two different origins. However in this model, displaced heavy strand is annealed to RNA (as opposed to mtSSB binding) which is then removed during the light strand synthesis (Reyes et al., 2013; Yasukawa et al., 2006). Contrary to the first two models of replication, SCD proposes that replication initiates from multiple origins across a broad zone, Ori_Z, spanning the region of mt-Cyb, mt-ND6 and mt-ND5 genes (Bowmaker et al., 2003). According to SCD, mtDNA replication of heavy and light strand is synchronised and progresses bidirectionally (Bowmaker et al., 2003; Holt et al., 2000).

1.3 Mitochondrial transcription

Mitochondrial transcription differs from cytosolic in many aspects. One of the major differences stems from the organisation of mtDNA, with very few noncoding regions and

the absence of introns. Heavy and light strands of the mtDNA are transcribed from HSP and LSP promoters, respectively, both located in the major non-coding region of the mtDNA (Figure 1.2). There are reports of two heavy strand promoters in the mtDNA (Chang and Clayton, 1984; Montoya et al., 1982, 1983), however based on the more recent studies it is questionable whether HSP2 is a functional promoter (Litonin et al., 2010; Terzioglu et al., 2013). From each strand, transcription machinery produces one long polycistronic transcript that needs to be processed post-transcriptionally (described in section 1.4). For the majority of eukaryotes including humans, the enzyme responsible for synthesis of RNAs, DNA-dependant RNA polymerase POLRMT, is a single subunit phage-type polymerase (Shutt and Gray, 2006). The only exception seems to be a multi subunit bacteria-type RNA polymerase of Jakobid protists (Burger et al., 2013).

Transcription by POLRMT requires accessory protein factors. During transcription initiation TFAM recruits POLRMT to the promoter DNA by its N-terminal extension and TFB2M induces structural rearrangements in POLRMT leading to promoter opening and stabilising the open promoter (Hillen et al., 2017). Dependence on TFAM differs between the LSP and HSP promoters, with the C-terminal tail of TFAM being important for initiation only from HSP (Uchida et al., 2017). Furthermore, transcription from HSP can be initiated with only TFB2M and POLRMT, however only for a single round of transcription and TFAM is required for reinitiation (Uchida et al., 2017).

Transcription elongation factor, TEFM, is important for efficient production of mitochondrial transcripts and its downregulation primarily affects promoter-distal transcripts (Minczuk et al., 2011). In absence of TEFM a substantial fraction of transcripts initiated from LSP terminates on conserved sequence box II in the D-loop region of mtDNA and this short RNA was proposed to function as a primer for heavy strand replication (Pham et al., 2006). Addition of TFEM to the *in vitro* transcription reaction abolishes premature transcription termination and stimulates POLRMT processivity (Posse et al., 2015).

Mitochondrial transcription termination from LSP relies on mitochondrial termination factor 1 (MTERF1) which binds the mtDNA within the $tRNA^{Leu}$ gene. Here MTERF1 induces termination by unwinding the DNA and flipping three nucleotides (Yakubovskaya et al., 2010). This prevents synthesis of antisense rRNAs that could interfere with ribosome biogenesis (Terzioglu et al., 2013). The mechanism by which HSP driven transcription is terminated is currently unknown.

1.4 Post-transcriptional RNA processing

Mitochondrial DNA is transcribed into two long polycistronic transcripts. These transcripts need to be processed to release individual mRNAs, tRNAs and rRNAs. For rRNAs and most mRNAs an accepted model of excision from the primary transcript is the "tRNA punctuation model" (Ojala et al., 1981). As apparent from the sequence of the mtDNA (Figure 1.2), apart from few exceptions, mRNA and rRNA genes are flanked by a tRNA. According to the tRNA punctuation model, tRNA cloverleaf structures serve as a recognition element for cleaving the primary RNA transcript and release of individual RNA molecules (Figure 1.3).

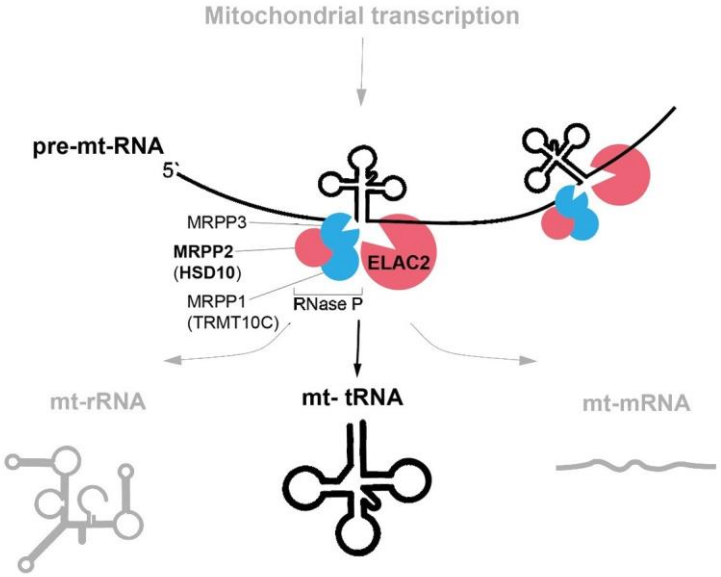


Figure 1.3: Nucleolytic processing of primary mtDNA transcript
mtDNA is transcribed into long polycistronic transcripts, where majority of mRNAs and rRNAs are separated by tRNA genes. Secondary structure of tRNAs is recognised by mitochondrial RNase P and ELAC2 for 5' and 3' processing of precursor RNAs, leading to release of tRNA, mRNA and rRNA molecules that are subjected to further maturation steps and posttranscriptional modifications. Adapted from (Powell et al., 2015).

The enzymes responsible for endonucleolytic cleavage at the 5' and 3' ends of tRNA are RNase P and ELAC2, respectively (Brzezniak et al., 2011; Holzmann et al., 2008). As opposed to many other known RNase P enzymes, mammalian mitochondrial RNase P does not contain an RNA component and is composed of three proteins: MRPP1, MRPP2 and MRPP3 (Holzmann et al., 2008). First, MRPP1 and a tetramer of MRPP2 form a subcomplex that binds to the conserved elements on tRNA, which is followed by MRPP3 binding and tRNA cleavage (Bhatta et al., 2021). Two MRPP1 and RNAs can bind to the MRPP2 tetramer, suggesting that two tRNAs can be processed simultaneously (Bhatta et al., 2021). Processing at the tRNA 5' ends by RNase P precedes the 3' processing by ERAL2. It has been proposed that 5' processed tRNA remains bound to the MRPP1 and MRPP2

which serve as a platform for ELAC2 processing as well as addition of 3' CCA (Reinhard et al., 2017). It is not entirely clear how processing of the primary transcript happens on sites where mRNA molecules are not surrounded by tRNAs. Members of Fas-activated serine/threonine kinase (FASTK) protein family have been proposed to have a role in processing of mitochondrial RNAs on these sites, however endonucleolytic activity of these protein has not been shown (Jourdain et al., 2017).

Once released from the primary transcript, mitochondrial RNAs are subjected to further maturation processes. These include polyadenylation of 3' ends of mRNAs, addition of 3'CCA and aminoacylation of tRNAs as well as posttranscriptional modifications of all classes of mitochondrial RNAs. There are various types of RNA modifications that can occur on either the nucleobase or at the ribose of ribonucleoside and collection of all RNA modification is termed the epitranscriptome (Rebelo-Guiomar et al., 2018).

1.4.1 mRNA maturation and modifications

The main maturation step of mRNAs is addition of polyA tail to their 3' end. Mitochondrial mRNAs are polyadenylated by mitochondrial poly(A) polymerase (mtPAP) (Tomecki et al., 2004). Similarly to other organisms, polyadenylation of mitochondrial RNAs affects their stability. The mechanism by which the stability is affected by the polyA differs between transcripts, with knockdown of mtPAP leading to either upregulation, downregulation or no change in transcript levels (Wilson et al., 2014). Furthermore, polyadenylation plays an important role for proper translation of seven mtDNA encoded proteins, since addition of polyA completes the stop codons of their mRNAs (Gagliardi et al., 2004). Other enzymes affect the length of the polyA, such as PDE12 which removes polyA extensions from mitochondrial RNAs (Rorbach et al., 2011). Inactivation of PDE12 leads to aberrant polyadenylation of mitochondrial tRNAs and rRNAs (Pearce et al., 2017).

Posttranscriptionally modified bases that were detected on mitochondrial mRNAs to date are pseudouridine (Ψ) and N^1-methyladenosine (m^1A) (Antonicka et al., 2017). Pseudouridylation is present on two mRNAs from Complex IV, mt-CO1 and mt-CO2, and its levels decrease upon depletion of pseudouridine synthetases TRUB2 and RPUSD3 (Antonicka et al., 2017). Two studies using recently developed next generation sequencing based techniques that allow identification of m^1A at single base resolution showed presence of this modification on mitochondrial mRNAs (Li et al., 2017; Safra et al., 2017). These studies differed in the number of identified m^1A site, with both identifying modification of mt-ND5 to interfere with its translation by a yet unknown mechanism.

1.4.2 tRNA maturation and modifications

Mitochondrial tRNAs are extensively posttranscriptionally modified. The CCA trinucleotide is added to the 3' end of all tRNAs by TRNT1 as it is required for charging of tRNA with amino acid (Nagaike et al., 2001). Nucleotide modifications on tRNAs are important for proper folding of tRNAs, in stabilising their structure or play a role in codon recognition and binding (Rebelo-Guiomar et al., 2018; Suzuki et al., 2020). To date, 18 different RNA modifications were identified at 137 positions of human mitochondrial tRNAs, with 34 enzymes implicated in deposition of these modifications (Suzuki et al., 2020). The anticodon loop of tRNAs bears large number of modifications, with the highest variability at the first position of the anticodon (Suzuki, 2021). This position interacts with the third base of the mRNA codon. Binding between the third codon and first anticodon positions is less stringent as usually the same tRNA needs to recognise multiple codons which differ in their third base (Figure 1.4). There are several cases of codons where the amino acid is determined by the presence of purine (A and G) or pyrimidine (U and C) at the third position. Modifications of first anticodon position play a role in specificity of decoding (Suzuki, 2021). For example, $tRNA^{Trp}$ and $tRNA^{LeuUUR}$ have 5-taurinomethyluridine and $tRNA^{Gln}$, $tRNA^{Lys}$ and $tRNA^{Glu}$ have 5-taurinomethyl-2-thiouridine at their first anticodon position, increasing their base pairing specifically with purines (Suzuki et al., 2002).

Before tRNA molecules can be used in translation, their cognate amino acid must be attached to their 3'CCA extension. This is performed by enzymes called aminoacyl-tRNA synthetases (aaRS). Each tRNA apart from $tRNA^{Gln}$ is charged by a tRNA specific aaRS. As there is no mitochondrial $tRNA^{Gln}$ specific aaRS, this tRNA is charged indirectly by first being misaminoacylated with glutamate aaRS EARS2 followed by modification of glutamate to glutamine by the GatCAB complex (Nagao et al., 2009). Some tRNAs can be polyadenylated which interferes with aminoacylation and therefore translation. The polyA needs to be removed from tRNAs by the PDE12 enzyme before they can be processed correctly (Pearce et al., 2017).

Mitochondrial genetic code

	U		C		A		G		
U	UUU	Phe (F)	UCU	Ser (S)	UAU	Tyr (Y)	UGU	Cys (C)	**U**
	UUC		UCC		UAC		UGC		**C**
	UUA	Leu (L)	UCA		UAA	STOP	**UGA**	**Trp (W)**	**A**
	UUG		UCG		UAG		UGG		**G**
C	CUU	Leu (L)	CCU	Pro (P)	CAU	His (H)	CGU	Arg (R)	**U**
	CUC		CCC		CAC		CGC		**C**
	CUA		CCA		CAA	Gln (Q)	CGA		**A**
	CUG		CCG		CAG		CGG		**G**
A	AUU	Ile (I)	ACU	Thr (T)	AAU	Asn (N)	AGU	Ser (S)	**U**
	AUC		ACC		AAC		AGC		**C**
	AUA	**Met (M)**	ACA		AAA	Lys (K)	**AGA**	**STOP**	**A**
	AUG		ACG		AAG		**AGG**		**G**
G	GUU	Val (V)	GCU	Ala (A)	GAU	Asp (D)	GGU	Gly (G)	**U**
	GUC		GCC		GAC		GGC		**C**
	GUA		GCA		GAA	Glu (E)	GGA		**A**
	GUG		GCG		GAG		GGG		**G**

Standard genetic code

	U		C		A		G		
U	UUU	Phe (F)	UCU	Ser (S)	UAU	Tyr (Y)	UGU	Cys (C)	**U**
	UUC		UCC		UAC		UGC		**C**
	UUA	Leu (L)	UCA		UAA	STOP	**UGA**	**STOP**	**A**
	UUG		UCG		UAG		UGG	Trp (W)	**G**
C	CUU	Leu (L)	CCU	Pro (P)	CAU	His (H)	CGU	Arg (R)	**U**
	CUC		CCC		CAC		CGC		**C**
	CUA		CCA		CAA	Gln (Q)	CGA		**A**
	CUG		CCG		CAG		CGG		**G**
A	AUU	**Ile (I)**	ACU	Thr (T)	AAU	Asn (N)	AGU	Ser (S)	**U**
	AUC		ACC		AAC		AGC		**C**
	AUA		ACA		AAA	Lys (K)	**AGA**	**Arg (R)**	**A**
	AUG	Met (M)	ACG		AAG		**AGG**		**G**
G	GUU	Val (V)	GCU	Ala (A)	GAU	Asp (D)	GGU	Gly (G)	**U**
	GUC		GCC		GAC		GGC		**C**
	GUA		GCA		GAA	Glu (E)	GGA		**A**
	GUG		GCG		GAG		GGG		**G**

Figure 1.4: Mitochondrial and cytosolic genetic code
Tables of mitochondrial and standard genetic code recognised by tRNAs in the mitochondria and in the cytosol, respectively. Codon sequence and three and single letter amino acid labels are shown. Initiation codon boxes are highlighted in green and stop codon boxes in orange. Differences between the two genetic codes are shown in bold. Mitochondrial DNA encodes 22 tRNAs, codons for 18 amino acids are recognised by a single tRNA each and codons for two amino acids are recognised by two different tRNAs (leucine – shades of blue, serine – shades of yellow).

1.4.3 rRNA modifications

Human mtDNA encodes two rRNAs, 12S and 16S that are structural components of small and large mitoribosomal subunits, respectively. Compared to cytosolic ribosomal RNAs, which are posttranscriptionally modified on more than 200 sites (Taoka et al., 2018), human mitochondrial rRNAs are modified on only 10 positions, five on the 12S and five on the 16S rRNAs (Rebelo-Guiomar et al., 2018). These modifications include base

methylations, pseudouridylation and 2'-*O* ribose methylations (Figure 1.5). All enzymes that modify mtDNA encoded rRNAs has been already identified. These are TRMT2B, METTL15, NSUN4 and TFB1M for 12S rRNA and TRMT61B, MRM1, MRM2, MRM3 and RPUSD4 for 16S rRNA.

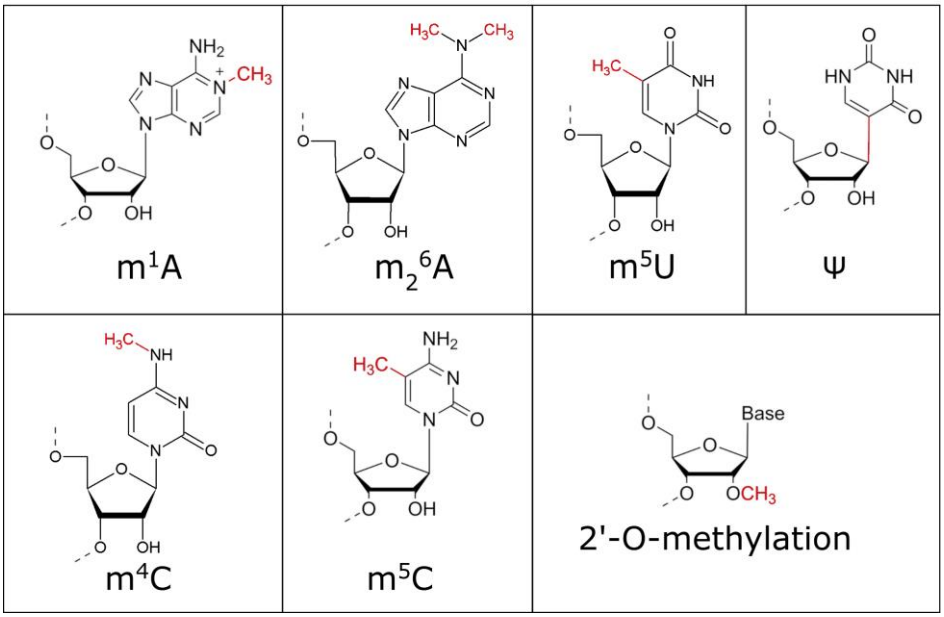


Figure 1.5: Posttranscriptional modifications present in mitochondrial rRNAs
A total of 10 posttranscriptional modifications are present on human mtDNA encoded rRNAs. These include base methylations (m^1A, two m_2^6A, m^4C, m^5C, m^5U), isomerisation of uracil to pseudouridine (Ψ) and three 2'-O ribose methylations. Modifications are shown in red. Molecule images adapted from (Wang and He, 2014) (m^1A, m^5C, Ψ, 2'-O-methylation) and Modomics database (Boccaletto et al., 2018) (m_2^6A, m^4C, m^5U).

TRMT2B protein is a methyltransferase responsible for m^5U429 modification of 12S rRNA (Laptev et al., 2020a; Powell and Minczuk, 2020). Both mouse and human cells where TRMT2B has been inactivated are generally healthy with no change in mitochondrial translation (Laptev et al., 2020a; Powell and Minczuk, 2020). Mouse cells with a knockout in TRMT2B showed statistically significant decrease in activity of Complexes I, III and IV (Laptev et al., 2020a).

Two enzymes, METTL15 and NSUN4 modify two nucleotides located near each other, m^4C839 and m^5C841, respectively. Knockout of NSUN4 in mouse is embryonically lethal and conditional knockout in heart leads to changes in mitochondrial function, including decreased translation (Metodiev et al., 2014). Apart from modifying C839 in 12S rRNA, NSUN4 also functions in a heterodimeric complex with MTERF4 as an assembly factor of mitochondrial ribosome (Cámara et al., 2011; Metodiev et al., 2014). Inactivation of METTL15 in either human or mouse cells has a negative effect on mitochondrial function, leading to decreased respiratory capacity, reduced mitochondrial translation and impaired assembly of mitoribosomal small subunit (Chen et al., 2020; Laptev et al., 2020b; Van Haute et al., 2019). The m^4C839 and m^5C841 modifications are interdependent as

inactivation of METTL15 also leads to decreased levels of m^5C841 modification (Chen et al., 2020; Laptev et al., 2020b; Van Haute et al., 2019).

Last known methyltransferase acting on 12S rRNA is TFB1M, which dimethylates positions A936 and A937 (Seidel-Rogol et al., 2003). Inactivation of TFB1M leads to mouse embryonic lethality and its disruption in heart results in impaired assembly of the mitochondrial ribosome and translation (Metodiev et al., 2009). TFB1M was identified as a type 2 diabetes risk factor (Koeck et al., 2011). ß-cell-specific knockout of TFB1M in mouse resulted in decreased insulin release from these cells and development of diabetes (Sharoyko et al., 2014).

The mitoribosomal large subunit 16S rRNA contains five modified nucleotides. Adenosine at position 947 is methylated by TRMT61B (Bar-Yaacov et al., 2016), which also methylates A58 at mitochondrial tRNAs (Chujo and Suzuki, 2012). Due to the charged nature of this modification it might be stabilising the rRNA structure by electrostatic interactions with the negatively charged phosphate rRNA backbone (Bar-Yaacov et al., 2016).

Three 2'-*O*-ribose methylations are present in the 16S rRNA. Gm1145 is deposited by MRM1 (Lee and Bogenhagen, 2014), which was found to interact with mitochondrial nucleoid (Lee et al., 2013). The function of MRM1 and its homologs in other organisms is discussed in detail in section 4.1.1. Other two 2'-*O*-ribose methyltransferases, MRM2 and MRM3, modify adjacent nucleotides U1369 and G1370, respectively (Lee and Bogenhagen, 2014). Previous studies showed that both MRM2 and MRM3 proteins are important for the biogenesis of large subunit of mitochondrial ribosome (Rorbach et al., 2014). Mutations in MRM2 gene lead to a human disease with clinical presentation similar to the mitochondrial encephalopathy, lactic acidosis and stroke-like episodes (MELAS) syndrome (Garone et al., 2017). Cells with a knockout of the MRM2 gene show a strong mitochondrial phenotype with severe disruption of mitochondrial translation (Rebelo-Guiomar et al., 2021). Similarly, mitochondrial translation is dysfunctional in MRM3 knockdown cells due to aberrant assembly of the mitoribosome large subunit (Rorbach et al., 2014).

RPUSD4 protein is responsible for pseudouridine formation at position 1397 of the 16S rRNA (Antonicka et al., 2017). Depletion of RPSUD4 leads to decreased mitochondrial translation, reduces levels of 16S rRNA as well as causes defects in mitoribosome assembly (Antonicka et al., 2017; Zaganelli et al., 2017). RPUSD4 is also responsible for Ψ39 in mitochondrial $tRNA^{Phe}$, however absence of this modification upon depletion of RPUSD4 does not affect steady-state levels or aminoacylation of this tRNA (Zaganelli et al., 2017).

1.5 Mitochondrial translation

In humans, the mitochondrial proteome contains around 1000 to 1500 proteins (Calvo et al., 2016; Smith and Robinson, 2019). Only thirteen of these proteins are encoded in the mtDNA and translated within mitochondria. Mitochondrial translation differs from the cytosolic translation systems both in protein machinery translating the mtDNA encoded proteins and in the genetic code used (Figure 1.4).

1.5.1 Mitochondrial ribosome

The central player in translation is a 55S mitochondrial ribosome, a large ribonucleoprotein complex composed of 28S small and 39S large subunits. The mitochondrial ribosome has diverged substantially compared to bacterial and cytosolic ribosomes, and it also differs between eukaryotic species (Table 1.1). Ribosomes of the majority of organisms are RNA rich, with typical ratio of RNA to proteins being 2:1. In mammalian mitochondria, the ratio is reversed with proteins constituting two thirds of the mitoribosomal mass (de Vries and van der Koogh-Schuuring, 1973). Human mitoribosome contains 82 nuclear encoded proteins and three mtDNA encoded RNAs (two rRNAs and one structural tRNA).

The mitoribosomal small subunit (mtSSU) serves as a platform for mRNA binding and decoding (De Silva et al., 2015). It is composed of 12S rRNA and 30 proteins, out of which 14 are specific to mitochondria (Figure 1.6) (Amunts et al., 2015; Greber et al., 2015). mRNA passes through mtSSU via an mRNA channel that starts at the entrance formed by uS5m and uS3m protein (Amunts et al., 2015). Due to the diameter of the channel only a single stranded mRNA can enter. The exit from the mRNA channel is formed by bS1m, bS21m and mS37 proteins (Amunts et al., 2015).

The role of large subunit of the mitoribosome (mtLSU) is catalysis of a peptidyl transferase reaction (De Silva et al., 2015). To be able to do this, mtLSU contains three binding sites for aminoacyl-tRNA (A), peptidyl-tRNA (P) and exit (E) site for deacylated-tRNA. A total of 52 proteins, 22 mitochondria specific and 30 with bacterial homologs, and two RNAs form the mtLSU (Amunts et al., 2015). The length of 16S rRNA of mtLSU is about half compared with bacterial 23S rRNA (Brown et al., 2014). The second RNA component of the large subunit is $tRNA^{Phe}$ or $tRNA^{Val}$ (Brown et al., 2014; Greber et al., 2015). Which of the two tRNAs is present in mtLSU differs between mammalian species, with $tRNA^{Val}$ in human and $tRNA^{Phe}$ in porcine mtLSU (Brown et al., 2014; Greber et al., 2015). It has been suggested that the reason why these two tRNAs and no others are part

of the mtLSU is because they are adjacent to the rRNA genes in the mtDNA (Figure 1.2) (Brown et al., 2014; Rorbach et al., 2016). Interestingly, in situations when $tRNA^{Val}$ is depleted, the human mitoribosome remains translation competent by incorporating $tRNA^{Phe}$ instead (Rorbach et al., 2016). This structural tRNA is a part of the central protuberance of the mtLSU, which interacts with the head of mtSSU and tRNAs in the mitoribosome (Brown et al., 2014). Polypeptides that formed during translation depart mtLSU via a polypeptide exit tunnel which has been remodelled for translation of hydrophobic membrane proteins (Brown et al., 2014). The two ribosomal subunits are connected by three protein-protein and six protein-RNA bridges (Amunts et al., 2015). The majority of mitoribosomes are associated with inner mitochondrial membrane and this interaction is mediated by the mL45 protein (Englmeier et al., 2017).

Table 1.1: Comparison of ribosomes between organisms and cell compartments

		E. coli	***S. cerevisiae***	**Human**	
		ribosome	**mitochondrial ribosome**	**cytosolic ribosome**	**mitochondrial ribosome**
Sedimentation coefficient		70S	74S	80S	55S
Molecular weight (MDa)		2.3	3	4	2.8
RNA : protein ratio		2:1	1:1	2:1	1:2
Small subunit	sedimentation coefficient	30S	37S	40S	28S
	RNA components	16S rRNA	15S rRNA	18S rRNA	12S rRNA
	number of proteins	21	34	32	30
Large subunit	sedimentation coefficient	50S	54S	60S	39S rRNA
	RNA components	23S rRNA; 5S rRNA	21S rRNA	28S rRNA; 5S rRNA; 5.8S rRNA	16S rRNA; $tRNA^{Val}$
	number of proteins	33	39	47	52
References		(Dunkle et al., 2011)	(Amunts et al., 2014; Desai et al., 2017)	(Anger et al., 2013)	(Amunts et al., 2015)

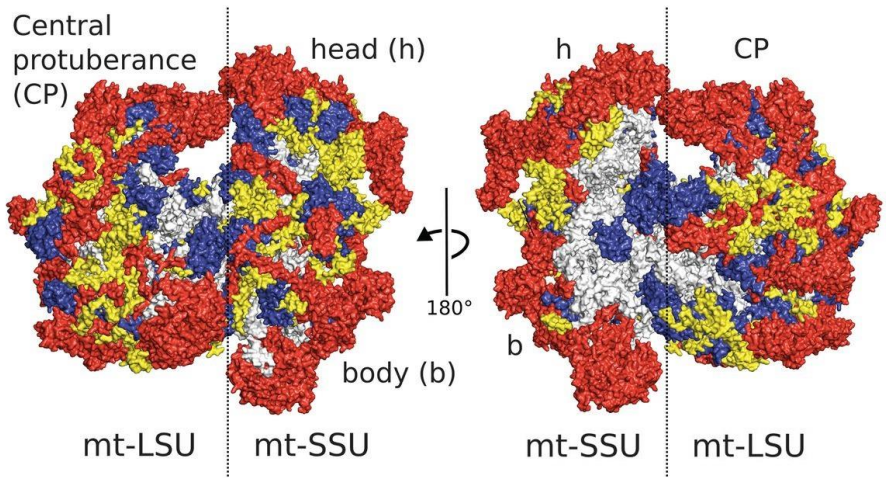


Figure 1.6: Structure of the human mitochondrial ribosome
Surface representation of the overall structure of the human mitoribosome. Colours represent protein conservation with proteins conserved between human and bacteria in blue, extensions of homologous proteins in yellow and mitochondria-specific proteins in red. rRNA is shown in grey. The morphological features of mtSSU (head and body) and of mtLSU (central protuberance) are labelled. Figure adapted from (Amunts et al., 2015).

1.5.2 Mitoribosome assembly

Assembly of a mitochondrial ribosome is a stepwise process in which small and large subunits are first assembled separately, followed by their joining into a translating monosome. Mitoribosomal proteins have been grouped into early, intermediate and late proteins based on their assembly kinetics analysed by pulse SILAC (stable isotope labelling by amino acids in cell culture) (Figure 1.7) (Bogenhagen et al., 2018). Clusters of proteins that share extensive protein-protein interactions join in a coordinated manner. Proteins from the mtSSU early binding group bind to the outer surface of the subunit, away from the interface with the mtLSU, whereas late binding proteins localise towards the interface of subunits (Bogenhagen et al., 2018). Similarly, early mtLSU proteins occupy outer surface of the mtLSU, avoiding the subunit interface. Number of early LSU proteins interact with the structural $tRNA^{Val}$. Intermediate binding proteins have a scattered distribution on the subunit and late binding proteins tend to be positioned near the subunit interface. The process of mitoribosome assembly takes approximately 2 to 3 hours (Bogenhagen et al., 2018). Assembly of the mtSSU and mtLSU starts in the nucleoid (Bogenhagen et al., 2014; Dalla Rosa et al., 2014). Later stages continue in a discrete foci located in close proximity to the nucleoids known as the mitochondrial RNA granules (MRGs) (Antonicka and Shoubridge, 2015; Tu and Barrientos, 2015). During translation, mitoribosomes are associated with the inner mitochondrial membrane and translated proteins are inserted into the membrane co-translationally (Englmeier et al., 2017; Itoh et al., 2021).

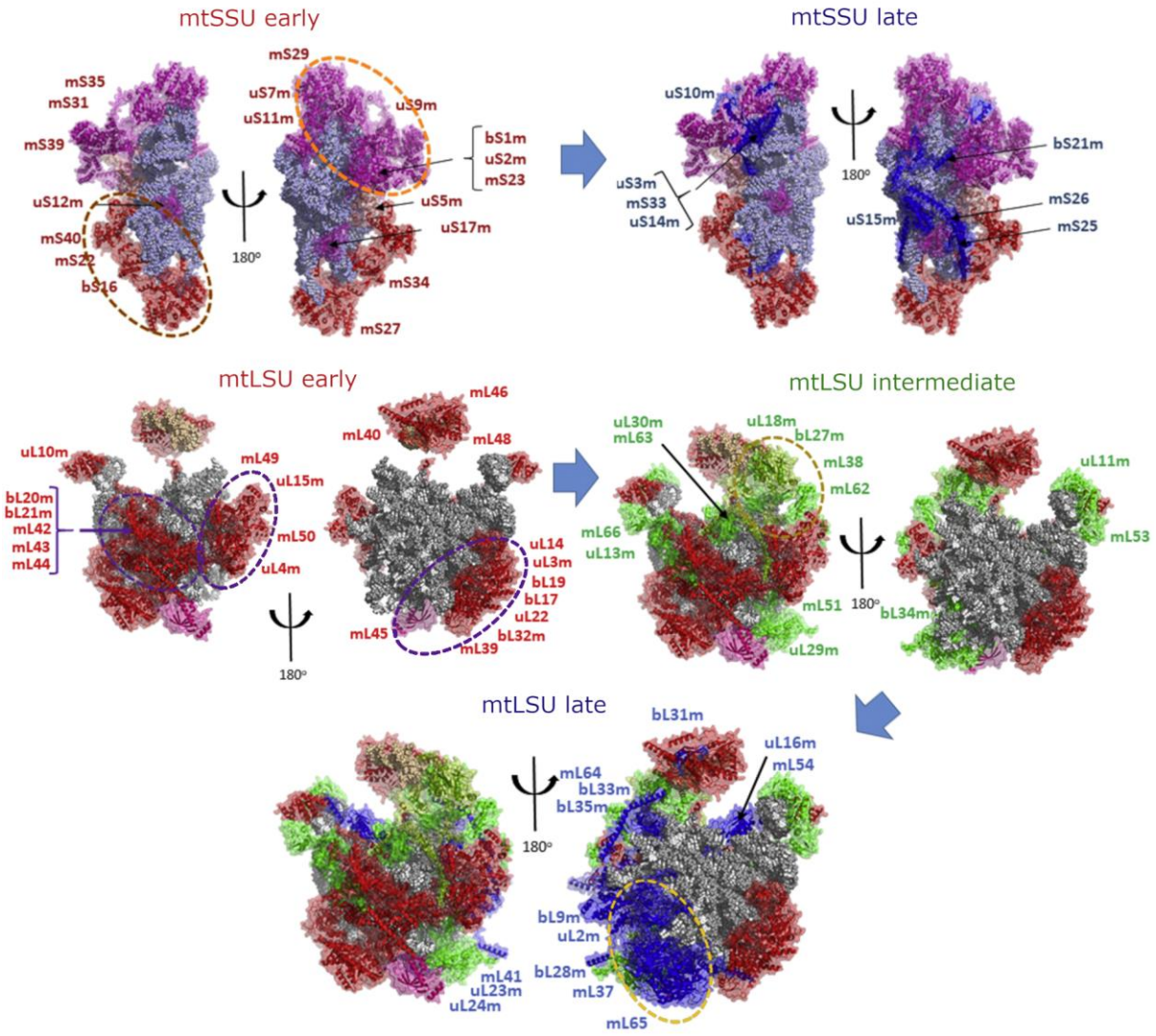


Figure 1.7: Order of assembly of mitoribosomal proteins
Mitoribosomal proteins were grouped into early, intermediate, and late assembly categories (Bogenhagen et al., 2018). 12S rRNA is shown in light blue spheres, 16S rRNA in grey and $tRNA^{Val}$ in tan. Early binding proteins are shown in red, intermediate in green and late in dark blue as cartoon structures within transparent surfaces. Clusters of interacting proteins which assembly is coordinated are highlighted by circles. mtSSU early binding proteins bS16m, mS22 and mS40 in the lower body interact with the 5' domain of 12S rRNA (mtSSU early, brown circle) and proteins mS34 and mS27 interact with the 3' domain of 12S rRNA. Proteins uS7m, uS9m and mS29 bind in the head region of mtSSU and have extensive RNA contacts (mtSSU early, orange circle). Subgroups of early assembly mtLSU proteins have strong interactions within each group, but only few strong interactions between the subgroups (mtLSU early, purple circles). tRNA is placed between two groups of proteins, with mL40, mL46 and mL48 binding slightly earlier than uL18m, bL27m and mL38 (mtLSU intermediate, green circle). Proteins uL2m, mL37 and mL65 join the mtLSU at a late assembly stage most likely in a coordinated manner (mtLSU late, yellow circle). Adapted from (Bogenhagen et al., 2018).

The process of the mitoribosome assembly is assisted by numerous assembly factors. These include RNA modifying enzymes (TRMT2B, METTL15, NSUN4, TFB1M, TRMT61B, MRM1, MRM2, MRM3 and RPUSD4; see section 1.4.3), GTPases (ERAL1, NOA1, GTPBP5, GTPBP6, GTPBP7, GTPBP10; see section 5.1.1), RNA helicases (DDX28, DHX30), anti-association factors (MALSU1, L0R8F8 and mtACP) and other chaperones (RBFA, $RCC1L^{V3}$, YBEY, FASTKD2, MPV17L2, MTERF3, MTERF4, NGRN, PTCD1, $RCC1L^{V1}$) (reviewed in Hilander et al., 2021; Lopez Sanchez et al., 2021).

Assembly factors play a role throughout the stages of the mitoribosome biogenesis. For example, MRM1 protein that was found to associate with nucleoids (Lee et al., 2013) and RNA helicase DDX28 that localises to the nucleoid adjacent MRGs (Antonicka and Shoubridge, 2015; Tu and Barrientos, 2015) are likely to function during early assembly. DDX28 interacts with 16S rRNA in a region between positions 1030-1480. Depletion of DDX28 leads to reduced levels of both 16S rRNA and mtLSU (Antonicka and Shoubridge, 2015; Tu and Barrientos, 2015). Its helicase activity is important for correct mtLSU assembly and it likely functions either on naked rRNA or during early assembly steps (Tu and Barrientos, 2015). Other factors such as MRM2 or the MALSU1, L0R8F8, mtACP anti-association module proteins are important in the late stages of the mitoribosome assembly (Brown et al., 2017; Rebelo-Guiomar et al., 2021). MRM2 was identified in mitoribosome late assembly intermediates by recent cryo-electron microscopy studies (Cipullo et al., 2021a; Hillen et al., 2021; Lenarčič et al., 2021) and in its absence late stage mtLSU assembly is stalled with the anti-association module bound to the mtLSU (Rebelo-Guiomar et al., 2021).

1.5.3 Process of mitochondrial translation

Translation can be divided into four main parts, initiation, elongation, termination, and ribosome recycling (Figure 1.8). In recent years progress has been made in characterising these steps, with valuable insight provided by high resolution cryo-EM structures of mitoribosomes at various translation stages (Aibara et al., 2020; Khawaja et al., 2020; Kummer et al., 2018). Nevertheless, the process of human mitochondrial translation is still not fully understood.

Translation initiation comprises binding of mRNA to the mtSSU and initiator tRNA into the P site of the ribosome (Kummer and Ban, 2021). Translation typically starts from a methionine amino acid. Since only one $tRNA^{Met}$ is encoded in mtDNA, methionyl-tRNA (Met-$tRNA^{Met}$) is marked for initiation by formylation (fMet-$tRNA^{Met}$). Mitochondrial translation initiation is assisted by only two initiation factors, mtIF2 and mtIF3. mtIF3 has been proposed to form a distinct pre-initiation complex with mtSSU (Khawaja et al., 2020). In this pre-initiation complex, position of mtIF3 C-terminal domain overlaps with the binding site of fMet-$tRNA^{Met}$. Therefore, the binding of mtIF3 and fMet-$tRNA^{Met}$ has been proposed to be mutually exclusive. In the presence of mtIF3, mitoribosomal protein mS37 stabilises the mtSSU head in a favourable position for subsequent accommodation of mtIF2 (Khawaja et al., 2020). Compared to the bacterial IF2 enzyme, mtIF2 has a 37 amino acids long insertion that substitutes the role of second translation initiation factor

(IF1) essential for bacterial translation (Gaur et al., 2008). mtIF2 blocks binding of tRNAs into the A site of mitoribosome and promotes binding of fMet-tRNA^{Met} in the P site (Kummer et al., 2018). Domain IV of the mtIF2 is bound to the 3' CCA of the fMet-tRNA^{Met}, with stable tRNA binding being dependent on hydrogen bonding between formylmethionine and mtIF2 (Kummer et al., 2018). mtIF2 and mtIF3 were observed to bind to the mtSSU simultaneously in absence of mRNA and initiator tRNA (Khawaja et al., 2020). Whether the mtLSU joins the mtSSU to form a monosome before binding of mRNA and tRNA or these RNAs bind to the mtSSU after mtIF3 departure and prior to subunit joining is currently unclear (Khawaja et al., 2020).

During elongation, mitochondrial elongation factor mtEFTu delivers aminoacylated tRNAs into the A site of the mitoribosome. mtEFTu is a guanine-nucleotide binding protein and GTP hydrolysis is required for this step. After delivery, guanine exchange factor mtEFTs recycles mtEFTu by exchange of GDP with GTP (Schwartzbach and Spremulli, 1989). Next, elongation factor G1 (mtEFG1) catalyses simultaneous movement of mRNA, peptidyl-tRNA from A site to P site and removal of deacylated tRNA from P site to allow decoding of next codon (Tsuboi et al., 2009). mRNAs are delivered to the mitoribosome for translation by a complex of leucine-rich PPR motif-containing protein (LRPPRC) and Stem-Loop Interacting RNA binding Protein (SLIRP) that bind to mS39 protein of the mtSSU (Aibara et al., 2020).

Translation is terminated once the ribosome reaches the end of an open reading frame with a stop codon in the A site. Stop codons UAA and UAG are recognised by a release factor mtRF1a which hydrolyses the peptidyl-tRNA bond, releasing the synthesised peptide (Soleimanpour-Lichaei et al., 2007). However, mitochondrial transcripts also contain non-standard stop codons AGA and AGG. For termination, these transcripts undergo a -1 frameshift, leading to positioning of UAG into the A site, which is subsequently recognised by mtRF1a and translation is terminated (Temperley et al., 2010).

The process of splitting mitoribosome subunits to make them available for translation of next mRNA is called ribosome recycling. This is performed by mtEFG2 (paralogue of elongation factor mtEFG1) and mitochondrial ribosome recycling factor mtRRF (Rorbach et al., 2008; Tsuboi et al., 2009). Recently a third protein, GTPBP6, has been proposed to play a role in the ribosome recycling, since its overexpression leads to increased amount of individual mtSSU and mtLSU subunits (Lavdovskaia et al., 2020).

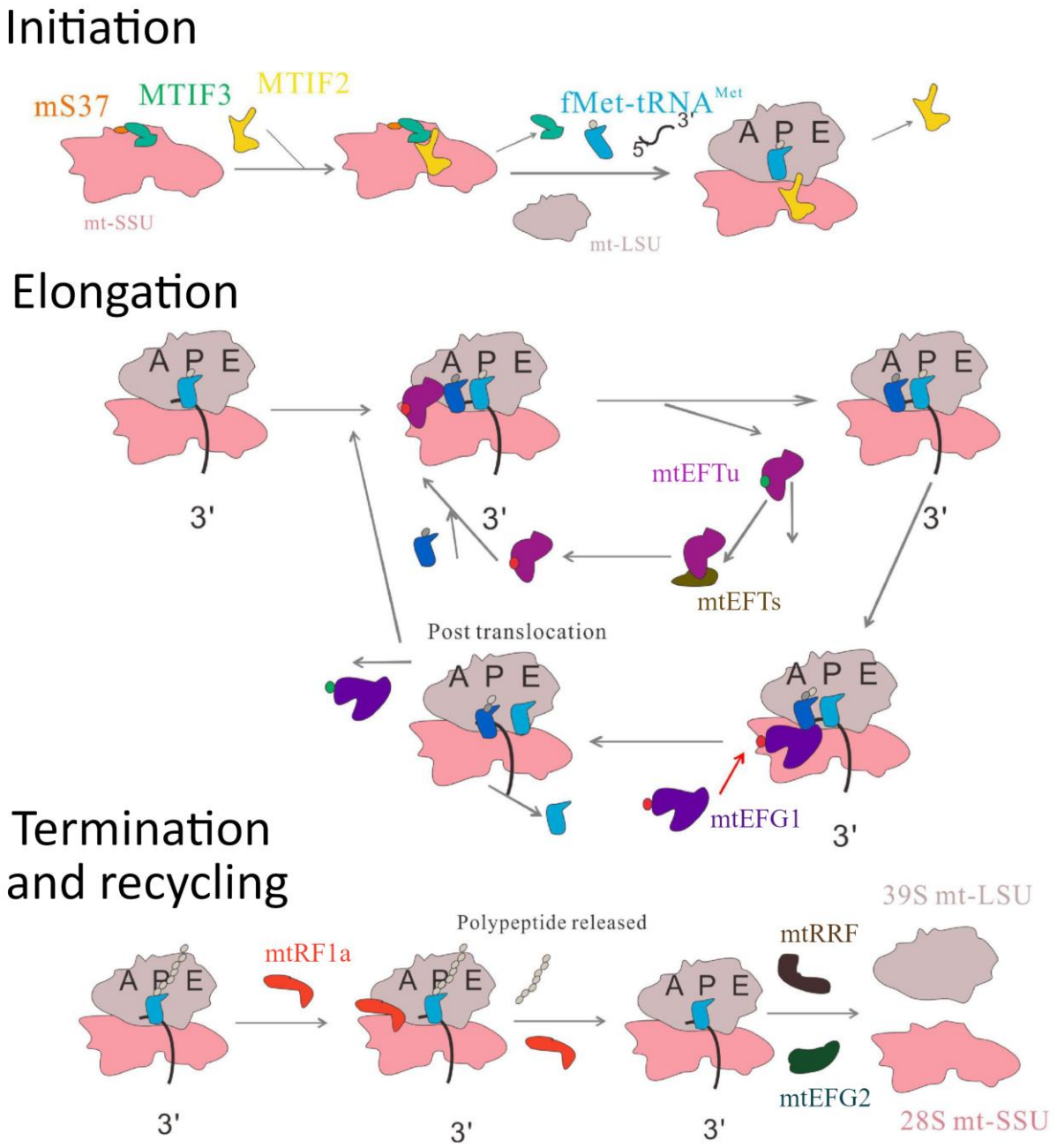


Figure 1.8: Translation in mammalian mitochondria
Translation can be divided into four stages – initiation, elongation, termination, and recycling. During initiation mtIF3 and mtIF2 bind to mtSSU forming a preinitiation complex. Binding of mtIF3 and initiator tRNA are mutually exclusive, therefore, this initiation factor dissociates from the complex before mtLSU, formylmethionyl tRNA^{Met} (fMet-tRNA^{Met}) and mRNA bind to create a complete initiation complex. During elongation, aminoacylated tRNAs are delivered to the ribosome by a ternary complex of mtEFTu, GTP and aminoacyl-tRNA. After delivery, mtEFTu is recycled by mtEFTs which exchanges GDP with GTP. Translocation reaction of mRNA and tRNAs from A site to P site and from P site to E site is catalysed by mtEFG1. Once stop codon reaches the A site of the mitoribosome, the cleavage of synthesised polypeptide is catalysed by a release factor mtRF1a and mitoribosome is recycled by mtRRF and mtEFG2 which split the monosome into mtLSU and mtSSU. Adapted from (Wang et al., 2021).

1.6 Aims of the thesis

As described throughout this section, mitochondrial gene expression is a complicated process that involves numerous protein factors and requires components with both mitochondrial and nuclear origin. Its study is often limited by the available experimental methods and tools to analyse data produced by these experiments. Therefore, one of the aims of this study was to introduce or improve methodology used for study of mitochondrial gene expression and mitoribosome integrity and assembly. Next, we aimed to apply this methodology in a study of selected proteins with a predicted role in biogenesis and/or function of mitochondrial ribosome in order to better understand their role in the cell. Two proteins were selected for study, mitochondrial rRNA methyltransferase 1 (MRM1) and GTP binding protein 8 (GTPBP8).

More specifically, we aimed to:

1. Streamline the analysis of SILAC complexome profiling and qDGMS experiments, a mass spectrometry-based methods for study of protein complexes (such as OXPHOS and mitoribosome) in near native state, by:
    - Creating a pipeline for analysis of SILAC complexome profiling data, from peptide tables to visualisation of results for proteins
    - Implementing the pipeline as an open-source software package, allowing future development and integration with other tools for analysis of mass spectrometry data
    - Making the analysis user friendly via a graphical user interface
2. Design and perform a CRIPSR/Cas9 screen aimed at identification of genetic interactions in mitochondrial gene expression machinery and use this approach to better understand the function of interacting genes. We selected MRM1 KO cell line for this study, since inactivation of MRM1 did not lead to mitochondrial defect in previous study (D'Souza, 2018). This aim was achieved by:
    - Aggregating a list of genes with known or predicted function in mitochondrial gene expression (MGE) and creating a custom gRNA library targeting MGE genes
    - Investigating genetic interactions of MRM1 by performing a CRISPR/Cas9 screen with this library in a wild type and MRM1 KO cells
    - Creating a list of candidate genetic interactions between MRM1 and other genes of MGE
    - Validating selected candidates by creating double knockout cell line

- Characterising double knockout cell line to better understand function of MRM1 protein in the cell

3. Characterise GTPBP8 protein that has been predicted to function in mitochondria. Several closely related GTP binding proteins were shown to be involved in the assembly of the mitochondrial ribosome. Therefore, we aimed to investigate whether this protein also plays a role in mitoribosome biogenesis and/or function by:
    - Creating a GTPBP8 knock out cell line (GTPBP8 KO)
    - Characterising mitochondrial function in GTPBP8 KO
    - Characterising mitochondrial translation in GTPBP8 KO
    - Studying mitochondrial ribosome in GTPBP8 KO by the qDGMS method in combination with the developed analysis pipeline

2 Material and methods

2.1 Materials

2.1.1 Chemicals and Reagents

[^{35}S]-L-Methionine was purchased from PerkinElmer.

1 kb plus DNA Ladder, Hygromycin, Lipofectamine 2000, Lipofectamine RNAiMax, NuPAGE MES SDS running buffers, NuPAGE LDS sample buffer, SimplyBlue Safe Stain, SYBR Safe were purchased from Invitrogen.

All restriction enzymes, Phusion polymerase were purchased from New England Biolabs.

DC assay kit and Precision Plus Protein Kaleidoscope prestained protein standard were purchased from Bio-Rad.

Alpha Select Silver Efficiency Competent Cells was purchased from Bioline Reagents.

DNA oligonucleotides, doxycycline, emetine dihydrochloride, glucose, galactose, L-glutamine, mannitol and uridine were purchased from Sigma-Aldrich.

Blasticidin and Zeocin were purchased from Invivogen.

Pierce™ BCA Protein Assay Kit was purchased from Thermo Scientific.

DTT was purchased from Melford Laboratories.

DMSO, EDTA, glycerol, glycine, HEPES, hydrochloric acid, isopropanol, magnesium chloride, methanol, potassium chloride, sodium chloride, sucrose and Triton-X100 were purchased from Fisher Scientific.

BSA, sodium hydroxide and urea were purchased from VWR International.

DMEM (Dulbecco's Modiefied Eagle's Medium), Glutamax, Opti-MEM, D-PBS (phosphate buffered saline), Penicillin-Streptomycin (10,000 U/ml), sodium pyruvate and trypsin-EDTA (0.25%) were purchased from Gibco.

DNeasy Blood and Tissue kits, Omniscript reverse transcription kits, QIAquick Gel Extraction kits, QIAquick PCR Purification kits and QIAprep Miniprep kits purchased from QIAGEN.

ECL and ECL prime were purchased from GE Healthcare.

RNaseZap, TRIzol and TRIzol LS were purchased from Ambion.

High Sensitivity D1000 Screentape was purchased from Agilent Technologies.

Milk Powder was purchased from Marvel.

RNasin Ribonuclease Inhibitor was purchased from Promega.

Proteinase Inhibitor tablets were purchased from Roche.

Tetracycline-free fetal bovine serum (FBS) was purchased from Biochrom AG.

Oligo pool for CRISPR screen was purchased from CustomArray.

X-ray films were purchased from Fujifilm.

Quick-DNA Miniprep Plus Kit was purchased from Zymo Research

2.1.2 Solutions

CONTINUOUS GRADIENT BUFFER: 50 mM Tris (pH 7.4), 20 mM MgAc, 40 mM NH_4Cl, 25 mM KCl

GRADIENT LYSIS BUFFER: 50 mM Tris (pH 7.4), 20 mM MgAc, 40 mM NH_4Cl, 25 mM KCl, 1 x Proteinase Inhibitor, 1% Triton X-100

HYPOTONIC BUFFER: 20 mM HEPES (pH 7.8), 5 mM KCl, 1.5 mM MgCl2, 1 mg/ml BSA, 1 x Proteinase Inhibitor

LB AGAR: 1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1 % (w/v) NaCl, 1.5 % (w/v) agar (pH 7)

MSH (1X): 210 mM mannitol, 70 mM sucrose, 20 mM HEPES (pH 7.8), 2 mM EDTA, 1 x Proteinase Inhibitor

MSH (2.5X): 525 mM mannitol, 175 mM sucrose, 20 mM HEPES (pH 7.8), 5 mM EDTA, 1 x Proteinase Inhibitor

PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4

PBS-T: 1 x PBS, 0.1 % (v/v) Tween 20

SOC MEDIUM: 0.5 % (w/v) yeast extract, 2 % (w/v) tryptone, 10 mM NaCl, 2.5 mM 10 mM MgCl2, 20 mM MgSO4, 20 mM glucose.

STEP GRADIENT BUFFER: 10 mM HEPES (pH 7.8), 5 mM EDTA, 0.5/1/1.5 M sucrose

TBE BUFFER: 89 mM tris base, 89 mM boric acid, 2 mM EDTA (pH 8.0)

TG LYSIS BUFFER: 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 1 mM EDTA, 1 % Triton X-100, 10 % Glycerol, 1.5 mM MgCl2, 1 x Proteinase Inhibitor

Cell culture media:

STANDARD DMEM: Dulbecco's modified eagle's media (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin solution (Gibco)

DMEM TET-: DMEM (Gibco) supplemented with 10% tetracycline free FBS (Biosera) and 1% penicillin-streptomycin solution (Gibco)

DMEM U: DMEM (Gibco) supplemented with 10% tetracycline free FBS (Biosera), 1% penicillin-streptomycin solution (Gibco) and 50 µg/ml uridine

SILAC MEDIUM: DMEM for SILAC (Thermo Fisher Scientific, 88364), 10% dialysed FBS, 1% penicillin-streptomycin solution (Gibco), 200 mg/L proline, 0.398 mM arginine, 0.798 mM lysine (Light-R_0K_0, Heavy-$R_{10}K_8$)

ZB DMEM: STANDARD DMEM containing 15 µg/ml blasticidin and 100 µg/ml zeocin

HB DMEM: DMEM TET- containing 15 µg/ml blasticidin and 50 µg/ml hygromycin B

NO GLUCOSE DMEM: DMEM, no glucose (Gibco), supplemented with 10% FBS (Biosera) and 1% penicillin-streptomycin solution (Gibco)

GLUCOSE MEDIUM: DMEM U containing 25 ng/ml doxycycline

GALACTOSE MEDIUM: DMEM, no glucose (Gibco), supplemented with 10% dialysed FBS and 1% penicillin-streptomycin solution (Gibco), 25 ng/ml doxycycline

2.1.3 Primers

Table 2.1: List of primers used in this work

Number	Name	Sequence 5' -> 3'	Length	Source
1	1-KpnI-GTPBP8-F	GTAAGGTACCATGGCGGCGCCCGGGCTG	28	This work
2	2-XhoI-GTPBP8-R	CTGACTCGAGGTCAAGACTTCCTGTTACACTGGC	34	This work
3	3-GTPBP8-KO-F1	CGGTGCTAGAGCGACTGAGC	20	This work
5	5-GTPBP8-KO-R1	CAGGTGAACCGTGAGAATCGCC	22	This work
25	25-XhoI-GTPBP8-F	TAAACTCGAGATGGCGGCGCCCGGGCTG	28	This work
26	26-GTPBP8-FLAG-XbaI-R	CATATCTAGACTACTTATCGTCGTCATCCTTGT AATCGTCAAGACTTCCTGTTACAC	57	This work
27	27-Lib-PS1-F	AGGCACTTGCTCGTACGACG	20	(Doench et al., 2016)
28	28-Lib-PS1-R	TTAAGGTGCCGGGCCCACAT	20	(Doench et al., 2016)
29	29-pKLV-gRNA-F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG TCTTGTGGAAAGGACGAAACACCG	57	This work
30	30-pKLV-gRNA-R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACA GCTAAAGCGCATGCTCCAGAC	54	This work
70	70_MRM2-KO-F2	CTGTGTAGCCTTATCAGTTCACC	23	This work
71	71_MRM2-KO-R2	CCGCATCCACTGGCAAAG	18	This work
73	73-MRM1-F-pWPXLd	CGAGACTAGCCTCGAGGTTTATGGCATTGCTCTCGACC	38	This work
75	75-MRM1-R_pWPXLd	CCGCGGATCCCGTAGTTTTCAGCCCTCATTTTGCCTCTC	39	This work
80	80-MRM1-E259A-F	GTGCTGGGGAATGCAGGCTCAGGTC	25	This work
81	81-MRM1-E259A-R	GACCTGAGCCTGCATTCCCCAGCAC	25	This work
82	82-MRM1-S289A-F	CCTGGACTTGAGGCATTGAACGTCTCTGTGGC	32	This work
83	83-MRM1-S259A-R	GCCACAGAGACGTTCAATGCCTCAAGTCCAGG	32	This work

2.2 Molecular cloning

For cloning of plasmid construct used for ICC experiments, following protocol was used:

GTPBP8 cDNA was purchased from CUSABIO (IMAGE: 4620206). cDNA was cloned into pcDNA5/FRT/TO FLAG/StrepII plasmid. cDNA was amplified with primers number 1 and 2 (for sequences see table in 2.1.3, PCR performed according to 2.2.1) and cloned into pcDNA5 vector using KpnI and XhoI restriction sites. This plasmid was used for immunocytochemistry performed according to 2.5.4.

For cloning of plasmid construct used for stable transfection, following protocol was used:

Wild type GTPBP8 cDNA and GTPBP8(S124A) cDNA (gift from Joanna Rorbach) was cloned into pINT plasmid. cDNA was PCR amplified with primers number 25 and 26 and cloned into pINT vector using XhoI and XbaI restriction sites. These plasmids were used for stable transfection according to 2.5.3.

2.2.1 PCR

PCR reactions were performed with Novagen KOD Hot Start DNA polymerase according to the manufacturer's instructions. Standard reaction composition is below.

PCR reaction (50 µl):

Volume	Reagent
5 µl	10X buffer
2 µl	dNTPs (2 mM each)
3 µl	25 mM $MgSO_4$
1 µl	KOD polymerase
1.5 µl	Forward primer (10 µM)
1.5 µl	Reverse primer (10 µM)
X µl	DNA template (50 ng)
To 50 µl	dH2O

PCR conditions:

1 x	95°C	2 min
35 x	95°C	20 s
	Primer dependent	10 s
	70°C	Target size dependent
	4°C	hold

2.2.2 Sanger sequencing

Plasmid DNA and PCR fragments were sequenced by Source Biosciences.

2.2.3 Restriction digestion

Restriction digestions of DNA were performed according to the manufacturer's instructions (New England Biolabs). Typically, 1-4 µg of PCR product or plasmid DNA were restricted with the required endonuclease for 1 hour. Digested plasmid was resolved on 1% agarose gel electrophoresis at 80V for 45-60 minutes. Band corresponding to the product of the digestion was excised and purified from the gel with QIAquick Gel Extraction kit according to the manufacturer's protocol.

2.2.4 DNA ligation

Endonuclease cleaved vector (100 ng) was mixed with PCR generated and endonuclease cleaved DNA insert at a molar ratio of 1:3. Ligation reaction was prepared by mixing the DNA with 2 µl of 10x Ligation buffer, 1 µl of T4 DNA ligase and water in a 20 µl total reaction volume. Reaction was incubated for 90 minutes at room temperature and used for transformation of chemically competent bacteria.

2.2.5 Gibson Assembly

Gibson assembly was performed to combine pWPLXd:IRES:PuroR plasmid linearised with PmeI restriction endonuclease with PCR amplified MRM1 cDNA fragments (primers for each construct specified in Table 2.2). 100 ng of linearised vector (corresponding to 0.0143 pmol) were combined with 5-fold excess of DNA inserts (1:5 molar ratio of vector:inserts), 10 µl of NEBuildr master mix and water to a final volume of 20 µl. Reaction was incubated at thermocycler at 50°C for 60 minutes. 5 µl of PmeI was added to the Gibson assembly reaction, to digest any vector that did not incorporate inserts, and incubated for 15 minutes at 37°C followed by 15 min heat inactivation at 65°C. 2 µl of the assembly reaction were used for bacterial transformation according to section 2.2.6.

Table 2.2: Primers for amplification of DNA fragments for Gibson Assembly

Construct	PCR primer pairs for amplification of DNA inserts
pWPLXd:IRES:PuroR+MRM1cDNA	73+75
pWPLXd:IRES:PuroR+MRM1(E259A)cDNA	80+75; 73+81
pWPLXd:IRES:PuroR+MRM1(S289A)cDNA	82+75;73+83

2.2.6 Transformation of chemically competent *E. coli*

Chemically competent bacteria were thawed on ice. 2 µl of the plasmid was added to the competent cells and mix gently by the pipette tip. Mixture was incubated on ice for 30 minutes. Heat shock was performed by placing the tube with bacteria into the 42°C water bath for 30 seconds. Tubes were transferred tubes to ice for 2 minutes. 950 µl of room-temperature SOC MEDIA was added to the tube and tubes were placed into the shaking incubator at 37°C for 60 minutes. After incubation,100 µl of the cells was spread onto the LB AGAR plates containing appropriate antibiotic. Plates were incubated overnight at 37°C.

2.3 SDS-PAGE

Whole cell pellets were lysed in TG LYSIS BUFFER and concentration of proteins in the lysate or in lysed isolated mitochondrial samples was determined by Pierce BCA protein assay. ~ 20 µg of proteins were mixed NuPAGE LDS 4 x sample buffer containing 200 mM DTT and boiled at 95ºC for 5 minutes. Afterwards, samples were loaded onto 4-12% Bis-Tris NuPAGE polyacrylamide gels and run at 200 V for 30 minutes. At this point, gels were either used for western blot analysis as per section 2.4 or prepared for mass spectrometry analysis as follows. Gels were rinsed with water for 5 minutes and stained using SimplyBlue Safestain for 2 hours, exchanging the stain after 1 hour. Gels were destained for 3x15 minutes in 10% acetic acid, 20% methanol. Image of the gel was taken, and gel was prepared for mass spectrometry as described in section 2.10.

2.4 Western blotting

Proteins resolved on 4-12% Bis-Tris NuPAGE polyacrylamide gels (Thermo Fischer Scientific) were transferred onto nitrocellulose membrane using the iBlot 2 Dry Blotting System (Thermo Fischer Scientific) using the protocol 'P0'. The membrane was blocked in 5% non-fat milk in PBS-T for one hour in room temperature, incubated with primary antibodies (Table 2.3) overnight at 4°C. After incubation, membrane was washed in PBS-T for 15 minutes 3 times, followed by an incubation with secondary antibody (Table 2.3) for one hour at room temperature. The membrane was washed in PBS-T 3 times for 15 minutes and developed using ECL (Amersham) and exposed to X-ray film in a dark room.

Table 2.3: Antibodies used for western blotting

Primary antibodies

Antibody	Dilution	Source	Reference	Species
MRM2	1:100	MyBioSource	MBS120390	mouse
MRM1	1:1000	Atlas Antibodies	HPA021598	rabbit
OXPHOS cocktail	1:1000	Abcam	ab110411	mouse
Beta actin	1:50000	Sigma	A2228	mouse
GTPBP8	1:1000	Atlas Antibodies	HPA034831	rabbit
MRPL3	1:2000	Proteintech	16584-1-AP	rabbit
MRPS17	1:2000	Proteintech	18881-1-AP	rabbit

Secondary antibodies

Antibody	Dilution	Source	Reference
Goat anti-rabbit IgG HRP	1:1000	Promega	W4011
Goat anti-mouse IgG HRP	1:2000	Promega	W4021

2.5 Cell culture

2.5.1 Maintenance

Cell lines used in this study are listed in Table 2.4. All cell lines were maintained in humidified incubators at 5 % CO2 and 37 °C. Cell were cultured in STANDARD DMEM. For cell lines where Cas9 was integrated into the FRT site a DMEM TET- was used. MRM1/2 KO and derived cell lines were maintained in DMEM U medium.

To freeze the cells for storage, cells were collected from confluent T175 or T75 flask, pelleted by centrifugation at 300g for 5 minutes and resuspended in 2 ml (for T175) or 1ml (T75) of FBS supplemented with 10% DMSO. 1 ml aliquots were transferred to a cryovial and frozen in a Mr Frosty container and placed at -80°C. Cells were transferred to a liquid nitrogen for long term storage.

Table 2.4: List of cell lines used in this work
First column specifies name used throughout this thesis. Plasmid constructs used for construction of cell lines and experiments in which a cell line has been used are listed in description column where appropriate.

Name used in this thesis	Description	Parental cell line	Source
143B	parental; used for immunocytochemistry experiments	NA	MBU cell bank
HEK293T	parental; used for viral production	NA	MBU cell bank
HEK293	Flp-in T-Rex™ HEK293 with Cas9 in FRT site; used throughout chapter 4	Flp-in T-Rex™ HEK293	Invitrogen; Cas9 integrated into FRT site by Aaron D'Souza
HEK293	HEK-attP/FRT with Cas9 in FRT site; used throughout chapter 5	HEK-attP/FRT	Waldner et al., 2011; Cas9 integrated into FRT site by Christopher Powell
MRM1 KO	chapter 4; stable integration of Cas9 into FRT site	MRM1 knockout in Flp-in T-Rex™ HEK293 background (D'Souza, 2018)	**This work**
MRM2 KO	chapter 4	Flp-in T-Rex™ HEK293	Rebelo-Guiomar et al., 2021
MRM1/2 KO	chapter 4; MRM2 knockout by CRIPSR/Cas9	MRM1 KO	**This work**
MRM1/2 KO + MRM1	chapter 4; lentiviral transduction of pWPLXd:IRES:PuroR+MRM1cDNA	MRM1/2 KO	**This work**
MRM1/2 KO + MRM1(E259A)	chapter 4; lentiviral transduction of pWPLXd:IRES:PuroR+MRM1(E259A)cDNA	MRM1/2 KO	**This work**
MRM1/2 KO + MRM1(S289A)	chapter 4; lentiviral transduction of pWPLXd:IRES:PuroR+MRM1(S289A)cDNA	MRM1/2 KO	**This work**
MRM1/2 KO + MRM2	chapter 4; lentiviral transduction of pWPLXd:IRES:PuroR+MRM2cDNA	MRM1/2 KO	**This work**
MRM1/2 KO +EV	chapter 4; lentiviral transduction of pWPLXd:IRES:PuroR	MRM1/2 KO	**This work**
HEK293 + MRM1	chapter 4; lentiviral transduction of pWPLXd:IRES:PuroR+MRM1cDNA	HEK293	**This work**
HEK293 + MRM2	chapter 4; lentiviral transduction of pWPLXd:IRES:PuroR+MRM2cDNA	HEK293	**This work**
HEK293 + EV	chapter 4; lentiviral transduction of pWPLXd:IRES:PuroR	HEK293	**This work**
GTPBP8 KO	chapter 5; knockout by CRISPR/Cas9	HEK293	**This work**
G8 KO + G8	chapter 5; stable integration of GTPBP8 into attP site	GTPBP8 KO	**This work**
G8 KO + G8(S124A)	chapter 5; stable integration of GTPBP8(S124A) into attP site	GTPBP8 KO	**This work**

2.5.2 CRISPR/Cas9 mediated gene knock-out

Cas9 expression in cell lines used for knock-out generation was induced by 50 ng/ml doxycycline 24 hours prior to transfection. $1x10^5$ cells were plated per well of a 6-well plate and transfected with the following reagents:

Volume	Reagent
7.5 µl	Lipofectamine RNAiMAX (Invitrogen)
625 ng	tracrRNA
625 ng	crRNA (combined for all 3 crRNAs per gene, Table 2.5)
In a total volume of 300 µl	Opti-MEM medium (Gibco)

This reaction was incubated at room temperature for 20 minutes prior addition to the cells. After 48 hours, 90% of cells were pelleted for DNA isolation to check cleavage efficiency and the remaining cells were diluted and split into single cells in a 96 well plate. Colonies were isolated, expanded and screened by PCR (primers for MRM2 KO: 70+72; primers for GTPBP8 KO: 3+5) and by western blotting (section 2.4).

Table 2.5: crRNA sequences used for knock-out production

Targeted gene	crRNA sequence
GTPBP8	GCGGTTCCTCGCCCCCTAC
GTPBP8	CGTATCTTTGACCCAAGCC
GTPBP8	GCGTGGTCGATACGGACGG
MRM2	GTGGCTGACCCGACATCTCA
MRM2	ACGAGAGGCACCAGATTCTG
MRM2	GCCTTCGGGTGTTAGACTGT

2.5.3 Stable transfection and selection

For integration of Cas9 into the FRT site of MRM1 KO cell line, following protocol was used:

Parental HEK293 Flp-in T-Rex cells were cultured in ZB DMEM. 24 hours before transfection cells were plated into 6 well tissue culture plate in non-selective STANDARD DMEM. Using Lipofectamine 2000 (Invitrogen) 0.25 µg pcDNA5/FRT/TO construct and 2.25 µg pOG44 vector were transfected into cells. After 24 hours, cells were transferred into 15 cm dish with STANDARD DMEM. 24 hours later, the medium was exchanged for selective HB DMEM. Once colonies were formed clones were isolated. Activity of Cas9 protein was assessed as described in section 2.6.2.

For integration of wild type GTPBP8, and GTPBP8(S124A) cDNA into the attP site of GTPBP8 KO cell line, following protocol was used:

GTPBP8 KO cells in the HEK-attP/FRT background were cultured in DMEM TET- with 500 µg/ml G418. 24 hours before transfection cells were plated into 6 well tissue culture plate in non-selective DMEM TET-. Using Lipofectamine 2000 (Invitrogen) 0.15 µg pINT constructs and 1.35 µg CMV_Int vector were transfected into cells. After 24 hours, cells were transferred into 15 cm dish with DMEM TET-. 24 hours later, the medium was exchanged for a selective DMEM TET- with 1 µg/ml puromycin. Once colonies were formed clones were isolated. Selected clones were induced with 100 ng/ml doxycycline for 24 hours and screened for expression of transgene by western blotting.

2.5.4 Transient DNA transfection and immunocytochemistry imaging

143B cells were transfected with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instruction. 24 hours later, cells were fixed using 4% formaldehyde in PBS and permeabilised by 1% Triton X-100 in PBS. Cells were incubated with the following antibodies: mouse anti-FLAG M2 (Sigma, 1:2000) primary, anti-mouse IgG Alexa Fluor 488 (Life Technologies, 1:1000) secondary; rabbit anti-TOM20 (Santa Cruz, 1:500) primary, anti-rabbit IgG Alexa Fluor 594 (Life Technologies, 1:1000) secondary. Coverslips with stained cells were mounted on slides in Prolong Gold antifade reagent with DAPI. Imaging was done using the Zeiss LSM 800 confocal microscope.

2.5.5 Growth rate measurement using Incucyte

For growth curves presented in Chapter 4:

$1x10^5$ cells were seeded per well of a 6 well plate in 2 ml of NO GLUCOSE DMEM Supplemented with 4.5 g/l glucose or 10 mM galactose and with or without 50 µg/ml uridine. Cell confluency was measured in the Incucyte S3 Live-Cell Analysis Systems (Sartorius) at 1-hour intervals.

For growth curves presented in Chapter 5:

$1x10^4$ cells were seeded per well of a 6 well plate in 2 ml of NO GLUCOSE DMEM Supplemented with 4.5 g/l glucose, 0.45 g/l glucose or 10 mM galactose as a sole carbon source. Cell confluency was measured in the Incucyte Zoom Live-Cell Analysis Systems (Sartorius) at 2-hour intervals.

2.5.6 Production of lentiviral vectors

Lentiviral vectors were produced in HEK293T cells over a four-day period. 24 hours before transfection, $2.5x10^6$ cells were plated on 10cm plate, one plate per cDNA construct (or gRNA library). 2 hours before transfection, medium was exchanged for STANDARD DMEM with 25µM chloroquine. Transfection reaction was prepared, incubated for 30 minutes at room temperature and added dropwise to the plate with cells and mixed gently. Plates were returned to the incubator. After 8-12 hours medium was replaced with a fresh 8 ml of STANDARD DMEM. 48 hours after transfection, culture medium containing viral particles was collected by centrifugation at 1700g for 5 minutes at room temperature. After centrifugation, supernatant was filtered through a 0.45 µm sterile filter and either used directly for transduction or split into 1-2 ml aliquots and frozen at -80°C for future use.
Composition of transfection reaction was:

Volume	Reagent
10 µg	transfer vector
6.55 µg	packaging plasmid psPAX2 (Addgene ID: 12260)
3.5 µg	envelope plasmid pMD2.G (Addgene ID: 12259)
to a final volume of 940 µl	DMEM without additives
60 µl	Fugene HD

Transfer vectors used were:

pWPLXd:IRES:PuroR for lentiviral particles used for complementation of MRM1/2 KO; wild type MRM1 and MRM2 cDNA, and mutated MRM1 cDNA were cloned into this plasmid using Gibson assembly (section 2.2.5).

pKLV-U6-esgRNA_BsmBI-stuffer-pGK-Puro-2A-BFP for lentiviral particles used for CRIPSR/Cas9 screen (gift from Paul Lehner); PCR amplified linear dsDNA oligonucleotides were cloned into this plasmid using BsmBI site (sections 2.6.1.1 2.6.1.2)

pKLV-U6-esgRNA_BsmBI-b2M-pGK-Puro-2A-BFP used for Cas9 activity assessment (gift from Paul Lehner, section 2.6.2).

2.5.7 Lentiviral transduction

For complementation of MRM1/2 KO with wild type MRM1 and MRM2, MRM1(E259A) and MRM1(S289A) following protocol was used:

Lentiviral particles were produces as describe in section 2.5.6 using pWPLXd:IRES:PuroR transfer plasmid with respective cDNA construct. MRM1/2 KO and HEK293 wild type cell line were cultured in 10 cm dish in DMEM U. Before transduction, 8 µg/ml of polybrene was added to the medium. 1-2 ml of viral supernatant was added to the cells and plates were returned to the incubator. 48 hours after transduction medium was replaced with selective medium DMEM U with 1 µg/ml puromycin.

2.6 CRISPR/Cas9 screen

2.6.1 Cloning of gRNA plasmid library

Sequences of gRNAs for selected genes were retrieved and cloning of the gRNA library (section 2.6.1.3) was performed in collaboration with Dr Ildar Gabaev.

2.6.1.1 PCR amplification of oligonucleotide pool

Oligonucleotide library containing 4490 different oligonucleotides ordered from CustomArray was resuspended in 80 µl of Tris-EDTA buffer, with a concentration of dissolved oligonucleotides being 28.49 ng/µl. First, oligonucleotides were PCR amplified in the following reaction (primer sequences are listed in Table 2.1):

PCR reaction (50 µl):

Volume	Reagent
10 µl	Q5 buffer
1 µl	dNTPs (10 mM each)
0.5 µl	Q5 DNA polymerase
2.5 µl	Primer 27 (10 µM)
2.5 µl	Primer 28 (10 µM)
1.4 µl	DNA oligo pool (40 ng template)
32.1 µl	dH2O

PCR conditions:

1 x	98°C	30 s
20 x	98°C	15 s
	62°C	20 s
	72°C	20 s
1 x	72°C	2 min
	4°C	hold

PCR product was purified using QIAquick Nucleotide Removal kit (QIAGEN) according to the manufacturer's instructions and eluted in 40 µl of EB buffer. Correct size of clean product was verified by agarose gel electrophoresis on a 2% gel prepared in TBE BUFFER.

2.6.1.2 Restriction digestion of plasmid and PCR product and ligation

10 µg of pKLV-U6-esgRNA_BsmBI-stuffer-pGK-Puro-2A-BFP plasmid was digested with 5 µl of BsmBI restriction endonuclease, 1mM DTT and 5 µl of 10x Tango buffer in a 50 µl reaction. Reaction was incubated for 4 hours at 37°C. Immediately after the digestion reaction was resolved on 1% agarose gel prepared in TBE BUFFER. Band corresponding to the linearised plasmid was excised from the gel and DNA was extracted from the gel using QIAquick gel extraction kit, according to manufacturer's instructions.

20 ng of purified PCR product from 2.6.1.1 was digested with 1 µl of BsmBI restriction endonuclease, 1mM DTT and 1 µl of 10x Tango buffer in a 10 µl reaction. Reaction was incubated for 4 hours at 37°C. Ligation of PCR product and digested plasmid was performed straight after the digestion by adding following reagents to the reaction: 1 µl of BsmBI, 1 µl of 10 mM DTT, 1 µl of 10 mM ATP, 1 µl of 10x Tango buffer, 0.5 µl of T7 DNA ligase, 100 ng of digested plasmid and water to a total reaction volume of 20 µl. This reaction was incubated in a PCR cycler at following conditions:

6 x	37°C	5 min
	22°C	5 min
	4°C	hold

Once reaction finished and cooled, this reaction was directly used for transformation by electroporation.

2.6.1.3 Electroporation of bacteria and plasmid extraction

Four transformation reactions were used for the transformation of plasmid library, resulting from ligation in section 2.6.1.2, into Stbl4 electrocompetent cells (Thermo Fisher Scientific).

25 µl cells were gently pipetted into 1.5 ml Eppendorf tubes on ice and 0.5 µl of ligation reaction was added to the cells and stirred briefly with a pipette tip. 25 µL of DNA/cell mixture was pipetted into MicroPulser electroporation cuvettes, 0.1 cm gap (Bio Rad) and cuvettes were kept on ice.

Cuvette were electroporated one at a time using the settings: 1800 V, 25 µF, 200 Ω, 1 mm cuvette.

Immediately after electroporation 975 µL SOC MEDIUM was added to the cuvette and transferred to the sterile 15 mL falcon (in a total of 2.5 ml of SOC MEDIUM), transferred to 37 °C shaking incubator and incubated for 1 hour.

After incubation media from 4 tubes was combined, centrifuged at 1950g for 5 minutes. Supernatant was removed so that 1.2 ml media was left behind. Bacteria were resuspended in the remaining media and 150 µl of suspension was pipetted per 14 cm LB AGAR plate containing 50 mg/ml ampicillin (8 plates in total). Plates were incubated in 37 °C incubator overnight

Approximately 24 hours after electroporation bacteria were harvested from the plates and plasmids purified using 3 columns of NucleoBond® Xtra Midi EF kit (Macherey-Nagel) according to the manufacturer's instructions.

2.6.2 Cas9 activity assessment

To assess Cas9 activity in cell lines that were intended for the CRISPR/Cas9 screen, Flp-in T-Rex™ HEK293 with Cas9 integrated in FRT site (for doxycycline inducible expression), either wild type or MRM1 KO, were transduced with lentiviral particles containing gRNA targeting B2M gene. Transduction was performed by spinfection. To this end, 100 µl of viral supernatant was added to the suspension of 0.5M cells in 2 ml of DMEM TET- supplemented with 8 µg/ml of polybrene in 15 ml falcon tubes. Tubes were centrifuged at 1000g for 30 minutes at room temperature. After centrifugation, cells were resuspended in the medium and plated into 6 well plates and placed in an incubator for 4-7 days. Next, cells were stained with W6/32 antibody (gift from Paul Lehner) against MHC class I and analysed by FACS.

Staining was performed according to the following protocol: Cells were transferred into an Eppendorf tube and pelleted by centrifugation at 600g, 4 minutes. Pellet was washed with 1 ml of PBS, followed by centrifugation at 600g, 4 minutes. PBS was removed by flicking the tube upside down, leaving ~50 µl of the PBS in the tube. Cells were resuspended in the remaining PBS by vortexing and W6/32 antibody was added at 1:10 dilution (5 µl) and incubated with the cells for 5 minutes in room temperature. 1ml of PBS was added to the cells and samples were centrifuged at 600g, 4 minutes. PBS was removed by inverting the tube, cells resuspended by vortexing. 2 µl of Alexa Fluor 647-conjugated goat anti-mouse secondary antibody was added to the cells and incubated with for 5 minutes in room temperature. After incubation, samples were centrifuged at 600g, 4 minutes. PBS was removed by inverting the tube, cells were resuspended in 500 µl of PBS and transferred to a FACS tube. Samples were analysed on LSR Fortessa (BD Bioscience)

flow cytometry analyser. Experimental and control samples used in Cas9 activity assessment experiments as well as expected outcomes of detected fluorescence are indicated in Table 2.6.

Table 2.6: Summary of samples used in Cas9 activity assessment experiments.
Indicated for each sample are cell line, volume of lentiviral supernatant (LV volume), presence of doxycycline, presence of puromycin and antibodies used for sample staining. Columns BFP and Alexa647 indicate expected outcome in these channels of FACS analysis. Untransduced sample with no staining was used as a control for setting the FACS and gating (sample 1). Several control samples were used: i) samples that were transduced with the virus were expected to produce BFP positive signal, since viral particles contain BFP gene (samples 3, 7, 8, 9); ii) samples that were stained with both primary and secondary antibody and were either untransduced or for cell line that does not express Cas9 were expected to produce maximal possible signal in far-red channel (samples 2, 3, 4, 5); iii) sample stained with only secondary antibody represented minimal signal of Alexa fluor 647, and therefor maximal signal reduction that could be expected after gRNA induced knockout of B2M gene (sample 6). Experimental samples were transduced with the viral supernatant and grown either in the presence (sample 7 and 9) or absence (sample 8) of doxycycline to assess whether level of leaky expression of Cas9 protein is sufficient for production of Cas9 induced gene knockout.

Sample	Cell line	LV volume	Dox	Puro	Staining	BFP	Alexa647
1	HEK293	0 µl	+	-	No	-	No
2	HEK293 no Cas9	0 µl	+	-	W6/32 + Alexa647	-	MAX
3	HEK293 no Cas9	100 µl	+	+	W6/32 + Alexa647	+	MAX
4	HEK293	0 µl	+	-	W6/32 + Alexa647	-	MAX
5	MRM1 KO	0 µl	+	-	W6/32 + Alexa647	-	MAX
6	HEK293	0 µl	+	-	Alexa647	-	min
7	HEK293	100 µl	+	+	W6/32 + Alexa647	+	min
8	HEK293	100 µl	-	+	W6/32 + Alexa647	+	min
9	MRM1 KO	100 µl	+	+	W6/32 + Alexa647	+	min

2.6.3 Multiplicity of infection test

Series of 0.1 million cells were transduced with different volumes of viral supernatant to determine volume of supernatant needed for the CRISPR/Cas9 screen in order to achieve desired multiplicity of infection. Wild type Flp-In T-Rex™ HEK293 and MRM1 KO, both expressing Cas9, were transduced with 4 µl, 20 µl, 40 µl, 80 µl and 140 µl of viral supernatant in 1 ml of DMEM TET- supplemented with 8 µg/ml of polybrene in 15 ml falcon tubes. Tubes were centrifuged at 1000g, 30 minutes at room temperature. Following centrifugation, cells were resuspended in the medium and plated into 24 well plates and placed in an incubator for 4-7 days. Fraction of BFP positive cells in each sample was analysed by flow cytometry on LSR Fortessa (BD Bioscience) and volume of supernatant necessary was calculated.

2.6.4 CRISPR/Cas9 viability screen procedure

4.55 million cells were transduced with the viral supernatant corresponding to 0.3 multiplicity of infection (MOI) with lentiviral library containing 4490 gRNAs, corresponding to 300x coverage (i.e. each gRNA was present in the infected population 300 times). Mixture of cells in DMEM U medium, viral supernatant, and 8 μg/ml polybrene in 50 ml falcon tube was centrifuged at 1000g for 30 minutes. After centrifugation cells were resuspended in the same medium and plated into T175 flask.

Two days post infection, 25 ng/ml doxycycline 1 μg/ml puromycin were added to the cells to induce Cas9 expression and select transduced cells, respectively. Before addition of puromycin small fraction of cells was seeded into a separate plate with non-selective medium and used for verification of MOI by analysis on LSR Fortessa (BD Bioscience).

Four days post infection, cell pellets corresponding to the 300x library coverage (i.e. for the library of size 4490 gRNAs and coverage 300x, each sample contained ~1.35 million cells) were saved for genomic DNA extractions. Remaining cells were split into GLUCOSE MEDIUM and GALACTOSE MEDIUM supplemented with 1 μg/ml puromycin to ensure complete selection of only transduced cells. Cells were maintained in the presence of puromycin until day 7 post infection. Cell pellets from cells cultured in GLUCOSE MEDIUM and GALACTOSE MEDIUM were collected on days 10, 14 and 17 post infection and stored at -20°C for further processing.

2.6.5 Sequencing library preparation and next-generation sequencing

Genomic DNA was isolated from cell pellets using Quick-DNA Miniprep Plus Kit (Zymo Research) according to the manufacturer's instructions. Concentration of DNA was measured on Qubit 2.0 using Qubit™ dsDNA BR Assay Kit as per manufacturer's instructions. gRNA sequences from the genomic DNA were amplified by two rounds of PCR (Figure 2.1).

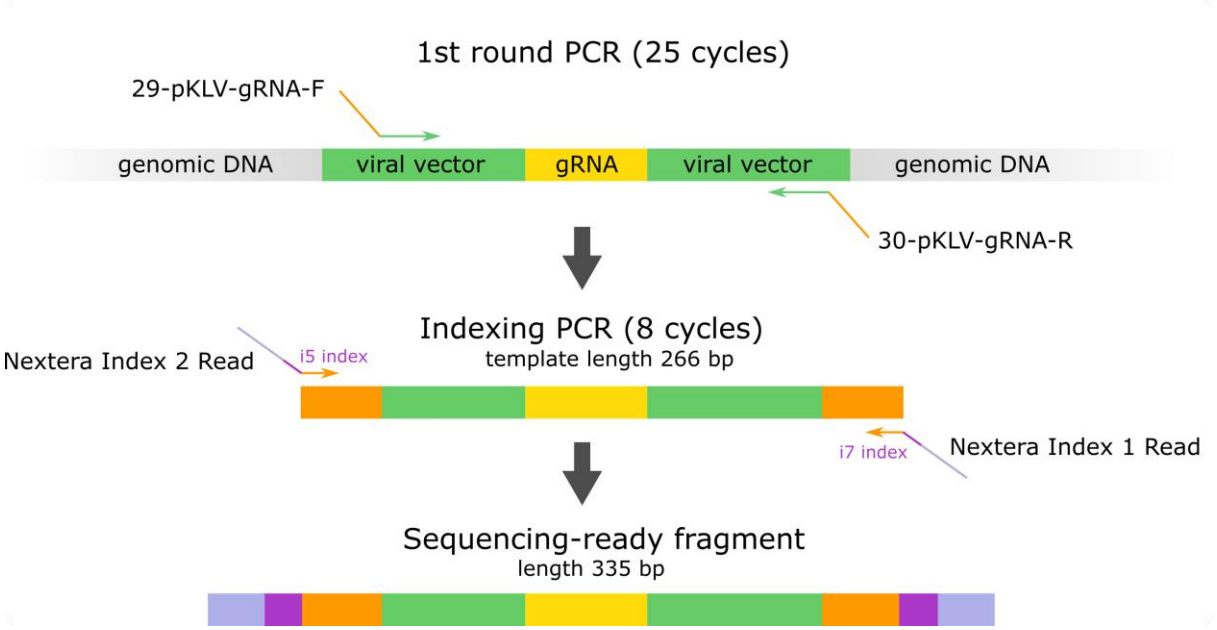


Figure 2.1: Amplification of gRNAs from genomic DNA
gRNA sequences were amplified from the genomic DNA using a pair of primers specific for viral vector sequence surrounding the gRNA integrated into the genomic DNA. Overhangs complementary to the Nextera Index Kit adapters were added by the first round of PCR, resulting in a PCR product with a length 266 bp. Second round of PCR was performed using Nextera Index Kit adaptors as primers to add sequencing indexes.

For each sample we aimed to perform the first round of PCR on a total amount of 8.9 µg of genomic DNA. This is an equivalent of a whole genome from 1.347 million diploid cells, which corresponds to the 300x coverage of gRNA library with the size of 4490 gRNAs. For some cell samples, the total amount of isolated genomic DNA was lower and actual coverage is summarised in Table 2.7. For each sample, multiple PCR reactions were performed, with ~3 µg of genomic DNA used as a template per one reaction.

PCR reaction (50 µl):

Volume	Reagent
10 µl	Q5 buffer
1 µl	dNTPs (10 mM each)
0.5 µl	Q5 DNA polymerase
2.5 µl	Primer 29 (10 µM)
2.5 µl	Primer 30 (10 µM)
-	Genomic DNA (2963.33 ng template)
Up to 50 µl	dH2O

PCR conditions:

1 x	98°C	30 s
25 x	98°C	10 s
	61°C	15 s
	72°C	20 s
1 x	72°C	2 min
	4°C	hold

PCR product was purified using QIAquick PCR Purification Kit (QIAGEN) according to the manufacturer's instructions. Concentration of purified PCR product was measured on nanodrop and diluted to a final concentration of 0.2 ng/µl. 1ng of PCR product was used as a template for second round indexing PCR with the Nextera XT Index Kit (Illumina).

PCR reaction (50 µl):

Volume	Reagent
25 µl	Phusion® Hot Start Flex 2X Master Mix
5 µl	Nextera Index Kit - Index 1 (i7) Adapter
5 µl	Nextera Index Kit - Index 2 (i5) Adapter
5 µl	1st round PCR product (0.2 ng/µl, 1 ng template)
10 µl	dH2O

PCR conditions:

1 x	95°C	1min 30 s
8 x	95°C	30 s
	55°C	30 s
	72°C	20 s
1 x	72°C	5 min
	4°C	hold

Product of indexing PCR was purified from excess primers with Ampure XP beads. Quality and concentration of the purified samples were assessed with a D1000 Screentape for TapeStation (Agilent Genomics). Samples were pooled into a final sequencing library with a 4nM concentration. Sequencing library was analysed on the Illumina MiSeq platform with 150 bp paired-end sequencing run. Paired-end sequencing was used to allow multiplexing of a higher number of samples.

Table 2.7: List of actual coverage of CRIPSR/Cas9 screen sequencing samples
To minimise technical biases, sufficient coverage of gRNA library must be maintained during the sample processing. This table summarises the coverage for each sample in each replicate in the step of PCR amplification of gRNA regions from genomic DNA. We aimed to maintain a 300x coverage for all sample. As at day 4 the puromycin selection of infected cells might not have been complete, the aim for these samples was 600x.

	Cell line:	**HEK293**		**MRM1 KO**	
	Growth medium:	**Glucose**	**Galactose**	**Glucose**	**Galactose**
Replicate 1	**Day 4**	580	NA	600	NA
	Day 10	458	300	300	221
	Day 14	304	581	281	451
	Day 17	300	196	300	413
Replicate 2	**Day 4**	600	NA	600	NA
	Day 10	286	281	300	300
	Day 14	300	300	300	300
	Day 17	300	300	300	300
Replicate 3	**Day 4**	464	NA	428	NA
	Day 10	300	300	300	300
	Day 14	300	300	300	300
	Day 17	300	300	300	300

2.6.6 Analysis of sequencing data

Read 1 FASTQ files were mapped to the gRNA library using a custom Python script to produce a count table containing number of occurrences of each gRNA in each sample. Python script for read mapping was adapted from the script written by Dr Felicity Allen. No mismatches between the read sequence and the gRNA library were allowed, achieving mapping for approximately 70% of the reads (Figure 2.2) and median sequencing coverage of 192x (Table 2.8).

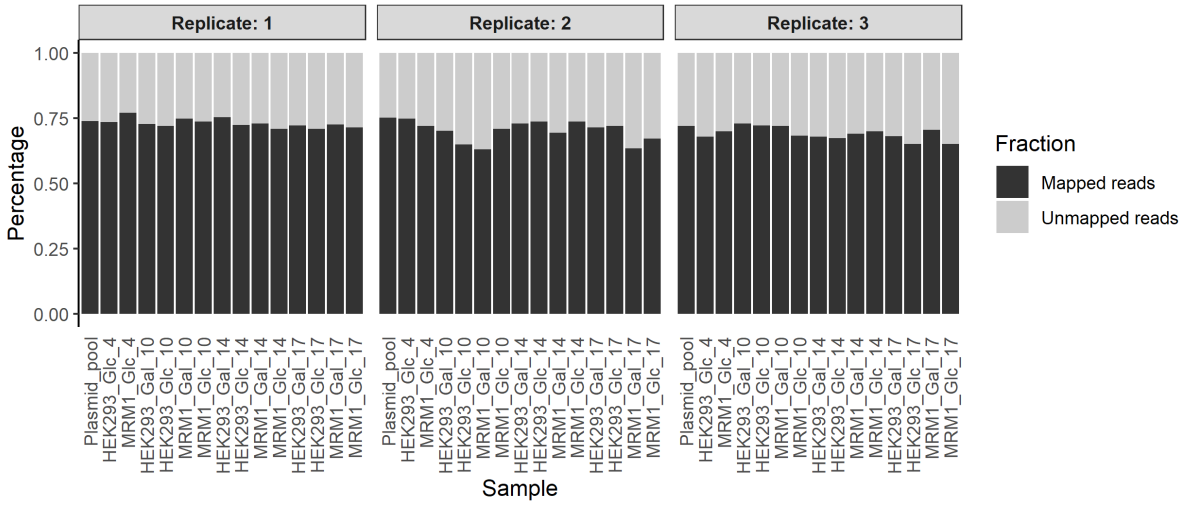


Figure 2.2: Fraction of reads that was mapped to the gRNA library
Raw reads for each sample were mapped to the gRNA library using a custom Python script. On average ~70% of reads were mapped to the reference library. Sample labels in the figure consist of cell line, growth medium and day of the sample collection from the start of the screen. "Plasmid_pool" label stands for plasmid library sample, reference sample that was used for production of lentiviral vectors.

Table 2.8: Sequencing coverage per sample per replicate
Sequencing coverage was calculated from gRNA reads that mapped to the reference gRNA library as a ratio between the number of mapped reads and the total number of gRNAs in the library. The median coverage for all samples and replicates was ~192x.

	Cell line:	**HEK293**		**MRM1 KO**	
	Growth medium:	**Glucose**	**Galactose**	**Glucose**	**Galactose**
Replicate 1	**Day 4**	225.82	NA	219.16	NA
	Day 10	185.11	355.62	188.45	213.19
	Day 14	161.08	387.57	189.65	400.79
	Day 17	188.88	400.70	201.41	379.73
Replicate 2	**Day 4**	99.36	NA	65.13	NA
	Day 10	43.33	59.26	83.53	76.11
	Day 14	106.06	103.98	134.93	60.56
	Day 17	85.03	82.44	79.20	74.70
Replicate 3	**Day 4**	200.40	NA	201.12	NA
	Day 10	217.65	168.41	216.99	196.97
	Day 14	226.30	195.16	212.75	196.83
	Day 17	180.54	192.96	193.94	200.06

For evaluation of differences between "adherent" and "suspension" populations gRNA read counts were normalised to adjust for differences in sequencing depth between the two compared samples (see Table 2.9). To normalise the read counts, read count for individual gRNA was divided by the median of reads per gRNA for genes belonging to the negative control category for a given sample. For each pair of samples (e.g. HEK293, day 10, suspension cells and HEK293, day 10, adherent cells), the count value for each gRNA was then multiplied by the average of the medians for negative control genes from these two samples. Reads were normalised in the same way for comparison of correlation between replicates.

Table 2.9: Normalisation of gRNA read counts

PRE-NORMALISATION

Sample	number of reads	total reads	% of total
HEK293 Day10 suspension	629883	1178671	53.44%
HEK293 Day10 adherent	548788		46.56%
HEK293 Day14 suspension	776024	1252446	61.96%
HEK293 Day14 adherent	476422		38.04%
HEK293 Day17 suspension	598768	1338744	44.73%
HEK293 Day17 adherent	739976		55.27%
MRM1 KO Day14 suspension	630219	1319584	47.76%
MRM1 KO Day14 adherent	689365		52.24%
MRM1 KO Day17 suspension	610536	1227759	49.73%
MRM1 KO Day17 adherent	617223		50.27%

POST-NORMALISATION

Sample	number of reads	total reads	% of total
HEK293 Day10 suspension	613160	1177335	52.08%
HEK293 Day10 adherent	564175		47.92%
HEK293 Day14 suspension	631118	1249528	50.51%
HEK293 Day14 adherent	618410		49.49%
HEK293 Day17 suspension	676630	1340242	50.49%
HEK293 Day17 adherent	663612		49.51%
MRM1 KO Day14 suspension	648892	1318974	49.20%
MRM1 KO Day14 adherent	670082		50.80%
MRM1 KO Day17 suspension	616086	1227798	50.18%
MRM1 KO Day17 adherent	611712		49.82%

Gene essentiality scores were calculated with the JACKS analysis tool using gRNA count tables as an input (Allen et al., 2019). Using JACKS, a log2 fold-change of each gene was calculated at days 4, 10, 14 and 17 compared to the plasmid pool, separately for each cell line and for each condition. Genetic interactions were identified based on calculation of a "diffScore" as a difference in gene essentiality in HEK293 and MRM1 KO cell lines. To identify genes with significantly stronger effect in MRM1 KO, normal distribution was fitted to the diffScore values of genes from negative control category. P-values for each gene were calculated as the area under the curve for more extreme values of diffScore

compared to the gene. Significance threshold of 0.01 was used and a Bonferroni correction applied to account for multiple testing.

2.7 RNA extraction, DNase treatment and cDNA synthesis

Cells from a confluent T75 flask were pelleted and resuspended in 1 ml of Ambion TRIzol™ reagent. Samples were either stored at -20°C or used for RNA extraction directly. RNA was extracted from the samples according to the manufacturer's protocol and stored at -80°C. For DNAse treatment, 4 µg of DNA in 8 µl of water were mixed with 1 µl of DNase buffer (Invitrogen) and 1 µl of DNase (Invitrogen) and incubated at 37°C for 1 hour. DNase was blocked with 1 µl of DNase Inactivation reagent (Invitrogen) and sample was incubated for 2 minutes at room temperature followed by centrifugation at 10 000 x g for 2 minutes. 5 µl were transferred into a new tube and used for cDNA synthesis. cDNA synthesis was performed using Omniscript Reverse Transcriptase (Qiagen) according to the manufacturer's instructions. Total reaction volume was 20 µl and random hexanucleotides were used as a primer. After cDNA synthesis the sample was diluted with 180 µl of water.

2.8 RT-qPCR

Total cellular DNA was extracted from the cells using DNeasy Blood and Tissue kit (QIAGEN) as per manufacturer's instruction. cDNA was prepared according to section 2.7. mtDNA copy number was quantified by measuring mtDNA amplicon relative to a nuclear control gene (B2M). RNA transcript levels of mt-CO1, mt-CO2, mt-ND1, 12S and 16S RNA were quantified relative to the GAPDH gene. qPCR was performed in biological triplicates, and each replicate was performed in technical triplicates using the TaqMan gene expression master mix with 100 ng of DNA or 5 µl of cDNA per sample as a template, 1.125µM of primers and 250 nM probe (Table 2.10).

Table 2.10: Sequences of RT-qPCR primers and probes

Primer/Probe name	Sequence
mtDNA forward	CACCCAAGAACAGGGTTTGT
mtDNA reverse	TGGCCATGGGTATGTTGTTAA
mtDNA probe	6-Fam-TTACCGGGCTCTGCCATCT-Tamra
B2M forward	TGCTGTCTCCATGTTTGATGTATCT
B2M reverse	TCTCTGCTCCCCACCTCTAAGT
B2M probe	6-Fam-TTGCTCCACAGGTAGCTCTAGGAGG-Tamra
GAPDH forward	GAAGGTGAAGGTCGGAGTCAAC
GAPDH reverse	CAGAGTTAAAAGCAGCCCTGGT
GAPDH probe	6-Fam-GTTTGGTCCGTATTGGGCGCCT-Tamra
mt-CO1 forward	CGATGCATACACCACATGAA
mt-CO1 reverse	AGCGAAGGCTTCTCAAATCA
mt-CO1 probe	6-Fam-GGCTCATTCATTTCTCTAACAGC-Tamra
mt-CO2 forward	CGTCTGAACTATCCTGCCCG
mt-CO2 reverse	TGGTAAGGGAGGGATCGTTG
mt-CO2 probe	6-Fam-CGCCCTCCCATCCCTACGCATC-Tamra
mt-ND1 forward	CCCTAAAACCCGCCACATCT
mt-ND1 reverse	AGAGCGATGGTGAGAGCTAAGG
mt-ND1 probe	6-Fam-ATCACCCTCTACATCACCGCCCCG-Tamra
16S rRNA forward	TTTGCAAGGAGAGCCAAAGC
16S rRNA reverse	AGACGGGTGTGCTCTTTTAGCT
16S rRNA probe	6-Fam-AGACCCCCGAAACCAGACGAGCTACC-Tamra
12S rRNA forward	CCCCAGGTTGGTCAATTTC
12S rRNA reverse	CGGCTTCTATGACTTGGGTTAA
12S rRNA probe	6-Fam-TGCAGCCACCGCGGTCA-Tamra

2.9 Mitochondrial isolation

Cells were harvested from the tissue culture flask with the use of trypsin and pelleted by centrifugation at 300g for 5 min. Samples were kept at 4°C for all subsequent steps. Pellets were resuspended in 3 ml of HYPOTONIC BUFFER per 1 g of wet cell mass and incubated on ice for 10 minutes. Balch homogeniser was washed with HYPOTONIC BUFFER containing the ball bearing with 12 µm clearance. The cells were passed through the homogeniser 5 times, followed by addition of MSH (2.5X), 2 ml of MSH (2.5X) per 3 mL of cells. The final cell suspension was diluted in 50 ml of MSH (1X). The suspension was centrifuged twice at 1000 g for 10 minutes. The pellet was discarded after each centrifugation. The supernatant was transferred into a new tube and centrifuged two times at 10,000g for 10 minutes. The two pellets were pooled and resuspended in MSH (1X). This sample was

loaded onto a STEP GRADIENT BUFFER. The tube was centrifuged at 26,000 rpm for 60 minutes in a SW40Ti rotor (tube number: 344060). The mitochondria formed a band in the interface between the 1 M and 1.5 M solutions. Mitochondria were purified from sucrose by dilution in 4 volumes of MSH (1X) and centrifuged at 10,000 g for 10 minutes at 4°C. The pellet was the isolated mitochondria.

2.10 SILAC proteomics analysis

Wild type and knock out cells were reciprocally labelled in heavy and light SILAC MEDIUM for 6 generations. After full labelling was achieved cells were harvested and pelleted by centrifugation at 500 g for 3 minutes and washed twice with PBS. Cell pellets were resuspended in 10 ml of PBS and quantified by the Pierce BCA protein assay kit as per manufacturer's instructions. Small aliquot of heavy cells was saved for verification of isotope incorporation. Equal proportion of cells was mixed based on the protein concentration. The cells were pelleted, and mitochondria were isolated as per section 2.9. Isolated mitochondria were lysed with 100-150 µl of GRADIENT LYSIS BUFFER and incubated on a roller at 4 °C for 15 minutes. Lysates were clarified by centrifugation at 10,000 g for 5 minutes at 4 °C. Protein concentration of the lysate was quantified by the Pierce BCA protein assay kit as per manufacturer's instructions. 20 µg of proteins were resolved on 4-12 % SDS-PAGE as per section 2.3. Each gel lane was cut into 10 pieces of 1x1 cm each (Figure 2.3) and analysed by LC-MS^2 by MBU Mass Spectrometry facility (see section 2.11).

Raw mass spectrometry files were analysed in MaxQuant, by comparing the proteins to the human UniProt database. For each protein, a ratio of heavy to light peptides was calculated. Quantification was reported only for protein for which at least two heavy-to-light ratios could be produced. Data were first processed in Perseus software where proteins only identified by site, that matched a decoy database of random peptides deliberately introduced to identify noise, and those peptides identified as contaminants were removed. Values were reversed, so that all ratios represent wild type to knockout protein amount, and log2 transformed. Visualisations were created in R.

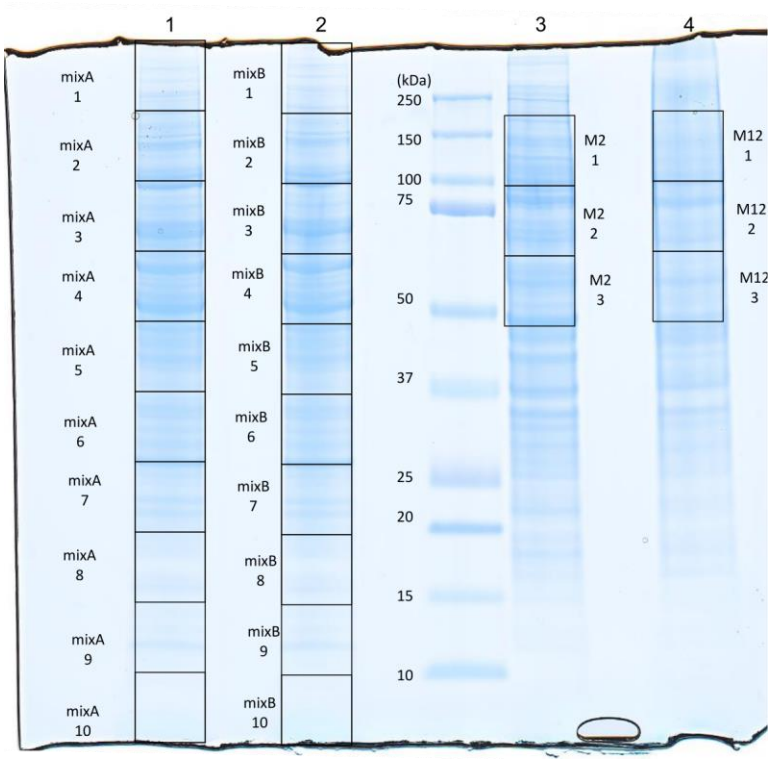


Figure 2.3: Gel sections for mass spectrometry analysis
~20 µg of proteins were resolved on 4-12 % Bis-Tris NuPAGE polyacrylamide gels. Gel lanes for SILAC proteomics were cut into 10 equal pieces of 1x1 cm each (lanes 1,2). To verify complete incorporation of heavy isotope, 3 gel sections were analysed from heavy-labelled sample for each experiment (lanes 3,4) As a marker od molecular weight 5µl of Precision Plus Protein™ Kaleidoscope™ Prestained Protein Standards was used.

2.11 Mass spectrometry analysis

Mass spectrometry analysis was performed by the MBU Mass Spectrometry facility. Samples were analysed on Thermo Q Exactive Plus coupled with a Proxeon EASY-nLC 1000 chromatography. Peptides were separated on 2 µm, 100 Å Acclaim PepMap C18 (50 µm x 150 mm, 300 nl/min) using a 84 min gradient of 5% to 40% acetonitrile with 0.1% formic acid, followed by an increase in acetonitrile concentration to 90% and re-equilibration with 5% acetonitrile, within a 105 minute period. Peptides were analysed by positive ion electrospray mass spectrometry using a method programmed to fragment the top 10 most abundant multiply charged peptide ions each second. Full scan MS data (400 to 1800 m/z) were recorded at a resolution of 70,000 with an automatic gain control (AGC) target of 1×10^6 ions and a maximum ion transfer of 20 ms. Ions selected for MS/MS were analysed using the following parameters: resolution 17,500; AGC target of 5×10^4; maximum ion transfer of 100 ms; 2 m/z isolation window; for HCD a normalised collision energy 28% was used; and dynamic exclusion of 30 s. A lock mass ion (polysiloxane, m/z = 445.1200) was used for internal MS calibration.

2.12 Analysis of mitoribosomes

2.12.1 Sucrose gradient fractionation

The 10-30% linear sucrose gradient in CONTINUOUS GRADIENT BUFFER was prepared in 11 x 35 mm thin wall polypropylene ultracentrifuge tubes using the Biocomp Gradient station. Whole cell or mitochondrial lysate was loaded onto the gradient and centrifuged for 2 hours 15 minutes at 100,000 g at 4°C (Beckman Coulter TLS-55 rotor). Equal fractions of 100 µl each were taken from the top of the gradient. Typically, 7 µl from each fraction was used for western blot, and 30-35 µl for qDGMS.

2.12.2 Quantitative density gradient analysis by mass spectrometry (qDGMS)

qDGMS analysis has been performed according to the protocol published in Páleníková et al. (2021a). Briefly, wild type and knockout cells were reciprocally labelled in SILAC MEDIUM and processed as described in section 2.10 up to mitochondrial lysis. At this point, instead of being resolved on SDS-PAGE, mitochondrial lysate was resolved on linear sucrose gradient as per section 2.12.1. Liquid samples were precipitated in 20 volumes of cold absolute ethanol overnight at -20 ºC. After precipitation samples were centrifuged at 4ºC. Protein pellet was digested in 1% (w/w) trypsin in 50 mM NH_4HCO_3 overnight at 37ºC and analysed by LC-MS^2 by MBU Mass Spectrometry facility.

Peptide data were analysed using the Proteome Discoverer™ (v 1.4) in combination with MASCOT database. Report of all peptides from all fractions was exported as a single tab delimited file and used as an input for ComPrAn. ComPrAn is described in chapter 3.

2.13 35S labelling

HEK293 cells were grown in 6 well plate to 80% confluency. Before starting the labelling assay, cells were washed twice with methionine and cysteine-free medium and incubated with this medium for 5 minutes at 37°C. The medium was replaced with DMEM supplemented with 2 mM Glutamax, 110 mg/L sodium pyruvate, 96 µg/ml cysteine and 5% dialysed FBS and the samples were incubated for 20 minutes at 37°C. 30 minutes before the assay, cytoplasmic translation was inhibited by addition of 100 µg/ml emetine. 166.6 µCi/ml of [35S]-L-methionine was added to the cells and incubated for 30 min at 37°C. The cells were harvested and centrifuged at 3450g for 3 min. The pellet was washed

with 1 x PBS and resuspended in 20 µl of PBS containing protease inhibitor. 0.1 % DDM and 50 U Benzonase was added, followed by incubation of samples on ice for 20 minutes. 1% SDS was added, the samples were vortexed vigorously and left on ice for 30 minutes. The proteins were quantified using the BioRAD DC assay kit as per manufacturer's instructions. 30 µg of protein was loaded on a 10-20% Tris-glycine gradient SDS polyacrylamide gel and run at 85V for 3 hours. The gel was washed with water, stained with SimplyBlue Safestain for 1 hour, destained with coomassie destain/fixer (20% methanol, 7% acetic acid, 3% glycerol) for 1 hour and dried using a Bio-Rad Model 583 gel dryer at 80 °C for 1 hour. The gel was visualised via autoradiography using the PhosphorScreen and Typhoon 9410 scanner.

3 Development of Complexome Profiling Analysis (ComPrAn) R package

3.1 Introduction

3.1.1 Complexome profiling and qDGMS

Many proteins in cells function as part of larger complexes. These complexes vary in size, ranging from small homodimers to large molecular machines such as ribosome. They can be composed of proteins only or contain other macromolecules like nucleic acids. Multiprotein complexes play an important role in the function of mitochondria. Inner mitochondrial membrane contains five oxidative phosphorylation (OXPHOS) complexes, that provide vital energy for cell survival. In humans, their size ranges from 4 subunits in the smallest Complex II to 45 subunits in the largest Complex I. Furthermore, individual complexes of the respiratory chain can join together to form supercomplexes (Letts and Sazanov, 2017; Schägger and Pfeiffer, 2000). Mitochondrial ribosome is an example of ribonucleoprotein complex functioning in mitochondria. Given the functional importance of protein complexes, it is desirable to be able to study their composition and assembly pathways. However, complexes are held together by weak non-covalent interactions, which can be disrupted during analysis of biological samples. Several methods have been developed that aim to either stabilise the interaction or process the sample in a near-native state, in order to allow investigation of a protein complex composition.

Complexome profiling is a method that aims to interrogate macromolecular complexes close to their native state. It combines separation of protein complexes in a non-denaturing conditions, typically by a blue-native polyacrylamide gel electrophoresis (BN-PAGE), with systematic analysis of entire gel lane by mass spectrometry (Giese et al., 2021; Heide et al., 2012; Wessels et al., 2009). As a result of such experiment, a protein amount in each fraction of the gel is reported to generate a migration profile of the protein. Proteins whose migration profiles peak in the same fraction are likely to be part of the same protein complex. Complexome profiling has been successfully applied in the study of

several mitochondrial OXPHOS complexes (Guerrero-Castillo et al., 2017; Heide et al., 2012; Protasoni et al., 2020; Senkler et al., 2017, 2018; Vidoni et al., 2017; Wessels et al., 2013). Since its development, several modifications to the method have been implemented, including cryo-slicing blue-native mass spectrometry approach to complexome profiling to achieve high resolution (Müller et al., 2016), use of alternative complex separation methods such as size exclusion chromatography (Havugimana et al., 2012; Larance et al., 2016) or addition of labelling to allow sample multiplexing (Guerrero-Castillo et al., 2021; Páleníková et al., 2021b; Protasoni et al., 2020; Vidoni et al., 2017). We recently developed and optimised **q**uantitative **D**ensity **G**radient analysis by **M**ass **S**pectrometry (qDGMS), which combines stable isotope labelling with amino acids in cell culture (SILAC) and density gradient ultracentrifugation, followed by fractionation and analysis by mass spectrometry (Figure 3.1, for details see Páleníková et al., 2021a). This technique has been applied to study mitoribosomes in several cell lines (D'Souza et al., 2021; Van Haute et al., 2019).

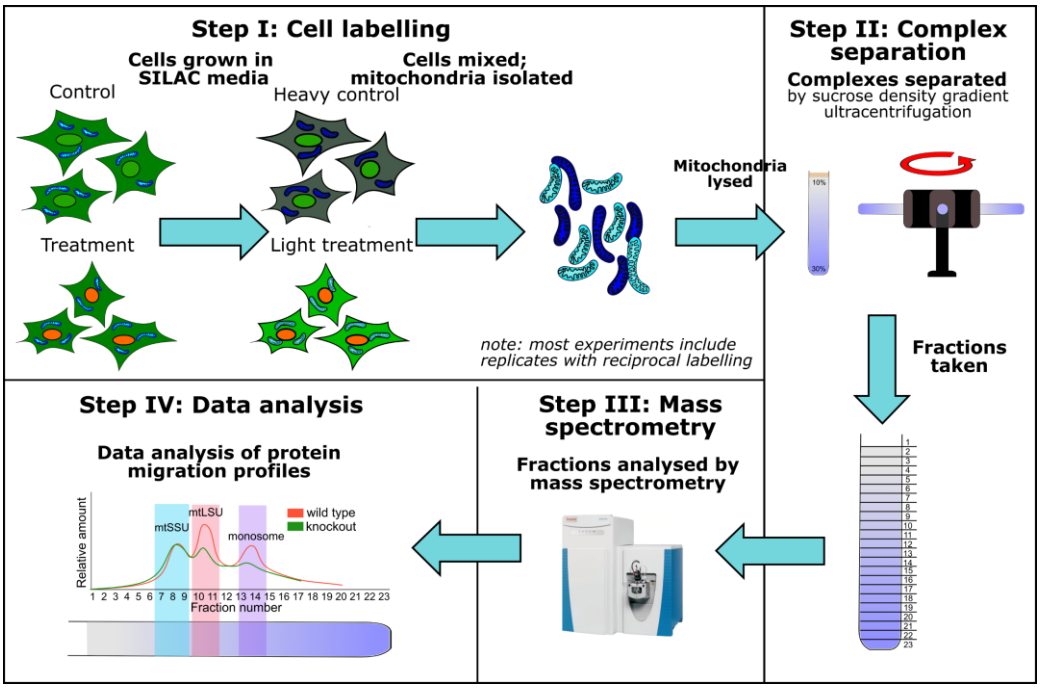


Figure 3.1: Schematic of a qDGMS experiment
There are 4 main steps in a typical qDGMS experiment. First, studied cell lines are reciprocally labelled with amino acids in cell culture (SILAC) when one cell line (e.g. wild type control) is grown in the presence of arginine and lysine containing heavy ^{13}C and ^{15}N isotopes and second cell line is grown in a medium with light ^{12}C and ^{14}N isotopes containing amino acids. Following labelling, two cell lines are mixed, mitochondria isolated and lysed. Second, protein complexes from mitochondrial lysate are resolved by sucrose density gradient ultracentrifugation. Fractions of equal volume are collected from the gradient post centrifugation. Third, liquid samples are precipitated and analysed by liquid chromatography-mass spectrometry. As a fourth step, mass spectrometry data are processed to produce migration profiles of proteins through the sucrose gradient.

3.1.2 Challenges in analysis of complexome profiling data

Proteomics techniques, such as complexome profiling, produce large amount of data and require specialised software for its analysis. Data analysis of complexome profiling experiments poses several challenges. Some are inherent to all proteomics data, such as that mass spectrometry provides data for peptides rather than proteins, whereas others are specific for complexome profiling. In proteomics experiments, the aim is typically to identify and/or quantify proteins from samples independently. However, complexome profiling aims to consecutively analyse a set of samples to generate a continuous migration profile throughout the blue-native gel lane or the sucrose gradient. A common approach to estimate protein abundance is to use the average of the three peptides with the highest ion counts. This can be problematic to achieve in complexome profiling in situations when different set of peptides is detected in different fractions.

Quantification in mass spectrometry is not straightforward. Detected amount of a peptide is not solely dependent on the quantity of protein from which it originated in the studied sample but is also dependent on factors such as peptide sequence, modifications, charge state, ionisability or fragmentation efficiency. Therefore, protein amount estimate might differ based on peptides used for its representation. When the aim is to compare multiple samples, labelling methods, such as SILAC (Ong and Mann, 2007), tandem-mass tags (TMT) (Thompson et al., 2003), isotope-coded affinity tags (ICAT) (Gygi et al., 1999) or isobaric tags for relative and absolute quantification (iTRAQ) (Wiese et al., 2007) can be used. As a metabolic labelling technique, SILAC provides an advantage since it allows mixing of samples at an early stage of the experiment, mitigating several technical biases that can be introduced when samples are processed separately. Typically, SILAC labelled samples are analysed in a software such as MaxQuant that outputs ratios of heavy to light samples. However, in qDGMS and SILAC complexome profiling, reporting only protein ratios is not desired since it would not allow to construct a complexome profile for all proteins within each sample. This approach would also lead to a potential loss of information about proteins that were detected in only one label state. Therefore, we devised a strategy where for each protein a same representative peptide is selected across all fractions and a chromatographic peak area is used as a proxy of the protein amount.

3.1.3 Data analysis tools for complexome profiling

Since introduction of complexome profiling, several tools have been developed for analysis and visualisation of data produced by this type of experiments. ComplexomeMap integrates results from investigation of the mitochondrial complexes of *Arabidopsis thaliana* (Senkler et al., 2017). NOVA software allows visualisation and cluster analysis of complexome profiles (Giese et al., 2015). Its clustering algorithms can be combined with different distance measures, and visualisations include 2D plots and heatmaps. ComplexFinder is a Python-based pipeline that uses machine learning to identify protein complexes in a peak-centric way (Nolte and Langer, 2021). Complexome Profiling Alignment (COPAL) tool has been developed to allow *in silico* correction for differences in protein migration, gel slicing process and total number of slices among experiments (Van Strien et al., 2019). This alignment enables comparison of protein migration between different samples and identification of most affected protein complexes. Results produced by complexome profiling experiments are a rich data resource, as typically information about hundreds to thousands of proteins is generated in each experiment. Original study usually focuses on only a narrow subset of produced data leaving a large portion of information unexplored. To facilitate reuse of such data, a ComplexomE profiling DAta Resource (CEDAR) has been set up recently (Van Strien et al., 2021).

At the point of development of qDGMS, to the best of our knowledge, there was no publicly available tool for analysis of the mass spectrometry data produced by this method. Studies that employed SILAC complexome profiling were analysed with custom scripts, that combined several software tools and required significant manual input from the user (Páleníková et al., 2021b). Therefore, we set to implement an easy-to-use software that will make the study of native protein complexes accessible to a broader audience. Here we present **Com**plexome **Pr**ofiling **An**alysis (ComPrAn) R package for analysis of SILAC complexome profiling data. ComPrAn uses peptide data identified and quantified by widely used proteomics software (e.g. the peptide search engine Mascot and the mass spectrometry pipeline application Proteome Discoverer, respectively) and includes tools for estimation of protein abundance from peptide chromatographic peak area, the normalisation of such protein abundance estimates, cluster analysis of protein abundance profiles and visualisation tools. The package is accompanied by a user-friendly graphical interface. Although ComPrAn has been developed for analysis of qDGMS experiments, it can also be used for any other type of protein complex chromatography methods that fractionate 2-state labelled samples.

Table 3.1: Comparison of complexome profiling software tools

Name	Description	Implementation	Visualisation functions	Graphical user interface	Reference
ComplexomeMap	Heatmap visualisation of complexome profiling data with searching and filtering functions	Webpage (HTML and Java Script)	+	N/A	(Senkler et al., 2017)
NOVA	Analysis and visualisation of protein level complexome profiling data with hierarchical clustering functions	Java	+	+	(Giese et al., 2015)
ComplexFinder	Analysis of protein level complexome profiling data that can predict novel protein-protein interactions. Profiles are represented by peak models allowing calculation of local profile similarities	Python	-	-	(Nolte and Langer, 2021)
COPAL	Alignment and normalisation of protein level complexome profiles	Python	-	+	(Van Strien et al., 2019)
CEDAR	Database of complexome profiling data in standardised format	Webpage (REST API)	+	N/A	(Van Strien et al., 2021)

3.2 ComPrAn – design and implementation

ComPrAn has been created as a dedicated tool for analysis of qDGMS and other complexome profiling experiments that employ two state labelling. It is written in R and contains a graphical user interface implemented with the use of Shiny R package (Chang et al., 2020). The workflow of ComPrAn is divided into two parts: (i) "peptide-to-protein" (Figure 3.2, blue box), which uses peptide-level data to estimate protein abundance and to normalise those data by scaling values between 0 and 1 for each identified protein, (ii) "protein workflow" (Figure 3.2, orange box), which allows for clustering and visualising of normalised data.

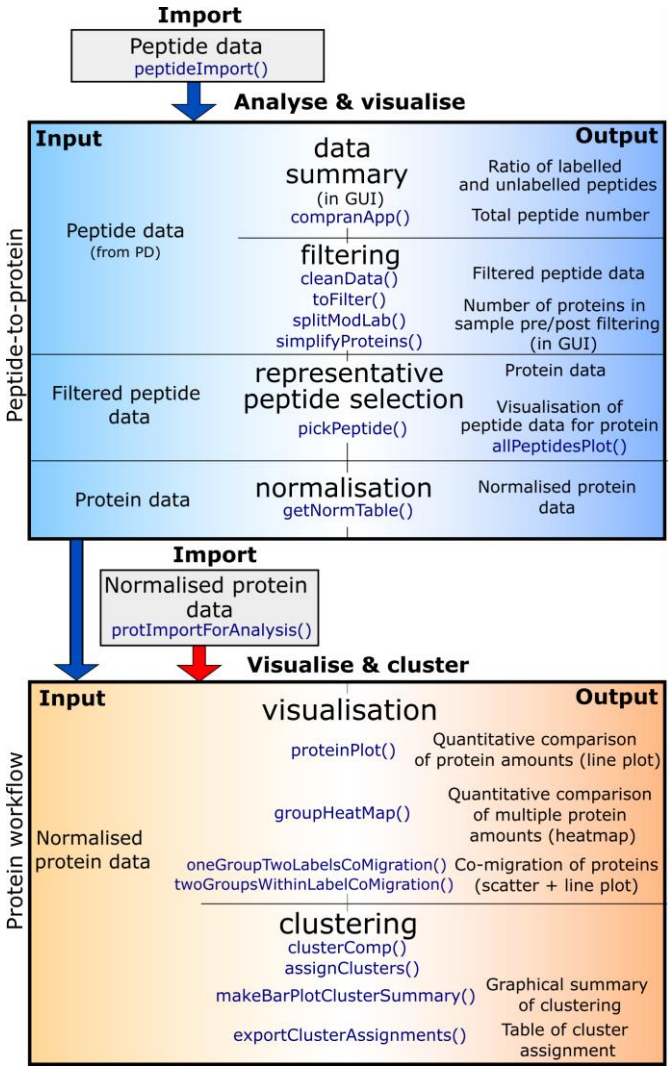


Figure 3.2: Schematic of ComPrAn workflow
ComPrAn functionality (middle), with input (left) and output (right) of each step performed for peptide and normalised protein data inputs. Either a peptide-level data or normalised protein data can be used as an input. ComPrAn functions are shown in blue. GUI = graphical user interface, PD = Proteome Discoverer™.

3.2.1 Installation

Current stable version of ComPrAn (v 1.0.0) is available on Bioconductor for R version 4.1 (https://bioconductor.org/packages/ComPrAn).

It can be installed by running the following code:

```
if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")

BiocManager::install("ComPrAn")
```

Development version of ComPrAn is available from Bioconductor and can be installed by running:

```
if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")

# The following initializes usage of Bioc devel
BiocManager::install(version='devel')

BiocManager::install("ComPrAn")
```

Source code of development version is available on GitHub. Package can be installed from GitHub by the following code:

```
# If devtools package is not already installed, install it from CRAN
install.packages("devtools")

# Install ComPrAn from GitHub
devtools::install_github("Scavetta/ComPrAn")
```

3.2.2 Data inputs

ComPrAn is not intended for processing of raw mass spectrometry data. Raw data first need to be interpreted in a software that can provide identity and quantity of each peptide. ComPrAn input files need to be in a tab delimited plain text format. Two types of input data can be used: i) peptide data or ii) normalised protein data. Example peptide data set present in the package was processed with Proteome Discoverer™ version 1.4 (Thermo Fisher Scientific) in combination with MASCOT database for peptide quantification and identification, respectively. Example normalised protein data set was produced by ComPrAn's "peptide-to-protein" pipeline. These data sets represent a subset of data from qDGMS Experiment A presented in section 5.2.5 of this thesis.

Peptide input file needs to be a text file with tab delimited values. Data are organised in a table format; first row contain column headers and each following row contains

information about a single peptide (Figure 3.3). Following columns must be present in the file, format of values of each column is also specified:

"**PSM Ambiguity**" – character/factor

"**Precursor Area**" – numeric

"**# Protein Groups**" – integer

"**Rank**" – integer

"**Confidence Level**" – character/factor

"**Protein Group Accessions**" – character/factor

"**Protein Descriptions**" – character

"**Modifications**" – character

"**Charge**" – integer

"**Sequence**" – character

"**Search ID**" – character (A-Z, AA-ZZ, AAA-ZZZ…; will be converted to integer)

Built-in peptide file can be accessed by:

```
# load ComPrAn package
library(ComPrAn)
# locate peptide data set
inputFile <- system.file("extData", "data.txt", package = "ComPrAn")
# read in example peptide data set using designated
# ComPrAn import function
peptides <- peptideImport(inputFile)
```

Normalised protein data need to be supplied in plain text file with tab delimited values. Data are organised in a table format; first row contain column headers and following rows contain information about proteins, one protein per row (Figure 3.4). This file must contain following columns, format of values of each column is also specified:

"**Protein Group Accessions**" – character/factor

"**Protein Descriptions**" – character

"**scenario**" – character/factor

"**label**" – logical

columns "**1**" to "**n**" – numeric, n represents total number of fractions in sample

Built-in normalised protein file can be accessed by:

```
# locate normalised protein data set

inputFile <- system.file("extData", "dataNormProts.txt",
package = "ComPrAn")

# read in file using dedicated ComPrAn function, this function
# automatically changes the structure of the table to the
# required format

forAnalysis <- protImportForAnalysis(inputFile)
```

Detailed description of input file columns can be found in the ComPrAn package vignette:

```
# open fileFormats vignette of ComPrAn package

vignette(topic = "fileFormats", package = "ComPrAn")
```

Confidence Level	Search ID	Sequence	PSM Ambiguity	Protein Descriptions	# Protein Groups	Protein Group Accessions	Modifications	Rank	Precursor Area	Charge
High	A	LNDFASTVR	Unambiguous	Cytochrome c oxida	1	P20674		1		2
High	A	LNDFASTVR	Unambiguous	Cytochrome c oxida	1	P20674		1	5573000	2
High	A	LNDFASTVR	Unambiguous	Cytochrome c oxida	1	P20674		1		2
High	A	LNDFASTVr	Unambiguous	Cytochrome c oxida	1	P20674	R9(Label:13C(6)15l	1	2980000	2
Middle	A	LNDFASTVr	Unambiguous	Cytochrome c oxida	1	P20674	R9(Label:13C(6)15l	1		2
Middle	A	LNDFASTVr	Unambiguous	Cytochrome c oxida	1	P20674	R9(Label:13C(6)15l	1		2
Low	A	DLILPTIQk	Unconsidered	Translational activa	1	Q92616	K9(Label:13C(6)15l	1		2
Low	A	DLILPTIQk	Unconsidered	Translational activa	1	Q92616	K9(Label:13C(6)15l	1		2
Low	A	DLILPTIQk	Unconsidered	Translational activa	1	Q92616	K9(Label:13C(6)15l	1		2
Low	A	DLILPTIQk	Unconsidered	Translational activa	1	Q92616	K9(Label:13C(6)15l	1		2
Low	A	DLILPTIQk	Unconsidered	Translational activa	1	Q92616	K9(Label:13C(6)15l	1		2
High	A	aEDMETKIK	Unambiguous	Cytochrome c oxida	1	P14854	N-Term(Acetyl)	1		2
High	A	aEDMETKIK	Unambiguous	Cytochrome c oxida	1	P14854	N-Term(Acetyl)	1		2
Low	A	aEDMETKIK	Unconsidered	Cytochrome c oxida	1	P14854	N-Term(Acetyl)	1		2

Figure 3.3: Example of minimal peptide input data
Peptide data table containing all necessary columns. In case additional columns are present in the peptide data file, these will be filtered during the processing.

Protein Group Accessions	Protein Descriptions	scenario	label	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
O00468	Agrin OS=Hor	A	TRUE	0	0	0	0.62	0	1	0.7	0.49	0.31	0.09	0.04	0.02	0.01	0.03	0	0.01	0.02	0	0	0	0.01	0.02	0.11
O00483	NADH dehyd	A	TRUE	0	0	0	0	0	0	0	0.47	0.56	0	0	0	0	0	0	0	0.24	0	0.12	0	0.17	0.38	1
O15145	Actin-related	A	TRUE	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O15457	MutS protein	A	TRUE	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O43291	Kunitz-type p	A	TRUE	0	0	0	1	0	0.81	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O43715	TP53-regulat	A	TRUE	0	0	0	0	1	0.47	0	0.28	0.33	0	0	0	0	0	0	0	0	0	0	0	0	0.09	0.86
O43819	Protein SCO2	A	TRUE	0	0	0	0.6	1	0.88	0	0.56	0.32	0.22	0.22	0	0	0.26	0	0.13	0.24	0.11	0.17	0.24	0.15	0.3	0.23
O60292	Signal-induce	A	TRUE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.85	0	0	0.62	0.89	1	0.64	0	0
O60318	80 kDa MCM3	A	TRUE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.67	0.72	0.8	0.49	0	1	0.59	0.87	0
O75394	39S ribosoma	A	TRUE	0	0	0	0	0	0	0	0	0	1	0.89	0.53	0	0	0	0	0	0	0	0	0	0	0
O75947	ATP synthase	A	TRUE	0	0	0	0.39	0.71	1	0.62	0.71	0.42	0.37	0.42	0.28	0.14	0.28	0.14	0.14	0.17	0.09	0.17	0.19	0.11	0.22	0.18
O75962	Triple functic	A	TRUE	0	0	0	0	0	0	0	0	0	1	0.27	0.74	0	0	0.29	0	0	0	0	0	0	0	0
O94973	AP-2 comple	A	TRUE	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O95755	Ras-related p	A	TRUE	0	0	0	0	0	0	0	0	0	0	0	0.78	0	1	0	0	0	0	0	0	0	0	0
O95865	N(G),N(G)-di	A	TRUE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.07	0	0	1

Figure 3.4: Example of normalised protein input data
In normalised protein data table, each row represents values for one protein. Columns 1 to n (n being the total number of fractions in the experiment), contain numeric values representing relative protein amount in a given fraction, that has been scaled between 0 and 1.

3.2.3 Estimation of protein abundance from peptide abundance

Mass spectrometry analysis can only provide information about peptides and this information is used during data processing to estimate protein abundance values. In qDGMS and other SILAC complexome profiling experiments, the aim is to characterise migration of protein complexes through sucrose gradient and between two studied samples from both qualitative and quantitative perspective. This is taken into consideration when protein abundance is estimated from peptide abundance in ComPrAn. The approach to protein abundance estimation used in ComPrAn is based on the principles used for the analysis of published SILAC complexome profiling experiments (Protasoni et al., 2020; Vidoni et al., 2017). Peptides are divided into "heavy" and "light" batches, based on their isotopic labelling status. For the purpose of ComPrAn data processing, peptides need to match in their sequence, charge state and modifications to be considered the same peptide species. To ensure that values are comparable between fractions and between samples, a single representative peptide is chosen for each protein as a proxy of protein abundance. The representative peptide is selected in two ways: one aimed to optimally estimate protein abundances separately for each sample (i.e. within each labelling state) and is intended for qualitative comparisons (henceforth Scenario A), while second is aimed to compare the protein abundance quantitatively between labelling states (Scenario B).

When the objective is to evaluate qualitative differences between samples, such as protein co-migration, it is advantageous to select a representative peptide independently for heavy and light samples (Scenario A). By treating the two labelling states independently, migration profiles can be defined for all identified proteins, even if some of

their peptides were not detected in both labelling states (Figure 3.5A). By comparing protein migration profiles, it is possible to assess changes in the composition of complexes and identify their potential interactors.

When the objective is to evaluate the difference in protein abundance between labelling states, for any given protein, the selected representative peptide is the same for both labelling states (Scenario B). By comparing peptides that match in their sequence, charge and modifications between heavy and light samples, changes in the protein amounts can be quantified between these states. Although only proteins for which at least one matched peptide was identified in both samples will be analysed, potentially restricting the number of assessed proteins (particularly those present at low abundance), this procedure ensures high reliability of quantitation (Figure 3.5B).

A)

Scenario A
(qualitative comparison of proteins)

		Heavy												Light										
	fraction:	1	2	3	4	5	6	7	8	9	#	sum		fraction:	1	2	3	4	5	6	7	8	9	#	sum
Protein A	Peptide 1	■	■			■		■			■	5	Protein A	**Peptide 1**	■			■	■		■		■	■	**6**
	Peptide 2	■	■	■		■		■		■	■	**7**		Peptide 2	■	■			■				■	■	5
	Peptide 3	■				■		■		■		4		Peptide 3											0
	Peptide 4					■						1		Peptide 4	■				■		■		■		4

B)

Scenario B
(quantitative comparison of proteins)

		Heavy											Light											H+L
	fraction:	1	2	3	4	5	6	7	8	9	#	sum		fraction:	1	2	3	4	5	6	7	8	9	#	sum	
Protein A	Peptide 1	■	■			■		■			■	5	Protein A	Peptide 1	■			■	■		■		■	■	6	11
	Peptide 2	■	■	■		■		■		■	■	7		**Peptide 2**	■	■			■				■	■	5	**12**
	Peptide 3	■				■		■		■		4		Peptide 3											0	4
	Peptide 4					■						1		Peptide 4	■				■		■		■		4	5

Figure 3.5: Selection of representative peptide in ComPrAn
Schematic example of the process of selecting peptides to represent protein abundance, for heavy and light labelling states (peptides are split into two groups according to their attributed labelling states). Green boxes represent that a given peptide was detected in a given fraction. A) Scenario A. The peptide present in the highest number of fractions within each labelling state is considered the representative peptide (in this example Peptide 2 for heavy and Peptide 1 for light). Thus, for the same protein, the peptide picked for one labelling state may be different from the peptide selected for the other. This scenario picks a representative peptide for all proteins, even if they are present only in one labelling state. B) Scenario B. The selected representative peptide is the one that is present in the highest number of fractions, for both labelling states considered together. Thus, for the same protein, the representative peptide selected is the same for both labelling states. In this scenario a representative peptide cannot be selected for proteins that do not have shared peptides or are not present in both labelling states. The peptide present in both samples and in the highest number of fractions (column H + L) is selected as representative (Peptide 2).

3.2.4 Normalisation of protein intensities

The range of abundance of different proteins in a cell spans several orders of magnitude (Beck et al., 2011) which may hinder an effective visualisation of the migration profiles. To ease identification of co-migrating proteins (Scenario A) or comparison of relative

quantities (Scenario B), protein abundance is scaled to values between 0 and 1. However, due to the differences in representative peptide selection between Scenarios A and B, also the methods of normalisation differ between these scenarios: protein abundance values are scaled by the maximum value either within each labelling state (Figure 3.6A), or across both labelling states (Figure 3.6B).

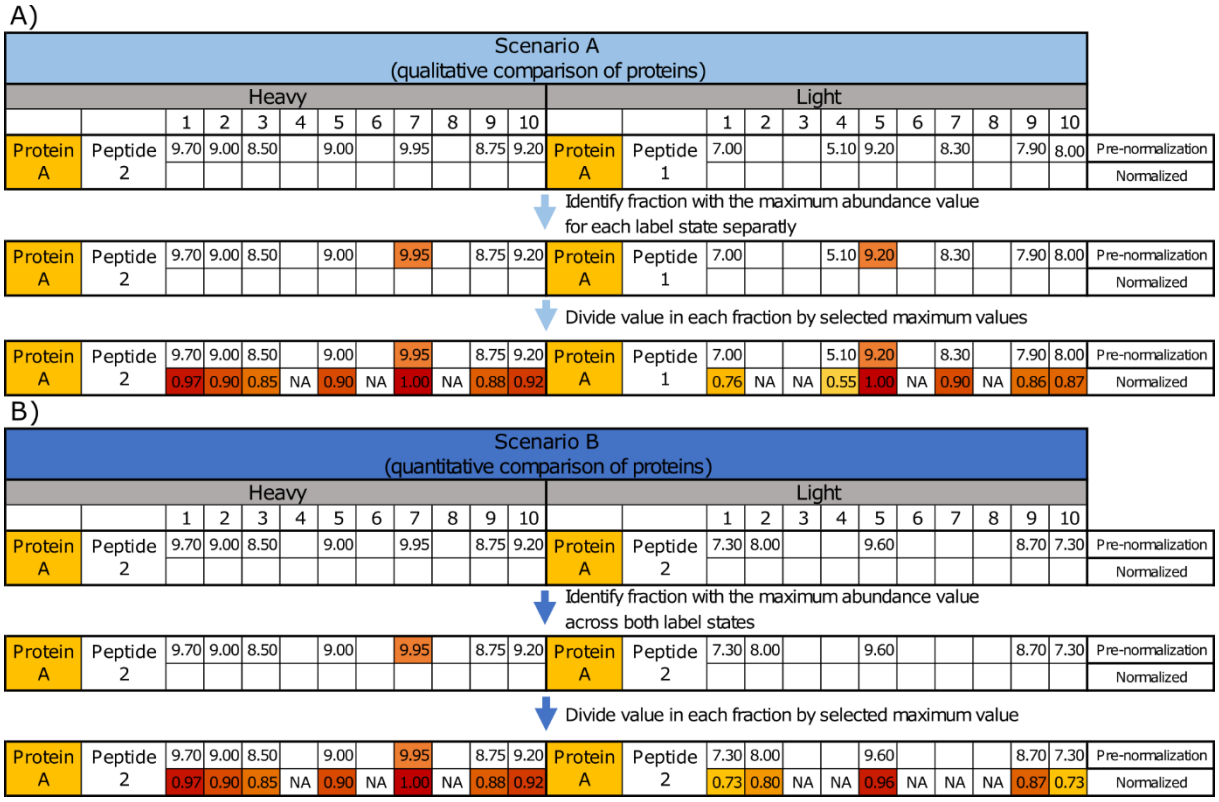


Figure 3.6: Normalisation of protein abundance values in ComPrAn
Schematic example of the process protein abundance normalisation. A) Schematic of a simplified case of normalisation according to scenario A (see also Figure 3.5A). First, the fraction with the highest value is identified, separately for heavy and light samples. Next, the value of each fraction is divided by this maximum value within each labelling state. This means that, for Protein A, there will be a value of 1 present at least in one fraction for the heavy sample and at least in one fraction for light sample. B) Schematic of a simplified case of normalisation according to scenario B (see also Figure 3.5B). First, the fraction with the highest value across both labelling states is identified. Next, the value of each fraction is divided by this maximum value. This means that, for Protein A, there will be a value of 1 present at least in one fraction in at least one of the heavy or light samples.

3.2.5 Hierarchical clustering

Hierarchical clustering of proteins can be performed as part of the ComPrAn workflow. While changes in the migration profile of proteins of interest can be verified manually, methods such as hierarchical clustering allow an assessment of overall changes in protein migration patterns across all detected proteins. In agglomerative hierarchical clustering, each protein is first assumed to be its own cluster and is progressively linked to others to form larger clusters. Proteins are assigned to clusters based on their "distance" from each other given a metric of similarity between their migration profiles. Several metrices can

be used as a distance measure, such as Pearson's correlation or the Euclidian distance in multidimensional space. In each step, clusters that are closest to each other are joined. The distance between clusters that are formed by more than one protein is determined by the choice of linkage method. In single linkage (also known as nearest neighbour), distance between clusters is the minimal distance between any two components of clusters. Conversely, complete linkage is defined as the maximum distance between any two components of clusters. Average linkage defines distance between clusters as the mean distance between all components of clusters (for further information on clustering techniques used in proteomics see Albaum et al., 2011; Meunier et al., 2007). Parameters available for cluster analysis in ComPrAn are based upon centred or uncentred Pearson's correlation as a distance measure with options of complete, average, or single linkage. The output is provided in a tabular format with cluster numbers associated with each protein, separately for each labelling state.

3.2.6 Graphical user interface

We aimed to make ComPrAn accessible to biologists with little to no programming experience. Therefore, a webpage-like graphical user interface ("ComPrAn app", Figure 3.7) has been implanted as part of ComPrAn. It was created using the Shiny R package, which enables building of interactive applications straight from the R code (Chang et al., 2020). After package installation, the app can be launched with the following command from within R:

```
# load ComPrAn R package
library(ComPrAn)
# launch the Shiny app
compranApp()
```

Analysis steps of the ComPrAn app are grouped into two parts. The user has a choice of inputting peptide data and analysing them by functions in Part 1. Alternatively, they can import normalised protein data produced in other software or exported previously after running Part1 of ComPrAn on peptide data and proceed directly to Part 2. The app is designed so that sections are available to the user only once necessary input is available (Figure 3.8).

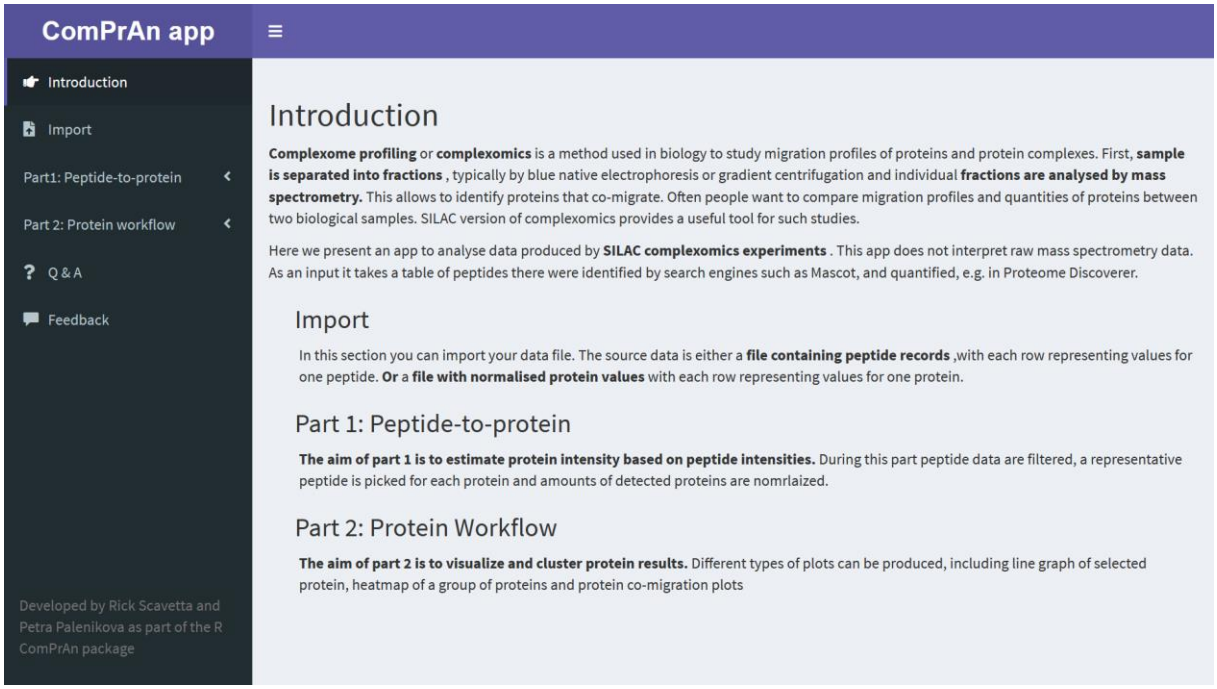


Figure 3.7: Introduction page of the ComPrAn graphical user interface
GUI can be launched from within R with a single command. It provides tools for analysis and visualisation of qDGMS results. Results of the analysis can be exported by the user in plain text format and visualisations in pdf format. App contains instructions for use as well as contact information to developers.

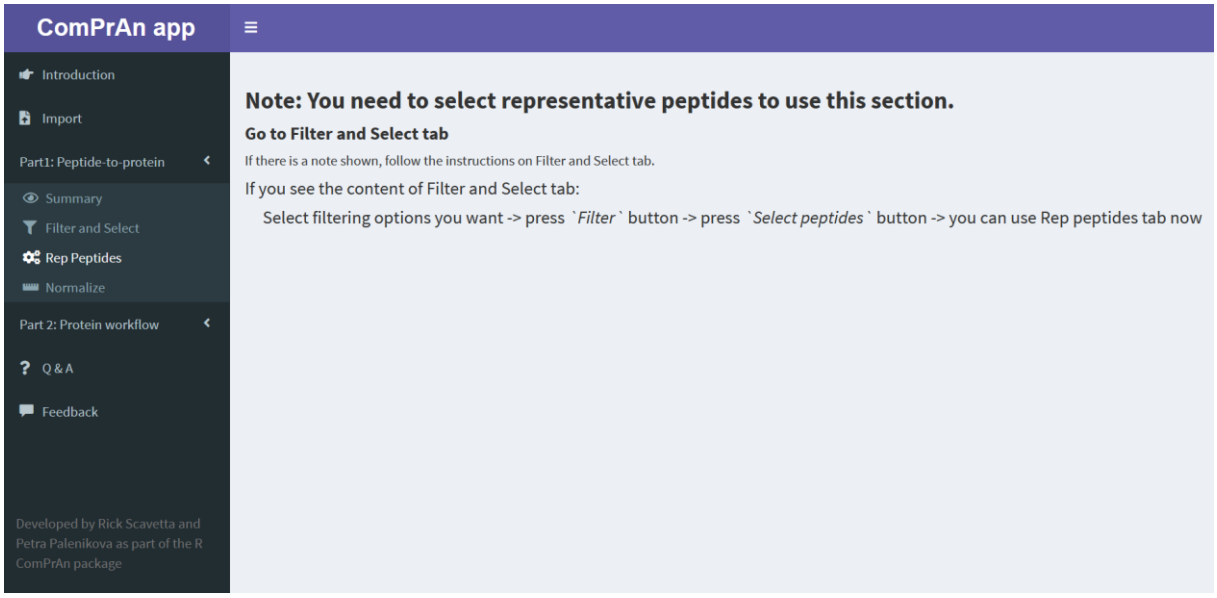


Figure 3.8: Conditional access to ComPrAn app sections
To reduce the chance of unexpected software behaviour, ComPrAn app is designed so that content of individual sections of the app is accessible only once the previous processing steps were performed and required input data are available. If users try to access parts of the app before performing all necessary steps, a note informing them of required actions is shown on screen. This example shows such note for "Rep Peptides" tab.

3.3 ComPrAn – outputs and visualisations

Two types of outputs can be generated with ComPrAn: plain text data tables or plots. ComPrAn visualisation functions produce ggplot objects (Wickham, 2016) and can be exported in many different formats including jpg, png, tif, pdf, svg etc. Visualisations exported from ComPrAn app are saved as pdf files.

ComPrAn outputs were designed with the aim to provide information necessary for answering specific biological questions. qDGMS experiments always compare two biological samples to identify differences between them. As a first indication about the differences, ComPrAn allows to create lists of proteins that were detected in both or only in one of the two samples. These lists are created from the peptide-level data. Peptide-level data tend to be large and their analysis is computationally intensive which makes it desirable to avoid multiple reanalysis. Therefore, intermediate output of ComPrAn pipeline is a table of normalised protein data (Figure 3.4). Two values are reported for each protein, one for Scenarios A and for Scenario B. These normalised protein data can be used as a direct input for ComPrAn visualisations without the need to reanalyse peptide data, as well as processed in other software. Lastly, a table of cluster assignments can be used to assess co-migration of all detected proteins.

Several types of visualisations can be created with ComPrAn. Two main plotting functions are provided for quantitative analyses. Firstly, a heatmap of selected groups of proteins allows to quantitatively compare large number of proteins (for examples see Figure 4.32, Figure 4.33, Figure 5.17, Figure 5.18). In the interactive app, heatmap can only be generated for the Scenario B and is intended for quantitative analysis. When used via command line, heatmaps can be generated from both Scenario A and Scenario B for qualitative and quantitative analysis, respectively. While heatmaps provide a good overview across a number of proteins of interest, it is often advantageous to be able to draw attention to individual proteins and to compare their profiles with line-plots. Therefore, this type of plot is included as a second visualisation alternative for quantitative data in ComPrAn (Figure 3.9). Changes in protein complex composition can be detected as shifts in the migration profile of its constituents. Furthermore, co-migration can be used as a prediction for putative protein interactions. To evaluate co-migration of known protein complexes or candidate assembly factors with the complex, list of proteins can be supplied to a function to generate average or median migration profile (for example see Figure 5.20B, D, F). For troubleshooting purposes, such as in case a certain protein is not present in quantitative plots or if it is showing unexpected behaviour in results, it is

possible to inspect quality of its detection by visualising all peptides that passed initial filtering (Figure 3.10). All these plotting functions are available both from the command line and as part of the ComPrAn app. Several customisation options are available for each plot in the app.

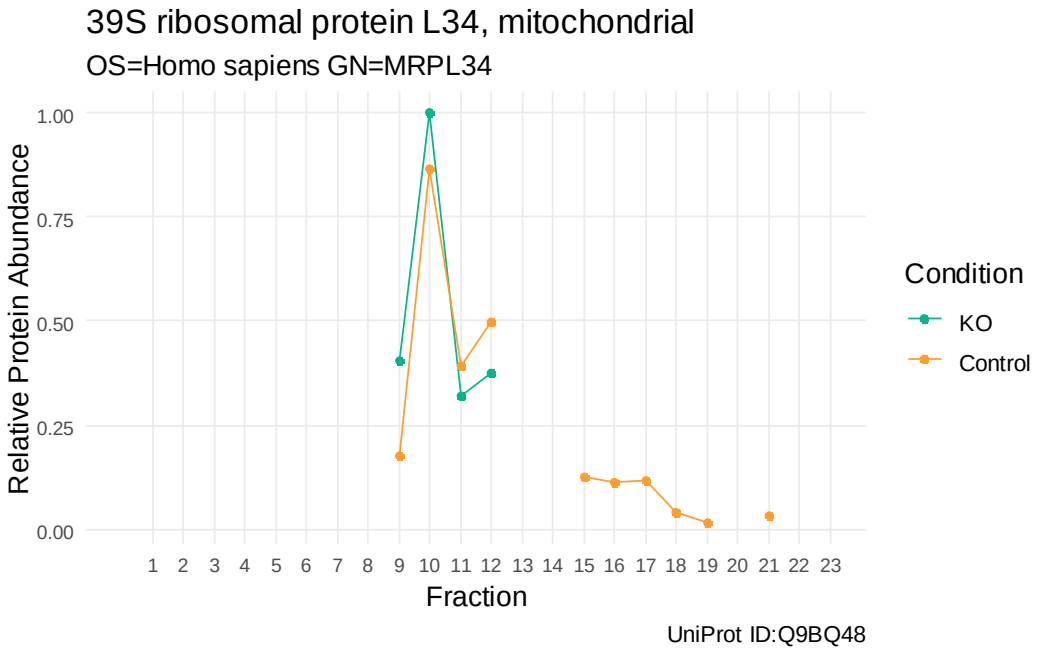


Figure 3.9: Quantitative analysis of a single protein in ComPrAn
Quantitative comparison of MRPL34 protein, component of large mitoribosomal subunit, between KO and Control samples from the ComPrAn example data set. Protein was not detected in fractions where no value is shown.

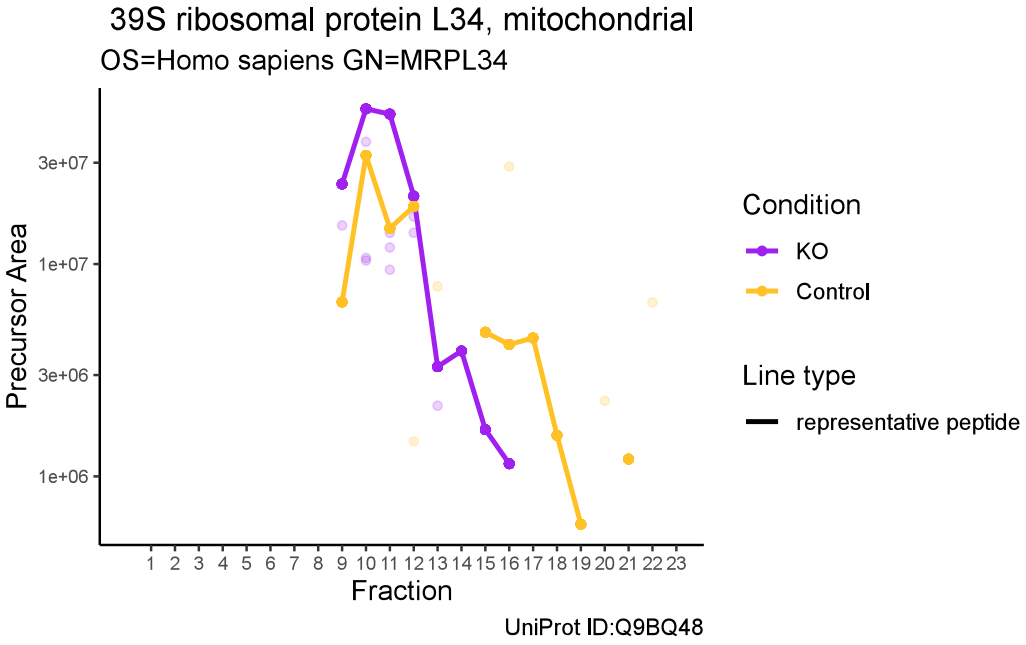


Figure 3.10: Visualisation of all peptides detected for a selected protein
All peptides detected for MRPL34 protein in KO and Control samples that passed the initial filtering in the ComPrAn example data set. Initial filtering removes peptides which were not assigned into a protein group had missing precursor area value or had a "Rejected" value of PSM (peptide spectrum match) ambiguity. Values of representative peptide for Scenario A are highlighted.

Additional visualisations summarising information about the data are available within the ComPrAn app. After importing peptide data, information about the total number of peptides in the file and how many of them passed the initial quality control is provided in the app section Summary (Figure 3.11A). Additionally, Filter and Select app section contains Venn Diagrams summarising numbers of proteins in the two biological samples (Figure 3.11B)

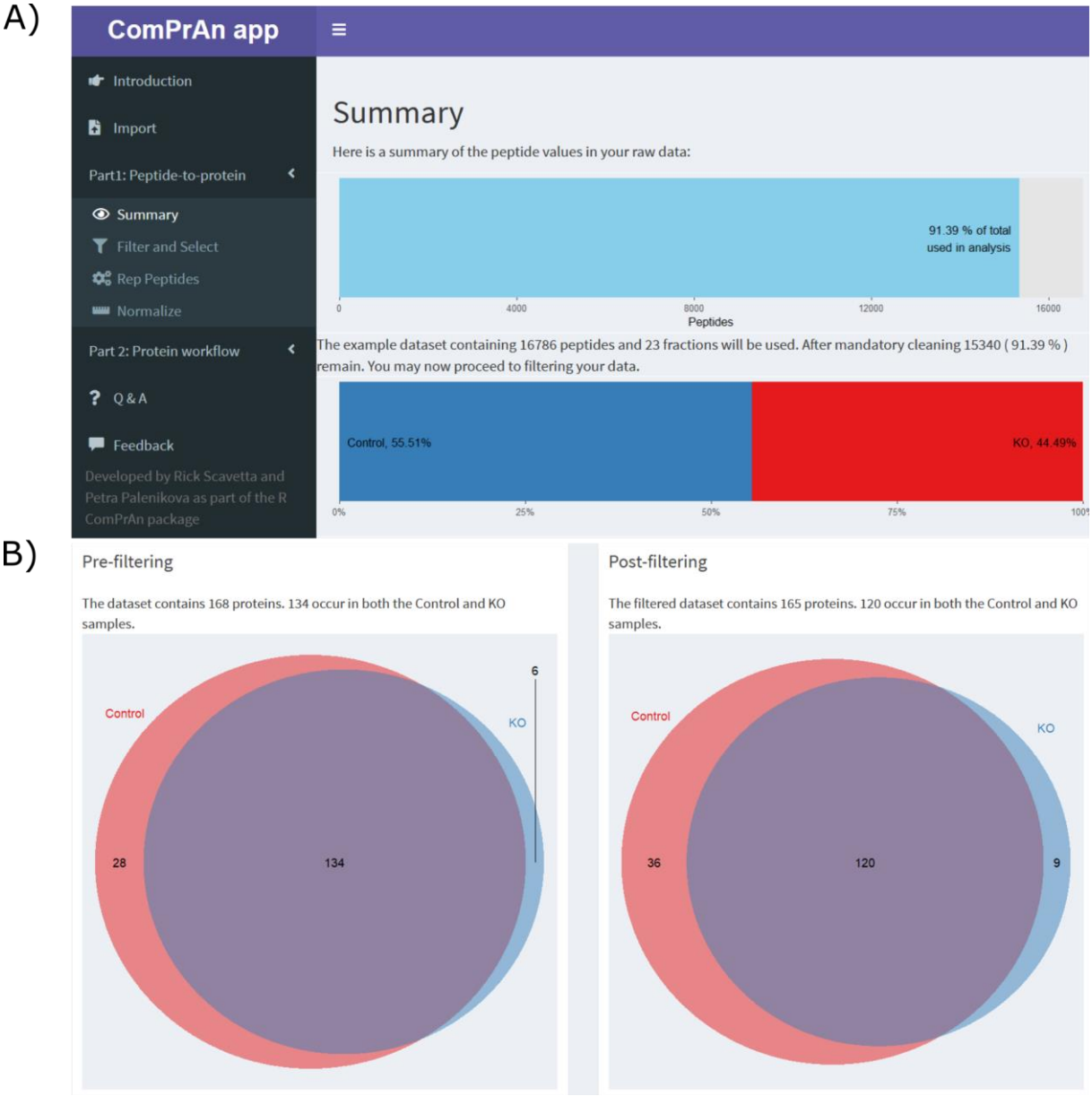


Figure 3.11: Summary plots available in ComPrAn Shiny app
A) Once data are imported into the ComPrAn app and after initial filtering of data, fraction of peptides used for analysis and distribution of peptides between "heavy" and "light" samples is shown as a rough indication of the sample pooling quality. B) Overall numbers of proteins detected in each sample is shown in a Venn diagram. Pre- and post- filtering values refer to additional optional filtering, parameters of which are chosen by the user. Note that these figures were created with the use of example data set, in real data, fraction of peptides passing mandatory cleaning is usually lower and overall number of detected proteins is higher.

3.4 Conclusions and future development

Complexome profiling is becoming a widely used method for study of mammalian mitochondrial protein complexes (Čunátová et al., 2021; Fernández-Vizarra et al., 2021; Lobo-Jarne et al., 2020; Protasoni et al., 2020) as well as protein complexes of other organisms (Berger et al., 2021; Evers et al., 2021; Maclean et al., 2021; Röhricht et al., 2021; Rugen et al., 2019, 2021). Recently a standardised format of reporting results of complexome profiling (minimal information about a complexome profiling experiment, MIACE) has been proposed by the research community (Van Strien et al., 2021). Concurrently, a data sharing platform has been established to serve as a repository of all published complexome profiling data (CEDAR database) in a standardised "CompTab" format (Van Strien et al., 2021). With a growing interest and use of complexome profiling experiments, the availability of user-friendly data analysis and visualisation tools will be crucial for achieving the full potential of this method.

ComPrAn R package is a dedicated data analysis tool with easy-to-use graphical user interface that helps in making qDGMS and SILAC complexome profiling more accessible. Due to its implementation as an open source R package, ComPrAn has a potential to be integrated and expanded with the breadth of software packages created by R developers. The current version of ComPrAn has been developed for 2-state labelled data, therefore allowing direct comparison of two samples within one biological repeat. Since each protein complex is represented by data from multiple proteins, even a single experiment can provide valuable information about the state of the studied complex. Integration of results from multiple repeats would require consideration of technical variability between experiments, such as variation in fractionation or differences in sample amount. One approach to compensate for this variability used in published software is normalisation of protein abundance values between experiments based on the total abundance of all detected proteins (Van Strien et al., 2019). Alternatively, use of metabolic labelling with "medium" isotopes or a NeuCode labelling (Potts et al., 2016) allows multiplexing of higher number of samples and therefore combining multiple biological repeats into one experiment. Multiplexing at the stage of metabolic labelling provides advantages over *in silico* normalisation, since it reduces technical variability introduced during sample processing. Approaches to allow streamlined combination of multiple repeats as well as analysis of multiplexed mass spectrometry data could be implemented into the future ComPrAn versions, as the outreach and demand for the method grows. Furthermore, introduction of the standardised CompTab file format (Van Strien et al., 2021) provides

scope for implementation of functions compatible with processing these files as well as functions to output protein data in the CompTab format as part of the ComPrAn pipeline to allow easy submission to the CEDAR database.

4 Identifying genetic interactions of MRM1

4.1 Introduction

4.1.1 The role of MRM1 in mitochondria

Subunits of the mitochondrial ribosome are composed of both proteins and RNA. 16S ribosomal RNA is part of the large mitoribosomal subunit (mtLSU). In humans, 16S rRNA contains five posttranscriptional modifications and the enzymes that deposit these modifications have been identified recently (for details see section 1.4.3). However, the precise role of these modification in the biogenesis and/or function of the mitoribosome is still unknown. This includes the modification deposited by mitochondria rRNA methyltransferase 1 (MRM1), an enzyme containing a 2'-*O*-ribose binding site and a SPOU methylase domain (Lee et al., 2013). Using a DNAzyme assay it was shown that MRM1 modifies 16S rRNA at position G1145 (Lee and Bogenhagen, 2014) which is located in the P-loop of the peptidyl transferase centre (Figure 4.1). MRM1 has been found to be associated with the nucleoid and to have only weak binding with the ribosome, suggesting it might function at an early stage of mitoribosome assembly (Lee et al., 2013).

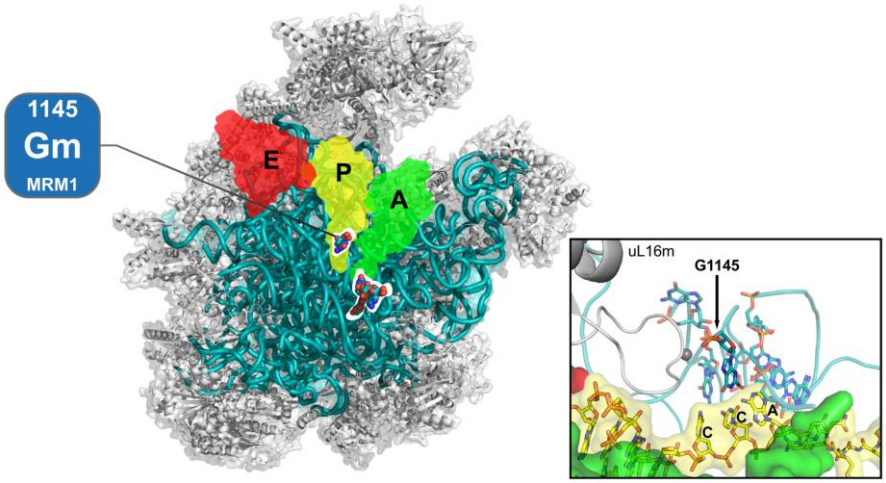


Figure 4.1: Position of the Gm1145 modification in the large mitoribosomal subunit
The Gm1145 modification, which is catalysed by MRM1 protein, is located in the P-site of the mtLSU. View of the mtLSU from the interface between mtSSU and mtLSU is shown. Three ribosome active sites (A, P and E) are highlighted. 16S rRNA is presented as a teal ribbon and mitoribosomal proteins as grey cartoons. Local environment of the modification is shown on the bottom right. Adapted from (Rebelo-Guiomar et al., 2018).

The modification deposited by MRM1 is highly evolutionary conserved. In *E. coli*, position G2251 on 23S rRNA (equivalent to G1145 in mitochondrial 16S rRNA) is modified by RlmB. This protein belongs to the SPOUT superfamily of methyltransferases with the AdoMet binding site located in an unique knotted region and it dimerises in solution (Michel et al., 2002). Mutation of *rlmB* gene has no effect on bacterial growth or ribosome integrity. The growth of a *ΔrlmB* strain is unaffected even under competitive growth conditions with a rlmB+ strain (Lövgren and Wikström, 2001).

In *S. cerevisiae*, equivalent position G2270 of 21S rRNA is modified by the Pet56p/Mrm1p methyltransferase (Sirum-Connolly and Mason, 1993). In contrast to bacteria, disruption of the promoter of *pet56* leads to inability of cell growth on a nonfermentable media and a petite phenotype. Furthermore, assembly of the mtLSU is disrupted as observed by a deficiency of the mature mtLSU particles and accumulation of slower sedimenting particles on the sucrose gradient (Sirum-Connolly et al., 1995). While purified Pet56p/Mrm1p enzyme exhibits methyltransferase activity on *in vitro* transcribed sense transcripts of domains IV and V of 21S rRNA from *S. cerevisiae* and 23S rRNA from *E. coli* (Sirum-Connolly et al., 1995), human MRM1 purified without an MTS using T7 expression system is monomeric and does not show convincing methyltransferase activity on *in vitro* transcribed 16S rRNA (Lee et al., 2013).

Given that Gm1145 is one of only five modifications present in the mtLSU and it is highly conserved, it is likely that MRM1 and the Gm1145 modification would play a role in the biogenesis and/or function of the human mitoribosome. However, study of a HEK293 cell line depleted of this protein did not show any detectable phenotype (D'Souza, 2018). MRM1 knockout cells were indistinguishable from wild type in terms of growth rate, OXPHOS function and mitochondrial translation. This cell line also did not show any changes in the assembly of the mitoribosomal subunits. MRM1 KO cells, however, showed a growth disadvantage compared to the wild type cells when cultured in medium with chloramphenicol (D'Souza et.al., unpublished). In the bacterial ribosome, chloramphenicol binds into the A-site of the peptidyl transferase centre and prevents peptidyl transferase activity (Bulkley et al., 2010; Svetlov et al., 2019). It is therefore possible that the role of MRM1 becomes apparent only in certain stress conditions. Previous studies in *S. cerevisiae* showed that for some tRNA modifying enzymes only deactivation of multiple enzymes at the same time has a detrimental effect on cell physiology (Alexandrov et al., 2006) or sensitises cells to certain compounds (Gustavsson and Ronne, 2008). Similar findings were shown also in mammalian systems (Okamoto et al., 2014). We hypothesised that an analogous effect might be true also for rRNA modifications, and that to observe an

effect of MRM1 knockout an inactivation of further proteins would be necessary. We therefore set to study genetic interactions between MRM1 and other genes with known or predicted functions in mitochondrial gene expression and maintenance by the means of CRIPSR/Cas9 screening methodology. Identification of such genetic interactions should enable better understanding of the role of MRM1 and Gm1145 modification in mitochondrial function.

4.1.2 Study of mitochondrial OXPHOS by large-scale genetic screens

CRISPR/Cas9 revolutionised gene editing technologies as it facilitates the creation of precise changes in the genome (Jinek et al., 2012). This technology requires only two components: a Cas9 nuclease and a single guide RNA (gRNA) to target the nuclease to the genomic region of interest. Since Cas9 is targeted by a relatively short RNA sequence, which can be delivered into a cell in a form of DNA that is transcribed by the cellular gene expression machinery, this method can be easily scaled up to target multiple genes in a single experiment. During the past decade experiments have been designed to target almost all genes encoded in the human genome. The integration of the effects of gene knockouts among hundreds of cancer cell lines within Project Achilles and Project Score enabled definition of core and context specific essential genes (Dempster et al., 2019; Pacini et al., 2021; Sharma et al., 2020).

Several large-scale studies aimed to identify a set of genes important for the function of oxidative phosphorylation. These studies often employ selective growth on media with galactose (without glucose) as mammalian cells with a defect in OXPHOS are unable to grow on this type of the culture media (Robinson et al., 1992). An RNA interference screen in mouse fibroblasts was used to identify new OXPHOS related genes by negative selection on glucose and galactose containing medium (Bayona-Bafaluy et al., 2011). The screen was performed with a library that covered ~30% of the transcripts encoded by the mouse genome and identified 643 putative OXPHOS related short hairpin RNAs (targeting a total of 519 genes). By combining these results with a coexpression analysis a collection of 33 genes with higher probability of being OXPHOS related was identified (Bayona-Bafaluy et al., 2011). Another study profiled 107 proteins out of the MitoCarta collection (Pagliarini et al., 2008) with known or predicted RNA-binding domains by an arrayed RNAi screen (Wolf and Mootha, 2014). Effects of protein knockdowns on mtDNA transcription were analysed by a MitoStrings assay that uses fluorescent probes to quantify mitochondrial

transcripts in a crude cell lysate without the need for RNA amplification. By applying hierarchical clustering to the results, they were able to group genes based on their roles in mtDNA transcription, RNA cleavage and stability (Wolf and Mootha, 2014). A compendium of 191 high confidence genes required for OXPHOS was identified by a positive selection CRIPSR/Cas9 screen (Arroyo et al., 2016). In this work, an approach to identify OXPHOS related genes was a positive selection of cells that expressed Annexin V, a marker of cell death, at 24 hours after swapping the culture media from glucose containing medium to one with galactose. Identified genes play a role in different processes in mitochondria, including mtDNA gene expression, oxidative phosphorylation, protein targeting to mitochondria and mitochondrial metabolism (Arroyo et al., 2016). Recently a comprehensive list of essentiality of nuclear encoded mitochondrial genes in various cancer cell lines has been created with the use of data from Project Achilles (Thomas et al., 2021). This study further investigated the role of mitochondrial genes in U2OS cell line grown under normoxic conditions in either glucose or galactose medium and under hypoxic conditions on glucose rich medium. As exemplified by these studies, large-scale screens are a useful tool to investigate nuclear encoded mitochondrial proteins.

4.1.3 Use of CRISPR/Cas9 screens for identification of genetic interactions and synthetic lethality

Genetic interaction refers to a situation when an effect of one gene is affected by another gene. In general, there are two types of genetic interactions, positive (buffering interaction, suppression) and negative (synthetic sick/lethal) (Figure 4.2). When two genes have a positive interaction, inactivation of the second gene is able to supress the effect of inactivation of the first. Conversely, genes with negative interaction have negative impact on cell growth only when inactivated at the same time. Synthetic lethality is an extreme case of negative genetic interaction when simultaneous inactivation of multiple genes causes cell death.

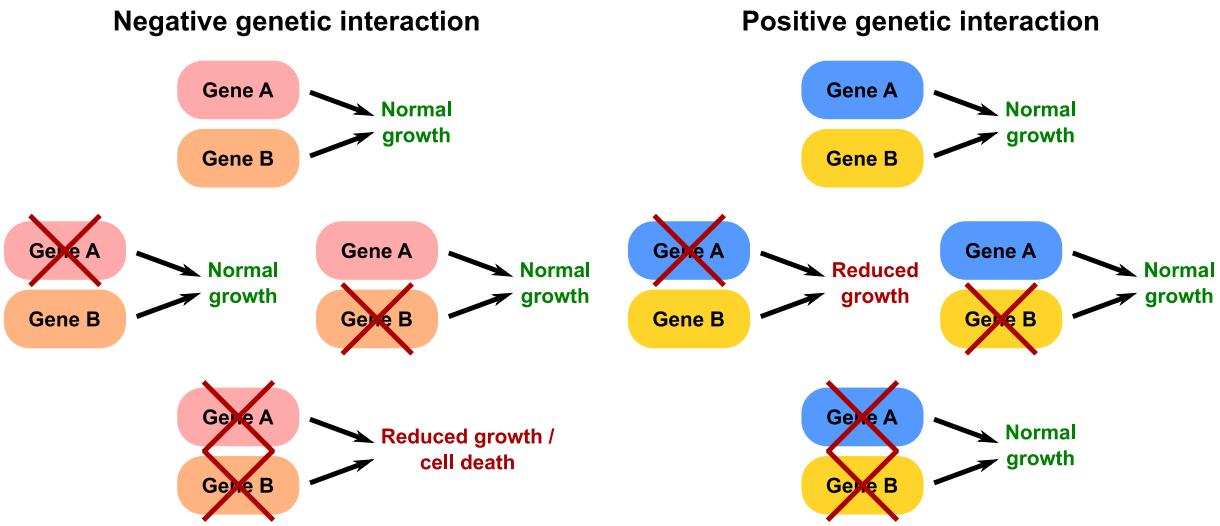


Figure 4.2: Schematic representation of genetic interactions
Genetic interactions occur when an effect of a gene is modified by a second gene. This illustration depicts a simple case of effects of Gene A and Gene B on cellular growth. A negative genetic interaction occurs if individual inactivation of either Gene A or Gene B have no effect on cell growth, but combined inactivation of both genes causes growth defect or cell death (left). A positive interaction between Genes A and B can occur if inactivation of both genes can supress the growth defect caused by individual inactivation of one of the genes (right).

Combinatorial screens targeting pairs of genes with dual gRNAs have been applied in a study of genetic interactions amongst genes that can be targeted by drugs (Han et al., 2017), and known cancer genes (Shen et al., 2017), providing insight into potential new therapeutic approaches. A human genetic interaction map for 222,784 gene pairs have been created by CRISPR interference screen, where catalytically dead Cas9 was targeted to 472x472 gene pairs in K562 and Jurkat cell lines (Horlbeck et al., 2018). This study showed that genetic interactions are informative about functional groups of genes. By clustering analysis, they revealed 33 distinct functional gene clusters spanning various cellular functions such as electron transport, mitochondrial translation, DNA replication and others. Apart from gene-gene interactions, information about relationships between gene functions can also be gathered by the study of chemical-genetic interactions. A study combining small molecule mitochondrial inhibitors and genome-wide CRIPSR/Cas9 screen reported 191 genetic modifiers of mitochondrial dysfunction (To et al., 2019). Interestingly, it appears that in mitochondrial dysfunction it can be beneficial to inhibit mitochondrial function at a second site. Cellular defects caused by oligomycin inhibition of Complex V can be supressed by loss of Complex I activity by either genetic inactivation or small molecule inhibitor (To et al., 2019).

The study of genetic interactions has proved useful in informing about the functions of poorly characterised genes. Several factors involved in the regulation of mitochondrial gene expression remain yet to be characterised in detail. Since a previous study of MRM1 protein by production and characterisation of MRM1 knockout cells did not provide

conclusive information about the role of this protein in the cell, we applied a targeted CRIPSR/Cas9 screen to identify its potential genetic interactions. Results of this screening approach were verified by characterisation of mitochondrial function in a double knockout cell line.

4.2 Results

4.2.1 Design of gRNA library targeting mitochondrial gene expression machinery

CRISPR screens can be performed using several approaches, and the best design depends on the research hypothesis. Screens can be performed in an arrayed or pooled format (reviewed in Agrotis and Ketteler, 2015). Arrayed screens are usually performed in multiwell plates, where each well contains cells transduced with a gRNAs targeting a single gene. In contrast, in pooled screens a population of cells is transduced by gRNAs targeting various genes. Arrayed screening allows for a direct readout of the effect of gene knockout, for example by evaluating cell growth in respective wells. However, it is technically difficult to perform this type of screening for a large number of genes. This limitation can be mitigated to a point by the use of automated plate handling systems. In the case of pooled screening, the effect of gene knockout is most often evaluated by next generation sequencing. Representation of the gRNAs in the pool is compared between the beginning of the experiment and the end point. Hits of the screen are manifested as an enrichment or depletion of gRNAs in positive or negative selection screens, respectively.

Another consideration for the screen design is the number of genes included in the library. Several genome-wide libraries have been used to study gene essentiality in various cell lines (Doench et al., 2016; Hart et al., 2015; Sanson et al., 2018; Tzelepis et al., 2016; Wang et al., 2015). The advantage of a genome wide approach is a fully unbiased study of all genes. This approach is however very demanding on resources. It increases costs of both cell culture and sample processing. Large numbers of cells need to be maintained to ensure gRNA representation is not affected by random depletion due to insufficient cell numbers. To obtain reliable results, sufficient coverage of the gRNA library also needs to be maintained during sample processing and higher depth of sequencing is required. When the research question targets specific cellular process, it proved to be useful to use only subsets of genes in the gRNA library (Birsoy et al., 2015; Grevet et al., 2018; Prolo et al., 2019). This can significantly reduce the cost of the experiment.

We were interested specifically in the study of genetic interactions in mitochondrial gene expression machinery. Therefore, we performed a pooled CRISPR/Cas9 screen with a targeted gRNA library. There is no commercially available library targeting genes with known or predicted function in mitochondrial gene expression. Therefore, we designed and

cloned a custom gRNA library. We selected a subset of 220 proteins with known functions in different parts of mitochondrial gene expression (MGE). In addition, we included 59 genes with predicted functions in mitochondrial gene expression. Genes for the inclusion in the library were assembled based on available published literature including genes with a function in mitochondrial transcription and translation (D'Souza and Minczuk, 2018), known and predicted tRNA modifying enzymes (Suzuki and Suzuki, 2014), rRNA modifying enzymes (Bohnsack and Sloan, 2018) or ribosome components and assembly factors (De Silva et al., 2015). In total, the targeted gRNA library was composed of 279 experimental genes (Figure 4.5, Appendix I).

To be able to assess whether the screening process works technically, we complemented the library with genes that showed a strong effect on cell survival in previous CRIPSR/Cas9 screens (positive control), genes that did not affect cell viability (negative control) and a set of non-targeting gRNAs (Wang et al., 2015). Control genes were selected based on gene effects assessed in previous studies for AVANA and GeCKOv2 genome wide libraries (Allen et al., 2019) and for Sabatini library (Wang et al., 2015). The first study reanalysed data from multiple pooled knockout screens performed with the specified libraries in more than 300 cell lines. The second work assessed essential genes across 4 cell lines.

As negative control we selected 100 genes with the smallest effect on growth from the above data sets using the following criteria: a) genes with the absolute mean, median and standard deviation values of log2 fold change in gRNA abundance below 0.11 across all cell lines; b) minimum and maximum log2 fold change in any one cell line between -0.4 and 0.4 (Figure 4.3). This gave a total of 105 negative control genes. To get a final set of 100 negative control genes, we removed 5 genes with the most negative effect on growth in the study by Morgens et.al. (2017) from the selected set. Criteria for selecting positive control genes were: a) mean and median values below -1 across all cell lines; b) not present in MitoCarta 2.0 (Calvo et al., 2016). This gave a final set of 20 positive control genes (Figure 4.4). We also included 500 gRNAs without a target in the genome as a control. Sequences of non-targeting gRNAs were taken from the study of (Wang et al., 2015). This resulted in a final custom gRNA library with 279 experimental genes, 120 control genes and 500 non targeting gRNAs (Figure 4.5).

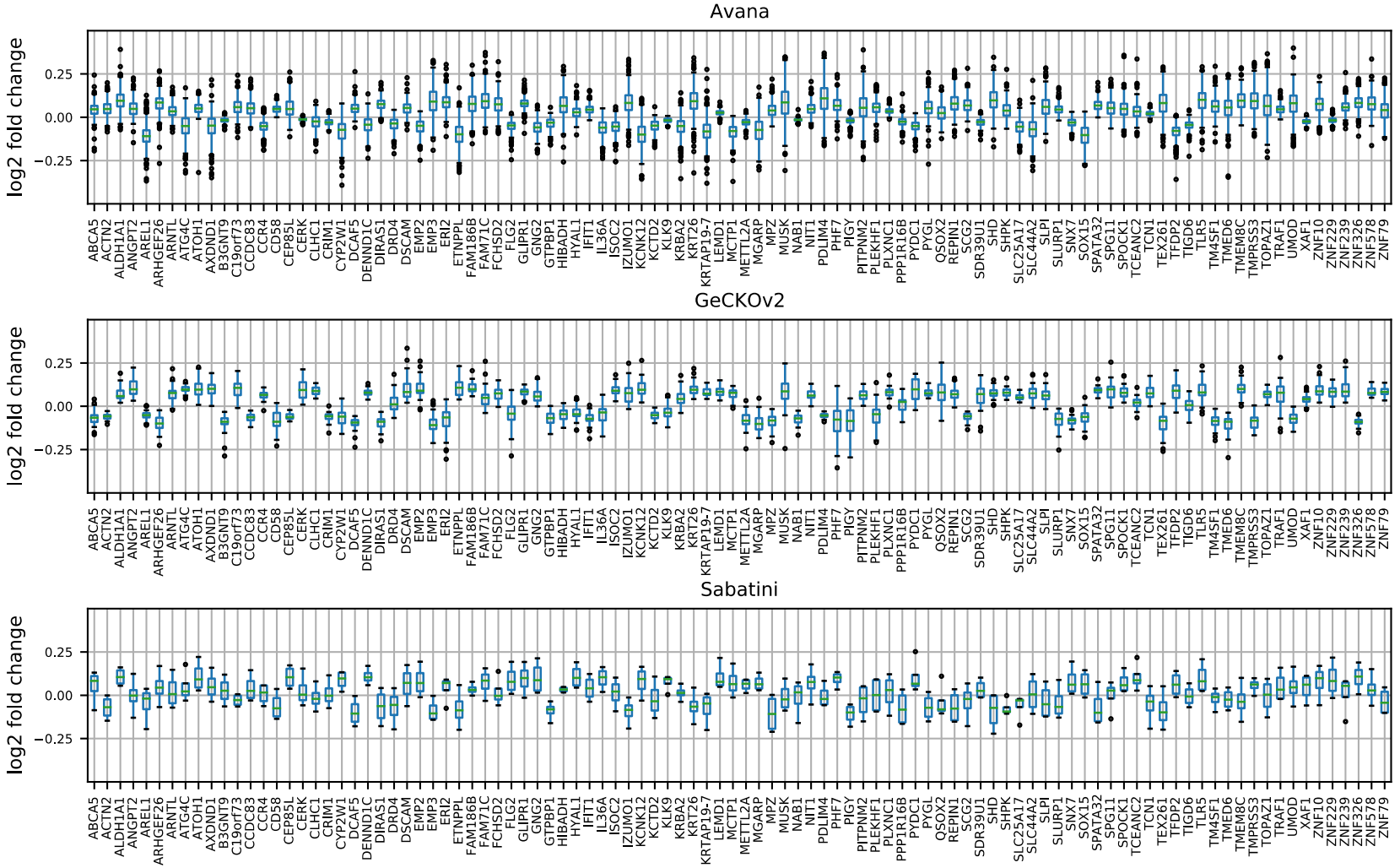


Figure 4.3: Selection of negative control genes

Bar plots showing the median and quartiles of log2 fold change values for individual genes in selected published genome wide CRISPR/Cas9 screens(Allen et al., (2019) for Avana and GeCKOv2, Wang et al., (2015) for Sabatini). All selected genes have absolute value of mean and median below 0.11 and standard deviation below 0.11 across all cell lines. The minimum and maximum log2 fold change in any one cell line is between -0.4 and 0.4. Boxplots show median (green line), first and third quartiles (box hinges) and whiskers extend to the highest or lowest point in each group within 1.5 times of inter-quartile range, any points above this threshold are shown as dots.

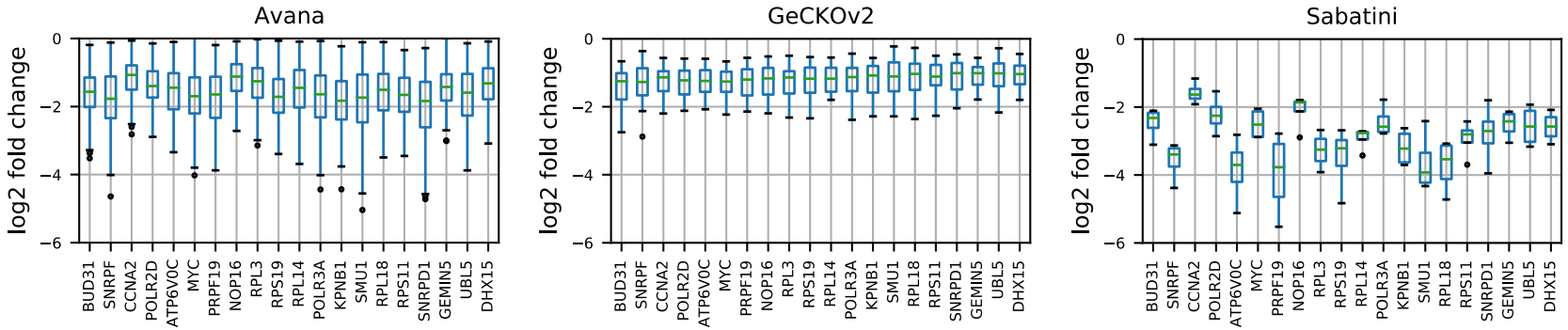


Figure 4.4: Selection of positive control genes

Bar plots showing the median and quartiles of log2 fold change values for individual genes in selected published genome wide CRISPR/Cas9 screens(Allen et al. (2019) for Avana and GeCKOv2, Wang et al. (2015) for Sabatini). All selected genes have value of mean and median below -1 across all cell lines. Boxplots show median (green line), first and third quartiles (box hinges) and whiskers extend to the highest or lowest point in each group within 1.5 times of inter-quartile range, any points above this are shown as dots.

Once full list of genes was assembled, 10 gRNA sequences per gene were picked from gRNA libraries designed by the Sabatini lab (Wang et al., 2015) and the Bassik lab (Morgens et al., 2017). For a small number of genes (RECQL4, GTPBP6, MRPS38, MTERF2), fewer than 10 gRNA sequences were taken from these studies due to the absence of the gene in these studies or the gRNA sequence not meeting exclusion criteria of DNA synthesis or cloning strategy (absence of single nucleotide stretches such as "TTTT" and "GGGGG" or absence of BsmBI restriction site, respectively). Missing gRNA sequences were taken from GeCKO v2 library (Sanjana et al., 2014).

The final custom library contained 4490 gRNAs. The library was synthesised as a pool of single stranded DNA oligonucleotides by CustomArray. These oligonucleotides were cloned into plasmid according to (Doench et al., 2016), resulting in a plasmid pool library which was used for the production of a lentiviral library.

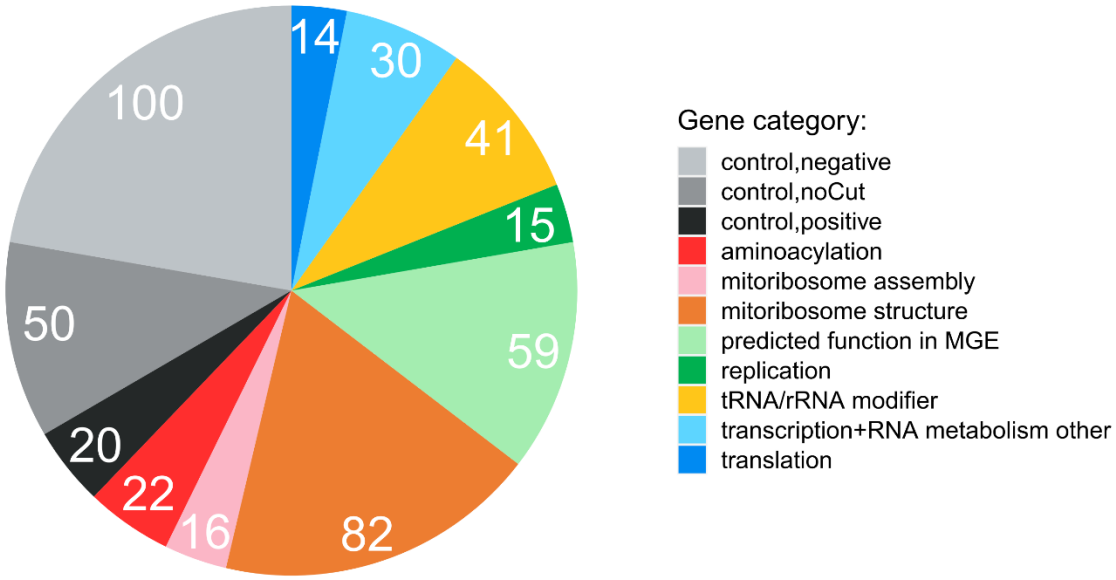


Figure 4.5: Composition of gRNA library
Genes selected for the gRNA library were assigned into different categories based on their function. Numbers on the pie chart represent numbers of individual genes in each of the categories. "control,noCut" category contained 500 gRNAs that do not have a target in the human genome, as each of the coding genes was targeted by 10 gRNAs, gRNAs in this category were randomly assigned into 50 groups with 10 gRNAs each. Full list of genes included in the library is available in Appendix I.

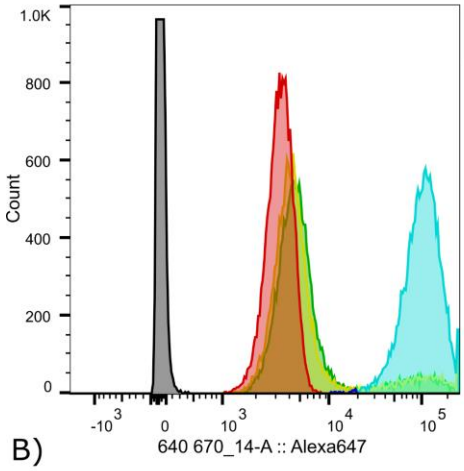
4.2.2 Generation of MRM1 KO cells stably expressing Cas9 protein

Expression of a Cas9 protein is a necessary component in CRIPSR screens. Two approaches have been used previously for the delivery of Cas9 into the cell line. In some studies, Cas9 has been delivered together with the gRNA as part of the lentiviral vector (Shalem et al., 2014). The second approach is to use a cell line with either constitutive or

inducible expression of the Cas9 protein (Koike-Yusa et al., 2014; Wang et al., 2014, respectively). Systems with Cas9 and gRNA in a single viral particle save experimental time as both gRNA and Cas9 integration happen at the same time. This approach might therefore be beneficial in cell lines that are difficult to culture. However, viral titters tend to be lower with one vector system (Sanjana et al., 2014). Prior integration of Cas9 into the genome followed by clonal selection also ensures uniform expression of this protein across all cells in the pooled screen. We took advantage of the FRT site presence in the Flp-In T-Rex HEK293 cell line which was used for production of MRM1 KO cell line (D'Souza, 2018). We integrated the Cas9 protein into the FRT site, allowing for control of expression levels via the Tet repressor system by inducing the protein expression using doxycycline. MRM1 KO cells were transfected with a plasmid carrying Cas9 protein and subsequently cultured in the medium containing hygromycin to select clones in which Cas9 has been integrated into the FRT site.

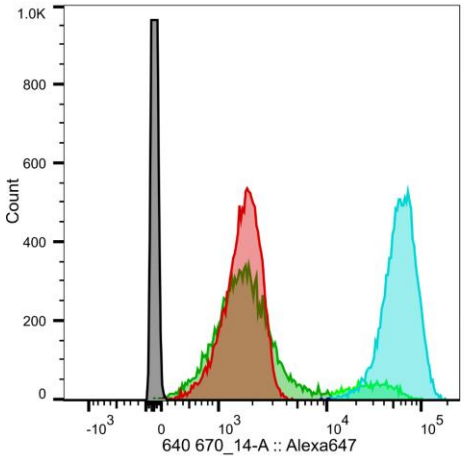
Once clones have been isolated, sufficient expression of Cas9 protein has to be verified. To this end, we transduced the cells with lentiviral vectors containing gRNA targeting the Beta-2-Microglobulin (B2M) gene. B2M functions as a chaperone for MHC class I, and in its absence, MHC I cannot be trafficked to the cell surface (Donaldson and Williams, 2009). By staining the cells with antibody against MHC class I followed by FACS analysis, we could observe the level of Cas9 expression efficiency (Figure 4.6). Cas9 activity was evaluated for MRM1 KO and wild type HEK293 cell lines. Experiments were performed in the presence and in absence of doxycycline induction. Approximately 90% reduction of fluorescence signal was observed in transduced cells, with or without doxycycline induction (Figure 4.6A). This suggests that leaky expression of Cas9 from the FRT site provides sufficient amount of the enzyme to effectively cut the B2M target site in most of the cells in the timeframe of one week. Unless otherwise stated, HEK293 and MRM1 KO in this chapter refer to the cell lines with Cas9 integrated into FRT site.

A)



	Subset Name	Count	Freq. of Parent	Workspace Annotation
	Alexa647-	17099	100	HEK293, no virus
	Alexa647+	0	0	HEK293, no virus
	Alexa647-	17689	99.9	HEK293, no virus, Alexa647
	Alexa647+	15.0	0.085	HEK293, no virus, Alexa647
	Alexa647-	107	0.65	HEK293, no virus, W6/32, Alexa647
	Alexa647+	16314	99.3	HEK293, no virus, W6/32, Alexa647
	Alexa647-	14904	88.1	HEK293, virus+, W6/32, Alexa647, Dox-
	Alexa647+	2009	11.9	HEK293, virus+, W6/32, Alexa647, Dox-
	Alexa647-	14492	88.4	HEK293, virus +, W6/32, Alexa647, Dox+
	Alexa647+	1908	11.6	HEK293, virus +, W6/32, Alexa647, Dox+

B)



	Subset Name	Count	Freq. of Parent	Workspace Annotation
	Alexa647-	14083	100	MRM1 KO, no virus
	Alexa647+	0	0	MRM1 KO, no virus
	Alexa647-	15120	100.0	MRM1 KO, virus+, Alexa647
	Alexa647+	4.00	0.026	MRM1 KO, virus+, Alexa647
	Alexa647-	66.0	0.48	MRM1 KO, no virus, W6/32, Alexa647
	Alexa647+	13792	99.5	MRM1 KO, no virus, W6/32, Alexa647
	Alexa647-	14268	89.2	MRM1 KO, virus+, W6/32, Alexa647
	Alexa647+	1729	10.8	MRM1 KO, virus+, W6/32, Alexa647

Figure 4.6: Cas9 activity check in HEK293 Cas9 and MRM1 KO Cas9 cells
The expression of Cas9 was evaluated in HEK293 Cas9 (A) and MRM1 KO Cas9 (B) cells. Cells were transduced with lentiviral particles with gRNA targeting B2M gene. Subsequently, cells were stained with W6/32 primary antibody and Alexa647-anti-mouse secondary antibody and analysed by FACS. Uninfected and unstained cells were used as a negative control. Uninfected cells stained with both primary and secondary antibodies were used as a control of the maximum level of MHC I. Cells stained with only secondary antibody were used as a control of maximum observable level of signal reduction. Cells transduced with the virus were analysed both with and without doxycycline induction of Cas9 expression.

4.2.3 CRISPR/Cas9 screen design

To identify genetic interactions between MRM1 and other genes of MGE, we performed a CRISPR/Cas9 screen in MRM1 KO and wild type HEK293 cells (Figure 4.7A). To be able to detect dysfunction of mitochondria, the screen was performed both on standard glucose-rich growth medium and on medium that forces the cells to rely on mitochondrial respiration for ATP production (galactose medium). Human cells with an OXPHOS deficiency are not able to survive in galactose medium (Robinson et al., 1992). Samples for analysis were collected at several time points after the transduction. Once all samples were collected, genomic DNA was isolated and used for PCR amplification of gRNAs integrated into the genome of the cells. Next generation sequencing was used to evaluate

representation of the gRNAs at individual time points. Gene essentiality was calculated by comparing the amount of the gRNA at the end compared to the plasmid pool used for production of lentiviral vectors (Figure 4.7A). Results were compared between HEK293 wild type cells and MRM1 KO cells to identify potential genetic interactions (Figure 4.7B)

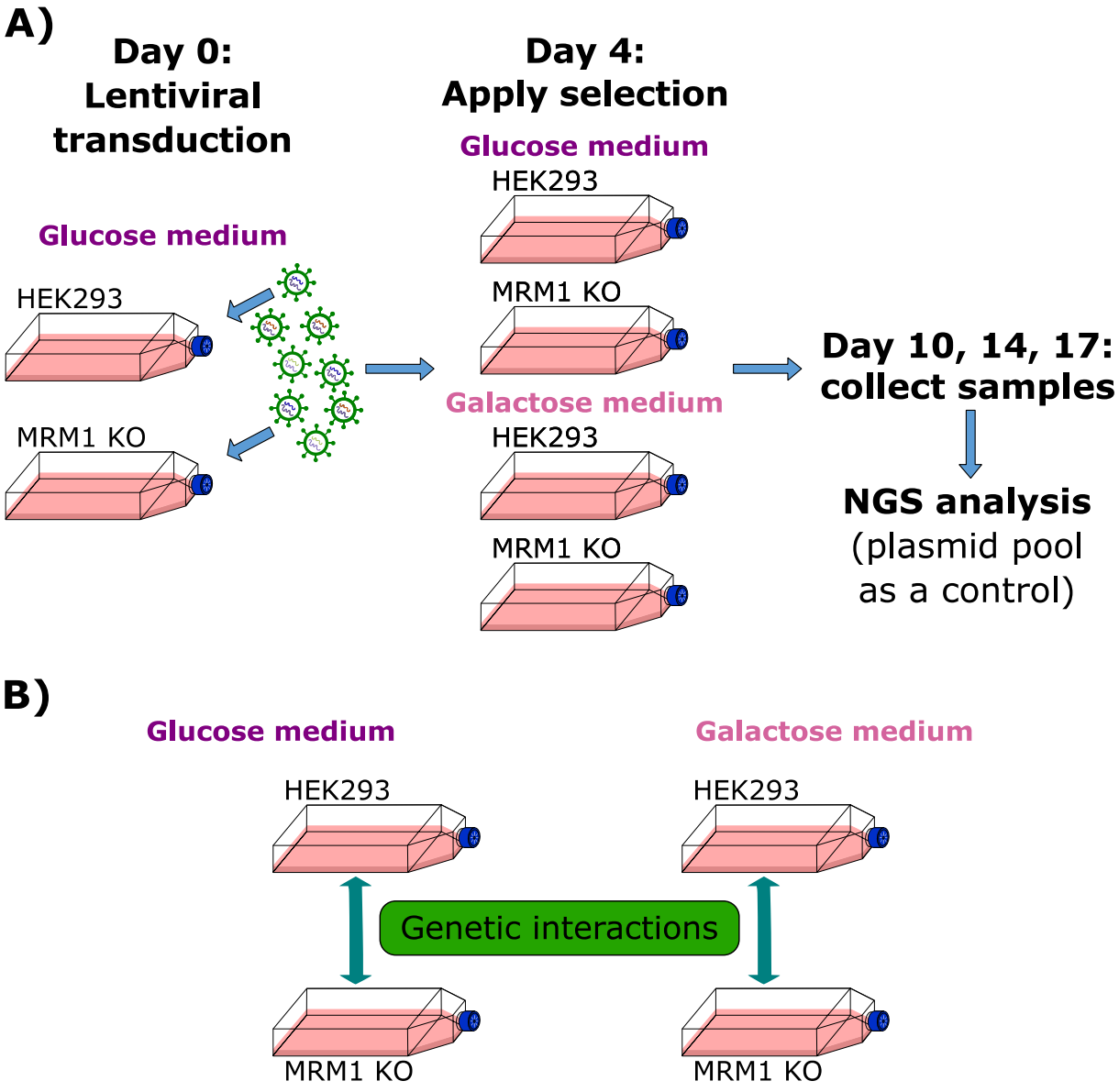


Figure 4.7: Schematic of the CRISPR/Cas9 screen design
A) HEK293 wild type and MRM1 KO cells were transduced with lentiviral library, while being cultured in a glucose medium supplemented with uridine. Four days post infection, cells were divided into glucose and galactose medium. Samples for assessment of gRNA representation were collected at days four, ten, fourteen and seventeen. Samples were analysed by next generation sequencing and the gRNA representation in each sample was compared to the plasmid pool. B) Gene essentiality was compared between HEK293 and MRM1 KO cells to identify potential genetic interactions.

To obtain a robust result from a CRIPSR/Cas9 screen, several parameters need to be considered (reviewed in Doench, 2018). To reduce changes in the gRNA distribution over time due to random chance, a sufficient number of cells need to be maintained for the

duration of an experiment as well as during sample processing. This is achieved by maintaining the number of cells in culture that correspond to several fold coverage of the screen library, typically in the range of 100 to 1000 x. A second point of consideration is the multiplicity of infection (MOI), which represents the number of lentiviral particles integrated per cell. Ideally, each cell in the pool should be transduced by a maximum of one viral particle. This is achieved by transducing the cells at an MOI of less than one. The fraction of cells transduced with zero, one and multiple viral vectors can be estimated from the Poisson distribution (Figure 4.8). Transducing cells at MOI less than one decreases the chances of the cell being transduced with multiple viral particles. Both the screen coverage and the MOI are set as a trade-off between the possible negative effects discussed above and the number of cells required for the experiment. The higher the coverage and lower the MOI, the higher number of cells necessary for the experiment. We aimed to perform the screen at 300 x coverage and MOI of 0.3 to decrease the chance of random depletion of gRNAs from the pool and limit multiple viral integration events, respectively. As there were 4490 gRNAs in the library, 4.5 million cells were used for transduction in each experiment. Puromycin selection was applied 2 days after transduction and maintained for 5-7 days thereafter to select only infected cells.

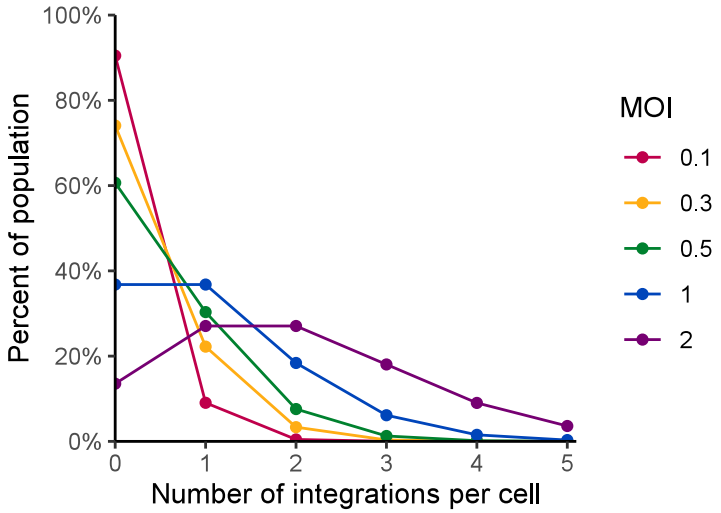


Figure 4.8: Fraction of infected cells modelled by Poisson distribution
With increasing multiplicity of infection (MOI) the fraction of uninfected cells decreases. At the same time, the fraction of cells with multiple viral integrations per cell increases.

Titration of viral supernatant has been performed to determine the volume necessary for transduction at desired MOI (Figure 4.9). Correct MOI as well as efficiency of puromycin selection were subsequently verified by FACS analysis. In transduced cells that were not subjected to puromycin selection, around 30% of cells were BFP positive, as expected with MOI of 0.3. Almost all cells were BFP positive after the puromycin selection was applied (Figure 4.10).

Viral batch	Supernatant volume	% BFP+
LV2A	4	8.7
LV2A	20	28.6
LV2A	40	45.8
LV2A	80	66.8
LV2A	140	84.1

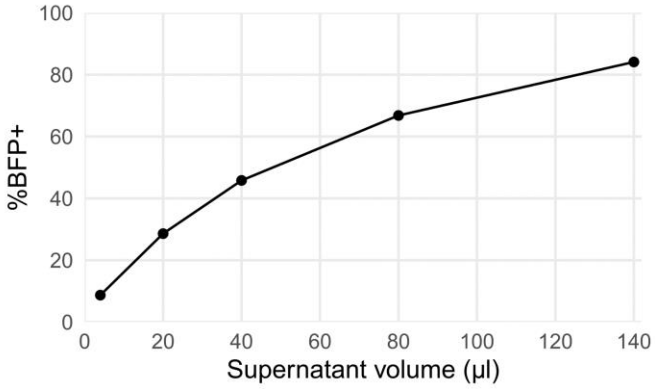
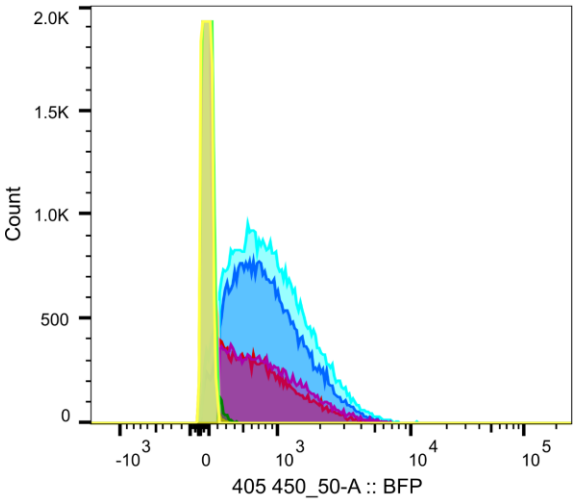


Figure 4.9: Titration of viral supernatant
Titration of viral supernatant to estimate volume necessary for 0.3 MOI. 0.1 million cells were transduced with 4-140 µl of viral supernatant. MOI was estimated by FACS measurement of fluorescence signal of BFP expression 5 day after transduction.



	Subset Name	Count	Freq. of Parent	Workspace Annotation
	BFP-	49117	99.6	HEK293
	BFP+	189	0.38	HEK293
	BFP-	47674	98.9	MRM1 KO
	BFP+	517	1.07	MRM1 KO
	BFP-	32720	66.3	HEK293, virus +
	BFP+	16658	33.7	HEK293, virus +
	BFP-	33027	69.0	MRM1 KO, virus +
	BFP+	14809	31.0	MRM1 KO, virus +
	BFP-	3637	7.36	HEK293, virus+, pruo
	BFP+	45768	92.6	HEK293, virus+, pruo
	BFP-	2498	6.45	MRM1 KO, virus+, puro
	BFP+	36249	93.6	MRM1 KO, virus+, puro

Figure 4.10: Evaluation of MOI
Whether a desired MOI was achieved was evaluated for each screen replicate. Small fraction of cells was kept aside after transduction without puromycin selection. One week after transduction, fraction of BFP positive cells was evaluated by FACS in the population of cells cultured without puromycin to estimate MOI. Untransduced cells were used as a negative control. Analysis of cells that were subjected to puromycin selection confirmed effectivity of puromycin selection.

4.2.4 CRISPR/Cas9 screen optimisation

During the pilot CRISPR/Cas9 experiment we encountered a few issues with the experimental pipeline that required optimisation. First, as the viral titration was performed in 6 well plate format with only approximately 2% of the necessary cells, the calculated volume of the viral supernatant needed for the transduction had to be scaled up for the purposes of the screen. FACS verification of the MOI of the pilot experiment showed that cells were transduced at MOI 0.5 rather than 0.3 (same approach as in Figure 4.10, data not shown). The same viral batch was used for further replicates of the screen, therefore, the volume of viral supernatant used was adjusted based on the results of the pilot experiment to achieve the MOI of 0.3.

The second issue we observed was a large number of infected cells cultured in galactose medium were found detached in the cell culture flask. HEK293 cells are typically grown in adhesion culture. Nevertheless, presence of these cells in suspension is unlikely to signify a contamination with a different cell line as HEK293 cells detach from the untreated surface of the cell culture vessel very easily. Although the presence of cells in suspension often signifies cell death, when these cells were re-plated into new vessel, they attached to the surface and proliferated further. This might suggest that suspension cells were not dead and that they could be included in the collected and passaged cell samples. Furthermore, CRISPR screens have been performed in human cell lines that grow in suspension (for example Horlbeck et al., 2018; Wang et al., 2015), which suggests that dead cells in suspension does not compromise the results of an experiment.

To clarify whether there are any differences between surface attached and detached cell states in the galactose medium, samples from adherent and suspension cells were collected and analysed separately in the first biological replicate of the screen. If there were differences in viability status between these two cell states it should be measurable by the changes in the gRNAs represented in the samples, specifically by decreased representation of positive control gRNAs (which target known essential genes) in the suspension population. We were able to isolate sufficient genomic DNA from both cell populations and compared the gRNA representation between them. The number of normalised reads per gRNA between suspension and adherent cell populations was well correlated (R between 0.96-0.99, Figure 4.11). Based on this information, in further biological replicates of the screen adherent and suspension cells were analysed together.

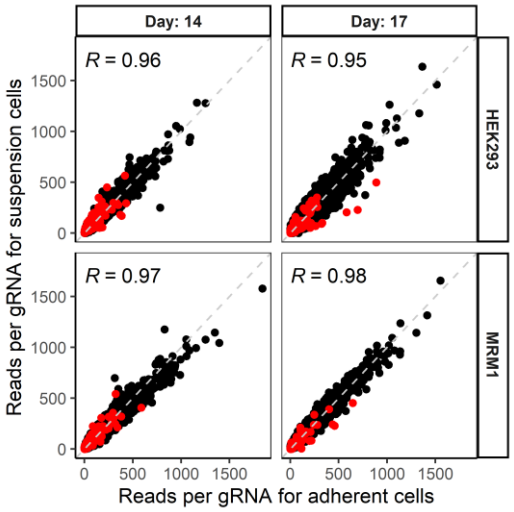


Figure 4.11: Normalised reads per gRNA in detached and adherent cell populations
Representation of individual gRNAs in suspension and adherent cell populations of HEK293 wild type and MRM1 KO cell lines was assessed separately. Read numbers were normalised to the median of gRNA counts of negative control genes in each sample. R number in individual plots represents Pearson's correlation between the suspension and adherent cell states. gRNAs of positive control genes are shown in red; gRNAs of all other genes are shown in black. Grey dashed line is defined by x = y.

4.2.5 Analysis of screen results

To evaluate the results of CRISPR/Cas9 screen in MRM1 KO and in HEK293 cells we used a combination of custom scripts and published software tools. To ensure reproducibility, a total of 3 independent biological replicates of the screen were performed. In a viability screen, such as ours, gene effects are evaluated based on the changes in representation of gRNAs in the cellular pool throughout the experiment. We assessed representation of gRNA by next generation sequencing. Sequencing reads were mapped to the gRNA library using a custom Python script, producing a count table of abundance of individual gRNAs in different conditions. We used JACKS tool to calculate log2 fold change of gRNA representation between plasmid pool and later timepoints, and to estimate gene essentiality (Allen et al., 2019). This tool summarises the information from all gRNAs per gene and all replicates into a single gene effect value. Before doing so, it is informative to verify whether major differences exist between screen replicates. The Pearson correlation analysis indicated good reproducibility among replicates (Figure 4.12). Analysis was performed for all time points, and samples from day 14 are shown as a representative example.

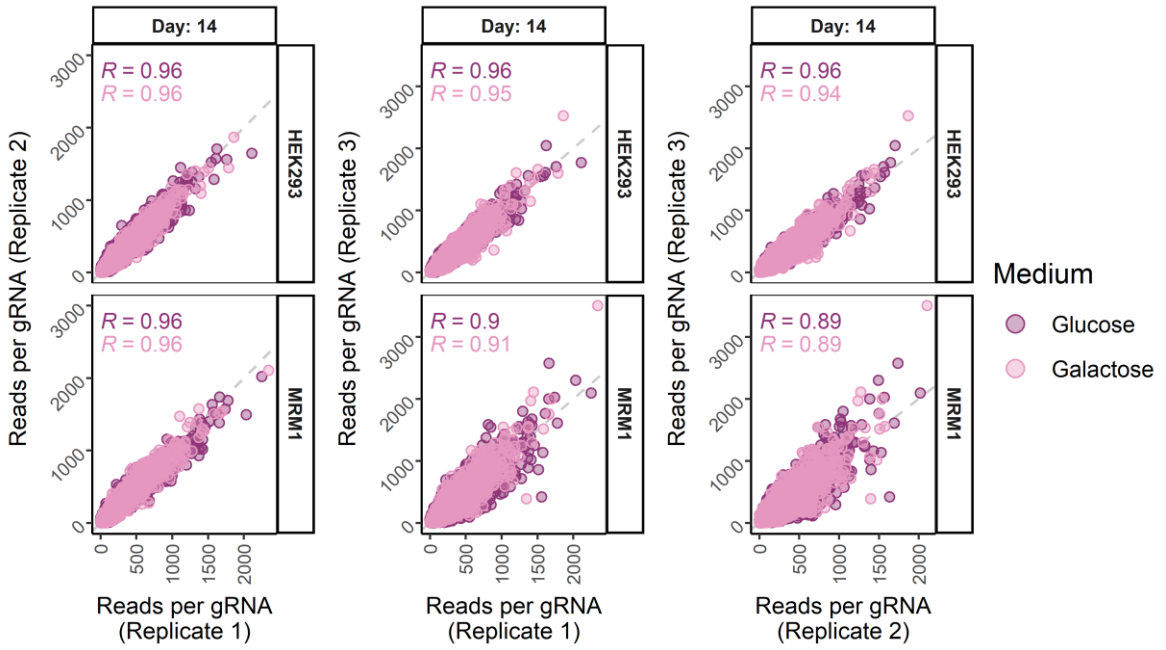


Figure 4.12: Comparison of normalised reads per gRNA among replicates
Representation of all gRNAs was assessed for each individual replicate. Read numbers were normalised to the median of gRNA counts of negative control genes in each sample. R number in individual plots represents Pearson's correlation between the replicates indicated by axis labels. Samples grown in glucose and galactose medium are shown together and differentiated by colour (purple glucose, pink galactose). Grey dashed line is defined by $x = y$.

4.2.5.1 Genetic interactions of MRM1

Genetic interaction is a term used for describing when a knockout of a combination of genes has a more severe effect on cell phenotype as compared to simple addition of effects of single knockouts. To identify genetic interactions of MRM1, we compared gene essentiality scores in MRM1 KO cell line and in HEK293 wild type cell line. We define a "diffScore" as a difference between the essentiality score in MRM1 KO cell line and in HEK293 KO cell line. Positive diffScore means that knockout of the gene had a more severe effect in MRM1 KO as compared to the wild type HEK293, and therefore suggests a genetic interaction of the gene with MRM1.

Several genes were identified as potential genetic interactors of MRM1 (Figure 4.13). Among the genes with the most significant difference in effect between MRM1 KO and HEK293 were structural components of mitochondrial ribosome, genes involved in mtDNA replication, modification of tRNA and rRNA, tRNA aminoacylation and translation. Several genes from the positive control group were identified as significantly more essential in MRM1 KO, especially in the early time point (Figure 4.13, Day 10). This might suggest that MRM1 KO has a slightly lower fitness than the wild type cells, that is in this case manifested by faster depletion of gRNAs targeting known essential genes from the pool. Mitochondrial rRNA methyltransferase 2 (MRM2) was identified as significantly more essential in MRM1 KO in all tested conditions. It had the highest diffScore at day 14 in both glucose and galactose medium, at day 17 in galactose medium and second highest score at day 17 in glucose medium. As the gene that consistently appeared in all tested conditions, this gene was selected for further characterisation of genetic interaction with MRM1. The full list of genes that were significantly more essential in MRM1 KO is provided in Table 4.1.

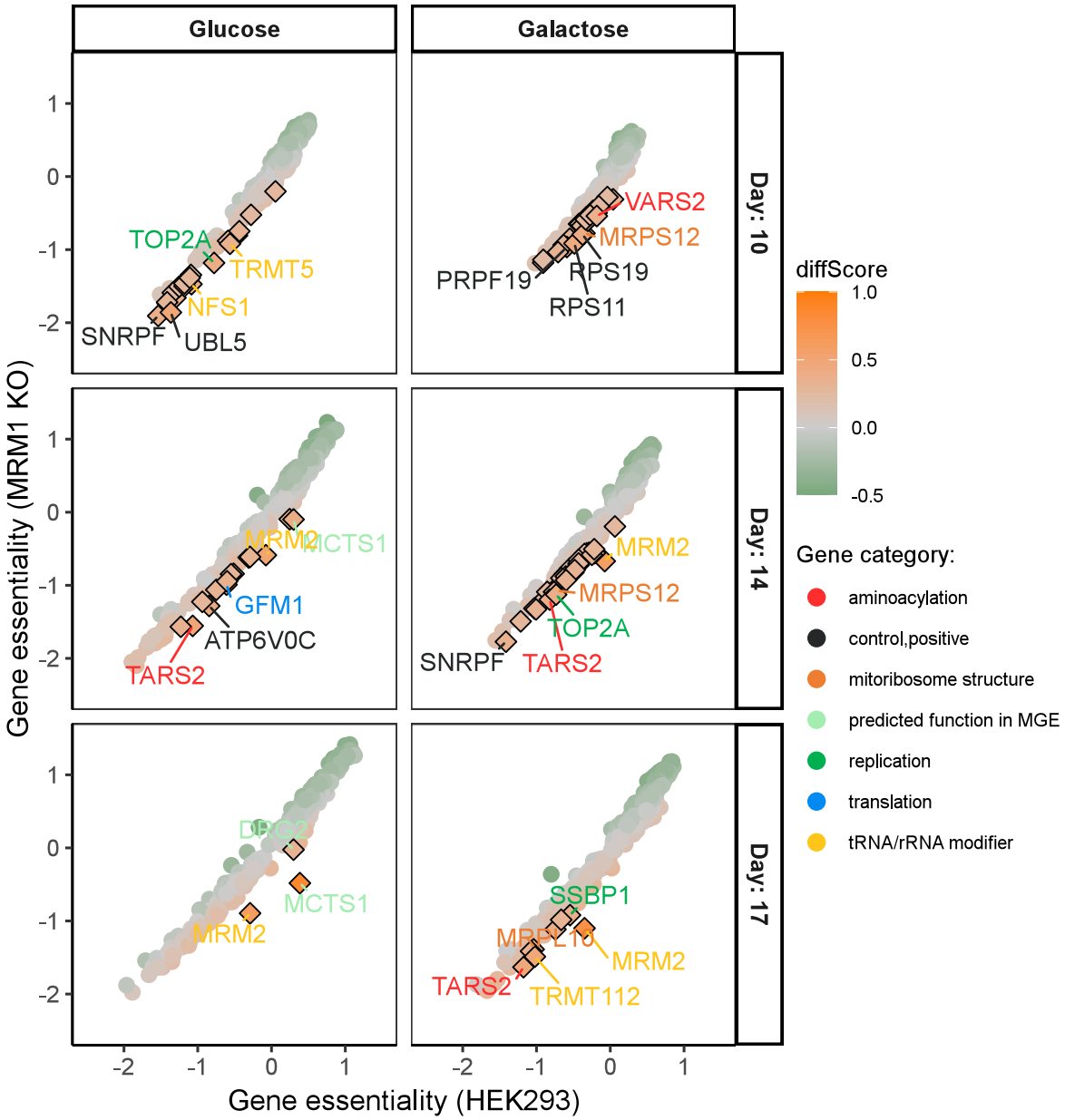


Figure 4.13: Genetic interactions of MRM1 and other genes of MGE
Genetic interactions were identified by comparing effects of gene knockout in MRM1 KO cell line to effects in wild type cells. Gene essentiality was calculated using JACKS software. We define a diffScore as a difference between gene essentiality in MRM1 KO and wild type. P-values were assigned based on the area under the curve of normal distribution fitted to the diffScore values of negative control genes. P-value threshold was set at 0.01 and Bonferroni correction for multiple comparisons was applied. Values of gene essentiality are shown for individual genes depicted either as circles or as diamonds for genes with significantly different diffScore and coloured by the diffScore value. Genes with five highest diffScore values are labelled, with gene names coloured based on their function category.

Table 4.1: Candidate genetic interactions of MRM1

Genes with statistically significant diffScores in at least one of the conditions. Conditions in which the gene reached significance are indicated by green box. Gal = galactose, Glc = glucose, DPI = day post infection.

Gene	Category	Gal DPI10	Gal DPI14	Gal DPI17	Glc DPI10	Glc DPI14	Glc DPI17
MRM2	16S rRNA modifier	■	■	■	■	■	■
SARS2	aminoacylation		■			■	
TARS2	aminoacylation	■	■	■	■	■	
VARS2	aminoacylation	■	■			■	
ATP6V0C	control,positive					■	
GEMIN5	control,positive		■				
KPNB1	control,positive	■			■		
NOP16	control,positive		■				
PRPF19	control,positive	■	■		■		
RPL14	control,positive		■		■		
RPL3	control,positive			■	■		
RPS11	control,positive	■			■		
RPS19	control,positive	■	■		■		
SMU1	control,positive				■		
SNRPD1	control,positive				■		
SNRPF	control,positive	■	■		■		
UBL5	control,positive	■			■		
GTPBP5	mitoribosome assembly	■	■	■			
RMND1	mitoribosome assembly						■
mS29	mitoribosome structure	■	■			■	
uL10m	mitoribosome structure	■	■	■		■	
uL15m	mitoribosome structure		■				
uL2m	mitoribosome structure		■				
bL28m	mitoribosome structure		■				
bL35m	mitoribosome structure	■	■				
mL39	mitoribosome structure	■	■				
mL43	mitoribosome structure	■	■				
mL45	mitoribosome structure	■					
uL29m	mitoribosome structure		■				
uS12m	mitoribosome structure	■	■		■		
mL66	mitoribosome structure				■		
mS40	mitoribosome structure				■		
bS21m	mitoribosome structure	■	■				
mS25	mitoribosome structure	■					
mS35	mitoribosome structure		■				
mS39	mitoribosome structure	■					
uS5m	mitoribosome structure	■	■				
bS6m	mitoribosome structure	■	■				
DRG2	predicted function in MGE					■	■
MCTS1	predicted function in MGE					■	■
NELFE	predicted function in MGE	■	■		■	■	
PSMA6	predicted function in MGE	■			■		
DNA2	replication		■				
SSBP1	replication	■	■	■		■	
TOP2A	replication	■	■	■	■		
GRSF1	RNA metabolism other		■				
KIAA0391	RNA metabolism other	■	■				
LRPPRC	RNA metabolism other	■	■				
PTCD1	RNA metabolism other	■	■			■	
SUPV3L1	RNA metabolism other	■	■			■	
GFM1	translation					■	
TRMT5	tRNA-modifier		■	■	■	■	
NFS1	tRNA-modifier candidate	■	■		■		
TRMT112	tRNA-modifier candidate		■	■	■		

4.2.6 MRM2 as a top candidate for MRM1 genetic interaction partner

The CRISPR/Cas9 screening approach allowed an unbiased search for genetic interactions of MRM1. From the identified candidate genes, we selected MRM2 for further verification as it consistently appeared among the top hits across the studied conditions. Furthermore, the role of MRM1 and MRM2 proteins in the cell are closely related. Both MRM1 and MRM2 are 2'-*O*-ribose methyltransferases acting on 16S rRNA. MRM2 is responsible for methylation of U1369 in the A loop (Lee and Bogenhagen, 2014) and G1145 modified by MRM1 is located in the P loop of the 16S rRNA. Both U1369 and G1145 residues are in the close proximity of the 3'CCA of the tRNAs when in the A and P ribosome sites, respectively. Knockout of MRM2 alone causes a severe translation deficiency and stalling of the assembly of the large mitoribosomal subunit in the late assembly state (Rebelo-Guiomar et al., 2021). MRM1 has been found to associate with the nucleoid (Lee et al., 2013) suggesting Gm1145 modification might be deposited early in the mitoribosome assembly. We hypothesised that simultaneous inactivation of both MRM1 and MRM2 would cause disruption of the mitoribosome in the earlier stages of the assembly.

4.2.6.1 Production of MRM1/2 double knockout cell line

To further characterise the potential genetic interaction between MRM1 and MRM2 genes, we produced an MRM1-MRM2 double knockout (MRM1/2 KO) cell line. As a parental cell line, we used MRM1 KO cells, in which we inactivated MRM2 with the use of CRISPR/Cas9 system. MRM1 KO cells were transfected with tracrRNA and three crRNAs targeting exon 2 of MRM2 gene. Following transfection, cell suspension was diluted and seeded into 96 well plates. Cells were maintained in a medium supplemented with uridine to prevent cell death in case of severe mitochondrial deficiency. Once colonies that originated from single cells could be observed, they were expanded and screened for the presence of MRM2 protein. MRM1/2 KO clones were identified by western blotting (Figure 4.14A) and sequence analysis of targeted region (Figure 4.14C). Only a single band was observed after the PCR amplification of the targeted region (Figure 4.14B). This could suggest a presence of only one allele of chromosome 7 in this clone of HEK293 cells, which are known to suffer from various chromosome aberrations (Lin et al., 2014). Alternatively, repair of the CRIPSR/Cas9 produced double strand break might have resulted in a large deletion, leading to the loss of one of the primer binding sites. In the identified allele, there was a two-nucleotide insertion in exon 2 of MRM2 gene, leading to a premature stop codon (Figure 4.14C).

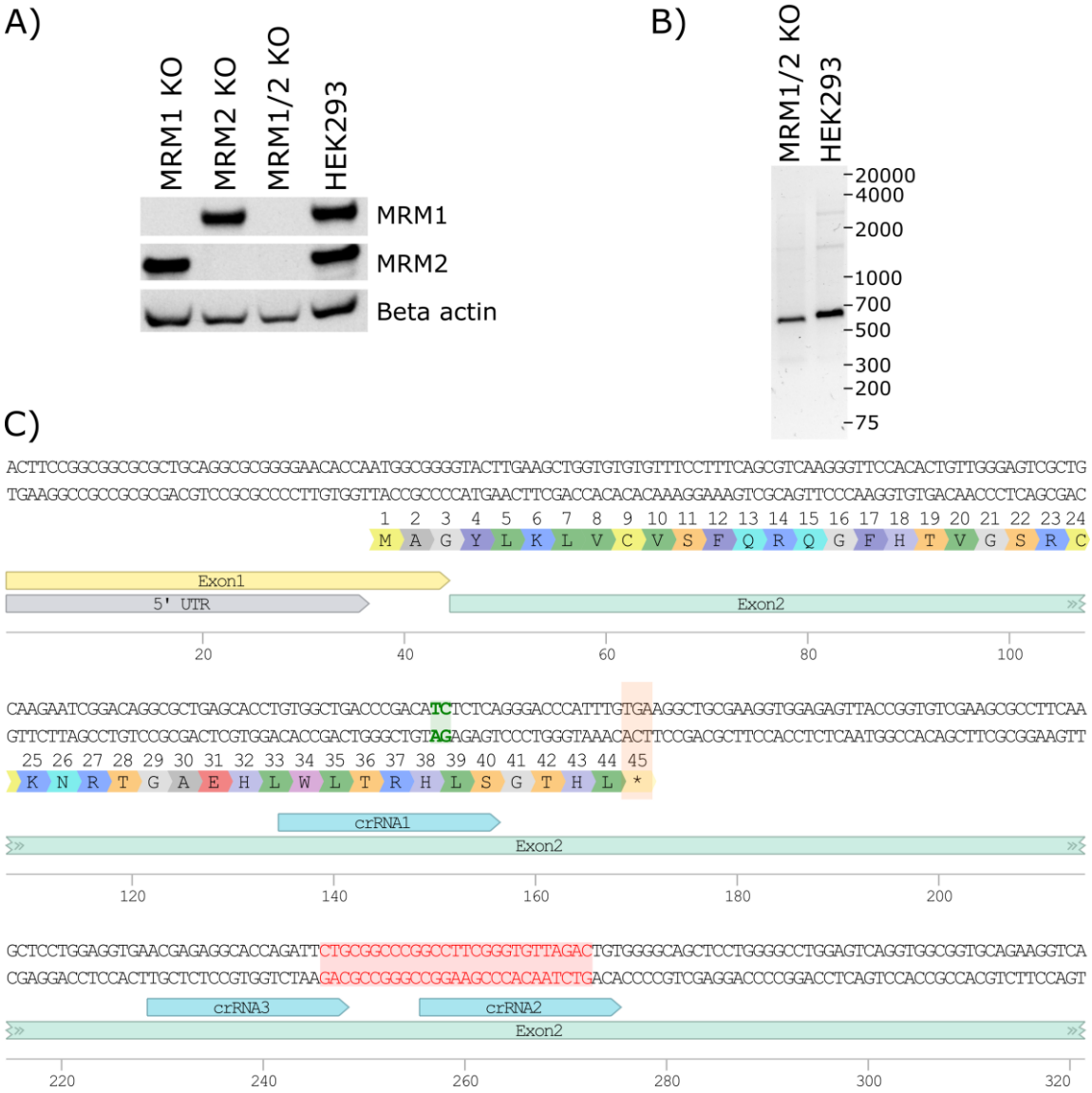


Figure 4.14: Production of MRM1/2 double knockout cell line

A) Western blot analysis of clone B1 of MRM1/2 double knockout cell line showed absence of both MRM1 and MRM2 proteins in this cell line. Beta actin was used as a loading control. B) PCR amplification of the region of MRM2 that was targeted by crRNAs showed a single band for both wild type and knockout cells. C) Sequence analysis of MRM1/2 double knockout cell line. DNA sequence, corresponding translation and sequence annotations are shown. Blue arrows with crRNA depict the region that was targeted for Cas9 restriction. A two-nucleotide insertion (bold green) leads to a premature stop codon in the exon 2 (star, orange box). This exon also contains a deletion of 27 nucleotides (red box). As premature stop codon appears upstream of this deletion, it does not influence the resulting protein. DNA sequence was visualised in Benchling and manually edited. (Benchling [Biology Software]. (2021). Retrieved from https://benchling.com)

4.2.6.2 Characterisation of MRM1/2 double knockout cell line

To understand the nature of interaction between MRM1 and MRM2 proteins, mitochondrial function of MRM1/2 KO was characterised and compared to the mitochondrial function of MRM1 and MRM2 single knockouts. In all following experiments MRM1 KO refers to the ZFN produced MRM1 KO cell line with Cas9 protein integrated into FRT site that was used for CRSPR/Cas9 screen (see section 4.2.2), and MRM2 KO to the cell line created by Pedro Rebelo-Guiomar (Rebelo-Guiomar et al., 2021).

Previous study of MRM1 KO cells did not show any detectable mitochondrial dysfunction (D'Souza, 2018), however knockout of MRM2 resulted in strong mitochondrial deficiency (Rebelo-Guiomar et al., 2021), therefore limiting the number of assays in which difference between MRM2 KO and MRM1/2 KO could be detected.

First, we assessed the effect of ablation of mitochondrial methyltransferases on the steady state levels of OXPHOS components. Strong decrease of components of Complexes I and IV was observed in MRM2 and MRM1/2 KO cells. Small decrease was also observed for Complex II component SDHB. Steady state levels of all studied OXPHOS components were comparable between wild type and MRM1 KO cells (Figure 4.15).

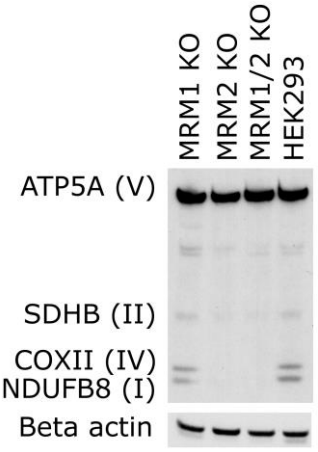


Figure 4.15: Analysis of OXPHOS in double KO cell line
Steady state levels of components of OXPHOS complexes were assessed by western blot analysis of whole cell lysates from wild type HEK293, MRM1 KO, MRM2 KO and MRM1/2 KO cells. Beta actin was used as a loading control. 3 biological replicates were performed, and representative result is shown.

Observed changes in steady state protein levels could have been caused by impaired translation or by perturbation of earlier stages of gene expression. Considering that both MRM1 and MRM2 are methyltransferases acting on 16S rRNA, dysfunctional translation seemed like the most likely explanation for the observed phenotype. To confirm this, we examined the steady state levels of mtDNA and RNA by RT-qPCR. A delta-delta Ct (ΔΔCt) method was used for quantification. In this approach, the Ct value, which represents the cycle number when the detected fluorescence reaches the threshold value, is determined for each studied gene and each sample. The Ct value of a reference gene is subtracted from the Ct of each experimental gene and subsequently Ct value of reference sample is subtracted from the Ct of each experimental sample. Results are usually presented in the format of $2^{-\Delta\Delta Ct}$ (Livak and Schmittgen, 2001). We used B2M as a reference gene for mtDNA copy number quantification and GAPDH as a reference for RNA levels. Each gene was quantified in biological triplicates, and each biological replicate was performed in technical triplicates. The average of the biological replicates for HEK293 cell line was used as a reference sample. No statistically significant change of mtDNA copy number was observed

(Figure 4.16A). mRNA levels of mitochondrially encoded Complex IV components (mt-CO1 and mt-CO2) were unchanged. Steady state level of mtND1 mRNA, which encodes for a protein of Complex I, was ~5 times upregulated in MRM1/2 KO, compared to only ~3 times upregulation in MRM2 KO and no change in MRM1 KO (Figure 4.16B). Therefore, the observed decrease of protein levels in MRM1/2 KO was not a downstream effect of changes in mtDNA or mRNA levels. The increase in mtND1 transcript level suggests a possible compensatory mechanism by which the cell tries to balance the decrease in translation. We further found a statistically significant decrease in 12S and 16S rRNA levels in MRM2 and MRM1/2 KO cell lines (Figure 4.16C). We again observed an additive effect when both MRM1 and MRM2 are knocked out as downregulation of rRNA levels in MRM1/2 KO cell line was about twice as large as in MRM2 KO.

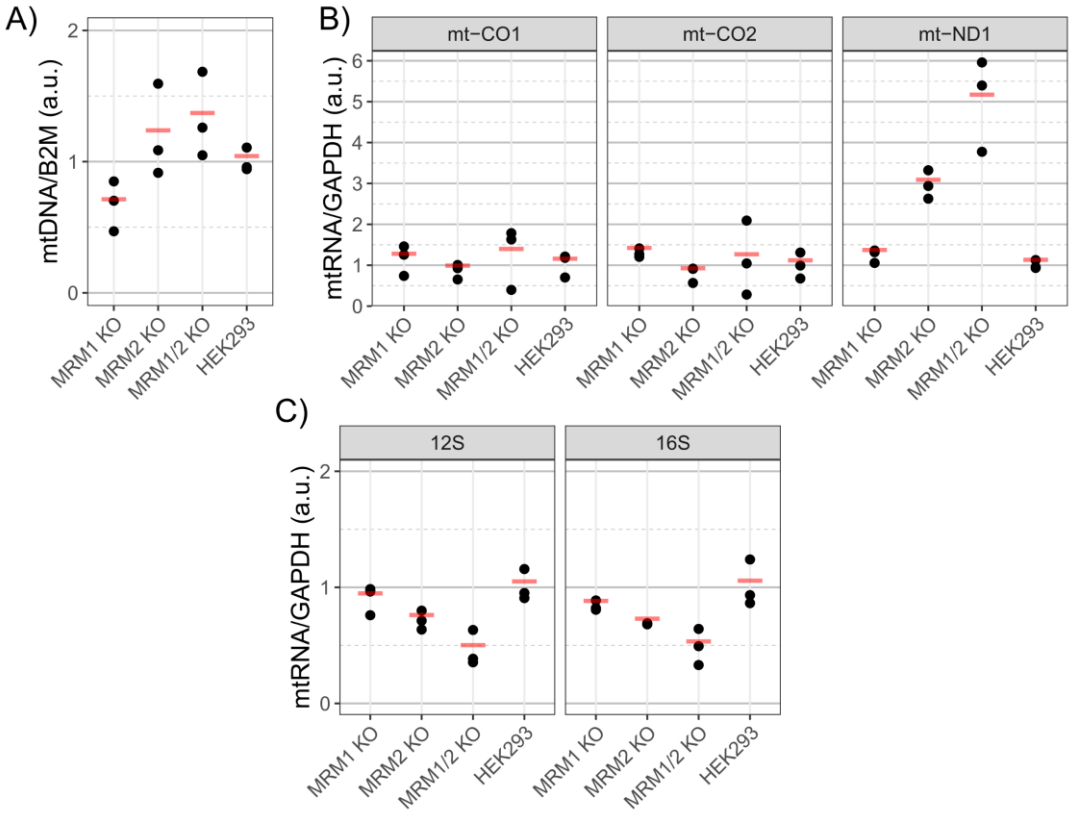


Figure 4.16: qPCR analysis of mtDNA copy number and RNA levels in MRM1/2-KO cell line
qPCR was performed on 100 ng total amount of template DNA or 5 µl of cDNA from wild type HEK293, MRM1 KO, MRM2 KO and MRM1/2 KO samples. A) mtDNA copy number normalised to B2M nuclear gene. No statistically significant change in mtDNA levels was observed. B) mt-mRNA levels compared with GAPDH nuclear gene. Mt-ND1 gene showed upregulation in MRM2 KO and MRM1/2 KO samples (p-values 0.01 and $7.5*10^{-5}$, respectively) C) mt-rRNA levels compared with GAPDH nuclear gene. Both 12S and 16S rRNA were downregulated in MRM2 and MRM1/2 KO cells (p-values 0.054 and 0.0017 for 12S rRNA for MRM2 KO and MRM1/2 KO, respectively; 0.035 and 0.0027 for 16S rRNA for MRM2 and MRM1/2 KO, respectively). Results of three biological replicates are shown in black dots, with mean of the three replicates depicted as a red line. P-values were calculated using Dunett's test.

Next, we assessed whether MRM1/2 KO cells are able to survive when grown in the medium with galactose as a sole carbon source. Cells were seeded into 6 well-plates at 5-10% initial confluency and their growth was monitored using IncuCyte imaging and analysis system. Unsurprisingly, MRM1/2 KO cells were unviable on galactose medium. In addition, we also did not observe growth in the standard glucose-containing growth medium that was devoid of uridine (Figure 4.17). To assess whether this phenotype is specific for the double knockout, we compared the growth rate of MRM1 KO, MRM2 KO and MRM1/2 KO cells in the presence and in the absence of uridine supplementation (Figure 4.18). MRM1/2 KO cells were unable to proliferate in the absence of uridine supplementation unlike MRM1 KO or MRM2 KO (Figure 4.18), a phenotype typically observed in cultured cells lacking mitochondrial DNA (Spadafora et al., 2016). Uridine is important for RNA synthesis and has a role in lipid, glucose and amino acid homeostasis (Zhang et al., 2020). Uridine dependency in cells with dysfunctional electron transport chain is linked to the deficiency of dihydroorotate dehydrogenase, a mitochondrial enzyme that is part of the *de novo* uridine biosynthetic pathway (Grégoire et al., 1984). This suggest that the level of OXPHOS function retained in MRM2 KO cells is substantially higher compared to the MRM1/2 double KO.

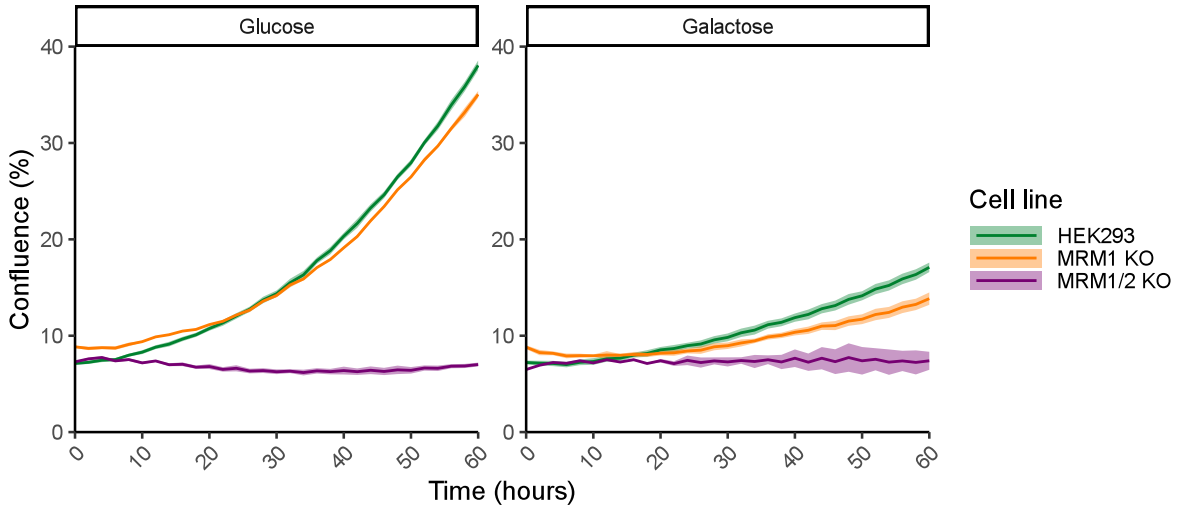


Figure 4.17: Growth curve of MRM1/2 KO in glucose and galactose medium.
Wild type HEK293, MRM1 KO, and MRM1/2 KO cells were grown in the media containing 4.5 g/L glucose or 10 mM galactose as a sole carbon source. MRM2 KO cells were not included in this experiment as it was previously shown that they are unable to grow in galactose medium (Rebelo-Guiomar et al., 2021). Each line is an average of three experiments and shaded areas represent standard deviation.

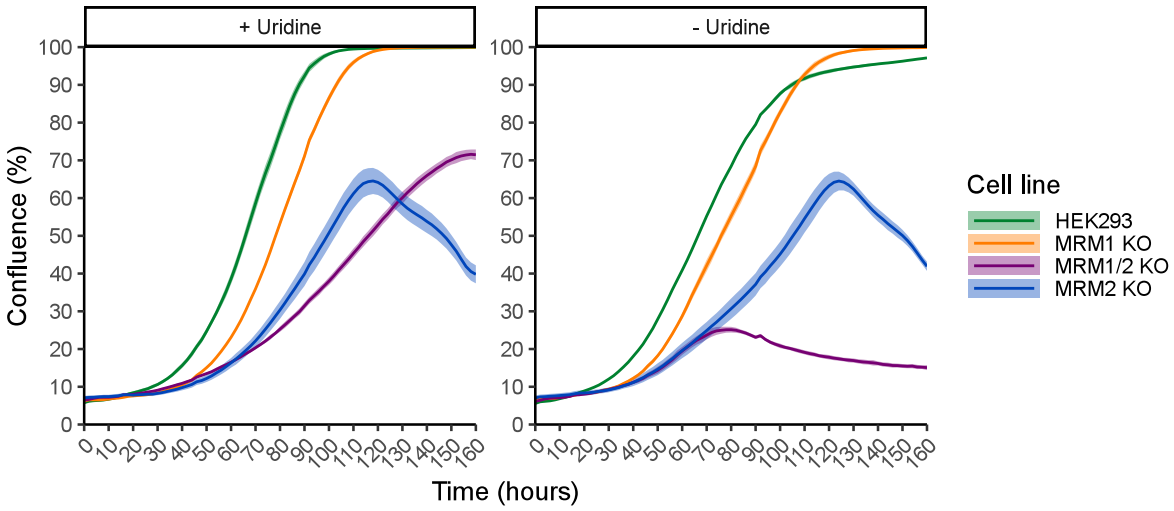


Figure 4.18: Growth curve of MRM1/2 KO in the presence or absence of uridine supplementation
Wild type HEK293, MRM1 KO, MRM2 KO and MRM1/2 KO cells were grown in the media containing 4.5 g/L glucose, 20% FBS and 50 µg/ml uridine (+ Uridine) or 10 % FBS without uridine (- Uridine). Double knockout cells were unable to grow without uridine supplementation. There was no difference in growth rate between 10% and 20% FBS with uridine supplementation (data not shown). Each line is an average of three experiments and shaded areas represent standard deviation.

4.2.6.3 Complementation of MRM1/2 double knockout

When knocking out a gene is selected as an experimental approach, it is important to verify that the observed phenotype is caused by the disruption of target gene, and not by off-target effects. In MRM1/2 KO cell line, two different methods were used for the disruption of target genes. MRM1 was knocked out with the use of ZFN (D'Souza, 2018), whereas MRM2 was knocked out using CRIPSR/Cas9 (see 4.2.6.1). Both methods can produce double stranded breaks in non-targeted regions of DNA (off-targets), which can lead to disruption of genes and cause unintended phenotypic defects. In the case of ZFNs, off-targets are produced when zinc finger proteins have low specificity to their target site. Overexpression of ZFNs on plasmids also increases the chance of off-target binding and cleavage. Specificity of CRIPSR/Cas9 system is mostly affected by the gRNA sequence. In recent years various improvements have been made to both ZFN and CRISPR/Cas9 gene editing tools to reduce off-target effects such as the use of enzymes that produce a single stranded nick as opposed to double stranded breaks (Chiang et al., 2016; Kim et al., 2012a; Ramirez et al., 2012; Ran et al., 2013; Shen et al., 2014; Wang et al., 2012).

To exclude the possibility that observed phenotype of MRM1/2 KO cell line is a result of off-target changes, we complemented this cell line with MRM1 or MRM2 proteins. Wild type MRM1 and MRM2 cDNAs were cloned into pWPLXd:IRES:puroR plasmid which was used for lentiviral production. The lentiviral particles were transduced into MRM1/2 KO and to HEK293 cells as a control. To evaluate the effect of viral transduction on cells, both

cell lines were also transduced with the lentiviral particles produced using a pWPLXd:IRES:puroR plasmid without a transgene. Puromycin antibiotic selection was used to identify transduced cells. Viral titer was not assessed, therefore varied number of integration events might have happened in each cell. Both HEK293 and MRM1/2KO cell lines were transduced with an equal amount of virus containing supernatant from the same batch. Due to random integration of lentiviruses, which might affect the phenotype of the cell based on the site of integration and the number of the lentiviral particles that transduced a cell, we did not perform clonal selection. Using a bulk population, where lentiviral particles integrated into various sites in different cells, should cancel out any clonal effects caused by potential disruption of genes by transgene integration. After integration, constitutive expression of cDNAs was driven by human elongation factor-1 α promoter. Expression was confirmed by western blot analysis (Figure 4.19).

We next assessed whether overexpression of either MRM1 or MRM2 in the double knockout cell line could rescue the observed phenotypic defects. When MRM1 protein is expressed in MRM1/2 KO cells, the expected phenotype should be equivalent to the MRM2 KO, therefore we expected these cells to show mitochondrial deficiency. Expression of MRM2 protein in MRM1/2 KO cell is expected to result in a cell line equivalent to MRM1 KO, with near wild type phenotype. First, we looked at the steady state levels of OXPHOS components (Figure 4.20). Consistent with the previous results (see Figure 4.15), we did not detect components of Complexes I, III and IV in either the MRM1/2 KO cells or cells transduced with the empty vector. These proteins were also not detected in cell line complemented with MRM1 protein. Only complementation with MRM2 cDNA led to the rescue of the OXPHOS protein expression to the levels comparable with the wild type HEK293.

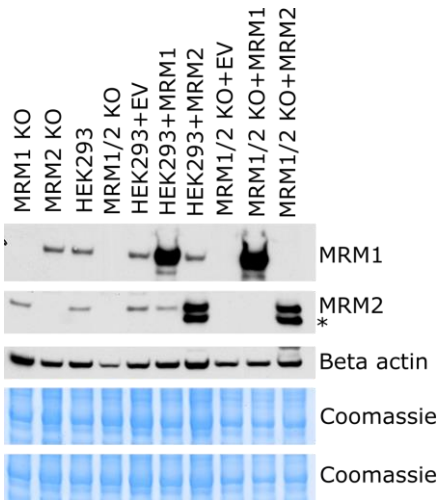


Figure 4.19: Western blot analysis of MRM1/2 KO complementation
MRM1/2 KO cells were complemented with wild type cDNAs of MRM1 or MRM2 proteins or with empty lentiviral vector (EV) as a control. Expression of MRM1 and MRM2 proteins was assessed by western blot analysis. Beta actin and SimplyBlue Safe Stain Coomassie staining were used as a loading control (top Coomassie gel corresponds to the membrane probed with MRM1 antibody, bottom to the membrane probed with MRM2 and Beta actin antibodies). Unspecific band when membrane was probed with MRM2 antibody is denoted by asterisk.

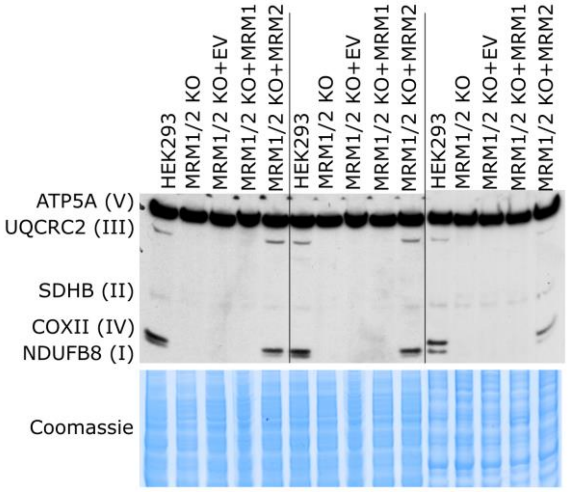


Figure 4.20: Western blot analysis of OXPHOS in MRM1/2 KO after complementation
Equal amounts of total protein lysates from wild type HEK293, MRM1/2 KO and MRM1/2 KO cells complemented with empty viral vector (EV), MRM1 or MRM2 proteins were analysed by western blotting. SimplyBlue Safe Stain Coomassie staining was used as a loading control. Results of three biological experiments are shown.

In the majority of experiments, we observed that combination of MRM1 and MRM2 knockouts led to a more severe phenotype as compared to the MRM2 KO alone. Uridine dependency was a phenotypic characteristic observed exclusively in the MRM1/2 KO cell line. We analysed growth of MRM1/2 KO after complementation in the presence and in the absence of uridine, to check whether this phenotype can be rescued by expression of wild type proteins (Figure 4.21A). Expression of both MRM1 and MRM2 in the double knockout rescued the uridine dependency, while MRM1/2 KO cells transduced with empty vector were unable to grow in the absence of uridine. Complementation with MRM1 cDNA did not have an effect on the growth rate. Cells complemented with MRM2 cDNA were

able to grow faster compared to the MRM1/2 KO, however their growth rate was still slower than the wild type HEK293 cells. Transduction with empty vector, MRM1 or MRM2 cDNAs had no effect on growth of HEK293 cells.

This observation of a phenotype specific for the disruption of MRM1 allowed us to ask whether the uridine dependency is caused by the absence of the protein or the absence of the 2'-*O* methyl modification deposited at G1145 of 16S rRNA by MRM1. There are no structural data with assignment of catalytic residue of the human MRM1 protein available. Therefore, we looked for such data in MRM1 homologs of other organisms. RlmB is a MRM1 homolog in bacteria, and a crystal structure of this protein in *E. coli* has been determined at 2.5 Å resolution (Michel et al., 2002). In this study they identified Glu198 and Ser224 of the bacterial enzyme as the most likely candidates for the catalytic residues. Based on personal communication with T.L. Mason they further state that mutation of equivalent Glu358 in the methyltransferase Pet56, which is a yeast homolog of MRM1, abolished its methyltransferase activity. We used Clustal Omega to align protein sequences of human MRM1, yeast Pet56 and bacterial rlmB (Goujon et al., 2010; Sievers et al., 2011). Residues that were identified as potential catalytic sites by Michel et al. are conserved between these three species and correspond to residues Glu259 and Ser289 in human MRM1 protein (Figure 4.22). We cloned MRM1 cDNA with either Glu259 to Ala (MRM1 E259A) or Ser289 to Ala (MRM1 S289A) mutation into pWPLXd:IRES:puroR viral transfer plasmid and used these for complementation (Figure 4.23). Neither E259A nor S289A mutant MRM1 proteins were able to rescue the uridine dependency of MRM1/2 KO cells (Figure 4.21B). This suggests that it is not the sole presence of MRM1 protein but its methyltransferase activity that is important for the mitochondrial function.

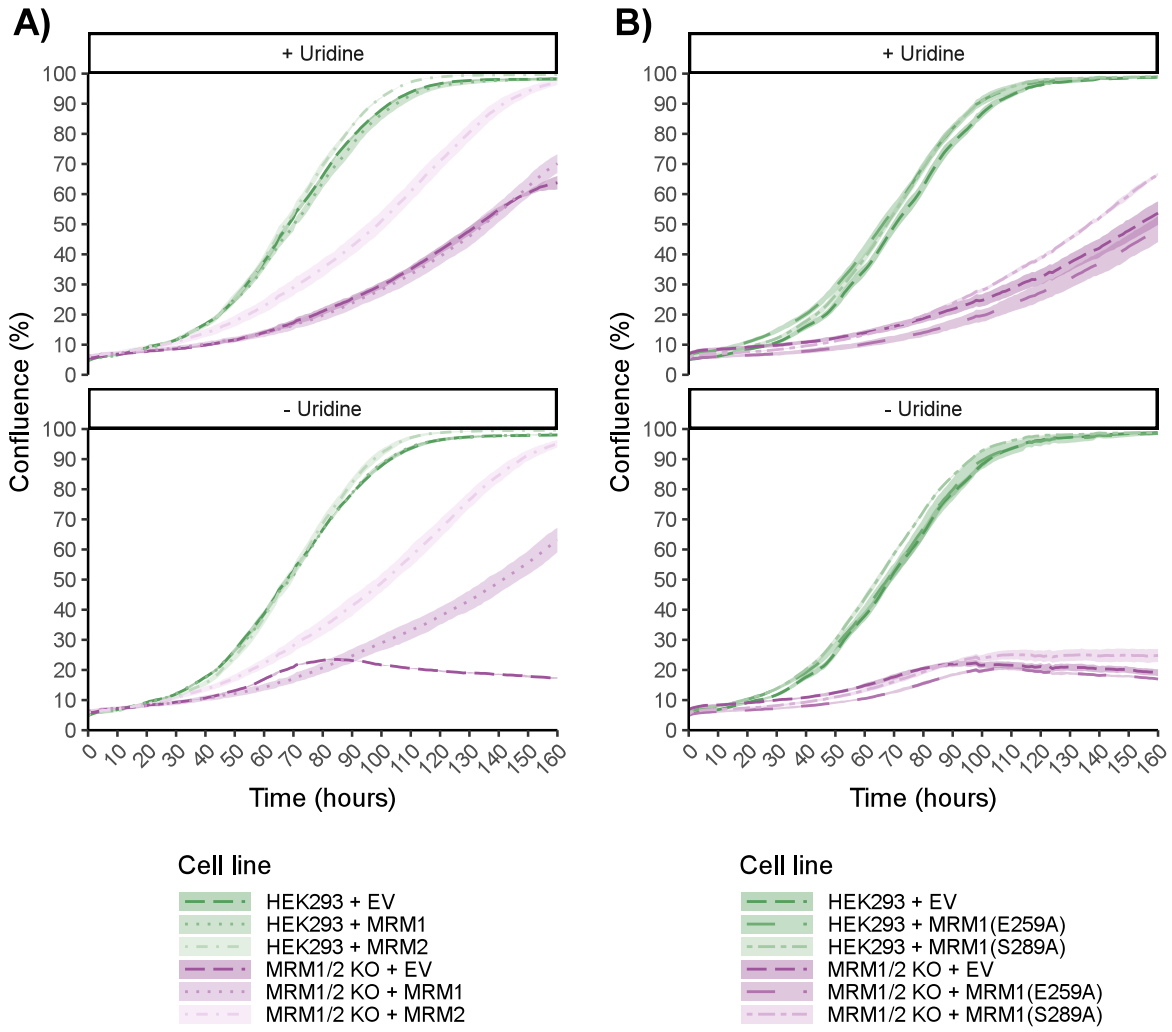


Figure 4.21: Growth curve of MRM1/2 KO after complementation
Cells were grown either in the presence of 50 µg/ml uridine (+ Uridine) or without uridine (- Uridine). Growth was monitored using IncuCyte imaging system. Each line is an average of three experiments and shaded areas represent standard deviation. A) MRM1/2 KO was complemented with wild type MRM1 or MRM2 proteins (MRM1/2 KO + MRM1, MRM1/2 KO + MRM2, respectively). Knockout cells complemented with empty vector (MRM1/2 KO + EV), and wild type HEK293 cells complemented with MRM1, MRM2 and EV were used as a control. B) Evaluation of importance of Gm1145 modification on uridine dependency. MRM1/2 KO cells were complemented with catalytically inactive mutant MRM1 protein, MRM1(E259A) or MRM1(S289A). As a control, transductions were also performed with empty viral vector and in wild type HEK293 cell line.

```
MRM1        DVFSTDDLTGFLQTK-AQQGWLVAGTVGCPSTED--PQSSEIPIMSCLEFLWERPTLLVL 256
rlmB        PLIRVTNLARTMRMLQEENIWIVGTAGEADHT--LYQ-----------SKMTGRLALVM 195
pet56       PIFYVDKPLEFFTKSQEMGGWTFITSHLANATSEKYTVGKTISMHDLNGLCNELPVVLVV 355
             :: . .    :        * .  :  .  *                        **:

MRM1        GNEGSGLSQEVQASCQLLLTILP------RRQLPPGLESLNVSVAAGILLHSICSQRKGF 310
rlmB        GAEGEGMRRLTREHCDELISIPMAGSV----------SSLNVSVATGICLFEAVRQRS-- 243
pet56       GNESQGVRTNLKMRSDFFVEIPFGGIEKGNRAPEPIVDSLNVSVATALLIDNILTCK--- 412
            * *..*:    :  .: :: *                .*******:.: : .    :

MRM1        PTEGERRQLLQDPQEPSARSEGLSMAQHPGLSSGPEKERQNEG  353
rlmB        ---------------------------------------------  243
pet56       ---------------------------------------------  412
```

Figure 4.22: Sequence alignment of human MRM1 protein and its homologs
*Sequence of MRM1 protein was compared to rlmB and pet56 from E. coli and S. cerevisiae, respectively. Homologs of human (UniProtKB/Swiss-Prot: Q6IN84-1), bacterial (UniProt/Swiss-Prot: P63177-1), and yeast (UniProt/Swiss-Prot: P25270-1) MRM1 proteins were aligned using Clustal Omega multiple sequence alignment tool. Symbols below the alignment represent fully conserved residues (asterisk, *), strongly similar properties (colon, :) and weakly similar properties (period, .) as defined by Clustal Omega. Predicted catalytic residues are marked by red arrows above the alignment.*

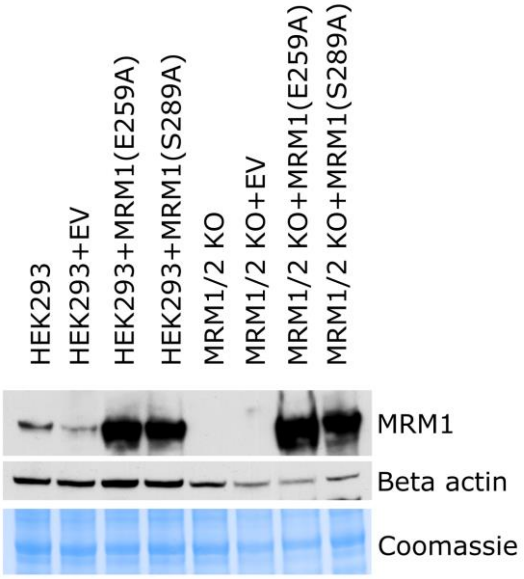


Figure 4.23: Western blot analysis of MRM1/2 KO complemented with catalytically inactive MRM1
MRM1/2 KO cells were complemented with MRM1 cDNAs bearing E259A or S289A amino acid mutations or with empty lentiviral vector (EV) as a control. Expression of MRM1 was assessed by western blot analysis. Beta actin and SimplyBlue Safe Stain Coomassie staining were used as a loading control.

4.2.6.4 Mitochondrial proteome in absence of methyltransferases

We observed a major disruption in mitochondrial function in MRM1/2 KO cells as manifested by lower amount of OXPHOS proteins and reduced growth rate (see section 4.2.6.2). For a comprehensive characterisation of the mitochondrial proteome in the absence of MRM1 and MRM2 proteins, we performed quantitative mass spectrometry analysis with the use of SILAC. In these experiments, cells were grown in a medium with lysine and arginine containing "light" (^{12}C and ^{14}N) or "heavy" (^{13}C and ^{15}N) isotopes of

carbon and nitrogen. Cells were cultured in labelling medium for at least six generations to ensure full incorporation of labelled amino acids into all proteins. Once complete labelling was achieved, differentially labelled MRM1/2 KO and HEK293 cells were mixed in equal proportion. Equal mixing was achieved by measuring protein content in a whole cell lysate of a small portion of each sample by BCA assay and combining sample volumes with equivalent protein content.

Following sample mixing, mitochondria were isolated by differential centrifugation and further purified by step sucrose gradient ultracentrifugation. Isolated mitochondria were lysed and ~20 μg of mitochondrial lysate was resolved on polyacrylamide gel electrophoresis. Gel lanes were divided into 10 equal pieces, trypsin digested into peptides and analysed by mass spectrometry. Raw mass spectrometry data were analysed using MaxQuant software to calculate a heavy to light ratio for each protein. In every experiment, a small fraction of heavy sample was saved before mixing to verify the level of heavy isotope incorporation. Only experiments where labelling achieved >95% for both arginine and lysine containing peptides were analysed further. A total of three biological replicates comparing the mitochondrial proteome of MRM1/2 KO and HEK293 were performed. In two experiments HEK293 cells were labelled with heavy amino acids and MRM1/2 KO with light amino acids (experiment A and experiment C). The third biological replicate was performed in reciprocal orientation (experiment B).

To evaluate overall quality and reproducibility between experiments, we first looked at the total number of detected proteins. The number of all detected proteins ranged from 1872 in Experiment A to 2226 in Experiment C. The majority of the proteins (1594) were identified in all experiments (Figure 4.24). We next compared the list of identified proteins to the MitoCarta 3.0 (Rath et al., 2021) to identify the fraction of mitochondrial proteome from all detected proteins (Figure 4.24). A total of 764 mitochondrial proteins were identified by at least one of the three experiments, with 620 identified in all three replicates, which represents 67.25% and 54.58% of mitochondrial proteome as defined by MitoCarta 3.0, respectively. Out of all detected proteins the mitochondrial proteins comprised around 30-35% (Figure 4.25).

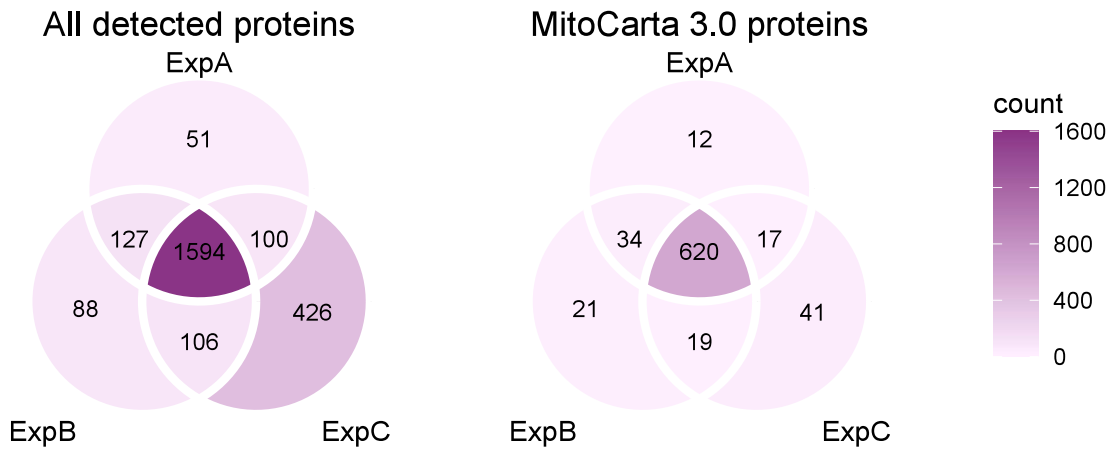


Figure 4.24: Number of detected proteins in SILAC mass spectrometry experiments
Total of three replicates of SILAC quantitative mass spectrometry analysis of mitochondrial proteome differences between MRM1/2 KO and HEK293 cell lines were performed (ExpA = experiment A, ExpB = experiment B, ExpC = experiment C). Comparison of numbers of detected proteins between the biological replicates is shown for all detected proteins and for the fraction of mitochondrial proteins listed in MitoCarta 3.0.

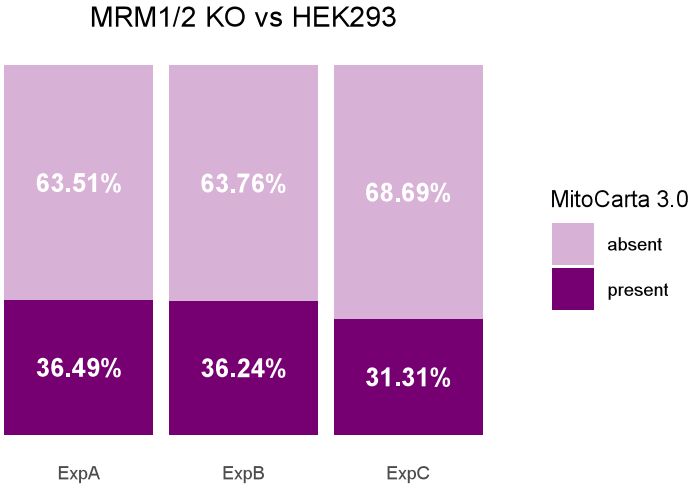


Figure 4.25: Fraction of mitochondrial proteins in SILAC proteomics experiments
For each replicate of SILAC mass spectrometry experiment ratio of mitochondrial and non-mitochondrial proteins has been evaluated. List of mitochondrial proteins was obtained from MitoCarta 3.0.

We observed substantial differences between the proteomes of MRM1/2 KO and wild type cells. For all mitochondrial OXPHOS complexes, at least some subunits were detected at significantly different levels (Figure 4.26). Complex I is the largest OXPHOS complex and is organised in six modules (N, Q, ND1, ND2, ND4 and ND5) (Signes and Fernandez-Vizarra, 2018). Five out of six proteins from Q module (NDUFA5, NDUFS2, NDUFS3, NDUFS7 and NDUFS8) were downregulated more than twofold. Other significantly downregulated proteins were NDUFA13 from ND1-module, NDUFB4 from ND4-module, NDUFS5 from ND2-module, NDUFA2 and NDUFS1 from N-module and NDFB8 that is a part of the ND5-module. The only significantly upregulated protein was NDUFB11 from ND4 module. It is important to note that the values presented are an average of three biological replicates and only proteins that were identified and quantified in all replicates

were included in this analysis. We observed downregulation of the steady state levels of all Complex II components, with SDHA being the least affected. Six out of eleven components of Complex III were detected in all three biological repeats. Of these, UQCRB was the most downregulated protein, with log2 fold change of about -2.8. CYC1 and UQCR10 proteins were significantly upregulated in MRM1/2 KO compared to the HEK293 cells. These two proteins were previously shown to be the only Complex III proteins with unchanged levels in MT-CYB null mutant (Protasoni et al., 2020), and to accumulate in assembly intermediates in cybrids with in-frame deletion in MT-CYB (Páleníková et al., 2021b). For mitochondrial Complex IV we detected both down- and up- regulated proteins. MT-CO2 was the only mtDNA encoded protein that was detected in all three biological replicates with a log2 fold change of approximately -3.9, which corresponds to almost 15-times downregulation. This provides further evidence of severe disruption of mitochondrial translation in MRM1/2 KO cells. Two protein components of Complex IV, COX5A and COX4I1, were upregulated in the double knockout. These proteins together with HIGD1A form an early intermediate in the assembly of Complex IV (Vidoni et al., 2017). Complex V components were largely unchanged, with only downregulation of USMG5 protein falling within the 10% false discovery rate. This protein is part of the membrane domain of the ATP synthase and assembles at the very late stage (He et al., 2018).

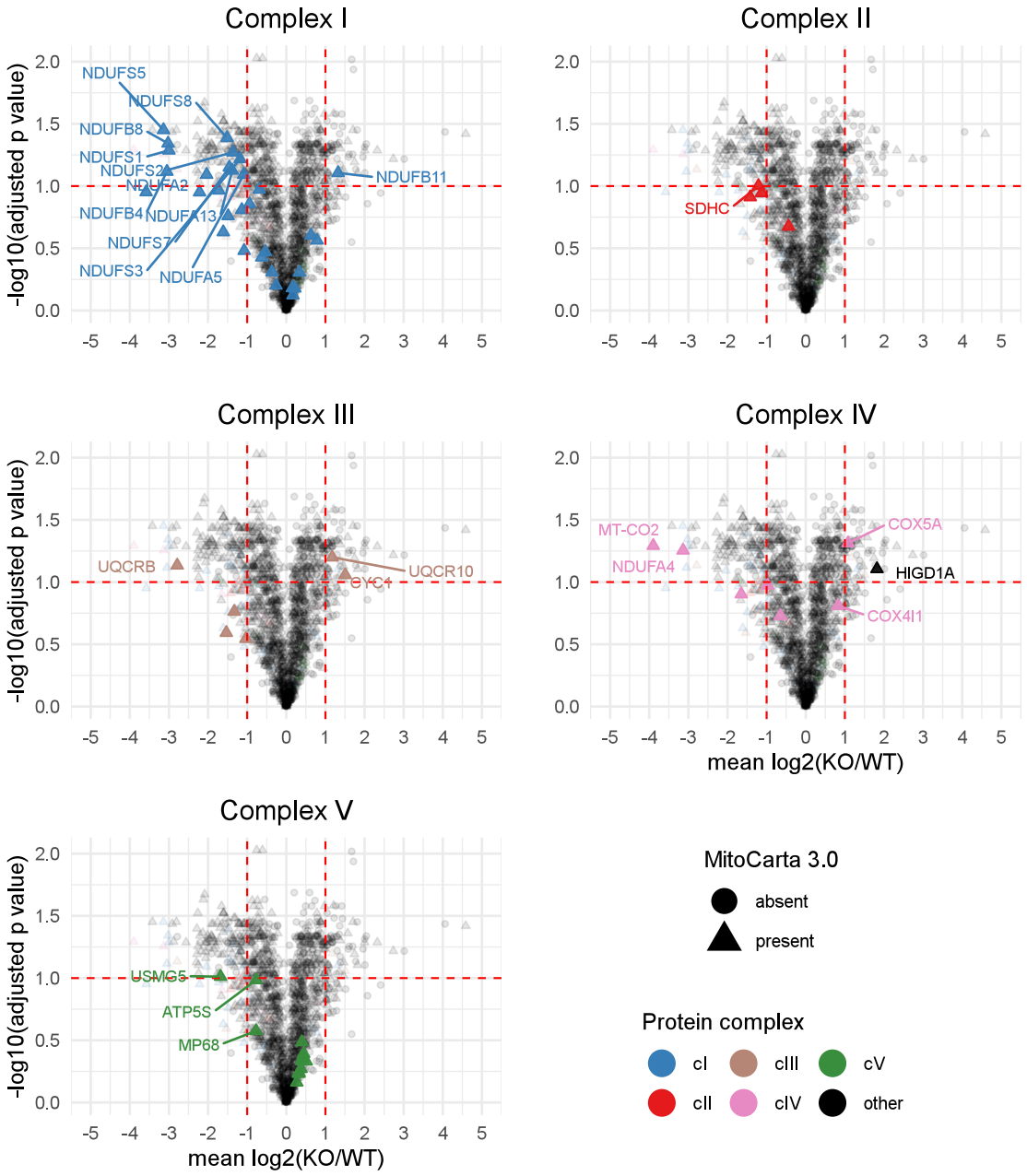


Figure 4.26: Changes in mitochondrial proteome in MRM1/2 KO compared to HEK293 cell line *Mitochondrial lysate from a mixture of differentially SILAC labelled MRM1/2 KO (KO) and HEK293 (WT) cells was analysed by LC-MS/MS. Ratios of labelled peptides between heavy and light samples were identified using MaxQuant. Average log2(KO/WT) from three biological repeats is shown in the plot. P-values were calculated using one sample t-test and FDR adjusted. Red horizontal line - 10% FDR, red vertical lines - 2-fold change. Only proteins that were quantified in all three biological replicates are shown. Each of the plots highlights proteins from selected complex of mitochondrial oxidative phosphorylation. Labelled proteins were up- or down- regulated at least two times and fall within 10% FDR (Complex I, Complex II and Complex III). For Complex IV, COX4I and HIGD41A are labelled in addition, these two proteins form a complex together with COX5A at the very early stage of Complex IV assembly. For Complex V all downregulated proteins are labelled.*

Changes in the proteome of MRM1/2 KO could be attributed to the disruption of MRM1, MRM2 or the combination of these two genes. To evaluate the effect of disruption of each methyltransferase on the mitochondrial proteome of the double knockout, we performed a SILAC proteomics analysis of MRM1 KO and MRM2 KO cell lines. Each cell

line was compared to the HEK293 as a reference in reciprocal labelling experiments. Steady state levels of detected proteins were largely unchanged in MRM1 KO (Figure 4.27). In MRM2 KO cell line, many proteins showed a major decrease compared to the wild type cells, with many of these proteins being part of the OXPHOS complexes (Figure 4.27). Next, using these data we compared changes in amounts of individual components of OXPHOS complexes among MRM1 KO, MRM2 KO and MRM1/2 KO cell lines (Figure 4.28). We did not observe major changes in any of these proteins in MRM1 KO. Many components of complexes I, III and IV were downregulated in both MRM2 and MRM1/2 KO. Nine out of thirteen mitochondrially encoded proteins were detected in at least one of the studied cell lines. Mt-CYB was identified and quantified only in MRM1 KO. The amount of mtDNA encoded components of Complex I, mt-ND1, mt-ND2, mt-ND4 and mt-ND5 were unchanged in MRM1 KO, decreased in MRM2 KO and not detected in MRM1/2 KO. Multiple factors affect whether a protein is detected by mass spectrometry. Apart from protein abundance within the cell, this can be affected by factors such as protein fragmentation, peptide ionisability, sample complexity or specific settings of the mass spectrometry analysis. Therefore, non-detection of a protein in mass spectrometry experiment does not prove absence of this protein in the sample. This is exemplified by the absence of mt-ND3, mt-ND4L, mt-ND6 and mt-CO3 in all three knockout samples (Figure 4.28). Considering that all detected OXPHOS proteins in MRM1 KO had very similar steady state levels to the wild type cells, it is most likely that these proteins were not detected due to the technical limitations, rather than their absence from the sample. Nevertheless, the fact that a lower number of mitochondrially encoded proteins were detected in MRM1/2 KO compared to MRM2 KO (three and eight, respectively) and that proteins that were detected in both cell lines were more downregulated in MRM1/2 KO support the conclusion that translation is more affected in MRM1/2 KO compared with the MRM2 KO. As discussed previously several nuclear encoded OXPHOS components were upregulated in MRM1/2 KO (Figure 4.26). We did not observe similar upregulation in neither MRM1 KO nor MRM2 KO (Figure 4.28), suggesting that this potential compensatory mechanism is activated only when mitochondrial translation drops below a certain threshold. Severe mitochondrial dysfunction in MRM1/2 KO is further highlighted by the downregulation of Complex II components. Complex II is usually unaffected by defects of mitochondrial translation as it does not contain any mtDNA encoded subunits.

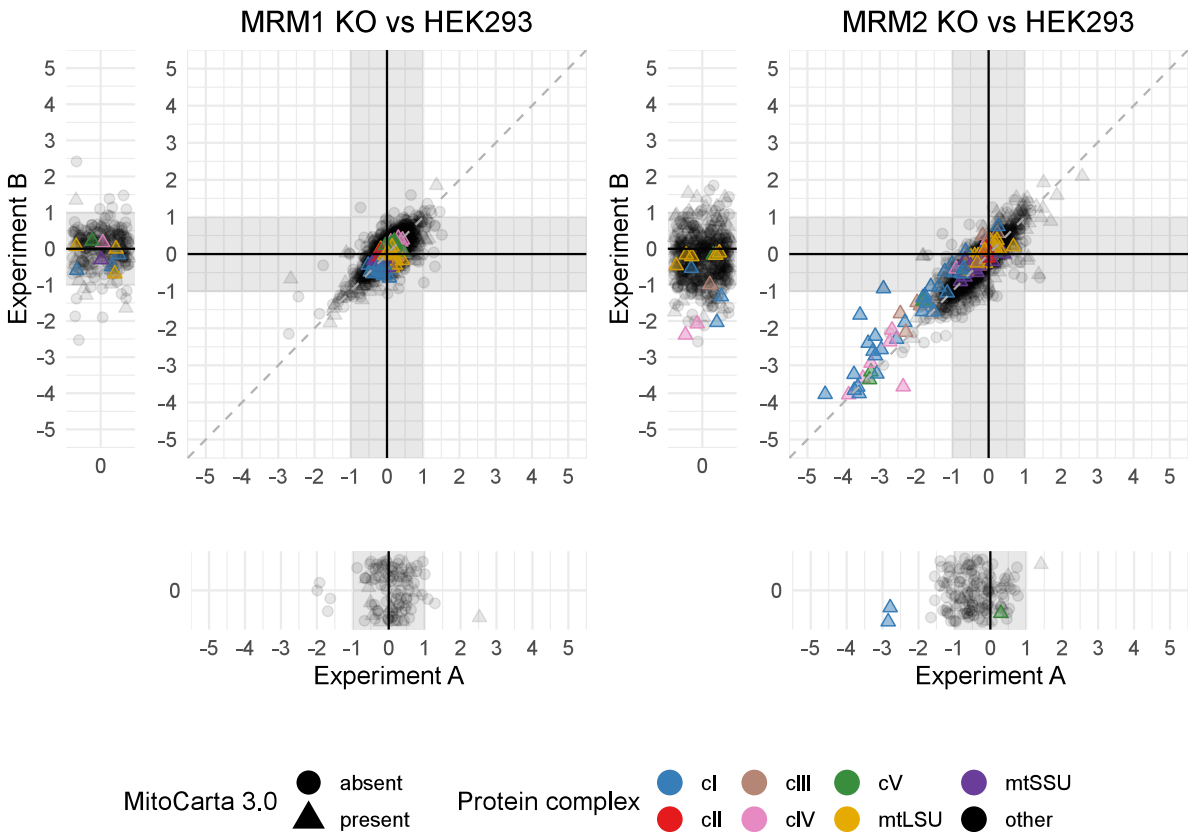


Figure 4.27: Quantitative analysis of mitochondrial proteome comparing MRM1 KO to HEK293 and MRM2 KO to HEK293
Mitochondrial lysates from a mixture of differentially SILAC labelled MRM1 KO and wild type or MRM2 KO and wild type cells were analysed by LC-MS/MS. Ratios of labelled peptides between heavy and light samples were identified using MaxQuant. Proteins that were identified in both reciprocal labelling experiments are shown in the middle plot and proteins that were identified in only one of the two replicates are shown in the side plots. Points are coloured based on the protein complex to which they belong and presence or absence of the protein in MitoCarta 3.0 is indicated by shape. Grey dashed line is defined by x=y.

To further investigate MRM1/2 KO specific phenotype, we performed a SILAC proteomics experiment directly comparing MRM1/2 KO to MRM2 KO cells (Figure 4.29). Several mitochondrial proteins were upregulated specifically in the MRM1/2 KO, with the most notable examples being FTL1, FTH and HMOX1 proteins that participate in the iron homeostasis. Unexpectedly, we observed many of mtSSU components being downregulated in MRM1/2 KO (Figure 4.29, purple triangles).

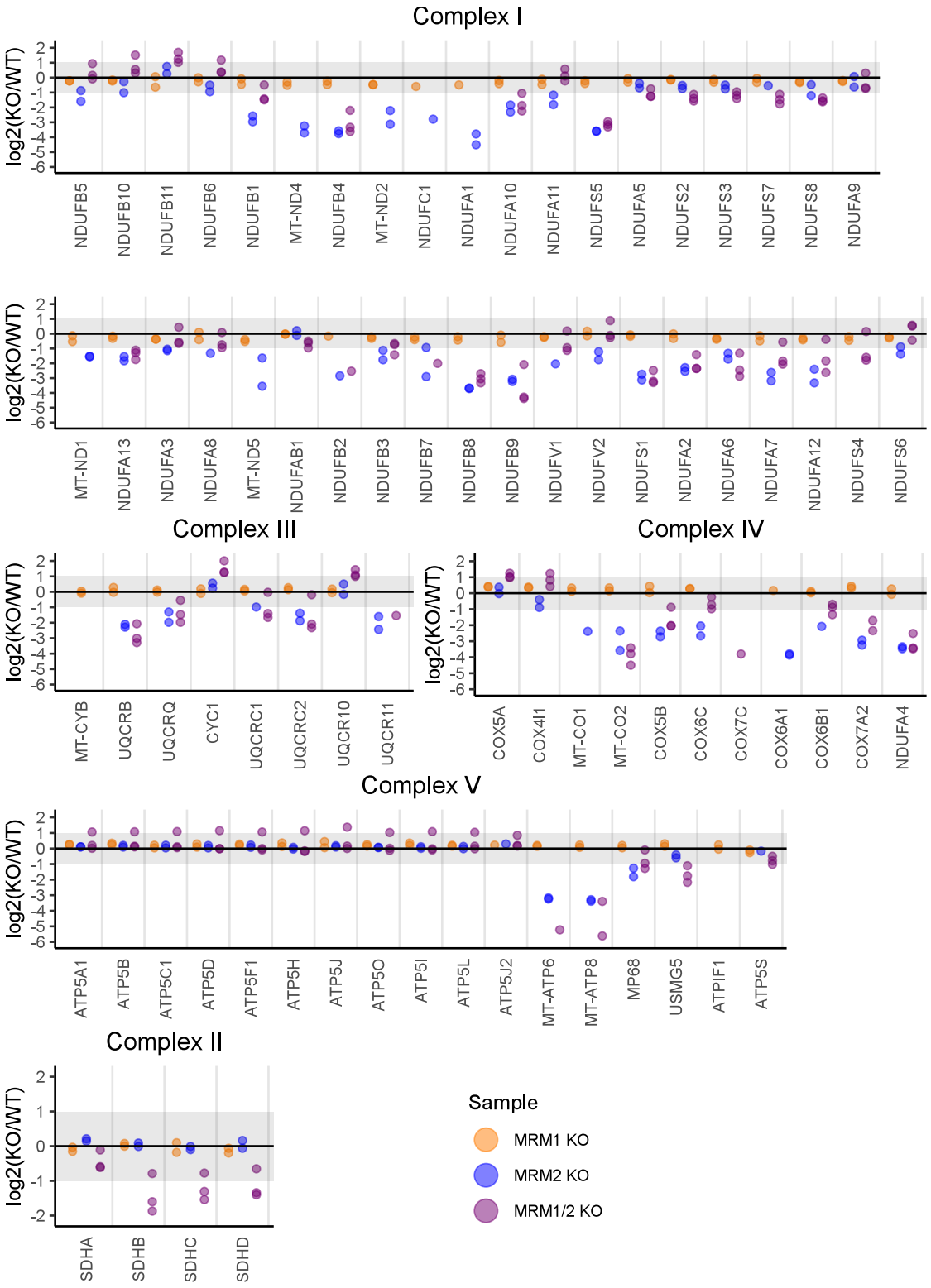


Figure 4.28: Quantitative comparison of steady state levels of OXPHOS proteins in the absence of mitochondrial methyltransferases

Mitochondrial lysate from a mixture of differentially SILAC labelled KO and WT cells was analysed by LC-MS/MS. Ratios of labelled peptides between heavy and light samples were quantified using MaxQuant. Either two (for MRM1 KO and MRM2 KO) or three (for MRM1/2 KO) biological replicates with reciprocal labelling were performed. Steady-state levels of OXPHOS proteins are shown. Only proteins that were detected at least in one replicate in at least one cell line are included.

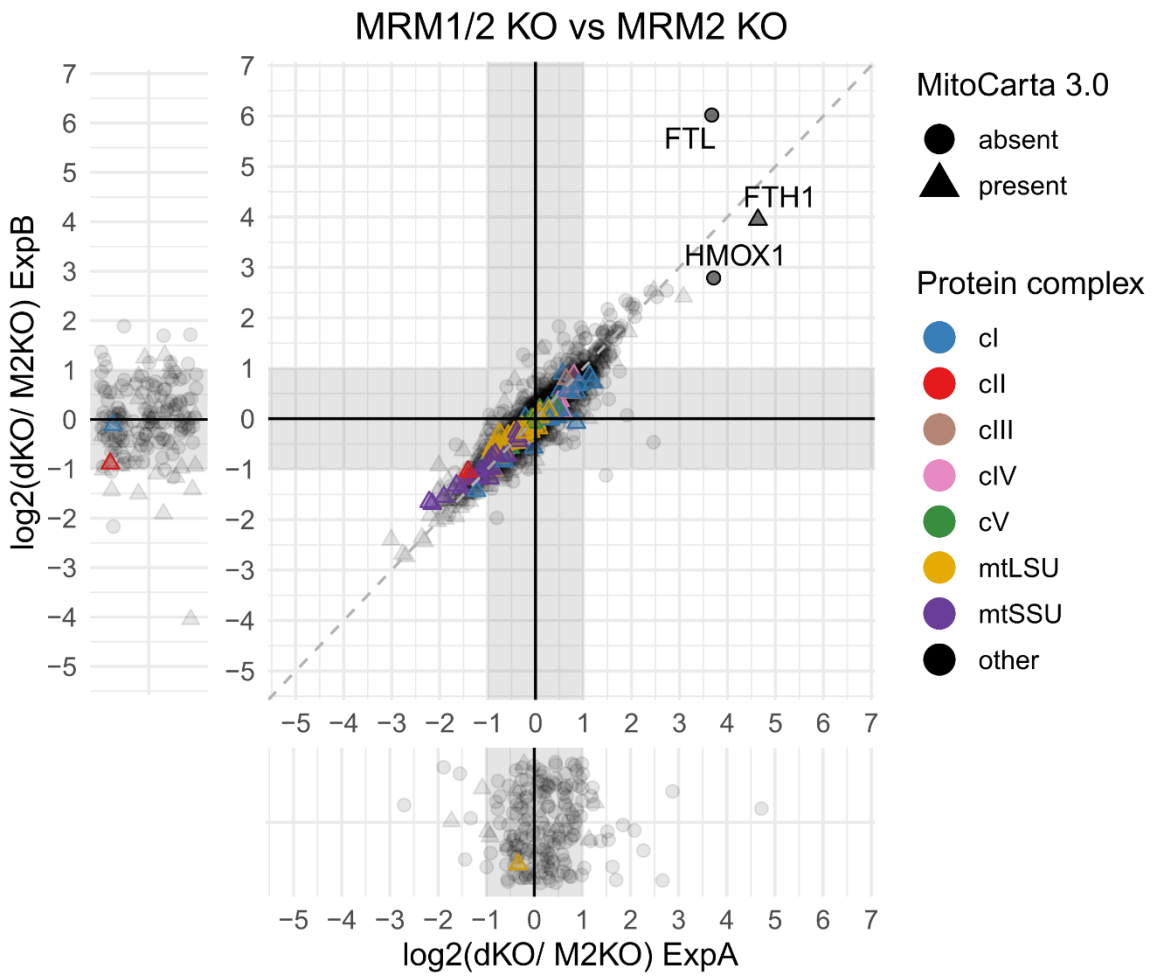


Figure 4.29: Quantitative comparison of mitochondrial proteome between MRM1/2 KO and MRM2 KO cell lines

Mitochondrial lysate from a mixture of differentially SILAC labelled MRM1/2 KO (dKO) and MRM2 KO (M2KO) cells was analysed by LC-MS/MS. Ratios of labelled peptides between heavy and light samples were quantified using MaxQuant. Proteins that were identified in both reciprocal labelling experiments are shown in the middle plot and proteins that were identified in only one of the two replicates are shown in the side plots. Points are coloured based on their protein complex and presence or absence of the protein in MitoCarta 3.0 is indicated by shape. Grey dashed line is defined by x=y.

4.2.6.5 Mitochondrial ribosome in MRM1/2 double knockout cell line

MRM1 and MRM2 proteins are methyltransferases that act on mitochondrial 16S rRNA. As 16S rRNA is an important structural component of the mitoribosome, we investigated in more detail the behaviour of mitoribosomal proteins by SILAC proteomics. Analysis of steady state levels of mitoribosome components in MRM1/2 KO showed decreased amounts of majority of the mtLSU proteins compared to the HEK293 cells, however these changes were not statistically significant (Figure 4.30, right). Analysis of mtLSU amounts in whole mitochondrial lysates of MRM1 KO and MRM2 KO also did not show major changes (Figure 4.31). From the three studied knockout cell lines mtLSU was perturbed to the largest extent in the MRM1/2 KO, however the changes in protein amounts were variable between experiments (Figure 4.31). Unexpectedly, many components of the mtSSU were significantly downregulated with some proteins (uS14m, bS21m, mS33) showing almost five-fold decrease in the double knockout (Figure 4.30, left). This change in the steady state levels of mtSSU proteins was specific for the MRM1/2 KO (Figure 4.31).

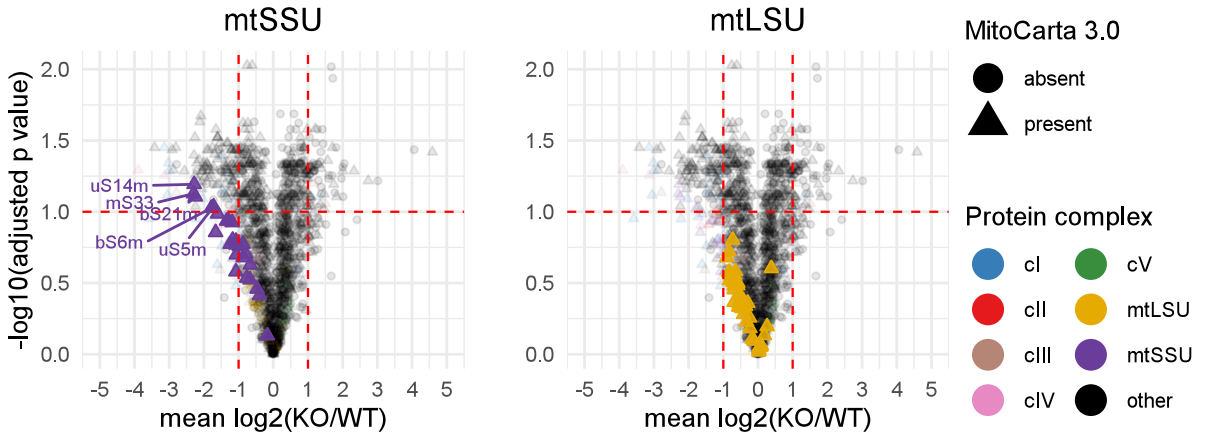


Figure 4.30: Quantitative comparison of components of mitochondrial ribosome between MRM1/2 KO cell line and HEK293 cell line

Mitochondrial lysate from a mixture of differentially SILAC labelled MRM1/2 KO (KO) and HEK293 (WT) cells was analysed by LC-MS/MS. Ratios of labelled peptides between heavy and light samples were quantified using MaxQuant. Average log2(KO/WT) from three biological repeats is shown in the plot. P-value was calculated using one sample t-test and FDR adjusted. Red horizontal line - 10% FDR, red vertical lines - 2-fold change. Only proteins that were quantified in all three biological replicates are shown. Protein components of mtSSU and mtLSU, respectively, are highlighted in the plots. Names are shown for proteins that were downregulated at least two times and fall within 10% FDR.

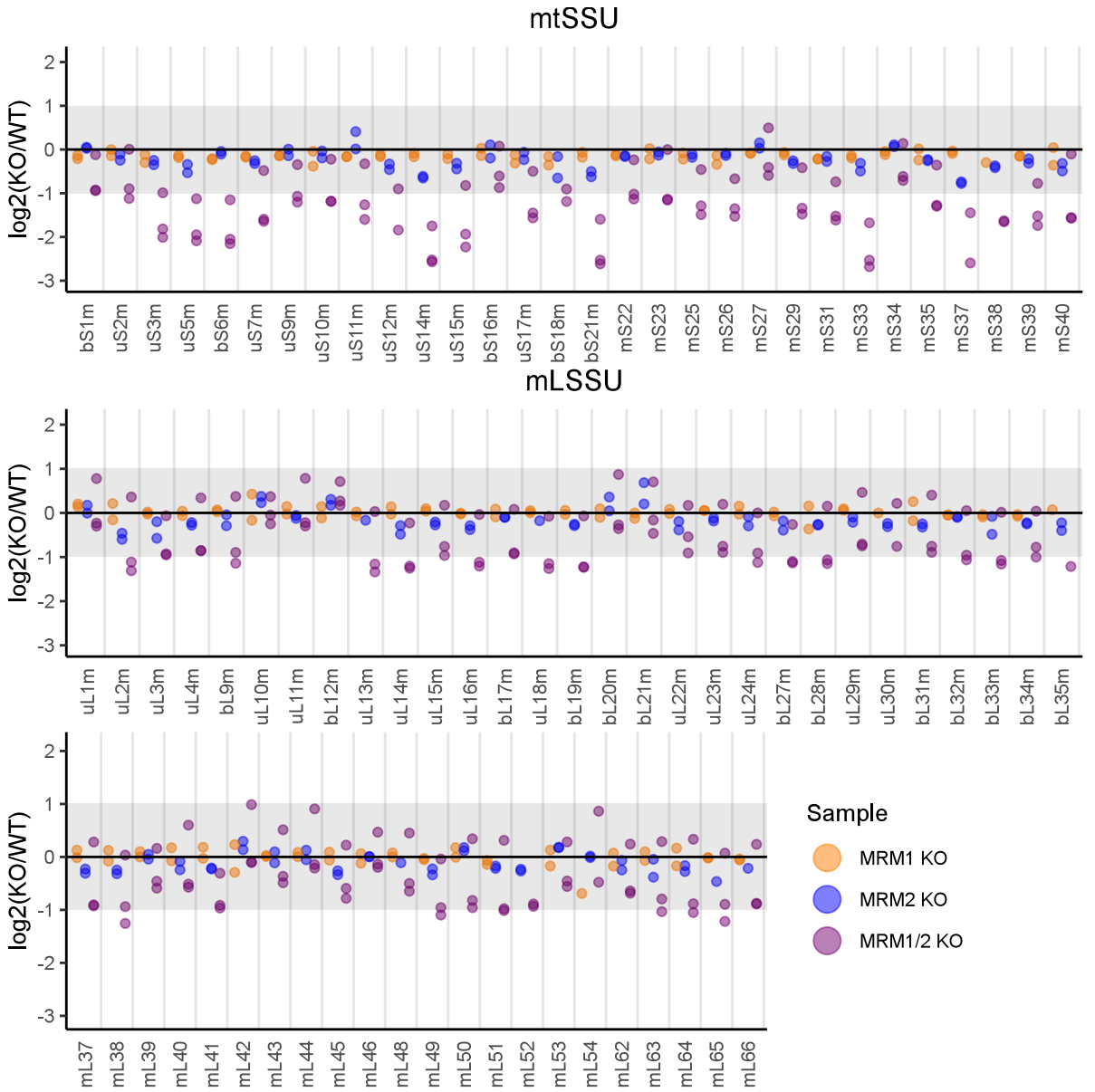


Figure 4.31: Changes in steady state levels of mitochondrial proteins in MRM1 KO, MRM2 KO and MRM1/2 KO cell lines

Quantitative comparison of steady state levels of proteins from mitochondrial small and large subunits between HEK293 (WT) cell line and cell lines lacking mitochondrial methyltransferases (KO). Mitochondrial lysate from a mixture of differentially SILAC labelled KO and WT cells was analysed by LC-MS/MS. Ratios of labelled peptides between heavy and light samples were identified using MaxQuant. Either two (for MRM1 KO and MRM2 KO) or three (for MRM1/2 KO) biological replicates with reciprocal labelling were performed. Steady-state levels of OXPHOS proteins are shown. bL36m is omitted from the figure as this protein was not detected in any of the experiments.

To further investigate the effect of simultaneous inactivation of MRM1 and MRM2, we performed a quantitative density gradient analysis by mass spectrometry (qDGMS) in MRM1/2 KO cells. In this method, protein complexes are separated on a sucrose gradient in near native state, which allows for the study of complex assembly intermediates and identification of potential protein interactions and complex assembly factors. Samples are SILAC labelled, allowing for quantitative comparison of protein complex components between the heavy and light labelled samples. Raw mass spectrometry data were analysed

in Proteome Discoverer, followed by a qDGMS-specific pipeline that has been developed in house and has been implemented as an R package called ComPrAn. Detailed description of qDGMS and ComPrAn has been published recently and development of the R package is discussed in Chapter 3 (Páleníková et al., 2021a). Data were analysed in two ways; one approach was aimed to compare protein migration qualitatively and second to compare heavy and light samples in a quantitative manner. In the qualitative approach, heavy and light samples are treated separately, and the aim is to identify co-migration of proteins within the label state. The aim of the quantitative approach is to directly compare the amount of individual proteins between two label states (for details see Chapter 3 sections 3.2.3 and 3.2.4).

First, we quantitatively compared components of mtSSU and mtLSU between MRM1/2 KO and HEK293. This analysis further confirmed a substantial downregulation of mtSSU in the double knockout (Figure 4.32). On the other hand, we observed an accumulation of mtLSU in the MRM1/2 KO (Figure 4.33). When MRM1/2 KO was directly compared to MRM2 KO by qDGMS, both mtSSU and mtLSU were downregulated in the double knockout compared to the single MRM2 KO (Figure 4.34). This indicated that although mtLSU accumulates in MRM1/2 KO compared to the wild type cells, fewer mtLSU particles are formed in the MRM1/2 KO than in the MRM2 KO.

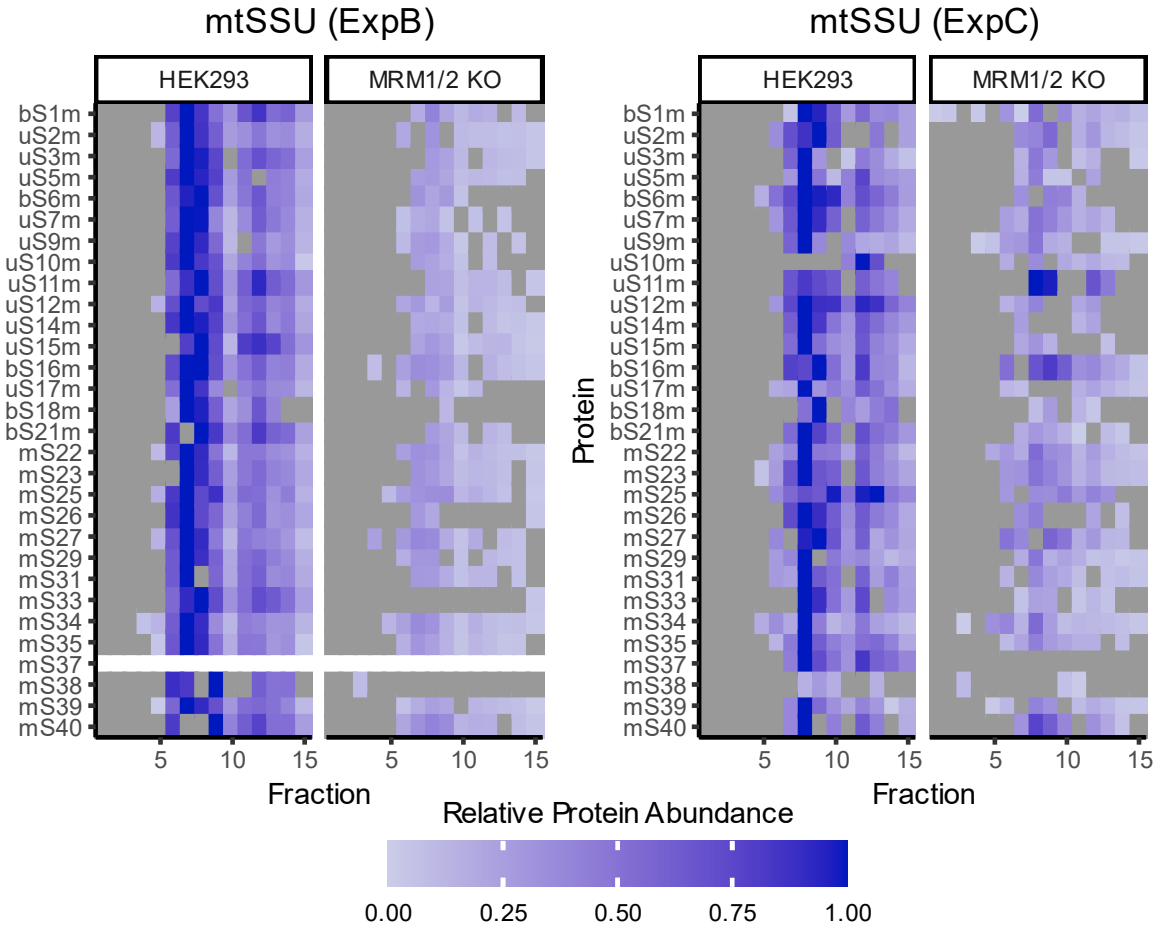


Figure 4.32: mtSSU in MRM1/2 KO compared with HEK293 wild type
Two replicates of qDGMS experiment (ExpB = Experiment B, ExpC = Experiment C) were performed with reciprocal labelling. ExpB and ExpC samples correspond to the ExpB and ExpC SILAC proteomics samples comparing MRM1/2 KO to wild type HEK293 (section 4.2.6.4). Reciprocal labelling of MRM1/2 KO and wild type HEK293 cells was performed, with 5-10% of mixed mitochondrial lysates used for direct mass spectrometry analysis of mitochondrial proteome and remaining 90-95% used for qDGMS. Data were analysed and visualised with the use of ComPrAn R package. Results of quantitative data analysis approach are shown, where proteins were normalised to the highest abundance value across both labelling states, allowing to compare up or down regulation of proteins between cell lines. Grey squares indicate fractions in which protein was not detected. For proteins that were not detected in any fraction a white line is shown.

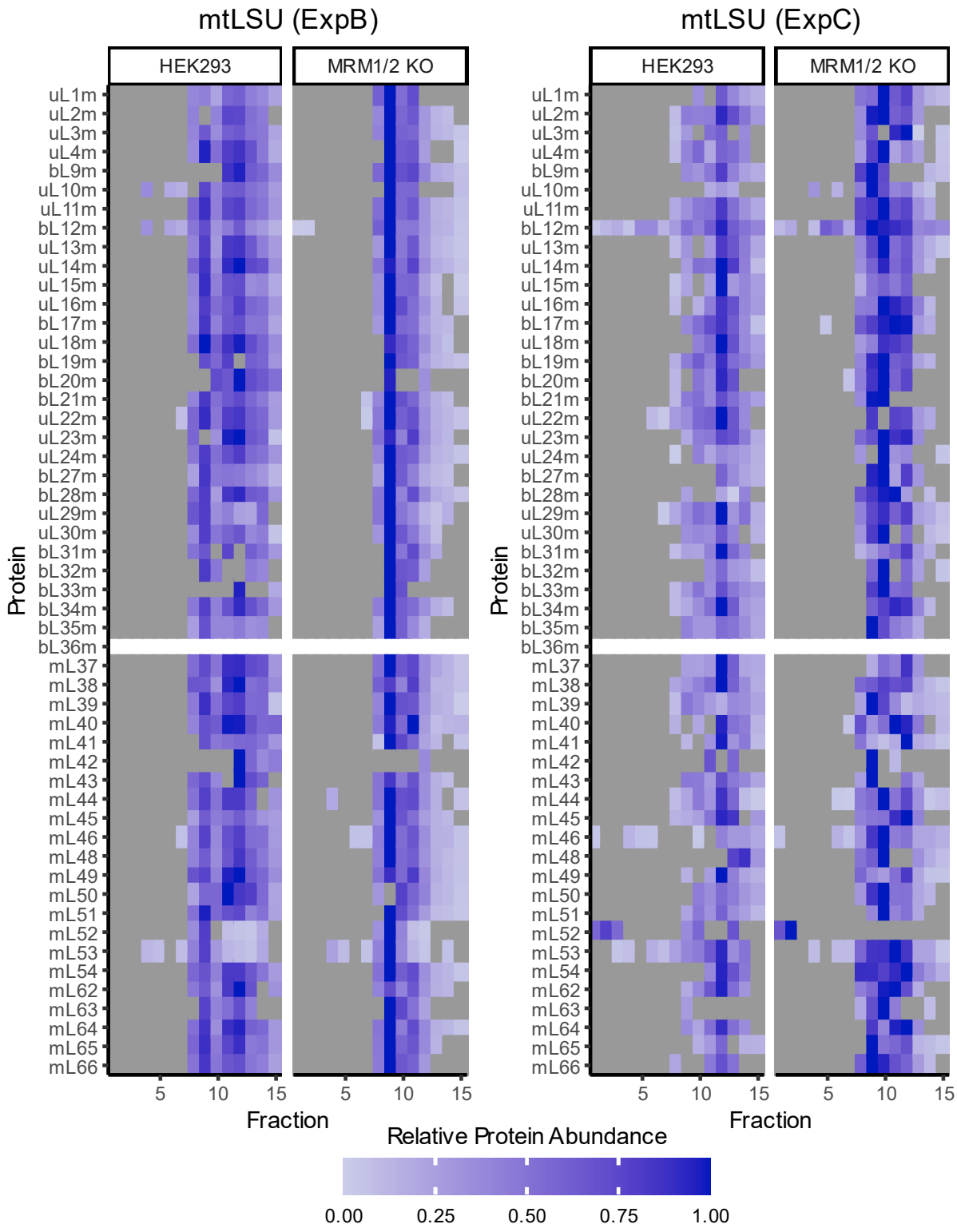


Figure 4.33: mtLSU in MRM1/2KO compared with HEK293 wild type
Data were processed and visualised in the same way as Figure 4.32, but for large mitoribosomal subunit.

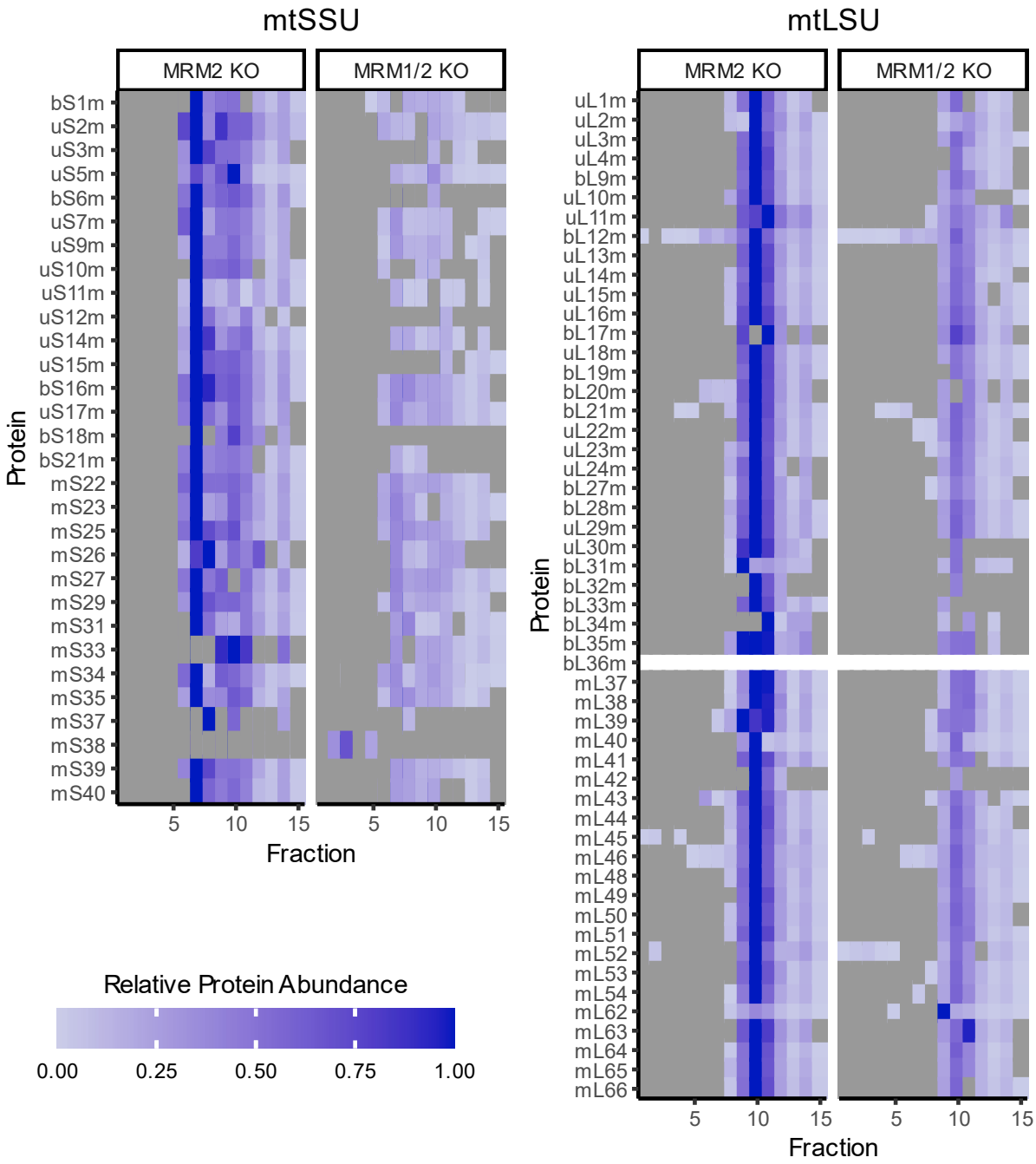


Figure 4.34: Quantitative comparison of mitoribosome between MRM1/2 KO and MRM2 KO by qDGMS

Two replicates of qDGMS experiment comparing MRM1/2 KO to MRM2 KO were performed with reciprocal labelling and a representative experiment with MRM1/2 KO heavy and MRM2 KO light is shown. Data were processed and visualised in the same way as Figure 4.32.

Next, we used the qualitative data to assess whether there are differences in the assembly of the mitoribosome subunits or the monosome. We did not observe accumulation of any assembly intermediates in either wild type or MRM1/2 KO cell line (Figure 4.35). The majority of the mtSSU proteins had a maximum in fraction 7 for both cell lines and mtLSU proteins had a maximum in fraction 9. Based on the co-migration of mtSSU and

mtLSU proteins in fractions 11 and 12 we can assume that these fractions correspond to the monosome. In HEK293, the median of all large subunit proteins is of similar value in mtLSU and monosome fractions (Figure 4.35A). However, in MRM1/2 KO, the median of all mtLSU proteins is highest in mtLSU fraction (Figure 4.35B), suggesting an accumulation of unassembled mtLSU in the double knockout and ineffective monosome formation. This is consistent with the defect of the monosome assembly in MRM2 KO (Rebelo-Guiomar et al., 2021).

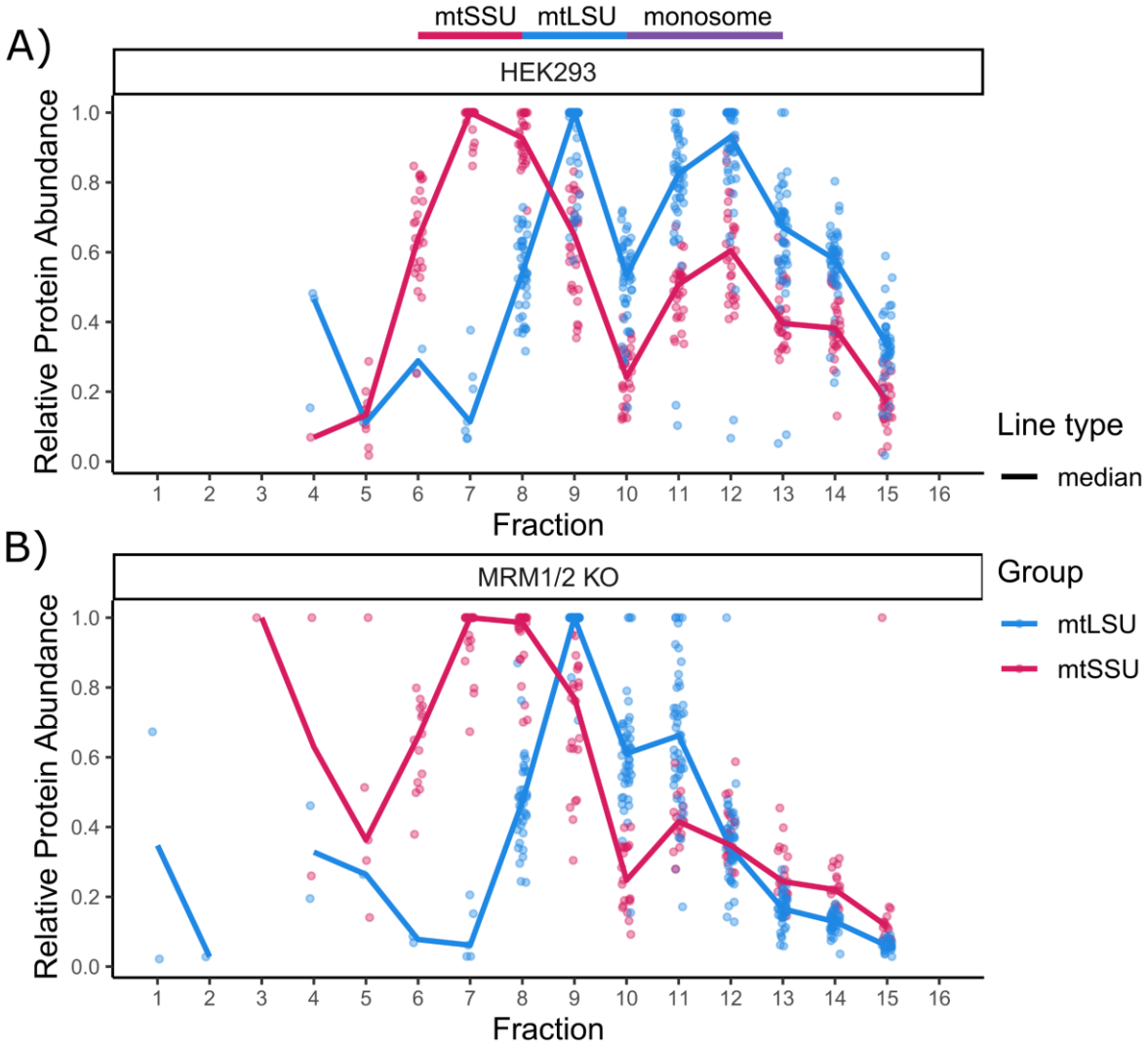


Figure 4.35: Migration profile of mtSSU and mtLSU proteins in MRM1/2KO and HEK293 cell lines
Two replicates of qDGMS experiment were performed with reciprocal labelling and results of Experiment B are shown as a representative example. Data were analysed and visualised with the use of ComPrAn R package. Results of the qualitative data analysis approach are shown, where proteins were normalised to the highest abundance value separately for HEK293 (A) and MRM1/2 KO (B) samples. Relative protein abundance values for each detected protein component of mtSSU and mtLSU are shown as circles. Median of all proteins for each subunit is also shown. Fractions corresponding to migration of mtSSU, mtLSU and monosome are indicated.

We were interested in further investigating the observed downregulation of the mtSSU in the double knockout. This disruption of mtSSU could have been caused by either a defect in the subunit assembly or by faster degradation of mtSSU that is not part of the monosome. Mitoribosome assembly is a complicated process and the methods available for its study are limited. However, a pulse SILAC study that investigated kinetics of mitoribosome assembly has been published recently (Bogenhagen et al., 2018). In this study, wild type HeLa cells were labelled with heavy amino acids for a short period of time, followed by an isolation of mitoribosomes and mass spectrometry analysis of ratios between newly synthesised heavy peptides and pre-existing light peptides in the mature subunits. This led to an establishment of a model of mitoribosome assembly kinetics and categorisation of proteins of mitoribosome small and large subunits into early, intermediate and late assembly groups (Bogenhagen et al., 2018). Applying this experimental method in wild type HEK293 and MRM1/2 KO to directly compare assembly kinetics would be technically challenging due to the large number of cells needed for SILAC based mass spectrometry. Furthermore, assembly kinetics is modelled with the use of doubling time of the cell line. The growth rate is very different between MRM1/2 KO and HEK293 (see Figure 4.18), which would make it difficult to distinguish between the effect of substantially longer doubling time of the double knockout and slower assembly kinetics. However, if we assume that the steady state levels of mitoribosome components in MRM1/2 KO are composed of a mixture of partially and fully assembled mitoribosomes, then comparing the ratios of protein levels in MRM1/2 KO and HEK293 could inform on the accumulation of certain intermediates. If the assembly kinetics were the same between wild type and the double knockout, all proteins should be either unaffected or up- or down-regulated by a similar amount. If the assembly of the subunit was perturbed in the double knockout, we would expect to see a broader range of up- or down- regulation of mitoribosome components. Perturbation of the mtLSU components in MRM1/2 KO was in a range from 1.5-fold upregulated proteins to about 2-fold downregulated proteins. In case of mtSSU, the range of downregulation was larger, from almost unaffected proteins to about 5-fold downregulation (Figure 4.36). This wide range of protein downregulation especially in the mtSSU might suggest a problem in the assembly of this subunit.

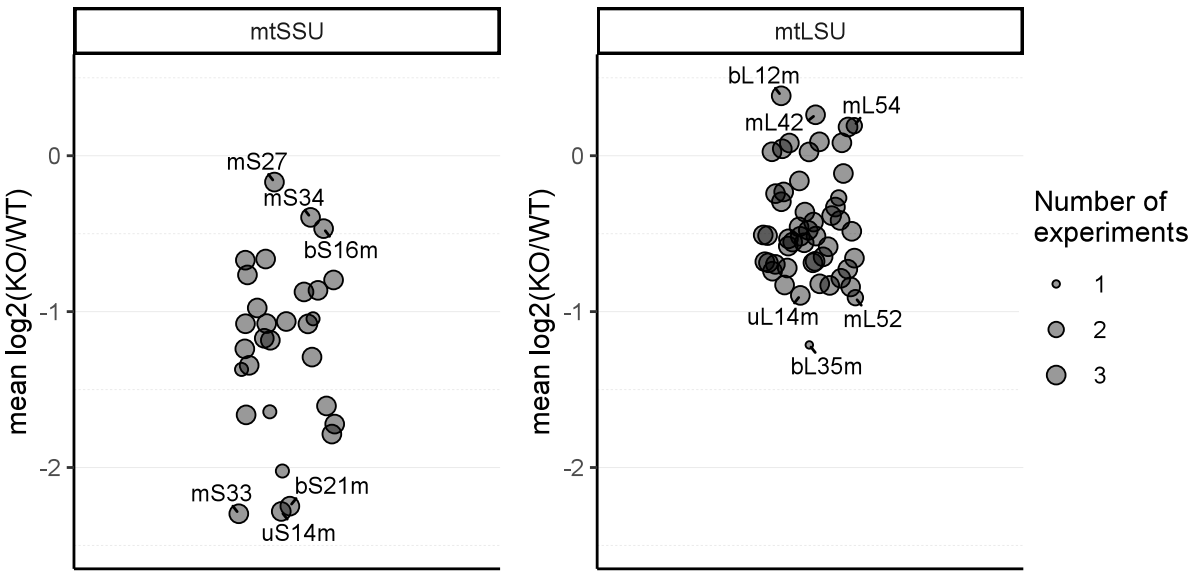


Figure 4.36: Perturbation of mitoribosomal proteins in mitochondrial lysate
Average log2 ratio of protein amount in mitochondrial lysate of MRM1/2 KO (KO) compared with HEK293 (WT). The experiment was performed in 3 biological replicates, average value of up to three replicates is shown for each protein, with the size of the point corresponding to the number of experiments in which the protein was identified and quantified. Proteins with three highest and lowest values of average log2 ratio are labelled.

We next compared our measured SILAC ratios of MRM1/2 KO and HEK293 to the SILAC ratios reported by Bogenhagen and colleagues. Using the reasoning as described before, in case any assembly intermediates would be accumulating in the MRM1/2 KO, we should observe a correlation with the partially labelled mitoribosomes. This was indeed the case for both mtSSU and mtLSU proteins (Figure 4.37). Early assembly proteins were least downregulated, whereas most downregulated proteins belonged to the late assembly category. These findings point towards defects in the assembly of mitoribosomal subunits in MRM1/2 KO. As the substrate for both MRM1 and MRM2 is large subunit 16S rRNA, it is likely that the proper assembly of mtSSU relies on feedback from the assembling mtLSU.

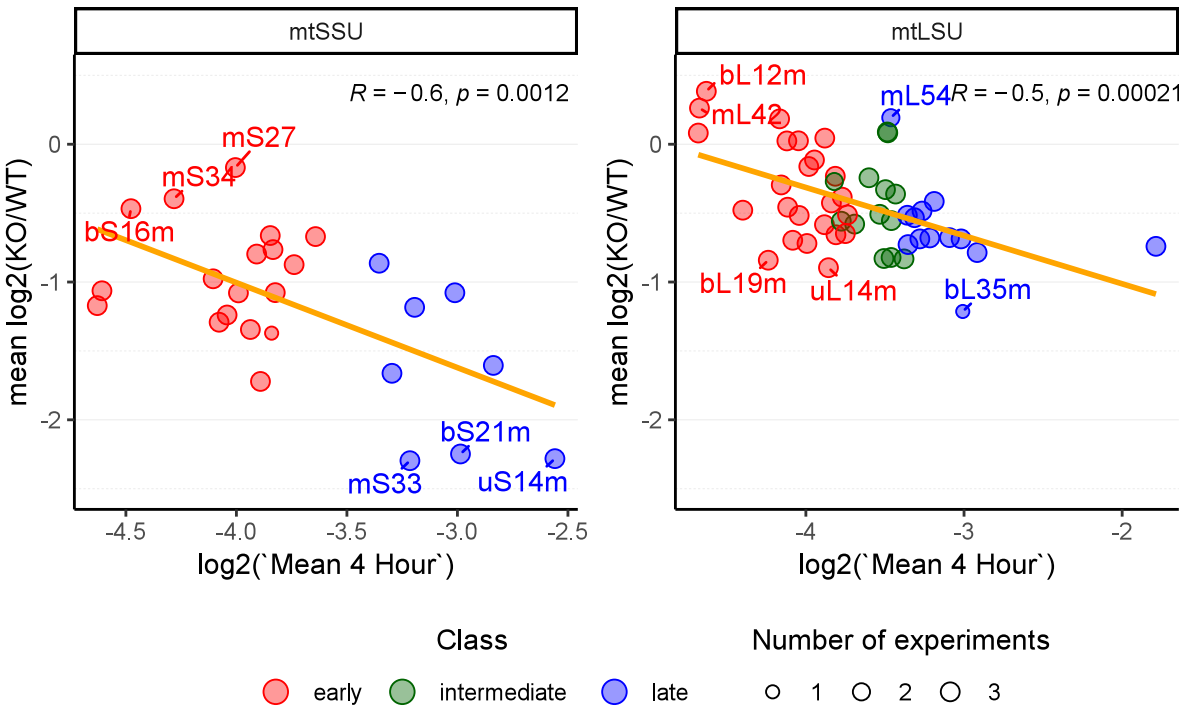


Figure 4.37: Evaluation of relationship between protein downregulation and mtSSU and mtLSU assembly sequence

Comparison of protein ratios between MRM1/2 KO (KO) and HEK293 (WT) to SILAC ratios from partial mitoribosome labelling experiments from (Bogenhagen et al., 2018). Average ratio of KO to WT protein abundance in a mitochondrial lysate from three biological replicates is shown on y-axis. Not all proteins were identified in all biological replicates, size of the dots indicates number of experiments in which the protein was quantified. On x-axis values of heavy to light ratios from 4-hour pulse SILAC labelling of mitoribosomes by Bogenhagen and colleagues is shown. Proteins are coloured based on the assembly class as assigned by Bogenhagen et al and proteins that were not assigned to a category in this study are not shown. Proteins with three highest and lowest values of average KO/WT log2 ratio are labelled. A linear model was fitted to the points (orange line) and Pearson correlation R number and p-value are shown.

4.3 Discussion

Targeted CRIPSR/Cas9 screen can identify genetic interactions

Almost 300 proteins have been estimated to be involved in the maintenance and expression of mitochondrial DNA. Translation is the last step of gene expression with the mitochondrial ribosome being a central player. Posttranscriptional RNA modifications play a crucial role in its assembly and function. MRM1 is responsible for depositing one of only five modifications on the human 16S rRNA; however its inactivation does not have an observable negative impact on the mitochondrial function (D'Souza, 2018). We hypothesised that identification of genetic interactions of this gene will provide insight into the importance of MRM1 for mitochondrial function. We created a custom gRNA library targeting genes with known or predicted functions in mitochondrial gene expression as of the time this study was performed. Using a CRISPR/Cas9 screen employing this custom library, we were able to identify candidate genetic interactions of MRM1 (Table 4.1).

With our screening approach we were able to identify candidate genetic interactions between MRM1 and genes involved at various stages of mitochondrial gene expression and maintenance. Including only 399 genes allowed for a reduction in the number of cells, scale of sample processing and sequencing depth compared to a genome-wide approach. However, this also limited findings since we might have missed genetic interactions with genes that were not yet identified as mitochondrial. Our knowledge of the mitochondrial proteome keeps improving, as exemplified by the revised version of MitoCarta3.0 (Rath et al., 2021), in which 78 genes were added and 100 removed compared to the previous version (Calvo et al., 2016). A major drawback of genome-wide library is its high demand on resources. Improvements have been made recently by optimising gRNA libraries so each gene can be targeted by fewer gRNAs (Gonçalves et al., 2021), making large scale CRIPSR/Cas9 screens more accessible. Despite the mentioned limitations, we were able to identify several genes that are likely to be true genetic interactions of MRM1.

Candidate genetic interactions identified by CRISPR/Cas9 screen function in various parts of mitochondrial gene expression, including replication (DNA2, SSBP1, TOP2A), tRNA aminoacylation (SARS2, VARS2, TARS2), posttranscriptional RNA modification (MRM2, TRMT5, NFS1, TRMT112) as well as structural components of the mitoribosome. Spatial proximity and dependence of components of mitochondrial nucleoids,

mitochondrial RNA granules and assembling ribosome have been shown previously (Antonicka and Shoubridge, 2015; Hensen et al., 2019; Jourdain et al., 2016), as well as colocalisation of MRM1 and nucleoids (Lee et al., 2013). Identification of genes involved in mtDNA replication as candidate genetic interactions of MRM1 may further support close relationship of these processes of gene expression in mitochondria. The VARS2 gene was amongst the identified aminoacyl tRNA synthetases. Mammalian mitochondrial ribosome does not contain 5S RNA, but a structural tRNA instead, which in human is mt-tRNA^{Val} (Brown et al., 2014; Rorbach et al., 2016). At present it is not clear which posttranscriptional modifications are present on this structural tRNA. If aminoacylation of this tRNA is necessary for its efficient incorporation into the mitoribosome, this could explain genetic interaction between MRM1 and VARS2. Several mitoribosomal proteins were also identified as potential genetic interactors of MRM1. When mapped onto the structure of the human mitoribosome, these proteins do not seem to cluster (Figure 4.38) and they include proteins that both interact directly with 12S rRNA (e.g. uS5m, uS12m, mS40) and 16S rRNA (e.g. uL2m, uL15m, mL66) as well as proteins without direct RNA interactions (e.g. mS39, mL39). Whether there is specific genetic interaction between MRM1 and these proteins or whether this is a general result of mild disruption of mitochondrial ribosome structure due to the absence of Gm1145 requires further investigation. Among all genetic interactors, MRM2 gene was identified in all of the studied conditions. This methyltransferase is responsible for 2'-*O* methylation of U1369 in the A-site of the mtLSU (Lee and Bogenhagen, 2014; Lee et al., 2013; Rorbach et al., 2014). Since this protein appeared in all tested conditions, it was selected for further characterisation of genetic interaction with MRM1 by production of a double knockout cell line.

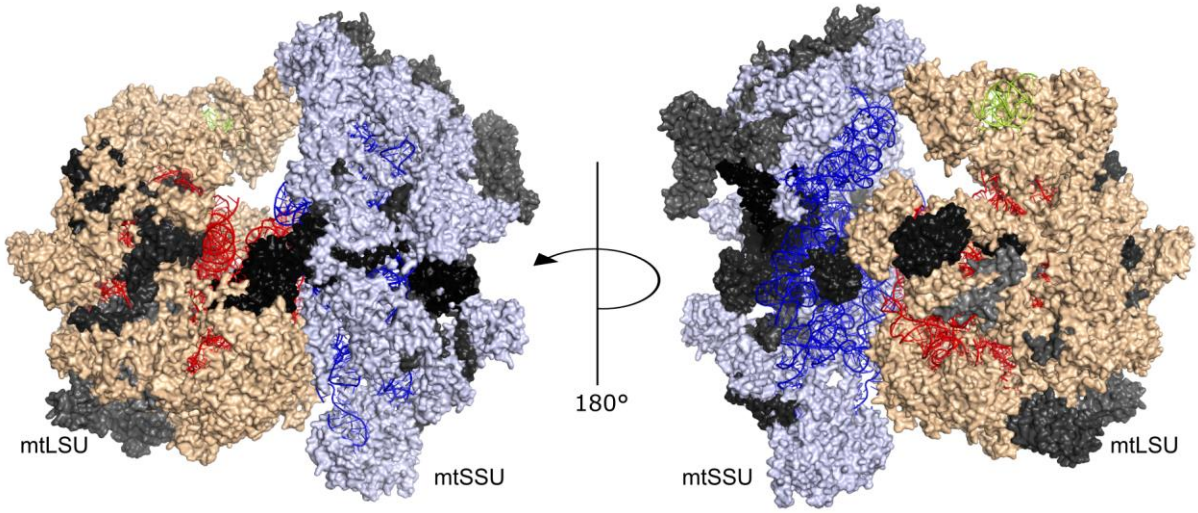


Figure 4.38: Mitoribosomal proteins predicted to have genetic interaction with MRM1
Proteins that were identified to have genetic interaction with MRM1 by CRISPR/Cas9 screen are highlighted on the structure of human mitoribosome (PDB: 3J9M). 12S rRNA, 16S rRNAs and tRNA^{Val} are shown as cartoon in blue, red and yellow, respectively. mtSSU and mtLSU proteins are shown as surface in light blue and light orange, respectively. Candidate genetic interactors are shown in shades of grey.

Simultaneous knockout of MRM1 and MRM2 leads to reduced mtSSU assembly

Apart from Gm1145, four other posttranscriptional modifications have been identified on the 16S rRNA: m_1A947, Um1369, Gm1370 and Ψ1397, deposited by TRMT61B, MRM2, MRM3 and RPUSD4, respectively (reviewed in Rebelo-Guiomar et al., 2018). Out of these, only MRM2 has been identified as a candidate for MRM1 genetic interaction, suggesting that this interaction is specific and not a general cooperation among epitranscriptomic marks of 16S rRNA.

Nucleotides modified by MRM1 and MRM2 are positioned in the P-loop and A-loop of the 16S rRNA, respectively. As opposed to the G1145 nucleotide, that is modified by MRM1, which interacts with the 3' CCA of the tRNA in the P-site (Figure 4.1), U1369 modified by MRM2 does not directly interact with the 3' CCA in the A-site. Phenotypic effects of inactivation of MRM1 and MRM2 in human cells are very different. While MRM1 KO cells show normal mitochondrial function (D'Souza, 2018), MRM2 KO cells exhibit severe mitochondrial deficiency (Rebelo-Guiomar et al., 2021). Mutation in MRM2 was also identified as a cause of MELAS-like syndrome in humans (Garone et al., 2017). Mitochondrial deficiency in MRM2 is due to the dysfunctional mitochondrial translation. However, when the integrity of the mitoribosome in MRM2 KO was analysed by sucrose density gradient and western blotting or qDGMS, mtLSU accumulated with only bL36m protein being downregulated in mtLSU fractions (Rebelo Guiomar, 2020). MRM2 was

identified in recent cryo-EM structural data showing late stages of mitoribosomal assembly (Cipullo et al., 2021a; Hillen et al., 2021; Lenarčič et al., 2021). In the absence of MRM2, mtLSU assembly is stalled with anti-association complex (MALSU1, L0R8F8 and mtACP), missing bL36m and unstructured RNA at the intersubunit interface (Rebelo-Guiomar et al., 2021).

To evaluate how simultaneous inactivation of MRM1 and MRM2 affects mitoribosome biogenesis, we performed qDGMS experiments in MRM1/2 KO and compared the results to wild type and MRM2 KO cells. We did not observe the formation of any mtLSU assembly intermediates on the sucrose gradient in the double knockout (Figure 4.33, Figure 4.35B). The amount of the mtLSU was higher in the MRM1/2 KO compared to the wild type (Figure 4.33). However, this accumulation was to the lower extent as compared to the MRM2 KO in qDGMS experiments (Figure 4.34). Surprisingly, although both MRM1 and MRM2 act on the 16S rRNA, which is a structural component of mtLSU, we observed a severe downregulation of mtSSU in the MRM1/2 KO cell line (Figure 4.30, Figure 4.32). Since SILAC proteomics and qDGMS inform only about steady-state levels of proteins, this downregulation could result either from reduced assembly or increased degradation of the mtSSU.

Knowledge about the assembly and degradation processes of the mammalian mitochondrial ribosome is still limited. To date, 9 and 22 protein factors have been implicated in the assembly of mtSSU and mtLSU of mammalian mitoribosomes, respectively (Hilander et al., 2021). A recent study of the kinetics of mitoribosome assembly by pulse-SILAC divided mitoribosomal proteins into early, intermediate, and late categories (Bogenhagen et al., 2018). We used this information to gain insight into the cause of downregulation of mtSSU in MRM1/2 KO cell line. Quantitative proteomic analysis of mitoribosomal proteins from MRM1/2 KO cells showed that the most and least downregulated proteins were from late (uS14m, bS21m, mS33) and early assembly (bS16m, mS27, mS34), respectively (Figure 4.36). Furthermore, SILAC ratios of mtSSU components of MRM1/2 KO to wild type cells were well correlated with the pulse-SILAC ratios of mtSSU assembly kinetics (Figure 4.37). This suggests that the assembly of mitoribosomal subunit is affected in the absence of both MRM1 and MRM2 proteins.

Another mitochondrial protein, NSUN4, has been shown to have a dual role, both as a 12S rRNA methyltransferase and in complex with MTERF4 as an assembly factor of mtLSU (Metodiev et al., 2014). Similarly, observed downregulation of mtSSU in MRM1/2 KO could be a result of dual role of MRM1 and/or MRM2 in both methylation of 16S rRNA and assembly of mtSSU.

Information about molecular targets of protein function can be gained by identifying its interaction partners. Immunoprecipitation experiments in HEK293 cell line overexpressing MRM1 did not find an interaction between MRM1 and mitoribosomal components, however a weak interaction with the mtSSU biogenesis factor, ERAL1, was identified (D'Souza, 2018). In the same work, MRM1 was found to migrate as a free protein in sucrose gradient as analysed by western blotting. We did not detect MRM1 protein in the mass spectrometry data set corresponding to wild type cells in qDGMS experiments presented here. However, we took advantage of a growing number of published qDGMS experiments (D'Souza et al., 2021; Rebelo-Guiomar et al., 2021; Van Haute et al., 2019), and inspected the presence of MRM1 in these data sets. MRM1 protein was present in some of these data sets (Table 4.2). Since assembly of the mitoribosome is perturbed in knockout cell lines, we focused on the migration of MRM1 in the wild type data sets. In data from Rebelo-Guiomar et al. (2021), only one MRM1 peptide has been detected in fraction 6, which corresponds to the region of the gradient where mtSSU migrated (data not shown). In Van Haute et al. (2019) data set, MRM1 was present in fractions adjacent to the mitoribosomal small subunit (Figure 4.39A, B). In data from D'Souza et al. (2021), MRM1 migrates in lower fractions compared to the mitoribosome (Figure 4.39C). Co-migration of proteins in sucrose gradient suggests protein interaction. The presence of MRM1 in the same fractions as mtSSU indicates MRM1 might function in the assembly of this mitoribosomal subunit. However, multiple protein complexes can migrate at similar regions of the gradient, therefore it should not be used as a sole indicator of assigning protein into protein complex. Similarly, the absence of protein co-migration on the sucrose gradient does not exclude the possibility that proteins are part of the same protein complex or are involved in an assembly of the protein complex. Such interactions might be weak and transient, leading to their disruption during sample processing. Approaches like immunoprecipitation combined with crosslinking or proximity-dependent biotin identification, BioID (Roux et al., 2012), could be applied to further evaluate weak or transient interactions of MRM1. With the currently available data, a dual role of MRM1 as both 16S rRNA methyltransferase and assembly factor of mtSSU cannot be excluded. It is unclear why this function would manifest only during simultaneous inactivation of MRM1 and MRM2 and not in single MRM1 knockout cells. A possible explanation could be feedback provided between the assembly of mtSSU and mtLSU (discussed later).

Table 4.2: MRM1 protein in published qDGMS experiments
The presence of MRM1 was evaluated in published qDGMS proteomics data. All studies concerned investigated assembly of mitoribosome in cell lines with knockout in a selected gene of interest. Gene of interest that was investigated in each study is listed together with cell line in which the knockout was produced. For each replicate of each study, it is indicated whether MRM1 was present (tick mark) or absent (cross) in data corresponding to the wild type and knockout.

Study	Replicate 1		Replicate 2	
	Wild type	KO	Wild type	KO
Van Haute et al. 2019, METTL15 KO (HAP1)	✓	✓	✓	✓
D'Souza et al. 2021, YbeY KO (HAP1)	✓	✓	✗	✗
Rebelo-Guiomar et al. 2021, MRM2 KO (HEK293)	✓	✗	✗	✓

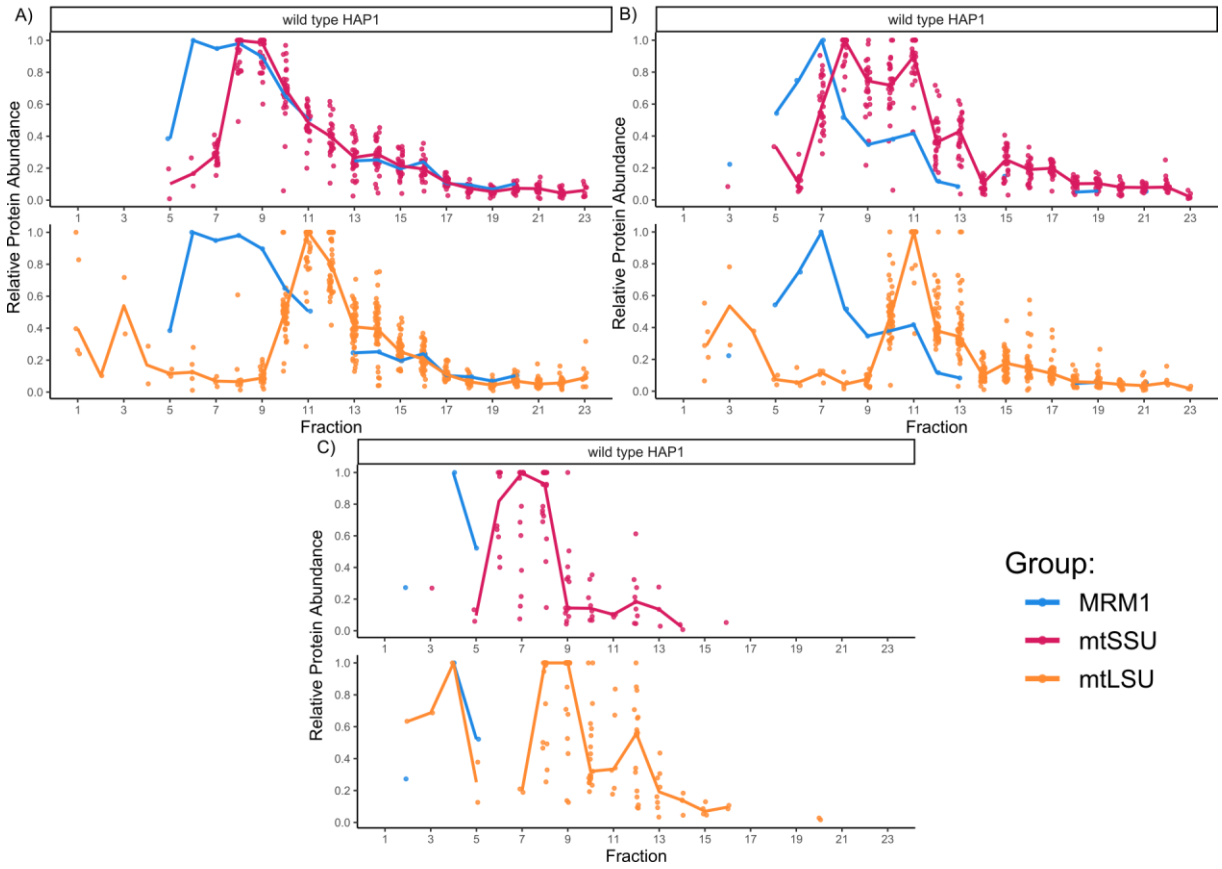


Figure 4.39: Migration of MRM1 protein in published qDGMS data
qDGMS data from Van Haute et al. (2019) and D'Souza et al. (2021) were reanalysed in ComPrAn and co-migration of MRM1 with mitoribosome was analysed. Qualitative data set (Scenario A, see section 3.2.3 for details) was used for this analysis. Results from two replicates with reciprocal labelling from Van Haute et al (A, B) and one replicate from D'Souza et al. (C) are shown with individual values (points) and median (line) for all detected components of mtSSU (red), mtLSU (orange) and MRM1 (blue).

In MRM2 KO cells, mtSSU seems unaffected in terms of both amount and assembly (Rebelo-Guiomar et al., 2021). No mtSSU proteins were reproducibly enriched in immunoprecipitation of MRM2 interactors (Rebelo Guiomar, 2020). Additionally, mtSSU proteins did not appear among the most enriched MRM2 proximal proteins identified by BioID (Rebelo Guiomar, 2020). MRM2 appears to migrate in the sucrose gradient in the regions of mtLSU and, to the lower extent, monosome region (Lee et al., 2013; Rorbach et al., 2014). Recent cryo-EM structures showed MRM2 as part of the late assembly of mtLSU (Cipullo et al., 2021a; Lenarčič et al., 2021). Taken together, it is unlikely that MRM2 would function as an assembly factor of the mtSSU.

Disruption of mtSSU assembly in MRM1/2 KO could also be a result of a connection between the assembly pathways of mitoribosomal subunits. It has been suggested previously that the assembly of the small subunit and large subunit might be linked (Dalla Rosa et al., 2014; D'Souza et al., 2021; Maiti et al., 2018). Small subunit assembly factor YbeY interacts with several proteins of the mitoribosomal large subunit (D'Souza et al., 2021; Summer et al., 2020). Mitoribosomal small subunit was impaired in cells where MPV17L2 protein, which interacts with mtLSU and monosome, was silenced (Dalla Rosa et al., 2014). This is thought to be due to prolonged interaction of mtSSU with nucleoids in situations when mtLSU assembly is perturbed (Dalla Rosa et al., 2014). Another mtLSU biogenesis factor, GTPBP10, also plays a role in the biogenesis of mtSSU by affecting processing of 12S-16S rRNA precursor (Maiti et al., 2018). Levels of mature 12S and 16S rRNA were reduced to about 70% of wild type in GTPBP10 KO, while 12S-16S rRNA precursor accumulated. Our analyses of steady-state levels of rRNAs in MRM1/2 KO by RT-qPCR showed a reduction in both 12S and 16S rRNA to about 50% of wild type level (Figure 4.16C). It is not possible to distinguish between processed rRNAs and unprocessed precursor 12S-16S rRNA by qPCR. It would be interesting to investigate whether unprocessed species accumulate in the MRM1/2 double knockout. In MPV17L2 knockdown, downregulation of mtSSU is stronger than that of the mtLSU (Dalla Rosa et al., 2014), in GTPBP10 KO the mtSSU is less affected (Maiti et al., 2018) and in both cases mtLSU and monosome are both downregulated. In contrast, in MRM1/2 KO we observed mild accumulation of mtLSU and a decrease in the monosome (Figure 4.33, Figure 4.35). This represents yet another mode of connection between mtSSU and mtLSU assembly pathways.

MRM1-MRM2 double knockout cell line shows strong mitochondrial function deficiency

Simultaneous inactivation of both MRM1 and MRM2 in HEK293 cells led to impaired mitochondrial translation, resulting in a severe mitochondrial deficiency. Except for a few proteins, all respiratory chain complexes I through IV had decreased steady state levels in the double knockout (Figure 4.15, Figure 4.26). Complexes I, III and IV contain mitochondrial encoded proteins and their defects have been shown in other cells lines with defective mitochondrial translation (D'Souza et al., 2021; Metodiev et al., 2009; Pearce et al., 2017; Van Haute et al., 2019) including in MRM2 knockout (Rebelo-Guiomar et al., 2021). Levels of Complex II are often unchanged in these cell lines since all its components are encoded in the nucleus. Nevertheless, Complex II contains heme *b* and iron-sulphur clusters, and synthesis of these cofactors relies on functional mitochondria (Braymer and Lill, 2017; Piel et al., 2019). In mammalian cells a mutation affecting heme binding in SDHC lead to destabilisation of Complex II and reduction of steady state levels of SDHB, SDHC, and SDHD (Lemarie and Grimm, 2009). Therefore, mitochondrial dysfunction that would impair processes of heme and/or iron-sulphur cluster biogenesis could lead to disruption of Complex II. We indeed observed proteins involved in iron homeostasis (FTL1, FTH, HMOX1) (Gozzelino and Soares, 2014) being the most upregulated in the MRM1/2 double knockout compared to the MRM2 KO (Figure 4.29) as well as to wild type HEK293 cells. The exact reason and mechanism of upregulation of these genes requires further investigation.

Although mitochondrial OXPHOS did not function properly in MRM1/2 KO due to dysfunctional mitochondrial translation, a few proteins from Complexes I, III and IV were significantly upregulated (Figure 4.26). In Complex I, NDUFB11 was upregulated ~2.5-times compared to the wild type (Figure 4.26, top left). This protein is part of an ND4-module initial assembly intermediate together with NDUFB5, NDUFB10 and NDUFB6 (Guerrero-Castillo et al., 2017; Signes and Fernandez-Vizarra, 2018). All four proteins of initial ND4-module were upregulated in the MRM1/2 KO (Figure 4.28). Assembly is likely to stall at this point in MRM1/2 KO, as the next step would be joining of mitochondrial encoded mt-ND4 protein. We observed more than two-fold upregulation of UQCR10 and CYC1 proteins from Complex III (Figure 4.26, middle left). These proteins were recently shown to interact with various intermediates of dimeric Complex III in mt-CYB-deficient human cell lines (Páleníková et al., 2021b; Protasoni et al., 2020). Three Complex IV proteins that were upregulated in MRM1/2 KO, COX5A, COX4I1 and HIGD1A, form the

initial Complex IV assembly (Vidoni et al., 2017). In a mt-CO3 mutant cybrid cell line these three proteins were the least downregulated from the Complex IV components (Vidoni et al., 2017). In MRM1/2 KO we detected upregulation of COX5A, HIGD1A and to the lower extent of COX4I (Figure 4.26, middle right). Accumulation of proteins that form early assembly intermediates of Complexes I, III and IV in MRM1/2 KO suggest a compensatory effect, when the cells are trying to compensate for an OXPHOS deficiency by upregulating the synthesis of individual components. As a very low amount of mtDNA encoded proteins are synthesised in the double knockout, only early assembly intermediates can be formed and proteins that cannot be incorporated into the OXPHOS complexes are most likely degraded.

As a further consequence of impaired OXPHOS function, MRM1/2 KO cells were unable to survive without uridine supplementation in the culture media (Figure 4.18). This phenotype was unique to the double knockout as both MRM1 KO, which had growth rate similar to wild type, and MRM2 KO, which had reduced growth rate, were able to proliferate without uridine supplementation. Uridine dependency is typically observed in cells lacking mitochondrial DNA, as pyrimidine synthesis is dependent on the activity of dihydroorotate dehydrogenase, which is dependent on the function of respiratory complexes (Grégoire et al., 1984). It appears that even a very low level of OXPHOS function can sustain the function of *de novo* uridine synthesis, since MRM2 KO cells were able to proliferate without uridine supplementation (Figure 4.18), although activity of OXPHOS enzymes in this cell line is very low or even below the detection limit when analysed by in-gel activity assay (Rebelo Guiomar, 2020). Activity of OXPHOS enzymes has not been analysed in MRM1/2 KO, but as we observed uridine dependency it is likely to be even lower than in the MRM2 KO and therefore below the detection limit of the available methods.

To examine the specificity of the observed mitochondrial defects in MRM1/2 KO we performed complementation experiments, re-expressing MRM1 or MRM2 proteins in this cell line. As expected, complementation of double knockout with MRM2, which should in effect produce a cell line equivalent to MRM1 KO, rescued OXPHOS function (Figure 4.20), as well as improved growth rate and rescued uridine dependency of the cell line (Figure 4.21A). Expression of MRM1 in the double knockout protein did not have an effect on OXPHOS function (Figure 4.20), however it rescued the uridine dependency of the cell line (Figure 4.21A). This confirmed that the mitochondrial dysfunction of MRM1/2 KO was indeed caused by the absence of MRM1 and MRM2 proteins and not by possible off target effects. Observation of a phenotype that could be rescued by expression of the wild type

MRM1 protein enabled us to examine the importance of 2'-*O* ribose methylation deposited by this enzyme. Complementation of MRM1/2 KO with two catalytically inactive variants of MRM1 did not lead to cell proliferation in the absence of uridine (Figure 4.21B). Based on this we conclude that methylation of G1145 by MRM1 is important for proper mitochondrial function. This is in contrast to the function of MRM2, where the presence of the catalytically inactive enzyme in MRM2 KO cells is able to rescue mitochondrial translation (Rebelo-Guiomar et al., 2021). Considering that MRM1 protein function manifested only once the additional methyltransferase was inactivated, it would be interesting to investigate effects of expression of catalytically inactive MRM2 in the MRM1/2 double knockout cell line.

5 The role of GTPBP8 protein in mitochondrial translation

5.1 Introduction

5.1.1 GTPases and their function in the assembly of mitochondrial ribosome

Assembly of mammalian mitochondrial ribosome is a complex process that requires multiple assembly factors. These protein factors can play various roles, including rRNA modification and processing (Lopez Sanchez et al., 2020; Rackham et al., 2016) and assembly of mitoribosomal subunits (Hilander et al., 2021). GTPases (guanosine triphosphate hydrolases), are a large group of proteins that can bind and hydrolase GTP. Group of GTPases that are involved in translation, signal transduction, cell motility and intracellular transport belong to the TRAFAC class (**tra**nslation **fac**tors) (Leipe et al., 2002). GTPases contain a G domain that is characterised by five regions important for GDP/GTP exchange, GTP-induced conformational change and GTP hydrolysis (Figure 5.1) (Bourne et al., 1991; Leipe et al., 2002; Verstraeten et al., 2011). Most of these regions contain a conserved amino acid sequence. The G1 domain contains a Walker A motif (or P-loop) with the sequence [$GX_4CK(S/T)$] and is involved in binding of α- and ß- phosphates of GTP or GDP. The G2 region in TRAFAC class of GTPases contains a conserved threonine residue. A Mg^{2+} ion that is essential for GTP hydrolysis is coordinated to oxygens of the ß- and γ- phosphates of GTP and to the side chain of threonine from this region as well as serine from region G1. The G3 region contains a Walker B motif [DX_2G] which plays a role in coordination of the catalytic Mg^{2+} and binding of the γ- phosphate. Regions G2 and G3 show large conformational changes between the GDP- and GTP- bound states and are therefore called switch I and switch II regions, respectively. In the G4 region, four hydrophobic or apolar amino acids are followed by (N/T)(K/Q)XD sequence motif. This region determines specificity for the guanine nucleotide by forming hydrogen bonds with guanine ring. The G5 region interacts with the guanine nucleotide mostly indirectly

through hydrogen bonds that stabilize side chains of amino acids in region G4 and is not strictly conserved (Bourne et al., 1991; Leipe et al., 2002; Verstraeten et al., 2011).

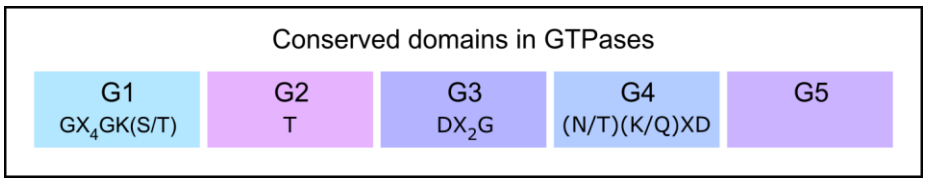


Figure 5.1: Conserved domains in TRAFAC class of GTPases
Schematic of conserved domain that play a role in GTP binding and hydrolysis of GTPases. A conserved sequence motif is shown for each domain. Commonly used names for G1, G2 and G3 domains are P-loop /Walker A motif, switch I and switch II /Walker B motif, respectively. G5 domain does not contain a conserved sequence (Bourne et al., 1991; Leipe et al., 2002).

All GTPases that have been implicated in the assembly of the mammalian mitochondrial ribosome to date belong to the TRAFAC class. These proteins play a role at various stages of the assembly of both small and large subunits (ERAL1, NOA1 and GTPBP5, GTPBP6, GTPBP7, GTPBP10, respectively) (Maiti et al., 2021).

To date, two GTPases have been identified to play a role in the assembly of the small mitoribosomal subunit. NOA1 interacts with the mtSSU (He et al., 2012) and knockout of this protein causes defects in mitochondrial translation and in oxidative phosphorylation (Kolanczyk et al., 2011). Furthermore, mitoribosomes from NOA1 knockout show altered sedimentation on sucrose gradient, with deficient assembly of 55S monosome and mtLSU shifted to the lower density region of the gradient (Kolanczyk et al., 2011). Second mtSSU GTPase, ERAL1, localises to the mitochondrial matrix and is associated with mitoribosomal proteins and 12S rRNA (Uchiumi et al., 2010). It binds to the 3' terminal region of 12S rRNA and the absence of this protein leads to a rapid degradation of 12S rRNA (Dennerlein et al., 2010).

Multiple GTPases are involved in the assembly of the large mitoribosomal subunit. GTPBP5 and GTPBP10 belong to the highly conserved Obg class of GTPases. GTPBP5 localises to mitochondria (Hirano et al., 2006), where it interacts with the inner mitochondrial membrane and is exposed to the matrix site (Kotani et al., 2013). Knockdown of GTPBP5 in HeLa cells leads to elongated mitochondria and abnormal nuclear morphologies (Hirano et al., 2006). In the knockdown cells synthesis of mtDNA encoded proteins is downregulated, with the exception of ATP6 and ATP8 which are upregulated (Kotani et al., 2013). In the GTPBP5 knockout cells, mitochondrial translation is reduced to about 50% of the wild type levels (Cipullo et al., 2021b; Maiti et al., 2020). GTPBP5 co-fractionates with mtLSU on sucrose gradient (Kotani et al., 2013; Maiti et al., 2020). GTPBP5 interacts with the late-stage assembly of mtLSU but not with the mtSSU (Cipullo et al., 2021b; Maiti et al., 2020). It also interacts with several late

assembly proteins including MRM2, GTPBP7, MALSU1, L0R8F8, MTERF4 and NSUN4. Levels of 55S monosome are severely decreased in the GTPBP5 knockout cells. The role of the GTPBP5 in the late stage of the mtLSU assembly and monosome formation is believed to be in facilitating methylation of 16S rRNA by MRM2 (Cipullo et al., 2021b; Maiti et al., 2020).

Second Obg class GTPase, GTPBP10, localises to the mitochondria where it is important for proper function of translation (Lavdovskaia et al., 2018; Maiti et al., 2018). GTPBP10 interacts with 16S rRNA, mtLSU and mitochondrial monosome (Lavdovskaia et al., 2018; Maiti et al., 2018), and its GTPase domain is important for mtLSU binding (Lavdovskaia et al., 2018). In the full knockout of GTPBP10, both mtLSU and mtSSU were downregulated (Maiti et al., 2018). This was accompanied by reduced levels of 12S and 16S rRNA and by increased amount of 12S-16S precursor, suggesting a secondary role of GTPBP10 in mtSSU assembly by stimulating rRNA processing (Maiti et al., 2018).

GTPBP6 belongs to the HflX class of GTPases. In humans it is a pseudoautosomal gene located at the terminal region of the short arms of the sex chromosomes and it contains an N-terminal mitochondrial targeting sequence (Gianfrancesco et al., 2001). GTPBP6 functions as a ribosome recycling factor in human mitochondria, similarly to its bacterial homolog HflX (Lavdovskaia et al., 2020). However, in human mitochondria it also has a role in the assembly of the mitochondrial ribosome. In its absence, late assembly intermediates containing MTERF4, NSUN4, MALSU1 as well as GTPBP5, GTPBP7 and GTPBP10 accumulate (Lavdovskaia et al., 2020). Based on cryo-EM analysis, GTPBP6 plays a role in the maturation of peptidyl transferase centre of the mitoribosome (Hillen et al., 2021).

GTPBP7 is a member of RbgA GTPase subfamily (Kim and Barrientos, 2018) that localises to mitochondria (Barrientos et al., 2003). Similarly to the GTPBP5, it interacts with the inner mitochondrial membrane and is exposed on the matrix site (Kim and Barrientos, 2018; Kotani et al., 2013). Depletion of GTPBP7 levels by RNAi in HeLa cells leads to decreased mitochondrial translation accompanied by reduction in Complex I and Complex IV formation (Kotani et al., 2013). Dysfunctional mitochondrial translation and OXPHOS function is also observed in HEK293T cells depleted of GTPBP7 protein (Kim and Barrientos, 2018). In these cells, sedimentation of mtLSU and mtSSU on sucrose gradient is unaffected, but 55S monosome is virtually absent. GTPBP7 interacts with 16S rRNA, specifically in the domain VI region (Kim and Barrientos, 2018). It further interacts with uL19 and mS27 proteins that form mB6 intersubunit bridge, with mS27 acting as a guanine nucleotide exchange factor in release of GTPBP7 GDP from the mitoribosome

during monosome formation (Kim and Barrientos, 2018). GTPBP7 was recently identified in cryo-EM structure of late assembly intermediate of mtLSU where it was proposed to monitor the folding status of helixes 68-71 of 16S rRNA and prevent early subunit joining (Cipullo et al., 2021a). All mtLSU GTPases function in the late stages of the mitoribosome large subunit assembly. Recently, a synthetic model of the late-stage maturation of mtLSU was created based on several new cryo-EM structures (Cheng et al., 2021; Cipullo et al., 2021a; Hillen et al., 2021; Lenarčič et al., 2021), proposing the order of function of GTPases to be GTPP10 → GTPBP5 → GTPBP7 → GTPBP6 (Sissler and Hashem, 2021).

5.1.2 Function of homologs of GTPBP8 in other organisms

Apart from GTPases discussed in the previous section, GTPBP8 is another TRAFAC class GTPase known in mammalian cells. This is a conserved protein with homologs present in bacteria as well as yeast (Figure 5.2). Its bacterial homolog EngB (also called YsxC) is an essential protein that interacts with the bacterial ribosome and plays a role in its assembly (Cooper et al., 2009; Ni et al., 2016; Schaefer et al., 2006; Wicker-Planquart et al., 2008). In *Staphylococcus aureus* it was found to associate with the large, 50S, ribosomal subunit (Cooper et al., 2009). Mature 70S ribosomes were decreased upon depletion of EngB protein, suggesting its role in the assembly and/or stability of the ribosome in *S. aureus* (Cooper et al., 2009). Similarly, EngB also interacts with 50S large ribosomal subunit in *Bacillus subtilis*, and is able to interact with this subunit in a nucleotide-free state (Wicker-Planquart et al., 2008). Mutation of conserved serine in Walker A motif (S37A) abolishes GTPase activity, however this mutant is still able to bind the 50S ribosomal particle (Wicker-Planquart and Jault, 2015). GTPase activity of wild type EngB protein is stimulated by the 50S ribosome (Wicker-Planquart and Jault, 2015). EngB interacts with both 23S and 16S rRNA that are structural parts of the large and the small ribosomal subunits, respectively (Wicker-Planquart and Jault, 2015). Depletion of EngB in *B. subtilis* leads to accumulation of 44.5S particles on sucrose gradient which lack, as a minimum, proteins uL16 and bL36. These particles sediment slightly slower compared to cells depleted of RbgA, homolog of human GTPBP7 protein, suggesting a distinct assembly intermediate (Schaefer et al., 2006). According to quantitative mass spectrometry analysis of ribosomes from EngB depleted cells, amount of proteins uL16, bL27, bL28, bL33, bL35 and bL36 is strongly decreased in the 44.5S intermediate (Ni et al., 2016). Cryo-EM analysis of 44.5S particles from EngB depleted cells revealed that EngB plays a role in the

maturation of the central protuberance, GTPase associated region and helices in the A, P and E sites of the 50S subunit (Ni et al., 2016).

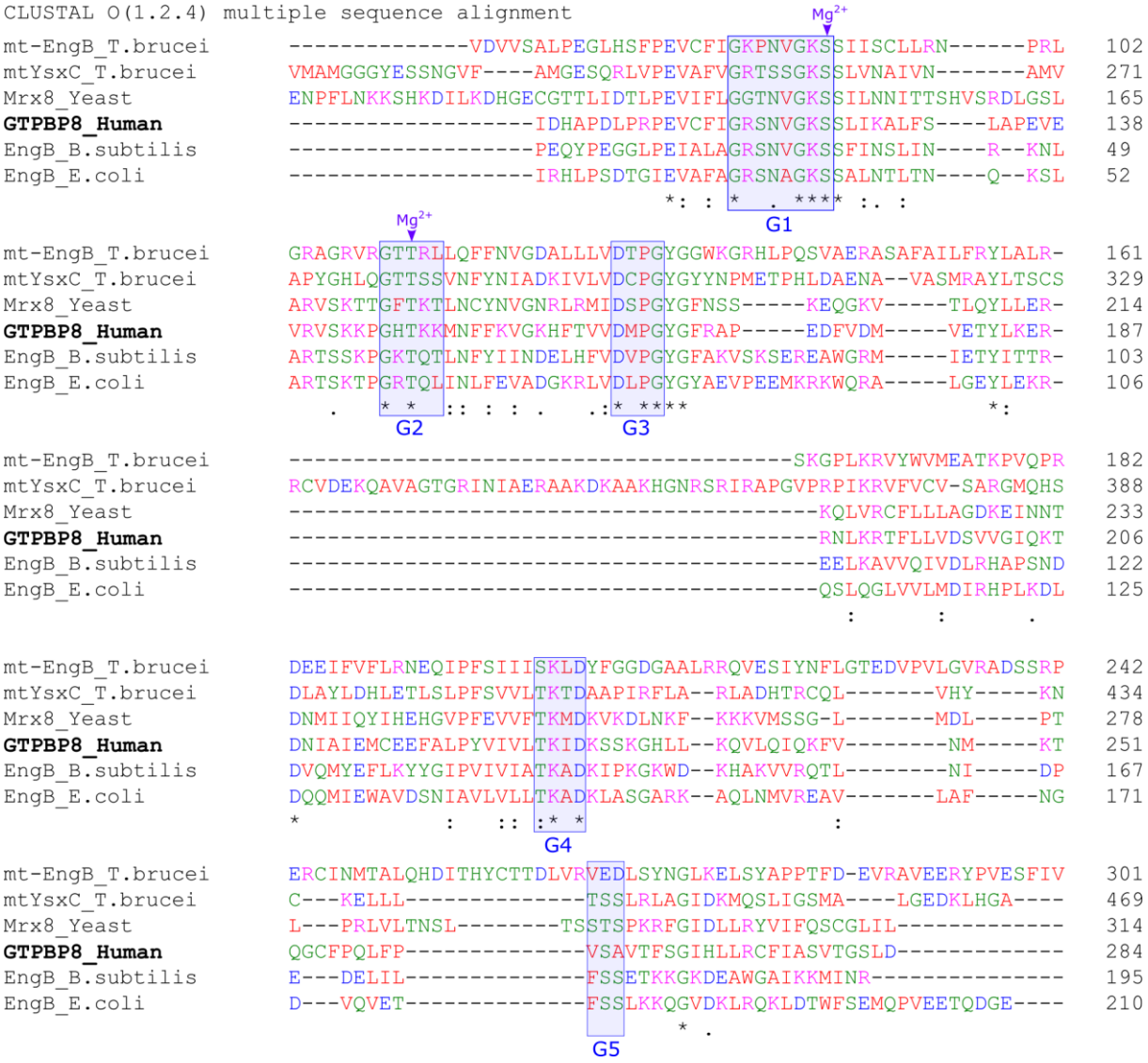


Figure 5.2: Sequence alignment of GTPBP8 from human and its homologs
*Alignment of protein sequences of the human GTPBP8 protein (UniProtKB: Q8N3Z3-1), with its homologs in S. cerevisiae, Mrx8 (UniProtKB: Q05473-1), B. subtilis, EngB (UniProtKB: P38424-1), E. coli, EngB (UniProtKB: P0A6P7-1), T. brucei, mt-EngB (UniProtKB: Q380Y8-1), and T. brucei, mtYsxC (UniProtKB: Q57XB3-1). Alignment was produced using Clustal Omega multiple sequence alignment tool (Goujon et al., 2010; Sievers et al., 2011). Symbols below the alignment represent fully conserved residues (asterisk, *), strongly similar properties (colon, :) and weakly similar properties (period, .) as defined by the Clustal Omega. Conserved G domains are highlighted by blue rectangles and residues that are important for binding of catalytic Mg^{2+} are indicated. G domains and metal binding sites were assigned according to the domains listed for B. subtilis protein in PROSITE (Hulo et al., 2006). Only part of protein sequences with homology is shown. Numbers on the right correspond to the position of amino acid in a whole protein. Full length of proteins is: 284 (human), 314 (yeast), 195 (B. subtilis), 210 (E. coli), 477 (T. brucei, mtYsxC), 655 (T. brucei, mt-EngB).*

Outside of bacteria, homologs of GTPBP8 in other organisms are poorly studied. Two proteins from *Trypanosoma brucei*, Tb927.11.16990 and Tb927.8.2760, were annotated as homolog of bacterial EngB and termed mt-EngB and mtYsxC, respectively (Jaskolowski et al., 2020; Týč et al., 2017). Depletion of mtYsxC in *T. brucei* by RNAi has only minor

effect on cell growth on both glucose-containing and glucose-free growth medium (Týč et al., 2017). However, the effectiveness of RNAi in this experiment was limited since mRNA levels of mtYsxC were downregulated by only 30%. Nevertheless, mitochondrial translation was inhibited in these cells (Týč et al., 2017). Recent cryo-EM structure of mitoribosomal large subunit intermediates from *T. brucei* identified mt-EngB bound to the structure where it interacts with the partially immature rRNA and might play a role in its stabilisation (Jaskolowski et al., 2020).

Homolog of GTPBP8 in *Saccharomyces cerevisiae,* Mrx8 (YDR336W), interacts with the mitoribosome under low-salt conditions (Kehrein et al., 2015). Very recently Mrx8 was further characterised in *S. cerevisiae*, where it was found to be tightly attached to the inner mitochondrial membrane facing the matrix site (Verma et al., 2021). *Δmrx8* cells showed reduced growth in glycerol respiratory media (glucose-free), especially at cold stress conditions. Steady state protein levels of COX1, COX2 and COX3 were reduced in *Δmrx8* cultured at 16°C*,* and COX1 showed lower *de novo* translation (Verma et al., 2021). COX1 translation was affected at the level of both initiation and elongation. Mrx8 was associated with the 54S large mitoribosomal subunit either on its own or as a part of 74S monosome (Verma et al., 2021).

Given the role of several TRAFAC class GTPases in the assembly of mammalian mitoribosome and function of GTPBP8 homologs in the ribosome assembly in other organisms, we investigated whether GTPBP8 also plays a role in the assembly of mammalian mitoribosome.

5.2 Results

5.2.1 Localisation of GTPBP8 into mitochondria

Considering that multiple GTP binding proteins are involved in the mitochondrial function and assembly of the mitochondrial ribosome, we studied whether GTPBP8 is also a mitochondrial protein. To investigate its localisation within the cell, we first used several online tools to assess the probability of mitochondrial targeting sequence on the N-terminus of the human GTPBP8 protein (Table 5.1). Probability score of targeting sequence varied between different tools, ranging from no mitochondrial presequence predicted by MitoFates (Fukasawa et al., 2015) to 0.815 in TargetP (Emanuelsson et al., 2000). Furthermore, two databases of mitochondrial proteome, MitoMiner v4.0 (Smith and Robinson, 2019) and MitoCarta 3.0 (Rath et al., 2021), also provided inconsistent information, with GTPBP8 being assigned as predicted mitochondrial protein only in MitoMiner v4.0.

Table 5.1: Prediction of GTPBP8 localisation
Protein sequence of GTPBP8 (UniProtKB/Swiss-Prot: Q8N3Z3.1) was analysed in online tools predicting N-terminal mitochondrial targeting sequence. Scores presented were calculated by TargetP (Emanuelsson et al., 2000), Predotar (Small et al., 2004) and MitoFates (Fukasawa et al., 2015) software tools. The presence of the protein in MitoCarta 3.0 (Rath et al., 2021) and MitoMiner v4.0 (Smith and Robinson, 2019) is also indicated.

Protein	TargetP	Predotar	MitoFates	MitoCarta 3.0	MitoMiner v4.0
GTPBP8	0.815	0.6	0	No	Yes

As predictions whether GTPBP8 is a mitochondrial protein provided mixed information, we further assessed its localisation by immunocytochemistry (ICC). To be able to achieve a clear signal with low background by ICC a high-quality antibody is needed. Antibody against endogenous GTPBP8, when used for western blot analysis, detected multiple unspecific bands, some of which were more pronounced that the band corresponding to the GTPBP8 (data not shown). Therefore, we performed ICC in 143B cells transiently transfected with a plasmid that allowed expression of GTPBP8 with FLAG-tag on C-terminus (Figure 5.3). FLAG-tag was attached to the C-terminus to prevent its interference with a potential mitochondrial import of the protein. This experiment showed a clear colocalisation of GTPBP8 with the mitochondrial marker TOM20 (Figure 5.3). Therefore, we concluded that GTPBP8 colocalises with mitochondria.

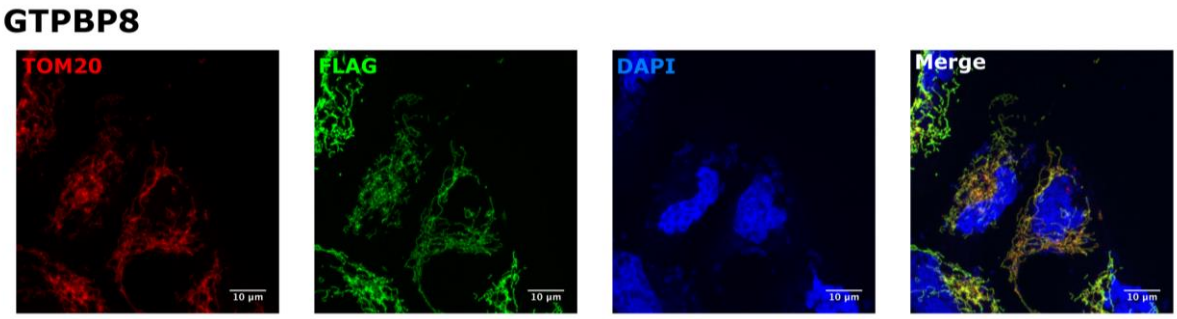


Figure 5.3: Localisation of GTPBP8 by immunocytochemistry
143B cells were transiently transfected with FLAG-tagged GTPBP8 protein and images acquired after fixation and staining. TOM20 was used as a mitochondrial marker (red), FLAG tag indicates GTPBP8 (green) and DNA was stained with DAPI (blue).

5.2.2 Production of GTPBP8 KO cell line by CRISPR/Cas9

To evaluate the role of GTPBP8 in mitochondria, we produced a knockout cell line with the use of CRISRP/Cas9 technology. Knockout cells were produced in a modified HEK293 cell line that contains FRT and attP recombination sites in its genome, providing two loci for site-specific transgene integration (Waldner et al., 2011). The Cas9 protein was integrated into the FRT site previously (performed by C. Powell). This enabled us to produce the knockout with direct transfection of this cell line with tracrRNA and three GTPBP8 specific crRNAs (Table 2.5). Presence of an empty second integration site (attP) was beneficial for complementation of this cell line in follow up experiments. Single clones were isolated following transfection with tracr- and crRNAs and screened for gene disruption by PCR and for presence of GTPBP8 protein by western blot analysis. We were able to isolate two clones, A5 and B1, which showed disruption in the first exon of GTPBP8 gene (Figure 5.4). No protein product was detected in either clone A5 or clone B1 as confirmed by western blotting (Figure 5.5).

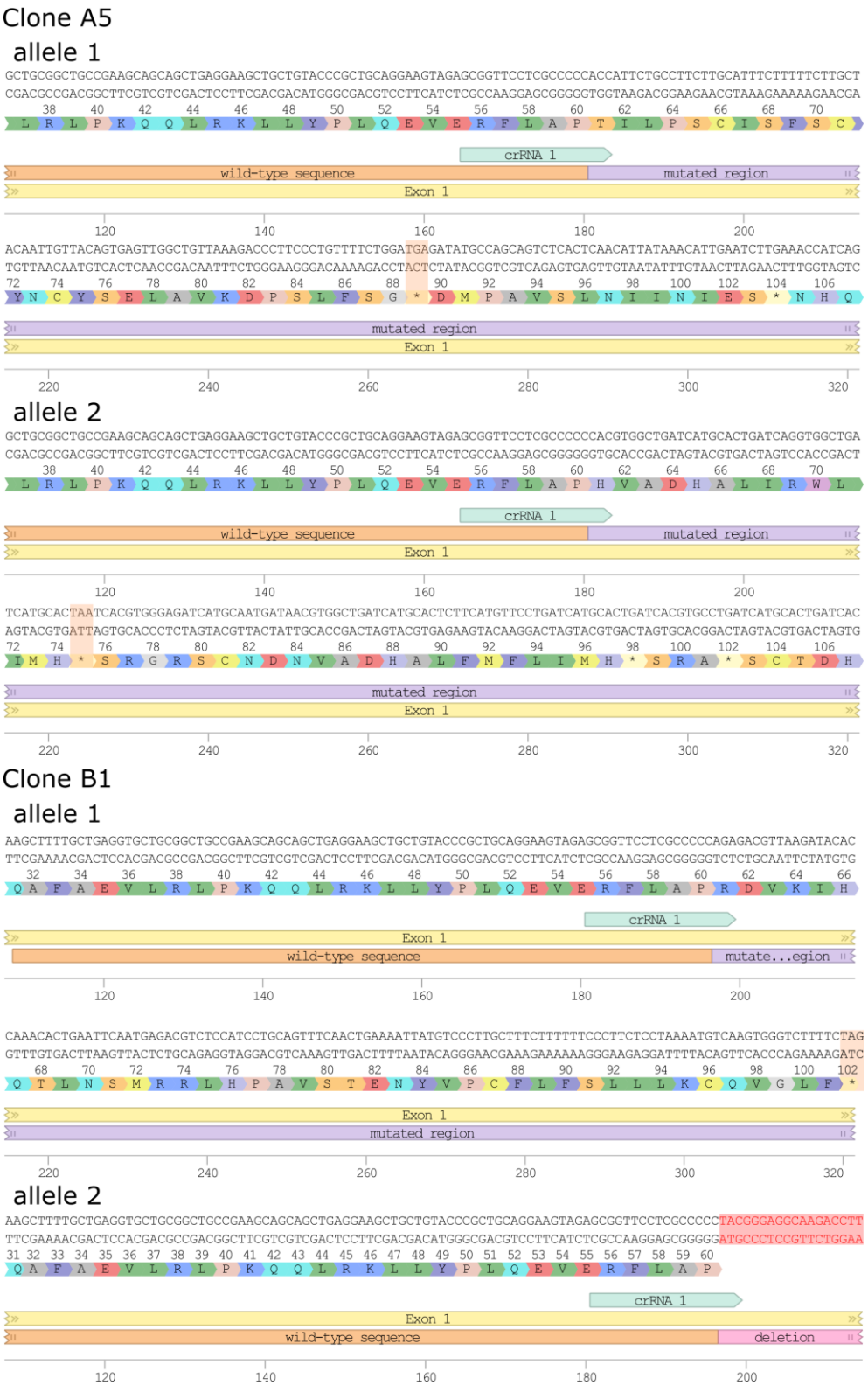


Figure 5.4: Sequence analysis of GTPBP8 knockout clones A5 and B1
Region of exon 1 targeted by three different crRNAs was PCR amplified and analysed by Sanger sequencing. DNA sequence, corresponding protein translation and sequence annotations are shown. Light green arrows (crRNA) depict the region that was targeted for Cas9 restriction. Both alleles in clone A5 contained insertions (mutated region, purple) leading to a premature stop codon (star, orange box). In clone B1 one allele contained insertion (mutated region, purple) with stop codon (star, orange box), and second allele contained large deletion (red box). DNA sequence was visualised in Benchling and manually edited. (Benchling [Biology Software]. (2021). Retrieved from https://benchling.com).

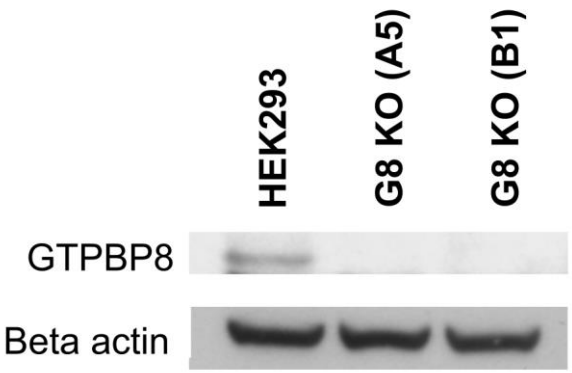


Figure 5.5: Analysis of GTPBP8 knockout by western blot
Presence of GTPBP8 protein in two clones (A5 and B1) was evaluated by western blot analysis of whole cell lysates. No protein was detected in either clone. Beta actin was used as a loading control. (G8 KO = GTPBP8 knockout).

5.2.3 Characterisation of cells lacking GTPBP8

Once we obtained HEK293 clones with a full knockout of GTPBP8, we set to characterise the mitochondrial physiology of the GTPBP8 KO cells. Production of ATP and maintenance of the membrane potential via the process of oxidative phosphorylation are major functions of mitochondria. As proper function of OXPHOS complexes is crucial for these mitochondrial functions, we first investigated changes in the steady state levels of components of OXPHOS. Large decrease in levels of complexes that contain mtDNA encoded proteins was observed (Figure 5.6). For NDUF88 and COXII (Complex I and IV, respectively) the signal was completely absent. Small downregulation was also observed for SDHB protein, component of the Complex II which contains only nuclear encoded proteins. This downregulation of Complex II might be a general response of the cell to the dysfunction of OXPHOS. As behaviour of clones A5 and B1 was very similar, we selected clone A5 as representative and performed all following experiments in this clone.

Next, we looked into the levels of mtDNA as a proxy of mitochondrial mass. Copy number of mtDNA was evaluated using a qPCR analysis of cellular DNA. In GTPBP8 KO levels of mtDNA were mildly increased compared to the wild type cells (average of 1.64 and 1.03, respectively), however this change was not statistically significant (Figure 5.7). This suggests a slight increase of mitochondrial mass in GTPBP8 KO cells, which might act as a compensation of decreased OXPHOS function in these cells.

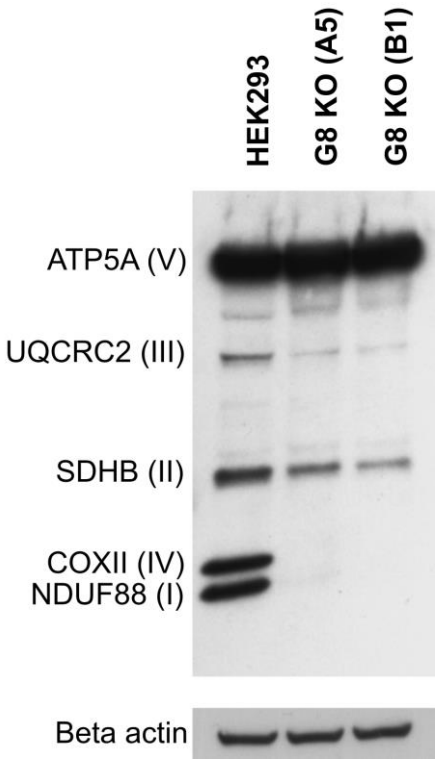


Figure 5.6: Western blot analysis of OXPHOS complexes in GTPBP8 KO
Steady state levels of components of OXPHOS complexes were assessed by western blot analysis of whole cell lysates from wild type HEK293, and GTPBP8 KO clones A5 and B1. Beta actin was used as a loading control. (G8 KO = GTPBP8 knockout)

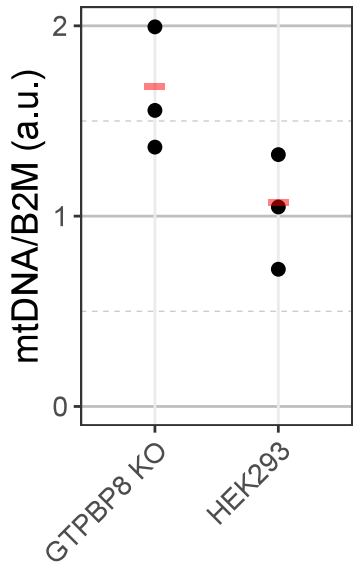


Figure 5.7: mtDNA copy number in GTPBP8 KO cells
mtDNA levels were evaluated by qPCR analysis of whole cell DNA. mtDNA was quantified using delta-delta Ct (ΔΔCt) method and results are shown as $2^{-\Delta\Delta Ct}$. Copy number is expresses as a relative level compared to the levels of B2M gene. Experiment was performed in three biological replicates, each performed in technical triplicates. Values for individual biological replicates (black dots) and the average (red line) are shown. Significance was calculated using two-sided Student's t-test (p-value = 0.076).

Human cells with OXPHOS deficiency are unable to proliferate in the medium with galactose as a sole carbon source (Robinson et al., 1992). Since we observed downregulation of several OXPHOS proteins in GTPBP8 KO cells, we evaluated their growth in medium with high glucose, low glucose, or galactose as a carbon source (Figure 5.8). Growth rate of GTPBP8 KO was much slower compared to the growth of the wild type in both glucose media. Furthermore, knockout cells were unable to survive in galactose medium. This provides additional evidence of importance of GTPBP8 protein for proper OXPHOS function.

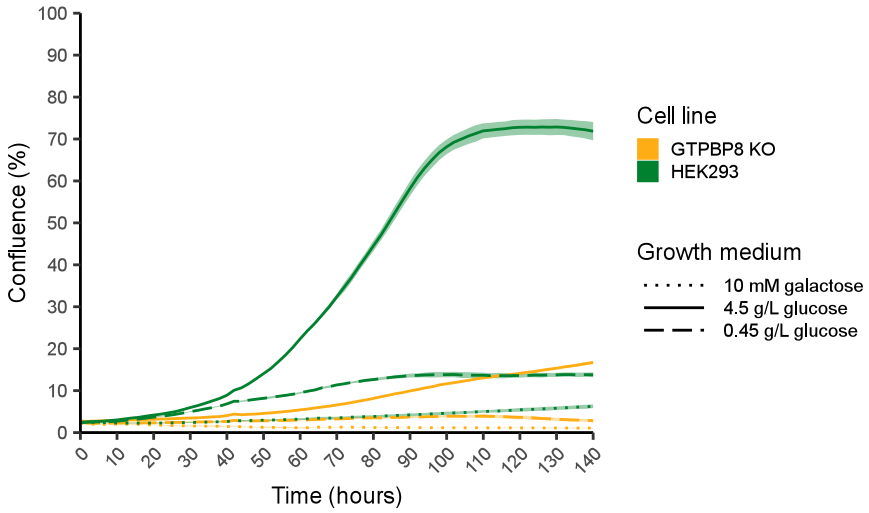


Figure 5.8: Growth of GTPBP8 KO cells in media with various carbon source
Wild type HEK293 (green) and GTPBP8 KO (yellow) cells were grown in medium containing high (4.5 g/L) or low (0.45g/L) glucose concentration or galactose (10mM) as a carbon source. Cells were seeded into a 6-well plate at 3-5% confluency and their growth was analysed by IncuCyte system over 7 day period. Graph shows mean values of confluency from 3 to 6 replicates and shaded areas represent standard deviation.

In previous experiments we observed changes in protein levels of OXPHOS components in GTPBP8 KO (Figure 5.6). We used RT-qPCR analysis of selected mRNAs to evaluate whether this change is also present at the RNA level (Figure 5.9A). mtCO1 and mtCO2 mRNAs were present at a similar level as in wild type, suggesting that the observed absence of COXII protein in GTPBP8 KO (Figure 5.6) is due to the translation defect. On the other hand, mtND1 levels were ~5 times upregulated in the knockout, indicating a compensatory effect of Complex I defect at the level of transcription (Metodiev et al., 2009; Rorbach et al., 2011).

All above results pointed towards dysfunctional translation in GTPBP8 KO cells. To test this directly, we performed metabolic [^{35}S]-methionine labelling of newly synthesised mtDNA-encoded proteins. Cytosolic translation was inhibited by emetine, to prevent it from interfering with the signal of mitochondrial translation. After a 30-minute labelling period we could not detect any translation products in the knockout cells. This confirms that there is no or only very low level of mitochondrial translation in the cells lacking GTPBP8 protein.

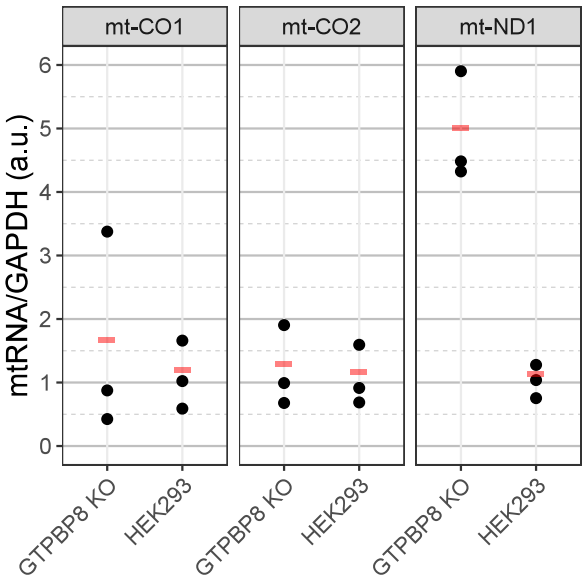


Figure 5.9: mRNA levels in GTPBP8 KO
Levels of mtDNA encoded mRNA transcripts were evaluated by RT-qPCR analysis of whole cell RNA. Experiment was performed in biological triplicates (black dots) and each replicate is an average of three technical replicates. Average of biological replicates is shown as a red line. RNA was quantified using delta-delta Ct (ΔΔCt) method and results are shown as $2^{-\Delta\Delta Ct}$. GAPDH was used as a reference gene and the average of the HEK293 replicates as a reference sample. Significance was calculated using two-sided Student's t-test (mtND1 p-value = 0.0018).

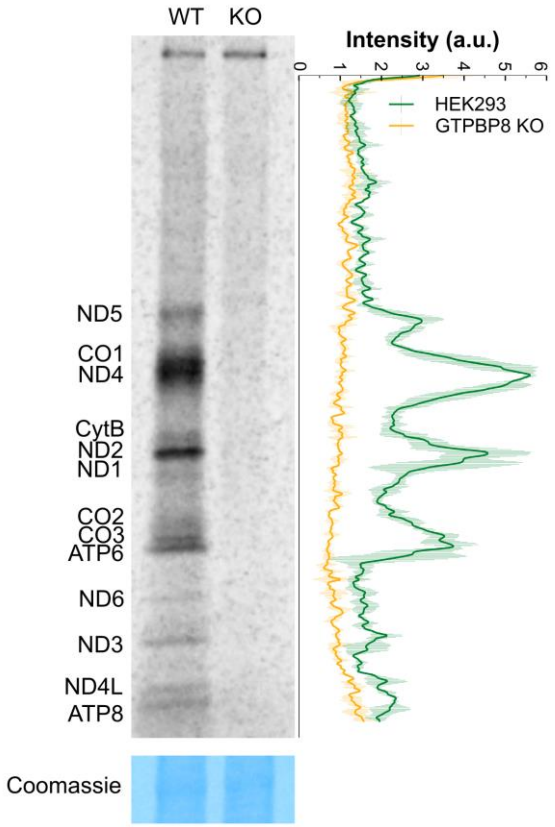


Figure 5.10: Metabolic labelling of mitochondrial translation products
Newly synthesised mtDNA encoded proteins were labelled with [^{35}S]-methionine over a 30-minute period. Half an hour before labelling, cytosolic translation was blocked with an inhibitor emetine. Experiment was performed in triplicate and representative image of a phosphor screen is shown (left). Signal has been quantified using ImageJ and average with standard deviation is shown (right). SimplyBlue Safe Stain Coomassie staining was used as a loading control (Coomassie, bottom).

5.2.4 Complementation of GTPBP8 knockout

GTPBP8 KO cells showed major mitochondrial phenotypes including inability to grow on galactose, decreased steady state OXPHOS protein levels and translation deficiency. To check whether these changes can be unambiguously attributed to the inactivation of GTPBP8, we complemented the knockout with a cDNA encoding the wild type and mutated GTPBP8(S124A) protein. Serine 124 was selected for mutation as a candidate for GTPase activity disruption because it is part of the G1 GTPase domain of GTPBP8 within the conserved $GX_4GK(S/T)$ Walker A motif (Walker et al., 1982). Furthermore, it is predicted to be a binding site of Mg^{2+} ion (UniProtKB: Q8N3Z3). Mutated GTPBP8(S124A) cDNA was a kind gift from Joanna Rorbach. Both wild type and mutated GTPBP8 cDNAs were cloned into pINT plasmid that allows integration into attP site of the knockout cells. Complementation was performed by co-transfection of pINT and pCMV_Int plasmids. Integration of the transgene into attP site renders cells susceptible to neomycin and resistant to puromycin (Waldner et al., 2011). After transfection we isolated single clones and cultured cells with puromycin to select cells with GTPBP8 cDNA integrated into the genome.

First, we confirmed that the GTPBP8 protein is expressed in the complemented cell lines (Figure 5.11A). Expression of GTPBP8 was observed in both complemented cell lines and the amount of protein was substantially higher compared to the wild type HEK293 cells. This overexpression was expected since expression of transgenes in attP docking site is driven by a strong CMV promoter.

Steady state levels of OXPHOS proteins were comparable between HEK293 and knockout cells complemented with the wild type GTPBP8 (Figure 5.11B). However, overexpression of GTPBP8(S124A) did not rescue the phenotype, suggesting that GTPase activity of this protein is important for mitochondrial function. Similarly, growth rate of the knockout expressing wild type GTPBP8 was only slightly slower compared to the HEK293 cells (Figure 5.12A), while cells complemented with GTPBP8(S124A) showed a severe growth defect (Figure 5.12B).

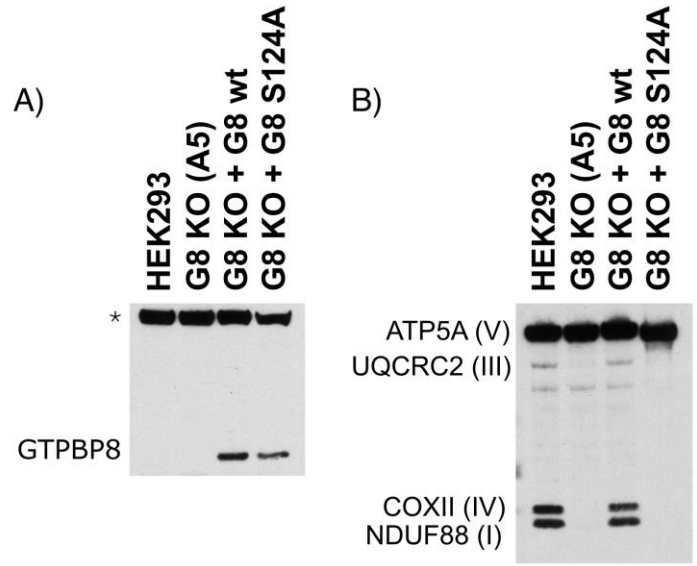


Figure 5.11: Complementation of GTPBP8
*A) GTPBP8 knockout cells were complemented with a wild type and a catalytic mutant of GTPBP8, and the presence of this protein was evaluated by probing with the GTPBP8 antibody. B) Steady state levels of OXPHOS complexes in whole cell lysates of wild type, knockout, and knockout after complementation were evaluated by probing with Total OXPHOS Human WB Antibody Cocktail. Antibody and corresponding complex are indicated. (G8 KO = GTPBP8 knockout; G8 KO + G8 wt = GTPBP8 KO complemented with wild type GTPBP8; G8 KO + G8 S124A = GTPBP8 KO complemented with GTPBP8 with mutation of serine 124 to alanine; * = unspecific signal of GTPBP8 antibody, indicative of loading)*

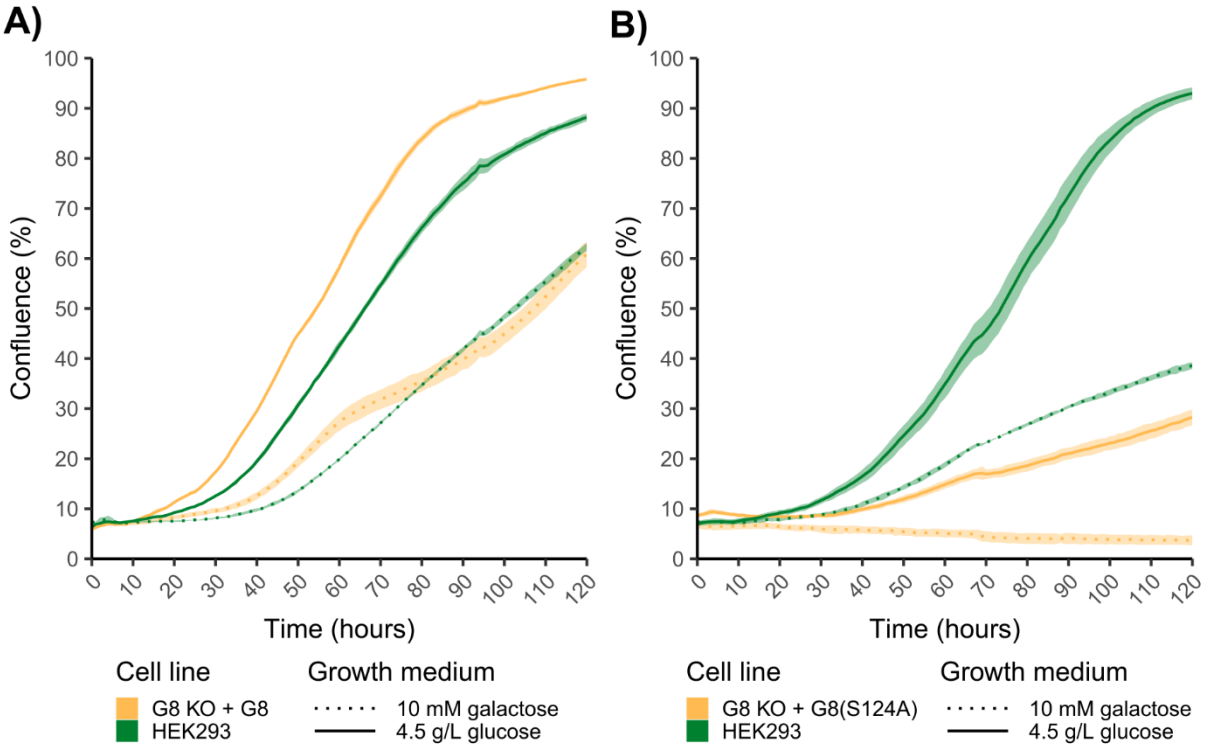


Figure 5.12: Effect of GTPBP8 KO complementation on growth rate
Wild type and knockout cells upon complementation were grown in the medium with glucose or galactose as a sole carbon source. Growth rate was monitored using IncuCyte imaging system. Average of three biological replicates is shown with standard deviation represented by shaded areas. A) Comparison of growth rate between wild type (HEK293) and GTPBP8 KO complemented with wild type GTPBP8 protein (G8 KO + G8). B) Comparison of growth rate between wild type (HEK293) and GTPBP8 KO complemented with mutant GTPBP8 protein (G8 KO + G8(S124A)).

Finally, we evaluated whether complementation rescues the translation deficiency of the GTPBP8 KO. Metabolic labelling of mtDNA encoded proteins showed no difference between parental HEK293 cells and GTPBP8 KO complemented with wild type GTPBP8 (Figure 5.13). This confirms that observed mitochondrial phenotypes are a result of dysfunctional translation due to the absence of GTPBP8 protein.

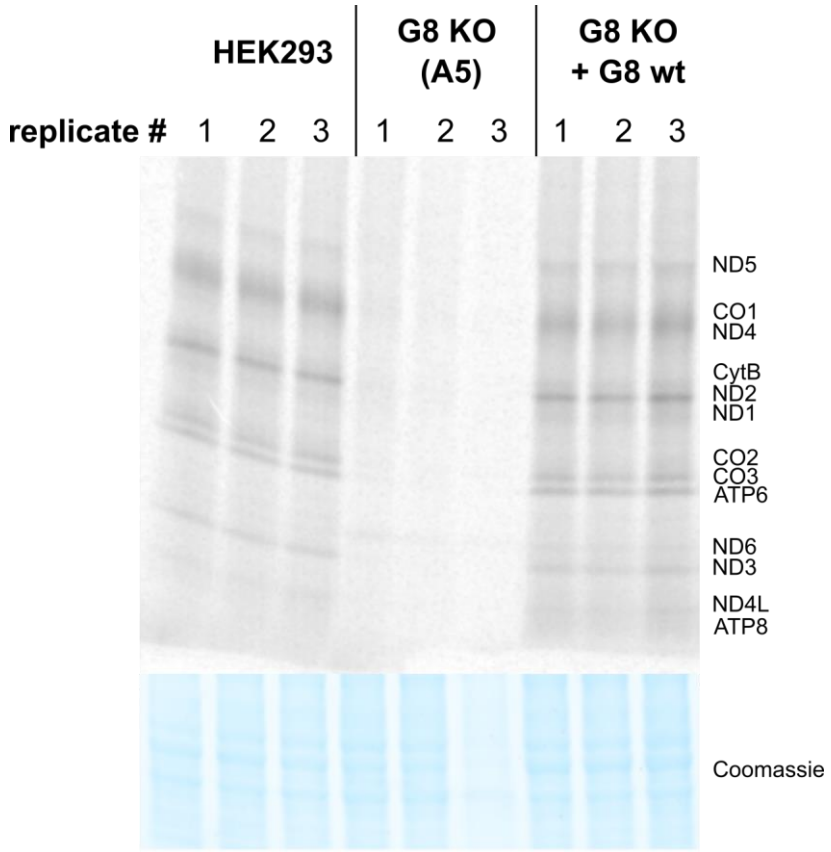


Figure 5.13: Mitochondrial translation upon complementation
Labelling of newly synthesised mtDNA encoded proteins with [^{35}S]-methionine over a 30-minute period was performed in parental HEK293 cell line, GTPBP8 KO (G8 KO) and knockout complemented with wild type GTPBP8 (G8 KO + G8 wt). Half an hour before labelling, cytosolic translation was blocked with an inhibitor emetine. Results of three biological replicates of the experiment are shown. SimplyBlue Safe Stain Coomassie staining was used as a loading control (Coomassie, bottom).

5.2.5 Mitochondrial ribosome in the absence of GTPBP8

Knockout of GTPBP8 led to a dysfunction in mitochondrial translation. Mitochondrial translation is a complex process that, as a minimum, involves steps of mitoribosome assembly, translation initiation, elongation and termination, and ribosome recycling. Furthermore, it requires coordination of nuclear and mitochondrial gene expression, as RNA components of the mitoribosome are encoded in the mtDNA whereas all protein components as well as many assembly factors are encoded in the nucleus. Potentially, role of GTPBP8 in any of these processes could lead to the observed mitochondrial translation

defect. Recently, several other mitochondrial GTPases have been implicated in the assembly of small (Dennerlein et al., 2010; He et al., 2012; Kolanczyk et al., 2011; Uchiumi et al., 2010) and large (Cipullo et al., 2021b; Hirano et al., 2006; Kotani et al., 2013; Lavdovskaia et al., 2018, 2020; Maiti et al., 2018, 2020) mitoribosomal subunits. Bacterial homolog of GTPBP8, EngB plays a role in the biogenesis of the large ribosomal subunit (Britton, 2009). Based on this information, we investigated the state of mitochondrial ribosome in GTPBP8 KO cells as a most likely target of the GTPBP8 function.

First, we analysed levels of rRNA transcripts in the GTPBP8 KO by RT-qPCR. Mitochondrial ribosome contains two rRNAs, 12S and 16S, that are component of small and large mitoribosomal subunits, respectively. Measured rRNA amounts were variable between experiments and no up or down regulation was observed between GTPBP8 KO and wild type (Figure 5.14).

Next, we assessed steady state levels of mitoribosomal proteins (MRPs) in the knockout cell line. Western blot analysis of whole cell lysate showed downregulation of components of mtSSU and mtLSU (uS17m and uL3m, respectively; Figure 5.15). Studied proteins were downregulated to about 25% of the wild type. However, as MRPs are synthesised and imported into mitochondria in excess (Bogenhagen et al., 2018), this change does not necessarily inform about the levels of the mitoribosome in the GTPBP8 KO. For this, the assembly state of the mitoribosome needs to be assessed and the amount of proteins compared in the assembled mitoribosomal subunits.

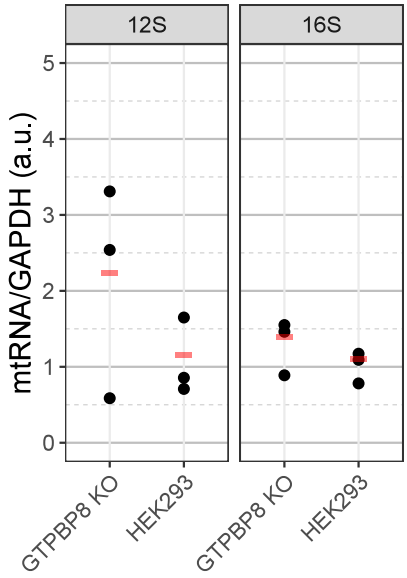


Figure 5.14: RT-qPCR analysis of rRNA levels in GTPBP8 KO
Levels of mtDNA encoded rRNA transcripts were evaluated by RT-qPCR analysis of whole cell RNA. Experiment was performed in biological triplicates and each replicate is an average of three technical replicates. RNA was quantified using delta-delta Ct (ΔΔCt) method and results are shown as $2^{-\Delta\Delta Ct}$. GAPDH was used as a reference gene and the average of the HEK293 replicates as a reference sample. Two-sided Student's t-test was used to evaluate differences between samples and no significant differences were detected.

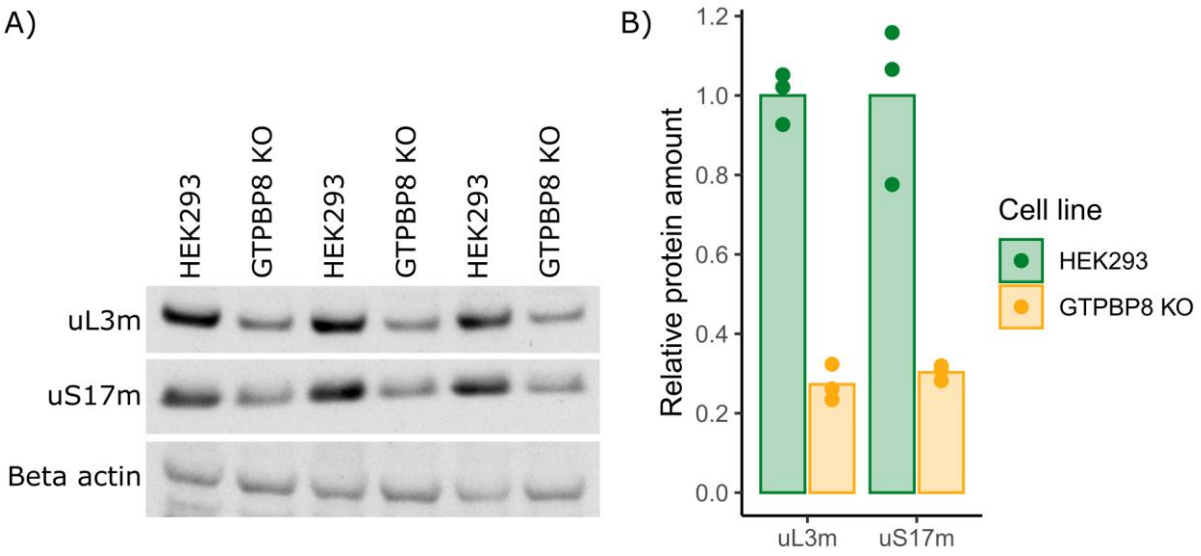


Figure 5.15: Analysis of steady state levels of mitoribosomal proteins in GTPBP8 KO
Whole cell lysates were used for western blot analysis of components of mitochondrial ribosome. A) Three biological replicates were performed with uS17m as a marker of mtSSU and uL3m as a marker of mtLSU. Beta actin was used as a loading control. B) Quantification of western blot in A using ImageJ. Values were adjusted for loading differences using beta actin as a reference and normalised to the average of the protein signal in HEK293. Points represent individual experiments and bars show the average.

To characterise mitoribosome assembly in GTPBP8 KO we resolved its subunits on sucrose gradient by ultracentrifugation followed by western blot analysis. In this experiment, whole cell lysates were resolved on 10-30% linear sucrose gradient. Fractions of equal volume were collected from top of the gradient and analysed by western blotting (Figure 5.16A). We did not observe any major changes in the migration profile of either of the mitoribosomal subunits (Figure 5.16B). However, this result is based on only two selected MRPs (uS17m and uL3m). Furthermore, samples were analysed on two separate sucrose gradients, which could introduce further variation during gradient preparation or sample collection. Even if equal amounts of each sample are loaded onto the gradient, quantitative comparison between these samples is imprecise, since there can be variation introduced during the process of western blotting. Therefore, an alternative approach is needed for precise quantification and evaluation of mitoribosome assembly.

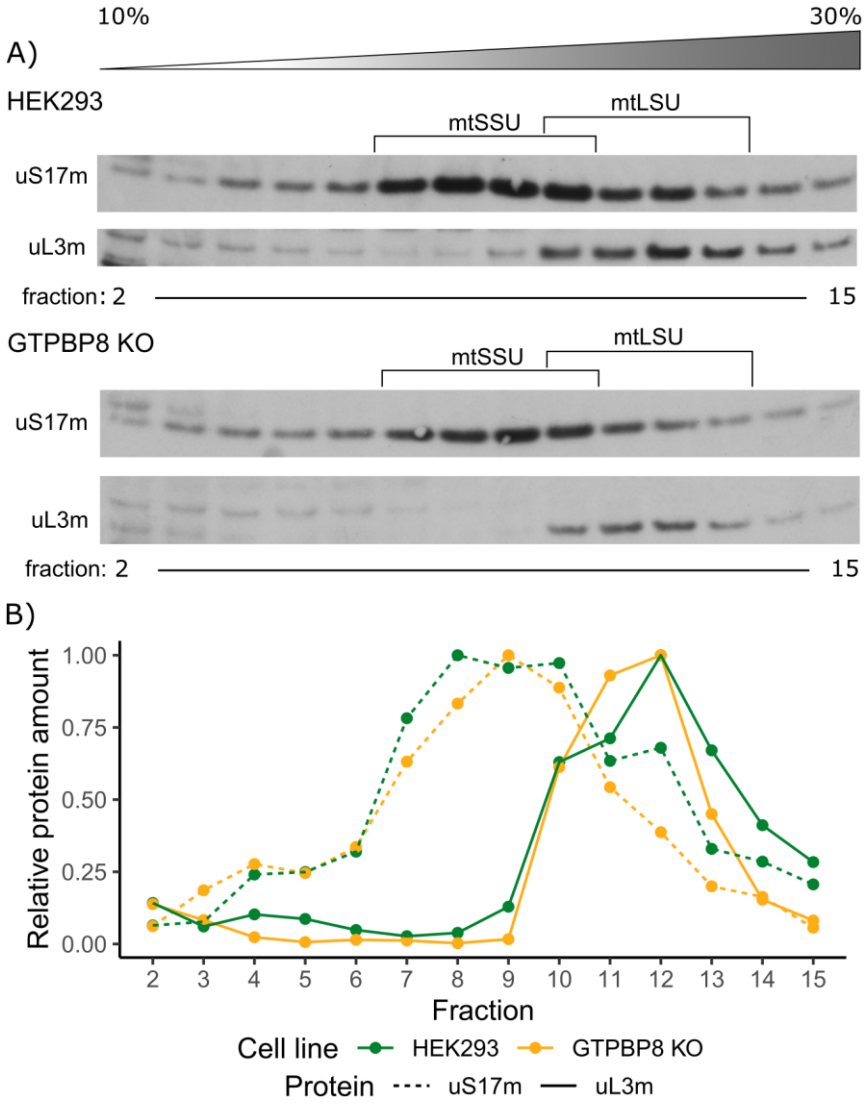


Figure 5.16: Analysis of mitoribosome in GTPBP8 KO cells
A) Whole cell lysates of parental HEK293 cells and GTPBP8 KO cells were resolved on 10-30% linear sucrose gradient. Migration of the subunits of mitochondrial ribosome was visualised by western blotting with antibodies against uS17m (mtSSU) and uL3m (mtLSU). B) Quantification of western blot from A using ImageJ. Protein amount values were normalised to the fraction with maximum signal, separately for each protein and cell line.

We recently developed a qDGMS method for qualitative and quantitative evaluation of mitoribosome assembly (Páleníková et al., 2021a, Chapter 3). This method combines metabolic labelling of proteins in cell culture (SILAC) with mass spectrometry analysis. In this method the two studied cell lines (knockout and wild type) are differentially labelled, which allows for them to be mixed at the early stages of sample processing and analysed on a single sucrose gradient, mitigating technical issues discussed above. SILAC labelling further allows for precise quantification of changes in protein amounts between the samples.

Using qDGMS method in combination with the data analysis R package ComPrAn, we compared mitoribosomes between HEK293 and GTPBP8 KO cells. Total of three

biological replicates were performed. In two replicates wild type cells were cultured in medium with heavy (containing ^{13}C and ^{15}N isotopes) and knockout cells with light (containing ^{12}C and ^{14}N isotopes) arginine and lysine (experiments A and C) and one biological replicate was performed with reverse orientation (experiment B). Representative results from a reciprocal labelling are shown. All proteins of the mitoribosomal small subunit were detected in both cell lines (Figure 5.17). Only bL36m protein was not detected in either of the replicates of the qDGMS experiment and mL42 was missing in one of the replicates (Figure 5.18). bL36m is a small protein with only 11.78 kDa mass that is difficult to detect by mass spectrometry analysis (Bogenhagen et al., 2018). mL42 protein was absent in quantitative analysis of one of the replicates, but this protein was still detected in the GTPBP8 KO sample. Protein was absent from the quantitative analysis because no common peptide between the two studied cell lines was detected. Overall, these data provide almost complete coverage of the structural components of the mitochondrial ribosome. Migration of proteins was similar between wild type and knockout cells and no assembly intermediates were observed (Figure 5.17, Figure 5.18). This suggests that both mtSSU and mtLSU are assembled correctly in the absence of the GTPBP8 or that the assembly is disrupted only in late stages. Small change in mass caused by absence of only a few proteins might be below the resolution provided with the current setup of the qDGMS. However, we observed accumulation of both mtSSU and mtLSU in the GTPBP8 KO (Figure 5.17, Figure 5.18). For each mtSSU and mtLSU protein we calculated an average relative protein abundance in selected fraction across three experiment replicates. Median abundance value for all mtSSU components was 0.73 in GTPBP8 KO compared to 0.44 in HEK293 (Figure 5.19A). For mtLSU, median relative protein abundance in fractions where this subunit peaks were 0.78 and 0.56 for knockout and wild type, respectively (Figure 5.19B). Considering that mitochondrial translation is impaired in GTPBP8 KO cells, this upregulation might be a compensatory mechanism of the cell striving to produce more ribosome particles. Mild decrease of median relative protein abundance of both mtSSU and mtLSU components in the gradient zone with monosome particles might indicate that the accumulated mtSSU and mtLSU particles are not fully mature or cannot assemble properly (Figure 5.19A, B). However, as majority of proteins migrated in gradient zones of individual mtSSU and mtLSU subunits and only small proportion was found in the monosome zone further validation is needed to conclusively assess whether monosome formation is impaired in the GTPBP8 KO.

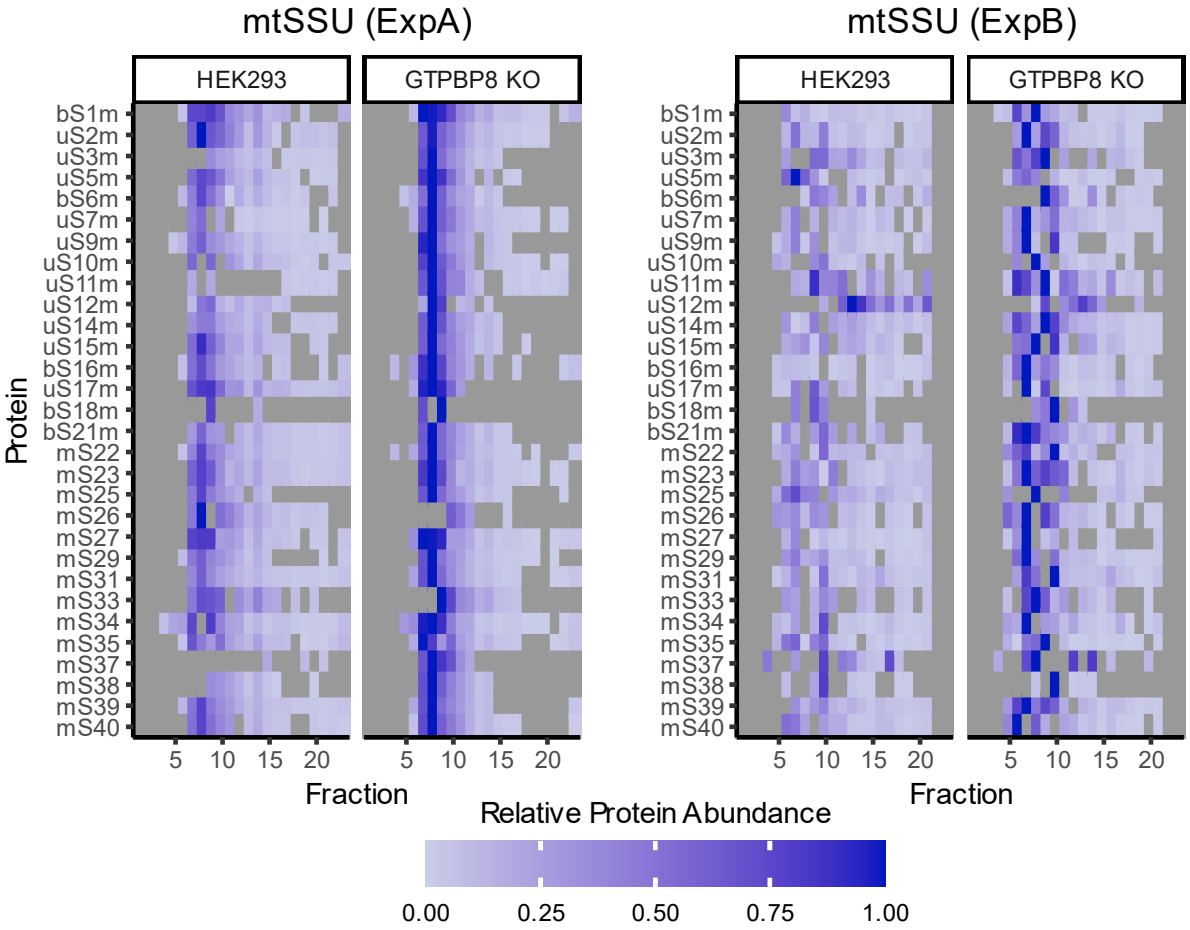


Figure 5.17: qDGMS analysis of mitoribosomal small subunit in GTPBP8 KO cells
Three biological replicates of qDGMS experiment were performed and results of experiments A (ExpA) and B (ExpB) with reciprocal labelling are shown (wild type – heavy, knockout – light in ExpA; wild type – light, knockout – heavy in ExpB). Data were analysed and visualised with the use of ComPrAn R package. Results of quantitative data analysis approach are shown, where proteins were normalised to the highest abundance value across both labelling states, allowing to compare up or down regulation of proteins between cell lines. Grey squares indicate fractions in which protein was not detected.

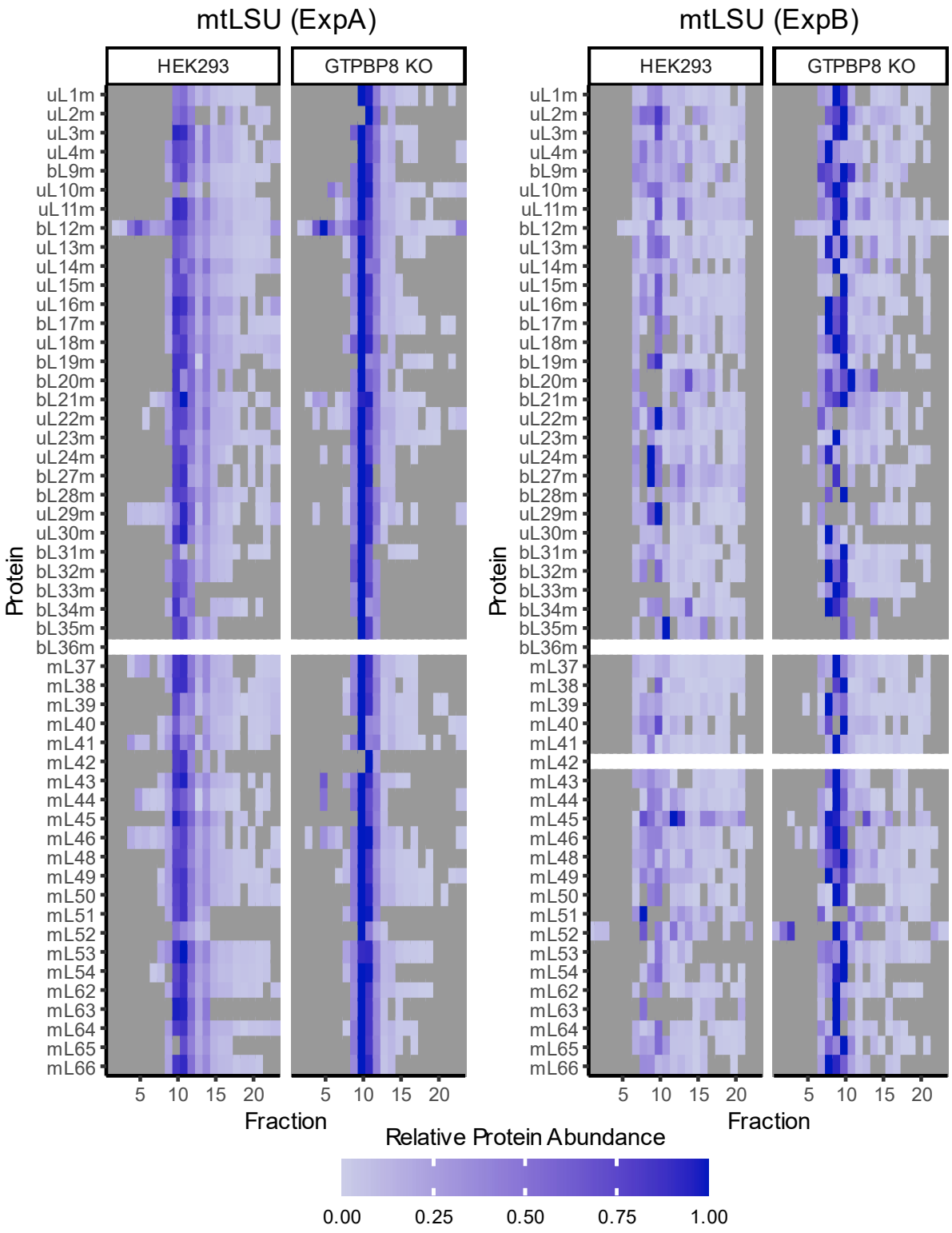


Figure 5.18: qDGMS analysis of mitoribosomal large subunit
Data were processed and visualised in a same way as Figure 5.17, but for large mitoribosomal subunit. (ExpA: wild type – heavy, knockout – light; ExpB: wild type – light, knockout – heavy). Grey squares indicate fractions in which protein was not detected. For proteins that were not detected in any fraction a white line is shown.

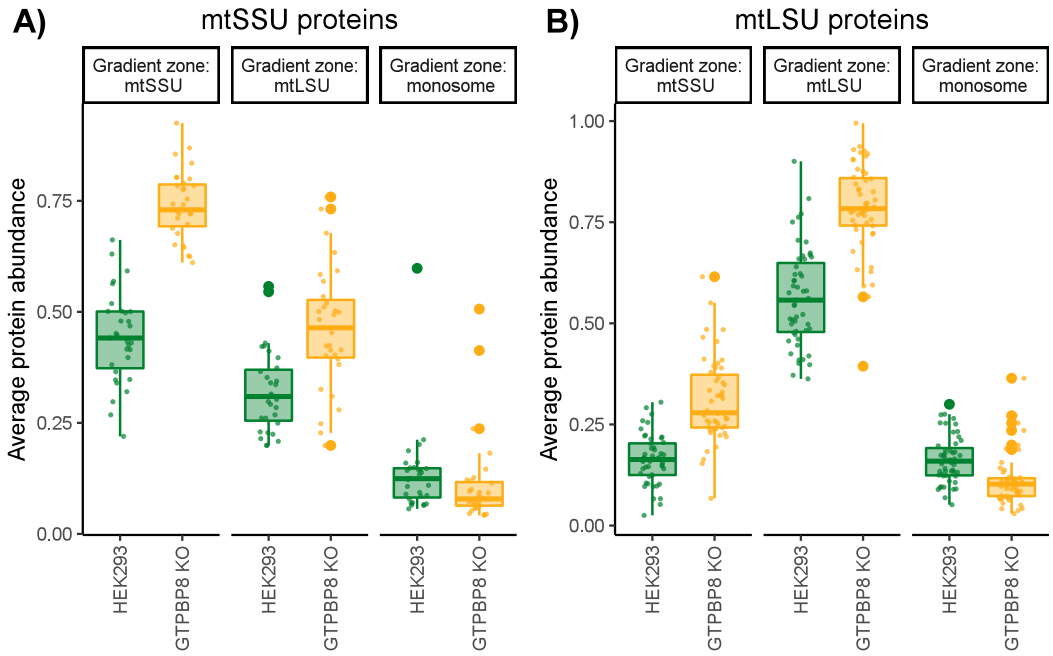


Figure 5.19: Quantification of mitoribosomal subunits in sucrose gradient
Summary of qDGMS data for mitoribosomal proteins to evaluate general change in subunit amounts in the GTPBP8 KO cells compared to wild type. For each protein, an average of relative protein abundance value was calculated in each gradient zone (mtSSU, mtLSU and monosome). Average relative abundance for mtSSU was calculated from fractions 7 to 9, 6 to 7 and 6 to 7 for experiments A, B and C, respectively. Average relative abundance for mtLSU was calculated from fractions 10 to 11, 9 to 10 and 8 to 10 for experiments A, B and C, respectively. Average relative abundance for monosome was calculated from fractions 14, 13 to 14 and 13 to 14 for experiments A, B and C, respectively. For each protein, single abundance value was calculated for each gradient zone as an average of the three experiments. Average relative protein abundance in the three gradient zones for all protein components of mtSSU(A) and mtLSU (B) is shown for wild type HEK293(green) and GTPBP8 KO (yellow). Values for individual proteins are shown in small dots. Boxplots depict median (thick horizontal line), first and third quartiles (box hinges) and whiskers extend to the highest or lowest point in each group within 1.5 times of inter-quartile range, any points above this are shown as larger dots.

Given the accumulation of MRPs in GTPBP8 KO we investigated whether migration through sucrose gradient and amount of mitoribosome assembly or translation factors is affected in GTPBP8 KO cells. To do this, we evaluated translation associated factors in qDGMS data. Quantitative analysis of factors that were detected in at least one of the three experiment replicates is shown in Figure 5.20 (A, C and E). Migration of translation associated factors was compared to the median migration profile of mtSSU and mtLSU proteins (Figure 5.20B, D and E), largely showing co-migration of protein factors with expected mitoribosome subunit. Several proteins were upregulated in GTPBP8 KO, including 12S rRNA methyltransferases METTL15 and TFB1M, mtSSU assembly factor YbeY and mtLSU assembly factor GTPBP10 (Figure 5.20A, C and E). We further observed accumulation of translation initiation factors mtIF2 and mtIF3 as well as anti-association factors MALSU1, L0R8F8 and mtACP (Figure 5.20A, C and E) which are bound to mtLSU in the late stages of the ribosome assembly (Brown et al., 2017). Conversely, ribosome recycling factors EFG2 and MRRF were downregulated in the knockout cells (Figure 5.20A, C and E).

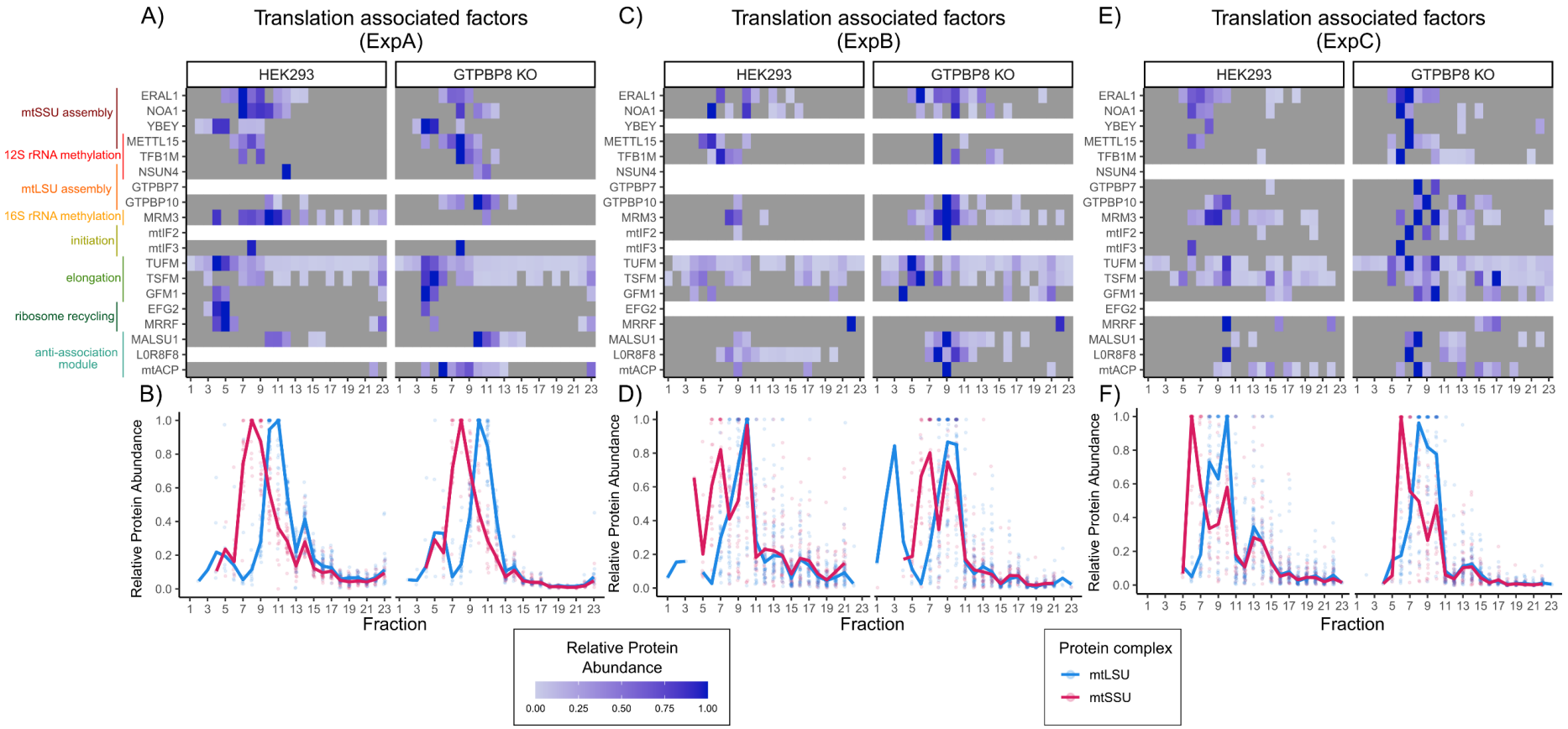


Figure 5.20: Evaluation of protein factors associated with translation in GTPBP8 KO cell line
Quantitative (A, C, E) and qualitative (B, D, E) analysis of qDGMS data from three biological replicates comparing GTPBP8 KO and wild type cells. A, C, E) Migration of translation associated factors was evaluated and is depicted as heatmap. Proteins were normalised to the highest abundance value across both labelling states, allowing to compare up or down regulation of proteins between cell lines. Grey squares indicate fractions in which protein was not detected. For proteins that were not detected in any fraction a white line is shown. B, D, E) Results of qualitative data analysis approach for mtSSU and mtLSU proteins are shown as a reference to indicate fractions of the sucrose gradient where mtSSU and mtLSU peak. Values for each protein were normalised to the highest abundance value separately for HEK293 and GTPBP8 KO samples. Relative protein abundance values for each detected protein component of mtSSU and mtLSU are shown as circles. Median of all proteins for each subunit is also shown. Data were analysed and visualised with the use of ComPrAn R package

5.3 Discussion

GTPBP8 is a mitochondrial protein important for OXPHOS function and mitochondrial translation

Mitochondria are a cellular organelle that provides cells with vital energy for their survival. Many nuclear encoded proteins are transported into mitochondria where they perform various functions. Our knowledge and understanding of the processes these proteins perform continues to expand. Here we characterised a putative GTP binding protein 8 (GTPBP8). Using an immunocytochemistry approach, we localised GTPBP8 to mitochondria (Figure 5.3). This is consistent with a very recent study where mitochondrial localisation of GTPBP8 was reported for HEK293 cells (Verma et al., 2021).

To characterise the role of GTPBP8 in the cell, we created a GTPBP8 knockout cell line (Figure 5.4, Figure 5.5). Importance of GTPBP8 homologs for cellular growth differs between organisms. The bacterial homolog of GTPBP8, EngB, is essential for growth (Cooper et al., 2009; Dassain et al., 1999; Prágai and Harwood, 2000). In *S. cerevisiae*, GTPBP8 homolog Mrx8 is required for growth on glucose-free medium, but only under the cold stress conditions (Verma et al., 2021). In *T. brucei,* depletion of mtYsxC has only a minor effect on growth (Týč et al., 2017). We showed that in human HEK293 cells, inactivation of GTPBP8 protein led to markedly reduced growth on glucose-containing medium and inability to proliferate in the absence of glucose (Figure 5.8). Variability in effect of gene knockout among different organisms is not surprising as it has been previously observed for other genes (D'Souza, 2018; Lövgren and Wikström, 2001; Sirum-Connolly et al., 1995).

Inability to survive on glucose-free galactose-containing medium is typical for cells with dysfunction in mitochondrial oxidative phosphorylation (Robinson et al., 1992). This was indeed the case in GTPBP8 KO, where steady state levels of Complexes I, III and IV were markedly decreased. (Figure 5.6). As opposed to Complex II, which was mostly unaffected in GTPBP8 KO (Figure 5.6), these complexes contain mitochondrially encoded components. This suggested that GTPBP8 might be important for the mitochondrial translation. We investigated *de novo* mitochondrial protein synthesis and did not detect any translation product within a 30-minute translation assay (Figure 5.10). All mtDNA encoded proteins were affected equally, which is in contrast with the translation defect observed in yeast *Δmrx8* cells, where translation was specifically inhibited for Cox1 (Verma et al., 2021). In *S. cerevisiae,* translation of Cox1 is regulated by translation factors

Mss51 and Pet309, but this mechanism is not conserved in mammals (Dennerlein and Rehling, 2015; Perez-Martinez et al., 2009; Zamudio-Ochoa et al., 2014). It is therefore unlikely that GTPBP8 would influence translation of a single mRNA in a similar manner to the yeast system. Our results indicate a general translation impairment in GTPBP8 KO rather than an mRNA specific defect.

Levels of mtDNA in GTPBP8 KO were slightly upregulated (Figure 5.7) and mRNA of mt-ND1 was significantly upregulated in the absence of GTPBP8 (Figure 5.9). This upregulation might be a mechanism by which the cell tries to compensate for a deficiency in translation, as upregulation of mitochondrially encoded RNA in cells as well as in mouse models with impaired mitochondrial translation has been observed previously (Cámara et al., 2011; Metodiev et al., 2009; Rorbach et al., 2011; Van Haute et al., 2016).

Complementation of GTPBP8 KO with wild type GTPBP8 protein confirmed that the observed mitochondrial defects are due to the absence of this protein (Figure 5.12, Figure 5.13). Rescue of mitochondrial defect was dependent on the GTPase activity of the GTPBP8, since expression of the protein with a mutated catalytic residue had no effect on steady-state levels of OXPHOS protein components and cell growth (Figure 5.11B and Figure 5.12, respectively). This is consistent with studies in yeast, where only expression of a catalytically active protein restored wild type phenotype (Verma et al., 2021). Interestingly, human GTPBP8 with a mitochondrial targeting sequence from yeast MTG3 protein fused to its N-terminus was able to partially rescue growth defect of *Δmrx8* cells on glycerol medium at 16°C (Verma et al., 2021). This functional conservation suggests a universal role of GTPBP8 in the mitochondrial translation among eukaryotic species.

Small and large mitoribosomal subunits, but not the monosome, accumulate in absence of GTPBP8

Bacterial homologs of GTPBP8 are important for the assembly of the bacterial ribosome (Cooper et al., 2009; Ni et al., 2016; Schaefer et al., 2006; Wicker-Planquart and Jault, 2015). As mitochondrial translation is impaired in GTPBP8 KO, we investigated whether this is caused by disruption of assembly of the mitochondrial ribosome. Migration of small and large mitoribosomal subunits was unaffected by the GTPBP8 inactivation (Figure 5.16). However, the amounts of the mitoribosomal components were perturbed in the knockout cell line. On the whole-cell level, amount of mtSSU and mtLSU were downregulated in the GTPBP8 KO (Figure 5.15). Contrary to this, the amount of assembled mitoribosomal subunits was higher in the knockout (Figure 5.16, Figure 5.17). Protein amounts on the whole-cell level were analysed only for selected proteins (uL3m

for mtLSU and uS17m for mtSSU). Comprehensive quantitative comparison of mitoribosomal proteins (MRPs) between wild type and GTPP8 KO, for example by SILAC mass spectrometry, will inform about the extent of perturbation across all MRPs. The seemingly contradictory result between quantification on a whole-cell level and on the level of assembled mitoribosome might be explained by the excessive synthesis of mitoribosomal proteins under normal conditions (Bogenhagen et al., 2018). Bogenhagen and colleagues analysed kinetics of the assembly of mitochondrial ribosome by pulse-chase SILAC mass spectrometry and found that MRPs are synthesised and imported into mitochondria in excess, with the unassembled protein copies being quickly degraded. Decrease of the overall synthesis and/or import rate of the MRPs in GTPBP8 KO could lead to the observed downregulation of steady state level of proteins on the whole cell level. Since under normal conditions there is an excess of free MRPs, it is possible that the amount of MRPs available in the knockout is still sufficient to sustain the levels of assembled mitoribosome. Experimental validation comparing synthesis and degradation rates of MRPs in GTPBP8 KO is necessary to conclusively resolve this inconsistency in the observed amounts of MRPs.

In bacteria, assembly intermediates of large ribosomal subunit accumulate in the absence of the homolog of GTPBP8 (Schaefer et al., 2006). As discussed previously, we did not detect assembly intermediates in the GTPBP8 KO, but observed an accumulation of both mtSSU and mtLSU, accompanied by a possible decrease in monosome formation (Figure 5.19). A block in the late stage of the mitoribosome assembly pathway could lead to accumulation of individual mitoribosomal subunits (Rebelo-Guiomar et al., 2021). Increased levels of both mtSSU and mtLSU accompanied by unstable monosome were observed previously in mutants of MTERF4, GTPBP6 and NSUN4 (Cámara et al., 2011; Lavdovskaia et al., 2020; Metodiev et al., 2014, respectively). MTERF4 is a mitoribosome assembly factor that forms a heterodimer with NSUN4 which binds to the large mitoribosomal subunit (Cámara et al., 2011). Inactivation of either MTERF4 or NSUN4 caused dysfunction of mitochondrial translation due to impaired monosome formation (Cámara et al., 2011; Metodiev et al., 2014). We detected NSUN4 in one of the replicates of qDGMS, where NSUN4 comigrated with the mtLSU and was downregulated in GTPBP8 KO compared to the wild type HEK293 cells (Figure 5.20A, B). This decrease of NSUN4 might be a secondary effect of GTPBP8 KO, further exacerbating the observed mitochondrial translation deficiency. MTERF4 interacts mainly with 16S rRNA but also showed binding to 12S rRNA in RNA immunoprecipitation experiment (Cámara et al., 2011). Increased biogenesis of the subunits of mitochondrial ribosome in MTERF4

knockout mouse hearts was supported by the observation of increased levels of TFB1M (Cámara et al., 2011). TFB1M is a methyltransferase that deposits m_2^6A(936/937) on 12S rRNA and this modification is important for stability of mtSSU (Metodiev et al., 2009). Similarly, TFB1M was upregulated also in GTPBP8 KO cells (Figure 5.20). GTPBP6 is another mitoribosome biogenesis factor which knockout leads to decrease in 55S monosome and increase in 28S mtSSU and 39S mtLSU (Lavdovskaia et al., 2020). Protein factors that associate with mtLSU in late stages of the assembly, MALSU1, GTPBP7 and GTPBP10 were upregulated in both GTPBP6 knockout (Lavdovskaia et al., 2020) and it GTPBP8 knockout presented here (Figure 5.20).

Taken together, our data indicate that GTPBP8 is a mitochondrial protein with a role in the biogenesis of mitochondrial ribosome. As our data are consistent with findings from cell lines/organisms where late assembly factors were inactivated, it is likely GTPBP8 is involved in the very late stage of the subunit assembly or joining of mtSSU and mtLSU into a monosome. At present, it is not clear what is the molecular target of GTPBP8. Identification of GTPBP8 interaction partners, either via immunoprecipitation or BioID-type of experiment, should provide insight into the exact molecular function of GTPBP8.

6 Conclusions

This work describes the development and the use of molecular biology and bioinformatics tools for study of mitochondrial gene expression, with a focus on translation and mitoribosome biogenesis. It expands our knowledge of function of two proteins, MRM1 and GTPBP8, proposed to be involved in the mitoribosome biogenesis by two distinct mechanisms, as a methyltransferase and as an assembly factor, respectively.

Function of genes is often investigated by their inactivation and observation of defects it causes in cellular phenotypes. However, for some genes such inactivation does not influence cell fitness. This was the case also for MRM1 knockout cell line (D'Souza, 2018). In order to characterise its importance for mitochondrial function, we therefore subjected this cell line to a pooled CRIPSR/Cas9 screen to look for potential genetic interactions of MRM1. The screen was performed with a custom gRNA library, designed and cloned in house, targeting genes with known or predicted function in mitochondrial gene expression. This approach led to identification of several candidate genetic interactions, with MRM2 being the most promising hit. Genetic interaction between MRM1 and MRM2 was verified by production of a double knockout cell line. Simultaneous inactivation of MRM1 and MRM2 genes led most often to a phenotype that was more severe compared to inactivation of MRM2 on its own. The unique mitochondrial defect observed only in MRM1/2 KO was a strong dependency on uridine supplementation. This indicates a severe defect in mitochondrial OXPHOS function in the double knockout. Since mitochondrial translation is impaired in MRM2 KO cells (Rebelo-Guiomar et al., 2021), mtDNA transcript levels were not reduced in MRM1/2 KO cells, but no or very low levels of mtDNA encoded proteins were detected in MRM1/2 KO, we concluded that the OXPHOS dysfunction is caused by impaired mitochondrial translation.

Mitochondrial ribosome is a central player in the mitochondrial translation and therefore we applied our recently developed qDGMS method to study its quantity and assembly status in MRM1/2 KO cells. As a proteomics-based method, each qDGMS experiment produces a large amount of data. Analysis of such data often presents a burden that limits the implementation of the method by wider scientific community. In order to make the method more accessible as well as to streamline analysis of our future experiments, we developed a ComPrAn R package. ComPrAn integrates most of the analysis steps and allows for interactive visualisation of results. With this method, we

analysed mitoribosomes from MRM1/2 KO and discovered an accumulation of mtLSU and disruption of mtSSU in this cell line. This was an unexpected finding, as both MRM1 and MRM2 modify 16S rRNA of the mtLSU, and it further highlights the complexity of regulation of mitoribosome biogenesis.

Furthermore, we characterised a previously poorly studied putative GTPase, GTPBP8. We showed that GTPBP8 colocalizes with mitochondria, it is important for mitochondrial translation and its inactivation leads to OXPHOS dysfunction. The qDGMS method provided valuable insight into the effects of inactivation of GTPBP8, as we identified accumulation of both mitoribosomal subunits in the knockout cell line. This result combined with the observed upregulation of late assembly factor GTPBP10, anti-association module proteins (MALSU1, L0R8F8 and mtACP), and translation initiation factors (mtIF2, mtIF3), narrowed down the probable window of GTPBP8 action to late stage of mitoribosome subunit assembly, subunit joining or translation initiation.

In conclusion, this work contributes to the efforts to understand the process of mitoribosome biogenesis and mitochondrial translation by characterising protein factors that are important for proper mitochondrial function. Still, our knowledge about these processes is far from complete. Assembly of the mitoribosome and mitochondrial translation diverged significantly from their bacterial and cytosolic counterparts. The process and regulation of mitoribosome assembly and translation also differ among eukaryotic organism, further complicating their study. Mitochondrial translation is unique as it combines components of two distinct genetic origins, and how the expression from nuclear and mtDNA genome are coordinated is currently unknown. Understanding the details of mitochondrial function and the molecular processes that govern it is a first step in deciphering the causes of mitochondrial dysfunction and moving towards developing therapies for mitochondrial diseases.

7 References

Agrotis, A., and Ketteler, R. (2015). A new age in functional genomics using CRISPR/Cas9 in arrayed library screening. Front. Genet. *6*.

Aibara, S., Singh, V., Modelska, A., and Amunts, A. (2020). Structural basis of mitochondrial translation. ELife *9*, e58362.

Albaum, S.P., Hahne, H., Otto, A., Haußmann, U., Becher, D., Poetsch, A., Goesmann, A., and Nattkemper, T.W. (2011). A guide through the computational analysis of isotope-labeled mass spectrometry-based quantitative proteomics data: an application study. Proteome Sci. *9*, 30.

Alexandrov, A., Chernyakov, I., Gu, W., Hiley, S.L., Hughes, T.R., Grayhack, E.J., and Phizicky, E.M. (2006). Rapid tRNA Decay Can Result from Lack of Nonessential Modifications. Mol. Cell *21*, 87–96.

Allen, F., Behan, F., Khodak, A., Iorio, F., Yusa, K., Garnett, M., and Parts, L. (2019). JACKS: joint analysis of CRISPR/Cas9 knockout screens. Genome Res. *29*, 464–471.

Amunts, A., Brown, A., Bai, X., Llácer, J.L., Hussain, T., Emsley, P., Long, F., Murshudov, G., Scheres, S.H.W., and Ramakrishnan, V. (2014). Structure of the Yeast Mitochondrial Large Ribosomal Subunit. Science *343*, 1485–1489.

Amunts, A., Brown, A., Toots, J., Scheres, S.H.W., and Ramakrishnan, V. (2015). The structure of the human mitochondrial ribosome. Science *348*, 95–98.

Anger, A.M., Armache, J.-P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D.N., and Beckmann, R. (2013). Structures of the human and Drosophila 80S ribosome. Nature *497*, 80–85.

Antonicka, H., and Shoubridge, E.A. (2015). Mitochondrial RNA Granules Are Centers for Posttranscriptional RNA Processing and Ribosome Biogenesis. Cell Rep. *10*, 920–932.

Antonicka, H., Choquet, K., Lin, Z., Gingras, A., Kleinman, C.L., and Shoubridge, E.A. (2017). A pseudouridine synthase module is essential for mitochondrial protein synthesis and cell viability. EMBO Rep. *18*, 28–38.

Arroyo, J.D., Jourdain, A.A., Calvo, S.E., Ballarano, C.A., Doench, J.G., Root, D.E., and Mootha, V.K. (2016). A Genome-wide CRISPR Death Screen Identifies Genes Essential for Oxidative Phosphorylation. Cell Metab. *24*, 875–885.

Barrientos, A., Korr, D., Barwell, K.J., Sjulsen, C., Gajewski, C.D., Manfredi, G., Ackerman, S., and Tzagoloff, A. (2003). MTG1 Codes for a Conserved Protein Required for Mitochondrial Translation. Mol. Biol. Cell *14*, 2292–2302.

Bar-Yaacov, D., Frumkin, I., Yashiro, Y., Chujo, T., Ishigami, Y., Chemla, Y., Blumberg, A., Schlesinger, O., Bieri, P., Greber, B., et al. (2016). Mitochondrial 16S rRNA Is

Methylated by tRNA Methyltransferase TRMT61B in All Vertebrates. PLOS Biol. *14*, e1002557.

Bayona-Bafaluy, M.P., Sánchez-Cabo, F., Fernández-Silva, P., Pérez-Martos, A., and Enríquez, J.A. (2011). A genome-wide shRNA screen for new OxPhos related genes. Mitochondrion *11*, 467–475.

Beck, M., Schmidt, A., Malmstroem, J., Claassen, M., Ori, A., Szymborska, A., Herzog, F., Rinner, O., Ellenberg, J., and Aebersold, R. (2011). The quantitative proteome of a human cell line. Mol. Syst. Biol. *7*, 549.

Berger, S., Cabrera-Orefice, A., Jetten, M.S.M., Brandt, U., and Welte, C.U. (2021). Investigation of central energy metabolism-related protein complexes of ANME-2d methanotrophic archaea by complexome profiling. Biochim. Biophys. Acta BBA - Bioenerg. *1862*, 148308.

Bhatta, A., Dienemann, C., Cramer, P., and Hillen, H.S. (2021). Structural basis of RNA processing by human mitochondrial RNase P. Nat. Struct. Mol. Biol. *28*, 1–11.

Birsoy, K., Wang, T., Chen, W.W., Freinkman, E., Abu-Remaileh, M., and Sabatini, D.M. (2015). An Essential Role of the Mitochondrial Electron Transport Chain in Cell Proliferation Is to Enable Aspartate Synthesis. Cell *162*, 540–551.

Boccaletto, P., Machnicka, M.A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T.K., de Crécy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A., et al. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res. *46*, D303–D307.

Bock, F.J., and Tait, S.W.G. (2020). Mitochondria as multifaceted regulators of cell death. Nat. Rev. Mol. Cell Biol. *21*, 85–100.

Bogenhagen, D.F., Martin, D.W., and Koller, A. (2014). Initial Steps in RNA Processing and Ribosome Assembly Occur at Mitochondrial DNA Nucleoids. Cell Metab. *19*, 618–629.

Bogenhagen, D.F., Ostermeyer-Fay, A.G., Haley, J.D., and Garcia-Diaz, M. (2018). Kinetics and Mechanism of Mammalian Mitochondrial Ribosome Assembly. Cell Rep. *22*, 1935–1944.

Bohnsack, M.T., and Sloan, K.E. (2018). The mitochondrial epitranscriptome: the roles of RNA modifications in mitochondrial translation and human disease. Cell. Mol. Life Sci. *75*, 241–260.

Bourne, H.R., Sanders, D.A., and McCormick, F. (1991). The GTPase superfamily: conserved structure and molecular mechanism. Nature *349*, 117–127.

Bowmaker, M., Yang, M.Y., Yasukawa, T., Reyes, A., Jacobs, H.T., Huberman, J.A., and Holt, I.J. (2003). Mammalian Mitochondrial DNA Replicates Bidirectionally from an Initiation Zone*. J. Biol. Chem. *278*, 50961–50969.

Braymer, J.J., and Lill, R. (2017). Iron–sulfur cluster biogenesis and trafficking in mitochondria. J. Biol. Chem. *292*, 12754–12763.

Britton, R.A. (2009). Role of GTPases in Bacterial Ribosome Assembly. Annu. Rev. Microbiol. *63*, 155–176.

Brown, A., Amunts, A., Bai, X., Sugimoto, Y., Edwards, P.C., Murshudov, G., Scheres, S.H.W., and Ramakrishnan, V. (2014). Structure of the large ribosomal subunit from human mitochondria. Science *346*, 718–722.

Brown, A., Rathore, S., Kimanius, D., Aibara, S., Bai, X., Rorbach, J., Amunts, A., and Ramakrishnan, V. (2017). Structures of the human mitochondrial ribosome in native states of assembly. Nat. Struct. Mol. Biol. *24*, 866–869.

Brzezniak, L.K., Bijata, M., Szczesny, R.J., and Stepien, P.P. (2011). Involvement of human ELAC2 gene product in 3' end processing of mitochondrial tRNAs. RNA Biol. *8*, 616–626.

Bulkley, D., Innis, C.A., Blaha, G., and Steitz, T.A. (2010). Revisiting the structures of several antibiotics bound to the bacterial ribosome. Proc. Natl. Acad. Sci. *107*, 17158–17163.

Burger, G., Gray, M.W., Forget, L., and Lang, B.F. (2013). Strikingly Bacteria-Like and Gene-Rich Mitochondrial Genomes throughout Jakobid Protists. Genome Biol. Evol. *5*, 418–438.

Calvo, S.E., Clauser, K.R., and Mootha, V.K. (2016). MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res. *44*, D1251–D1257.

Cámara, Y., Asin-Cayuela, J., Park, C.B., Metodiev, M.D., Shi, Y., Ruzzenente, B., Kukat, C., Habermann, B., Wibom, R., Hultenby, K., et al. (2011). MTERF4 Regulates Translation by Targeting the Methyltransferase NSUN4 to the Mammalian Mitochondrial Ribosome. Cell Metab. *13*, 527–539.

Chang, D.D., and Clayton, D.A. (1984). Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. Cell *36*, 635–643.

Chang, W., Cheng, J., Allaire, J.J., Xie, Y., and McPherson, J. (2020). shiny: Web Application Framework for R. R package version 1.5.0. https://CRAN.R-project.org/package=shiny.

Chen, H., Shi, Z., Guo, J., Chang, K., Chen, Q., Yao, C.-H., Haigis, M.C., and Shi, Y. (2020). The human mitochondrial 12S rRNA m4C methyltransferase METTL15 is required for mitochondrial function. J. Biol. Chem. *295*, 8505–8513.

Cheng, J., Berninghausen, O., and Beckmann, R. (2021). A distinct assembly pathway of the human 39S late pre-mitoribosome. Nat. Commun. *12*, 4544.

Chiang, T.-W.W., le Sage, C., Larrieu, D., Demir, M., and Jackson, S.P. (2016). CRISPR-Cas9 D10A nickase-based genotypic and phenotypic screening to enhance genome editing. Sci. Rep. *6*, 24356.

Chujo, T., and Suzuki, T. (2012). Trmt61B is a methyltransferase responsible for 1-methyladenosine at position 58 of human mitochondrial tRNAs. RNA *18*, 2269–2276.

Cipullo, M., Gesé, G.V., Khawaja, A., Hällberg, B.M., and Rorbach, J. (2021a). Structural basis for late maturation steps of the human mitoribosomal large subunit. Nat. Commun. *12*, 3673.

Cipullo, M., Pearce, S.F., Lopez Sanchez, I.G., Gopalakrishna, S., Krüger, A., Schober, F., Busch, J.D., Li, X., Wredenberg, A., Atanassov, I., et al. (2021b). Human GTPBP5 is involved in the late stage of mitoribosome large subunit assembly. Nucleic Acids Res. *49*, 354–370.

Cooper, E.L., García-Lara, J., and Foster, S.J. (2009). YsxC, an essential protein in Staphylococcus aureus crucial for ribosome assembly/stability. BMC Microbiol. *9*, 266.

Corneo, G., Zardi, L., and Polli, E. (1968). Human mitochondrial DNA. J. Mol. Biol. *36*, 419–423.

Čunátová, K., Reguera, D.P., Vrbacký, M., Fernández-Vizarra, E., Ding, S., Fearnley, I.M., Zeviani, M., Houštěk, J., Mráček, T., and Pecina, P. (2021). Loss of COX4I1 Leads to Combined Respiratory Chain Deficiency and Impaired Mitochondrial Protein Synthesis. Cells *10*, 369.

Dalla Rosa, I., Durigon, R., Pearce, S.F., Rorbach, J., Hirst, E.M.A., Vidoni, S., Reyes, A., Brea-Calvo, G., Minczuk, M., Woellhaf, M.W., et al. (2014). MPV17L2 is required for ribosome assembly in mitochondria. Nucleic Acids Res. *42*, 8500–8515.

Dassain, M., Leroy, A., Colosetti, L., Carolé, S., and Bouché, J.-P. (1999). A new essential gene of the `minimal genome' affecting cell division. Biochimie *81*, 889–895.

De Silva, D., Tu, Y.-T., Amunts, A., Fontanesi, F., and Barrientos, A. (2015). Mitochondrial ribosome assembly in health and disease. Cell Cycle *14*, 2226–2250.

Dempster, J.M., Pacini, C., Pantel, S., Behan, F.M., Green, T., Krill-Burger, J., Beaver, C.M., Younger, S.T., Zhivich, V., Najgebauer, H., et al. (2019). Agreement between two large pan-cancer CRISPR-Cas9 gene dependency data sets. Nat. Commun. *10*, 5817.

Dennerlein, S., and Rehling, P. (2015). Human mitochondrial COX1 assembly into cytochrome c oxidase at a glance. J. Cell Sci. *128*, 833–837.

Dennerlein, S., Rozanska, A., Wydro, M., Chrzanowska-Lightowlers, Z.M.A., and Lightowlers, R.N. (2010). Human ERAL1 is a mitochondrial RNA chaperone involved in the assembly of the 28S small mitochondrial ribosomal subunit. Biochem. J. *430*, 551–558.

Desai, N., Brown, A., Amunts, A., and Ramakrishnan, V. (2017). The structure of the yeast mitochondrial ribosome. Science *355*, 528–531.

Doench, J.G. (2018). Am I ready for CRISPR? A user's guide to genetic screens. Nat. Rev. Genet. *19*, 67–80.

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol. *34*, 184–191.

Donaldson, J.G., and Williams, D.B. (2009). Intracellular Assembly and Trafficking of MHC Class I Molecules. Traffic Cph. Den. *10*, 1745–1752.

D'Souza, A.R. (2018). Protein factors involved in the biogenesis of the mitochondrial ribosome. Thesis. University of Cambridge.

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

D'Souza, A.R., Van Haute, L., Powell, C.A., Mutti, C.D., Páleníková, P., Rebelo-Guiomar, P., Rorbach, J., and Minczuk, M. (2021). YbeY is required for ribosome small subunit assembly and tRNA processing in human mitochondria. Nucleic Acids Res. *49*, 5798–5812.

Dunkle, J.A., Wang, L., Feldman, M.B., Pulk, A., Chen, V.B., Kapral, G.J., Noeske, J., Richardson, J.S., Blanchard, S.C., and Cate, J.H.D. (2011). Structures of the Bacterial Ribosome in Classical and Hybrid States of tRNA Binding. Science *332*, 981–984.

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting Subcellular Localization of Proteins Based on their N-terminal Amino Acid Sequence. J. Mol. Biol. *300*, 1005–1016.

Englmeier, R., Pfeffer, S., and Förster, F. (2017). Structure of the Human Mitochondrial Ribosome Studied In Situ by Cryoelectron Tomography. Structure *25*, 1574-1581.e2.

Evers, F., Cabrera-Orefice, A., Elurbe, D.M., Kea-te Lindert, M., Boltryk, S.D., Voss, T.S., Huynen, M.A., Brandt, U., and Kooij, T.W.A. (2021). Composition and stage dynamics of mitochondrial complexes in Plasmodium falciparum. Nat. Commun. *12*, 3820.

Falkenberg, M. (2018). Mitochondrial DNA replication in mammalian cells: overview of the pathway. Essays Biochem. *62*, 287–296.

Fernández-Vizarra, E., López-Calcerrada, S., Formosa, L.E., Pérez-Pérez, R., Ding, S., Fearnley, I.M., Arenas, J., Martín, M.A., Zeviani, M., Ryan, M.T., et al. (2021). SILAC-based complexome profiling dissects the structural organization of the human respiratory supercomplexes in SCAFIKO cells. Biochim. Biophys. Acta BBA - Bioenerg. *1862*, 148414.

Flegontov, P., Michálek, J., Janouškovec, J., Lai, D.-H., Jirků, M., Hajdušková, E., Tomčala, A., Otto, T.D., Keeling, P.J., Pain, A., et al. (2015). Divergent mitochondrial respiratory chains in phototrophic relatives of apicomplexan parasites. Mol. Biol. Evol. *32*, 1115–1131.

Fukasawa, Y., Tsuji, J., Fu, S.-C., Tomii, K., Horton, P., and Imai, K. (2015). MitoFates: Improved Prediction of Mitochondrial Targeting Sequences and Their Cleavage Sites. Mol. Cell. Proteomics MCP *14*, 1113–1126.

Fuss, J.O., Tsai, C.-L., Ishida, J.P., and Tainer, J.A. (2015). Emerging critical roles of Fe–S clusters in DNA replication and repair. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1853*, 1253–1271.

Fusté, J.M., Shi, Y., Wanrooij, S., Zhu, X., Jemt, E., Persson, Ö., Sabouri, N., Gustafsson, C.M., and Falkenberg, M. (2014). In Vivo Occupancy of Mitochondrial Single-Stranded DNA Binding Protein Supports the Strand Displacement Mode of DNA Replication. PLOS Genet. *10*, e1004832.

Gabaldón, T., and Huynen, M.A. (2007). From Endosymbiont to Host-Controlled Organelle: The Hijacking of Mitochondrial Protein Synthesis and Metabolism. PLOS Comput. Biol. *3*, e219.

Gagliardi, D., Stepien, P.P., Temperley, R.J., Lightowlers, R.N., and Chrzanowska-Lightowlers, Z.M.A. (2004). Messenger RNA stability in mitochondria: different means to an end. Trends Genet. *20*, 260–267.

Garone, C., D'Souza, A.R., Dallabona, C., Lodi, T., Rebelo-Guiomar, P., Rorbach, J., Donati, M.A., Procopio, E., Montomoli, M., Guerrini, R., et al. (2017). Defective mitochondrial rRNA methyltransferase MRM2 causes MELAS-like clinical syndrome. Hum. Mol. Genet. *26*, 4257–4266.

Gaur, R., Grasso, D., Datta, P.P., Krishna, P.D.V., Das, G., Spencer, A., Agrawal, R.K., Spremulli, L., and Varshney, U. (2008). A Single Mammalian Mitochondrial Translation Initiation Factor Functionally Replaces Two Bacterial Factors. Mol. Cell *29*, 180–190.

Gianfrancesco, F., Sanges, R., Esposito, T., Tempesta, S., Rao, E., Rappold, G., Archidiacono, N., Graves, J.A.M., Forabosco, A., and D'Urso, M. (2001). Differential Divergence of Three Human Pseudoautosomal Genes and Their Mouse Homologs: Implications for Sex Chromosome Evolution. Genome Res. *11*, 2095–2100.

Giese, H., Ackermann, J., Heide, H., Bleier, L., Dröse, S., Wittig, I., Brandt, U., and Koch, I. (2015). NOVA: a software to analyze complexome profiling data. Bioinformatics *31*, 440–441.

Giese, H., Meisterknecht, J., Heidler, J., and Wittig, I. (2021). Mitochondrial Complexome Profiling. In Mitochondrial Gene Expression: Methods and Protocols, M. Minczuk, and J. Rorbach, eds. (New York, NY: Springer US), pp. 269–285.

Giorgi, C., Marchi, S., and Pinton, P. (2018). The machineries, regulation and cellular functions of mitochondrial calcium. Nat. Rev. Mol. Cell Biol. *19*, 713–730.

Gonçalves, E., Thomas, M., Behan, F.M., Picco, G., Pacini, C., Allen, F., Vinceti, A., Sharma, M., Jackson, D.A., Price, S., et al. (2021). Minimal genome-wide human CRISPR-Cas9 library. Genome Biol. *22*, 40.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL–EBI. Nucleic Acids Res. *38*, W695–W699.

Gozzelino, R., and Soares, M.P. (2014). Coupling Heme and Iron Metabolism via Ferritin H Chain. Antioxid. Redox Signal. *20*, 1754–1769.

Greber, B.J., Bieri, P., Leibundgut, M., Leitner, A., Aebersold, R., Boehringer, D., and Ban, N. (2015). The complete structure of the 55S mammalian mitochondrial ribosome. Science *348*, 303–308.

Grégoire, M., Morais, R., Quilliam, M.A., and Gravel, D. (1984). On auxotrophy for pyrimidines of respiration-deficient chick embryo cells. Eur. J. Biochem. *142*, 49–55.

Grevet, J.D., Lan, X., Hamagami, N., Edwards, C.R., Sankaranarayanan, L., Ji, X., Bhardwaj, S.K., Face, C.J., Posocco, D.F., Abdulmalik, O., et al. (2018). Domain-focused CRISPR screen identifies HRI as a fetal hemoglobin regulator in human erythroid cells. Science *361*, 285–290.

Guerrero-Castillo, S., Baertling, F., Kownatzki, D., Wessels, H.J., Arnold, S., Brandt, U., and Nijtmans, L. (2017). The Assembly Pathway of Mitochondrial Respiratory Chain Complex I. Cell Metab. *25*, 128–139.

Guerrero-Castillo, S., Krisp, C., Küchler, K., Arnold, S., Schlüter, H., and Gersting, S.W. (2021). Multiplexed complexome profiling using tandem mass tags. Biochim. Biophys. Acta BBA - Bioenerg. *1862*, 148448.

Gustavsson, M., and Ronne, H. (2008). Evidence that tRNA modifying enzymes are important in vivo targets for 5-fluorouracil in yeast. RNA *14*, 666–674.

Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat. Biotechnol. *17*, 994–999.

Han, K., Jeng, E.E., Hess, G.T., Morgens, D.W., Li, A., and Bassik, M.C. (2017). Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. Nat. Biotechnol. *35*, 463–474.

Hart, T., Chandrashekhar, M., Aregger, M., Steinhart, Z., Brown, K.R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S., et al. (2015). High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell *163*, 1515–1526.

Havugimana, P.C., Hart, G.T., Nepusz, T., Yang, H., Turinsky, A.L., Li, Z., Wang, P.I., Boutz, D.R., Fong, V., Phanse, S., et al. (2012). A Census of Human Soluble Protein Complexes. Cell *150*, 1068–1081.

He, J., Cooper, H.M., Reyes, A., Di Re, M., Kazak, L., Wood, S.R., Mao, C.C., Fearnley, I.M., Walker, J.E., and Holt, I.J. (2012). Human C4orf14 interacts with the mitochondrial nucleoid and is involved in the biogenesis of the small mitochondrial ribosomal subunit. Nucleic Acids Res. *40*, 6097–6108.

He, J., Ford, H.C., Carroll, J., Douglas, C., Gonzales, E., Ding, S., Fearnley, I.M., and Walker, J.E. (2018). Assembly of the membrane domain of ATP synthase in human mitochondria. Proc. Natl. Acad. Sci. *115*, 2988–2993.

Heide, H., Bleier, L., Steger, M., Ackermann, J., Dröse, S., Schwamb, B., Zörnig, M., Reichert, A.S., Koch, I., Wittig, I., et al. (2012). Complexome Profiling Identifies TMEM126B as a Component of the Mitochondrial Complex I Assembly Complex. Cell Metab. *16*, 538–549.

Hensen, F., Potter, A., van Esveld, S.L., Tarrés-Solé, A., Chakraborty, A., Solà, M., and Spelbrink, J.N. (2019). Mitochondrial RNA granules are critically dependent on mtDNA replication factors Twinkle and mtSSB. Nucleic Acids Res. *47*, 3680–3698.

Hilander, T., Jackson, C.B., Robciuc, M., Bashir, T., and Zhao, H. (2021). The Roles of Assembly Factors in Mammalian Mitoribosome Biogenesis. Mitochondrion.

Hillen, H.S., Morozov, Y.I., Sarfallah, A., Temiakov, D., and Cramer, P. (2017). Structural Basis of Mitochondrial Transcription Initiation. Cell *171*, 1072-1081.e10.

Hillen, H.S., Lavdovskaia, E., Nadler, F., Hanitsch, E., Linden, A., Bohnsack, K.E., Urlaub, H., and Richter-Dennerlein, R. (2021). Structural basis of GTPase-mediated mitochondrial ribosome biogenesis and recycling. Nat. Commun. *12*, 3672.

Hirano, Y., Ohniwa, R.L., Wada, C., Yoshimura, S.H., and Takeyasu, K. (2006). Human small G proteins, ObgH1, and ObgH2, participate in the maintenance of mitochondria and nucleolar architectures. Genes Cells *11*, 1295–1304.

Holt, I.J., Lorimer, H.E., and Jacobs, H.T. (2000). Coupled Leading- and Lagging-Strand Synthesis of Mammalian Mitochondrial DNA. Cell *100*, 515–524.

Holzmann, J., Frank, P., Löffler, E., Bennett, K.L., Gerner, C., and Rossmanith, W. (2008). RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell *135*, 462–474.

Horlbeck, M.A., Xu, A., Wang, M., Bennett, N.K., Park, C.Y., Bogdanoff, D., Adamson, B., Chow, E.D., Kampmann, M., Peterson, T.R., et al. (2018). Mapping the Genetic Landscape of Human Cells. Cell *174*, 953-967.e22.

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P.S., Pagni, M., and Sigrist, C.J.A. (2006). The PROSITE database. Nucleic Acids Res. *34*, D227–D230.

Itoh, Y., Andréll, J., Choi, A., Richter, U., Maiti, P., Best, R.B., Barrientos, A., Battersby, B.J., and Amunts, A. (2021). Mechanism of membrane-tethered mitochondrial protein synthesis. Science *371*, 846–849.

Jaskolowski, M., Ramrath, D.J.F., Bieri, P., Niemann, M., Mattei, S., Calderaro, S., Leibundgut, M., Horn, E.K., Boehringer, D., Schneider, A., et al. (2020). Structural Insights into the Mechanism of Mitoribosomal Large Subunit Biogenesis. Mol. Cell *79*, 629-644.e4.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science *337*, 816–821.

Jourdain, A.A., Boehm, E., Maundrell, K., and Martinou, J.-C. (2016). Mitochondrial RNA granules: Compartmentalizing mitochondrial gene expression. J. Cell Biol. *212*, 611–614.

Jourdain, A.A., Popow, J., de la Fuente, M.A., Martinou, J.-C., Anderson, P., and Simarro, M. (2017). The FASTK family of proteins: emerging regulators of mitochondrial RNA biology. Nucleic Acids Res. *45*, 10941–10947.

Karnkowska, A., Vacek, V., Zubáčová, Z., Treitli, S.C., Petrželková, R., Eme, L., Novák, L., Žárský, V., Barlow, L.D., Herman, E.K., et al. (2016). A Eukaryote without a Mitochondrial Organelle. Curr. Biol. *26*, 1274–1284.

Kehrein, K., Schilling, R., Möller-Hergt, B.V., Wurm, C.A., Jakobs, S., Lamkemeyer, T., Langer, T., and Ott, M. (2015). Organization of Mitochondrial Gene Expression in Two Distinct Ribosome-Containing Assemblies. Cell Rep. *10*, 843–853.

Khawaja, A., Itoh, Y., Remes, C., Spåhr, H., Yukhnovets, O., Höfig, H., Amunts, A., and Rorbach, J. (2020). Distinct pre-initiation steps in human mitochondrial translation. Nat. Commun. *11*, 2932.

Kim, H.-J., and Barrientos, A. (2018). MTG1 couples mitoribosome large subunit assembly with intersubunit bridge formation. Nucleic Acids Res. *46*, 8435–8453.

Kim, E., Kim, S., Kim, D.H., Choi, B.-S., Choi, I.-Y., and Kim, J.-S. (2012a). Precision genome engineering with programmable DNA-nicking enzymes. Genome Res. *22*, 1327–1333.

Kim, H.J., Khalimonchuk, O., Smith, P.M., and Winge, D.R. (2012b). Structure, function, and assembly of heme centers in mitochondrial respiratory complexes. Biochim. Biophys. Acta *1823*, 1604–1616.

Koeck, T., Olsson, A.H., Nitert, M.D., Sharoyko, V.V., Ladenvall, C., Kotova, O., Reiling, E., Rönn, T., Parikh, H., Taneera, J., et al. (2011). A common variant in TFB1M is associated with reduced insulin secretion and increased future risk of type 2 diabetes. Cell Metab. *13*, 80–91.

Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M.D.C., and Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat. Biotechnol. *32*, 267–273.

Kolanczyk, M., Pech, M., Zemojtel, T., Yamamoto, H., Mikula, I., Calvaruso, M.-A., van den Brand, M., Richter, R., Fischer, B., Ritz, A., et al. (2011). NOA1 is an essential GTPase required for mitochondrial protein synthesis. Mol. Biol. Cell *22*, 1–11.

Korhonen, J.A., Pham, X.H., Pellegrini, M., and Falkenberg, M. (2004). Reconstitution of a minimal mtDNA replisome in vitro. EMBO J. *23*, 2423–2429.

Kotani, T., Akabane, S., Takeyasu, K., Ueda, T., and Takeuchi, N. (2013). Human G-proteins, ObgH1 and Mtg1, associate with the large mitochondrial ribosome subunit and are involved in translation and assembly of respiratory complexes. Nucleic Acids Res. *41*, 3713–3722.

Kukat, C., Wurm, C.A., Spåhr, H., Falkenberg, M., Larsson, N.-G., and Jakobs, S. (2011). Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. Proc. Natl. Acad. Sci. U. S. A. *108*, 13534–13539.

Kummer, E., and Ban, N. (2021). Mechanisms and regulation of protein synthesis in mitochondria. Nat. Rev. Mol. Cell Biol. 1–19.

Kummer, E., Leibundgut, M., Rackham, O., Lee, R.G., Boehringer, D., Filipovska, A., and Ban, N. (2018). Unique features of mammalian mitochondrial translation initiation revealed by cryo-EM. Nature *560*, 263–267.

Laptev, I., Shvetsova, E., Levitskii, S., Serebryakova, M., Rubtsova, M., Bogdanov, A., Kamenski, P., Sergiev, P., and Dontsova, O. (2020a). Mouse Trmt2B protein is a dual specific mitochondrial metyltransferase responsible for m5U formation in both tRNA and rRNA. RNA Biol. *17*, 441–450.

Laptev, I., Shvetsova, E., Levitskii, S., Serebryakova, M., Rubtsova, M., Zgoda, V., Bogdanov, A., Kamenski, P., Sergiev, P., and Dontsova, O. (2020b). METTL15 interacts with the assembly intermediate of murine mitochondrial small ribosomal subunit to form m4C840 12S rRNA residue. Nucleic Acids Res. *48*, 8022–8034.

Larance, M., Kirkwood, K.J., Tinti, M., Murillo, A.B., Ferguson, M.A.J., and Lamond, A.I. (2016). Global Membrane Protein Interactome Analysis using In vivo Crosslinking and Mass Spectrometry-based Protein Correlation Profiling. Mol. Cell. Proteomics *15*, 2476–2490.

Lavdovskaia, E., Kolander, E., Steube, E., Mai, M.M.-Q., Urlaub, H., and Richter-Dennerlein, R. (2018). The human Obg protein GTPBP10 is involved in mitoribosomal biogenesis. Nucleic Acids Res. *46*, 8471–8482.

Lavdovskaia, E., Denks, K., Nadler, F., Steube, E., Linden, A., Urlaub, H., Rodnina, M.V., and Richter-Dennerlein, R. (2020). Dual function of GTPBP6 in biogenesis and recycling of human mitochondrial ribosomes. Nucleic Acids Res. *48*, 12929–12942.

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

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

Leipe, D.D., Wolf, Y.I., Koonin, E.V., and Aravind, L. (2002). Classification and evolution of P-loop GTPases and related ATPases11Edited by J. Thornton. J. Mol. Biol. *317*, 41–72.

Lemarie, A., and Grimm, S. (2009). Mutations in the heme b-binding residue of SDHC inhibit assembly of respiratory chain complex II in mammalian cells. Mitochondrion *9*, 254–260.

Lenarčič, T., Jaskolowski, M., Leibundgut, M., Scaiola, A., Schönhut, T., Saurer, M., Lee, R.G., Rackham, O., Filipovska, A., and Ban, N. (2021). Stepwise maturation of the peptidyl transferase region of human mitoribosomes. Nat. Commun. *12*, 3671.

Letts, J.A., and Sazanov, L.A. (2017). Clarifying the supercomplex: the higher-order organization of the mitochondrial electron transport chain. Nat. Struct. Mol. Biol. *24*, 800–808.

Lewis, M.T., Kasper, J.D., Bazil, J.N., Frisbee, J.C., and Wiseman, R.W. (2019). Quantification of Mitochondrial Oxidative Phosphorylation in Metabolic Disease: Application to Type 2 Diabetes. Int. J. Mol. Sci. *20*, 5271.

Li, X., Xiong, X., Zhang, M., Wang, K., Chen, Y., Zhou, J., Mao, Y., Lv, J., Yi, D., Chen, X.-W., et al. (2017). Base-Resolution Mapping Reveals Distinct m1A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts. Mol. Cell *68*, 993-1005.e9.

Lin, Y.-C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., et al. (2014). Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. Nat. Commun. *5*, 4767.

Litonin, D., Sologub, M., Shi, Y., Savkina, M., Anikin, M., Falkenberg, M., Gustafsson, C.M., and Temiakov, D. (2010). Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes in vitro. J. Biol. Chem. *285*, 18129–18133.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods *25*, 402–408.

Lobo-Jarne, T., Pérez-Pérez, R., Fontanesi, F., Timón-Gómez, A., Wittig, I., Peñas, A., Serrano-Lorenzo, P., García-Consuegra, I., Arenas, J., Martín, M.A., et al. (2020). Multiple pathways coordinate assembly of human mitochondrial complex IV and stabilization of respiratory supercomplexes. EMBO J. *39*, e103912.

Lopez Sanchez, M.I.G., Cipullo, M., Gopalakrishna, S., Khawaja, A., and Rorbach, J. (2020). Methylation of Ribosomal RNA: A Mitochondrial Perspective. Front. Genet. *11*, 761.

Lopez Sanchez, M.I.G., Krüger, A., Shiriaev, D.I., Liu, Y., and Rorbach, J. (2021). Human Mitoribosome Biogenesis and Its Emerging Links to Disease. Int. J. Mol. Sci. *22*, 3827.

Lövgren, J.M., and Wikström, P.M. (2001). The rlmB Gene Is Essential for Formation of Gm2251 in 23S rRNA but Not for Ribosome Maturation inEscherichia coli. J. Bacteriol. *183*, 6957–6960.

Maclean, A.E., Bridges, H.R., Silva, M.F., Ding, S., Ovciarikova, J., Hirst, J., and Sheiner, L. (2021). Complexome profile of Toxoplasma gondii mitochondria identifies divergent subunits of respiratory chain complexes including new subunits of cytochrome bc1 complex. PLOS Pathog. *17*, e1009301.

Maiti, P., Kim, H.-J., Tu, Y.-T., and Barrientos, A. (2018). Human GTPBP10 is required for mitoribosome maturation. Nucleic Acids Res. *46*, 11423–11437.

Maiti, P., Antonicka, H., Gingras, A.-C., Shoubridge, E.A., and Barrientos, A. (2020). Human GTPBP5 (MTG2) fuels mitoribosome large subunit maturation by facilitating 16S rRNA methylation. Nucleic Acids Res. *48*, 7924–7943.

Maiti, P., Lavdovskaia, E., Barrientos, A., and Richter-Dennerlein, R. (2021). Role of GTPases in Driving Mitoribosome Assembly. Trends Cell Biol.

Mense, S.M., and Zhang, L. (2006). Heme: a versatile signaling molecule controlling the activities of diverse regulators ranging from transcription factors to MAP kinases. Cell Res. *16*, 681–692.

Metodiev, M.D., Lesko, N., Park, C.B., Cámara, Y., Shi, Y., Wibom, R., Hultenby, K., Gustafsson, C.M., and Larsson, N.-G. (2009). Methylation of 12S rRNA Is Necessary for In Vivo Stability of the Small Subunit of the Mammalian Mitochondrial Ribosome. Cell Metab. *9*, 386–397.

Metodiev, M.D., Spåhr, H., Polosa, P.L., Meharg, C., Becker, C., Altmueller, J., Habermann, B., Larsson, N.-G., and Ruzzenente, B. (2014). NSUN4 Is a Dual Function Mitochondrial Protein Required for Both Methylation of 12S rRNA and Coordination of Mitoribosomal Assembly. PLOS Genet. *10*, e1004110.

Meunier, B., Dumas, E., Piec, I., Béchet, D., Hébraud, M., and Hocquette, J.-F. (2007). Assessment of Hierarchical Clustering Methodologies for Proteomic Data Mining. J. Proteome Res. *6*, 358–366.

Michel, G., Sauvé, V., Larocque, R., Li, Y., Matte, A., and Cygler, M. (2002). The Structure of the RlmB 23S rRNA Methyltransferase Reveals a New Methyltransferase Fold with a Unique Knot. Structure *10*, 1303–1315.

Minczuk, M., He, J., Duch, A.M., Ettema, T.J., Chlebowski, A., Dzionek, K., Nijtmans, L.G.J., Huynen, M.A., and Holt, I.J. (2011). TEFM (c17orf42) is necessary for transcription of human mtDNA. Nucleic Acids Res. *39*, 4284–4299.

Montoya, J., Christianson, T., Levens, D., Rabinowitz, M., and Attardi, G. (1982). Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. Proc. Natl. Acad. Sci. U. S. A. *79*, 7195–7199.

Montoya, J., Gaines, G.L., and Attardi, G. (1983). The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. Cell *34*, 151–159.

Morgens, D.W., Wainberg, M., Boyle, E.A., Ursu, O., Araya, C.L., Tsui, C.K., Haney, M.S., Hess, G.T., Han, K., Jeng, E.E., et al. (2017). Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. Nat. Commun. *8*, 15178.

Müller, C.S., Bildl, W., Haupt, A., Ellenrieder, L., Becker, T., Hunte, C., Fakler, B., and Schulte, U. (2016). Cryo-slicing Blue Native-Mass Spectrometry (csBN-MS), a Novel Technology for High Resolution Complexome Profiling. Mol. Cell. Proteomics *15*, 669–681.

Nagaike, T., Suzuki, T., Tomari, Y., Takemoto-Hori, C., Negayama, F., Watanabe, K., and Ueda, T. (2001). Identification and Characterization of Mammalian Mitochondrial tRNA nucleotidyltransferases*. J. Biol. Chem. *276*, 40041–40049.

Nagao, A., Suzuki, T., Katoh, T., Sakaguchi, Y., and Suzuki, T. (2009). Biogenesis of glutaminyl-mt tRNAGln in human mitochondria. Proc. Natl. Acad. Sci. U. S. A. *106*, 16209–16214.

Ni, X., Davis, J.H., Jain, N., Razi, A., Benlekbir, S., McArthur, A.G., Rubinstein, J.L., Britton, R.A., Williamson, J.R., and Ortega, J. (2016). YphC and YsxC GTPases assist the maturation of the central protuberance, GTPase associated region and functional core of the 50S ribosomal subunit. Nucleic Acids Res. *44*, 8442–8455.

Nicholls, T.J., and Minczuk, M. (2014). In D-loop: 40 years of mitochondrial 7S DNA. Exp. Gerontol. *56*, 175–181.

Nolte, H., and Langer, T. (2021). ComplexFinder: A software package for the analysis of native protein complex fractionation experiments. Biochim. Biophys. Acta BBA - Bioenerg. *1862*, 148444.

Ojala, D., Montoya, J., and Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. Nature *290*, 470–474.

Okamoto, M., Fujiwara, M., Hori, M., Okada, K., Yazama, F., Konishi, H., Xiao, Y., Qi, G., Shimamoto, F., Ota, T., et al. (2014). tRNA Modifying Enzymes, NSUN2 and METTL1, Determine Sensitivity to 5-Fluorouracil in HeLa Cells. PLoS Genet. *10*, e1004639.

Ong, S.-E., and Mann, M. (2007). Stable Isotope Labeling by Amino Acids in Cell Culture for Quantitative Proteomics. In Quantitative Proteomics by Mass Spectrometry, S. Sechi, ed. (Totowa, NJ: Humana Press), pp. 37–52.

Pacini, C., Dempster, J.M., Boyle, I., Gonçalves, E., Najgebauer, H., Karakoc, E., van der Meer, D., Barthorpe, A., Lightfoot, H., Jaaks, P., et al. (2021). Integrated cross-study datasets of genetic dependencies in cancer. Nat. Commun. *12*, 1661.

Pagliarini, D.J., Calvo, S.E., Chang, B., Sheth, S.A., Vafai, S.B., Ong, S.-E., Walford, G.A., Sugiana, C., Boneh, A., Chen, W.K., et al. (2008). A mitochondrial protein compendium elucidates complex I disease biology. Cell *134*, 112–123.

Páleníková, P., Harbour, M.E., Ding, S., Fearnley, I.M., Van Haute, L., Rorbach, J., Scavetta, R., Minczuk, M., and Rebelo-Guiomar, P. (2021a). Quantitative density gradient analysis by mass spectrometry (qDGMS) and complexome profiling analysis (ComPrAn) R package for the study of macromolecular complexes. Biochim. Biophys. Acta BBA - Bioenerg. *1862*, 148399.

Páleníková, P., Harbour, M.E., Prodi, F., Minczuk, M., Zeviani, M., Ghelli, A., and Fernández-Vizarra, E. (2021b). Duplexing complexome profiling with SILAC to study human respiratory chain assembly defects. Biochim. Biophys. Acta BBA - Bioenerg. *1862*, 148395.

Payne, B.A.I., Wilson, I.J., Yu-Wai-Man, P., Coxhead, J., Deehan, D., Horvath, R., Taylor, R.W., Samuels, D.C., Santibanez-Koref, M., and Chinnery, P.F. (2013). Universal heteroplasmy of human mitochondrial DNA. Hum. Mol. Genet. *22*, 384–390.

Pearce, S.F., Rorbach, J., Haute, L.V., D'Souza, A.R., Rebelo-Guiomar, P., Powell, C.A., Brierley, I., Firth, A.E., and Minczuk, M. (2017). Maturation of selected human mitochondrial tRNAs requires deadenylation. ELife *6*, e27596.

Perez-Martinez, X., Butler, C.A., Shingu-Vazquez, M., and Fox, T.D. (2009). Dual Functions of Mss51 Couple Synthesis of Cox1 to Assembly of Cytochrome c Oxidase in Saccharomyces cerevisiae Mitochondria. Mol. Biol. Cell *20*, 4371–4380.

Pham, X.H., Farge, G., Shi, Y., Gaspari, M., Gustafsson, C.M., and Falkenberg, M. (2006). Conserved Sequence Box II Directs Transcription Termination and Primer Formation in Mitochondria*. J. Biol. Chem. *281*, 24647–24652.

Piel, R.B., Dailey, H.A., and Medlock, A.E. (2019). The Mitochondrial Heme Metabolon: Insights into the Complex(ity) of Heme Synthesis and Distribution. Mol. Genet. Metab. *128*, 198–203.

Poole, A.M., and Gribaldo, S. (2014). Eukaryotic Origins: How and When Was the Mitochondrion Acquired? Cold Spring Harb. Perspect. Biol. *6*, a015990.

Posse, V., Shahzad, S., Falkenberg, M., Hällberg, B.M., and Gustafsson, C.M. (2015). TEFM is a potent stimulator of mitochondrial transcription elongation in vitro. Nucleic Acids Res. *43*, 2615–2624.

Potts, G.K., Voigt, E.A., Bailey, D.J., Rose, C.M., Westphall, M.S., Hebert, A.S., Yin, J., and Coon, J.J. (2016). Neucode Labels for Multiplexed, Absolute Protein Quantification. Anal. Chem. *88*, 3295–3303.

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

Powell, C.A., Nicholls, T.J., and Minczuk, M. (2015). Nuclear-encoded factors involved in post-transcriptional processing and modification of mitochondrial tRNAs in human disease. Front. Genet. *6*, 79.

Prágai, Z., and Harwood, C.R. (2000). YsxC, a Putative GTP-Binding Protein Essential for Growth of Bacillus subtilis 168. J. Bacteriol. *182*, 6819–6823.

Prolo, L.M., Li, A., Owen, S.F., Parker, J.J., Foshay, K., Nitta, R.T., Morgens, D.W., Bolin, S., Wilson, C.M., Vega L, J.C.M., et al. (2019). Targeted genomic CRISPR-Cas9 screen identifies MAP4K4 as essential for glioblastoma invasion. Sci. Rep. *9*, 14020.

Protasoni, M., Pérez-Pérez, R., Lobo-Jarne, T., Harbour, M.E., Ding, S., Peñas, A., Diaz, F., Moraes, C.T., Fearnley, I.M., Zeviani, M., et al. (2020). Respiratory supercomplexes act as a platform for complex III-mediated maturation of human mitochondrial complexes I and IV. EMBO J. *39*, e102817.

Rackham, O., Busch, J.D., Matic, S., Siira, S.J., Kuznetsova, I., Atanassov, I., Ermer, J.A., Shearwood, A.-M.J., Richman, T.R., Stewart, J.B., et al. (2016). Hierarchical RNA Processing Is Required for Mitochondrial Ribosome Assembly. Cell Rep. *16*, 1874–1890.

Ramirez, C.L., Certo, M.T., Mussolino, C., Goodwin, M.J., Cradick, T.J., McCaffrey, A.P., Cathomen, T., Scharenberg, A.M., and Joung, J.K. (2012). Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. Nucleic Acids Res. *40*, 5560–5568.

Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. (2013). Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell *154*, 1380–1389.

Rath, S., Sharma, R., Gupta, R., Ast, T., Chan, C., Durham, T.J., Goodman, R.P., Grabarek, Z., Haas, M.E., Hung, W.H.W., et al. (2021). MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. Nucleic Acids Res. *49*, D1541–D1547.

Rebelo Guiomar, P. (2020). The role of RNA modifications in mitochondrial translation. Thesis. University of Cambridge.

Rebelo-Guiomar, P., Powell, C.A., Van Haute, L., and Minczuk, M. (2018). The mammalian mitochondrial epitranscriptome. Biochim. Biophys. Acta BBA - Gene Regul. Mech.

Rebelo-Guiomar, P., Pellegrino, S., Dent, K.C., Sas-Chen, A., Miller-Fleming, L., Garone, C., Haute, L.V., Rogan, J.F., Dinan, A., Firth, A.E., et al. (2021). A late-stage assembly checkpoint of the human mitochondrial ribosome large subunit. BioRxiv 2021.06.09.447755.

Reinhard, L., Sridhara, S., and Hällberg, B.M. (2017). The MRPP1/MRPP2 complex is a tRNA-maturation platform in human mitochondria. Nucleic Acids Res. *45*, 12469–12480.

Reyes, A., Kazak, L., Wood, S.R., Yasukawa, T., Jacobs, H.T., and Holt, I.J. (2013). Mitochondrial DNA replication proceeds via a 'bootlace' mechanism involving the incorporation of processed transcripts. Nucleic Acids Res. *41*, 5837–5850.

Robinson, B.H., Petrova-Benedict, R., Buncic, J.R., and Wallace, D.C. (1992). Nonviability of cells with oxidative defects in galactose medium: a screening test for affected patient fibroblasts. Biochem. Med. Metab. Biol. *48*, 122–126.

Roger, A.J., Muñoz-Gómez, S.A., and Kamikawa, R. (2017). The Origin and Diversification of Mitochondria. Curr. Biol. *27*, R1177–R1192.

Röhricht, H., Schwartzmann, J., and Meyer, E.H. (2021). Complexome profiling reveals novel insights into the composition and assembly of the mitochondrial ATP synthase of Arabidopsis thaliana. Biochim. Biophys. Acta BBA - Bioenerg. *1862*, 148425.

Rorbach, J., Richter, R., Wessels, H.J., Wydro, M., Pekalski, M., Farhoud, M., Kühl, I., Gaisne, M., Bonnefoy, N., Smeitink, J.A., et al. (2008). The human mitochondrial ribosome recycling factor is essential for cell viability. Nucleic Acids Res. *36*, 5787–5799.

Rorbach, J., Nicholls, T.J.J., and Minczuk, M. (2011). PDE12 removes mitochondrial RNA poly(A) tails and controls translation in human mitochondria. Nucleic Acids Res. *39*, 7750–7763.

Rorbach, J., Boesch, P., Gammage, P.A., Nicholls, T.J.J., Pearce, S.F., Patel, D., Hauser, A., Perocchi, F., and Minczuk, M. (2014). MRM2 and MRM3 are involved in biogenesis of the large subunit of the mitochondrial ribosome. Mol. Biol. Cell *25*, 2542–2555.

Rorbach, J., Gao, F., Powell, C.A., D'Souza, A., Lightowlers, R.N., Minczuk, M., and Chrzanowska-Lightowlers, Z.M. (2016). Human mitochondrial ribosomes can switch their structural RNA composition. Proc. Natl. Acad. Sci. U. S. A. *113*, 12198–12201.

Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J. Cell Biol. *196*, 801–810.

Rugen, N., Straube, H., Franken, L.E., Braun, H.-P., and Eubel, H. (2019). Complexome profiling reveals association of PPR proteins with ribosomes in the mitochondria of plants. Mol. Cell. Proteomics.

Rugen, N., Schaarschmidt, F., Eirich, J., Finkemeier, I., Braun, H.-P., and Eubel, H. (2021). Protein interaction patterns in Arabidopsis thaliana leaf mitochondria change in dependence to light. Biochim. Biophys. Acta BBA - Bioenerg. *1862*, 148443.

Safra, M., Sas-Chen, A., Nir, R., Winkler, R., Nachshon, A., Bar-Yaacov, D., Erlacher, M., Rossmanith, W., Stern-Ginossar, N., and Schwartz, S. (2017). The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature *551*, 251–255.

Sagan, L. (1967). On the origin of mitosing cells. J. Theor. Biol. *14*, 225-IN6.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nat. Methods *11*, 783–784.

Sanson, K.R., Hanna, R.E., Hegde, M., Donovan, K.F., Strand, C., Sullender, M.E., Vaimberg, E.W., Goodale, A., Root, D.E., Piccioni, F., et al. (2018). Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. Nat. Commun. *9*, 5416.

Schaefer, L., Uicker, W.C., Wicker-Planquart, C., Foucher, A.-E., Jault, J.-M., and Britton, R.A. (2006). Multiple GTPases Participate in the Assembly of the Large Ribosomal Subunit in Bacillus subtilis. J. Bacteriol. *188*, 8252–8258.

Schägger, H., and Pfeiffer, K. (2000). Supercomplexes in the respiratory chains of yeast and mammalian mitochondria. EMBO J. *19*, 1777–1783.

Schwartzbach, C.J., and Spremulli, L.L. (1989). Bovine mitochondrial protein synthesis elongation factors: Identification and initial characterization of an elongation factor Tu-elongation factor Ts complex *. J. Biol. Chem. *264*, 19125–19131.

Seidel-Rogol, B.L., McCulloch, V., and Shadel, G.S. (2003). Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop. Nat. Genet. *33*, 23–24.

Senkler, J., Senkler, M., Eubel, H., Hildebrandt, T., Lengwenus, C., Schertl, P., Schwarzländer, M., Wagner, S., Wittig, I., and Braun, H.-P. (2017). The mitochondrial complexome of Arabidopsis thaliana. Plant J. *89*, 1079–1092.

Senkler, J., Rugen, N., Eubel, H., Hegermann, J., and Braun, H.-P. (2018). Absence of Complex I Implicates Rearrangement of the Respiratory Chain in European Mistletoe. Curr. Biol. *28*, 1606-1613.e4.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science *343*, 84–87.

Sharma, S., Dincer, C., Weidemüller, P., Wright, G.J., and Petsalaki, E. (2020). CEN-tools: an integrative platform to identify the contexts of essential genes. Mol. Syst. Biol. *16*, e9698.

Sharoyko, V.V., Abels, M., Sun, J., Nicholas, L.M., Mollet, I.G., Stamenkovic, J.A., Göhring, I., Malmgren, S., Storm, P., Fadista, J., et al. (2014). Loss of TFB1M results in mitochondrial dysfunction that leads to impaired insulin secretion and diabetes. Hum. Mol. Genet. *23*, 5733–5749.

Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X., et al. (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat. Methods *11*, 399–402.

Shen, J.P., Zhao, D., Sasik, R., Luebeck, J., Birmingham, A., Bojorquez-Gomez, A., Licon, K., Klepper, K., Pekin, D., Beckett, A.N., et al. (2017). Combinatorial CRISPR–Cas9 screens for de novo mapping of genetic interactions. Nat. Methods *14*, 573–576.

Shutt, T.E., and Gray, M.W. (2006). Bacteriophage origins of mitochondrial replication and transcription proteins. Trends Genet. *22*, 90–95.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

Signes, A., and Fernandez-Vizarra, E. (2018). Assembly of mammalian oxidative phosphorylation complexes I-V and supercomplexes. Essays Biochem. *62*, 255–270.

Sirum-Connolly, K., and Mason, T.L. (1993). Functional requirement of a site-specific ribose methylation in ribosomal RNA. Science *262*, 1886–1889.

Sirum-Connolly, K., Peltier, J.M., Crain, P.F., McCloskey, J.A., and Mason, T.L. (1995). Implications of a functional large ribosomal RNA with only three modified nucleotides. Biochimie *77*, 30–39.

Sissler, M., and Hashem, Y. (2021). Mitoribosome assembly comes into view. Nat. Struct. Mol. Biol. *28*, 631–633.

Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004). Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. PROTEOMICS *4*, 1581–1590.

Smith, A.C., and Robinson, A.J. (2019). MitoMiner v4.0: an updated database of mitochondrial localization evidence, phenotypes and diseases. Nucleic Acids Res. *47*, D1225–D1228.

Soleimanpour-Lichaei, H.R., Kühl, I., Gaisne, M., Passos, J.F., Wydro, M., Rorbach, J., Temperley, R., Bonnefoy, N., Tate, W., Lightowlers, R., et al. (2007). mtRF1a is a human mitochondrial translation release factor decoding the major termination codons UAA and UAG. Mol. Cell *27*, 745–757.

Spadafora, D., Kozhukhar, N., Chouljenko, V.N., Kousoulas, K.G., and Alexeyev, M.F. (2016). Methods for Efficient Elimination of Mitochondrial DNA from Cultured Cells. PLOS ONE *11*, e0154684.

Summer, S., Smirnova, A., Gabriele, A., Toth, U., Fasemore, A.M., Förstner, K.U., Kuhn, L., Chicher, J., Hammann, P., Mitulović, G., et al. (2020). YBEY is an essential biogenesis factor for mitochondrial ribosomes. Nucleic Acids Res. *48*, 9762–9786.

Suzuki, T. (2021). The expanding world of tRNA modifications and their disease relevance. Nat. Rev. Mol. Cell Biol. *22*, 375–392.

Suzuki, T., and Suzuki, T. (2014). A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs. Nucleic Acids Res. *42*, 7346–7357.

Suzuki, T., Suzuki, T., Wada, T., Saigo, K., and Watanabe, K. (2002). Taurine as a constituent of mitochondrial tRNAs: new insights into the functions of taurine and human mitochondrial diseases. EMBO J. *21*, 6581–6589.

Suzuki, T., Yashiro, Y., Kikuchi, I., Ishigami, Y., Saito, H., Matsuzawa, I., Okada, S., Mito, M., Iwasaki, S., Ma, D., et al. (2020). Complete chemical structures of human mitochondrial tRNAs. Nat. Commun. *11*, 4269.

Svetlov, M.S., Plessa, E., Chen, C.-W., Bougas, A., Krokidis, M.G., Dinos, G.P., and Polikanov, Y.S. (2019). High-resolution crystal structures of ribosome-bound

chloramphenicol and erythromycin provide the ultimate basis for their competition. RNA *25*, 600–606.

Swenson, S.A., Moore, C.M., Marcero, J.R., Medlock, A.E., Reddi, A.R., and Khalimonchuk, O. (2020). From Synthesis to Utilization: The Ins and Outs of Mitochondrial Heme. Cells *9*.

Taoka, M., Nobe, Y., Yamaki, Y., Sato, K., Ishikawa, H., Izumikawa, K., Yamauchi, Y., Hirota, K., Nakayama, H., Takahashi, N., et al. (2018). Landscape of the complete RNA chemical modifications in the human 80S ribosome. Nucleic Acids Res. *46*, 9289–9298.

Temperley, R., Richter, R., Dennerlein, S., Lightowlers, R.N., and Chrzanowska-Lightowlers, Z.M. (2010). Hungry codons promote frameshifting in human mitochondrial ribosomes. Science *327*, 301.

Terzioglu, M., Ruzzenente, B., Harmel, J., Mourier, A., Jemt, E., López, M.D., Kukat, C., Stewart, J.B., Wibom, R., Meharg, C., et al. (2013). MTERF1 binds mtDNA to prevent transcriptional interference at the light-strand promoter but is dispensable for rRNA gene transcription regulation. Cell Metab. *17*, 618–626.

Thomas, L.W., Esposito, C., Morgan, R.E., Price, S., Young, J., Williams, S.P., Maddalena, L.A., McDermott, U., and Ashcroft, M. (2021). Genome-wide CRISPR/Cas9 deletion screen defines mitochondrial gene essentiality and identifies routes for tumour cell viability in hypoxia. Commun. Biol. *4*, 1–12.

Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., and Hamon, C. (2003). Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. Anal. Chem. *75*, 1895–1904.

To, T.-L., Cuadros, A.M., Shah, H., Hung, W.H.W., Li, Y., Kim, S.H., Rubin, D.H.F., Boe, R.H., Rath, S., Eaton, J.K., et al. (2019). A Compendium of Genetic Modifiers of Mitochondrial Dysfunction Reveals Intra-organelle Buffering. Cell *179*, 1222-1238.e17.

Tomecki, R., Dmochowska, A., Gewartowski, K., Dziembowski, A., and Stepien, P.P. (2004). Identification of a novel human nuclear-encoded mitochondrial poly(A) polymerase. Nucleic Acids Res. *32*, 6001–6014.

Tsuboi, M., Morita, H., Nozaki, Y., Akama, K., Ueda, T., Ito, K., Nierhaus, K.H., and Takeuchi, N. (2009). EF-G2mt is an exclusive recycling factor in mammalian mitochondrial protein synthesis. Mol. Cell *35*, 502–510.

Tu, Y.-T., and Barrientos, A. (2015). The Human Mitochondrial DEAD-Box Protein DDX28 Resides in RNA Granules and Functions in Mitoribosome Assembly. Cell Rep. *10*, 854–864.

Týč, J., Novotná, L., Peña-Diaz, P., Maslov, D.A., and Lukeš, J. (2017). RSM22, mtYsxC and PNKD-like proteins are required for mitochondrial translation in Trypanosoma brucei. Mitochondrion *34*, 67–74.

Tzelepis, K., Koike-Yusa, H., De Braekeleer, E., Li, Y., Metzakopian, E., Dovey, O.M., Mupo, A., Grinkevich, V., Li, M., Mazan, M., et al. (2016). A CRISPR Dropout Screen

Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. Cell Rep. *17*, 1193–1205.

Uchida, A., Murugesapillai, D., Kastner, M., Wang, Y., Lodeiro, M.F., Prabhakar, S., Oliver, G.V., Arnold, J.J., Maher, L.J., III, Williams, M.C., et al. (2017). Unexpected sequences and structures of mtDNA required for efficient transcription from the first heavy-strand promoter. ELife *6*, e27283.

Uchiumi, T., Ohgaki, K., Yagi, M., Aoki, Y., Sakai, A., Matsumoto, S., and Kang, D. (2010). ERAL1 is associated with mitochondrial ribosome and elimination of ERAL1 leads to mitochondrial dysfunction and growth retardation. Nucleic Acids Res. *38*, 5554–5568.

Upadhyay, A.S., Vonderstein, K., Pichlmair, A., Stehling, O., Bennett, K.L., Dobler, G., Guo, J.-T., Superti-Furga, G., Lill, R., Överby, A.K., et al. (2014). Viperin is an iron-sulfur protein that inhibits genome synthesis of tick-borne encephalitis virus via radical SAM domain activity. Cell. Microbiol. *16*, 834–848.

Van Haute, L., Dietmann, S., Kremer, L., Hussain, S., Pearce, S.F., Powell, C.A., Rorbach, J., Lantaff, R., Blanco, S., Sauer, S., et al. (2016). Deficient methylation and formylation of mt-tRNAMet wobble cytosine in a patient carrying mutations in NSUN3. Nat. Commun. *7*.

Van Haute, L., Hendrick, A.G., D'Souza, A.R., Powell, C.A., Rebelo-Guiomar, P., Harbour, M.E., Ding, S., Fearnley, I.M., Andrews, B., and Minczuk, M. (2019). METTL15 introduces N4-methylcytidine into human mitochondrial 12S rRNA and is required for mitoribosome biogenesis. Nucleic Acids Res. *47*, 10267–10281.

Van Strien, J., Guerrero-Castillo, S., Chatzispyrou, I.A., Houtkooper, R.H., Brandt, U., and Huynen, M.A. (2019). COmplexome Profiling ALignment (COPAL) reveals remodeling of mitochondrial protein complexes in Barth syndrome. Bioinformatics *35*, 3083–3091.

Van Strien, J., Haupt, A., Schulte, U., Braun, H.-P., Cabrera-Orefice, A., Choudhary, J.S., Evers, F., Fernandez-Vizarra, E., Guerrero-Castillo, S., Kooij, T.W.A., et al. (2021). CEDAR, an online resource for the reporting and exploration of complexome profiling data. Biochim. Biophys. Acta BBA - Bioenerg. *1862*, 148411.

Verma, Y., Mehra, U., Pandey, D.K., Kar, J., Pérez-Martinez, X., Jana, S.S., and Datta, K. (2021). MRX8, the conserved mitochondrial YihA GTPase family member is required for de novo Cox1 synthesis at suboptimal temperatures in Saccharomyces cerevisiae. Mol. Biol. Cell mbcE20070457.

Verstraeten, N., Fauvart, M., Versées, W., and Michiels, J. (2011). The Universally Conserved Prokaryotic GTPases. Microbiol. Mol. Biol. Rev. MMBR *75*, 507–542.

Vidoni, S., Harbour, M.E., Guerrero-Castillo, S., Signes, A., Ding, S., Fearnley, I.M., Taylor, R.W., Tiranti, V., Arnold, S., Fernandez-Vizarra, E., et al. (2017). MR-1S Interacts with PET100 and PET117 in Module-Based Assembly of Human Cytochrome c Oxidase. Cell Rep. *18*, 1727–1738.

de Vries, H., and van der Koogh-Schuuring, R. (1973). Physicochemical characteristics of isolated 55-S mitochondrial ribosomes from rat-liver. Biochem. Biophys. Res. Commun. *54*, 308–314.

Walberg, M.W., and Clayton, D.A. (1981). Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Nucleic Acids Res. *9*, 5411–5421.

Waldner, C., Rempel, O., Schütte, F., Yanik, M., Solomentsew, N., and Ryffel, G.U. (2011). Double conditional human embryonic kidney cell line based on FLP and ΦC31 mediated transgene integration. BMC Res. Notes *4*, 420.

Walker, J.E. (2013). The ATP synthase: the understood, the uncertain and the unknown. Biochem. Soc. Trans. *41*, 1–16.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. *1*, 945–951.

Wang, X., and He, C. (2014). Dynamic RNA Modifications in Posttranscriptional Regulation. Mol. Cell *56*, 5–12.

Wang, Z., and Wu, M. (2015). An integrated phylogenomic approach toward pinpointing the origin of mitochondria. Sci. Rep. *5*.

Wang, F., Zhang, D., Zhang, D., Li, P., and Gao, Y. (2021). Mitochondrial Protein Translation: Emerging Roles and Clinical Significance in Disease. Front. Cell Dev. Biol. *9*, 1737.

Wang, J., Friedman, G., Doyon, Y., Wang, N.S., Li, C.J., Miller, J.C., Hua, K.L., Yan, J.J., Babiarz, J.E., Gregory, P.D., et al. (2012). Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. Genome Res. *22*, 1316–1326.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic Screens in Human Cells Using the CRISPR-Cas9 System. Science *343*, 80–84.

Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015). Identification and characterization of essential genes in the human genome. Science *350*, 1096–1101.

Wessels, H.J.C.T., Vogel, R.O., Heuvel, L. van den, Smeitink, J.A., Rodenburg, R.J., Nijtmans, L.G., and Farhoud, M.H. (2009). LC-MS/MS as an alternative for SDS-PAGE in blue native analysis of protein complexes. PROTEOMICS *9*, 4221–4228.

Wessels, H.J.C.T., Vogel, R.O., Lightowlers, R.N., Spelbrink, J.N., Rodenburg, R.J., Heuvel, L.P. van den, Gool, A.J. van, Gloerich, J., Smeitink, J.A.M., and Nijtmans, L.G. (2013). Analysis of 953 Human Proteins from a Mitochondrial HEK293 Fraction by Complexome Profiling. PLOS ONE *8*, e68340.

Wicker-Planquart, C., and Jault, J.-M. (2015). Interaction between Bacillus subtilis YsxC and ribosomes (or rRNAs). FEBS Lett. *589*, 1026–1032.

Wicker-Planquart, C., Foucher, A.-E., Louwagie, M., Britton, R.A., and Jault, J.-M. (2008). Interactions of an Essential Bacillus subtilis GTPase, YsxC, with Ribosomes. J. Bacteriol. *190*, 681–690.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New York).

Wiese, S., Reidegeld, K.A., Meyer, H.E., and Warscheid, B. (2007). Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. PROTEOMICS *7*, 340–350.

Wilson, W.C., Hornig-Do, H.-T., Bruni, F., Chang, J.H., Jourdain, A.A., Martinou, J.-C., Falkenberg, M., Spåhr, H., Larsson, N.-G., Lewis, R.J., et al. (2014). A human mitochondrial poly(A) polymerase mutation reveals the complexities of post-transcriptional mitochondrial gene expression. Hum. Mol. Genet. *23*, 6345–6355.

Wolf, A.R., and Mootha, V.K. (2014). Functional genomic analysis of human mitochondrial RNA processing. Cell Rep. *7*, 918–931.

Yakubovskaya, E., Chen, Z., Carrodeguas, J.A., Kisker, C., and Bogenhagen, D.F. (2006). Functional Human Mitochondrial DNA Polymerase γ Forms a Heterotrimer*. J. Biol. Chem. *281*, 374–382.

Yakubovskaya, E., Mejia, E., Byrnes, J., Hambardjieva, E., and Garcia-Diaz, M. (2010). Helix Unwinding and Base Flipping Enable Human MTERF1 to Terminate Mitochondrial Transcription. Cell *141*, 982–993.

Yasukawa, T., Reyes, A., Cluett, T.J., Yang, M.-Y., Bowmaker, M., Jacobs, H.T., and Holt, I.J. (2006). Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. EMBO J. *25*, 5358–5371.

Zaganelli, S., Rebelo-Guiomar, P., Maundrell, K., Rozanska, A., Pierredon, S., Powell, C.A., Jourdain, A.A., Hulo, N., Lightowlers, R.N., Chrzanowska-Lightowlers, Z.M., et al. (2017). The Pseudouridine Synthase RPUSD4 Is an Essential Component of Mitochondrial RNA Granules. J. Biol. Chem. *292*, 4519–4532.

Zamudio-Ochoa, A., Camacho-Villasana, Y., García-Guerrero, A.E., and Pérez-Martínez, X. (2014). The Pet309 pentatricopeptide repeat motifs mediate efficient binding to the mitochondrial COX1 transcript in yeast. RNA Biol. *11*, 953–967.

Zaremba-Niedzwiedzka, K., Caceres, E.F., Saw, J.H., Bäckström, D., Juzokaite, L., Vancaester, E., Seitz, K.W., Anantharaman, K., Starnawski, P., Kjeldsen, K.U., et al. (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature *541*, 353–358.

Zhang, Y., Guo, S., Xie, C., and Fang, J. (2020). Uridine Metabolism and Its Role in Glucose, Lipid, and Amino Acid Homeostasis. BioMed Res. Int. *2020*, 7091718.

Appendix I

Summary of genes in the gRNA library

Category	# of genes in category	Broader category	# of genes in broader category
control,negative	100	control,negative	100
control,noCut	50	control,noCut	50
control,positive	20	control,positive	20
aminoacylation	22	aminoacylation	22
mitoribosome assembly	16	mitoribosome assembly	16
mitoribosome structure	82	mitoribosome structure	82
predicted function in MGE	59	predicted function in MGE	59
replication	15	replication	15
transcription	4	transcription+RNA metabolism other	28
RNA metabolism other	24		
12S rRNA modifier	3	tRNA/rRNA modifier	41
16S rRNA modifier	3		
tRNA-modifier candidate	13		
tRNA-modifier	22		
translation	14	translation	14

List of all genes targeted by the gRNA library used for CRISPR/Cas9 screen

Gene	Broader category	Category
AARS2	aminoacylation	aminoacylation
CARS2	aminoacylation	aminoacylation
DARS2	aminoacylation	aminoacylation
EARS2	aminoacylation	aminoacylation
FARS2	aminoacylation	aminoacylation
GARS	aminoacylation	aminoacylation
GATB	aminoacylation	aminoacylation
GATC	aminoacylation	aminoacylation
HARS2	aminoacylation	aminoacylation
IARS2	aminoacylation	aminoacylation
KARS	aminoacylation	aminoacylation
LARS2	aminoacylation	aminoacylation
MARS2	aminoacylation	aminoacylation
NARS2	aminoacylation	aminoacylation
PARS2	aminoacylation	aminoacylation
QRSL1	aminoacylation	aminoacylation

RARS2	aminoacylation	aminoacylation
SARS2	aminoacylation	aminoacylation
TARS2	aminoacylation	aminoacylation
VARS2	aminoacylation	aminoacylation
WARS2	aminoacylation	aminoacylation
YARS2	aminoacylation	aminoacylation
ABCA5	control,negative	control,negative
ACTN2	control,negative	control,negative
ALDH1A1	control,negative	control,negative
ANGPT2	control,negative	control,negative
AREL1	control,negative	control,negative
ARHGEF26	control,negative	control,negative
ARNTL	control,negative	control,negative
ATG4C	control,negative	control,negative
ATOH1	control,negative	control,negative
AXDND1	control,negative	control,negative
B3GNT9	control,negative	control,negative
C19orf73	control,negative	control,negative
CCDC83	control,negative	control,negative
CCR4	control,negative	control,negative
CD58	control,negative	control,negative
CEP85L	control,negative	control,negative
CERK	control,negative	control,negative
CLHC1	control,negative	control,negative
CRIM1	control,negative	control,negative
CYP2W1	control,negative	control,negative
DCAF5	control,negative	control,negative
DENND1C	control,negative	control,negative
DIRAS1	control,negative	control,negative
DRD4	control,negative	control,negative
DSCAM	control,negative	control,negative
EMP2	control,negative	control,negative
EMP3	control,negative	control,negative
ERI2	control,negative	control,negative
ETNPPL	control,negative	control,negative
FAM186B	control,negative	control,negative
FAM71C	control,negative	control,negative
FCHSD2	control,negative	control,negative
FLG2	control,negative	control,negative
GLIPR1	control,negative	control,negative
GNG2	control,negative	control,negative
GTPBP1	control,negative	control,negative
HIBADH	control,negative	control,negative
HYAL1	control,negative	control,negative
IFIT1	control,negative	control,negative
IL36A	control,negative	control,negative

ISOC2	control,negative	control,negative
IZUMO1	control,negative	control,negative
KCNK12	control,negative	control,negative
KCTD2	control,negative	control,negative
KLK9	control,negative	control,negative
KRBA2	control,negative	control,negative
KRT26	control,negative	control,negative
KRTAP19-7	control,negative	control,negative
LEMD1	control,negative	control,negative
MCTP1	control,negative	control,negative
METTL2A	control,negative	control,negative
MGARP	control,negative	control,negative
MPZ	control,negative	control,negative
MUSK	control,negative	control,negative
NAB1	control,negative	control,negative
NIT1	control,negative	control,negative
PDLIM4	control,negative	control,negative
PHF7	control,negative	control,negative
PIGY	control,negative	control,negative
PITPNM2	control,negative	control,negative
PLEKHF1	control,negative	control,negative
PLXNC1	control,negative	control,negative
PPP1R16B	control,negative	control,negative
PYDC1	control,negative	control,negative
PYGL	control,negative	control,negative
QSOX2	control,negative	control,negative
REPIN1	control,negative	control,negative
SCG2	control,negative	control,negative
SDR39U1	control,negative	control,negative
SHD	control,negative	control,negative
SHPK	control,negative	control,negative
SLC25A17	control,negative	control,negative
SLC44A2	control,negative	control,negative
SLPI	control,negative	control,negative
SLURP1	control,negative	control,negative
SNX7	control,negative	control,negative
SOX15	control,negative	control,negative
SPATA32	control,negative	control,negative
SPG11	control,negative	control,negative
SPOCK1	control,negative	control,negative
TCEANC2	control,negative	control,negative
TCN1	control,negative	control,negative
TEX261	control,negative	control,negative
TFDP2	control,negative	control,negative
TIGD6	control,negative	control,negative

TLR5	control,negative	control,negative
TM4SF1	control,negative	control,negative
TMED6	control,negative	control,negative
TMEM8C	control,negative	control,negative
TMPRSS3	control,negative	control,negative
TOPAZ1	control,negative	control,negative
TRAF1	control,negative	control,negative
UMOD	control,negative	control,negative
XAF1	control,negative	control,negative
ZNF10	control,negative	control,negative
ZNF229	control,negative	control,negative
ZNF239	control,negative	control,negative
ZNF326	control,negative	control,negative
ZNF578	control,negative	control,negative
ZNF79	control,negative	control,negative
CTRL01	control,noCut	control,noCut
CTRL02	control,noCut	control,noCut
CTRL03	control,noCut	control,noCut
CTRL04	control,noCut	control,noCut
CTRL05	control,noCut	control,noCut
CTRL06	control,noCut	control,noCut
CTRL07	control,noCut	control,noCut
CTRL08	control,noCut	control,noCut
CTRL09	control,noCut	control,noCut
CTRL10	control,noCut	control,noCut
CTRL11	control,noCut	control,noCut
CTRL12	control,noCut	control,noCut
CTRL13	control,noCut	control,noCut
CTRL14	control,noCut	control,noCut
CTRL15	control,noCut	control,noCut
CTRL16	control,noCut	control,noCut
CTRL17	control,noCut	control,noCut
CTRL18	control,noCut	control,noCut
CTRL19	control,noCut	control,noCut
CTRL20	control,noCut	control,noCut
CTRL21	control,noCut	control,noCut
CTRL22	control,noCut	control,noCut
CTRL23	control,noCut	control,noCut
CTRL24	control,noCut	control,noCut
CTRL25	control,noCut	control,noCut
CTRL26	control,noCut	control,noCut
CTRL27	control,noCut	control,noCut
CTRL28	control,noCut	control,noCut
CTRL29	control,noCut	control,noCut
CTRL30	control,noCut	control,noCut
CTRL31	control,noCut	control,noCut

CTRL32	control,noCut	control,noCut
CTRL33	control,noCut	control,noCut
CTRL34	control,noCut	control,noCut
CTRL35	control,noCut	control,noCut
CTRL36	control,noCut	control,noCut
CTRL37	control,noCut	control,noCut
CTRL38	control,noCut	control,noCut
CTRL39	control,noCut	control,noCut
CTRL40	control,noCut	control,noCut
CTRL41	control,noCut	control,noCut
CTRL42	control,noCut	control,noCut
CTRL43	control,noCut	control,noCut
CTRL44	control,noCut	control,noCut
CTRL45	control,noCut	control,noCut
CTRL46	control,noCut	control,noCut
CTRL47	control,noCut	control,noCut
CTRL48	control,noCut	control,noCut
CTRL49	control,noCut	control,noCut
CTRL50	control,noCut	control,noCut
ATP6V0C	control,positive	control,positive
BUD31	control,positive	control,positive
CCNA2	control,positive	control,positive
DHX15	control,positive	control,positive
GEMIN5	control,positive	control,positive
KPNB1	control,positive	control,positive
MYC	control,positive	control,positive
NOP16	control,positive	control,positive
POLR2D	control,positive	control,positive
POLR3A	control,positive	control,positive
PRPF19	control,positive	control,positive
RPL14	control,positive	control,positive
RPL18	control,positive	control,positive
RPL3	control,positive	control,positive
RPS11	control,positive	control,positive
RPS19	control,positive	control,positive
SMU1	control,positive	control,positive
SNRPD1	control,positive	control,positive
SNRPF	control,positive	control,positive
UBL5	control,positive	control,positive
DDX28	mitoribosome assembly	mitoribosome assembly
DHX30	mitoribosome assembly	mitoribosome assembly
ERAL1	mitoribosome assembly	mitoribosome assembly
FASTKD2	mitoribosome assembly	mitoribosome assembly
GTPBP5	mitoribosome assembly	mitoribosome assembly
L0R8F8	mitoribosome assembly	mitoribosome assembly
MALSU1	mitoribosome assembly	mitoribosome assembly

MPV17L2	mitoribosome assembly	mitoribosome assembly
MTERF3	mitoribosome assembly	mitoribosome assembly
MTERFD2	mitoribosome assembly	mitoribosome assembly
MTG1	mitoribosome assembly	mitoribosome assembly
NGRN	mitoribosome assembly	mitoribosome assembly
NOA1	mitoribosome assembly	mitoribosome assembly
RMND1	mitoribosome assembly	mitoribosome assembly
WBSCR16	mitoribosome assembly	mitoribosome assembly
YBEY	mitoribosome assembly	mitoribosome assembly
DAP3	mitoribosome structure	mitoribosome structure
MRPL1	mitoribosome structure	mitoribosome structure
MRPL10	mitoribosome structure	mitoribosome structure
MRPL11	mitoribosome structure	mitoribosome structure
MRPL12	mitoribosome structure	mitoribosome structure
MRPL13	mitoribosome structure	mitoribosome structure
MRPL14	mitoribosome structure	mitoribosome structure
MRPL15	mitoribosome structure	mitoribosome structure
MRPL16	mitoribosome structure	mitoribosome structure
MRPL17	mitoribosome structure	mitoribosome structure
MRPL18	mitoribosome structure	mitoribosome structure
MRPL19	mitoribosome structure	mitoribosome structure
MRPL2	mitoribosome structure	mitoribosome structure
MRPL20	mitoribosome structure	mitoribosome structure
MRPL21	mitoribosome structure	mitoribosome structure
MRPL22	mitoribosome structure	mitoribosome structure
MRPL23	mitoribosome structure	mitoribosome structure
MRPL24	mitoribosome structure	mitoribosome structure
MRPL27	mitoribosome structure	mitoribosome structure
MRPL28	mitoribosome structure	mitoribosome structure
MRPL3	mitoribosome structure	mitoribosome structure
MRPL30	mitoribosome structure	mitoribosome structure
MRPL32	mitoribosome structure	mitoribosome structure
MRPL33	mitoribosome structure	mitoribosome structure
MRPL34	mitoribosome structure	mitoribosome structure
MRPL35	mitoribosome structure	mitoribosome structure
MRPL36	mitoribosome structure	mitoribosome structure
MRPL37	mitoribosome structure	mitoribosome structure
MRPL38	mitoribosome structure	mitoribosome structure
MRPL39	mitoribosome structure	mitoribosome structure
MRPL4	mitoribosome structure	mitoribosome structure
MRPL40	mitoribosome structure	mitoribosome structure
MRPL41	mitoribosome structure	mitoribosome structure
MRPL42	mitoribosome structure	mitoribosome structure
MRPL43	mitoribosome structure	mitoribosome structure
MRPL44	mitoribosome structure	mitoribosome structure
MRPL45	mitoribosome structure	mitoribosome structure

MRPL46	mitoribosome structure	mitoribosome structure
MRPL47	mitoribosome structure	mitoribosome structure
MRPL48	mitoribosome structure	mitoribosome structure
MRPL49	mitoribosome structure	mitoribosome structure
MRPL50	mitoribosome structure	mitoribosome structure
MRPL51	mitoribosome structure	mitoribosome structure
MRPL52	mitoribosome structure	mitoribosome structure
MRPL53	mitoribosome structure	mitoribosome structure
MRPL54	mitoribosome structure	mitoribosome structure
MRPL55	mitoribosome structure	mitoribosome structure
MRPL57	mitoribosome structure	mitoribosome structure
MRPL58	mitoribosome structure	mitoribosome structure
MRPL59	mitoribosome structure	mitoribosome structure
MRPL9	mitoribosome structure	mitoribosome structure
MRPS10	mitoribosome structure	mitoribosome structure
MRPS11	mitoribosome structure	mitoribosome structure
MRPS12	mitoribosome structure	mitoribosome structure
MRPS14	mitoribosome structure	mitoribosome structure
MRPS15	mitoribosome structure	mitoribosome structure
MRPS16	mitoribosome structure	mitoribosome structure
MRPS17	mitoribosome structure	mitoribosome structure
MRPS18A	mitoribosome structure	mitoribosome structure
MRPS18B	mitoribosome structure	mitoribosome structure
MRPS18C	mitoribosome structure	mitoribosome structure
MRPS2	mitoribosome structure	mitoribosome structure
MRPS21	mitoribosome structure	mitoribosome structure
MRPS22	mitoribosome structure	mitoribosome structure
MRPS23	mitoribosome structure	mitoribosome structure
MRPS24	mitoribosome structure	mitoribosome structure
MRPS25	mitoribosome structure	mitoribosome structure
MRPS26	mitoribosome structure	mitoribosome structure
MRPS27	mitoribosome structure	mitoribosome structure
MRPS28	mitoribosome structure	mitoribosome structure
MRPS30	mitoribosome structure	mitoribosome structure
MRPS31	mitoribosome structure	mitoribosome structure
MRPS33	mitoribosome structure	mitoribosome structure
MRPS34	mitoribosome structure	mitoribosome structure
MRPS35	mitoribosome structure	mitoribosome structure
MRPS37	mitoribosome structure	mitoribosome structure
MRPS38	mitoribosome structure	mitoribosome structure
MRPS39	mitoribosome structure	mitoribosome structure
MRPS5	mitoribosome structure	mitoribosome structure
MRPS6	mitoribosome structure	mitoribosome structure
MRPS7	mitoribosome structure	mitoribosome structure
MRPS9	mitoribosome structure	mitoribosome structure
AK4	predicted function in MGE	predicted function in MGE

AKAP1	predicted function in MGE	predicted function in MGE
ALKBH7	predicted function in MGE	predicted function in MGE
ANGEL2	predicted function in MGE	predicted function in MGE
ASMTL	predicted function in MGE	predicted function in MGE
AUH	predicted function in MGE	predicted function in MGE
C16orf62	predicted function in MGE	predicted function in MGE
C19orf52	predicted function in MGE	predicted function in MGE
C19orf70	predicted function in MGE	predicted function in MGE
C21orf33	predicted function in MGE	predicted function in MGE
C3orf33	predicted function in MGE	predicted function in MGE
CCBL2	predicted function in MGE	predicted function in MGE
CHCHD2	predicted function in MGE	predicted function in MGE
COMTD1	predicted function in MGE	predicted function in MGE
DAZAP1	predicted function in MGE	predicted function in MGE
DHX29	predicted function in MGE	predicted function in MGE
DNAJA3	predicted function in MGE	predicted function in MGE
DNLZ	predicted function in MGE	predicted function in MGE
DRG2	predicted function in MGE	predicted function in MGE
EEFSEC	predicted function in MGE	predicted function in MGE
FASN	predicted function in MGE	predicted function in MGE
FBXO18	predicted function in MGE	predicted function in MGE
FBXO4	predicted function in MGE	predicted function in MGE
FRAT2	predicted function in MGE	predicted function in MGE
GLUD1	predicted function in MGE	predicted function in MGE
GTPBP10	predicted function in MGE	predicted function in MGE
GTPBP6	predicted function in MGE	predicted function in MGE
GTPBP8	predicted function in MGE	predicted function in MGE
HDDC2	predicted function in MGE	predicted function in MGE
HEMK1	predicted function in MGE	predicted function in MGE
HNRNPK	predicted function in MGE	predicted function in MGE
LRRC59	predicted function in MGE	predicted function in MGE
MCTS1	predicted function in MGE	predicted function in MGE
METTL17	predicted function in MGE	predicted function in MGE
NELFE	predicted function in MGE	predicted function in MGE
NFXL1	predicted function in MGE	predicted function in MGE
NHP2	predicted function in MGE	predicted function in MGE
NIF3L1	predicted function in MGE	predicted function in MGE
NOCT	predicted function in MGE	predicted function in MGE
NOXIN	predicted function in MGE	predicted function in MGE
NTHL1	predicted function in MGE	predicted function in MGE
NUDT2	predicted function in MGE	predicted function in MGE
PDIP38	predicted function in MGE	predicted function in MGE
PSMA6	predicted function in MGE	predicted function in MGE
PTRF	predicted function in MGE	predicted function in MGE
PTRH1	predicted function in MGE	predicted function in MGE

PTRH2	predicted function in MGE	predicted function in MGE
PTRHD1	predicted function in MGE	predicted function in MGE
RECQL4	predicted function in MGE	predicted function in MGE
RNASEL	predicted function in MGE	predicted function in MGE
RRNAD1	predicted function in MGE	predicted function in MGE
RSPH9	predicted function in MGE	predicted function in MGE
SECISBP2	predicted function in MGE	predicted function in MGE
SND1	predicted function in MGE	predicted function in MGE
SPRYD4	predicted function in MGE	predicted function in MGE
TATDN3	predicted function in MGE	predicted function in MGE
TDRKH	predicted function in MGE	predicted function in MGE
TSHZ3	predicted function in MGE	predicted function in MGE
VWA8	predicted function in MGE	predicted function in MGE
APEX1	replication	replication
DNA2	replication	replication
LIG3	replication	replication
MGME1	replication	replication
MPV17	replication	replication
NIPSNAP1	replication	replication
OGG1	replication	replication
POLG	replication	replication
POLG2	replication	replication
RNASEH1	replication	replication
SSBP1	replication	replication
TOP1MT	replication	replication
TOP2A	replication	replication
TWNK	replication	replication
UNG	replication	replication
ATAD3A	transcription+RNA metabolism other	RNA metabolism other
ATAD3B	transcription+RNA metabolism other	RNA metabolism other
C1QBP	transcription+RNA metabolism other	RNA metabolism other
ELAC2	transcription+RNA metabolism other	RNA metabolism other
FASTK	transcription+RNA metabolism other	RNA metabolism other
FASTKD1	transcription+RNA metabolism other	RNA metabolism other
FASTKD3	transcription+RNA metabolism other	RNA metabolism other
FASTKD5	transcription+RNA metabolism other	RNA metabolism other
GRSF1	transcription+RNA metabolism other	RNA metabolism other
HSD17B10	transcription+RNA metabolism other	RNA metabolism other
KIAA0391	transcription+RNA metabolism other	RNA metabolism other
LACTB2	transcription+RNA metabolism other	RNA metabolism other
LRPPRC	transcription+RNA metabolism other	RNA metabolism other
MTERF	transcription+RNA metabolism other	transcription
MTERF2	transcription+RNA metabolism other	transcription
MTPAP	transcription+RNA metabolism other	RNA metabolism other
PABPC5	transcription+RNA metabolism other	RNA metabolism other
PAPD5	transcription+RNA metabolism other	RNA metabolism other

PDE12	transcription+RNA metabolism other	RNA metabolism other
PNPT1	transcription+RNA metabolism other	RNA metabolism other
POLRMT	transcription+RNA metabolism other	transcription
PTCD1	transcription+RNA metabolism other	RNA metabolism other
PTCD2	transcription+RNA metabolism other	RNA metabolism other
REXO2	transcription+RNA metabolism other	RNA metabolism other
SLIRP	transcription+RNA metabolism other	RNA metabolism other
SUPV3L1	transcription+RNA metabolism other	RNA metabolism other
TBRG4	transcription+RNA metabolism other	RNA metabolism other
TEFM	transcription+RNA metabolism other	transcription
TFAM	transcription+RNA metabolism other	transcription
TFB2M	transcription+RNA metabolism other	transcription
C12orf65	translation	translation
GFM1	translation	translation
GFM2	translation	translation
GUF1	translation	translation
MRRF	translation	translation
MTFMT	translation	translation
MTIF2	translation	translation
MTIF3	translation	translation
MTRF1	translation	translation
MTRF1L	translation	translation
OXA1L	translation	translation
TACO1	translation	translation
TSFM	translation	translation
TUFM	translation	translation
ALKBH1	tRNA/rRNA modifier	tRNA-modifier
CDK5RAP1	tRNA/rRNA modifier	tRNA-modifier
DUS2	tRNA/rRNA modifier	tRNA-modifier candidate
GTPBP3	tRNA/rRNA modifier	tRNA-modifier
METTL15	tRNA/rRNA modifier	12S rRNA modifier
METTL2B	tRNA/rRNA modifier	tRNA-modifier candidate
METTL8	tRNA/rRNA modifier	tRNA-modifier
MRM1	tRNA/rRNA modifier	16S rRNA modifier
MRM2	tRNA/rRNA modifier	16S rRNA modifier
MRM3	tRNA/rRNA modifier	16S rRNA modifier
NFS1	tRNA/rRNA modifier	tRNA-modifier candidate
NSUN2	tRNA/rRNA modifier	tRNA-modifier
NSUN3	tRNA/rRNA modifier	tRNA-modifier
NSUN4	tRNA/rRNA modifier	12S rRNA modifier
OSGEPL1	tRNA/rRNA modifier	tRNA-modifier
PUS1	tRNA/rRNA modifier	tRNA-modifier
PUS3	tRNA/rRNA modifier	tRNA-modifier candidate
PUSL1	tRNA/rRNA modifier	tRNA-modifier
PYURF	tRNA/rRNA modifier	tRNA-modifier candidate
QTRT1	tRNA/rRNA modifier	tRNA-modifier

QTRTD1	tRNA/rRNA modifier	tRNA-modifier
RPUSD1	tRNA/rRNA modifier	tRNA-modifier candidate
RPUSD2	tRNA/rRNA modifier	tRNA-modifier candidate
RPUSD3	tRNA/rRNA modifier	tRNA-modifier candidate
RPUSD4	tRNA/rRNA modifier	tRNA-modifier|16S rRNA modifier
TFB1M	tRNA/rRNA modifier	12S rRNA modifier
THG1L	tRNA/rRNA modifier	tRNA-modifier
THUMPD2	tRNA/rRNA modifier	tRNA-modifier candidate
THUMPD3	tRNA/rRNA modifier	tRNA-modifier candidate
TRIT1	tRNA/rRNA modifier	tRNA-modifier
TRMT1	tRNA/rRNA modifier	tRNA-modifier
TRMT10C	tRNA/rRNA modifier	tRNA-modifier
TRMT11	tRNA/rRNA modifier	tRNA-modifier candidate
TRMT112	tRNA/rRNA modifier	tRNA-modifier candidate
TRMT2B	tRNA/rRNA modifier	tRNA-modifier|12S rRNA modifier
TRMT5	tRNA/rRNA modifier	tRNA-modifier
TRMT61B	tRNA/rRNA modifier	tRNA-modifier|16S rRNA modifier
TRMU	tRNA/rRNA modifier	tRNA-modifier
TRNT1	tRNA/rRNA modifier	tRNA-modifier
TRUB2	tRNA/rRNA modifier	tRNA-modifier candidate
YRDC	tRNA/rRNA modifier	tRNA-modifier