Thesis declaration

I, Stefan Johannes Siira, certify that:

This thesis has been substantially accomplished during enrolment in the degree.

This thesis does not contain material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution.

No part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of The University of Western Australia and where applicable, any partner institution responsible for the joint-award of this degree.

This thesis does not contain any material previously published or written by another person, except where due reference has been made in the text.

The work(s) are not in any way a violation or infringement of any copyright, trademark, patent, or other rights whatsoever of any person.

The research involving animal data reported in this thesis was assessed and approved by The University of Western Australia Animal Ethics Committee. Approval #: RA/3/100/1256, RA/3/100/1390, RA/3/100/1477, RA/3/300/75, RA/3/300/92, RA/3/300/98

The research involving animals reported in this thesis followed The University of Western Australia and national standards for the care and use of laboratory animals.

The work described in this thesis was funded by the National Health and Medical Research Council (APP1058442, APP1045677, APP1041582, APP1023460, APP1005030, APP1043978), and the Australian Research Council.

This thesis contains published work and/or work prepared for publication, some of which has been co-authored.

Signature:

Date:
Acknowledgements

I would like to thank my supervisors Aleksandra Filipovska and Oliver Rackham for their steadfast support and encouragement through difficult times. There were many times I didn't believe I could make it through, but they always did everything they could to help me keep going and push through and I am immensely grateful. I would also like to thank the other members of the lab, for making it such a friendly work environment with little in the way of office drama. And lastly, I would like to thank my parents, who have always pushed me to succeed and who have always been reliable in their support when I have needed it. I couldn't have done it without you Mum and Dad.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.
List of Publications


* co-first author


Conference Proceedings


<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>LRPPRC</td>
<td>leucine-rich pentatricopeptide cassette protein</td>
</tr>
<tr>
<td>miRNA</td>
<td>microRNA</td>
</tr>
<tr>
<td>Mitoribosome</td>
<td>mitochondrial ribosome</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>MRPP1</td>
<td>mitochondrial RNase P protein 1</td>
</tr>
<tr>
<td>MRPP2</td>
<td>mitochondrial RNase P protein 2</td>
</tr>
<tr>
<td>MRPP3</td>
<td>mitochondrial RNase P protein 3</td>
</tr>
<tr>
<td>mtDNA</td>
<td>mitochondrial DNA</td>
</tr>
<tr>
<td>mtPAP</td>
<td>mitochondrial polyadenosine phosphylase</td>
</tr>
<tr>
<td>mt-tRNA</td>
<td>mitochondrial transfer RNA</td>
</tr>
<tr>
<td>ncRNA</td>
<td>non-coding RNA</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>OXPHOS</td>
<td>oxidative phosphorylation</td>
</tr>
<tr>
<td>PAR-CLIP</td>
<td>photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation</td>
</tr>
<tr>
<td>PARE</td>
<td>parallel analysis of RNA ends</td>
</tr>
<tr>
<td>POLR3A</td>
<td>RNA polymerase III, subunit A</td>
</tr>
<tr>
<td>POLRMT</td>
<td>mitochondrial RNA polymerase</td>
</tr>
<tr>
<td>PPR</td>
<td>pentatricopeptide repeat</td>
</tr>
<tr>
<td>PTCD1</td>
<td>pentatricopeptide repeat containing domain protein 1</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>SLIRP</td>
<td>SRA stem-loop-interacting RNA-binding protein</td>
</tr>
<tr>
<td>snRNA</td>
<td>small nuclear RNA</td>
</tr>
<tr>
<td>TFAM</td>
<td>transcription factor A, mitochondrial</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
</tbody>
</table>
Abstract

The computational analysis of transcriptome-wide RNA sequencing data sets provides the opportunity to investigate a wide array of post-transcriptional processes that modulate RNA metabolism and facilitate gene expression in the cell. In this thesis, I have applied various analytic methods to examine the consequences on gene expression of pathogenic mutations that disrupt proteins involved in the regulation of the nuclear and mitochondrial transcriptomes. The expression of genes from both of these genomes is integral to cellular function and their disruption in models of disease has been analysed computationally to uncover how they function.

The nuclear RNA polymerase III (Pol III) transcribes tRNAs and regulatory non-coding RNAs and mutations in the largest subunit of this complex, subunit 3A, give rise to neurological disorders. A mutation in Pol III in three patients was identified with an intronic mutation in this gene that presented an unusual phenotype and performed differential gene expression analyses to investigate the pathological consequences of the mutation. I found an overall reduction in relative tRNA abundance and dysregulated expression of other Pol III transcripts involved in tRNA maturation, transcriptional regulation of Pol II and translation. In addition, I identified a downstream reduction of protein-coding expression in genes involved in pre-mRNA splicing, rRNA maturation and proteostasis that have been previously linked to neurodegeneration. These results suggest that the impairment of Pol III leads to impaired gene expression and consequent imbalance of tRNAs and non-coding RNAs that have further downstream effects on the expression of genes involved in RNA metabolism and proteostasis.

RNA metabolism within mitochondria differs from that operating on the nuclear transcriptome due to differences in genomic features such as its lack of introns, its compact size, low gene number and organisation. Mutations in the mitochondrial genome sequence, or in nuclear genes encoding proteins imported into the mitochondrion, have been shown to disrupt mitochondrial RNA processing and result in disease. I captured and mapped the 5′ ends of mitochondrial RNAs in two cell lines harbouring mutations in mitochondrial tRNA\textsuperscript{Leu(UUR)}, confirming impaired processing of this transcript and indicating further downstream effects on the processing of other transcripts in one mutant. Furthermore, I identified numerous unidentified sites within
mRNA, rRNA and non-coding regions where RNA 5′ ends were localised, that were proximal to regions of secondary structure and may indicate the action of promiscuous nucleases. These results reveal the variable effects on RNA processing that these mutations have and provide evidence for novel processing sites that demonstrate the need for further investigating the lifecycle of RNAs in this organelle.

The LRPPRC/SLIRP complex is essential for the stability, polyadenylation and coordinated translation of mitochondrial mRNAs, however the sites where this complex binds RNA are unknown. I discovered putative footprints in isolated mitochondrial RNA from Lrpprc and Slirp knockout mice using digital RNase footprinting and discovered that these proteins have a preference for binding mRNA and rRNA transcripts. My findings were consistent with my analyses of the location of binding sites identified by PAR-CLIP when I used mouse embryonic fibroblast cell lines. Furthermore I identified that the binding sites of these proteins were conserved in human cells by carrying out PAR-CLIP analyses in these cells. I examined secondary structure changes to show that the binding of this complex prevents the formation of secondary structure and exposes RNA for maturation and translation.

The use of computational methods to analyse the transcriptome is a valuable approach for investigating all stages of gene expression, from transcription, to splicing and processing, nucleotide modification and RNA:protein interactions. In this work the techniques that I have developed and established helped me to investigate the transcriptome in models of disease and to understand the mechanisms of gene regulation.
# Table of Contents

Declaration.................................................................................................................................ii
Acknowledgments.....................................................................................................................iii
List of Publications....................................................................................................................iv
Conference Proceedings.........................................................................................................vi
Abbreviations............................................................................................................................viii
Abstract......................................................................................................................................ix

## Table of Contents

1: Introduction.............................................................................................................................1

1.1 Transcriptome-wide analyses and applications to models of disease.............................5
1.2 The RNA polymerase III transcriptome.................................................................10
1.3 Mitochondrial gene expression and RNA metabolism..............................................10
    1.3.1 Mitochondrial RNA processing...............................................................14
    1.3.2 Mitochondrial RNA maturation and stability...........................................16
1.4 Aims of Study....................................................................................................................18

2: Results and Discussion.........................................................................................................19

2.1 Paper I..........................................................................................................................19
2.2 Paper II.........................................................................................................................22
2.3 Paper III......................................................................................................................26

3: General Discussion...............................................................................................................30

4: References...........................................................................................................................34

5. Papers

5.1 Paper I..........................................................................................................................41
5.2 Paper II........................................................................................................................77
5.3 Paper III.........................................................................................................................103

6: Appendices

A...........................................................................................................................................145
B...........................................................................................................................................177
C...........................................................................................................................................209
D...........................................................................................................................................220
E...........................................................................................................................................237
Declaration for Theses Containing Published and/or Work Prepared for Publication

This thesis contains co-authored published work and work prepared for publication. The bibliographic details of the work and contribution of authors is stated below.

Paper I

Stefan Siira's contribution is 50% and the other co-first author identified the mutation by whole exome sequencing and provided the description of the clinical findings. I carried out all the transcriptomic studies presented in this work, developed new bioinformatics pipelines to analyse the data, visualised the data, wrote the manuscript that was edited by the senior authors.

Paper II

Stefan Siira's contribution is at least 90% and the contribution of the remaining authors is 10%. I carried out all the work in the paper, developed bioinformatics pipelines, analysed the data and wrote the manuscript. The senior authors edited the manuscript before submission.

Paper III
Stefan Siira's contribution is at least 90% and the contribution of the remaining authors is 10%. I carried out all the work in the paper, developed bioinformatics pipelines, analysed the data and wrote the manuscript. The senior authors edited the manuscript before submission.

Appendix A


Stefan Siira's contribution is at least 15% and the contribution of the remaining authors is 85%. I carried out the transcriptomic analyses, developed the pipelines, wrote the methods and results section related to the bioinformatic data and results. I edited the manuscript before submission.

Appendix B


Stefan Siira's contribution is at least 35% and the contribution of the remaining authors is 65%. I carried out the transcriptomic analyses, developed new methods to analyse the presented data, wrote the methods and results section related to the bioinformatic data and results. I edited the manuscript before submission.

Appendix C

Stefan Siira's contribution is at least 20% and the contribution of the remaining authors is 80%. I carried out the transcriptomic analyses, developed new methods to analyse the presented data, wrote the methods and results sections related to the bioinformatic data and results. I edited the manuscript before submission.

**Appendix D**

Stefan Siira's contribution is at least 20% and the contribution of the remaining authors is 80%. I carried out the protein preparation and purification optimisation, analysed the presented data, wrote the methods and results section related to my contribution. I edited the manuscript before submission.

**Appendix E**

Stefan Siira's contribution is at least 20% and the contribution of the remaining authors is 80%. I carried out the transcriptomic analyses, developed new methods to analyse the presented data, wrote the methods and results section related to the bioinformatic data and results. I edited the manuscript before submission.

Student signature:
Date:

I, Aleksandra Filipovska, certify that the student statements regarding their contribution to each of the works listed above are correct.

Coordinating supervisor signature:
Date:
1. Introduction

The era of genetics gave way to that of genomics with the sequencing of the human genome (Lander et al., 2001), which promoted a rapid development of high-throughput sequencing technology and led to the discovery of a great number of novel genes and isoforms. Today, after the widespread proliferation of genome sequencing, highly detailed genome annotations of model species and large-scale databases cataloguing a multitude of genetic variants, the study of the complement of ribonucleic acid produced from a genome, transcriptomics, has become a popular new frontier to uncover functional insights into the complexity of gene expression. As our understanding of the genome has improved, it has become clear that many disease-causing mutations are not simply disrupting the production of specific proteins but also by affecting the regulatory systems that operate within the cell such as splicing, maturation, transcription and translation. By investigating the transcriptome we can reveal new insight into how it is expressed and regulated, which will provide the necessary conceptual basis for developing an understanding of how it is disrupted in disease.

This proliferation of high-throughput transcriptome data sets has created a bottleneck from the ever-increasing amount of data whilst the number of people with necessary expertise for analysis has lagged behind. Addressing this problem requires a deep biochemistry and molecular biology knowledge in addition to a high degree of informatics and computer literacy skills, including programming. There is also a need for the development of, and establishing consensus on, the use of standardised bioinformatic pipelines to allow for results obtained from different models and research groups to be comparable and integrable. This is complicated by the further development of new high-throughput methods, which can allow for or require new analytic approaches, and developments in computer science and statistics, which can rapidly change the preferred tools of use.

1.1 Transcriptome-wide analyses and applications

The development of high-throughput RNA sequencing technologies, especially the recent advances in short read technology, have resulted in the development of a range of sequencing methodologies from wide spectrum preparations to those geared towards
specific processes. The current range of sequencing platforms, often called next-generation sequencing, typically involves the isolation and fragmentation of RNA samples, ligation of adapter sequences, conversion into complementary DNA (cDNA) by reverse transcriptase and sequencing by short read sequencers (reviewed in (Wang et al., 2009)). Many different variations on this procedure have been developed to accommodate a variety of RNA species or to capture products of specific transcriptional processes.

For investigating widespread gene expression, of messenger RNAs (mRNAs) and non-coding RNAs (ncRNAs), total RNA sequencing library preparation methods have become most frequently used (Figure 1). These methods capture mRNAs and multiple types of ncRNAs, selectively depleting nuclear and/or mitochondrial ribosomal RNAs (rRNAs) if desired, and produces a data set containing thousands or millions of short sequences that represent fragmented portions of the transcriptome. Varying the library preparation can alter the composition of the resulting library and enrich for particular types of transcripts. For instance, incorporating a size-selection step for small molecules into the RNA preparation will produce a library that contains a much higher proportion of reads derived from endogenous small transcripts. This aids the investigation of transcripts such as transfer RNAs (tRNAs) and micro RNAs (miRNAs), and the identification of new small RNA transcripts. The analysis of these types of sequencing data sets has developed to encompass methods for many different purposes. Applications such as DESeq/DESeq2 (Anders and Huber, 2010; Love et al., 2014) and edgeR (Robinson et al., 2010) have been developed and applied to the analysis of differential gene expression, particularly in microbial and organellar genomes. In brief, these applications apply statistical approaches to the counts of reads mapped to regions overlapping genomic features (such as exons) between sample groups to normalise and test for differences in expression level. More complex techniques include transcriptome assembly and isoform abundance estimation pipelines, such as the Cufflinks/CuffDiff pipeline (Trapnell et al., 2012), that produce gene expression estimates for isoforms as well as genes or other groups of features, and are useful for transcriptome-wide analysis of isoforms, but require well-annotated genomes.
For some specific type of genes, certain analytic approaches should be taken to account for characteristics that may affect quantification. For instance, many non-coding genes occur in repeated clusters and this can complicated the process of read alignment as a single read can be mapped to more than one location in the genome, introducing ambiguity to its true location. Often these reads have been simply filtered out of the data set but in the case of small ncRNAs, especially those short enough to be fully contained within single reads, this would result in the loss of a large amount of data and under-represent the expression level of these genes. In order to address this, several approaches have been developed, including ERANGE (Mortazavi et al., 2008), Rcount (Schmid and Grossniklaus, 2015), ShortStack (Axtell et al., 2013; Johnson et al., 2016) and HOMER (Heinz et al., 2010). ERANGE, Rcount and ShortStack use similar approaches to locally weight the placement probabilistically based on unique or spliced reads aligned to the locus or surrounding region. ERANGE and Rcount were primarily developed for mRNA analysis, while ShortStack is a sRNA-focused implementation of
this approach, although its focus is towards small regulatory RNAs such as miRNAs, siRNAs and piRNAs. These approaches improve precision but rely on probabilistic assumptions that may not always hold true, such as an active copy from a large, repetitive family located in a gene desert region. An alternative approach is employed by HOMER, which involves analysing repeats in families while ensuring that only one alignment per read is considered for gene expression, utilising the mapping statistics produced from the aligner to select the placement. This is approximate to a certain degree due to inherent uncertainty resulting from nucleotide modifications, polymorphisms or sequencing errors but avoids loss of data or probabilistic approaches.

Figure 2: Parallel analysis of RNA ends (PARE) library preparation (Rackham and Filipovska, 2014)

Many transcriptional processes result in detectable changes to the specific RNAs that they affect and which can be investigated by the use of particular RNA library preparation approaches. The parallel analysis of RNA ends (PARE) method (Figure 2)
was developed to identify the mRNA target sites of miRNAs (German et al., 2009). It utilises the free 5′ monophosphate residue found on the cleaved products of miRNA-targeted mRNAs to attach an adapter containing an MmeI recognition site. Once incubated with MmeI, the cDNA will be cleaved approximately 20 nt downstream of the recognition site producing short sequences whose 5′ terminal nucleotide correspond the 5′ end of the miRNA-cleaved product. After alignment against a reference genome sequence, the location of miRNA targets can be identified. Although developed for this purpose, the principles underlying it mean it can be re-purposed to investigate any other processes that produce 5′ ligation-capable RNA products.

Another important aspect of gene expression is the regulation of RNA by bound proteins and the identification of specific binding sites is an important step in the elucidation of the proteins function. In this respect, many methods have been developed including photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (Hafner et al., 2010) (PAR-CLIP) and digital ribonuclease (RNase) footprinting (Liu et al., 2013). Analysis of RNA-binding by PAR-CLIP involves the ultraviolet (UV)-mediated crosslinking of proteins bound to RNAs, followed by enzymatic degradation of the unbound regions and the incorporation of the nucleotide analogue 4-thiouridine (Figure 3). This procedure induces numerous thymidine (T) to cytidine (C) transitions at the sites bound by protein in the cDNA, and after sequencing and alignment of these RNA fragments binding sites can be identified by the read coverage and T-C transition rate. Digital RNase footprinting is an alternative method for identifying sites on RNA bound to proteins and can be considered a high-throughput version of classical RNase footprinting. Like classical footprinting, it utilises the overlapping cleavage specificities of several RNases in the presence and absence of a protein of interest compared to an undigested sample to identify regions that are more accessible to RNase degradation it its absence. With RNA sequencing and bioinformatics the throughput of this approach can be improved and statistically evaluated.
1.2 The RNA polymerase III transcriptome

In the nucleus of mammalian cells, gene transcription is performed by one of three discrete RNA polymerase complexes (Roeder and Rutter, 1969, 1970). RNA polymerase I (Pol I) constitutively transcribes the large rRNAs genes arranged as clusters of tandem repeats. RNA polymerase II (Pol II) is responsible for transcribing the protein-coding genes into messenger RNAs (mRNAs) as well as several other types of genes, including miRNAs and small nuclear RNAs (snRNAs). RNA polymerase III (Pol III) is targeted to the promoters of tRNAs and a collection of other house-keeping non-coding genes. The three polymerase complexes share a common core of proteins in addition to polymerase-specific factors that mediate interaction with and expression of specific classes of genes (Vannini and Cramer, 2012).
Pol III is a large complex composed of 17 proteins and has been predominantly associated with the constitutive transcription of nuclear tRNAs and several types of ncRNAs, including the ribosomal RNAs (5S rRNA), RNAs involved in splicing (U6 snRNA), RNA maturation (i.e. the nuclear ribonuclease P and mitochondrial RNA-processing endonuclease RNA) and regulation of Pol II transcription (7SK RNA), among others (reviewed in (Nikitina et al., 2011)). Recent studies, however, continue to indicate additional functions of the Pol III complex that have yet to be fully characterised, including unknown new target genes (Canella et al., 2010), involvement in large-scale genome organisation and regulatory functions (Canella et al., 2012). The collection of genes under transcriptional control of Pol III share particular characteristics, including their seeming exclusivity for non-coding genes, the promoter types used and the transcription factors regulating their expression (reviewed in (White, 2011). Multiple mutations have been identified which disrupt Pol III and cause pathology (Bernard and Vanderver, 2012), in particular the two largest protein subunits of Pol III, POLR3A (RNA polymerase III, subunit A) and RNA polymerase III, subunit B, that generally result in hypomyelinating, neurodegenerative diseases (Saitsu et al., 2011).

1.3 Mitochondrial gene expression and RNA metabolism

Mitochondria are ubiquitous eukaryotic organelles and the site of cellular respiration via oxidative phosphorylation (OXPHOS), which provides the cell with the majority of its supply of adenosine triphosphate (ATP). An interesting result of their evolutionary origin as a free-living α-proteobacteria remains evident in their discrete, compartmentalised and co-regulated circular genome. In mammals, this genome has become greatly condensed over evolutionary time and as such exhibits several unique characteristics related to the expression of its genes. The mitochondrial DNA (mtDNA) of humans is a relatively short (~16.6 kb) circular molecule (Figure 4), that contains 13 protein-coding genes (on 11 mRNAs), 2 mitochondrial rRNAs and 22 mitochondrial tRNAs (mt-tRNAs) arranged as two large polycistrons across both strands of the DNA (Anderson et al., 1981). The two strands are denoted the heavy- and light strands according to their different buoyant densities on a caesium chloride gradient which corresponds to their guanine (G)+T and relative gene content (Kasamatsu and Vinograd,
The majority of the mitochondrial proteome are nuclear-encoded proteins that are imported into the mitochondrion, including all the proteins required for mitochondrial gene expression (Pagliarini et al., 2008).

![Figure 4: The mammalian mitochondrial genome (Rackham et al., 2016).](image)

Transcription of the heavy- and light strands of mtDNA begin from their respective promoters, the heavy-strand promoter and light-strand promoter, located in the ~1 kb non-coding region. Mitochondria only possesses two transcription factors (Litonin et al., 2010), unlike the multitude observed in the nucleus. In the early stages of transcription, the promoter is bound by the mitochondrial transcription factor and nucleoid protein, TFAM (transcription factor A, mitochondrial), a member of the high mobility group of box proteins (reviewed in (Gustafsson et al., 2016)). When mtDNA is bound by TFAM, it induces a conformational change in the molecule which promotes the recruitment of the mitochondrial RNA polymerase (POLRMT). The recruitment of POLRMT to the TFAM:mtDNA complex forms a pre-initiation complex (pre-IC) which becomes transcriptionally active upon binding of the second mitochondrial transcription factor, mitochondrial transcription factor B2, which melts the promoter sequence with TFAM (Ramachandran et al., 2016), initiates transcription and dissociates.

After transcription begins, the polymerase complex elongates the nascent RNA molecule as it proceeds along the length of the genome, before terminating at protein-bound sites. POLRMT produces two polycistronic RNA molecules; the heavy-strand
RNA that contains 12 protein-coding genes, 2 rRNAs and 14 tRNAs, which terminates at a site downstream of the distally-encoded gene, tRNA$_{Pro}$, and the light-strand RNA that contains 1 mRNA and 8 tRNAs, and is terminated by mitochondrial transcription termination factor 1 bound to a sequence located within the mt-tRNA$_{Leu(UUR)}$ gene region (Terzioglu et al., 2013). Each polycistronic precursor transcript must be processed for the multiple, contiguous coding sequences to be released and made available for maturation (Figure 5).

Figure 5: Mitochondrial gene expression in animals (Small, Rackham and Filipovska, 2013).
The maturation process modifies and adds additional nucleotides to the RNAs which ensures they are translation-capable and promotes their stability. Mature mRNAs are translated by the mitochondrial ribosome (mitoribosome), a mitochondria-specific complex with a much greater protein:RNA content than the nuclear ribosome, a highly coordinated process with much yet to elucidate. The final stage of RNA metabolism in mitochondria is degradation by a complex formed from the human homologue of the yeast Suv3 protein, the RNA helicase Suv3-like protein 1, polynucleotide phosphorylase proteins (reviewed in (Levy and Schuster, 2016)), a process which requires the presence of a 3′ polyadenosine (poly(A)) extension (Chujo et al., 2012).

1.3.1 Mitochondrial RNA processing

The processing of polycistronic precursor mt-RNAs typically occurs at the junctions of mt-tRNA genes which are cleaved from the surrounding RNA, known as the “tRNA punctuation” model (Ojala et al., 1981), although several exceptions exist. This process is mediated by an RNA-binding protein complex known as RNase P (Holzmann et al., 2008), which cleaves at the 5′ end of mt-tRNAs, and the zinc-binding phosphodiesterase, ELAC2 (ElaC homolog protein 2) which cleaves at the 3′ end of tRNAs (Brzezniak et al., 2011; Lopez Sanchez et al., 2011). In addition to this, there are four junctions which do not harbour mt-tRNA sequences at their 5′ and 3′ ends and are processed by alternative means.

The mitochondrial RNase P complex is composed of three protein subunits, denoted mitochondrial ribonuclease P proteins 1, 2 and 3 (MRPP1, MRPP2 and MRPP3). The MRPP1 protein is an m1G9 methyltransferase of the tRNA methyltransferase homolog C (TRMT10) family that catalyses an evolutionarily conserved modification on mt-tRNAs, the methylation of A- or G-residues at position 9 which contributes to their structural stability (Vilardo et al., 2012). MRPP2 is a short-chain dehydrogenase with a wide substrate specificity but a preference for short-chain fatty acids and forms a methylation-capable active subcomplex with MRPP1. The MRPP3 subunit contains the metallonuclease domain and performs the nucleolytic cleavage of target RNA. This subunit harbours RNA-binding activity due to the presence of pentatricopeptide (PPR) domains in its structure. In mammals, these RNA-binding domains are only found in the mitochondria and all on proteins involved in multiple aspects of mitochondrial RNA
As mt-tRNAs are well distributed between the various other mitochondrial transcripts, in most cases RNase P and ELAC2 are sufficient to release most the interspersed mt-RNAs. However, four mt-mRNA junctions are not bordered by mt-tRNA sequences their processing uses alternate complexes. The regions flanking the 5′ end of CO1 (cytochrome c oxidase subunit 1) and CYTB (cytochrome b) are non-coding and antisense mt-RNA sequences, the 5′ end of CO3 (cytochrome c oxidase subunit 3) is proximal to the 3′ end of the ATP8/6 (ATP synthase subunit 8/ATP synthase subunit 6) bicistronic coding sequence and the 3′ end of ND6 (NADH dehydrogenase subunit 6) flanks a non-coding region. Recent studies have identified proteins of the Fas-activated serine/threonine kinase family as responsible for the processing at these junctions. Processing at the non-canonical heavy-strand junctions was inhibited and transcripts crossing these junctions accumulated upon the knockdown of mitochondrial FAST kinase domain-containing protein 5 (Antonicka and Shoubridge, 2015), whereas processing at the 3′ end of ND6 was specifically lost when the Fas-activated serine/threonine kinase gene was ablated in cultured cells (Jourdain et al., 2015).

Mutations in nuclear and mitochondrial genes have been observed to disrupt mitochondrial RNA processing and result in disease. In the mitochondrial genome, numerous pathological mutations have been identified in mitochondria tRNA genes and several are known to disrupt normal processing. Two adenosine to guanosine transitions mt-tRNA\textsubscript{Leu(UUR)}, m.3243A>G and m.3302A>G, are known to affect RNA processing and result in the accumulation of a stable precursor, known as RNA 19 (Hess et al., 1991; King et al., 1992; Maniura-Weber et al., 2006). While the pathological role of RNA 19 is unknown, both mutations are known to cause pathology. The m.3243A>G mutation is seen in the majority of cases of mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes (reviewed in (Kirino and Suzuki, 2005)). This mutation appears to have multiple effects on the tRNA, including impairment of normal maturation processes such as wobble base modification (Kirino et al., 2004) and aminoacylation (Borner, 2000), as well as structural (Wittenhagen and Kelley, 2002) and translational effects (Sasarman et al., 2008). The m.3302A>G mutation is often associated with mitochondrial myopathy (Bindoff et al., 1993) and is less-well studied, but has been reported to disrupt respiration by impairing efficient processing and
affecting aminoacylation (Maniura-Weber et al., 2006).

1.3.2 Mitochondrial RNA maturation and stability

The three classes of mitochondrial RNAs must undergo post-transcriptional modification before they will become fully functional, including the modification of individual nucleotides and nucleotide additions at their terminal ends. Mt-tRNAs undergo a great number of modifications, which confer structural support and alter codon recognition characteristics. Mt-tRNAs are shorter than nuclear tRNAs and must accommodate the structural characteristics of the mitoribosome, which is facilitated by several types of nucleotide modification. For instance, most mt-tRNAs are methylated at position 9, a modification performed by the MRPP1/MRPP2 subcomplex, which appears to increase its structural rigidity (Vilardo et al., 2012). Since the mitochondrial genome has been so densely compacted the genetic code has become degenerate, requiring mt-tRNAs to recognise an expanded repertoire of codons via modification of the $3^{rd}$ anticodon “wobble” base. The unique structural characteristics of mt-tRNAs also require the modification of the nucleotide located +1 to the $3^{rd}$ anticodon base. Nucleotide modification can also function to alter anticodon recognition. The taurinylation of the wobble base in mt-tRNA$_{Leu(UUR)}$ alters its codon-recognition profile in favour of purines (Kirino et al., 2004), resulting in stronger preference for UUG codons.

All mitochondrial mRNAs are polyadenylated, with the exception of ND6, but to a much lesser extent when compared to nuclear mRNAs (~45 – 55 nt) and unlike their nuclear counterparts mt-mRNAs are not capped at their 5' ends. Polyadenylation is catalysed by the mitochondrial poly(A) polymerase (mtPAP) and is often necessary to complete the terminal stop codon, UAA, owing to the condensation of the genome. The effects of polyadenylation are varied and can both increase and decrease RNA stability, and its role in mitochondrial RNA metabolism is not yet well understood. The leucine-rich pentatricopeptide cassette protein (LRPPRC), and its cognate partner SLIRP (SRA stem-loop-interacting RNA-binding protein), have been shown to increase the processivity of mtPAP and promote mRNA stability, polyadenylation and coordinated translation (Ruzzenente et al., 2012; Wilson et al., 2014). LRPPRC is an RNA-binding protein that consists of a series of up to 30 – 33 PPR domains along its length that is
stabilised in a complex by the binding of SLIRP through a novel protein-protein interaction mediated by PPR and RNA recognition motif domains (Spåhr et al., 2016). The LRPPRC/SLIRP complex shows a broad sequence specificity its loss in cells results in decreased mt-mRNA stability, impaired polyadenylation and dysregulated coordination of translation. Mutations that disrupt LRPPRC are known to cause a rare French-Canadian form of Leigh syndrome with a strongly associated cytochrome c oxidase (COX) deficiency (Merante et al., 1993; Morin et al., 1993). Patients with this form of Leigh syndrome present with metabolic symptoms such as lactic acidosis in addition to recurrent neurological episodes and degeneration involving the brain stem and basal ganglia structures (Debray et al., 2011).

Disruption of mitochondrial RNA metabolism any stage can result in disease, due to the highly coupled nature of gene expression mechanisms in mitochondria. Mutations that disrupt mt-tRNA excision (Bindoff et al., 1993), mt-rRNA (Cotney et al., 2009) and mt-tRNA modification (Kirino et al., 2004), polyadenylation and RNA stability (Chujo et al., 2012) have all been observed to result in pathological states. In many cases, however, the exact mechanisms underlying these defects or the downstream effects that result in disease are poorly understood. The elucidation of these mechanisms will require considerable study of multiple mitochondrial processes and their interactions in order to deconvolute the direct and indirect effects of their disruption.
1.4 Aims of Study

This research aims to investigate the proteins and processes involved in the expression and regulation of transcriptomes by analysing cellular and animal models of disease using RNA sequencing and bioinformatics approaches. The specific aims were:

I. To investigate the transcriptome-wide effects on gene expression of a pathogenic nuclear mutation causing a neurological disease.

II. To elucidate the role of pathogenic mutations in a gene causing mitochondrial disease in cell models.

III. To identify the mechanism of action of RNA-binding proteins in animal models of disease.
2. Results and Discussion

2.1 Paper I

Transcriptome-wide effects of a POLR3A gene mutation in patients with an unusual phenotype of striatal involvement

The effects of a pathogenic mutation affecting the POLR3A gene were investigated in three patients with a neurological condition displaying an unusual phenotype. Three patients from two unrelated families from the same endogamous Roma subisolate were assessed and found to possess nearly identical symptoms including dystonia, dysarthria and gait ataxia that developed around the age of seven or eight years. Dentition and gonadal function were normal. Examination by MRI revealed atrophy and signal changes in the caudate nucleus, putamen and red nucleus, with no white matter involvement unlike other pathogenic POLR3A mutations. Whole exome sequencing was performed using the TruSeq capture method and pathogenic variants were searched for using a step-wise filtering procedure for polymorphisms from the HapMap collection, under the assumption of autosomal recessive inheritance. A single homozygous variant was identified that fulfilled all criteria, c.1771-6C>G, located 6 nt upstream of the junction between intron 13 and exon 14, in a large shared region of homozygosity on chromosome 10. The frequency of this variant in the Roma population was estimated as 1.4%, compared to about 5% in the subisolate examined. Predictive analysis of the variant suggested possible disruptive effects on a SRp40 exonic splicing enhancer element located in intron 13, which was confirmed with RT-PCR and sequencing revealing an aberrant POLR3A product skipping exon 14 with a premature stop codon subject to nonsense-mediated decay and lowered full-length POLR3A transcript levels.

I investigated the effects of the c.1771-6C>G mutation on gene expression were analysed by TruSeq and small RNA-Seq of RNA obtained from the three patients and three healthy controls. Small RNA analysis identified many classes of small regulatory ncRNAs and the most significant changes were observed in the tRNA and snRNA families. My analysis revealed a modest decrease in transcript levels for most tRNA isoacceptor families but most significantly affecting tRNAs Gly, Met, Leu, Lys and His. Other Pol III transcripts also appear to be dysregulated as we observed a reduction in 7SL RNA, a component of the signal recognition particle that is involved in co-
translational translocation (Akopian et al., 2013), in the patients, but an increase in other transcripts involved in metabolism of rRNAs and tRNAs and in RNA polymerase regulation, such as the 7SK RNA, RNase P RNA and MRP RNA. Nuclear RNase P and MRP are closely related catalytic ribonucleoproteins (RNPs) that share protein components and are involved in tRNA and rRNA metabolism. RNase P has also been previously reported to interact with and exert regulatory effects on transcription by both Pol I and Pol III. 7SK RNA is a component of the positive transcription elongation factor b, which couples transcriptional elongation of Pol II with alternative splicing (Barboric et al., 2009), indicating a potential mechanism for general disruption of mRNA production. An increase in 5S rRNA was also observed that may be a compensatory response to low tRNA levels and preferential recruitment of Pol III to these loci, further showing a change in regulation and not just a reduction in expression resulting from decreased polymerase.

Next, I investigated the transcriptome-wide changes of overall gene expression to elucidate downstream effects of the mutation. The results showed a reduction in the expression of genes encoding proteins that have been implicated in neurological disorders involved in transcriptional and post-transcriptional processes including pre-mRNA splicing and proteostasis. These include the heterogeneous nuclear ribonucleoprotein H2, the heat shock proteins DnaJ homolog subfamily A member 1 and DnaJ homolog subfamily B member 2, heat shock protein 90kDa and ubiquitin. An increase in several snRNAs and small nucleolar RNAs, transcribed by Pol II and involved in splicing and nuclear RNA metabolism, was also observed. Small nucleolar RNAs have been previously implicated in neurodegenerative disease (Sahoo et al., 2008) and their promoters have been observed to interact with the 7SK ribonucleoproteins, mediated by the RNA helicase, DEAD box protein 21 (Calo et al., 2014).

These results characterise c.1771-6C>G as a pathogenic mutation that disrupts an exonic splicing enhancer element causing aberrant splicing of the POLR3A transcript and reduced levels of the full length transcript. These patients showed no involvement of white matter in the brain, or defects in dentition or hormonal regulation, instead displaying atrophy of the striatum and red nucleus. RNA sequencing analysis revealed an overall decline in tRNA levels and dysregulated expression of other Pol III
transcribed genes that participate in transcriptional and translation regulation of Pol II and Pol III. In addition, ncRNAs involved in tRNA and rRNA maturation and protein-coding genes that regulate pre-mRNA splicing and proteostasis were affected, many of which have been associated with other neurodegenerative diseases. Together, these results suggest that impaired Pol III transcription leads to an imbalance of tRNAs and ncRNAs that has downstream effects on Pol II and Pol III target genes involved in RNA metabolism and proteostasis.

Such differential regulatory effects of RNA polymerase have also been identified in the mitochondria in an analysis of Polrmt knockout mice generated by the Larsson group, where I used my new bioinformatics pipelines as a basis to develop similar techniques to characterise differential heavy and light strand regulation (Kühl et al., 2016, and appendix A). This study sought to investigate the role of POLRMT in the generation of replication primers from the light strand of mtDNA, and as part of that investigation I analysed the effects of Polrmt loss on differential gene expression of mitochondrial preparations isolated from heart. I identified that the expression of heavy strand transcripts were found to be more severely affected by the loss of Polrmt than for their light strand counterparts, indicating that under low concentrations the preservation of replicative ability takes priority over gene expression in mitochondria. These results helped to elucidate mechanisms of coordinated regulation of gene expression and DNA replication in mammalian mitochondria.
Defects in mitochondrial RNA metabolism in disease

The processing of mitochondrial RNAs is not fully understood, but an essential step in the expression of proteins required for OXPHOS and disruption of this process causes disease. I investigated the effects of two known pathogenic mt-tRNA mutations, m.3243A>G and m.3302A>G in mt-tRNA\textsuperscript{Leu}\textsubscript{UUR}, that have been reported to affect RNA processing (Bindoff et al., 1993; Rossmanith and Karwan, 1992) using PARE to profile the location and relative abundance of the precise 5′ ends of mitochondrial RNAs in mutant cell lines. Both of these mutations have been shown to result in the accumulation of a long-lived RNA precursor known as RNA 19 (Bindoff et al., 1993; Koga et al., 1995), which comprises the 16S mt-rRNA, mt-tRNA\textsuperscript{Leu(UUR)} and MT-ND1 coding regions in a contiguous sequence (see Figure 4), and investigations of their effects on processing efficiency have shown similarities and differences in their effects on mt-tRNA\textsuperscript{Leu(UUR)} cleavage (Rossmanith and Karwan, 1992). My results showed a marked reduction in the abundance of sequenced reads that mapped to the 5′ ends of MT-TL1 and MT-ND1, the two internal junctions within RNA 19, in both mutants which is consistent with previous observations of 3′ as well as 5′ defects resulting from these mutations. Furthermore, the abundance of MT-ND1 5′ end reads was lower than for MT-TL1 in both mutants, a situation that is reversed in the control suggesting a greater 3′ impairment or compounding downstream effects.

No accumulation of PARE reads was observed at the non-canonical 5′ junctions of MT-CO1 and MT-CYB, which suggested to that these RNAs may lack the 5′-monophosphate required for PARE capture or the 5′ terminal nucleotide is occluded by protein-binding or the presence of secondary structure, whereas a peak was observed at the 5′ end of MT-CO3 in the m.3243A>G mutant only. To investigate this further, 5′-phosphate-dependent exonuclease digestion was performed on isolated mitochondrial RNAs from 143B cells, which revealed that while MT-ATP8/6 and MT-CO2 were susceptible to degradation the MT-CO3 transcript was resistant, suggesting it lacked a 5′ terminal monophosphate residue. The appearance of captured 5′ ends in the m.3243A>G sample, while absent in the control and m.3302A>G mutant, suggests downstream effects of the m.3243A>G that influence the junction at MT-CO3 possibly by altering secondary
structure. The 3’ end of MT-ND6 is also a non-canonical junction that, while not directly assayed by PARE, its cleavage will produce downstream RNAs from the precursor with 5’-monophosphates that can be detected. An accumulation of PARE reads was detected whose 5’ ends mapped to a position 501 nt downstream of the MT-ND6 open reading frame (ORF) in all samples, and possessed 3’ non-template adenosine residues. Since polyadenylation and degradation are linked, these may be an indication that this junction is processed by degradation processes as recently suggested (Jourdain et al., 2015).

In addition to changes at known junctions, I searched for novel sites of prominent 5′ PARE abundance outside of known RNA junctions. Multiple such sites were found in the m.3243A>G mutant only, further suggesting the ability of this mutation to cause downstream effects outside of the processing of mt-tRNA\textsuperscript{Leu\textsubscript{UUR}}. These include a cluster of sites observed approximately centrally within the MT-CO1 coding region, and single sites within the MT-ND4 and MT-RNRI gene regions. None of these were observed in the m.3302A>G mutant. However, we also identified several new sites that appeared in all three samples in MT-CO2, and in non-coding regions of the light strand transcript. The MT-CO2 site is seen towards the middle of the MT-CO2 coding region and appears to be unaffected by the reduction of abundance at the 5′ end of MT-CO2 itself, while the ratio of internal to 5′ end reads was increased in the mutants. This implies that the internal cleavage occurs independently of the cleavage at the 5′ end of the containing transcript, and thus occurs on the unprocessed precursor containing MT-CO2. The sites appearing on the light strand include a cluster located downstream of the WANCY tRNA cluster and another close upstream of the MT-ATP8/6-MT-CO3 junction region. These may be the result of RNA produced by promiscuous processing enzymes that cleave sites with strong secondary structure or are derived from processing intermediates of the light strand.

This study has investigated the effects of two pathogenic mt-tRNA mutations on RNA processing in the mitochondrial transcriptome by mapping the 5′ ends of mt-RNAs transcriptome-wide. Both mutations had reduced abundance of reads aligned to the 5′ ends of MT-TL1 and MT-ND1, consistent with impaired processing of RNA 19, and a greater proportionate effect on MT-ND1 5′ end abundance. Processing at non-canonical 5′ sites was only detected in the m.3243A>G for MT-CO3 which was resistant to 5′-
phosphate-dependent exonuclease degradation, suggesting this transcript lacks a 5′-monophosphate and that its normal processing is disrupted by this mutation allowing its detection. *MT-ND4* and *MT-RNRI* also harboured 5′ end peaks that did not correspond to any known transcripts, implying that this disruption may be widespread. PARE reads with 5′ terminal ends mapped to a location ~500 nt downstream of the *MT-ND6* ORF containing 3′ non-templated adenosines were seen in all samples and may be products of the degradation-dependent 3′ processing as has been suggested for this junction and another prominent peak was observed in the *MT-CO2* coding region. These results reveal the effects of the m.3243A>G mutation on global RNA processing in mitochondria and describes novel sites of RNA processing.

To elucidate the order of events in mitochondrial RNA processing, I analysed global RNA processing *in vivo* using *Mrpp3* knockout mice, generated by the Filipovska group, which showed a severe impairment of RNA processing (Rackham et al., 2016, and appendix B). The computational analysis used paired end RNA sequencing data to elucidate the hierarchy of processing events with implications for mitoribosome assembly. The preparation of RNA used a library preparation protocol that retains long fragments is advantageous for the analysis of mitochondrial RNA processing as small RNAs, such as tRNAs, are not captured and given the polycistronic organisation of the mitochondrial genome the coverage we observe across these regions can be taken to derive from processing intermediates. Combined with PARE, it was found that the 5′ processing of tRNAs precedes 3′ processing and that the loss of RNase P was unable to be rescued by a moderate increase of the 3′ processing enzyme, ELAC2.

I applied similar bioinformatics techniques to identify changes in the relative precursor abundance in human cells, from infant patients who died from severe mitochondrial disease due to mutations in the m1G9 methyltransferase component of RNase P, MRPP1 (Metodiev et al., 2016, and appendix C). These patients presented with symptoms of lactic acidosis, hypotonia, feeding difficulties and deafness however the abundance of mitochondrial RNAs was not found to be substantially different compared to controls, despite the moderate increase in the long-lived precursor RNA 19 suggesting that these mutations affected RNA processing. Using computational analyses, I identified an increase in the relative abundance of covering from processing intermediates across the transcriptome that did not affect the tRNA methyltransferase activity of MRPP1, which
could be observed by the unchanged mismatch rate at positions corresponding to the nucleotide modified by this protein. These investigations illustrate how the use of bioinformatics allows the identification of subtle transcriptome-wide changes not easily detectable by other means as a diagnostic tool to identify dysfunctional processes in disease.
2.3 Paper III

Protein-mediated RNA-folding in a massively reduced transcriptome

The life cycle of mRNAs is dependent upon interaction with a wide range of RNA-binding proteins, which modulate their transcription, processing, translation and decay. I analysed the in vivo binding sites of the RNA-binding protein LRPPRC and its binding partner SLIRP, using digital RNase footprinting and PAR-CLIP to reveal a conserved preference for mRNAs and effects on the secondary structure and stability of the transcriptome upon its loss in mouse and human. The effects of its loss on secondary structure indicate the importance of RNA-protein interactions in the folding of mitochondrial RNAs and suggests that LRPPRC binding is important for RNA maturation, stability and translation.

Digital RNase footprinting involves the treatment of isolated mitochondria with endonucleases RNase A, T1 and If that cleave with different specificities, from which RNA is extracted and sequenced along with a mock-digested control. After aligning the sequenced reads from all samples to the mouse mitochondrial genome, the RNase accessibility for each position was determined according to a cleavage score (C score) that describes the abundance of 5′ ends in RNase treated versus untreated samples. Footprints are identified by scanning the genome for short regions with an average C score lower than its surrounding flanking regions and calculating their footprint score (F score) that describes the depth of the footprint relative to its flanks. Footprints from the control mice were compared to the same region from knockout mice and the change in F score compared against an empirical null model obtained from random shuffling of the C score in the genome.

The binding of the pentatricopeptide repeat containing domain protein 1 (PTCD1) has previously been investigated in human cells lines with this technique (Liu et al., 2013, and appendix D) since this protein is a member of the PPR family of proteins which have important roles in the regulation of mitochondrial gene expression, and PTCD1 functions in a coordinated manner with other enzymes such as LRPPRC to release and mature nascent mitochondrial RNA. The expression of PTCD1 was knocked down using siRNA and identified binding site footprints by comparison to a mock-treated
control, which revealed five PTCD1-specific RNA footprints at the 3’ end of tRNAs validating the role of PTCD1 in mt-tRNA metabolism (Rackham et al., 2009). Given the apparent lack of nuclease activity of PTCD1 (Rackham and Filipovska, 2011), this protein may function via direction interaction with RNA to regulate or facilitate the cleavage of mt-tRNAs in conjunction with ELAC2.

A similar mechanism may be at play with the LRPPRC/SLIRP complex, which is almost entire composed of PPR domain. Using conditional Lrpprc and Slirp knockout mice generated by the Larsson group, I analysed digital RNase footprinting data from isolated mitochondria to investigate its binding mode. The analysis found 178 footprints throughout the transcriptome, concentrated on mRNAs and rRNAs, with few found on tRNAs or in the regulatory D-loop region, which is consistent with the reported role of LRPPRC in polyadenylation and stability (Ruzzenente et al., 2012). The footprints were particularly localised on mRNAs and rRNAs, indicating that mt-tRNAs are not direct targets of LRPPRC in vivo, consistent with the roles of LRPPRC in polyadenylation and translation. These results suggest that LRPPRC binds to regions of mRNAs involved in RNA stability and ribosome recruitment. It was recently shown that RNA processing and ribosome assembly to occur co-transcriptionally (Rackham et al., 2016, and appendix B), and the presence of footprints across the two rRNAs suggesting that LRPPRC may act as an RNA chaperone to facilitate RNA recognition by ribosomal proteins. The footprinting analysis of the Slirp knockout revealed no significant footprint sites, indicating that SLIRP does not bind RNA in vivo as has been recently reported (Spåhr et al., 2016).

To examine downstream changes in secondary structure or RNA-bound sites, I used the footprinting pipeline to search for footprints in the Lrpprc knockout mice data and identify regions of reduced RNase accessibility in the absence of LRPPRC compared to the control. This identified 124 footprints that indicate the presence of RBPs or secondary structural changes that render the site resistant to cleavage, far fewer than the number of LRPPRC footprints found, suggesting a general RNA chaperone role. Analysis of Slirp knockout mice did not identify any significant footprints. Next, I calculated an R score for each position along the genome to describe its propensity for cleavage by double-stranded over single-stranded RNases. It was found that loss of LRPPRC results in a net increase in the secondary structure of mitochondrial transcripts.
and that regions bound by and flanking LRPPRC were less structured in its presence. These results suggest that the LRPPRC/SLIRP complex acts to disrupt secondary structure formation within mitochondrial RNAs and expose single-stranded RNA for modification or translation.

To further validate the binding sites of LRPPRC, I analysed PAR-CLIP data from mouse embryonic fibroblast cells obtained from Lrpprc-FLAG transgenic mice from the Larsson group. Numerous binding sites were identified, with the majority concentrated on mRNAs and to a lesser degree in rRNAs. Most monocistronic transcripts showed a stronger from the 3’ ends suggesting that binding may be required for polyadenylation in these cases, promoting RNA stability and translation. Binding was also prominent within bicistronic transcripts, towards the internal start and stop codons, which could indicate that LRPPRC exposes these regions to the mitoribosome to regulate translation initiation and termination. Binding sites were also found towards the 5’ and 3’ ends of 16S rRNA, indicating a potential role in the maturation of rRNA, the assembly of the ribosome and translation. To compare the evolutionary conservation of LRPPRC/SLIRP binding between mouse and human mitochondria, I performed an identical analysis on data from HeLa cells expressing FLAG-tagged LRPPRC. The results revealed a highly similar binding profile across both species and show that the location of preferred binding sites is maintained through evolution.

The lack of identified SLIRP footprints is consistent with previous investigations of the role of SLIRP in ribosomal RNA fractions that provided evidence that SLIRP is involved in regulating the coordinated translation of mRNAs (Lagouge et al., 2015, and appendix E). In that study, I used differential expression analysis tools to examine sucrose gradient fractions containing mitochondrial RNAs bound to ribosomal proteins to identify changes in translating transcripts. This revealed a reduction in the relative abundance of mRNAs, except mt-Nd6, associated with ribosomal subunits, in cells from Slirp knockout mice from the Larsson group that was most prominent for the fractions containing unassembled ribosomal subunits. This was consistent with the observed reduction in translation rate and in line with reduced mitoribosome engagement, while polyadenylation appeared unaffected. Taken together, these implicate the SLIRP component of the LRPPRC/SLIRP complex in regulating the association of mRNA with the mitoribosome for translation, and the localisation of LRPPRC footprints within
mRNAs and rRNA supports this hypothesis.

Over evolutionary time, the regulation of the mitochondrial transcriptome has shifted to one predominantly carried out by RNA-binding proteins. In the mitochondria, PPR proteins are prominently involved in RNA metabolism and the LRPPRC/SLIRP complex has been implicated in polyadenylation, RNA stability and the coordination of translation. This work has revealed widespread binding across the transcriptome predominantly located towards the 5′ and 3′ ends of mRNAs and rRNA, and indicate that LRPPRC binding disrupts secondary structure that may be important for exposing RNAs for maturation and translation.
2.4 General Discussion

The field of transcriptomics is expanding and continuously developing new methods to investigate the functional relationship between regulators of gene expression. These new approaches require the analyses of large-scale transcriptome-wide data sets using computational methods. Given the flexibility of short read technology, different preparations can yield libraries with populations of RNA species geared towards the analysis of specific classes of RNAs or processes in RNA metabolism. Elucidating the mechanisms of RNA-binding proteins that regulate this metabolism is essential to understand how mutations in these genes exert their pathogenic effects. Through the use of several transcriptomic approaches, I have investigated the effects of disease-causing mutations on gene expression and RNA metabolism in mammalian cells.

RNA polymerase III is central to the expression of the genome due its role in transcribing tRNAs required for translation and regulatory ncRNAs that modulate the transcription, maturation and translation of transcripts. I examined the transcriptome-wide effects on gene expression resulting from a disease-causing mutation in its largest subunit, POLR3A, in order to elucidate genes and processes involved in its pathogenesis. A significant imbalance of Pol III transcripts was observed, including a general decrease in tRNAs and mixed increases and decreases for other ncRNAs, supporting previous hypotheses regarding altered availability of Pol III transcripts in patients with POLR3A mutations (Bernard et al., 2011) and indicating that the loss of Pol III changes the regulatory balance of its target genes. Furthermore, the downstream effects on genes involved in pre-mRNA splicing, rRNA maturation and proteostasis implicate these processes in the pathology of this mutation. These patients were initially assessed as having probable mitochondrial disease due to the effect on basal ganglia structures, which are a common feature of primary mitochondrial disorders, however our analysis found a nuclear mutation was responsible, highlighting the utility of genome- and transcriptome-wide analyses for elucidating the cause and consequences of pathological mutations.

Expression of the mitochondrial genome is an often overlooked aspect of cellular biology, but is essential for the production of OXPHOS proteins and sufficient ATP levels for the entire cell. In contrast to the nuclear genome, mitochondrial transcription
is carried out by a single RNA polymerase, POLRMT, which possesses an RNA-binding PPR domain and which we found regulates the switch between replication and transcription through the generation of RNA primers (Kühl et al., 2016, and appendix A). The association of mt-tRNA mutations with disease is strong and disrupted RNA processing is frequently observed (Koga et al., 2003). I performed PARE to map the precise location of mitochondrial 5′ ends in mutant cell lines harbouring two A-G transition mutations within mt-tRNA\textsubscript{LabUUR} and examine their effects on RNA processing, using techniques we also applied to investigate Mrpp3 knockout mice (Rackham et al., 2016 and appendix B) and cells from patients with MRPP1 mutations (Metodiev et al., 2016, and appendix C). This work revealed similar effects on the processing of this particular tRNA, consistent with previously reported effects on the accumulation of RNA 19, but divergent effects on downstream processing. In particular, numerous sites of RNA cleavage were observed only in the m.3243A>G sample and the 5′ end of MT-CO3 was detected despite 5′-phosphate-dependent exonuclease digestion experiments suggesting the lack of a 5′ terminal monophosphate on this transcript. In addition, several novel sites were also observed within mRNA, rRNA and non-coding regions that correspond to no known transcript ends but were seen in all samples and are close in proximity to regions of predicted secondary structure. These results demonstrate the utility of this 5′ end capture protocol for investigating the effects of disrupted RNA processing in mitochondria and reveal varied effects of two mutations affecting the same tRNA gene on global RNA processing, which may contribute to their different pathology.

RNA processing in mitochondria releases the nascent transcripts from polycistronic precursors and makes them available for maturation. For mRNAs, this requires the addition of a poly(A) tail to mRNAs, mediated by mtPAP, which in many cases completes the terminal stop codon of the ORF. The association of the LRPPRC/SLIRP complex with efficient polyadenylation and RNA stability has already been reported (Lagouge et al., 2015; Ruzzenente et al., 2012), however its binding position along mitochondrial transcripts has not been established. I analysed a combination of digital RNase footprinting, which utilises RNases digestion to identify increases in accessibility in the absence of your protein of interest that we have previously applied to mitochondrial RNA-binding proteins (Liu et al., 2013, and appendix D), and PAR-CLIP data, which uses the incorporation of 4-thiouridine to induce T-C transitions at bound
sites and mark these regions for identification, to examine the binding profile of LRPPRC/SLIRP and examine structural changes. Many LRPPRC footprints were found distributed across the transcriptome, concentrated on mRNAs and rRNAs which was consistent with the bound regions identified by PAR-CLIP, while none were found for SLIRP confirming its lack of RNA-binding activity previously described (Spåhr et al., 2016). An overall increase in the propensity for cleavage by a double-stranded over a single-stranded RNase in the \textit{Lrpprc} knockout revealed that in the absence of protein the footprint regions and surrounding flanks were more structured. This shows that the binding of LRPPRC/SLIRP complex disrupts secondary structure formation and the widespread distribution with a preference towards transcript ends suggests that this serves to expose surrounding regions for modification or translation. This supports other investigations into SLIRP and its role within the LRPPRC/SLIRP complex, which revealed that SLIRP is required for proper association of mRNAs to the mitochondrial ribosome, but dispensable for polyadenylation (Lagouge et al., 2015, and appendix E). The presence of PAR-CLIP sites and footprints within the rRNAs also indicate a potential role in the stability and maturation of rRNA and mitoribosome assembly given these processes occur co-transcriptionally (Rackham et al., 2016, and appendix B).

My investigation of the transcriptome-wide effects on gene expression that result from pathogenic and RNA-binding protein mutations used RNA sequencing and computational analysis to characterise alterations in transcription, RNA processing and RNA stability in nuclear and mitochondrial transcriptomes. This included the differential expression analysis of protein-coding and non-coding genes in total and small RNA sequencing libraries, the analysis of RNA processing in mitochondria with PARE and RNA-Seq and the identification of mitochondrial RBP binding sites with RNase footprinting and PAR-CLIP, carried out with RNA samples obtained from patient cells, cell lines and animal models. The findings have revealed evidence for dysregulation of genes involved in the coordinated regulation of transcription and translation of nuclear RNA polymerases in patients with a POLR3A mutation, compared and contrasted the downstream effects observed in cell lines harbouring tRNA mutations on RNA processing, localised the binding sites of the LRPPRC/SLIRP complex and characterised the effects of its loss on RNA structure. They provide valuable insight into the mechanisms of gene expression in the nucleus and mitochondrion, providing valuable insight into their \textit{in vivo} function and physiological
importance for health and disease.
References


Saitsu, H., Osaka, H., Sasaki, M., Takanashi, J., Hamada, K., Yamashita, A.,


Paper I
Transcriptome-wide effects of a POLR3A gene mutation in patients with an unusual phenotype of striatal involvement

Dimitar N. Azmanov, Stefan J. Siiro, Teodora Chamova, Ara Kaprelyan, Velina Guerguelcheva, Anne-Marie J. Shearwood, Ganqiang Liu, Bharti Morar, Oliver Rackham, Michael Bynevelt, Margarita Grudkova, Zdravko Kamenov, Vassil Svechtarov, Ivailo Tournev, Luba Kalaydjieva and Aleksandra Filipovska

1Harry Perkins Institute of Medical Research and Centre for Medical Research, Level 7, QQ Block, QEII Medical Centre, 6 Verdun Street, and School of Chemistry and Biochemistry, The University of Western Australia, Crawley, WA 6009, Australia
2Department of Diagnostic Genomics, PathWest, Perth, WA 6009, Australia
3Sofia Medical University, Faculty of Medicine, Department of Neurology, University Hospital Alexandrovskia, Sofia, Bulgaria
4Department of Neurology, Medical University, Varna, Bulgaria
5Neurogenomics Lab and Parkinson Personalized Medicine Program, Harvard Medical School and Brigham & Women’s Hospital, Cambridge, MA 02139, USA
6Department of Surgery and School of Medicine, The University of Western Australia; Sir Charles Gairdner Hospital, Neurological Intervention and Imaging Service of WA
7Sofia Medical University, Faculty of Medicine, Department of Internal Medicine, Endocrinology Clinic, University Hospital Alexandrovskia, Sofia, Bulgaria
8Sofia Medical University, Faculty of Dental Medicine, Department of Oral and Maxillofacial Surgery, Sofia, Bulgaria
9Department of Cognitive Science and Psychology, New Bulgarian University, Sofia, Bulgaria

*These authors have contributed equally to the work.

Correspondence to: Aleksandra Filipovska. Address: The School of Chemistry and Biochemistry and Harry Perkins Institute for Medical Research, Level 7, QQ Block, QEII
Medical Centre, 6 Verdun Street, Nedlands WA 6009, Australia. Email:

aleksandra.filipovska@uwa.edu.au

Keywords: RNA polymerase III, non-coding RNAs, tRNAs, Roma/Gypsy founder mutation
Abbreviations: POLR3A, RNA polymerase III (DNA directed) polypeptide A; RNA-Seq, RNA sequencing
Abstract

RNA polymerase III is essential for the transcription of non-coding RNAs, including tRNAs. Mutations in the genes encoding its largest subunits are known to cause hypomyelinating leukodystrophies (HLD7) with pathogenetic mechanisms hypothesised to involve impaired availability of tRNAs. We have identified a founder mutation in the POLB3A gene that leads to aberrant splicing, a premature termination codon and partial deficiency of the canonical full-length transcript. Our clinical and imaging data showed no evidence of the previously reported white matter or cerebellar involvement; instead the affected brain structures included the striatum and red nuclei with the ensuing clinical manifestations. Our transcriptome-wide investigations revealed an overall decrease in the levels of Pol III-transcribed tRNAs and an imbalance in the levels of regulatory ncRNAs such as small nuclear and nucleolar RNAs (snRNAs and snoRNAs). In addition, the Pol III mutation was found to exert complex downstream effects on the Pol II transcriptome, affecting the general regulation of RNA metabolism.
INTRODUCTION

Mutations in POLR3A, the gene encoding the largest subunit of RNA polymerase III (Pol III), have been identified as the cause of Hypomyelinating Leukodystrophy 7 (HLD7) (MIM607694) (reviewed in (1)). The disorders classified under this heading show substantial inter- and intrafamilial phenotypic variation and, in addition to the major manifestations of white matter involvement, may present with impaired dentition and hypogonadotropic hypogonadism. No cases of homozygosity for null mutations have been reported; all known defects are predicted to result in partial POLR3A deficiency.

Pol III is composed of 17 protein subunits and POLR3A and POLR3B are the largest subunits required for the catalytic activity of this enzyme complex. In the classical concept of eukaryotic transcription, Pol III is responsible for the constitutive transcription of housekeeping genes including transfer RNAs (tRNAs) and several non-coding RNAs (ncRNAs) (2). However, an increasing number of recent studies point to the complexity of the Pol III transcriptome, its multiple target ncRNA genes of unknown function, possible roles in genome organisation and epigenetic regulation, and interaction with Pol II (3-6). Hypotheses on the pathogenetic mechanisms in HDL7 include dysregulation of the expression of tRNAs leading to perturbed cytoplasmic protein synthesis in the brain (1), as well as impaired function of Pol III-transcribed ncRNAs important for myelin development and maintenance (7).

Here we describe a founder mutation in the POLR3A gene in patients of Roma ethnicity who presented with an unusual clinical and neuroimaging phenotype. Our analysis of gene expression revealed wide-ranging changes that affect the balance of Pol III-transcribed tRNAs and ncRNAs, as well as the expression of some Pol II target genes involved in transcription, splicing and translation regulation.
RESULTS

A unique disease phenotype caused by a new founder mutation in the POLR3A gene

The three patients belonged to two unrelated families from the same endogamous Roma group (Figure 1). Written informed consent was obtained from all participants. The study complied with the ethical guidelines of the institutions involved.

Pregnancy and delivery were uneventful in all cases. Early psychomotor development was normal. All patients completed year 8 at school. The individual clinical course of the disorder is described below. The findings at the most recent examination are summarised in Table 1.

Patient 1 (ShII-1)

A 56-year old male patient with clinical onset at the age of 8 years, when dysarthria became obvious. Gait ataxia (wide-based gate with frequent falls) developed at age 11 years. Stiffness in the lower limbs developed a year later. These complaints progressed over time, leading to severely impaired ambulation at the age of 50 years. At the last examination, at age 52 years, ataxia of stance and gait was prominent (SARA 24). Gaze-evoked nystagmus and impaired smooth pursuits were present. Tendon reflexes were brisk at the arms, with bilaterally positive Babinski sign. Extrapyramidal signs included symmetric rigidity and dystonic movements in the right hand and lower limbs, combined with cogwheel phenomenon in the right wrist. Severe speech disturbance, due to a combination of dystonia and cerebellar dysarthria, was present. Brain MRI showed a mild small brain, with bilateral symmetric atrophy of the caudate nucleus and putamen and associated increased signal. The region of the medial red nucleus and the intra-axial course of the 3rd cranial nerve showed focal symmetric signal change. The white matter was of normal volume and signal. No evidence of oligodentia or hypogonadotropic hypogonadism was found. The IQ at 52 years was 95.
Patient 2 (IzII-1)

A 31-year old male, who presented with initial speech disturbances at the age of 7 years. At 8 years, dystonic movements in the lower limbs, more pronounced on the right, became apparent. At age 10-11 years, the gait became unstable and he started having problems with swallowing. Ataxia, dystonia and dysarthria progressed over time. The most recent examination at 27 years showed moderate ataxia of stance and gait (SARA 15). Tendon reflexes were normal, but an upgoing extensor plantar response was observed. Symmetric rigidity and dystonic movements in the lower limbs, more pronounced on the right, and cogwheel phenomenon in the right wrist were found. Moderate dysarthria, as well as dysphagia with preserved pharyngeal reflexes, were present. Limb deformities, such as bilateral pes cavus were observed. Brain MRI showed bilateral symmetric atrophy and increased signal of the caudate nucleus and putamen (Figure 2A and B), with prominence of the lateral ventricular frontal horns as a consequence, and focal bilateral symmetric signal change in the region of the medial red nucleus intra-axial course of the 3rd cranial nerve (Figure 2C). Positron emission tomography - computed tomography with fluorodeoxyglucose showed dramatic bilateral hypometabolism in the striatum (Figure 2D) compared to that of a control age and sex matched individual (Figure 2E). The white matter was normal. Dentition and the function of the gonadal axis were normal. The IQ was 85.

Patient 3 (IzII-2)

A 27-years old male with clinical onset at 8 years, when dystonic movements in the distal parts of the lower limbs became apparent. At age 14 years, his speech became dysarthric and the patient started having swallowing problems. Gait ataxia was developed at 18 years. At the most recent examination at age 23 years, very mild ataxia of stance and gait were present (SARA 6). Normoreflexia was present in the four limbs, but an upgoing extensor plantar response was elicited. Symmetric dystonic movements in the distal parts of the lower limbs
were observed. The dysarthria and dysphagia were mild, the pharyngeal reflexes were preserved. Bilateral pes cavus was observed. Upon brain MRI, the caudate nucleus and putamen showed mild changes. There was abnormal focal signal change in the medial red nucleus and the intra-axial course of the 3rd cranial nerve. The white matter was normal. The IQ was 80.

We performed whole exome sequencing (WES) using the TruSeq capture system and the HiSeq2000 platform on DNA samples from the three affected subjects and the parents from family Iz (Figure 1). Based on pedigree structure, origin from a genetic isolate, and sharing of an identical rare phenotype, we assumed autosomal recessive inheritance with an ancestral deleterious mutation present in the homozygous state in all three patients and heterozygous in the parents. The WES data were analysed as described previously (8, 9).

In a preliminary analysis, we used 64,973 polymorphisms at HapMap Phase II SNP positions, extracted from the WES data (10), to estimate inbreeding (11, 12) and perform homozygosity mapping (13). The inbreeding coefficients ranged between 0.056 and 0.085, suggesting cryptic consanguinity and supporting our assumption of autosomal recessive inheritance. Homozygosity mapping identified 24 regions of homozygosity shared by the affected subjects.

Our step-wise filtering strategy (8, 9), driven by the inbreeding estimates, identified a single homozygous variant exome-wide that satisfied all criteria (Figure S1). The variant, NC_000010.10: g.79769439G>C; NM_007055.3: c.1771-6C>G, was located in intron 13 of the POLR3A gene, in the largest region of homozygosity at 10q22.3 (hg19 chr10:73772999, 92.27cM to hg19 chr10:80704258, 98.32cM). Sanger sequencing (Table S1) confirmed the presence and expected segregation of the variant (Figure S2).

Frequency in the general Roma population was analysed using a Taqman assay (Table S1) in a panel of 703 healthy population-matched controls, 202 representing Vlax Roma
groups and 501 Balkan and Central/Western European Roma. We identified 10 heterozygotes, all confined to the Vlax Roma population, in agreement with the origin of the affected families. The overall carrier rate was estimated at 1.4%, and that in Vlax groups was close to 5%, compatible with previous observations on autosomal recessive disorders in the Roma (14).

The POLR3A mutation causes exon skipping
The c.1771-6C>G change is located six nucleotides upstream of the intron 13/exon 14 junction of POLR3A (Figure 3A). This change is one of the three variant alleles of SNP, rs115020338 (T/A/G), all present at very low frequency in the general population (Exome Aggregation Consortium (ExAC), Cambridge, MA (URL: http://exac.broadinstitute.org); accessed 09/02/2016). Amongst more than 60,000 exomes in the ExAC browser, there were four heterozygous carriers of the C>G variant, 76 carriers of the C>A variant, and 144 carriers of the C>T variant, with no homozygotes reported for any of the three. Analysis with the Human Splicing Finder prediction tool (15) showed a minor effect on the splicing score of either c.1771-6C>G or C>A and no change in the presence of C>T (not shown). By contrast, ESEfinder (16) predicted abolition of an SRp40 exonic splicing enhancer element (ESE) by only c.1771-6C>G, in contrast to no effect on SRp40 exerted by the other alleles. Removal of the ESE element could result in exon 14 skipping and a premature stop codon after eight amino acid residues in the new open reading frame (i.e. NM_007055.3: p.(Pro591Metfs*9)). RT-PCR of POLR3A transcripts in cultured patient and control fibroblasts with primers in exons 13 and 15 (Table S1) revealed a product of the expected size in the control and the presence of two transcripts in mutant cells (Figure 3B left panel). Sequencing of the RT-PCR products (Table S1) revealed an aberrant transcript skipping exon 14 in the mutant cells, produced in parallel to the canonical full-length mRNA (Figure 3C). Quantitative RT-PCR
(qRT-PCR) (Table S1) of POLR3A transcripts showed that the full-length transcript was reduced by 37±3% in the patient compared to the control. The transcript lacking exon 14 was detected at low levels, suggesting nonsense-mediated decay (NMD). POLR3A transcript analysis in the presence of the NMD inhibitor cycloheximide showed a significant increase in the aberrant transcript, confirming that it was targeted by NMD (Figure 3B right panel). The observed reduction in the full-length transcript is similar to previous reports (1) and can be predicted to lead to partial Pol III deficit.

The POLR3A mutation causes imbalance in Pol III produced transcripts

The effects of the c.1771-6C>G mutation on gene expression were analysed by directional sequencing of RNA extracted from blood from the three patients and three healthy controls. We used small RNA sequencing (RNA-Seq) to capture Pol III transcripts such as tRNAs, and other short regulatory ncRNAs such as small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs), and TruSeq libraries to investigate the changes in longer ncRNAs and in mRNAs as a downstream consequence of the POLR3A mutation. In the small RNA-Seq datasets the library size ranged from 21 to 30 million reads and 67 - 77 % of total reads mapped to the human genome in control and patient datasets. The TruSeq library size ranged from ~32 to 40 million reads and ~76 % of total reads mapped to the human genome in control and patient datasets. Basic alignment and mapping statistics from patients and controls are shown in Table S2.

We investigated differences in transcripts produced by Pol III between controls and patients (Figure 4). We extracted and identified different classes of small, regulatory ncRNAs from the small RNAseq datasets such as tRNAs and snRNAs, as well as transcripts with repetitive sequences such as the 5S RNA, long interspersed elements (LINEs) and short interspersed elements (SINEs) (Figure 4A). We observed the most significant changes in the
overall levels of tRNA and snRNA families, which were decreased in the patients compared to controls (Figure 4A). Since changes in the availability of tRNAs have been proposed as the major pathogenetic mechanism in POLR3A mutants (1), we investigated the steady state tRNA levels transcriptome-wide in the small RNA datasets. Nuclear tRNA isoacceptors are encoded by many loci on different chromosomes, therefore we analyzed the average expression of each family of tRNA isoacceptor (Figure 4B). We observed an overall decrease in tRNA abundance in the patients dataset, most significant for tRNA^{59}, but also other tRNAs including the initiator tRNA^{Met}, tRNA^{Leu} and tRNA^{His} (Figure 4B). Although the differences for specific tRNA families were not always statistically significant there was an overall trend towards decreased levels of most tRNA families in the patients compared to controls (Figure 4B), indicating that the POLR3A mutation lowers the levels of tRNAs.

In addition to tRNAs, Pol III transcribes 5S rRNA, and various non-coding RNAs including 7SL and 7SK RNAs, RNA P RNA, vault RNAs and Y RNAs (2). Therefore, we considered the effect of the mutation on these transcripts. In contrast to the general decrease in tRNAs, we observed an imbalance in levels of different Pol III-transcribed mRNA in the patient dataset (Figure 4C). For example, we found decreased levels of the snRNA U14 involved in splicing and in the RNA component of the signal recognition particle (SRP), RN7SL1, and increased levels of the transcriptional regulator 7SK RNA and the processing regulatory RNAs RMPI and RPPH1 in the patients (Figure 4C). We used qRT-PCR to confirm these changes in the blood from patients and controls (Figure 4D). The decrease in 7SL RNA, a component of the signal recognition particle that mediates co-translational translocation across membranes, and increase in 7SK RNA, a component of the 7SK RNP that couples transcriptional elongation with alternate splicing, supports the presence of both translational and transcriptional consequences of the POLR3A mutation. Changes in the levels of Pol III transcripts may exert downstream effects on mRNAs, for example an increase in
7SK RNA may have a negative effect on Pol II-mediated transcription, by repressing the positive Pol II transcription elongation factor P-TEFb via its sequestration within the 7SK RNP (16). In addition, the increased levels in 7SK RNA may affect the expression of different genes that undergo alternative splicing in the patients. RPPH1 encodes the catalytic RNA component of the ribonucleoprotein complex, RNase P, responsible for the cleavage and maturation of the 5’ termini of pre-tRNA molecules (17). RNase P has been shown to influence Pol I and III transcription (18), suggesting that in the patients the increased levels of RPPH1 may be in response to the decrease in tRNA levels.

We observed an increase in 5S rRNA expression (Figures 4C and 4D) that may be a compensatory response to decreased tRNA levels and preferential recruitment of Pol III to the promoters of the 5S rRNA, possibly complemented by perturbed translation regulation resulting from the reduced levels of 7SL RNA. Nucleosome positioning and chromatin context are known to affect the transcription of Pol III genes (5), in particular U6 sRNA and 7SK RNA (19, 20) and may explain the increased expression of specific Pol III transcripts by preferential recruitment of Pol III to their promoters.

Transcriptome-wide changes caused by the POLR3J mutation

Taking into account the increasing evidence of wide-ranging effects of Pol III target genes on overall gene expression (3-6), we searched for transcriptome-wide changes in c.1771-6C>G homozygotes. We compared gene expression in patients and in controls using DESeq2; identified differences at p < 0.05 are listed in Table S3. We observed dysregulated expression of genes involved in transcription regulation particularly in splicing and proteostasis that have been implicated in neurological disorders (Figure 5).

Next, we investigated changes in the expression of genes involved in ncRNA-mediated processes in the central nervous system such as snoRNAs and regulatory RNAs that
are components of spliceosomes and RNP complexes involved in transcription and translation regulation (Figure 5A) (21). Two groups of dysregulated genes in c.1771-6C>G homozygotes included small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) that have been implicated in a range of human diseases, including neurodegenerative disorders (22, 23). SnRNAs that are components of the spliceosomal complexes required for RNA splicing and their levels are imbalanced in the patients (Figure 5A and B) as observed for the Pol III-generated transcripts that are part of spliceosomal complexes (Figure 4C and D). Although Pol III has been observed to transcribe snoRNAs in C. elegans, D. melanogaster and S. cerevisiae, no such Pol III-dependent snoRNAs have been described in humans to date (2).

We observed that the POLR3A mutation significantly affected the levels of several different snoRNAs involved in pre-rRNA splicing and maturation (Figure 5A and B), suggesting that some of them may be transcribed by Pol III or their levels are affected indirectly by the changes in the ribonucleoprotein complexes that snoRNAs form with other recently identified Pol III-transcribed RNAs including vault RNA1-2, 7SK and 7SL RNAs (24). The altered levels of snoRNAs may contribute to the neurological phenotype observed in the patients, similarly to other snoRNAs-mediated neurological disease (25). Furthermore the decrease of the splicing factor HNRNPH2 (Figure 5A and 5B) which binds G-rich sequence motifs and regulates pre-mRNA splicing (26) points to potential effects on splicing in the patients.

The POLR3A mutation also leads to decreased expression of mRNAs coding for proteins involved in protein folding (Figure 5A and 5C). Hsp90 involved in protein folding and stress, encoded by HSP90AA1 is decreased in the patients and reduced phosphorylation of this protein in the brain has been correlated with frontotemporal lobe degeneration (27). The heat shock proteins required for protein folding and import and clearance of aggregated tau (28), such as DNAJA1 and the heat shock protein DNAJB2, which is expressed primarily in the neuronal layers of the brain (29) and are both decreased in the patients. These proteins
have protective neuronal role by suppressing the aggregation of proteins that cause
neurological diseases such as ALS (30) and Huntington’s disease (31) or regulating their
degradation such as that of ataxin-3 (32).

Decreased expression of other genes in the patients may contribute to the
neurodegenerative phenotype, such as decrease in ubiquitin (Figure 5), which is required for
protein degradation as its loss has been implicated in hypothalamic neurodegeneration (33).
Decreased ubiquitin that regulates the transcriptional activity of FOXO4 (34), a forkhead box
transcription factor required for cell cycle regulation and neural differentiation (35) may also
contribute to the observed FOXO4 decrease in the patients. We observe a significant decrease
in the iron-binding glycoprotein lactoferrin that protects from prion protein-induced cell death
(36) (Figure 5), and its decrease may contribute to the neurodegeneration observed in the
patients. These data indicate that perturbed Pol III RNA metabolism can have downstream
effects on proteostasis and protein quality control, processes that are known to be commonly
altered in neurodegenerative diseases.

DISCUSSION

Our findings characterize c.1771-6C>G in POLR3A as a pathogenic mutation leading to
aberrant splicing, reduced levels of the full-length transcript and expected partial protein
deficiency. Mutations in POLR3A and POLR3B, the genes encoding the largest subunits of
DNA-directed RNA polymerase III, cause several allelic autosomal recessive disorders (1),
now grouped under the heading Hypomyelinating Leukodystrophy 7 (MIM607694). Phenotypic heterogeneity in this group of disorders is manifested not only by the presence or
absence of defects in dentition and hormonal regulation, but also by the extent of white matter
versus cerebellar involvement (1, 7). Our patients showed no MRI evidence of white matter
involvement, and did not replicate previous findings of thinned corpus callosum, vermian
cerebellar atrophy, prominent cerebellar folia, and brainstem atrophy. Instead, they presented
with striatal and red nucleus involvement and clinical manifestations corresponding to the
unusual localization of the pathological changes (Table 1). While some clinical features seen
in our patients are also present in Hypomyelinating Leukodystrophy 7, their origins and
pathogenesis are different. Dysphagia and dysarthria are thought to result from bulbar
weakness in HLD7 (37), whereas the preserved pharyngeal reflexes in our patients point to
dystonia of the pharyngeal and laryngeal muscles as a likely cause. Similarly, the ataxia in
HDL7 is explained by the cerebellar atrophy and white matter changes (1, 37-39) whereas in
our patients it appears to be due to dorsal midbrain involvement, as suggested by the imaging
data. The differences between the phenotype described here and those reported previously
suggest that the clinical spectrum of POLR3 deficiency is broader than currently appreciated.
Its better understanding will depend on expanding the clinical indications for genetic analysis
beyond the leukodystrophies.

Our study of gene expression aimed to examine the effects of the mutation on both Pol
III target genes and Pol II-mediated expression. Wide-ranging between-tissue differences in
overall gene expression, as well as in Pol III transcript levels are well known (40). In our
experiments, we used whole blood RNA as previous studies have seen correlation between
brain and blood in the expression of specific gene clusters involved in transcription
regulation, protein synthesis, synapse formation and energy metabolism (41). These same
clusters showed significant changes in our patients. Understanding the brain-specific effects
of POLR3 mutations will await analyses of gene expression in affected brains, with studies in
patients with divergent pathologies being of particular interest. While individual gene changes
in the brain may not exactly mirror those described here, our findings already point to
significant derangement in Pol III transcription leading to an imbalance, not just a decrease, in
tRNAs and ncRNAs. Such changes are likely to be present in the brain, supporting the
previously hypothesized altered availability of tRNAs and ncRNAs in patients with POLR3
mutations (1, 7). Moreover, our data suggest that imbalance in Pol III-transcribed RNAs as a
result of the Pol III mutation has downstream effects on the promoter recognition of other Pol
III transcripts, as it has been shown recently (42) or on Pol II target genes involved in the
regulation of RNA metabolism (transcription, splicing and translation) and proteostasis,
processes that are required for brain development and have been implicated in
neurodegeneration. The dysregulation in these pathways likely contributes to the pathology of
the disorder and provides a resource for elucidating the regulatory mechanisms underlying Pol
III transcription.

MATERIALS AND METHODS

Methods and Subjects
Two families with three affected individuals and four unaffected relatives participated in the
study (Figure 1). The two families resided in different geographical areas of Bulgaria and
were unaware of each other’s existence. Both belonged to a young population sub-isolate, a
Roma group known as the Bowlmakers, characterized by small founding size, limited genetic
diversity and a high prevalence of autosomal recessive disorders (43, 44). Carrier rates were
investigated in a panel of 703 healthy population controls from diverse Roma/Gypsy sub-
isolates. Written informed consent was obtained from all participants. The study complied
with the guidelines of the institutions involved.

Clinical investigations
The affected subjects underwent neurological, ophthalmological, endocrinological, dental and
radiological examinations at the University Hospitals in Sofia and Varna. Clinical data were
also collected from. Information on the evolution of the disease was obtained from previous hospital records and interviews with care-providers.

Ataxia severity was evaluated with the Scale for the Assessment and Rating of Ataxia (SARA) (45). Cognitive performance was evaluated using the Mini-Mental State Examination (MMSE), and formal neuropsychological assessment of general intelligence, memory, word fluency and executive function. The endocrinological investigations included basal hormonal levels and stimulation tests of the gonadal axis with gonadotropin releasing hormone and clomiphene citrate (Supplementary Methods On-line). The dental status was assessed through detailed examinations and orthopantomograms (OPG).

Nerve conduction studies (NCS) and electromyography (EMG) using a Dantec-Keypoint portable electromyograph (Natus, Copenhagen, Denmark) were performed in all patients.

Magnetic resonance imaging (MRI) of the brain was performed on a 1.5T MR imager (MR Signa HDst, GE Healthcare Milwaukee USA). Patient IzII-1 also had positron emission tomography - computed tomography with fluorodeoxyglucose (FDG PET-CT) of the brain, using Gemmini TF 16, Philips equipment.

**Exome sequencing and data analysis**

Whole exome sequencing (WES) was performed on DNA samples from the three affected subjects and the parents in the Iz family (Figure 1). The analysis was done by Axeq Technologies (Seoul, Korea) using the TruSeq capture system and the HiSeq2000 platform (Illumina, San Diego, CA, USA). The initial data processing included alignment to the hg19 reference genome (46), variant calling in SAMtools (47) using default parameters, and identification of variants in dbSNP135 (http://www.ncbi.nlm.nih.gov/projects/SNP/). Variants were annotated to the UCSC Known Genes by ANNOVAR (48) version 23 Oct 2012, and ANNOVAR-formatted databases based on the UCSC Known Gene (“hg19_knownGene”),
the 1000 Genomes project (“hg19_ALL.sites.2012_02”), and the NHLBI Exome Sequencing Project (http://evs.gs.washington.edu/EVS/) (“hg19-esp6500_all”).

The WES data were used to extract 64,973 polymorphisms at HapMap Phase II SNP positions (8, 10). These were used for homozygosity mapping in AutoSNPa (13) with a cutoff ≥10 consecutive SNPs. A selected subset of 5,534 markers (in approximate linkage equilibrium, 0.15 cM, average heterozygosity 0.42) was used to calculate inbreeding coefficients in FEstim (11), with allele frequencies from HapMap-CEU data. Relatedness between the two families was estimated in PLINK (12) with allele frequencies obtained from an in-house dataset of 28 Roma exomes.

The search for the disease-causing mutation focused on variants with a quality score ≥20 and coverage ≥4X, which were located outside of segmental duplications and simple repeats. Based on the unusual phenotype, the endogamous nature of the population and pedigree structures, the search was based on the assumption of a rare deleterious variant for which all three patients are homozygous, the parents are heterozygous and no homozygotes are present among population controls. We used a step-wise filtering strategy (Supplementary Figure S1) including: a) removal of variants with allele frequency ≥1% in the 1000Genomes (http://www.1000genomes.org/) or NHLBI Exome Sequencing (http://evs.gs.washington.edu/EVS/) projects; b) “deleteriousness” predictions, retaining non-synonymous variants with Polyphen2 (49) scores >0.8 and SIFT (50) scores ≤0.05, splice-site, nonsense, non-stop, and small in-frame or frame-shift in/dels (including exonic and splicing variants in mRNAs); c) homozygosity in all affected individuals; d) heterozygosity in the parents; e) no homozygous subjects in a larger in-house database of 45 Roma exomes.

**Mutation verification**

The presence of the POLR3A mutation was confirmed with bidirectional Sanger sequencing (Australian Genome Research Facility, AGRF, Perth Node) of amplified fragments (Primers
listed in Supplementary Table S1). Evolutionary constraint was assessed using Genomic Evolutionary Rate Profiling (GERP) scores (51, 52). The functional impact on splicing was predicted using the Human Splicing Finder (HSF) (15), the Expasy translate tool (http://web.expasy.org/translate/), and the ESEfinder 3.0 (53). Mutation screening in the panel of healthy, ethnically-matched controls was performed with a custom designed TaqMan assay (Applied Biosystems), following the manufacturer’s protocol (Supplementary Table S1). Carrier status was confirmed by Sanger sequencing.

Cell culture
Fibroblasts were established from Sh II-1 and a healthy control from skin biopsies as previously described (54) and cultured at 37°C under humidified 95% air/5% CO2 in Dulbecco’s modified Eagle’s medium (DMEM, Gibco, Life Technologies) containing glucose (4.5 g/L), 2 mM glutamine, penicillin (100 U/ml), streptomycin sulfate (100 µg/ml) and 10 % fetal bovine serum (FBS). Fibroblasts were grown in the presence and absence of 100 ng/µl cycloheximide (CHX) for 24 h to inhibit nonsense-mediated decay before RNA was isolated and analyzed by qRT-PCR.

RNA Isolation and quantitative reverse transcription PCR
Total RNA was extracted from primary fibroblast cultures (patient Sh II-1 and a healthy control), and from whole blood (from Iz II-1 and Iz II-2 and two healthy controls), using the miRNeasy Mini kit (Qiagen) and the PAXgene Blood RNA system that included a DNase treatment step, respectively, following the manufacturer’s protocols. cDNA was prepared with the QuantiTect Reverse Transcription kit (Qiagen), used as a template in the subsequent PCR for mRNAs and ncRNAs. Reactions were performed using a Corbett Rotorgene 6000 using SensiMix SYBR mix (Bioline) and normalised to 18S rRNA.

The effect of the mutation on POLR3A mRNA processing was analysed by PCR amplification and product separation by 2% agarose/TAE gel electrophoresis. Visible bands
were excised, extracted from gel pieces using the GeneJET PCR Purification Kit (Fermentas) and Sanger sequenced. The abundance of normal and aberrant transcripts was assessed by qRT-PCR as described above.

**RNA sequencing**

RNA sequencing libraries were prepared using 4 µg of total RNA from blood of the three patients and three, aged matched healthy controls, using the TruSeq Small RNA Prep kit, and the small RNA-seq kit (Illumina) for which we size selected fragments ranging in size from 50-200 bp. RNA library preparation and strand-specific RNA sequencing was carried out by the Cologne Genome Sequencing Centre, Germany on a Hi-Seq Illumina sequencer.

**Alignment and differential expression analysis**

*Small RNA library*

Adapter sequences were trimmed from the 3’ end with cutadapt 1.7.1 (55) (–e 0.1 -n 2 -O 6 -m 16 –match-read-wildcards). Successfully trimmed reads were aligned to the human genome (hg19) with bowtie2 v2.2.5 (56) (-N 1) and repeat expression analysed with HOMER v4.7 (57) (makeTagDirectory -keepOne -sspe; analyzeRepeats.pl repeats hg19 -rplm).

*Total RNA library*

Quality and adapter trimming of raw reads was performed with Trimomatic 0.32 (ILLUMINACLIP:TruSeq3-PE:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36). Paired and unpaired trimmed reads were aligned with TopHat 2.0.12 (–b2-very-sensitive –library-type fr-firststrand –read-realign-edit-dist 0 –mate-inner-dist 37 –mate-std-dev 56 -G genes.gtf) against the human genome (hg19). Gene-specific fragment counts were generated with HTSeq 0.6.1p1 (–s yes) and the Ensembl 75 gene annotation, and differential expression was tested with DESeq2 1.6.2 using standard parameters (parametric dispersion estimates, Wald statistic and Benjamini-Hochberg multiple testing correction).
Acknowledgements:

We thank the affected families for participating in this study and Alistair Forrest for advice on bioinformatics analyses. This project was supported by fellowships and project grants (to A.F and O.R.) from the National Health and Medical Research Council (APP1058442, APP1045677, APP1041582, APP1023460, APP1005030), the Australian Research Council (FT0991008, FT0991113) and the University of Western Australia. S.S. is supported by an APA scholarship from the Australian Government.
References:


12. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., et al. (2007) PLINK: A Tool Set for Whole-
Genome Association and Population-Based Linkage Analyses. The American Journal of Human Genetics, 81, 17–17.


Figure Legends

Figure 1. Pedigree structures of families Sh and Iz with Sanger sequencing genotypes for the POLR3A c.1771-6C>G splicing mutation. Individuals with available whole exome sequencing (WES) data are indicated.

Figure 2. T2 FLAIR MRI axial (A) and coronal (B) demonstrating marked bilateral atrophy of the corpus striatum. (C) TSE T2 MRI axial illustrating the signal change (white arrow) in the region of the red nuclei and third cranial nerve tracts. (D) FDG PET-CT of subject Iz II-1 demonstrating dramatic bilateral hypometabolism in the striatum. (E) FDG PET-CT of a healthy age and sex matched control. The axial PET-CT images are acquired of the basal ganglia.

Figure 3. The POLR3A mutation causes a splicing defect. (A) Schematic representation of the predicted functional impact of the POLR3A c.1771-6C>G mutation. Vertical bars designate exons, exonic sequences are shown with capital letters, intronic with small letters (5' to 3' gene orientation, gene encoded on the minus strand). The SRp40 exonic splicing enhancer element (ESE) is highlighted. The codons of the predicted open reading frame of the transcript skipping exon 14 are indicated with brackets; the premature termination codon is boxed. (B) RT-PCR with primers in exons 13 and 15 reveals exon skipping in the POLR3A mRNA in patient fibroblasts carrying the POLR3A c.1771-6C>G mutation. Addition of cyclohexamide (CHX) stabilises the exon skipped mRNA, indicating that it is normally subject to nonsense-mediated decay (NMD). (C) Sanger sequencing of RT-PCR products confirms the identity of the exon skipped mRNA and the predicted translation reveals premature termination codons within the new reading frame.

Figure 4. Mutation in the POLR3A gene causes an imbalance in Pol III transcripts.
(A) RNAseq was carried out on blood RNA from the three patients and three age matched healthy control and the datasets were analyzed for expression and represented as rmkp of the patients compared to the controls for tRNAs and other regulatory and ncRNA transcribed by Pol III. (B) The expression values for tRNAs was summarized by amino acid and anticodon. Normalized counts were summed for all expressed loci from each sample that encoded tRNAs of the same amino acid type or all copies of the same tRNA isotype. (C) Differential expression of non-tRNA Pol III transcripts is shown; for multi-copy genes, normalized counts for all expressed loci were summed prior to calculating log2 fold change. (D) RNA isolated from blood was analyzed by qRT-PCR for the Pol III produced transcripts, normalized to 18S rRNA expression and the statistical significance was calculated using the student t-test (*p<0.05 and **p<0.01).

Figure 5. The POLR3A mutation leads to specific changes in genes involved in transcription regulation, splicing, translation and proteostasis. (A) Summary of the differentially expressed genes in the TruSeq datasets that are affected by the Pol III mutation; the affected cellular processes and their relevance to neurological diseases are also shown. (B) The levels of regulatory ncRNAs involved in splicing were measured by qRT-PCR in RNA from the patients’ and controls’ blood to corroborate the changes identified by RNAseq. Amplification of the ncRNA transcripts is shown, normalized to 18S rRNA as means ± SD of three independent biological experiments. (C) RNA isolated from blood was analyzed by qRT-PCR for differentially expressed mRNAs involved in protein folding, degradation and cell death. The data are normalized to 18S rRNA as means ± SD of three independent biological experiments; the statistical significance was calculated using the student t-test (*p<0.05 and **p<0.01).
Table 1. Clinical features of the three patients.

<table>
<thead>
<tr>
<th>Patients</th>
<th>Shh-1</th>
<th>Izll-1</th>
<th>Izll-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year of birth</td>
<td>1959</td>
<td>1985</td>
<td>1989</td>
</tr>
<tr>
<td>Sex</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Age at onset</td>
<td>8</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Age at assessment</td>
<td>52</td>
<td>27</td>
<td>23</td>
</tr>
<tr>
<td><strong>Upper Motor Neuron Signs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muscle weakness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spasticity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brisk tendon reflexes</td>
<td>Upper limbs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Babinski sign</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>Extrapyramidal signs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rigidity</td>
<td>LL, R wrist</td>
<td>LL</td>
<td></td>
</tr>
<tr>
<td>Dystonia</td>
<td>R hand, LL</td>
<td>LL predominantly R</td>
<td>Distal LL</td>
</tr>
<tr>
<td>Tremor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ataxia (SARA scores)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gait (0-8)</td>
<td>6</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Stance (0-6)</td>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Sitting (0-4)</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Speech (0-6)**</td>
<td>5</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Total score (0-40)</td>
<td>24</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>Dysphagia</td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Gaze-evoked nystagmus</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abnormal Smooth Pursuits</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td>95</td>
<td>85</td>
<td>80</td>
</tr>
<tr>
<td><strong>Neuroimaging</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atrophy of striatal structures</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Signal changes in red nucleus</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Limb deformities</td>
<td>Bilateral pes cavus</td>
<td>Bilateral pes cavus</td>
<td></td>
</tr>
<tr>
<td>Hypodontia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypogonadotropic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypogonadism</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* SARA - Scale for the Assessment and Rating of Ataxia (29): 0 indicates lack of impairment, and higher scores indicate increasing severity. The scores were low or normal for the remaining SARA items (not shown).

** The speech disturbance is due to a combination of dystonia and cerebellar dysarthria.

The following abbreviations are used: LL, lower limbs; R, right; L, left.
Figure 1
208x164mm (96 x 96 DPI)
Figure 3

197x289mm (300 x 300 DPI)
Figure 5

A

<table>
<thead>
<tr>
<th>Small nuclear RNA</th>
<th>Case</th>
<th>Control</th>
<th>log2 Fold</th>
<th>P-value</th>
<th>Associated Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>U7TERT/RHODES</td>
<td>Splicing</td>
<td>+1.01</td>
<td>0.025483</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U6SNUR/7</td>
<td>Splicing</td>
<td>+0.58</td>
<td>0.012964</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNORD7</td>
<td>Translation</td>
<td>+1.29</td>
<td>0.029277</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNORD10</td>
<td>Translation</td>
<td>+1.14</td>
<td>0.017864</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B

C

Figure 5

207x213mm (300 x 300 DPI)
Supplementary Figure S1. Step-wise filtering of the variants identified by exome sequencing in the affected individuals. The search for the disease mutation was based on the assumption of a rare/unique variant homozygous in all three patients, heterozygous in the available parents from family Iz and not homozygous in any of the 45 Roma exomes available in-house. *Primary quality control keeping for further analysis variants with quality scores ≥20 and coverage≥4X, located outside of segmental duplications or simple repeats; ^Variants defined as deleterious included non-synonymous amino acid substitutions with a Polyphen2 score >0.8 and SIFT score ≤ 0.05, splice-site, nonsense and non-stop changes, as well as small in-frame or frameshift insertion/deletions.

<table>
<thead>
<tr>
<th></th>
<th>Iz II-1</th>
<th>Iz II-2</th>
<th>Sh II-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial QC*</td>
<td>69,329</td>
<td>67,685</td>
<td>67,586</td>
</tr>
<tr>
<td>Frequency in 1000Genomes/NHLBI exomes &lt;0.01</td>
<td>7,599</td>
<td>7,271</td>
<td>7,303</td>
</tr>
<tr>
<td>Deleterious^</td>
<td>1,563</td>
<td>1,491</td>
<td>1,460</td>
</tr>
<tr>
<td>Homozygous</td>
<td>499</td>
<td>472</td>
<td>456</td>
</tr>
<tr>
<td>Homozygous in the three affected individuals</td>
<td>230</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterozygous in the available parents</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Not homozygous in 45 control Roma exomes</td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Figure S2. POLR3A c.1771-6C>G (NM_007055.3) mutation identification and verification. Top panel: Integrative Genomics Viewer snapshot of the short reads alignment from the exome sequencing aligned to the human genome reference hg19 and relative to the (+) strand of chromosome 10. Sequences matching the reference are given in grey and the variant is coloured and marked by the base letter. The genomic position of the variant is indicated by an arrow. Bottom panel: confirmatory Sanger sequencing in the transcript direction (- strand).
Defects in mitochondrial RNA metabolism in disease

Stefan J. Siira¹, Anne-Marie J. Shearwood¹, Cameron P. Bracken², Oliver Rackham¹,³ and Aleksandra Filipovska¹,³*

¹ Harry Perkins Institute of Medical Research and Centre for Medical Research, Level 7 QQ Block, QEII Medical Centre, 6 Verdun Street, Nedlands WA 6009 Australia;
² Division of Human Immunology, Centre for Cancer Biology, SA Pathology, Adelaide SA 5000, Australia
³ School of Chemistry and Biochemistry, The University of Western Australia, Nedlands, Western Australia 6009, Australia.

*Correspondence: aleksandra.filipovska@uwa.edu.au; Tel.: +61-8-6151-0732
Abstract

The expression of mitochondrially-encoded genes requires the efficient processing of long precursor RNAs at the 5’ and 3’ ends of tRNAs, a process which, when disrupted, results in disease. Two such mutations reside within mt-tRNA\(^{Leu(UUR)}\); a m.3243A>G transition, which is the most common cause of MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes), and m.3302A>G which often causes mitochondrial myopathy (MM). We used parallel analysis of RNA ends (PARE) that captures the 5’ terminal end of 5’-monophosphorylated mitochondrial RNAs to compare the effects of the m.3243A>G and m.3302A>G mutations on mitochondrial tRNA processing and downstream RNA metabolism. We confirmed previously identified RNA processing defects, identified common internal cleavage sites and new sites unique to the m.3243A>G mutants that do not correspond to transcript ends. These sites occur in regions of predicted RNA secondary structure, or are in close proximity to such regions, and may identify regions of importance to the processing of mtRNAs.

Key words: RNA metabolism; mitochondrial disease; tRNA; PARE
1. Introduction

Mitochondria contain their own discrete genomes, a characteristic that is borne from their α-proteobacterial origin. The human mitochondrial genome is a circular molecule 16,569 bp in length that encodes 13 protein components of the electron transport chain, 2 mitochondrial rRNAs and 22 mitochondrial tRNAs. Mitochondrial RNAs are initially transcribed as long, polycistronic transcripts from both strands of the genome denoted the heavy and light strands due to their high and low purine content, respectively, which are processed to excise interspersed tRNAs thus releasing the other transcripts (reviewed in (Rackham et al., 2012)). Both strands are transcribed from promoters located within the non-coding region containing the D-loop, known as the control region, with light strand transcript termination occurring at the MTERF1 site within the MT-TL1 mtDNA region while the heavy-strand transcript is almost genome-length and terminates upstream of the D-loop (reviewed in (Gustafsson et al., 2016)). Following transcription of mitochondrial DNA (mtDNA), the 5’ and 3’ ends of tRNAs residing within the precursor transcripts are cleaved by the mitochondrial RNase P and RNase Z (ELAC2) (Brzezniak et al., 2011; Holzmann et al., 2008; Rackham et al., 2016; Sanchez et al., 2011), respectively to release individual rRNAs, mRNAs and tRNAs. These mitochondrial RNAs (mtRNAs) undergo maturation processes involving various post-transcriptional modifications such as nucleotide modifications of tRNAs and rRNAs, 3’ CCA trinucleotide additions to tRNAs and polyadenylation of mRNAs (reviewed in (Rackham et al., 2012)). Degradation of mitochondrial RNAs is mediated by the hSUV3-hPNPase complex in a process that appears to be functionally coupled with polyadenylation, although many aspects of mitochondrial degradation remain unknown (reviewed in (Levy and Schuster, 2016)).

In the mitochondrial genome, tRNA genes have been noted as hotspots for pathological mutations and MT-TL1 in particular (Moraes et al., 1993), perhaps due to its dual importance in the processing of rRNA-containing precursor RNA and as the location of an MTERF1-binding site that mediates transcriptional termination. An A to G transition at nucleotide position 3243, which resides within MT-TL1, is one of the most studied mitochondrial mutations due to its frequency and association with a number of mitochondrial diseases, most notably mitochondrial myopathy, encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS) where it is responsible for ~80% of cases, but also myoclonus epilepsy with ragged-red fibres (MERRF), mitochondrial myopathy (MM), Leigh syndrome (LS), chronic progressive external ophthalmoplegia (CPEO) and maternally inherited diabetes and deafness (MIDD) (reviewed in (Kirino and Suzuki, 2005)). High load of this mutation results in a decreased rate of protein synthesis, compromised cellular respiration and reduced steady-state levels of mitochondrially-encoded respiratory chain components (King et al., 1992).
At a molecular level, the mutation has been shown to have multiple effects on the tRNA itself including impaired post-transcriptional modification such as aminoacylation (Börner et al., 2000), methylation (Helm et al., 1999) and taurine modification of the anticodon wobble base (Kirino et al., 2004), alterations to its tertiary structure leading to dimerization (Wittenhagen and Kelley, 2002), and misincorporation of amino acids in elongating peptides leading to elevated degradation (Sasarman et al., 2008). Furthermore, reduced efficiency of 5’ and 3’ MT-TL1 cleavage disrupts the processing of mitochondrial precursor RNA, as is evident by the accumulation of an intermediate RNA product known as RNA19 (Hess et al., 1991), that contains the MT-RNR2, MT-TL1 and MT-ND1 transcripts, observed both in patient muscle and transmitochondrial cybrids close to homoplasmy (King et al., 1992; Maniura-Weber et al., 2006).

The accumulation of RNA19 as a consequence of another mutation at position m.3302A>G in MT-TL1 has been associated most often with mitochondrial myopathy (Bindoff et al., 1993), but has also been linked to MELAS (Goto et al., 2014) as well as other disorders (van den Bosch et al., 2004). The mutation of this highly conserved nucleotide impairs 3’ processing of MT-TL1 (Levinger et al., 2004) leading to a reduction in steady-state levels and partially impaired aminoacylation, the accumulation of stable RNA19 (Maniura-Weber et al., 2006) and a respiratory defect that predominantly affects complex I (Bindoff et al., 1993).

Parallel analysis of RNA ends (PARE) is an approach that was developed to investigate RNAs with a monophosphate at the 5’ end such as those found in miRNAs, RNA degradation products (German et al., 2009) and mitochondrial RNAs (Mercer et al., 2011; Rackham et al., 2016). The 5’ monophosphate of cleaved RNAs allows the ligation of an adapter containing a MmeI recognition site, which after reverse transcription and cDNA synthesis the restriction enzyme MmeI is used to cleave the cDNA ~20 nt downstream of its recognition site to fragments that are amplified and sequenced. Mitochondrial transcripts possess 5’ terminal monophosphates (Rackham et al., 2016), in contrast to the 5’ m3G protective caps present on nuclear mRNAs, making PARE amenable for investigation of RNA processing in mitochondria. Analysis of the abundance and distribution of the terminal 5’ nucleotide of these tags produces a profile of both specific and non-specific cleavage events, such as degradation, which appear as large single nucleotide peaks or as series of stepped peaks in close proximity, respectively (Rackham et al., 2016).

In this study, we used PARE to investigate the effects of the m.3243A>G and m.3302A>G mutations on mitochondrial RNAs and identify any transcriptome-wide differences in their processing or degradation as a consequence.

2. Materials and Methods
2.1 Cell culture
Cybrid cell lines were cultured at 37 °C under humidified 95% air/5% CO\textsubscript{2} in Dulbecco's modified Eagle's medium (DMEM, Gibco, Life Technologies) containing glucose (4.5 g/l\textsuperscript{−1}), 1 mM pyruvate, 50 μg/ml uridine, 2 mM glutamine, penicillin (100 U/ml\textsuperscript{−1}), streptomycin sulfate (100 μg/ml\textsuperscript{−1}) and 10% fetal bovine serum (FBS).

2.2 Mitochondrial isolation
Mitochondria were prepared from 10\textsuperscript{7} cells grown overnight in 15 cm\textsuperscript{2} dishes and isolated as described previously [56], with some modifications. Mitochondria were lysed for 30 min in buffer containing 250 mM sucrose, 100 mM KCl, 20 mM magnesium acetate, 10 mM Tris–HCl pH 7.5, 0.5% Triton X-100 and EDTA-free Complete protease inhibitor cocktail (Roche).

2.3 RNA isolation, digestion, northern blotting and qRT-PCR
RNA was isolated from mitochondria using the miRNeasy Mini kit (Qiagen) incorporating an on-column RNase-free DNase digestion to remove all DNA. RNA (5 μg) was resolved on 1.2% agarose formaldehyde gels, then transferred to 0.45 μm Hybond-N\textsuperscript{+} nitrocellulose membrane (GE Lifesciences) and hybridized with biotinylated oligonucleotide probes specific to human mitochondrial mRNAs, rRNAs and tRNAs. Hybridizations were carried out overnight at 50°C in 5x SSC, 20 mM Na\textsubscript{2}HPO\textsubscript{4}, 7% SDS and 100 μg.ml\textsuperscript{−1} heparin, followed by washing. The signal was detected using either streptavidin-linked horseradish peroxidase or streptavidin-linked infrared-labelled antibody (diluted 1: 2000 in 3x SSC, 5% SDS, 25 mM Na\textsubscript{2}HPO\textsubscript{4} pH 7.5) by enhanced chemiluminescence (GE Lifesciences) or using an Odyssey Infrared Imaging System (Li-Cor), respectively.

Complementary DNA (cDNA) was prepared using the QuantiTect Reverse Transcription Kit (Qiagen) and used as a template in the subsequent PCR that was performed using a Corbett Rotorgene 6000 using SensiMix SYBR mix (Bioline) and normalized to 18S rRNA. Isolated mitochondrial RNA (2.5 μg) was incubated in the presence or absence of Terminator exonuclease (Gene Target Solutions) in Buffer B for 30 min at 42°C in the presence of RNaseOUT inhibitor (ThermoScientific) according to the manufacturer’s instructions followed by northern blotting as described above.

2.4 PARE and alignments
RNA sequencing was performed on mitochondrial RNA isolated from cybrid cells using the Illumina MiSeq platform according to the PARE protocol [6,45,46].
Raw reads were processed as follows: 3’ adaptor sequences were trimmed with cutadapt 1.9.1 (Martin, 2011) (-e 0.1 -n 2 -O 6 --quality-base=64 --match-read-wildcards -a TCGTATGCGCCTCTCTGCTTG), followed by library splitting by index sequence (5’-ATACG-3’ and 5’-CGATGT-3’) with fastx_barcode_splitter.pl 0.0.13 from the FASTX-Toolkit (-eol -exact), then index sequence trimming with cutadapt (-e 0.0 -n 1 -O 6 --quality-base=64). Reads that were 20 or 21 nt in length after trimming were extracted with awk and aligned uniquely to the human mitochondrial genome (hg38) with Bowtie 1.1.2 (Ben B Langmead et al., 2008) (-yS -m 1 –phred64-quals). Strand-specific 5’ end profiles normalised to library size (reads per million, RPM) were generated with BEDtools 2.25.0 (Quinlan and Hall, 2010) (-d -5 -strand [-/+] -scale [1,000,000/total mapped reads]) and converted to bedGraph format for visualisation. PARE data for 143B cybrids containing wild type mtDNA was retrieved from the Gene Expression Omnibus (GEO: GSM769736) and processed as above, with the exception of specifying phred-33 rather than phred-64 at trimming and alignment steps to account for sequence quality format differences. The data generated in this study have been submitted to GEO with accession GSE85549.

3. Results

3.1. The effects of m.3243A>G and m.3302A>G mutations on the processing of RNA19

We investigated the effects of the mutations in both m.3243A>G and m.3302A>G cell lines to identify that the 5’ ends of MT-TL1 and the MT-ND1 mRNA were both dramatically reduced relative to control cells (Figs 1 and 2; Supplementary Table 1). Additionally, the ratio of MT-TL1:MT-ND1 abundance in the mutants was opposite to the wild type, such that the abundance at the MT-ND1 5’ end was twice that at the 5’ end of MT-TL1 whereas in the mutant cells the 5’ end of MT-TL1 is twice as abundant as MT-ND1, indicating a greater 3’ end processing impairment relative to 5’ end processing that was consistent with previous reports (Levinger et al., 2004; Maniura-Weber et al., 2006; Rackham et al., 2016; Rossmanith and Karwan, 1998). However, the 5’ end abundance of MT-TL1, while reduced relative to the wild type cells, is not markedly different between the two mutant cell lines despite the large difference in reported 5’ processing efficiency for the m.3243A>G and m.3302A>G mutations indicating that both mutations lead to processing defects of MT-TL1 and consequent RNA19 accumulation (Rossmanith and Karwan, 1998).

Interestingly, we observed that the 5’ end of MT-RNR2, which corresponds to the 5’ end of RNA19 was approximately twice as abundant in the m.3243A>G cells compared to both wild type and m.3302A>G cells (Supplementary Table 1), although the significance of this is unclear. In addition to an overall reduction in abundance compared to the wild type cells, the reduction in MT-ND1 5’ end abundance relative to MT-TL1 5’ end abundance in the mutant cells is consistent with impaired
3’ tRNA processing previously reported for these mutations (Levinger et al., 2004). This indicates that impaired 5’ cleavage has downstream effects on 3’ processing and further establishes that 5’ tRNA processing precedes 3’ cleavage in mitochondria as it has been identified recently in vivo (Rackham et al., 2016). We found that the 5’ end of MT-TV were also considerably more abundant in the wild type cells (~320,000 RPM) compared to either of the mutant cells (~60,000 – 69,000 RPM) suggesting that somehow the 5’ end cleavage of RNA19 was affected by the mutations in the MT-TL1 (Figs 1 and 2; Supplementary Table 1). In contrast, the 5’ ends of MT-TF and MT-RNR1 upstream of RNA19 were considerably more abundant in the mutant cells than in the wild type cells, indicating variable processing efficiency when there is a mutation in MT-TL1 (Fig 1; Supplementary Table 1).

3.2. The cleavage of non-canonical sites is altered in MT-TL1 mutant cell lines

Mitochondrial rRNAs and mRNAs are typically released from their respective precursors RNAs via the excision of interspersed tRNAs, however several junctions are not flanked by tRNAs or are flanked by antisense tRNA sequences, namely the 3’ end of MT-ND6, the junction between MT-ATP8/6 and MT-CO3, the 5’ end of MT-COI and the junction between MT-ND5 and MT-CYB (Figs 3 and 4). In our data two of these four junctions lacked any distinctive 5’ peaks, MT-ND5/MT-CYB and the 5’ end of the COI mRNA (Fig 3A and 3B), located 3 nt upstream of the start codon. The MT-ATP8/6-MT-CO3 junction, which lacks an intervening tRNA, was not observed in the wild type or m.3302A>G cells, however the m.3243A>G cells contained a very large peak of ~32,000 RPM (Fig 4A). Since PARE requires a 5’ monophosphate to capture RNAs, we performed 5’-phosphate-dependent exonuclease digestion of isolated 143B mitochondrial RNA and revealed that MT-CO3 is unaffected (Fig 4B), suggesting the absence of this residue and, consequently, its presence in the m.3243A>G sample. The MT-CO3 mRNA is known to co-exist within a tricistronic transcript with MT-ATP8/6, known as RNA15, which is susceptible to the exonuclease digestion, much like the other mt-RNAs indicating that the 5’ of MT-CO3 is uniquely protected from digestion. The processing of the MT-CO3 mRNA from the bicistronic MT-ATP8/6 is currently poorly understood. It is possible that the FASTKD5 protein may play a role in this site since knockdown of this protein results in an increased accumulation of the RNA15 transcript along with other non-canonical junctions (Antonicka and Shoubridge, 2015). It is interesting to note that the m.3243A>G cells are the only ones identified to date to contain a distinctive 5’ peak at the MT-ATP8/6-MT-CO3 junction suggesting that this mutation may have downstream effects on the secondary structure of this transcript.
A peak 34 or 35 nt downstream of the 3’ end of the MT-ND6 open reading frame (ORF), as would be expected for the 3’ UTR was not observed. However, small but distinctive peaks were observed in all three samples at position 13,647, which, if taken as the ND6 UTR 3’ processing junction implies a 3’ terminal nucleotide at position 13,648, 501 nt downstream of the MT-ND6 ORF 3’ end consistent with the region described by (Slomovic et al., 2005) (Fig 4C). The 3’ ends of several of these PARE tags terminated with non-templated adenosine residues, however MT-ND6 is not known to be polyadenylated suggesting that these nucleotides may be degradation-related additions. The processing of the MT-ND6 ORF 3’ end appears decreased in both mutant cell lines indicating potential downstream effects of these mutations on the expression or processing of the ND6 transcript.

3.3. Mutations in MT-TL1 lead to altered mitochondrial RNA metabolism

Next we investigated the presence of 5’ peaks across the mitochondrial transcriptomes in wild-type and mutant cell lines to determine if there were any unique and non-canonical sites that are present or downstream consequences of the mutations in the MT-TL1 gene. We observed multiple large peaks located within the coding regions of several mRNAs, many of which were unique to the m.3243A>G cell line (Table 1). In this cell line MT-COI contains a peak of ~38,000 RPM at position 6,534, surrounded by four smaller peaks of ~1,400 – 2,700 RPM, in addition to a sixth peak ~200 nt downstream (Fig 3B). There was also a single peak of ~1,500 RPM within the MT-ND4 ORF at position 11,567 (Fig 5A) and another similarly sized peak at position 2,585 within MT-RNR1 (Fig 2) observed only in the m.3243A>G cells. These may be internal cleavage sites that are a consequence of unstable mRNAs that are downstream consequence of the m.3243A>G mutation. Mitochondrial protein synthesis in the m.3243A>G cell line is entirely lost as a result and the increased degradation of mitochondrial RNAs may be a consequence of impaired translation of all 13 mitochondria encoded polypeptides (Supplementary Fig 1). Although m.3302A>G causes impaired processing of RNA19 we do not observe an overall defect in mitochondrial translation (Supplementary Fig 1) and this may explain the lack of internal cleavage sites in mitochondrial RNAs from these cells.

We observed several instances of distinctive peaks that appear in all three samples, but which do not correspond to any known transcript ends. For example, a cluster of peaks was observed on the light strand antisense to MT-ND2 (Fig 5B), the largest and most upstream of which was located at position 5,292, and was much more abundant in the wild-type compared to mutant cell lines. Additionally, a second cluster of peaks with similar characteristics with a maximum abundance up to 18,000 RPM was observed on the light strand, in a region proximal to the MT-
ATP8/6-MT-CO3 junction region on the heavy strand (Fig 6A), which corresponds with a previously identified region of predicted RNA secondary structure (Mercer et al., 2011). A third peak of ~1,000 – 2,300 RPM was observed in all samples within the MT-CO2 coding region (Fig 6B) and although its abundance was only slightly increased in the m.3243A>G mutant, relative to the 5’ MT-CO2 peak the ratio of internal to 5’ end abundance was increased in both mutants by a factor of ~3 – 6 (Table 1, Supplementary Table 1). These cleavage sites may be a product of RNA processing enzymes some of which can be promiscuous and cleave secondary structures within transcripts that resemble tRNAs as observed before (Rackham et al., 2011). The mutations in MT-TL1 contribute to an overall defect in the hierarchy of RNA processing and may contribute to the changes in cleavage at these secondary sites.

4. Discussion

We investigated the effect of two mutations in MT-TL1 in the mitochondrial transcriptome by PARE. RNA19 is a mtRNA processing intermediate comprised of the complete MT-RNR2, MT-TL1 and MT-ND1 sequences as a single contiguous molecule. Both mutations investigated are known to impair the processing of MT-TL1 and lead to the accumulation of RNA19 (Bindoff et al., 1993; King et al., 1992; Maniura-Weber et al., 2006), although the 5’ processing defect has been reported to be more severe for the m.3243A>G mutation (Levinger et al., 2004). Our data directly sampled the 5’ ends of MT-TL1 and MT-ND1 and showed reduced abundance at both junctions relative to wild type, indicative of impaired RNA19 processing as the primary defect, however there was little difference in 5’ MT-TL1 abundance between the two mutants despite a report that the m.3243A>G mutation can also impair 5’ processing (Rossmann and Karwan, 1998). This suggests that, at least with this tRNA, the efficiency of processing of 5’ and 3’ ends is interdependent.

The 5’ and 3’ ends of mitochondrial RNAs are typically released from their polycistronic precursors via RNA processing enzymes acting on the tRNAs flanking the gene (reviewed in (Rackham et al., 2012)), however a number of exceptions exist: the 5’ ends of MT-CO1, MT-CO3 and MT-CYB, and the 3’ end of ND6. The 5’ end of MT-COI is flanked by the antisense of MT-TY; the 5’ of MT-CYB is likewise flanked by an antisense tRNA, MT-TE, but this tRNA forms a portion of the MT-ND5 3’ UTR and is not processed by the canonical machinery (Sanchez et al., 2011); and the remaining two non-canonical junctions are located between MT-ATP8/6 and MT-CO3 and the 3’ end of MT-ND6, lack tRNAs entirely and have been reported to be affected by the FASTK and FASTKD5 proteins (Antonicka and Shoubridge, 2015; Jourdain et al., 2015). In our datasets, the 5’ ends of MT-COI and MT-CYB were not detected or of very low abundance and were not appreciably different in appearance to the surrounding background. Furthermore, the MT-
ATP8/6-MT-CO3 junction was observed as a single large peak at the 5´ nucleotide of the MT-CO3 mRNA in the m.3243A>G sample only. Since PARE relies on the presence of a free 5´ monophosphate to capture RNAs, the alignment of tags at the 5´ end of MT-CO3 in the m.3243A>G cybrid only suggested that the 5´ end of this mRNA is lacking or obscured. Indeed, when isolated wild type 143B mitochondrial RNA is incubated with a 5´-phosphate-dependent exonuclease the MT-CO3 mRNA but not MT-ATP8/6 and MT-CO2, is resistant to digestion, suggesting it lacks the requisite monophosphate at its 5´ end unlike the other mRNAs. An alternate possibility is that the formation of an RNA secondary structure, as has been previously predicted (Mercer et al., 2011), may preclude the capture of the terminal region of this mRNA, although these possibilities are not mutually exclusive.

The location of the 3´ end of MT-ND6 has been reported at both ~500 nt (Slomovic et al., 2005) and ~33/34 nt 3´ UTRs (Mercer et al., 2011; Rackham et al., 2011). The identification of the 3´ junctions with this type of data is possible because of the compact, polycistronic nature of the mtRNA precursors where each 3´ cleavage event also creates a downstream RNA with a free 5´ monophosphate. In our data here, an accumulation of tags with non-template 3´ adenosines within the region described by (Slomovic et al., 2005) was observed. This might indicate that the predominant 3´ end of MT-ND6 occurs after ~500 nt 3´ UTR or, since PARE does not identify 3´ ends directly, that this peak result from a long-lived, downstream degradation intermediate in the discarded light strand non-coding RNA. Since MT-ND6 is the only mitochondrial mRNA that lacks a poly(A) tail the origin of these adenosine extensions may originate with the mtPAP and the degradation machinery. Indeed, such a potential mechanism for the processing of the 3´ end of MT-ND6 that involves the degradation machinery and the FASTK protein bound to the MT-ND6 ORF and 3´ UTR has been proposed (Jourdain et al., 2015).

The peaks observed within MT-ND4 and MT-RNR2 in the m.3243A>G sample do not correspond to any specific features, and their genesis is unknown. The cluster of peaks within MT-CO1, however, resides in a region predicted to harbour RNA secondary structure and within which a large central PARE peak in HEK cells is observed (Mercer et al., 2011). The significance of the RNA structure and cleavage within this region is unclear, but its specific appearance in the m.3243A>G sample suggests that changes in mitochondrial function resulting from the tRNA mutation may have changed its relative abundance and allowed its detection.

The other unidentified peaks are consistently observed in all three samples, although their origin and function are unclear. RNA structures have been predicted and cleavage detected previously (Mercer et al., 2011) within the genomic regions where peaks were observed in our data on the light strand antisense to MT-ND2 and MT-CO3. Given their staggered appearance and
location within non-coding genomic regions these peaks may indicate protein-binding sites, particularly those involved in the processing and degradation of the light strand transcript. The other distinctive feature observed in all samples was a central peak within the MT-CO2 coding region, which was relatively abundant compared to the MT-CO2 5’ end. Additionally, whilst the 5’ end abundance of MT-CO2 was reduced in both mutants the abundance of the internal peak was relatively stable or even slightly increased, suggesting that MT-CO2 5’ processing efficiency does not directly affect the abundance of the internal peak.

The numerous additional alternations in RNA cleavage beyond the mutated MT-TL1 indicate that there is some form of feedback causing uncoordinated RNA processing in these cells. The mechanism behind this observation is not known but it may result from the altered levels or activity of RNA processing enzymes or regulators. In support of this hypothesis, we have previously observed that PTCD1, which associates with ELAC2 and modulates processing of the 3′ ends of mitochondrial tRNAs (Sanchez et al., 2011), is increased in m.3243A>G and m.3302A>G cells. Further studies will be required to understand the coordinated regulation of mitochondrial RNA processing proteins.

Taken together we show that a single mutation in a mitochondrial tRNA can lead to a processing defect by that in turn can have downstream effects on global mitochondrial processing. Details of the order of mitochondrial RNA processing are still lacking and development of new next generation technologies can significantly improve our understanding of the mechanisms of RNA cleavage in mitochondria.

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgments: This project was supported by fellowships and project grants from the National Health and Medical Research Council (APP1058442, APP1045677, APP1041582, APP1023460, APP1005030, APP1043978 to A.F. and O.R.), Australian Research Council (to A.F. and O.R.), the Cancer Council of Western Australia (to O.R.). S.J.S. was supported by a University of Western Australia Australian Postgraduate Award.


Struct Biol 9, 586–590.
Figure Legends

Figure 1. Heatmap of canonical 5’ abundance.
Heatmap of the normalized PARE abundance at canonical 5’ mtRNA end positions in wild type, m.3243A>G and m.3302A>G cells (first panel) and the relative change compared to wild type and the other mutant cells (second panel). The histogram shows the relative distribution of 5’ ends of mt-tRNAs. Positions where PARE peaks were abundant across two consecutive nucleotides are labelled numerically.

Figure 2. The effects of m.3243A>G and m.3302A>G mutations on RNA19 processing.
PARE profile across the RNA19 region (n.p. 1,546-4,348), displayed in reads per million (RPM) on a three-way split scale. The areas highlighted in blue and red indicate junctions formed by rRNA 5’ and 3’ processing, respectively. The area highlighted in green shows an unidentified peak observed only in the m.3243A>G cybrid. The peaks highlighted in purple are only observed in the wild type cells and display characteristics of degradation products, such as a stepped appearance and non-template polyadenylated PARE tags.

Figure 3. The cleavage of non-canonical sites is altered in MT-TL1 mutant cell lines.
(A) PARE profile across the MT-ND5/MT-CYB junction (n.p. 13,548-16,049). The area highlighted in green displays the 5’ terminal mononucleotide of MT-CYB and its absence in all samples.
(B) PARE profile across the MT-CO1 region (n.p. 5,427-7,628). The areas highlighted in green show the cluster of internal peaks and downstream peak observed only in the m.3243A>G cells. The region highlighted in purple shows the absence of a peak at the 5’ terminal mononucleotide of MT-CO1 in all samples. The area highlighted in red shows the increased abundance of 3’ processing MT-TW products in the mutant cells, at a junction that lacks a proximal downstream gene. The area highlighted in blue shows the 5’ end of mt-tRNAs.

Figure 4. Mitochondrial RNAs with unusual 5’ or 3’ termini.
(A) PARE profile across the MT-ATP8/6-MT-CO3 junction (n.p. 8,165-10,166). The region highlighted in purple displays the 5’ terminal mononucleotide of MT-CO3 and a substantial peak that is only observed in the m.3243A>G cells. The area highlighted in blue shows the 5’ end and the area in red shows the 3’ end of mt-tRNAs.
(B) Terminator exonuclease was incubated with mtRNA for 15 or 30 min at 42°C to show that the canonical 5’ terminal mononucleotide of MT-CO3 is resistant to digestion compared to MT-ATP8/6 and MT-CO2.

(C) PARE profile across the MT-ND6 region (n.p. 12,258-14,759). The light strand annotation displays the ND6 coding region, and the long and short ND6 3’ UTR. The peaks highlighted in purple are located in the approximate region described by long UTR, and the PARE tags possess 3’ non-template adenosines. The area highlighted in blue shows the 5’ end and the area in red shows the 3’ end of mt-tRNAs.

Figure 5. Mutations in MT-TL1 lead to altered mitochondrial RNA metabolism.
(A) PARE profile across the MT-ND4L/4 region (n.p. 10,351-12,352). The peak highlighted in green is observed only in the m.3243A>G cells; the peaks highlighted in blue and red are junctions processed by RNase P and ELAC2, respectively.

(B) PARE profile across the MT-ND2 – WANCY region (n.p. 5,050-6,050). The region highlighted in green shows the cluster of peaks observed on the light strand antisense to the 3’ end of MT-ND2; the peaks highlighted in blue and red are junctions processed by RNase P and ELAC2, respectively.

Figure 6. Consequences of mutations in MT-TL1 on mitochondrial RNAs.
(A) PARE profile across the antisense MT-ATP8/6-MT-CO3 junction (n.p. 8,700-9,700). The peaks highlighted in green are observed in all samples and do not correspond with any known transcript end.

(B) PARE profile across the MT-CO2 region (n.p. 7,430-8,430). The peak highlighted in green does not correspond to any known transcript end and is observed in all samples; the peaks highlighted in blue and red are junctions processed by RNase P and ELAC2, respectively.
Table 1: PARE tag abundance at non-canonical positions across the mitochondrial genome.

Normalised abundances (reads per million; RPM) and relative changes (log2 fold change of RPM values plus a pseudocount of 1) of PARE tags at non-canonical positions.
Figure 1

Color Key and Histogram

Log10RPM

Color Key and Histogram

Log2FoldChange

Figure(s)
Figure 4
Table S1: PARE tag abundance at canonical 5’ mtRNA end positions.

Normalised abundances (reads per million, RPM) and relative changes (log2 fold change of RPM values plus a pseudocount of 1) of PARE tags at canonical 5’ mtRNA end positions. Positions where PARE peaks were abundant across two consecutive nucleotides are labelled with consecutive numerals.
Supplementary Figure 1. Mitochondrial protein synthesis in wild type and mutant cell lines.

Mitochondrial translation in wild type and mutant cells was measured by pulse incorporation of ³⁵S-labeled methionine and cysteine in the presence of cycloheximide. Equal amounts of cell lysates (50 mg) were separated by SDS-PAGE, immunoblotted for VDAC to show equal loading and visualized by autoradiography. Representative gels are shown of three independent biological experiments.
Paper III
LRPPRC-mediated folding of the mitochondrial transcriptome

Stefan J. Siira¹, Henrik Spåhr², Anne-Marie J. Shearwood¹, Benedetta Ruzzenente²,³, Nils-Göran Larsson²,⁴, Oliver Rackham¹,⁵† and Aleksandra Filipovska¹,⁵†

¹Harry Perkins Institute of Medical Research and Centre for Medical Research, The University of Western Australia, Nedlands, Western Australia 6009, Australia.
²Department of Mitochondrial Biology, Max Planck Institute for Biology of Ageing, D-50931 Cologne, Germany.
³Inserm U1163, Université Paris Descartes-Sorbonne Paris Cité, Institut Imagine, 75015 Paris, France
⁴Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm 17177, Sweden
⁵School of Molecular Sciences, The University of Western Australia, Nedlands, Western Australia 6009, Australia.

Running title: In vivo RNA chaperone activity of LRPPRC

†Correspondence should be addressed to: oliver.rackham@uwa.edu.au or aleksandra.filipovska@uwa.edu.au

Keywords: RNA metabolism, transcriptomics, PPR domains
Abbreviations used are: mitoribosome, mitochondrial ribosome; LRPPRC, leucine rich pentatricopeptide repeat containing protein; PPR, pentatricopeptide repeat; SLIRP, SRA stem-loop interacting RNA-binding protein.
Abstract

The roles of RNA-binding proteins as chaperones in the lifecycles of mRNAs are not well understood. The mammalian mitochondrial genome has been compressed over evolution to a size of just 16 kb, nevertheless the expression of its genes requires transcription, RNA processing, translation and RNA decay, much like the more complex chromosomal systems, providing an opportunity to use it as a model system to understand the fundamental aspects of gene expression. Here we combine RNase footprinting with PAR-CLIP at unprecedented depth to reveal the importance of RNA-protein interactions guided by the LRPPRC/SLIRP complex in dictating RNA folding within the mitochondrial transcriptome. We show that LRPPRC, in complex with its protein partner SLIRP, binds throughout the mitochondrial transcriptome, with a preference for mRNAs, and its loss affects the entire secondary structure and stability of the transcriptome. We demonstrate that the LRPPRC/SLIRP complex is a global RNA chaperone that stabilizes RNA structures to expose the required sites for translation, stabilization and polyadenylation. Our findings reveal a general mechanism where extensive RNA-protein interactions ensure that RNA is accessible for its biological functions.
Introduction

RNA-binding proteins (RBPs) regulate the lifecycles of RNAs from transcription to degradation and as such are important modulators of gene expression. This is particularly evident in mitochondria where gene expression is predominantly regulated by nuclear encoded mitochondria localized RBPs (mtRBPs)\(^1\,^2\). The importance of nuclear and mitochondrial RBPs is exemplified by the number of different disorders caused by mutations in genes encoding these proteins, including neurological conditions, metabolic diseases and cancer\(^3\,^4\).

Recent efforts have focused on transcriptome-wide identification of new RNA-binding proteins in different cell types and organisms, as well as validation of already known RNA-binding proteins or discovery of those with dual functions\(^5\). Consequently, an increasing number of these proteins have been assigned to different families/classes of RNA-binding proteins, thought to play a role in RNA metabolism based on their sequence or fold homologies. Although we have insight into the RNA targets of a few well-characterized RBPs, the targets of most RBPs identified by high-throughput techniques are not known. Considering that RBPs coat RNAs extensively and thereby regulate their stability, translation and localization, it is essential to understand not only how they bind but more importantly the transcriptome-wide consequences of their loss in cells and organisms. This has been investigated at a genome level where loss of transcription factors, histones or histone modifying enzymes affect genome organization and transcription\(^6\), but this is lacking for post-transcriptional regulation of gene expression by RBPs both in the cytosol and mitochondria. Currently, the major goals in this field are the in-depth identification of RNA targets for RBPs and particularly characterizing their mechanism of action \textit{in vivo}. Our study provides a model of transcriptome-wide consequences of mitochondrial RBP loss on RNA accessibility and secondary structure, revealing the critical role of a RBP as a RNA chaperone.
The sequence specific RNA-binding repeat proteins, including the Pumilio and FBF homology (PUF) and pentatricopeptide repeat (PPR) proteins, have been of particular interest because of their modular recognition of their RNA targets and their potential use as designer RBPs. Although the RNA targets of PUF proteins have been identified and validated biochemically, structurally and in vivo, the physiological targets of PPR proteins have remained largely unknown, limiting our understanding of their mechanistic roles in organelle gene expression and energy metabolism. This is particularly true for the mammalian PPR proteins, where there are only seven PPR domain proteins that are all localized to mitochondria. Cellular and in vivo studies have identified that these proteins regulate different aspects of mitochondrial RNA metabolism; from transcription via the mitochondrial RNA polymerase (POLRMT), to RNA processing (PPR domain containing proteins 1 and 2, PTCD1 and PTCD2, mitochondrial RNase P protein 3, MRPP3), maturation and stability (the leucine-rich PPR cassette protein, LRPPRC), and protein synthesis (the mitochondrial ribosomal protein of the small subunit 27, MRPS27, and PTCD3).

Although the RNA recognition code for plant PPR proteins has been identified computationally and using in vitro approaches, and based on these findings artificial PPR proteins have been designed to modulate gene expression, the RNA targets of mammalian PPR proteins remain elusive.

Here we use the mitochondrial transcriptome as a compact model to investigate and identify RNA-protein interactions at an unprecedented depth that has not been possible for the nuclear transcriptome. Mitochondrial genes coding for proteins and tRNAs are located on both the heavy and light strands of the mtDNA, and are transcribed as large polycistronic transcripts covering almost the entire length of each strand. This differs from nuclear gene transcription and requires several unique processing and maturation steps to form functional RNAs, all of which depend on mtRBPs. In nearly all cases, genes encoding protein or rRNA
are interspersed by one or more tRNAs, that act as “punctuation” marks for processing, recruiting the RNase P complex or RNase Z that carry out the cleavage at the 5' end or 3' end of tRNAs, respectively. Processing is followed by maturation of the RNAs by a range of mRBP including, the LRPPRC/SLIRP complex, poly(A) polymerase, methylases, tRNA modifying enzymes, and the CCA-addition enzyme, all necessary for the assembly of the tRNAs into mitochondrial ribosomes and translation of the mRNAs using tRNAs (reviewed in).

To understand the role of the LRPPRC/SLIRP RNA-protein complex in vivo, we sought to identify its RNA targets by photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) and high throughput RNase footprinting of mitochondrial transcriptomes from mice where each of the genes coding for these proteins are knocked out. LRPPRC was identified when a mutation in its gene was shown to cause a rare French-Canadian variant of Leigh syndrome with a cytochrome c oxidase deficiency. LRPPRC is one of the longest PPR proteins known with 30-33 PPRs and does not contain any other functional domains. LRPPRC forms a stable complex with SLIRP via a PPR-RRM protein interface, where SLIRP acts to protect LRPPRC from degradation in vivo. Both proteins within the complex are required to regulate the rate of mitochondrial protein synthesis and the stability of the poly(A) tails of mitochondrial mRNAs, however the binding sites of the LRPPRC/SLIRP complex and its molecular role are not known.

Here we have investigated transcriptome-wide footprints of both LRPPRC and SLIRP and the changes in molecular interactions and RNA secondary structure in their absence. We identify that single strand regions on mitochondrial mRNAs are the primary targets of the LRPPRC/SLIRP complex in vivo, indicating the role of this complex in RNA remodelling required to enable polyadenylation and translation. Furthermore we show transcriptome-wide changes in RNA interactions in the absence of these proteins indicating that the
LRPPRC/SLIRP complex acts as a global mt-RNA chaperone required to relax secondary structures. Our findings reveal the role of RBPs as chaperones for the lifecycle of mitochondrial mRNAs and provide insight into their recognition modes by investigating them at a much greater depth than previously possible.

RESULTS

High-throughput identification of LRPPRC and SLIRP in vivo footprints

To reveal the in vivo binding landscape of the LRPPRC/SLIRP complex we treated mitochondria isolated from hearts of conditional Lrpprc or whole body Slirp knockout mice, where these proteins are lost (Supplementary Fig. 1), with three endonucleases that have different cleavage specificities: (i) RNase A cleaves single- and double-stranded RNA after pyrimidine nucleotides, (ii) RNase T1 specifically cleaves single-stranded RNA after guanine residues and (iii) RNase If, which is a sequence-independent endonuclease that preferentially cleaves single-stranded RNA over double-stranded RNA. The RNA isolated from these preparations was sequenced to enable transcriptome-wide in vivo mapping of footprints of each RBP. Mock-digested mitochondrial preparations were used as undigested controls. The data sets were aligned to the mitochondrial genome and normalized by sequencing depth; a cleavage score (C score) was determined for each nucleotide across the mitochondrial transcriptome for all data sets from coverage profiles of the 5′ nucleotide of each read, as previously described. Based on these C scores, we identified regions within mitochondrial transcripts protected from endonuclease cleavage by LRPPRC or SLIRP in the control data sets. For each identified region a footprint score (F score) was calculated based on the central RNase accessibility relative to the flanking regions in both controls (LrpprcloxP/loxP or Slirp+/+) and knockout (KO; LrpprcloxPloxP.Cre+ or Slirp−/−) data sets. A significant increase in the F score in the KO relative to the WT datasets enabled us to identify 178 different footprints for
LRPPRC that were dispersed throughout mt-RNAs, the majority concentrated on mRNAs and to a lesser degree in rRNAs (Fig. 1 and Supplementary Fig. 2). There were fewer footprints in tRNAs and only two in the regulatory D-loop region of the mitochondrial genome (Supplementary Fig. 2), consistent with the reported roles of LRPPRC in mitochondrial mRNA polyadenylation and translation. This is also in agreement with previous in vitro data where recombinant LRPPRC preferentially binds single stranded mRNAs rather than tRNAs. Most of the footprints were localized to mRNAs, and some LRPPRC footprints were particularly enriched across the entire length of transcripts, such as the Co1 mRNA, as well as Atp8/6 and Co3 mRNAs that often form a tricistronic transcript (Fig. 1). These findings suggest that LRPPRC binds regions of the mt-RNAs that are involved in processes such as RNA maturation, stability and translation. We observed extensive footprints along the two rRNAs suggesting that LRPPRC may act as an RNA chaperone to facilitate rRNA processing and recognition by ribosomal proteins, since we now know that rRNA processing and ribosome assembly occur co-transcriptionally. The identified footprints suggest that these sites are regulated by LRPPRC to enable their efficient maturation and recognition by the ribosome for translation. We analyzed footprints in the SLIRP knockout data sets to discover that there was no significant enrichment of footprints (Supplementary Fig. 3), indicating that SLIRP does not bind RNA in vivo confirming previous in vitro biochemical findings.

An In vivo LRPPRC/SLIRP binding signature identified by PAR-CLIP

We performed PAR-CLIP on mouse embryonic fibroblast (MEF) cells isolated from Lrpprc-FLAG transgenic mice to identify the direct binding sites of LRPPRC and to compare them with those identified by our footprinting. The incubation of MEF cells with 4-thiouridine (4TU) and its incorporation into de novo transcripts, followed by crosslinking and
immunoprecipitation, enabled us to map the precise LRPPRC binding sites by identifying sites with thymidine (T) to cytidine (C) transition rates indicative of cross-linking in the sequenced data set. The specificity of LRPPRC binding was established when compared to a control PAR-CLIP experiment that did not result in specific alignment to the mitochondrial transcriptome (Supplementary Table 1). We identified a number of direct and specific LRPPRC binding sites, all located within mRNAs and rRNAs but not tRNAs, and the PAR-CLIP coverage across cross-linked sites is shown in Fig. 2a. Here again we find discrete binding sites of LRPPRC that are most often located throughout mRNAs, which require polyadenylation for stabilization and their coordinated translation, as shown before20. In the bicistronic mRNAs, Atp8/6 and Nd4l/4 we found binding sites across the middle of the transcripts in the 3′ ends of the Atp8 or Nd4l open reading frames (ORFs) and the 5′ ends of the Atp6 or Nd4 ORFs, suggesting that LRPPRC may be required to expose the internal start and stop codons of these transcripts to the ribosome for translation initiation or termination. Interestingly, we found LRPPRC binding sites within the 16S rRNA that were concentrated predominantly at either end of this transcript, suggesting that because of its affinity for transcript ends LRPPRC could facilitate rRNA processing, maturation and ribosome assembly. LRPPRC binding may facilitate the processing of the 16S rRNA from the RNA19 precursor transcript since loss of LRPPRC only decreases the levels of tRNA19 that is also contained within this precursor transcript20.

To validate the identified in vivo binding sites of this complex we carried out RNA electrophoretic mobility shift assays (RNA EMSAs) using targets identified in both the footprinting and PAR-CLIP analyses. We show that the mouse LRPPRC/SLIRP protein complex has a high affinity for a region within the mt-Co3 mRNA identified both in the footprinting and PAR-CLIP data sets (Fig 2b) and that the recombinant mouse SLIRP alone does not bind RNA (Supplementary Fig. 4) as previously identified for human SLIRP32.
Since the LRPPRC/SLIRP complex has high RNA-binding affinity with a large variety of RNA targets we included a competitor RNA to mimic the physiological conditions in vivo. We show that the LRPPRC/SLIRP complex binds with higher affinity to specific regions within transcripts compared to other target regions of mitochondrial transcripts in vivo (Supplementary Fig. 5). The binding of these targets is higher compared to control or less enriched binding targets indicating further that these are LRPPRC binding sites. Nevertheless we found that LRPPRC can bind to a wide variety of RNA targets in vitro and in vivo and might act as a general mRNA-binding factor in mitochondria, and that regions that are bound more frequently might be determined by a combination of binding affinity and accessibility.

Next we defined a set of high confidence direct RNA targets of LRPPRC by selecting a set of sites that overlap between our footprinting and PAR-CLIP datasets (Fig. 2c). The overlap between the two datasets was statistically significant and included the majority of PAR-CLIP sites (Supplementary Table 2). The direct binding sites were found predominantly in mRNAs and rRNAs, indicating that tRNAs are not direct targets of LRPPRC in vivo (Supplementary Fig. 6a). This was further confirmed by analysing the footprints that did not overlap with the high confidence binding sites identified by PAR-CLIP, which were mainly distributed on tRNAs and rRNA (Supplementary Fig. 6b), indicating that these sites are protected from cleavage upon LRPPRC binding through the recruitment of other proteins or secondary structure changes. Motif searching using MEME identified a consensus binding sequence for LRPPRC (shown in red in Fig. 2a). These enriched sequences were somewhat degenerate (Fig. 2d), reflecting the preponderance of different sites recognised throughout the mitochondrial transcriptome.

Transcriptome-wide consequences of LRPPRC loss
In contrast to the direct LRPPRC binding sites, which are enriched in mRNAs and almost entirely absent from tRNAs, the footprints defined by comparing RNase accessibility in knockout and wild-type animals are found in all mt-RNAs, suggesting that the binding of LRPPRC affects the RNA structure and accessibility to other RNA-binding proteins. Therefore to investigate transcriptome-wide consequences of LRPPRC loss we identified footprints in the LRPPRC knockout mice and examined their changes relative to control data sets (Fig. 3a). Here, we identified regions within mitochondrial transcripts protected from endonuclease cleavage by the absence of LRPPRC in the KO data sets. We identified 124 footprints with significant reduction of the F score in the knockout relative to the control samples, indicative of either RBP or secondary structure changes that render these specific sites resistant to RNase cleavage in the absence of LRPPRC (Fig. 3a). For example, we observe an appearance of a footprint in the mt-Nd5 mRNA in the Lrpprc KO mice that may be a result of a change in secondary RNA structure or caused by another RBP that associates with this mRNA in the absence of LRPPRC (Fig. 3b). The number of footprints in the KO data set is less than the LRPPRC footprints identified in Fig. 1, suggesting that LRPPRC has a general RNA chaperone role in regulating RNA maturation, stability and translation. The same analyses in the SLIRP knockout data set failed to identify any significantly enriched footprints further confirming that SLIRP does not exert any direct effect on RNAs.

Next we investigated if RNA secondary structure can affect in vivo RBP interactions with the mitochondrial transcriptome in the presence and absence of the LRPPRC/SLIRP complex. We treated isolated mitochondria from control, Lrpprc and Slirp knockout mice with RNase I to specifically degrade single-stranded RNA and calculated an R score for each nucleotide across the transcriptome, where a higher R score is indicative of a greater propensity for strong secondary structure. We identified that loss of LRPPRC results in a net increase in the secondary structure of mitochondrial transcripts whereas loss of SLIRP does
not affect them (Supplementary Table 3). We examined the distribution of changes in average R score within footprints that overlapped PAR-CLIP binding sites and identified that LRPPRC binding reduced the RNA secondary structure within its target sites (Fig. 3c). Furthermore the average R score of regions flanking LRPPRC binding sites was lower in the presence of LRPPRC (Supplementary Fig. 7a), therefore LRPPRC acts to prevent secondary structure formation where it binds and facilitates opening of flanking regions of RNAs, and preventing RNA within its binding site from interacting with complementary regions. We observed this effect of LRPPRC binding on mRNAs and rRNAs (Supplementary Fig. 7a-c).

To validate the in vivo finding that LRPPRC and the LRPPRC/SLIRP complex can act as RNA chaperones biochemically we incubated purified proteins with complementary RNA oligonucleotides. Over time we observed that the presence of the LRPPRC/SLIRP complex blocked hybridization between the complementary RNA oligonucleotides (Fig. 3d). LRPPRC/SLIRP did not facilitate the annealing of complementary RNA oligonucleotides (Fig. 3d). This likely reflects that LRPPRC adheres to the canonical mode of RNA recognition of PPR proteins were RNA adopts an extended conformation where the bases are recognised by direct hydrogen bonding interactions from consecutive repeats. To confirm this in vivo we investigated the effects of LRPPRC loss on mitochondrial RNAs by RNA-Seq. We identify decreased polyadenylation of mitochondrial mRNAs (Supplementary Fig. 8a) and confirmed the reduction in poly(A) tail length in the absence of LRPPRC using modified 3′-RACE (Supplementary Fig. 8b). The loss of polyadenylation causes reduced stability of all mRNAs except mt-Nd6 (Supplementary Fig. 9), confirming previous findings 20, and corroborating that Nd6 is the only mt-mRNA that is not polyadenylated (Supplementary Fig. 8a). Taken together our results suggest that the LRPPRC/SLIRP complex acts to stabilize regions of single strandedness in mitochondrial RNAs and expose them for modification and maturation that enables coordinated translation.
The evolutionary conservation of LRPPRC/SLIRP binding in mouse and man

Given LRPPRC’s broad substrate specificity we wondered whether the locations of its binding sites were conserved through evolution. We performed PAR-CLIP in HeLa cells stably expressing LRPPRC that was FLAG tagged at the C-terminus and identified the binding sites of this protein in human mitochondria. The last common ancestor shared between human and mouse existed 96 million years ago\(^3\) and their mtDNAs are 70% similar at the nucleotide level, such that most LRPPRC binding sites vary in primary sequence between them. Nevertheless we identified a very similar profile of binding sites when we compared the PAR-CLIP coverage profiles for LRPPRC in human cells to those identified in mouse cells (Fig. 4a). To determine if these similarities resulted from co-evolution of the LRPPRC proteins and their RNA targets or from restriction by other proteins or RNA secondary or tertiary structures, we assayed binding of RNA targets derived from a shared binding location in \(mt\text{-}Cyb\) that differed in sequence between human and mouse (Fig. 4b). We found that the mouse LRPPRC/SLIRP complex preferentially bound the mouse \(mt\text{-}Cyb\) RNA target, while the human LRPPRC/SLIRP complex preferred the human \(MT\text{-}CYB\) RNA target (Fig. 4c). These results show that the preferred locations of LRPPRC binding sites within mitochondrial RNAs are restrained through evolution, likely illustrating their importance in dictating the local RNA structures that are critical in the lifecycles of mitochondrial RNAs.

DISCUSSION

The compact organization of the mitochondrial genome has resulted in a shift in the regulation of its transcriptome, from the transcriptional regulation of its proteobacterial ancestor to regulation predominantly carried out at the post-transcriptional level by RBPs, similar to gene regulation in the eukaryotic cytosol. Indeed, here we found that mtRNAs were
extensively bound by RBPs, as intimated by previous studies in human mitochondria where the majority of the transcriptome was bound by mtRBPs. As well as recruiting enzymatic activities and other protein complexes to RNAs, RBPs are often required to facilitate structural rearrangements of RNAs. The best characterized of these are the RNA helicases, that use ATP to unwind RNA helices with high processivity. The importance of RNA helicases is well established in processes such as RNA splicing, where a number of helicases are required to rearrange base paired structures between splicesomal RNAs and their pre-mRNA targets. Less well studied, however, are ATP-independent RNA chaperones, whose activities do not require ATP, as they do not share a common evolutionary origin and rather are characterized by using strong direct interactions with RNAs to remodel their structures. Here we reveal that the LRPPRC complex fulfils this role in mammalian mitochondria.

LRPPRC binding sites predominantly decorated the mRNAs and resulted in disruption of local mitochondrial RNA secondary structures. This suggests that relaxing the RNA structure by LRPPRC/SLIRP at these sites facilitates the coordinated translation and stability of mRNAs via polyadenylation. Also at the 3′ ends of mRNAs, LRPPRC/SLIRP could be required to present single stranded RNA ends to the mitochondrial poly(A) polymerase (MTPAP) to facilitate polyadenylation. Evidence for this comes from in vitro experiments where addition of LRPPRC increased the processivity of MTPAP and that loss of LRPPRC causes a polyadenylation defect.

The mechanisms by which mitochondrial ribosomes recognize mRNAs and initiate translation are not currently understood. The start codons of mitochondrial mRNAs are typically found at the extreme 5′ ends with little or no 5′-UTR sequence. Our observation that the RNA chaperone, LRPPRC/SLIRP, relaxes RNA structures of mRNAs, may indicate that the mitoribosome requires the start codons to be exposed in order to initiate translation. Furthermore, it is possible that LRPPRC’s chaperone activity might assist in the production of
tRNAs. Since mRNAs are typically flanked by tRNAs, LRPPRC binding at the termini of mRNAs could prevent these flanking regions from interfering with tRNA folding, processing or modification. Indeed, previous analyses of the steady state levels of tRNAs found that a proportion were altered in the absence of LRPPRC\textsuperscript{20}.

LRPPRC has more PPRs than almost any other reported PPR protein, containing 33 predicted PPRs. Although structural information is only available for a few PPR proteins to date, it appears that they can often interact with their RNA targets in a binding mode where each PPR recognizes a single nucleotide base\textsuperscript{37}. This might be important for LRPPRC’s ability to bind many diverse RNAs, since only a subset of repeats would need to recognize individual RNA bases in order to achieve sufficient binding affinity for each individual RNA target. Nevertheless the specificity of LRPPRC’s PPRs is still important to target critical regions of the transcriptome, as exemplified by the co-variation we observed between LRPPRC and its RNA targets through the evolution of human and mouse.

Here we have shown that loss of a global RNA chaperone can lead to remodelling of secondary RNA structures as well as other RNA-protein interactions. Studies of the RNA binding sites of eukaryotic nuclear and cytosolic RBPs have revealed sparse distributions of binding sites across individual mRNAs. However given that there are an estimated 1,542 human RBPs\textsuperscript{38}, as an ensemble they likely act as LRPPRC does in the mitochondrial transcriptome to coat RNAs and assist in their folding. Therefore the binding of LRPPRC across the mitochondrial transcriptome represents a minimal system to study the importance of ATP-independent protein-mediated RNA folding. The massively deep coverage of the mitochondrial transcriptome in our sequencing experiments has enabled us to elucidate the in vivo binding modes of the LRPPRC/SLIRP complex and establish a model approach of how to investigate the effects of RBP loss on a global transcriptome-wide level.
EXPERIMENTAL PROCEDURES

Animals and Housing

LRPPRC and SLIRP knockout transgenic mice on a C57BL/6N background were housed in standard cages (45 cm × 29 cm × 12 cm) under a 12-h light/dark schedule (lights on 7 am to 7 pm) in controlled environmental conditions of 22 ± 2 °C and 50 ± 10% relative humidity and fed a normal chow diet and water were provided ad libitum. LRPPRC loss is embryonic lethal therefore there are no full-body Lerppc knockout mice and we used heart and skeletal muscle-specific knockout mice, as described previously\(^20\). Also SLIRP loss is not embryonic lethal and consequently we used hearts from the full-body Slirp knockout mice, as described previously \(^30\). The study was approved by the Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen and performed in accordance with the recommendations and guidelines of the Federation of European Laboratory Animal Science Associations (FELASA).

Mitochondrial isolation and RNase treatment

Mitochondria were isolated from homogenized hearts and isolated by differential centrifugation as described previously\(^30,39\), with some modifications. Hearts or livers were cut and washed three times with ice cold PBS, and once with mitochondrial isolation buffer (MIB) containing 310 mM sucrose, 10 mM Tris-HCl and 0.05% BSA (w/v) by centrifugation at 4,500 × g for 1 min at 4°C. Heart pieces were homogenized in 5 ml of fresh MIB using a Potter S pestle (Sartorius). The homogenate was centrifuged at 1,000 g for 10 min at 4°C and the supernatant was centrifuged at 4,500 g for 15 min at 4°C to isolate mitochondria. Crude mitochondrial pellets were suspended in MIB supplemented with 1x Complete EDTA-free protease inhibitor cocktail (Roche). Protein concentration was determined by the Bradford or
BCA method using BSA as a standard. Mitochondria (2 mg ml\(^{-1}\)) were lysed by addition of 200 μl of lysis buffer (100mM Tris-HCl, 100mM NaCl, 40mM MnCl\(_2\), 2mM dithiothreitol pH 7.5, 0.1% TritonX-100). The concentration of the RNase A (10 U μl\(^{-1}\)), RNase T1 (0.1 U μl\(^{-1}\)) or RNase If (0.01 U μl\(^{-1}\)) were optimized and added to each mitochondrial lysate or purified RNA to generate 15-55 nt size fragments. All incubations were carried out at 37˚C for up to 30 minutes and reactions were ended by addition of 700 μl Qiazol, followed by RNA isolation using the miRNeasy Mini Kit (Qiagen). We used mitochondria lysed and mock treated as controls for the footprinting assay.

**Immunoblotting**

Immunoblotting was carried out using mitochondrial proteins on mitochondria isolated from mouse hearts as previously described\(^30\).

**Library construction**

RNA concentration, purity and integrity was confirmed by BioAnalyser (Agilent, CA). The libraries were constructed using the Illumina TruSeq Small RNA Sample Prep Kit and deep sequencing of the mitochondrial small RNAs was performed by Australian Genomic Research Facility (Melbourne, Australia) on an Illumina GAII (Illumina, CA) according to the manufacturer’s instructions with one modification, sample isolation from the PAGE gel after adaptor ligation was performed with a modified set of size markers to facilitate capture of small RNAs between 15-55 nt.

**Mapping and identification of footprints**

Technical replicates were pooled and sequenced reads trimmed of adapter sequences with cutadapt v1.10, using default parameters, and aligned to the mouse genome (mm10) with
Bowtie2\(^{20}\) v2.2.9 with a seed length of 10 and reporting up to 20 alignments per read (-L10 -k 20). Paired alignments with multiply mapping reads that aligned once to the mitochondrial genome and at least once to a NUMT region were rescued. All properly paired alignments to the mitochondrial genome with an observed template length of 15 to 35 nt were retained and a subtractive alignment against nuclear tRNA and Illumina contaminant sequences was performed. Strand-specific fragment BED files were created and 5’ coverage profiles normalised to sequenced library size were produced with BEDtools\(^{41}\) v2.26.0. Footprints were identified as previously described\(^{33}\) with some modifications reported here. The RNase accessibility of each base, i, in the mitochondrial genome was quantified according to its C score, defined as:

\[
C_{score_i} = \log_{10} \left( \frac{\max(A_{i+1}, T_{1i+1}, I_{Fi+1}) + 1}{\text{Untreated}_{i+1} + 1} \right)
\]

where \(A_{i+1}, T_{1i+1}, I_{Fi+1}\) and \(\text{Untreated}_{i+1}\) represent the normalised 5’ coverage of the nucleotide immediately downstream of the inspected position in RNase A-, T1-, and If-treated and untreated samples, respectively. The footprint detection algorithm searches for a span of between 8 and 40 nucleotides with an average C score lower than the flanking left and right three nucleotides, the central footprinting region, and calculates its F score:

\[
F_{score} = \frac{10^C}{10^L} + \frac{10^C}{10^R}
\]

where C, L and R are the average C score of the central, left flanking and right flanking regions, respectively. We initially searched for footprints in the control data and for each footprint we calculated its \(F_{score_{control}}\) and the F score of the equivalent region in the
knockout sample, F score_{experiment}, and the log_2-transformed fold change of F scores (knockout/control). We also searched for novel RBP interactions or sites of decreased RNase accessibility by searching for footprints in the experimental data initially, before calculating F scores and extracting the data from the equivalent region in the control sample. To estimate if the F score fold change of a footprint is significant, we built an empirical null model by shuffling the C score of both samples within the mitochondrial genome 1,000 times and for each locus, calculating an F score fold change from each pair of shuffled data sets. For the candidate footprint loci, the central footprint region was required to have a C score_{experiment} greater than its C score_{control}, and the F score fold change was filtered to achieve an expected 5% false discovery rate (FDR) relative to a score obtained from random shuffling 1,000 times, as previously described. The propensity for secondary structure was quantified according to the ratio of double-stranded RNase cleavage over single-stranded, its R score:

\[ R_{score_i} = \log_2 \left( \frac{A_{i+1}}{F_{i+1}} \right) \]

Violin plots were generated with the vioplot package, and the log_2 fold change of R scores calculated as:

\[ \log_2 \left( \frac{2^{R_{score_{experiment}}}}{2^{R_{score_{control}}}} \right) \]

Control footprint analyses of the cytoplasmic 28S rRNA were carried out in an identical manner, except alignment was performed to the GENCODE vM14 reference transcript set combined with the 28S rRNA sequence downloaded from GenBank (NR_003279.1), and no significant footprints were identified (Supplementary Table 7).

**PAR-CLIP method and analyses**

Mouse embryonic fibroblasts from Lrpcc-FLAG transgenic mice generated as previously described, were maintained in DMEM medium (Thermo Fisher Scientific) supplemented with...
with GlutaMAX-I, 10% fetal bovine serum, 1% penicillin-streptomycin and 1X MEM non-essential amino acids. HeLa Tet-On cells stably transformed with pTRE-hLRPPRC-FLAG were generated as previously described and maintained in DMEM medium supplemented with GlutaMAX-I, 10% fetal bovine serum, 1% penicillin-streptomycin and 200 μg/ml hygromycin. Induction of LRPPRC-FLAG expression was achieved by addition of 2 μg/ml doxycycline to the incubation medium for 72 h. PAR-CLIP was performed according to Spitzer et al., with the exception that cells were lysed in lysis buffer (10 mM Tris-HCl (pH 7.5), 260 mM sucrose, 100 mM KCl, 20 mM MgCl2, 1% digitonin) in the presence of RNase inhibitor (New England Biolabs) and protease inhibitors (Roche). Lysates were clarified by centrifugation and diluted in lysis buffer without digitonin until the final digitonin concentration was 0.2%. Diluted lysates were incubated with anti-FLAG M2 magnetic beads for 2.5 h at 4°C. Beads were washed in lysis buffer with 0.1% digitonin and then in lysis buffer without digitonin. RNA library preparation and sequencing was performed by Vertis Biotechnologie (Freising, Germany) using unique molecular identifiers to identify PCR duplicates according to Kivioja et al.

The PAR-CLIP libraries were prepared using indexed adapter containing an 8 bp barcode followed by 8 random bases to allow multiplexing and unambiguous detection of PCR duplicates. Following de-multiplexing and de-duplication, adapter sequences were removed from PAR-CLIP reads with cutadapt, requiring a minimum length after trimming of 14 nt. As PAR-CLIP libraries detect crosslinked regions utilising photoactivatable ribonucleosides, the reads will contain a large number of sequence mismatches relative to the reference. To prevent misalignment of mitochondrial reads to nuclear mitochondrial sequences the mouse and human genome reference sequence (mm10 and hg38) was masked for these regions based on the mm9 and hg19 NumtS tracks from UCSC, with coordinates converted to mm10 and hg38 with UCSC liftOver. Trimmed reads
were aligned to the Numts-masked genome sequence with Bowtie\textsuperscript{49} v1.12 (-k 1 -n 2 -M 100 --
chunkmbs 512 --best --strata). Cross-linked regions were identified with BMix\textsuperscript{50}, using
default parameters (minimum coverage of 5, refinement coverage of 1, a minimum posterior
probability of crosslinking of 95% and strand-specific parameter calculation). MEME\textsuperscript{51} 4.11.2
was used to search for ungapped motifs within the binding sites identified by BMix (-rna
-time 18000 -maxsize 60000 -mod anr -nmotifs 5 -minw 6 -maxw 50). Strand-specific
coverage profiles were generated with BEDtools v2.26.0\textsuperscript{41}. To assess the regions of the
mitochondrial genome that were found to be overlapped by both PAR-CLIP and RNase
footprinting, the mitochondrial genome was binned into 50 nt windows with BEDtools
v2.26.0\textsuperscript{41}. The number of PAR-CLIP sites or RNase footprints overlapping each region was
calculated and the significance of the number of sites overlapped by both was tested with a
hypergeometric test.

**Protein purification**

Recombinant human and mouse LRPPRC and SLIRP proteins were expressed and purified as
previously described\textsuperscript{32}.

**In vitro RNA Binding Assays**

RNA Electrophoretic Mobility Shift Assays were performed as previously described\textsuperscript{32}. RNA
chaperone activity was assayed according to Rajkowitsch \textit{et al}\textsuperscript{31}. Briefly, the following
oligoribonucleotide 21R+: 5′-Cy3-AUGUGGAAAAUCUCUAGCAGU-3′ and its
complement 21R–: 5′-Cy5-ACUGCUAGAGAUUUUCCACAU-3′ (Dharmacon) were
hybridized in the presence or absence of 1 μM of mouse LRPPRC/SLIRP complex. Aliquots
were removed at specific time points, mixed on ice with stop buffer and separated on a 20%
native polyacrylamide gel at 4°C. Fluorescence was detected with a Typhoon FLA 9500 (GE
21
Capture of polyadenylated mRNA ends by PCR

Polyadenylation of mt-Nd1, mt-Cyb and mt-Nd2 was examined using a modified 3′-RACE method as previously described\textsuperscript{52}.

RNA-Seq analyses

RNA sequencing was performed on total RNA from four control and four \textit{Lrpprc} knockout mouse hearts on an Illumina HiSeq platform, according to the Illumina Tru-Seq protocol. We used random hexamer primers for cDNA library generation and carried out cytoplasmic rRNA depletion using the Ribo-Zero rRNA removal kit. Sequenced reads were aligned against the mouse transcriptome with RSEM v1.2.31 (--bowtie2 --strandedness reverse --paired-end) using the GENCODE vM13 transcript set combined with custom mitochondrial rRNA/mRNA sequences with an additional 50 adenine residues added to the 3′ ends to allow mapping of polyA tails. Gene-level counts were imported and analysed for differential expression with tximport v1.4.0\textsuperscript{53} and DESeq2 v1.16.1. Coverage profiles for each mitochondrial transcript were generated from full fragment BED files with BEDtools v2.26.0\textsuperscript{41}.  

\textsuperscript{134}
Acknowledgements: This project was supported by fellowships and project grants from the National Health and Medical Research Council (APP1058442, APP1045677, APP1041582, APP1023460, APP1005030, APP1043978 to AF and OR), Australian Research Council (to AF and OR), the Cancer Council of Western Australia (to OR), Swedish Research Council (Rådsprofessor 2015–0418 to NGL) and the Alexander von Humboldt Foundation (to AF, OR and NGL). SJS was supported by an Australian Postgraduate Awards (APA) Scholarship. We thank Dr Yuliya Karpievitch and Dr Ganqiang Liu for advice on biostatistics and analyses, Judith Ermer for technical assistance, and AGRF for advice and help with library construction.

Author contribution. S.S., O.R. and A.F. conceived the project. O.R. and A.F. designed the experiments. S.S., H.S., A.-M.J.S., B.R., N.-G.L., O.R. and A.F. conducted and analyzed the experiments. S.S., O.R. and A.F. wrote the manuscript and the other authors approved the manuscript.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. All sequencing data generated in this study are available at the NCBI Gene Expression Omnibus (GEO, accession number GSE100733). Additional data related to this paper may be requested from the authors.
References:


FIGURE LEGENDS

**Figure 1.** (a) Schematic showing the principle of high throughput RNA footprinting. (b) A circular representation of the mitochondrial genome (centre track) displaying LRPPRC footprints (exterior tracks) coloured by the log2 fold change of F score (scale: 0.0 – 2.5). Control samples were scanned for footprints and compared to *Lrpprc* knockout samples to identify LRPPRC binding sites. In the control samples we identify regions with F scores less than 2 that indicate lower RNase accessibility in the central footprint relative to the flanking regions. The control F scores are compared to the F scores identified for each footprint in the *Lrpprc* knockout samples and the significant log2 fold changes between the control and *Lrpprc* knockout mice are shown.

**Figure 2.** (a) Circular representation of the mitochondrial genome (centre track) displaying the PAR-CLIP binding site regions (interior tracks) and normalized coverage (log10 [reads per million]; scale: 1 – 100,000) across positions with at least a 95% posterior probability of being cross-linked (exterior tracks), as identified by BMix\textsuperscript{50}. (b) RNA EMSA of the LRPPRC/SLIRP complex, LRPPRC only or SLIRP only incubated with a RNA probe identified as a strong LRPPRC/SLIRP binding motif in vivo. (c) Schematic of shared binding sites among footprinting and PAR-CLIP datasets; the significance (P) of the overlap is indicated. (d) A sequence logo shows a predicted consensus binding motif for LRPPRC based on the binding sites that overlap between footprinting and PAR-CLIP datasets.

**Figure 3.** LRPPRC footprints show an increase in secondary structure propensity with its loss. (a) Circular representation of the mitochondrial genome (centre track) displaying footprints that are present in the *Lrpprc* knockout mice but not the controls (exterior tracks) coloured by the log2 fold change of the F score (scale: 0.0 – -2.5). The negative values
indicate that we have compared the footprints present when LRPPRC is lost to control samples. (b) A footprint located in mt-Nd5, showing the increased C score (scale: -4.3 – 3.7) and R score (scale: -0.5 – 1.0) across the footprint region in the knockout. (c) Violin plot showing the distribution of the log2 fold change in average R score across all LRPPRC footprints that overlap PAR-CLIP binding sites across the mitochondrial transcriptome, compared to the left and right 10 nt flanking regions. Loss of LRPPRC causes an increase in secondary structure propensity of mitochondrial RNAs. (d) LRPPRC demonstrates RNA chaperone activity in RNA annealing assays. Complementary oligoribonucleotides were hybridized in the presence or absence of the mouse LRPPRC/SLIRP complex.

Figure 4. LRPPRC/SLIRP complex binding is conserved in humans and mice. (a) Transcriptome-wide distribution of binding sites in HeLa and MEFs expressing LRPPRC-FLAG determined by PAR-CLIP. (b) The locations of binding sites of the LRPPRC/SLIRP complex within mt-Cyb are well conserved in human and mouse. (c) RNA EMSA indicating the preference of the LRPPRC/SLIRP complex for its target within each species.
Figure S1: Immunoblots showing that LRPPRC is lost in heart mitochondria from Lrpprc conditional heart and skeletal-knockout, LrpprcloxP/loxPcre+, mice compared to control LrpprcloxP/loxP mice (a) and SLIRP is lost in heart mitochondria from Slirp full body knockout mice, Slirp−/−, compared to control, Slirp+/+, mice (b). Porin was used as a loading control for each immunoblot.

Figure S2. Related to Figure 1 | Location of LRPPRC footprints in different classes of mt-RNAs.
Figure S3. Related to Figure 1 | SLRP does not bind mitochondrial RNAs in vivo.
Circular representation of the mitochondrial genome (centre track) displaying the putative footprint sites (grey bars) that showed differences in RNase protection between control and Slirp knockout mice, but these changes were not significant at a false discovery rate of ≤ 0.05.
### Table S1. Related to Figure 2 | PAR-CLIP library statistics for LRPPRC and control samples.

<table>
<thead>
<tr>
<th></th>
<th>LRPPRC (HeLa cells)</th>
<th>Control (HeLa cells)</th>
<th>LRPPRC (MEFs)</th>
<th>Control (MEFs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Reads</td>
<td>25,911,145</td>
<td>2,213,917</td>
<td>22,189,108</td>
<td>2,655,345</td>
</tr>
<tr>
<td>Mapped Reads</td>
<td>13,749,342</td>
<td>40,557</td>
<td>8,596,138</td>
<td>58,745</td>
</tr>
<tr>
<td>Unmapped Reads</td>
<td>12,161,803</td>
<td>2,173,360</td>
<td>13,592,970</td>
<td>2,596,600</td>
</tr>
<tr>
<td>Alignment Rate</td>
<td>53.06%</td>
<td>1.83%</td>
<td>38.74%</td>
<td>2.21%</td>
</tr>
</tbody>
</table>
Figure S4. Related to Figure 2 | Human SLIRP does not bind RNA in vitro. RNA EMSA of human SLIRP with hCYB nt 15598-15617 target RNA.

Figure S5. Related to Figure 2 | LRPPRC/SLIRP competitor assay. RNA EMSA of the mouse LRPPRC/SLIRP complex with mCo3 nt 9191-9210 target RNA (left panel) and mNd5 nt 12100-12125 non-target RNA (right panel). The RNA EMSAs were performed in the absence (a) or presence of 10x (b) and 100x (c) molar excess of unlabeled mNd5 12840-12865 non-target mRNA.
Table S2. Related to Figure 2c | High confidence binding sites of LRPPRC.

<table>
<thead>
<tr>
<th>Chrom</th>
<th>Start</th>
<th>End</th>
<th>Name</th>
<th>FDR</th>
<th>Strand</th>
</tr>
</thead>
<tbody>
<tr>
<td>chrM</td>
<td>906</td>
<td>943</td>
<td>footprints_37_0.129019_1.05681</td>
<td>0.0255</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1146</td>
<td>1187</td>
<td>footprints_11_0.009823_1.141524</td>
<td>0.0433</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1177</td>
<td>1202</td>
<td>footprints_25_0.047596_0.72678</td>
<td>0.0313</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1204</td>
<td>1214</td>
<td>footprints_10_0.015022_0.59324</td>
<td>0.0059</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1216</td>
<td>1240</td>
<td>footprints_24_0.003548_0.65438</td>
<td>0.0097</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1290</td>
<td>1306</td>
<td>footprints_16_0.046507_0.78884</td>
<td>0.0379</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1308</td>
<td>1316</td>
<td>footprints_8_0.107292_0.868081</td>
<td>0.0081</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1351</td>
<td>1360</td>
<td>footprints_9_0.153585_1.115471</td>
<td>0.0344</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1382</td>
<td>1389</td>
<td>footprints_17_0.047183_0.69359</td>
<td>0.0014</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1527</td>
<td>1538</td>
<td>footprints_9_-0.092514_0.646180</td>
<td>0.0028</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1598</td>
<td>1613</td>
<td>footprints_15_-0.066706_0.580301</td>
<td>0.0242</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>1786</td>
<td>1798</td>
<td>footprints_12_0.000000_0.63317</td>
<td>0.0227</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>2209</td>
<td>2219</td>
<td>footprints_10_-0.037586_0.621145</td>
<td>0.0359</td>
<td>-</td>
</tr>
<tr>
<td>chrM</td>
<td>2355</td>
<td>2363</td>
<td>footprints_12_-0.123076_1.243978</td>
<td>0.0311</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>2407</td>
<td>2485</td>
<td>footprints_28_0.042262_0.67909</td>
<td>0.0074</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>2499</td>
<td>2513</td>
<td>footprints_24_-0.272451_0.412981</td>
<td>0.003</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>2808</td>
<td>2826</td>
<td>footprints_18_0.020005_1.698620</td>
<td>0.0016</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>2910</td>
<td>2919</td>
<td>footprints_9_0.175478_1.283608</td>
<td>0.0356</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>3130</td>
<td>3148</td>
<td>footprints_18_0.184833_0.890067</td>
<td>0.043</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>3358</td>
<td>3370</td>
<td>footprints_12_0.073608_0.85233</td>
<td>0.0287</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>3381</td>
<td>3390</td>
<td>footprints_9_-0.459572_0.689690</td>
<td>0.0335</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>3465</td>
<td>3473</td>
<td>footprints_8_0.031186_0.69922</td>
<td>0.0208</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>3987</td>
<td>4004</td>
<td>footprints_17_0.336617_0.864999</td>
<td>0.0409</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>4057</td>
<td>4067</td>
<td>footprints_10_-0.058880_0.693090</td>
<td>0.0096</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>4593</td>
<td>4615</td>
<td>footprints_22_-0.022590_0.97061</td>
<td>0.0285</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>4682</td>
<td>4700</td>
<td>footprints_18_0.042832_0.75512</td>
<td>0.0312</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>5400</td>
<td>5410</td>
<td>footprints_10_-0.199968_0.679841</td>
<td>0.0267</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>5543</td>
<td>5566</td>
<td>footprints_23_0.159854_0.760140</td>
<td>0.0477</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>5623</td>
<td>5634</td>
<td>footprints_21_-0.064993_0.79747</td>
<td>0.0142</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>5646</td>
<td>5666</td>
<td>footprints_20_0.037364_0.47668</td>
<td>0.0067</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>5716</td>
<td>5724</td>
<td>footprints_6_-0.255098_0.659260</td>
<td>0.0117</td>
<td>+</td>
</tr>
<tr>
<td>chrM</td>
<td>start</td>
<td>stop</td>
<td>feature</td>
<td>start/stop</td>
<td>end/stop</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>chrM</td>
<td>5735</td>
<td>5747</td>
<td>footprints_12</td>
<td>0.122841/0.872596</td>
<td>0.0221</td>
</tr>
<tr>
<td>chrM</td>
<td>6602</td>
<td>6625</td>
<td>footprints_23</td>
<td>0.002798/0.858259</td>
<td>0.0358</td>
</tr>
<tr>
<td>chrM</td>
<td>6648</td>
<td>6656</td>
<td>footprints_8</td>
<td>0.034212/0.541813</td>
<td>0.002</td>
</tr>
<tr>
<td>chrM</td>
<td>6734</td>
<td>6742</td>
<td>footprints_8</td>
<td>0.000759/0.740846</td>
<td>0.0147</td>
</tr>
<tr>
<td>chrM</td>
<td>7181</td>
<td>7199</td>
<td>footprints_16</td>
<td>0.009354/0.642730</td>
<td>0.0261</td>
</tr>
<tr>
<td>chrM</td>
<td>7201</td>
<td>7213</td>
<td>footprints_12</td>
<td>0.011332/0.984710</td>
<td>0.0368</td>
</tr>
<tr>
<td>chrM</td>
<td>7323</td>
<td>7335</td>
<td>footprints_12</td>
<td>0.034212/0.541813</td>
<td>0.0369</td>
</tr>
<tr>
<td>chrM</td>
<td>7381</td>
<td>7391</td>
<td>footprints_10</td>
<td>0.006734/0.848416</td>
<td>0.003</td>
</tr>
<tr>
<td>chrM</td>
<td>7482</td>
<td>7491</td>
<td>footprints_13</td>
<td>0.024959/0.513339</td>
<td>0.038</td>
</tr>
<tr>
<td>chrM</td>
<td>7969</td>
<td>7977</td>
<td>footprints_19</td>
<td>0.009399/0.733068</td>
<td>0.0315</td>
</tr>
<tr>
<td>chrM</td>
<td>8113</td>
<td>8122</td>
<td>footprints_6</td>
<td>0.0266</td>
<td></td>
</tr>
<tr>
<td>chrM</td>
<td>8236</td>
<td>8276</td>
<td>footprints_40</td>
<td>0.0312</td>
<td></td>
</tr>
<tr>
<td>chrM</td>
<td>8296</td>
<td>8308</td>
<td>footprints_12</td>
<td>0.024959/0.513339</td>
<td>0.043</td>
</tr>
<tr>
<td>chrM</td>
<td>8384</td>
<td>8397</td>
<td>footprints_13</td>
<td>0.052203/0.518012</td>
<td>0.003</td>
</tr>
<tr>
<td>chrM</td>
<td>8438</td>
<td>8446</td>
<td>footprints_6</td>
<td>0.0026</td>
<td></td>
</tr>
<tr>
<td>chrM</td>
<td>8632</td>
<td>8642</td>
<td>footprints_10</td>
<td>0.006734/0.848416</td>
<td>0.0048</td>
</tr>
<tr>
<td>chrM</td>
<td>8763</td>
<td>8783</td>
<td>footprints_20</td>
<td>0.006734/0.848416</td>
<td>0.0059</td>
</tr>
<tr>
<td>chrM</td>
<td>8784</td>
<td>8792</td>
<td>footprints_8</td>
<td>0.054955/0.668197</td>
<td>0.0309</td>
</tr>
<tr>
<td>chrM</td>
<td>8934</td>
<td>8942</td>
<td>footprints_10</td>
<td>0.006734/0.848416</td>
<td>0.0048</td>
</tr>
<tr>
<td>chrM</td>
<td>9191</td>
<td>9210</td>
<td>footprints_19</td>
<td>0.0026</td>
<td></td>
</tr>
<tr>
<td>chrM</td>
<td>9246</td>
<td>9255</td>
<td>footprints_8</td>
<td>0.024959/0.513339</td>
<td>0.0407</td>
</tr>
<tr>
<td>chrM</td>
<td>9297</td>
<td>9308</td>
<td>footprints_11</td>
<td>0.024959/0.513339</td>
<td>0.0471</td>
</tr>
<tr>
<td>chrM</td>
<td>9505</td>
<td>9515</td>
<td>footprints_10</td>
<td>0.006734/0.848416</td>
<td>0.0074</td>
</tr>
<tr>
<td>chrM</td>
<td>9551</td>
<td>9561</td>
<td>footprints_19</td>
<td>0.0026</td>
<td></td>
</tr>
<tr>
<td>chrM</td>
<td>10217</td>
<td>10232</td>
<td>footprints_16</td>
<td>0.006734/0.848416</td>
<td>0.0233</td>
</tr>
<tr>
<td>chrM</td>
<td>10348</td>
<td>10365</td>
<td>footprints_17</td>
<td>0.0026</td>
<td></td>
</tr>
<tr>
<td>chrM</td>
<td>10578</td>
<td>10588</td>
<td>footprints_10</td>
<td>0.006734/0.848416</td>
<td>0.0314</td>
</tr>
<tr>
<td>chrM</td>
<td>10662</td>
<td>10691</td>
<td>footprints_29</td>
<td>0.0026</td>
<td></td>
</tr>
<tr>
<td>Chromosome</td>
<td>Start</td>
<td>End</td>
<td>Start of Footprints</td>
<td>End of Footprints</td>
<td>P-value</td>
</tr>
<tr>
<td>------------</td>
<td>-------</td>
<td>-----</td>
<td>---------------------</td>
<td>-------------------</td>
<td>---------</td>
</tr>
<tr>
<td>chrM</td>
<td>11147</td>
<td>11174</td>
<td>footprints_27</td>
<td>0.005570 0.949365</td>
<td>0.0083</td>
</tr>
<tr>
<td>chrM</td>
<td>11178</td>
<td>11218</td>
<td>footprints_40</td>
<td>0.090460 0.564475</td>
<td>0.0111</td>
</tr>
<tr>
<td>chrM</td>
<td>11406</td>
<td>11414</td>
<td>footprints_8</td>
<td>0.014280 0.988735</td>
<td>0.034</td>
</tr>
<tr>
<td>chrM</td>
<td>11699</td>
<td>11720</td>
<td>footprints_21</td>
<td>0.014532 1.099660</td>
<td>0.0402</td>
</tr>
<tr>
<td>chrM</td>
<td>11989</td>
<td>11996</td>
<td>footprints_8</td>
<td>0.153161 1.240775</td>
<td>0.0402</td>
</tr>
<tr>
<td>chrM</td>
<td>12567</td>
<td>12582</td>
<td>footprints_15</td>
<td>0.287680 0.852714</td>
<td>0.0368</td>
</tr>
<tr>
<td>chrM</td>
<td>12587</td>
<td>12608</td>
<td>footprints_11</td>
<td>0.040257 1.332191</td>
<td>0.0076</td>
</tr>
<tr>
<td>chrM</td>
<td>13385</td>
<td>13419</td>
<td>footprints_34</td>
<td>0.039015 0.819830</td>
<td>0.0262</td>
</tr>
<tr>
<td>chrM</td>
<td>13770</td>
<td>13799</td>
<td>footprints_29</td>
<td>0.074552 0.825320</td>
<td>0.0192</td>
</tr>
<tr>
<td>chrM</td>
<td>13824</td>
<td>13846</td>
<td>footprints_32</td>
<td>0.032980 0.754464</td>
<td>0.0142</td>
</tr>
<tr>
<td>chrM</td>
<td>14233</td>
<td>14242</td>
<td>footprints_9</td>
<td>0.442847 0.815809</td>
<td>0.0254</td>
</tr>
<tr>
<td>chrM</td>
<td>14333</td>
<td>14368</td>
<td>footprints_31</td>
<td>0.111271 0.826867</td>
<td>0.0229</td>
</tr>
<tr>
<td>chrM</td>
<td>14838</td>
<td>14872</td>
<td>footprints_34</td>
<td>0.014980 0.680558</td>
<td>0.0076</td>
</tr>
<tr>
<td>chrM</td>
<td>14995</td>
<td>15015</td>
<td>footprints_20</td>
<td>0.227673 0.634789</td>
<td>0.0195</td>
</tr>
<tr>
<td>chrM</td>
<td>15051</td>
<td>15065</td>
<td>footprints_14</td>
<td>0.050168 1.013145</td>
<td>0.0327</td>
</tr>
</tbody>
</table>
Figure S6. Related to Figure 2c. (a) Location of high confidence LRPPRC binding sites in different classes of mt-RNAs, determined by overlapping identified footprints and PAR-CLIP binding sites; and locations footprints that do not overlap with PAR-CLIP binding, normalized to transcript length (b).
LRPPRC/SLRP are necessary to relax secondary structures of mitochondrial RNAs to enable their maturation and polyadenylation.

(a) The distribution of average R scores across the left flanking, central footprint and right flanking regions (10 nt) of all footprints that overlap PAR-CLIP binding sites identified by BMix, in wild-type and knockout mice.
(b) The distribution of average R scores and changes in R score in mRNA-overlapping footprints. The average R score across the left flanking, central footprint and right flanking regions of all footprints that overlap mitochondrial mRNA coding sequences, in wild-type and knockout mice. The average log₂ fold change of the R score across the left flanking, central footprint and right flanking regions (10 nt) of all footprints that overlap mitochondrial mRNA-encoding regions.

(c) The distribution of average R scores and changes in R score in rRNA-overlapping footprints. The average R score across the left flanking, central footprint and right flanking regions (10 nt) of all footprints that overlap mitochondrial rRNA coding sequences, in wild-type and knockout mice. The average log₂ fold change of the R score across the left flanking, central footprint and right flanking regions of all footprints that overlap mitochondrial rRNA-encoding regions.
Figure S8. LRPPRC is required for polyadenylation of mitochondrial mRNAs. 
(a) Mapping of RNA-Seq reads to polyadenylated reference sequences reveals a reduction in polyadenylation in Lrpprc knockout hearts (blue line), compared to wild-type hearts (orange line). The canonical 3' region of each mRNA is shown in dark grey and the section of polyadenosine is shown in light grey.
(b) A modified 3′-RACE method confirms the reduced length of the 3′ poly(A) tails of mt-Nd1, mt-Nd2 and mt-Cyb mRNAs.
Figure S9. Loss of LRPPRC chaperone activity leads to significantly reduced stability of mitochondrial mRNAs, identified by DE-Seq analyses of RNA-Seq datasets from four control and four Lrpprc knockout mice.
Supplementary Table 3. WT called LRPPRC footprints with statistically significant increases in F score (indicating a reduction in footprint depth in the KO).

Supplementary Table 4. WT called SLIRP footprints with no statistically significant increases in F score.

Supplementary Table 5. KO called LRPPRC footprints with statistically significant increase in F score (indicating an increase in footprint depth in the KO).

Supplementary Table 6. Average R scores across rRNA- and mRNA-encoding regions in LRPPRC and SLIRP WT and KO samples.

Supplementary Table 7. Control footprint analyses of the cytoplasmic 28S rRNA in LRPPRC WT and KO samples.
Appendix A
Mitochondria are vital in providing cellular energy via their oxidative phosphorylation system, which requires the coordinated expression of genes encoded by both the nuclear and mitochondrial genomes (mtDNA). Transcription of the circular mammalian mtDNA depends on a single mitochondrial RNA polymerase (POLRMT). Although the transcription initiation process is well understood, it is debated whether POLRMT also serves as the primer for the initiation of mtDNA replication. In the nucleus, the RNA polymerases needed for gene expression have no such role. Conditional knockout of Polrmt in the heart results in severe mitochondrial dysfunction causing dilated cardiomyopathy in young mice. We further studied the molecular consequences of different expression levels of POLRMT and found that POLRMT is essential for primer synthesis to initiate mtDNA replication in vivo. Furthermore, transcription initiation for primer formation has priority over gene expression. Surprisingly, mitochondrial transcription factor A (TFAM) exists in an mtDNA-free pool in the Polrmt knockout mice. TFAM levels remain unchanged despite strong mtDNA depletion, and TFAM is thus protected from degradation of the AAA+ Lon protease in the absence of POLRMT. Last, we report that mitochondrial transcription elongation factor may compensate for a partial depletion of Polrmt in heterozygous Polrmt knockout mice, indicating a direct regulatory role of this factor in transcription. In conclusion, we present in vivo evidence that POLRMT has a key regulatory role in the replication of mammalian mtDNA and is part of a transcriptional mechanism that provides a switch between primer formation for mtDNA replication and mitochondrial gene expression.

INTRODUCTION

Mitochondria are essential for a variety of metabolic processes, including oxidative phosphorylation (OXPHOS), whereby energy is harvested from food nutrients to synthesize adenosine 5'-triphosphate (ATP). Mitochondrial dysfunction is associated with a number of genetic diseases and is heavily implicated in age-associated diseases, as well as in the aging process itself. Mitochondria are vital in providing cellular energy via their oxidative phosphorylation system, which requires the coordination of mitochondrial and nuclear gene expression. The initiation of mtDNA replication. In the nucleus, the RNA polymerases needed for gene expression have no such role. Conditional knockout of Polrmt in the heart results in severe mitochondrial dysfunction causing dilated cardiomyopathy in young mice. We further studied the molecular consequences of different expression levels of POLRMT and found that POLRMT is essential for primer synthesis to initiate mtDNA replication in vivo. Furthermore, transcription initiation for primer formation has priority over gene expression. Surprisingly, mitochondrial transcription factor A (TFAM) exists in an mtDNA-free pool in the Polrmt knockout mice. TFAM levels remain unchanged despite strong mtDNA depletion, and TFAM is thus protected from degradation of the AAA+ Lon protease in the absence of POLRMT. Last, we report that mitochondrial transcription elongation factor may compensate for a partial depletion of Polrmt in heterozygous Polrmt knockout mice, indicating a direct regulatory role of this factor in transcription. In conclusion, we present in vivo evidence that POLRMT has a key regulatory role in the replication of mammalian mtDNA and is part of a transcriptional mechanism that provides a switch between primer formation for mtDNA replication and mitochondrial gene expression.
RESEARCH ARTICLE

the elongation step, TEFM2 dissociates from the complex (11, 13). The formation of the pre-IC has been suggested to serve as an important point of regulation of transcription in mitochondria (19). Several in vitro studies suggest a difference in activity between HSP and LSP, but this question has never been addressed in vivo. In addition to its role in transcription initiation, TFAM is also a DNA packaging factor that coats the entire mtDNA in a sequence-independent manner and tangles it into compacted nucleoid structures (20, 21). TFAM is essential for mtDNA maintenance and is a key factor involved in directly regulating mtDNA copy number in mammals (22). Small variations in the TFAM-to-mtDNA ratio can regulate the number of genomes available for mitochondrial gene transcription and mtDNA replication (23). TFAM levels and turnover are reported to be regulated by the AAA+ Lon protease (LONP), which degrades TFAM but not bound to mtDNA (24, 25).

In vitro studies have suggested that POLRMT plays an essential role in mtDNA replication and acts together with critical nuclear-encoded factors of the mitochondrial replication machinery, that is, mitochondrial DNA polymerase γ (POLγ) (26, 27), the mitochondrial DNA helicase TWINKLE (28), and mitochondrial single-stranded DNA binding protein 1 (SSB1). It is hypothesized that transcription is prematurely terminated in close proximity to the O2I, which is located about 100 base pairs (bp) downstream of LSP. The RNA primer generated in this way is used by the mitochondrial replication machinery to initiate mtDNA replication at O2I (29–31). However, in the in vivo importance of POLRMT in replication primer formation is debated, because the mammalian mitochondrial primase/polymerase (Pim1/Pim2) is reported to play a role in mtDNA replication (32). Three models have been proposed to explain how mtDNA is primed and replicated (34, 36, 37). According to the strand displacement model, replication of the H strand is initiated at O2I and continues to displace the parental H strand. The replication machinery reaches the L-strand origin of replication (OL) after synthesis of two-thirds of the H strand (38). When this region of the L strand becomes single-stranded, it forms a stem-loop structure to which POLRMT can bind to produce the RNA primer necessary for initiation of mtDNA replication at O3I (39, 40). Most mammalian mtDNA replication events initiated at the leading strand are abortive and lead to the formation of a triple-stranded structure containing a prematurely terminated nascent H-strand DNA fragment of 680 nucleotides, which is called the displacement (D) loop (41). The abortive mtDNA replication product is denoted 7S DNA and remains bound to the parental L strand, thereby displacing the H strand. Mitochondrial transcription elongation factor (TEFM) was found to increase the processivity of POLRMT, thus facilitating the synthesis of long polycistronic RNA (42–44). Interaction between TEFM and POLRMT was reported to prevent the generation of replication primers and was thus suggested to serve as a molecular switch between replication and transcription (44). Further in vivo analysis is required to investigate how such a switch could be regulated in the mitochondria of actively metabolizing cells. The in vivo regulation of coordinated expression and replication of mtDNA still remains one of the most significant gaps in our understanding of mitochondrial function. POLRMT was suggested as a key factor in both of these processes. The role of POLRMT as a primase for mtDNA replication is debated, because nuclear DNA replication is dependent on specific primases distinct from the RNA polymerases needed for transcription of nuclear genes. To address this question, we generated a Polrmt knockout mouse model to elucidate the in vivo role of POLRMT in mtDNA replication.

RESULTS

Polrmt is essential for embryo development

We generated a conditional knockout allele for Polrmt by targeting exon 3 to disrupt the expression of POLRMT (fig. S1). Germline heterozygous knockout (Polrmt+/−) animals were obtained by breeding mice with a heterozygous floxed Polrmt allele (Polrmtfl/+)) to mice with ubiquitous expression of cre recombinase (β-actin-cre). Subsequent intercrosses of Polrmt−/− mice produced no viable heterozygous knockouts (genotyped offspring, n = 56; Polrmt−/−, 66%; Polrmt+/−, 34%; Polrmt++/−, 0%). We therefore proceeded to analyze staged embryos at embryonic day (E) 8.5 (n = 47) and found that 23% of embryos with a mutant appearance all had the genotype Polrmt−/−, whereas the normally appearing embryos were either Polrmt++/− (28%) or Polrmt++/− (49%). Loss of POLRMT thus leads to a severely mutant appearance and embryonic lethality at E8.5, similar to the phenotype of other mouse knockouts with germline disruption of genes essential for maintenance or expression of mtDNA (28, 45).

Disruption of Polrmt in heart causes dilated cardiomyopathy

To gain further in vivo insights into the role of Polrmt, we studied conditional knockout mice (genotype, Polrmtfl+/−, +/Cre−/−) with disruption of Polrmt in heart and skeletal muscle (46). Analysis of complementary DNA (cDNA) by reverse transcription polymerase chain reaction (RT-PCR) verified that sequences corresponding to exon 3 of the Polrmt mRNA were lacking in the knockout tissues (Fig. 1A). The conditional knockout mice had a drastically shortened life span, and all of them died before 6 weeks of age (Fig. 1B). We found a progressive enlargement of the heart with dilution of the left ventricular chamber without any apparent increase in the left ventricular wall thickness (Fig. 1C). The ratio of heart to body weight progressively increased during the first weeks of postnatal life (Fig. 1D). Electrocardiography (ECG) showed 40% reduction of the heart rate variability, defined as the variation between successive heart beats in late-stage conditional knockout mice, whereas the average heart rate was not different from controls (fig. S2).

Loss of POLRMT causes severe mitochondrial dysfunction in heart

Transmission electron microscopy of terminal-stage knockout hearts showed disruption of the tissue organization and abnormally appearing mitochondria with disorganized cristae (Fig. 2A) consistent with mitochondrial dysfunction. Analysis of OXPHOS capacity in Polrmt knock-out hearts showed that the enzyme activities of complex I, IV, and V were deficient, whereas the activity of the exclusively nuclear-encoded complex II was normal (Fig. 2B). Consistently, analysis of the levels of assembled OXPHOS complexes by blue native polyacrylamide gel electrophoresis (BN-PAGE) showed reduced levels of complexes I, IV, and V, whereas complex II was unchanged in the Polrmt conditional knockout hearts (Fig. 2C). In addition to a reduction of assembled complex V, increased levels of a subcomplex containing the Fo portion of ATP synthase was found (Fig. 2C). The observed pattern of deficient OXPHOS sparring complex II (Fig. 2B) is typically caused by reduced mtDNA expression. Therefore, we proceeded to analyze mitochondrial transcript levels.
POLRMT synthesizes primers for mtDNA replication initiation

We did not detect the 7S RNA transcript in Polrmt knockout hearts (Fig. 3A). The 7S RNA is the most promoter-proximal transcript formed by transcription initiation at the LSP of mtDNA (47). In vitro studies have suggested that transcription from LSP may also form primers for initiation of mtDNA replication. To further elucidate this possibility, we studied levels of D-loop strands referred to as 7S DNA (Fig. 3B). The levels of 7S DNA were profoundly reduced in relation to total mtDNA levels in Polrmt knockout hearts. The steady-state levels of mtDNA were also reduced in hearts lacking POLRMT expression (Fig. 3C and D). Finally, we determined whether the reduced levels of 7S DNA and mtDNA molecules were due to decreased formation or increased degradation by performing in organello mtDNA replication experiments (Fig. 3E and F). The de novo formation of 7S DNA, as well as mtDNA, was much reduced in the absence of POLRMT, consistent with an essential role of this protein in initiation of leading-strand mtDNA replication.

Loss of POLRMT leads to up-regulation of TWINKLE protein levels

To further characterize molecular changes associated with the mtDNA depletion phenotype caused by the absence of POLRMT, we studied the expression of known nuclear-encoded transcription and replication factors, as well as their intramitochondrial interactions. We followed the levels of TFR2M during the progressive depletion of POLRMT in heart and found a parallel decrease in both factors (Fig. 4A and figs. S3, A and B, and S4A), consistent with a previously reported direct molecular interaction between TFR2M and POLRMT (12, 19, 48). It is important to note that the decrease in TFR2M protein levels occurred despite normal expression of the Tfr2m mRNA (Fig. 4B), showing that TFR2M is dependent on POLRMT for its stability. In contrast, TEFM is stable or even slightly increased in the absence of POLRMT (fig. S4, B and C). We further assessed the levels of some important factors involved in mitochondrial RNA metabolism and translation. Both the zinc phosphodiesterase ELAC protein 2 (ELAC2) and the G-rich sequence factor 1 (GRSF1) were strongly increased in Polrmt knockout hearts (fig. S4,

RESEARCH ARTICLE

Fig. 1. Knockout of Polrmt in germline and heart. (A) RT-PCR analysis of Polrmt transcripts from control (L/L) and tissue-specific knockout mice (L/L, cre). Different primer sets were used as indicated; exon 3, 551 bp. UTR, untranslated region. (B) Survival curve of control (n = 40) and tissue-specific knockout (n = 47) mice. (C) Cardiac phenotype. Vertical (upper panels) and transverse (lower panels) sections through the midportion of hearts of control and tissue-specific knockout mouse hearts at 4 weeks of age. Scale bars, 2 mm. (D) Heart-to-body weight ratio of control (n = 42) and tissue-specific knockout mice (n = 37) at different time points. Error bars indicate ±SEM (**P < 0.001; two-tailed Student’s t test).
B and C), ELAC2 is important for RNA processing at the 3′ end of tRNAs (49, 50), and GRSF1 interacts specifically with ribonuclease (RNase) P and is suggested to be involved in early RNA processing events (51, 52).

The levels of the leucine-rich pentatricopeptide repeat-containing (LRPPRC) protein (53, 54) were much reduced (Fig. 4A and figs. S3, A and C, and S4A) despite normal levels of Lrpprc transcripts (Fig. 4B).

Next, we found that the levels of mitochondrial ribosomal proteins of the small and large ribosomal subunits (MRPS35 and MRPL37) were severely reduced in the absence of POLRMT (Fig. 4C and fig. S4, B and C). Noteworthy, the steady-state levels of the 39S ribosomal protein L12 (MRPL12), suggested to play a role in regulation of mitochondrial transcription (55), were not changed in Polrmt knockout hearts (fig. S4, B and C). We found significantly increased protein levels of TWINKLE (Fig. 4A and figs. S3, A and C, and S4A) despite normal levels of Twinkle transcripts in Polrmt knockout hearts (Fig. 4B), showing that this replicative helicase is stabilized when mtDNA replication is compromised.

An mtDNA-free pool of TFAM appears in absence of POLRMT

Despite severely reduced mtDNA levels, the protein levels of TFAM were unchanged in the absence of POLEMT (Fig. 4A and figs. S3, A and B, and S4A). Furthermore, the steady-state levels of Tfam transcripts were increased by ~50% in the knockout hearts, whereas the mtDNA expression of the other nucleus-encoded factors was not significantly altered (7b,25, Tfam, PolA, Bulb, Mtraf, Mterf4, and Lrpprc) (Fig. 4B). The AAA+ LONP, which has been reported to degrade TFAM if not bound to mtDNA (25), showed an increased expression (Fig. 4A and figs. S3A and S4A). The finding that TFAM protein levels do not follow the decreased mtDNA levels is very surprising because there are numerous reports of costabilization between TFAM and mtDNA (56). We therefore performed linear density glycerol gradient experiments to determine the pools of TFAM that were unbound or present together with mtDNA in the nucleoid (21, 57). The fractions of the gradient containing the nucleoid was determined by the presence
of mtDNA, verified by Southern blotting and deoxyribonuclease (DNase) treatment, and by the presence of replication factors, such as TWINKLE and POL γA (Fig. 4C and fig. S5A). In mitochondria isolated from Polrmt knockout hearts, there was a clear increase in TFAM protein levels in the mtDNA-free fractions (Fig. 4, D and E, and fig. S6A). The mtDNA depletion that results from deficient replication primer formation in the absence of POLRMT thus leads to an increased pool of mtDNA-free TFAM that is not degraded despite increased LONP protein levels.

LSP-initiated transcription is favored at low POLRMT levels Consistent with the previously observed profound reduction in de novo transcription of mtDNA in the absence of POLRMT (48), we found a global severe decrease in the steady-state levels of all analyzed mitochondrial mRNAs, rRNAs, and tRNAs (Fig. 5, A to C). Northern blot analysis showed that, in the absence of POLRMT, the levels of several transcripts encoded on the L-strand, that is, mt-ND6, mt-Tp, mt-Te, mt-Ts2, mt-Tc, mt-Tn, and mt-Tq, were less decreased than transcripts encoded on the H-strand (Fig. 5, A and B). Next, we used RNA sequencing (RNA-Seq) on isolated mitochondria from FO为民 knockout hearts at different ages by Northern blot on total RNA; loading, 18S rRNA. RNAs from hearts of MrPer4 conditional knockout mice (22) with increased 7S RNA levels were loaded as controls. (B) Southern blot analyses on mtDNA to assess 7S DNA levels of 4-week-old control and tissue-specific knockout mice. To allow relative comparison, the loaded amount of mtDNA from knockouts was higher than the amount loaded from control samples. (C) Southern blot analysis on total DNA to assess mtDNA levels of control and tissue-specific knockout mice at different ages; loading, 18S rDNA. (D) Quantification of Southern blots: mtDNA levels were normalized to 18S rDNA and presented as the percentage of controls. Error bars indicate ±SEM (*P < 0.05; two-tailed Student's t test). (E) Levels of de novo-synthesized DNA of isolated heart mitochondria of 4-week-old tissue-specific knockout and control mice. Equal input was ensured by Western blot analysis (voltage-dependent anion channel (VDAC)) on isolated mitochondria after labeling before mtDNA extraction. (F) Quantification of the results from (E).

Heterozygous Polrmt knockouts increase TEFM levels and maintain transcription To determine how a moderate reduction of POLRMT levels affects mtDNA expression, we studied Polrmt+/− mice. The Polrmt+/− mice showed decreased POLRMT protein levels in heart, skeletal muscle, and liver in accordance with a 50% reduction in Polrmt gene dosage (Fig. 6A and fig. S7, A and B). A global reduction of POLRMT expression does not cause any phenotype, and the heterozygous Knockout
Fig. 4. Loss of POLRMT results in an increased mtDNA-free pool of TFAM. (A) Steady-state protein levels of nuclear-encoded factors of mtDNA expression analyzed by Western blotting on mitochondrial extracts from hearts of control and tissue-specific knockout mice; loading, VDAC; asterisk, cross-reacting band (28); for quantification, see fig. S4A. (B) Quantitative RT-PCR (qRT-PCR) of transcript levels of nuclear-encoded mitochondrial proteins. Normalization, β-2M (β-2-microglobulin). Error bars indicate ±SEM (*P < 0.05 and ***P < 0.001; two-tailed Student’s t test; see table S1). (C and D) Linear glycerol density gradient fractionations of mitochondrial lysates from tissue-specific knockout and control mice followed by Western blot analysis; for quantification, see figs. S5 and S6. Samples taken from fractions 1 to 16 are of increasing density (that is, from top to bottom of the tube after separation by ultracentrifugation), as indicated by the schematic representation of the centrifuge tube to the left. Fractions were loaded from left to right on the gel, as indicated by the lane numbering input, aliquots of unfractionated lysates. The mtDNA content of the fractions was determined by Southern blotting. (E) Relative TFAM and POLRMT protein distribution across the gradient from control and knockout heart mitochondria.
Fig. 5. LSP and HSP show different sensitivities at low POLRMT concentrations. (A and B) Northern blot analyses of mitochondrial mRNAs, tRNAs, and rRNAs from hearts of 4-week-old control and tissue-specific knockout mice; loading, 18S rRNA. (C) Relative mitochondrial RNA abundance of mRNA and rRNA levels in hearts of 4-week-old tissue-specific knockout and control mice normalized to the upper quartile of the gene count distribution. The data analyzed are from three independent RNA-Seq experiments; all RNAs have \( P \leq 0.0001 \). Error bars indicate ±SEM. (D) In vitro transcription assay at different POLRMT levels. POLRMT was added at 128, 32, 8, 2, and 0.5 nM in lanes 1 to 5, respectively; lane 6, control without POLRMT; lane 7, molecular weight marker (New England Biolabs). (E) Quantification of the results from (D). The experiment was performed in triplicates, and HSP transcription levels were normalized to LSP for each POLRMT concentration; bars, mean values. Error bars indicate ±SD (n = 3).
mice are apparently viable, fertile, and healthy at 1 year of age. We proceeded to determine whether the moderately altered POLRMT levels have an effect on the steady-state levels of OXPHOS enzyme complexes. Western blot analysis of respiratory chain subunits showed normal levels of NDUFB8 (complex I), SDHA (complex II), UQRC2 (complex III), COX1 and COX2 (complex IV), and ATP5A1 (complex V) in hearts from 26- and 52-week-old Polrmt+/- mice (Fig. 6A and fig. S7, C and D). Both Northern blot analysis of steady-state levels of mitochondrial rRNAs, tRNAs, and mRNAs, and RNA-Seq analysis of the mitochondrial transcriptome of Polrmt+/- mice in heart showed a normal mitochondrial transcript abundance despite reduced POLRMT levels (Fig. 6B and fig. S8, A and B). Consistent with the normal mt-mRNA levels, the levels of LRPPRC were not changed in Polrmt+/- mice (Fig. 6C and fig. S9A). Noteworthy, RNA-Seq data showed that the mt-Nd6 transcript encoded on the L strand was slightly reduced, whereas all other mitochondrial rRNAs and mRNAs encoded on the H strand showed a tendency to be increased in Polrmt+/- hearts at 26 weeks of age (fig. S8B). We observed a strong increase in tTEFm protein levels (Fig. 6C and fig. S9A), which may provide a compensatory response to reduced transcription initiation in Polrmt+/- mice. This suggestion is supported by the finding of slightly increased de novo transcription in heart mitochondria from Polrmt+/- mice (Fig. 6D). Furthermore, the levels of 7S RNA are normal in Polrmt+/- mice (Fig. 6E), thus showing that promoter proximal transcription at LSP is sufficient.

Fig. 6. Characterization of heterozygous Polrmt knockout mice. (A) POLRMT steady-state protein levels in heart from wild-type (+/+) and heterozygous Polrmt knockout (+/-) mice; loading, VDAC; for quantification, see fig. S7C. (B) Steady-state levels of mitochondrial mRNAs, rRNAs, and tRNAs; loading, 18S rRNA; for quantification, see fig. S8A. (C) Steady-state protein levels of nuclear-encoded factors of mtDNA expression analyzed by Western blotting on mitochondrial heart extracts; loading, VDAC; for quantification, see fig. S9A; asterisk, cross-reacting band (28S). (D) De novo-synthesized mitochondrial transcripts from hearts of 52-week-old mice. Steady-state levels of individual mitochondrial transcripts were verified with a radiolabeled probe (mt-Co1), input: Western blot analysis (POLRMT and VDAC) after labeling. (E) 7S RNA levels in mouse hearts by Northern blotting on total RNA; loading, 18S rRNA. (F) Quantification of mtDNA by quantitative PCR (qPCR) with mt-Co1, mt-Nd1, and mt-Nd5 probes on mouse heart. Signals were normalized to the 18S signal; n = 3. Error bars indicate ±SEM. (G) De novo-synthesized DNA of isolated mitochondria from hearts of 12-week-old mice. The mtDNA was radiolabeled, isolated and boiled to release newly synthesized 7S DNA before Southern blotting; input: Western blotting (POLRMT and VDAC) after labeling; for quantification, see fig. S10B.
to maintain mtDNA replication and levels (Fig. 6F and fig. S10A). Consistent with the normal mtDNA levels, the levels of TFAM were not changed in Polrmt−/− mice (Fig. 6C and fig. S8A). Also, de novo formation of 7S DNA was not decreased when POLRMT was reduced (Fig. 6G and fig. S9B), and the normal levels of TWINKLE, POLA, and SSBP1 proteins (Fig. 6C and fig. S9, A and B) provide further support for normal mtDNA replication in Polrmt−/− mice. Thus, a moderate reduction of POLRMT expression in Polrmt−/− mice does not affect overall mtDNA transcription or maintenance.

**DISCUSSION**

Here, we show that POLRMT, in addition to its essential role in gene expression, is required for primer formation for initiation of mtDNA replication in vivo. Our data show that primer synthesis and initiation of mtDNA replication are prioritized over gene transcription because (i) there is a differential effect of POLRMT at the mtDNA promoters as illustrated in Fig. 7. At lower POLRMT levels, transcription is primarily initiated at LSP, which helps to ensure that primer synthesis can be maintained. (ii) Under normal conditions, about 95% of all replication initiation events at O2 are prematurely terminated, forming the 7S DNA. When POLRMT is depleted, 7S DNA is no longer formed, suggesting that all residual replicative events continue to fail-length replication. (iii) The increase in TWINKLE protein levels is likely a compensatory response aimed to promote productive mtDNA replication. The increased TWINKLE protein levels are most likely due to posttranscriptional regulation because the Twinkle mRNA levels remain unchanged. Although the exact mechanisms of TWINKLE stabilization and regulatory role in mtDNA replication remain unknown, TWINKLE can be loaded at the 3′ end of 7S DNA to promote full-length genomic replication (28). It is possible that abortive mtDNA replication is favored by an antihelicase activity at the end of the D loop (38, 39) and that increased TWINKLE levels will overcome this block. The increase in TWINKLE as a compensatory mechanism agrees with results from other mouse models with severe mtDNA depletion, suggesting an involvement of TWINKLE in the regulation of mtDNA replication in mammals. Surprisingly, the TFAM protein levels remain normal in the absence of POLRMT despite profound mtDNA depletion. The inability of this excess TFAM to bind mtDNA to form nucleoids results in an increased free pool of TFAM. There is strong experimental evidence that the amount of TFAM directly regulates mtDNA copy number and that mtDNA levels also microscopically affect TFAM replication. (iii) The increase in TWINKLE protein levels is likely a compensatory response to ensure a normal transcription rate and mitochondrial steady-state transcript levels even when the ratio of POLRMT protein per mtDNA molecule is reduced in the heterozygous Polrmt−/− mice. Thus, a compensatory mechanism that controls TFAM levels in mitochondria. To our knowledge, this discrepancy between mtDNA and TFAM levels has not been found in other knockout mouse models with reduced mtDNA levels (28), and it remains to be clarified whether this is specifically linked to the loss of POLRMT. LONP is known to play a role in degrading unfolded, unassembled, or oxidatively damaged proteins. Thus, the highly induced levels of LONP in the Polrmt knockout mice are likely due to an imbalance between mtDNA- and nucleus-encoded CYFIP1OS subunits, as already described in previous studies (61). Indeed, mechanisms that regulate transcript levels in mammalian mitochondria remain to be clarified, and our study of the heterozygous Polrmt knockout mice provides important insights. Despite a drop in POLRMT levels, the steady-state mitochondrial transcript levels in these mice are unchanged. We also observed increased TEFM protein levels, whereas all other nuclear-encoded factors involved in mtDNA expression, for example, LRPPRC, TFAM, and TWINKLE, were normal. This could be a compensatory response to ensure a normal transcription rate and mitochondrial steady-state transcript levels even when the ratio of POLRMT protein per mtDNA molecule is reduced in the heterozygous Polrmt knockout mice. In vitro studies recently reported that TEFM increases transcription processivity to allow mitochondria to increase transcription rates (43, 44). In contrast to TEFM, neither TFAM nor TFB2M levels were affected in the heterozygous Polrmt knockout mice. Both of these factors are directly involved in the transcription initiation mechanism, but apparently, their levels are not limiting in promoting increased levels of transcription. In the nucleus, differential activation of the large number of protein-coding genes is not controlled by varying combinations of bound transcription factors that regulate promoter specificity. In contrast to the nucleus, where protein-coding genes typically are present only in two copies per cell, there are thousands of copies of mtDNA per cell. This means that increased transcription of mtDNA does not necessarily require increased transcription from each promoter, but it can instead be achieved by engaging more mtDNA templates in
transcription. This mode of regulation will only be an efficient regulatory system if the number of target promoters is low, such as is the case for mtDNA that only has two promoters per genome. How the presence of TFEM may prevent early transcription termination and promote transcription beyond the immediate promoter region remains to be discovered. The genetic data we present here resolve the controversy concerning the role of POLRMT as a primase for mammalian mtDNA replication. This issue has been widely debated because nuclear DNA replication is dependent on specific primasins, which are distinct from the RNA polymerases needed for transcription of nuclear genes. Our mouse knockout data show that in the absence of POLRMT, mtDNA is no longer formed, and there is a severe reduction of mtDNA levels. In vitro data have shown that POLRMT can generate the RNA primers needed for mtDNA synthesis at both O and O′ (40), but in vivo importance of these findings has not been established (62). The 7S RNA, a polyadenylated transcript that is terminated near the 5′ end of the nascent D loop strand, is undetectable in the absence of POLRMT. The function of the 7S RNA is not clear, but it has been suggested that it is involved in primor formation for initiation of mtDNA synthesis at O (36, 58). In mammalian mitochondria, PrimPol was reported to have DNA and RNA primase activities and play a role in mtDNA replication (33). Our knockout data show that PrimPol or any other primase cannot complement for the absence of POLRMT when it comes to mtDNA replication initiation. Together, our study provides clear evidence that POLRMT functions as the primase for mtDNA replication in mammalian mitochondria in vivo. Deletion of Polrmt in heart results in a severe decrease in all mitochondrial transcripts, but we also report that the mitochondrial transcripts derived from the L strand, in particular the mt-Nd6 mRNA, are less reduced than those transcripts encoded on the H strand. This could be explained by different stabilities of the individual mitochondrial transcripts due to up-regulation/down-regulation of factors involved in mitochondrial transcript processing/stability. Previous studies reported that most L-strand transcripts, including mt-Nd6, were unaffected when silencing or knocking out the Lepr gene (54, 63). LRPPRC levels decrease during the absence of POLRMT. Levels of LRPPRC and mitochondrial mRNAs (64). The strongly increased GRSF1 levels in Polrmt knockouts are in line with the less reduced amounts of mt-Nd6 mRNA in comparison with levels of H-strand transcripts, because GRSF1 has been suggested to be involved in regulating and interacting with mt-Nd6 mRNA and its precursor transcript. For Northern blot analysis, total RNA was isolated from heart tissue using the DNeasy Blood & Tissue Kit (Qiagen) and subjected to Southern blot analysis as previously described (28). DNA (3 to 10 μg) was digested with Sac I and, after treatment (TURBO DNA-free; Ambion), RT-PCR was carried out on 4 ng of total DNA in a 7900HT Real-Time PCR System (Applied Biosystems) using TaqMan probes specific for the mt-Co1, mt-Nd1, mt-Nd6, and 18S rDNA (Applied Biosystems). Total DNA was isolated from heart tissue using the DNeasy Blood & Tissue Kit (Qiagen) and subjected to Southern blot analysis as previously described (28). DNA (3 to 10 μg) was digested with Sac I and, after treatment (TURBO DNA-free; Ambion), RT-PCR was carried out on 4 ng of total DNA in a 7900HT Real-Time PCR System (Applied Biosystems). In conclusion, we found that POLRMT is essential for embryonic development, and that whole-body knockout results in embryonic lethality due to early gestation, whereas a moderate reduction of POLRMT functions as the primase for mtDNA replication in mammalian mitochondria. At low POLRMT levels, we observed a significant discrepancy in the initiation of mitochondrial transcription between the HSP and LSP, suggesting that POLRMT is part of a mechanism that provides a switch between RNA primer formation for mtDNA replication and mtDNA transcription. This mode of regulation will only be an efficient regulatory system if the number of target promoters is low, such as is the case for mtDNA that only has two promoters per genome. How the presence of TFEM may prevent early transcription termination and promote transcription beyond the immediate promoter region remains to be discovered. The genetic data we present here resolve the controversy concerning the role of POLRMT as a primase for mammalian mtDNA replication. This issue has been widely debated because nuclear DNA replication is dependent on specific primasins, which are distinct from the RNA polymerases needed for transcription of nuclear genes. Our mouse knockout data show that in the absence of POLRMT, mtDNA is no longer formed, and there is a severe reduction of mtDNA levels. In vitro data have shown that POLRMT can generate the RNA primers needed for mtDNA synthesis at both O and O′ (40), but in vivo importance of these findings has not been established (62). The 7S RNA, a polyadenylated transcript that is terminated near the 5′ end of the nascent D loop strand, is undetectable in the absence of POLRMT. The function of the 7S RNA is not clear, but it has been suggested that it is involved in primor formation for initiation of mtDNA synthesis at O (36, 58). In mammalian mitochondria, PrimPol was reported to have DNA and RNA primase activities and play a role in mtDNA replication (33). Our knockout data show that PrimPol or any other primase cannot complement for the absence of POLRMT when it comes to mtDNA replication initiation. Together, our study provides clear evidence that POLRMT functions as the primase for mtDNA replication in mammalian mitochondria in vivo. Deletion of Polrmt in heart results in a severe decrease in all mitochondrial transcripts, but we also report that the mitochondrial transcripts derived from the L strand, in particular the mt-Nd6 mRNA, are less reduced than those transcripts encoded on the H strand. This could be explained by different stabilities of the individual mitochondrial transcripts due to up-regulation/down-regulation of factors involved in mitochondrial transcript processing/stability. Previous studies reported that most L-strand transcripts, including mt-Nd6, were unaffected when silencing or knocking out the Lepr gene (54, 63). LRPPRC levels decrease during the absence of POLRMT. Levels of LRPPRC and mitochondrial mRNAs (64). The strongly increased GRSF1 levels in Polrmt knockouts are in line with the less reduced amounts of mt-Nd6 mRNA in comparison with levels of H-strand transcripts, because GRSF1 has been suggested to be involved in regulating and interacting with mt-Nd6 mRNA and its precursor transcript. For Northern blot analysis, total RNA was isolated from heart tissue using the DNeasy Blood & Tissue Kit (Qiagen) and subjected to Southern blot analysis as previously described (28). DNA (3 to 10 μg) was digested with Sac I and, after treatment (TURBO DNA-free; Ambion), RT-PCR was carried out on 4 ng of total DNA in a 7900HT Real-Time PCR System (Applied Biosystems) using TaqMan probes specific for the mt-Co1, mt-Nd1, mt-Nd6, and 18S rDNA (Applied Biosystems). Total DNA was isolated from heart tissue using the DNeasy Blood & Tissue Kit (Qiagen) and subjected to Southern blot analysis as previously described (28). DNA (3 to 10 μg) was digested with Sac I and, after treatment (TURBO DNA-free; Ambion), RT-PCR was carried out on 4 ng of total DNA in a 7900HT Real-Time PCR System (Applied Biosystems) using TaqMan probes specific for the mt-Co1, mt-Nd1, mt-Nd6, and 18S rDNA (Applied Biosystems). Total DNA was isolated from heart tissue using the DNeasy Blood & Tissue Kit (Qiagen) and subjected to Southern blot analysis as previously described (28). DNA (3 to 10 μg) was digested with Sac I and, after treatment (TURBO DNA-free; Ambion), RT-PCR was carried out on 4 ng of total DNA in a 7900HT Real-Time PCR System (Applied Biosystems) using TaqMan probes specific for the mt-Co1, mt-Nd1, mt-Nd6, and 18S rDNA (Applied Biosystems).
blotting, 1 to 2 μg of total RNA were denatured in RNA Sample Loading Buffer (Sigma-Aldrich), separated on a 1 or 1.8% formaldehyde- MOPS [3-(N-morpholino) propanesulfonic acid] agarose gels before transfer onto Hybond-N membranes (GE Healthcare) overnight. After ultraviolet (UV) crosslinking, the blots were hybridized with various probes at 65°C in Rapid-hyb buffer (Amersham) and then washed in 2× and 0.2× SSC/0.1% SDS before exposure to film. Mitochondrial probes used for visualization of mt-tRNA and mt-dtRNA levels were restriction fragments labeled with [α-32P]dATP and a random priming kit (Agilent). Different mitochondrial RNAs and 75 RNA were detected using specific oligonucleotides labeled with [γ-32P]ATP. Radioactive signals were detected by autoradiography.

In organello transcription and replication assays
In organello transcription and replication assays were performed on mitochondria isolated from mouse hearts by differential centrifugation as previously described (28). In organello transcription assays were carried out as previously reported (64). For each in organello replication assay (modified from (65)), 1 mg of purified mitochondria was washed in 1 ml of incubation buffer [10 mM tris (pH 7.4), 25 mM sucrose, 75 mM K-acetate, 100 mM KCl, 10 mM K2HPO4, 30 μM EDTA, 5 mM MgCl2, 10 mM glutamate, 2.5 mM malate, bovine serum albumin (BSA; 1 mg/ml), and 1 mM adenosine 5′-diphosphate]; resuspended in 500 μl of incubation buffer supplemented with 20 μM [α-32P]dCTP, ATP (2′-deoxyadenosine 5′-triphosphate), dGTP (2′-deoxyguanosine 5′-triphosphate), and 20 μM of [α-32P]dTTP (PerkinElmer); and incubated for 2 hours at 37°C. After incubation, mitochondria were washed three times in 10 mM tris (pH 7.0), 0.15 mM MgCl2, and 10% glycerol. An aliquot of mitochondria was collected for immunoblotting with the VDAC antibody (Millipore) as a loading control. mtDNA was isolated by phenol/chloroform extractions or by Puregene Core Kit A (Qiagen), and radiolabeled replicated DNA was analyzed by DNA sequencing using the Illumina MinION Mini Kit (Qiagen), and the concentration, purity, and integrity were confirmed using a Bioanalyzer. RNA-Seq libraries were constructed using the Illumina TruSeq Sample Prep Kit. Paired-end deep sequencing of the mitochondrial RNAs was performed on an Illumina MiSeq according to the manufacturer’s instructions. RNA-Seq was performed on mitochondrial RNA from three different control and POLRMT tissue-specific knockout mice aged 3 to 4 weeks. Sequenced reads were aligned to the mouse genome (GRCm38) with HISAT 0.1.6 (28)–aligned to the mouse genome (GRCm38) with HISAT 0.1.6 (28)–an spliced alignment disabled to reflect the unspliced nature of the mitochondrial transcriptome and prevent the introduction of spurious splice junctions. Gene-specific counts were summarized with featureCounts 1.3.5-p4 (28)–an spliced alignment disabled to reflect the unspliced nature of the mitochondrial transcriptome and prevent the introduction of spurious splice junctions. Gene-specific counts were summarized with featureCounts 1.3.5-p4 (28)–transcriptome and prevent the introduction of spurious splice junctions. Generalized linear model approach with tagwise dispersion estimates, and sequencing depth, performed with EDASeq 3.2.2 (68). Differential expression analysis was performed in R 3.2.0 with edgeR 3.11.2 using a generalized linear model approach with tagwise dispersion estimates, and the offsets were generated by EDASeq (69).

In vitro transcription assay
Transcription reactions were in 25 μl total volume containing 25 mM tris- HCl (pH 8.0), 10 mM MgCl2, 64 mM NaCl, BSA (100 μg/ml), 1 mM dithiothreitol (DTT), 400 μM ATP, 150 μM guanosine-5′-triphosphate (GTP), 150 μM cytidine-5′-triphosphate (CTP), 10 μM uridine-5′-triphosphate (UTP), 0.02 μM [α-32P]dUTP, and 4 U RNase Inhibitor Murex (New England Biolabs). The transcription template was added at 4 nM and consisted of a restriction cut (Hind III/Eco RI), purified (QiAquick PCR Purification Kit) human LSP/HSP plasmid, where a fragment consisting of positions 325 to 742 of human mtDNA was cloned between the Sma I and Hind III sites of pUC18 (23). To each reaction, 128 nM human TFFM and 320 nM human TFAM (corresponding to 1 TFAM per 40 bp) were added. Human POLRMT was added at five different concentrations: 128, 32, 8, 2, and 0.5 nM, respectively. The reactions were incubated at 32°C for 30 min and stopped by the addition of stop buffer [10 mM tris- HCl (pH 8.0), 0.2 M Na2, 1 mM EDTA, and proteinase K (100 μg/ml)] followed by incubation at 4°C for 45 min. The transcription products were ethanol-precipitated, dissolved in 20 μl loading buffer (8% formaldehyde, 10 mM EDTA, 0.1% sodium lauroyl FPLC, and 0.025% bromophenol blue), and analyzed on a 4% denaturing polycrylamide gel (1× tris-borate EDTA and 7 M urea). Quantifications of transcript levels were performed using the program MultiGauge with images generated from a PhosphorImager instrument.

Western blots, BN-PAGE, antisera, and enzymatic activity
Mitochondria were isolated from mouse heart using differential centrifugation as previously described (28). Proteins were separated by SDS-PAGE (using 4 to 12% bis-tris protein gels; Invitrogen) and then transferred onto polyvinylidene difluoride membranes (GE Healthcare). Immunoblotting was performed according to standard techniques using enhanced chemiluminescence (Immun-Star HRP Western Luminol/Enhancer from Bio-Rad). BN electrophoresis was carried out on 100 μg of mitochondria solubilized with digitonin as previously described (28). BN gels were further subjected to immunoblotting or Coomassie Brilliant Blue R staining as indicated. The following antibodies were used: NDUFA9 (complex I) and SDHA (complex II) from R&D Systems; VDAC (porin) and COX1 (complex IV) from MitoSciences; LRPPRC, TWINKLE, TFB2M, LONP, and GRSF1, SSBP1, TEFM, MRPL12, MRPS9, MRPL17, ELAC2, and tubulin from Sigma-Aldrich. Further, polyclonal antisera were used to detect COX2, TFAM, TFBM, LSP, TWINKLE, LONP, and POLRMT proteins (23, 28, 46, 54, 30). The rabbit polyclonal antiserum against recombinant mouse TWINKLE, POLRMT, and TFBM protein that lack the mitochondrial targeting signal were generated by Agriera and subsequently affinity-purified using the corresponding recombinant protein. To measure mitochondrial respiratory chain complex activities, 15 to 50 μg of mitochondria were diluted in phosphate buffer [50 mM KH2PO4, (pH 7.4)], followed by spectrophotometric analysis of isolated respiratory chain complex activities at 37°C, using a Hitachi UV-3600 spectrophotometer. To follow citrate synthase activity, increase in absorbance at 412 nm was recorded after the addition of 0.1 mM acetyl-CoA, 0.5 mM oxaloacetate, and 0.1 mM 5,5′-dithiobis-2-nitroanilic acid. SDH activity was measured.
at 600 nm after the addition of 10 mM succinate, 35 μM dichlorofluoroscein-OH, and 1 mM KCN. NADH dehydrogenase activity was determined at 340 nm after the addition of 0.25 mM NADH, 0.25 mM decylubiquinone, and 1 mM KCN and controlling for rotenone sensitivity. Citochrome c oxidase activity was measured by standard tetramethyl-p-phenylenediamine ascorbate/KCN-sensitive assays. To assess the ATPase activity, frozen isolated mitochondria (65 μg/ml) were incubated at 37°C with 75 mM Tris(hydroxymethyl)aminomethane and 1 mM MgCl₂ (pH 8.9). Mitochondria were preincubated for 2 min with amethinicin (10 μg/ml) before addition of 2 mM ATP. Samples were removed every 2 min and precipitated in 7% HCO₃⁻, and 25 mM EDTA (50 μl). Phos- phate was quantified by incubating an aliquot with 5.34 mM molybdate (1 ml), 2.68 mM ferrous sulfate, and 0.75 N H₂SO₄ for 2 min. The ab- sorbance was assayed at 600 nm. In parallel, oligomycin (2.5 μg per ml protein) was added to the mitochondrial suspension to determine the oligomycin-insensitive ATPase activity. Each activity was normalized to milligram protein by using the Lowry-based Bio-Rad protein detergent-compatible kit. All chemicals were obtained from Sigma-Aldrich.

Linear density glycerol gradients

The mtDNA bound and unbound T4AMD pools were assayed using ultracentrifugation through a 10 to 45% linear density glycerol gradient modified from previous studies (71). Cudite mitochondria (1 to 3 mg) from 4- to 5-week-old mouse hearts were isolated by differential cen- trifugation as described earlier (20, pelleted by centrifugation (15 min, 9000g/1°C), and then lysed in buffer containing 5% glycerol, 20 mM NaCl, 30 mM HEPES (pH 7.4), 1 mM EDTA, 1 mM DTT, 0.2% Triton X-100, and EDTA-free Complete protease inhibitor cocktail (Roche); and centrifuged in an SW 41 rotor at 210,000 g for 3 h to a Beckman Coulter Optima L-100 XP ultracentrifuge. Gradients were prepared using Gradient material (Beckman Instruments Inc.). Fractions (750 μl) of each fraction was assayed using Ultracell (Beckman Instruments Inc.). Fractions (750 μl) were collected from the top of the gradients, and 20 μl of each fraction was analyzed by SDS-PAGE (Invitrogen) and Western blotting. For analy- sis of mtDNA sedimentation profiles, mtDNA was isolated from two- thirds of each fraction using phenol/chloroform extraction digested with Sal I and subjected to Southern blotting.

Morphological analysis (heart sections)

Hematolunin and resin stainings were performed on paraffinembedded (PFA)-fixed cryosections from 4-week-old mouse hearts that were im- mediately embedded in OCT Tissue-Tek in cooled methyl-butan. Images of heart sections were generated by stitching of several images taken with the Nikon Eclipse CI microscope. For transmission electron microscopy, pieces of the mouse heart apex were fixed in 2% glutar- aldehyde and 2% PFA in 0.1 M sodium cacodylate buffer (pH 7.4). Specimens were postfixed in 1% osmium tetroxide (in 0.1 M sodium cacodylate buffer (pH 7.4). After thorough washing with water, speci- mens were dehydrated in ethanol followed by acetone and embedded in medium-grade Agar Low Viscosity Resin (Plano). Ultrathin sections (70 to 80 nm) were cut with a Reichert-Jung Ultracut E Ultramicrotome, stained with 2% uranyl acetate in 70% ethanol, followed by lead citrate, and examined with a Hitachi H-7650 transmission electron microscope operating at 100 kV fitted with a microdigitized AMT XR41-M digital camera (Advanced Microscopy Techniques).

Statistical analysis

Experiments were performed at least in triplicates, and results are re- presentative of n = 3 independent biological experiments. All values are expressed as mean ± SEM, unless indicated differently. Statistical sig- nificance was assessed by using two-tailed unpaired Student’s t test. Differences were considered statistically significant at a value of P < 0.05 (P < 0.05, **P < 0.01, and ***P < 0.001).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi- content/full/4/18/eaat0686/DC1

REFERENCES AND NOTES

Y. Matsushima, Y.-i. Goto, L. S. Kaguni, Mitochondrial Lon protease regulates mitochondrial 
N. Hance, M. I. Ekstrand, A. Trifunovic, Mitochondrial DNA polymerase gamma is essential 
D. F. Bogenhagen, D. Rousseau, S. Burke, The layered structure of human mitochondrial 
Y. I. Morozov, A. V. Parshin, K. Agaronyan, A. C. M. Cheung, M. Anikin, P. Cramer, 
Kühl et al. RESEARCH ARTICLE
D. A. Clayton, Replication of animal mitochondrial DNA.
D. Bogenhagen, D. A. Clayton, Mechanism of mitochondrial DNA replication in mouse L-cells:
D. D. Chang, W. W. Hauswirth, D. A. Clayton, Replication priming and transcription initiate 
I. J. Holt, A. Reyes, Human mitochondrial DNA replication.
M. Falkenberg, N.-G. Larsson, Mitochondrial transcription factor A (TFAM).
C. M. Gustafsson, M. Falkenberg, Mitochondrial RNA polymerase is needed for activation of 
J. M. Fusté, S. Wanrooij, E. Jemt, C. E. Granycome, T. J. Cluett, Y. Shi, N. Atanassova, I. J. Holt, 
B. Xu, D. A. Clayton, A persistent RNA-DNA hybrid is formed during transcription at a phy-
C. M. Gustafsson, N.-G. Larsson, Mitochondrial transcription factor A regulates mtDNA copy 
D. F. Bogenhagen, D. Rousseau, S. Burke, The layered structure of human mitochondrial 
D. D. Chang, W. W. Hauswirth, D. A. Clayton, Replication priming and transcription initiate 
I. J. Holt, A. Reyes, Human mitochondrial DNA replication.
M. Falkenberg, N.-G. Larsson, TWINKLE is an essential mitochondrial helicase required for synthesis of nascent 
H. Antonicka, F. Sasarman, T. Nishimura, V. Paupe, E. A. Shoubridge, The mitochondrial 
M. Lagouge, A. Mourier, H. J. Lee, H. Spåhr, T. Wai, C. Kukat, E. S. Ramos, E. Motori, 
A. Reyes, L. Kazak, S. R. Wood, T. Yasukawa, H. T. Jacobs, I. J. Holt, Mitochondrial DNA 
R. Puzicha, M. Kippe, M. Wipplinger, C. D. Sleigh, R. Lightowlers, Z. M. Orzalesi-Tovarnik, I.-C. Martínez, qRT-PCR regulates mitochondrial RNA processing in human 
D. Litonin, M. Sologub, Y. Shi, M. Savkina, M. Anikin, M. Falkenberg, C. M. Gustafsson, 
E. J. Emot, Ö. Persson, Y. Shi, M. Medenovic, J. P. Uhr, M. David Lopez, C. Freyer, 
P. L. Polosa, S. Deceglie, M. Roberti, M. N. Gadaleta, P. Cantatore, Contrahelicase activity of 
A. M. Gillum, D. A. Clayton, Mechanism of mitochondrial DNA replication in mouse L-cells: 
A. M. Gillum, D. A. Clayton, Mechanism of mitochondrial DNA replication in mouse L-cells: 
K. Ichimura, T. Tanaka, S. Takeuchi, T. H. Kison, S. Sato, T. Hayakawa, K. M. Okano, 
51. T. Richman, J. S. Mattick, O. Rackham, A. Filipovska, RNA processing in human mitochon-
D. L. Maccioni, A. windowHeight, W. C. Gimenez, A. M. F. Boni, T. H. Kison, S. Sato, 
54. B. Ruzzenente, M. D. Metodiev, A. Wredenberg, A. Bratic, C. B. Park, Y. Cámara,
53. F. Sasarman, C. Brunel-Guitton, H. Antonicka, T. Wai, E. A. Shoubridge, LSFC Consortium, 
58. E. Jemt, Ö. Persson, Y. Shi, M. Medenovic, J. P. Uhr, M. David Lopez, C. Freyer, 
57. C. M. Gustafsson, M. Falkenberg, Mitochondrial DNA replication and transcription. 
55. P. L. Polosa, S. Deceglie, M. Roberti, M. N. Gadaleta, P. Cantatore, Contrahelicase activity of 
59. P. L. Polosa, S. Deceglie, M. Roberti, M. N. Gadaleta, P. Cantatore, Contrahelicase activity of 
58. E. Jemt, Ö. Persson, Y. Shi, M. Medenovic, J. P. Uhr, M. David Lopez, C. Freyer, 
48. D. Litonin, M. Sologub, Y. Shi, M. Savkina, M. Anikin, M. Falkenberg, C. M. Gustafsson, 
47. A. M. Gillum, D. A. Clayton, Mechanism of mitochondrial DNA replication in mouse L-cells: 
38. D. A. Clayton, Replication of animal mitochondrial DNA.
37. I. J. Holt, A. Reyes, Human mitochondrial DNA replication.
36. M. Falkenberg, N.-G. Larsson, DNA replication and transcription in mammalian 
35. H. H. Hurlburt, Y. J. Forgay, A. Pandi, G. Tkacz, W. C. Copeland, Holp is essential for mammalian 
34. J. H. Price, F. Farge, Y. Shi, M. G. Wohlgemuth, M. Falkenberg, Conserved sequence boxes direct 
33. B. Xu, D. A. Clayton, Persistent RNA-DNA hybrid is formed during transcription on a phy-
32. D. D. Chang, W. W. Hauswirth, D. A. Clayton, Replication priming and transcription initiate 
31. D. D. Chang, D. A. Clayton, Printing of human mitochondrial DNA replication occurs in the 
27. C. M. Gustafsson, M. Falkenberg, Mitochondrial DNA polymerase is needed for activation of the 
18. D. A. Clayton, Replication of animal mitochondrial DNA.
Acknowledgments: We thank N. Hochhard for technical assistance and C. K. Suzuki for providing the mouse LONP antibody. RNA library construction and sequencing were performed at the Cologne Center for Genomics. ECG analysis was performed with the ECG screening system of the pheno-typing core facility, Max Planck Institute for Biology of Ageing. We are grateful for the technical support from the CECAD Imaging Facility and the FACS & Imaging Core Facility of the Max Planck Institute for Biology of Ageing.

Funding: This study was supported by a European Research Council Advanced Investigator grant (268887), the Swedish Research Council (2015-0418–2013-365), and the Knut and Alice Wallenberg Foundation to N.-G.L. C.M.G. was funded by the Swedish Research Council (2012-2842), the Swedish Cancer Foundation, and the Knut and Alice Wallenberg Foundation. A.F. is a National Health and Medical Research Council Senior Research Fellow (APP 1058442), supported by an Alexander von Humboldt Fellowship for experienced researchers and by the National Health and Medical Research Council (APP105277). Author contributions: Conceived and designed the experiments: I.K., P.L.P., C.M.G., and N.-G.L. Performed experiments: I.K., M.M., V.P., D.M., A.M., N.A.B., U.N., P.L.P., and A.F. Analyzed the data: I.K., M.M., V.P., S.J.S., and A.F. Wrote the paper: I.K., C.M.G., and N.-G.L. Competing interests: The authors declare that they have no competing interests. Data and materials availability: RNA-Seq data have been deposited in the Gene Expression Omnibus repository under accession no. GSE83368. Additional data related to this paper may be requested from the authors. Knockout mouse models presented in the paper must be obtained through a material transfer agreement.

Submitted 2 May 2016
Accepted 1 July 2016
Published 5 August 2016
10.1126/sciadv.1600963

POLRMT regulates the switch between replication primer formation and gene expression of mammalian mtDNA

Inge Küh, Maria Miranda, Viktor Posse, Dusanka Milekovic, Arnaud Mourier, Stefan J. Siira, Nina A. Bonekamp, Ulla Neumann, Aleksandra Filipovska, Paola Loguercio Polosa, Claes M. Gustafsson and Nils-Göran Larsson (August 5, 2016) Sci Adv 2016, 2:
doi: 10.1126/sciadv.1600963

This article is published under a Creative Commons license. The specific license under which this article is published is noted on the first page.

For articles published under CC BY licenses, you may freely distribute, adapt, or reuse the article, including for commercial purposes, provided you give proper attribution.

For articles published under CC BY-NC licenses, you may distribute, adapt, or reuse the article for non-commercial purposes. Commercial use requires prior permission from the American Association for the Advancement of Science (AAAS). You may request permission by clicking here.

The following resources related to this article are available online at http://advances.sciencemag.org. [This information is current as of August 10, 2016]:

Updated information and services, including high-resolution figures, can be found in the online version of this article at: http://advances.sciencemag.org/content/2/8/e1600963.full

Supporting Online Material can be found at: http://advances.sciencemag.org/content/suppl/2016/08/01/2.8.e1600963.DC1

This article cites 71 articles, 33 of which you can access for free at: http://advances.sciencemag.org/content/2/8/e1600963#BIBL

Science Advances (ISSN 2375-2548) publishes new articles weekly. The journal is published by the American Association for the Advancement of Science (AAAS), 1200 New York Avenue NW, Washington, DC 20005. Copyright is held by the Authors unless stated otherwise. AAAS is the exclusive licensee. The title Science Advances is a registered trademark of AAAS.
POLRMT regulates the switch between replication primer formation and gene expression of mammalian mtDNA

Inge Kühl, Maria Miranda, Viktor Posse, Dusanka Milenkovic, Arnaud Mourier, Stefan J. Siira, Nina A. Bonekamp, Ulla Neumann, Aleksandra Filipovska, Paola Loguercio Polosa, Claes M. Gustafsson, Nils-Göran Larsson

DOI: 10.1126/sciadv.1600963

This PDF file includes:

- fig. S1. Targeting strategy for disrupting the Polrmt gene in mouse.
- fig. S2. ECG analysis of conditional Polrmt knockout mice.
- fig. S3. Total protein levels of nuclear-encoded factors involved in mtDNA expression.
- fig. S4. Steady-state mitochondrial protein levels of conditional Polrmt knockout mice.
- fig. S5. Relative mitochondrial protein distribution across a gradient.
- fig. S6. Relative mitochondrial protein distribution in Polrmt knockout mice.
- fig. S7. Respiratory chain complex protein levels in heterozygous Polrmt knockout mice.
- fig. S8. Relative mitochondrial RNA abundance in heterozygous Polrmt knockout mice.
- fig. S9. Steady-state mitochondrial protein levels in heterozygous Polrmt knockout hearts.
- fig. S10. mtDNA levels in heterozygous Polrmt knockout hearts.
- table S1. Summary statistics and P values of RT-qPCR presented in Fig. 4B.
fig. S1. Targeting strategy for disrupting the *Polrmt* gene in mouse. The mutated locus was transmitted through the mouse germ line resulting in heterozygous *Polrmt*+/loxP-puro mice. These mice were subsequently mated with transgenic mice ubiquitously expressing the flp recombinase, to remove the puromycin resistance cassette, which resulted in the generation of mice heterozygous for a loxP-flanked *Polrmt* allele. Next, these *Polrmt*+/loxP mice were crossed to β-actin-cre mice to obtain heterozygous *Polrmt* knockout mice (*Polrmt*+/−).
fig. S2. ECG analysis of conditional Polm

knockout mice. Electrocardiographs of knockout mice (L/L, cre; N = 14) and controls (L/L; N = 15) were assessed at 4 weeks of age. Heart rate (HR), heart rate variability (HRV) and rate-corrected QT interval (QTC).
fig. S3. Total protein levels of nuclear-encoded factors involved in mtDNA expression.
Western Blots were performed on total extracts isolated from hearts of control (L/L) and tissue-specific knockout (L/L, cre) mice at different ages. (A) Two different membranes are presented with their corresponding loading control (tubulin). (B to C) Quantification of protein levels in total extracts from knockout (L/L, cre) heart at different time points; values are in percent of control (L/L) values. Loading was normalized to tubulin. E18.5 = embryonic day 18.5.
fig. S4. Steady-state mitochondrial protein levels of conditional Polrmt knockout mice. (A) Quantification of western blot shown in Fig. 4A. (B) Western blot analysis were performed on mitochondria isolated from hearts of control (L/L) and tissue-specific knockout (L/L, cre) mice at different ages. Five different membranes are presented with their corresponding loading control (VDAC). (C) Quantification of B normalized to VDAC signals and controls (L/L). Error bars: ± SEM; * $P < 0.05$, ***$P < 0.01$, ****$P < 0.001$; two-tailed Student’s t-test.
fig. S5. Relative mitochondrial protein distribution across a gradient. Quantification of signals obtained in the 16 different gradient fractions from mitochondrial lysates of control mice (L/L) via western blotting shown in Fig. 4C. (A) Distribution of some proteins of the mitochondrial replication
machinery compared to TFAM levels. (B) Distribution of the proteins of the mitochondrial transcription machinery. (C) Distribution of the proteins involved in mtRNA metabolism and mitochondrial ribosomal proteins.

**fig. S6. Relative mitochondrial protein distribution in Polrmt knockout mice.** Quantification of signals obtained in the 16 different gradient fractions from mitochondrial lysates of Polrmt knockout mouse (L/L, cre) via western blotting shown in Fig. 4D. Distribution of some proteins of the mitochondrial replication machinery compared to TFAM levels.
fig. S7. Respiratory chain complex protein levels in heterozygous Polrmt knockout mice. (A)
Western blots were performed on mitochondria isolated from heart, skeletal muscle and liver of wild-type (+/+) and heterozygous (+/−) knockout mice at different ages; loading control: VDAC. (B)
Quantification. (C) Quantification of western blot shown in Fig. 6A; * P < 0.05; two-tailed Student’s t-test. (D) Western blots performed on mitochondria isolated from heart, skeletal muscle and liver of wild-type (+/+) and heterozygous (+/−) knockout mice at different ages; loading control: SDHA (complex II). Error bars: ± SEM.
fig. S8. Relative mitochondrial RNA abundance in heterozygous Polrmt knockout mice. (A)
Quantification of relative mitochondrial transcript levels normalized to wild-type (+/+) mice, see
northern blot in Fig. 6B. (B) RNA-Seq on mitochondrial mRNA and rRNA levels in hearts of 26 weeks-
old wild-type (+/+) and heterozygous (+/−) knockout mice normalized to the upper quartile of the gene
count distribution. The data analyzed are from three independent RNA-Seq experiments. Error bars: ±
SEM. (* P < 0.05).
fig. S9. Steady-state mitochondrial protein levels in heterozygous *Polrmt* knockout hearts. (A)
Combined quantification data of western blots shown in Fig. 6C and fig. S9B. (B) Western blot analysis of mitochondria isolated from heart of wild-type (+/+) and heterozygous (+/-) knockout mice at different ages; loading control: VDAC. Error bars: ± SEM.
fig. S10. mtDNA levels in heterozygous Polrmt knockout hearts. (A) Southern blots were performed on total DNA isolated from hearts of adult wild-type (+/+) and heterozygous (+/-) knockout mice; loading: 18S rDNA. (B) Quantification of relative de novo synthesized mtDNA normalized to L/L signal, shown in Fig. 6G. Error bars: ± SD.
table S1. Summary statistics and *P* values of RT-qPCR presented in Fig. 4B.

<table>
<thead>
<tr>
<th>gene</th>
<th>genotype</th>
<th>mean</th>
<th>SEM</th>
<th>N</th>
<th><em>P</em>-value</th>
<th>Type of t test</th>
<th>t test</th>
<th><em>P</em>-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>P0mt</td>
<td>L/L</td>
<td>1000</td>
<td>6.963</td>
<td>19</td>
<td>6.314E-10</td>
<td>2-tails, unequal variance</td>
<td>1.23E-10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>12,012</td>
<td>1.277</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>md-Atp6</td>
<td>L/L</td>
<td>1000</td>
<td>22.842</td>
<td>6</td>
<td>1.965E-65</td>
<td>2-tails, unequal variance</td>
<td>5.58E-03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>2,356</td>
<td>1.563</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16S rRNA</td>
<td>L/L</td>
<td>1000</td>
<td>16.988</td>
<td>6</td>
<td>2.776E-03</td>
<td>2-tails, unequal variance</td>
<td>2.14E-03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>12,998</td>
<td>3.327</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tbk1m</td>
<td>L/L</td>
<td>1000</td>
<td>7.831</td>
<td>5</td>
<td>0.686</td>
<td>2-tails, equal variance</td>
<td>0.329</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>112,956</td>
<td>9.715</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tfos</td>
<td>L/L</td>
<td>1000</td>
<td>9.502</td>
<td>19</td>
<td>0.270</td>
<td>2-tails, equal variance</td>
<td>0.574</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>128,966</td>
<td>11.734</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tiam</td>
<td>L/L</td>
<td>1000</td>
<td>6.199</td>
<td>5</td>
<td>0.591</td>
<td>2-tails, equal variance</td>
<td>4.00E-04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>151,956</td>
<td>6.215</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tmno</td>
<td>L/L</td>
<td>1000</td>
<td>9.715</td>
<td>5</td>
<td>0.025</td>
<td>2-tails, equal variance</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>121,453</td>
<td>1.277</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PrP</td>
<td>L/L</td>
<td>1000</td>
<td>11.639</td>
<td>16</td>
<td>0.023</td>
<td>2-tails, unequal variance</td>
<td>0.694</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>107,872</td>
<td>6.779</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sreb1</td>
<td>L/L</td>
<td>1000</td>
<td>7.019</td>
<td>19</td>
<td>0.949</td>
<td>2-tails, equal variance</td>
<td>0.193</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>112,785</td>
<td>6.591</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mvor2</td>
<td>L/L</td>
<td>1000</td>
<td>16.045</td>
<td>4</td>
<td>0.964</td>
<td>2-tails, equal variance</td>
<td>0.568</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>112,980</td>
<td>14.204</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mvor3</td>
<td>L/L</td>
<td>1000</td>
<td>8.657</td>
<td>9</td>
<td>0.048</td>
<td>2-tails, equal variance</td>
<td>0.782</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>96,735</td>
<td>7.710</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mvor4</td>
<td>L/L</td>
<td>1000</td>
<td>10.042</td>
<td>9</td>
<td>0.714</td>
<td>2-tails, equal variance</td>
<td>0.782</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>104,194</td>
<td>10.904</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lypic</td>
<td>L/L</td>
<td>1000</td>
<td>8.427</td>
<td>5</td>
<td>0.175</td>
<td>2-tails, equal variance</td>
<td>0.163</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L/L, cre</td>
<td>130,305</td>
<td>17.870</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B
Hierarchical RNA Processing Is Required for Mitochondrial Ribosome Assembly

Graphical Abstract

Highlights
- MRPP3 is essential and its activity is non-redundant in mitochondria
- 5’ tRNA cleavage precedes 3’ end processing in vivo
- RNA processing is required for biogenesis of mitoribosomes
- RNA processing links transcription to translation via mitoribosome assembly

Authors
Oliver Rackham, Jakob D. Busch, Stanka Matic, ..., Dusanka Milenkovic, Nils-Göran Larsson, Aleksandra Filipovska

Correspondence
aleksandra.filipovska@uwa.edu.au

In Brief
Rackham et al. generate conditional knockout mice lacking the endoribonuclease component of the mitochondrial RNase P and show it is essential. They show that 5’ tRNA cleavage precedes 3’ processing and that processing is required for mitoribosome biogenesis. They find that RNA processing links transcription to translation via mitoribosome assembly.

Accession Numbers
GSE83471
Hierarchical RNA Processing Is Required for Mitochondrial Ribosome Assembly

Oliver Rackham,1,4 Jakob D. Busch,2 Stanka Matic,2 Stefan J. Silra,1 Irina Kuznetsova,1 Ilian Atanassov,1 Judith A. Ermer,1 Anne-Marie J. Shearwood,3 Tara R. Richman,1 James B. Stewart,3 Arnaud Mourier,3,5 Dusanka Milenkovic,3 Nils-Goran Larsson,3,6 and Aleksandra Filipovska1,2,*

1Harry Perkins Institute of Medical Research and Centre for Medical Research
2School of Chemistry and Biochemistry
The University of Western Australia, Nedlands, WA 6009, Australia
3Department of Mitochondrial Biology, Max Planck Institute for Biology of Ageing, 50931 Cologne, Germany
4Proteomics Core Facility, Max Planck Institute for Biology of Ageing, D-50931 Cologne, Germany
5CNRS, Institut de Biochimie et Génétique Cellulaires UMR 5095, Saint-Saëns, 33077 Bordeaux, France
6Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm 17177, Sweden
*Correspondence: aleksandra.filipovska@uwa.edu.au

SUMMARY
The regulation of mitochondrial RNA processing and its importance for ribosome biogenesis and energy metabolism are not clear. We generated conditional knockout mice of the endonuclease component of the RNase P complex, MRPP3, and report that it is essential for life and that heart and skeletal-muscle-specific knockout leads to severe cardiomyopathy, indicating that its activity is non-redundant. Transcriptome-wide parallel analyses of RNA ends (PARE) and RNA-seq enabled us to identify that in vivo 5' RNA cleavage precedes 3' RNA processing, and this is required for the correct biogenesis of the mitochondrial ribosomal subunits. We identify that mitoribosomal biogenesis proceeds co-translationally because large mitoribosomal proteins can form a subcomplex on an unprocessed RNA containing the 16S rRNA. Taken together, our data show that RNA processing links transcription to translation via assembly of the mitoribosome.

INTRODUCTION
The mammalian mitochondrial genome is compact and relatively small, only ~16 kb in size, and encodes 11 mRNAs, as well as two rRNAs and 22 tRNAs required for translation (Figs. 1A and reviewed in Rackham et al., 2012). Mammalian mitochondrial mRNAs code for 13 proteins, which are all essential components of the respiratory chain and the ATP synthase that are responsible for energy maintenance by oxidative phosphorylation (OXPHOS). Although the mitochondrial OXPHOS complexes are predominately composed of polypeptides that are encoded by nuclear genes and imported into mitochondria from the cytosol, their assembly and function is dependent on the expression of the mitochondria-encoded polypeptides (Hallberg and Larsson, 2014). Furthermore, the mitochondrial genome is dependent on nuclear-encoded proteins for replication, repair, transcription, and translation (Hallberg and Larsson, 2014; Rackham et al., 2012). Consequently, the expression of both mitochondrial and nuclear genes requires cooperative regulation to enable balanced energy production. Nuclear gene expression is significantly correlated with mitochondrial gene expression, reflecting close coordination between nuclear and mitochondrial genomes in relation to the energy demands of different tissues (Mercer et al., 2011).

Mitochondrial genes coding for proteins and RNAs are located on both the heavy and light strands of the mtDNA, which are transcribed as large polycistronic transcripts covering almost the entire length of each strand (Hallberg and Larsson, 2014; Montoya et al., 1981). This differs significantly from nuclear gene transcription and requires several unique processing steps to form functional RNAs. In nearly all cases, genes encoding for protein or rRNA are interspersed by one or more tRNAs, which were proposed to act as “punctuation” marks for processing by Attardi and colleagues (Attardi et al., 1981). This processing involves cleavage at the 5' end of tRNAs by the RNase P complex (Holzmann et al., 2008) and cleavage of the 3' end by the mitochondrial RNase Z (Broszczyk et al., 2011; Sanchez et al., 2011). Processing is followed by maturation of the RNAs, the assembly of the RNAs into mitochondrial ribosomes, and translation of the mRNAs using tRNAs. Although previously not investigated in vivo, it has been proposed that regulation of the processing of mitochondrial RNAs might have profound effects on mitochondrial gene expression, the assembly of the protein synthesis machinery, and the overall level of translation, particularly in mitochondrial diseases caused by mutations in regulatory mtDNA-encoded genes such as RNAs and in mitochondrial RNA-binding proteins that regulate mitochondrial gene expression and are nuclear encoded (Sanchez et al., 2011; Rackham et al., 2013).

In animal mitochondria, the RNase P complex lacks the RNA catalytic component and is composed of only three protein subunits, mitochondrial RNase P protein 1 (MRPP1), MRPP2, and MRPP3 (Holzmann et al., 2008). The MRPP1 subunit of mitochondrial RNase P enzyme has a conserved guanine methyltransferase domain that modifies adenines or guanines at position 9 of mitochondrial RNAs in addition to its role as a subunit for replication, repair, transcription, and translation (Hallberg and Larsson, 2014; Rackham et al., 2012).

© 2016 The Author(s). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Mrpp3 mRNA levels relative to control and normalized to 18S rRNA

Heart to body weight ratio (%)

Wild type Myg3 locus

Targeting vector

Targeted allele after homologous recombination

Conditional KO allele after Flp recombination

Constitutive KO allele after Cre recombination

Legends on next page

Cell Reports 16, 1874–1890, August 16, 2016 1875
of RNase P (Sanchez et al., 2011; Vlardo et al., 2012). MRPP2 is a member of the short-chain dehydrogenase superfamily, and mutations in the gene encoding this protein cause mitochondrial disease as a result of impaired RNA processing, further indicating that it is required for RNase P activity (Vlardo and Rossmann, 2015; Zschocke, 2012). MRPP3 consists of a metallo-

**RESULTS**

MRPP3 Is Essential for Development, and Heart-

Specific Loss of MRPP3 Causes Cardiomyopathy

MRPP3 shares homology with the transcriptase RNase P (PRORP) enzymes that lack an RNA catalytic subunit and cleave the 5′ ends of RNAs in mitochondria, chloroplasts as well as the nucleus of plants and protists (Gibbert et al., 2010; Taichman et al., 2012). Despite having an endonuclease domain for the cat-

alytic cleavage of RNAs, the mitochondrial MRPP3 protein in animal mitochondria requires two additional proteins MRPP1 and MRPP2 for RNase P activity (Hoizmann et al., 2008). To un-

derstand the in vivo role of MRPP3 in animal mitochondria, we generated a conditional knockout allele of the mouse Mrpp3 gene by flanking exon 3 with loxP sites in embryonic stem (ES) cells (Figures 1B and 1C) and subsequent transmission through the germline to obtain heterozygous Mrpp3+/− mice. These mice were crossed with transgenic mice expressing the Fip recombinase to excise the puromycin cassette (Figure 1B), and the resulting Mrpp3+/− mice were mated to mice expressing Cre recombinase under the control of the β-actin promoter to generate heterozygous Mrpp3 knockout mice (Mrpp3+/−). Inter-

crossing the Mrpp3−/− mice produced Mrpp3−/− and Mrpp3+/− mice in Mendelian proportions (genotyped pups n = 115, Mrpp3−/− n = 66 and Mrpp3+/− n = 49); however, the homozy-
gous knockout mice (Mrpp3−/−) were not viable. Previous ana-
lyses of essential proteins involved in mitochondrial gene expression have observed embryonic lethality at E8.5 (Camara et al., 2011; Larsson et al., 1998; Metodiev et al., 2009, 2014; Park et al., 2007; Ruzzenente et al., 2012); therefore, we analyzed the embryos at this stage to identify that Mrpp3−/− em-

bryos had not developed normally compared to those observed for Mrpp3−/− and Mrpp3+/− mice, and hence the loss of MRPP3 was embryonic lethal (Figure 1D). Based on these findings, we conclude that MRPP3 is essential for embryo development and survival.

**Figure 1. Full-Body Knockout of MRPP3 Causes Embryonic Lethality and Cardiomyopathy in Heart-Specific Knockout Mice**

(A) Schematic representation of the gene organization in mammalian mtDNA.
(B) Schematic showing the disruption of the Mrpp3 gene (also known as MT00605550 and KIAA0017). LoxP sites flanking exon 3 of the Mrpp3 gene were inserted in the mouse genome by homologous recombination. The long homology arm (LHA), floxed regions (FR), and short homology arm (SHA) are shown.
(C) Homologous recombination at the Mrpp3 locus shown by Southern blotting.
(D) Morphology of the Mrpp3+/+ and Mrpp3−/− embryos at day E8.5. Scale bar, 1 mm.
(E) Heart weight to body weight ratio in control (L/L) and knockout mice (L/L, cre) at different ages. At 4 weeks, L/L, n = 8, L/L, cre n = 8; at 8 weeks, L/L, n = 12, L/L, cre n = 12; at 11 weeks, L/L, n = 12, L/L, cre n = 12. Error bars, SEM. *p < 0.01; **p < 0.001. Student’s t-test.
(F) H&E, NADH, and COX and trichrome gomorii staining of hearts from control (L/L) and knockout mice (L/L, cre).
(G) Levels of Mrpp3 transcript levels were determined in isolated total heart RNA from control (L/L) and knockout mice (L/L, cre) by qRT-PCR (n = 5). The data are expressed relative to control mice and normalized to 185 rRNA.
Next, we crossed Mrpp3<sup>loxP/loxP</sup> mice with transgenic mice expressing a Cre recombinase under the control of muscle creatine kinase promoter (Cmm-cre) to produce heart- and skeletal-muscle-specific Mrpp3 knock-out mice (Mrpp3<sup>Cre-cre</sup>, Cmm-cre). The MRPP3 mice with Cmm-cre-directed knock-out have a short lifespan and die by 11 weeks of cardiomyopathy, determined by the increased size of the heart, by measuring the heart weight to body weight ratio (Figure 1E) and by histology, indicating the accumulation of necrotic foci in the myocardium of these mice (Figure 1F). Furthermore, we observed that the transcription factors ATF4 and ATF5, that are elevated in cardiomyopathy as a result of mitochondrial dysfunction (Dogan et al., 2014), are significantly increased in the Mrpp3 knockout mice compared to their age and sex matched controls (Mrpp3<sup>loxP/loxP</sup>), and this increase is most apparent at 11 weeks (Figure 1G). These physiological changes in the knock-out mice are a consequence of the Mrpp3 deletion and loss of Mrp3 expression measured by Q-PCR over 4, 8, and 11 weeks (Figure 1H). We confirmed that loss of Mrp3 expression in skeletal muscle over the same time (Figure S1A) resulted in reduced muscle fiber appearance and reduction in complex I and complex IV staining (Figure S1B), without any effects on ATF4 and ATF5 expression (Figure S1C), indicating that cardiomyopathy is the major cause of death in these mice.

**Defects in RNA Processing Increase the Rate of Transcription and Impair Protein Synthesis**

We investigated the effects of MRPP3 loss on mitochondrial transcription by de novo pulse and chase labeling with [<sup>35</sup>S]radio-labeled uridine triphosphate (UTP) and identified that impaired RNA 5' end processing led to a significantly increased rate of transcription relative to controls (Figure 2A). It is interesting that the mature 12S and 16S rRNAs that most prominently featured in the controls are not observed in the Mrpp3 knockout mice, and instead we observe higher molecular weight unprocessed transcripts in these mice (Figure 2A). Northern blotting of these membranes confirmed that mitochondrial mRNAs, tRNAs, and rRNAs are contained within the higher molecular weight bands in the Mrpp3 knockout mice and that there are no mature transcripts produced in these mice compared to controls (Figure S2B). In contrast, de novo mitochondrial translation measured using [<sup>35</sup>S]-methionine and cytosine incorporation revealed significant reduction in protein synthesis of all mitochondrial encoded proteins in the Mrpp3 knockout mice compared to controls (Figure 2B). Immunoblotting of mtDNA- and nuclear-encoded mitochondrial supercomplexes revealed a significant decrease in their abundance of all proton translocating respiratory complexes in control and Mrpp3 knockout mice (Figure 2C). The activities of complexes I, III, and IV were all significantly decreased levels observed on BN-PAGE and immunoblotting (Figures 2D and S2C). Complex II activity was slightly increased (Figure 2E), which is typically observed when there is a general oxidative phosphorylation (OXPHOS) defect (Duff et al., 2015). Measurement of mitochondrial oxygen consumption showed that there was a significant reduction in the non-phosphorylated, phosphorylated, and uncoupled respiration state in the Mrpp3 knockout mice (Figure 2F). Therefore, loss of MRPP3 causes profound mitochondrial dysfunction by causing general OXPHOS defects, further indicating that RNA processing is essential for mitochondrial function.

**The Mitochondrial RNase P Regulates the Processing of Canonical RNA 5' Ends**

MRPP3 has been shown to cleave the 5' ends of mitochondrial tRNAs when it is part of the RNase P complex with MRPP1 and MRPP2 in vitro (Holzman et al., 2008; Li et al., 2015; Reinhardt et al., 2015). We have shown that the mitochondrial RNase P cleaves the 5' ends of mitochondrial tRNAs and can have promiscuous activity toward RNAs that have tRNA-like structures at their 5' ends in cells (Sanchez et al., 2015). However, all of these effects have been investigated in cells grown in high glucose where the MRPP3 levels have been knock-downed and given the high stability of the protein it is necessary to investigate the effects of MRPP3 loss in vivo. Furthermore the hierarchy of RNA processing has not been investigated previously; therefore, we analyzed the effects of the MRPP3 loss on the order of cleavage and abundance of steady-state mitochondrial RNAs in hearts by northern blotting over 4, 8, and 11 weeks. MRPP3 loss leads to gradual loss of mature mRNAs such as mt-Co1, mt-Nduf1, and mt-Atp8, and mt-Cyt b by 11 weeks and a converse increase in the accumulation and end processing of precursor transcripts (Figure 3A). In contrast, the mature mt-Nd5...
Figure 2. Loss of MRPP3 Impairs Mitochondrial Biogenesis and Increases Transcription

(A) In organello transcription was measured in heart mitochondria in the presence of 32P-labeled UTP, and the radiolabeled RNA was isolated and resolved on a 0.9% formaldehyde gel that was visualized by autoradiography. VDAC was used as a loading control. The data are representative of results obtained for 11-week-old L/L and L/L, cre mice (n = 6) in three independent biological experiments.

(B) Protein synthesis in heart mitochondria from 11-week-old control and knockout mice was measured by pulse incorporation of 35S-labeled methionine and cysteine. Equal amounts of mitochondrial protein (50 μg) were separated by SDS-PAGE, immunoblotted against antibodies to investigate the steady-state levels of nuclear and mitochondrial-encoded proteins. ATP5A was used as a loading control.

(C) Mitochondrial proteins (20 μg) from 4-, 8-, and 11-week-old heart mitochondria from control and knockout mice were separated by SDS-PAGE gels and immunoblotted against antibodies to investigate the steady-state levels of nuclear and mitochondrial-encoded proteins. ATP5A was used as a loading control.

(D) Heart mitochondria (75 μg) treated with 1% digitonin were separated on a 4%–30% BN-PAGE, and the respiratory complexes were visualized by Coomassie staining. In-gel activity stains were used for complex I and complex IV.

(E) Respiratory complex activities were measured spectrophotometrically and normalized to citrate synthase activity in mitochondria isolated from hearts of 11-week-old control and knockout mice (n = 6). Error bars, SEM; **p < 0.01; ***p < 0.001, Student’s t test.

(F) Loss of the MRPP3 causes profound reduction in mitochondrial respiration at complex I, II-III, and IV in the KO mice compared to controls. Non-phosphorylating (state 4), phosphorylating (state 3), and uncoupled respiration in the presence of 0.5 μM carbonyl cyanide 3-chlorophenylhydrazone (CCCP) was measured in heart mitochondria using an Oroboros oxygen electrode using pyruvate, glutamate, and malate, or succinate as substrates in the presence of substrates as indicated. Error bars, SEM; **p < 0.01; ***p < 0.001, Student’s t test.
levels are not dramatically affected by the loss of Mrpp3 (Figure S3A), which is likely a consequence of its organization within the mitochondrial genome (Figure 1A). The 3′ end of mt-tRNA is in effect the 5′ end of mt-ND5 that is cleaved by the mitochondrial RNase Z along with its 3′ end, thereby allowing it to be produced as a correctly processed mature RNA. The 3′ end of mt-tRNA is a “mirror tRNA” of mt-tRNA, which must be processed by either ELAC2 or another endonuclease because it seems to be unaffected by the loss of Mrpp3. Mrpp3 loss also leads to lack of mature RNAs as a consequence of their accumulation without unprocessed transcripts (Figure 3B). The levels of the mature mitochondrial RNAs are significantly decreased, whereas unprocessed transcripts accumulates in the Mrpp3 knockout mice relative to controls (Figure 3C). It is interesting to note that the organization of the 12S and 16S rRNAs should allow them to be processed by the RNase Z at the 3′ ends thereby leaving them with RNA5′ and RNA3′(N 5′,C-C) attached at their 5′ ends, respectively. However, instead we observe loss of each 12S and 16S RNA by 11 weeks (Figure 3C) providing the first experimental evidence that 3′ tRNA processing is dependent on functional 5′ RNA processing to release the RNAs.

We used qR-PCR to investigate processing events across every canonical junction in the mitochondrial transcriptome (Figure 3D) and to account for any junctions that are not processed by the RNase P complex that was not possible to assess by northern blotting because of the organization of the genome (Figure 1A) and the hierarchical RNA processing. There was significant accumulation of precursors in the Mrpp3 knockout mice compared to controls. The magnitude of changes in these precursors indicates the stabilities of their unprocessed forms as well as the efficiency of normal processing at these junctions (Figure 3D). For example the mt-Cyt b-tRNA precursor was the most abundant and likely most stable, compared to the mt-tRNA Met and mt-tRNA Val precursors that were least abundant because of the multiple ELAC2 sites that span mt-tRNA Met and mt-tRNA Val and do not require Mrpp3 for processing (Figure 3D). Interestingly, because the qRT-PCR examined individual junctions we were able to identify two sites, between the mt-ATP6/8 and mt-Co3 and the 5′ end of mt-ND5, that are not cleaved by RNase P. We have shown previously that the 5′ end of mt-ND5 is produced by cleavage at the 3′ end of mt-tRNA by the RNase Z, ELAC2 (Banchet et al., 2011), and FASTKD5 affects the processing of the mt-ATP6/8 and mt-Co3 precursor, commonly known as RNA44 (Arntsvich and Shubinina, 2011). We identified a similar increase in junction accumulation in skeletal muscle (Figure S2D), identifying that mitochondrial dysfunction (Figures S1A–S1C) was a result of impaired RNA processing due to loss of Mrpp3 in this tissue.

**Figure 3. Loss of Mrpp3 Causes Impaired RNA Processing**

(A–D) The abundance of mature mitochondrial RNAs (5′-tRNAs, 3′-tRNAs, and mt-tRNAs) and tRNAs (C) in hearts from 4-, 8-, and 11-week-old control and knockout mice were analyzed by northern blotting. 18S RNA was used as a loading control. The data are representative of results obtained from at least eight mice from each strain and three independent biological experiments.

(B) The canonical mitochondrial tRNA junctions were measured in total heart RNA from control (L/L) and knockout (L/L, cre) mice by qRT-PCR and normalized to 18S RNA. Error bars, SEM; *p < 0.05; **p < 0.01. Student’s t test.

(C) Frequency distribution of PARE reads from heavy mitochondrial RNA from three control (L/L, blue) and three knockout (L/L, red) mice. Windows are centered on reads that align 15 nt away from either side of all annotated 5′ ends of mitochondrial tRNAs from the heavy and light chain.
The absence of this motif results in the heavy and light strand (Figure S3 C). We find that the sites where the ELAC2 activity is increased in the knockout mRNAs that is not covered in our PARE datasets (Figure 4 D). This effect might be due to the increased stability or impaired processing and increased transcription, which produces a high load of unprocessed RNAs within mitochondria (Figure 4A). This was also apparent in the significant reduction of mature mtRNAs in the RNA-seq datasets (Figure 4A) as well as those found by northern blotting (Figures 3A–3C). We investigated the specificity of mtRNA degradation digitally by searching for motifs around the cleavage sites that are enriched in the absence of MRPP3 and correlated to the abundance of each nucleotide in the heavy and light strand of the mouse mitochondrial transcriptome (Figure S4B). We conclude that the degradation is carried out by an endonuclease that lacks sequence specificity.

Mitochondrial RNA Processing is Required for the Assembly of the Mitolobosomes and the Stability of Ribosomal Proteins

To investigate the changes in gene expression as a result of MRPP3 loss, we carried out transcriptome-wide RNA-seq and identified the most significant effects on heart and mitochondrial function (Figure 4A). We observed a significant decrease of cardiac markers such as cardiac actin and myosins as well as the ryanodine receptor 2 (Figure 2A; Table S2), consistent with the development of cardiomyopathy (Figures 1D and 1F). Transcripts encoding mitochondrial proteins, particularly those encoding the electron transport chain polypeptides, were most changed in response to MRPP3 loss and consequent mitochondrial dysfunction (Figure 5A; Table S3). Also, transcripts involved in translation were significantly affected, and those involved in mitochondrial translation were reduced, whereas transcripts involved in cytosolic translation were increased, likely as a way to stimulate mitochondrial biogenesis and compensate for dysfunctional energy metabolism (Figure 4B). We investigated the effects of MRPP3 loss on the mRNAs encoding mitochondrial ribosomal proteins and identified that their levels were largely reduced (Figure 5B), consistent with the loss of mitochondrial translation (Figure 5B).

Mitochondrial gene expression is predominantly regulated at the post-transcriptional level by RNA-binding proteins (Lu et al., 2013). Therefore, to investigate the consequences of impaired RNA 5’ end processing over 4, 8, and 11 weeks and validate our RNA-seq data, we analyzed the effects of MRPP3 loss on nuclear-encoded mitochondrial RNA-binding proteins.
by immunoblotting. We found that the increased rate of transcription in the Mrpp3 knockout mice (Figure 2A) was a consequence of increased POLRMT levels as a retrograde response of nuclear gene expression to impaired mitochondrial tRNA processing (Figure 5C). Similarly, we observed increased levels of the two mitochondrial transcription factors, TFAM and TFB2M, in Mrpp3 knockout mice at 8 and 11 weeks of age (Figure 5C). The levels of the other two RNase P proteins, MRPP1 and MRPP2, were also increased, as well as those of ELAC2 (Figure 5C), likely to compensate for the loss of MRPP3. Interestingly, the LRPPRC protein, required for the stability and translation of mt-mRNAs, is also increased at weeks 8 and 11, possibly as a consequence of the low levels of mature mt-mRNAs. The levels of POLRMT are consistently increased from 4 to 11 weeks in contrast to the other mitochondrial RNA-binding proteins that show an initial decrease in their levels at week 4 in the Mrpp3 knockout mice but increase their levels relative to controls by weeks 8 and 11 (Figure 5C). This suggests that the initial compensatory response to impaired RNA processing is to increase the rate of transcription, which has also been observed upon knockout of mitochondrial RNA-binding proteins that affect the stability of the ribosome such as TFB1M, MTERF3, and MTERF4 (Cámar et al., 2011; Metodiev et al., 2009; Park et al., 2007). This is in contrast to other proteins that regulate mitochondrial gene expression, such as LRPPRC, where the rate of transcription is not affected (Ruzzenente et al., 2012).

Since mitochondrial protein synthesis was significantly reduced in the Mrpp3 knockout mice, we investigated the effects of MPFRP3 loss on the steady-state levels of mitoribosomal proteins (Figure 5D). We found that the levels of some ribosomal proteins were increased in the Mrpp3 knockout mice, such as MRPS34 and MRPL12, whereas some were decreased such as MRPS35 and MRPL37 (Figure 5D), consistent with the differential expression and enrichment analyses of the mRNAs encoding mitoribosomal proteins found in our RNA-seq data (Figure 5A; Table S2). These findings indicate that impaired RNA processing affects mitoribosome stability, as we have observed before in cells (Sanchez et al., 2011).

To investigate the effects of impaired RNA processing on the assembly and stability of the ribosome, we resolved the mitoribosomal subunits and assembled ribosomes on 10%–30% sucrose gradients and examined them by immunoblotting against markers of the small and large ribosomal subunits (Figure 6A). We identify that loss of MRPP3 and hence tRNA 5′-end processing impairs 55S monosome formation (Figure 6A). Furthermore, we observed a shift of the small and large ribosomal subunit proteins toward the less dense sucrose fraction in the Mrpp3

Figure 5. Impaired RNA Processing Causes an Imbalance in Nuclear-Encoded Mitochondrial Proteins
Transcriptome-wide RNA-seq was carried out on total heart RNA from three control (L/L) and three knockout (L/L, cre) mice, and differential expression analyses were performed. (A and B) Significant differences (p < 0.05) in positively and negatively correlated transcripts relative to controls are based on specific gene ontologies are shown for cardiac markers and mitochondrial genes (B) and for oxidative phosphorylation, the translation machinery, and mitoribosomal and RNA-binding proteins (B). (C) The levels of nuclear-encoded mitochondrial RNA-binding proteins were measured by immunoblotting in heart mitochondria from 4-, 8-, and 11-week-old control (L/L) and knockout (L/L, cre) mice. (D) Mitoribosomal protein abundance was measured by immunoblotting in heart mitochondria from 4-, 8-, and 11-week-old L/L and L/L, cre mice.
Figure 6. Mitochondrial Ribosome Assembly Is Impaired as a Consequence of Impaired RNA Processing

(A) A continuous 10%–30% sucrose gradient was used to determine the distribution of the small and large ribosomal subunit and the monosome in the control and knockout mice. Mitochondrial ribosomal protein markers of the small (MRPS35) and large (MRPL37, MRPL44, MRPL11, and MRPL12) ribosomal subunits were detected by immunoblotting with specific antibodies. The input (10% of the total lysate) was used as a positive control. The data are representative of results from at least three independent biological experiments.

(B and C) The distribution of the mt-Co1 mRNA, 12S and 16S rRNAs (B), and 12S-16S rRNA junctions (C) in sucrose gradients were analyzed by qRT-PCR. The data are expressed as percentage of total RNA abundance and are representative of results from four independent biological experiments.
knockout mice compared to controls (Figure 6A). The shift in small ribosomal subunit proteins was dramatic, indicating that the assembly of these proteins on unprocessed 125S RNA is not possible or at least severely impaired. In contrast, the large ribosomal subunit proteins were shifted but still migrated in a large molecular weight complex, suggesting that a partial ribosomal subcomplex is formed. It is possible that an unprocessed rRNA transcript may be sufficient to stimulate the assembly of the early assembling ribosomal proteins that form a subcomplex but are unable to complete the assembly of the mature subunits. Interestingly, the MRPL37 and MRPL11 proteins, whose abundance was increased in the knockout (Figure 6D), and MRPL44, were found both in the subcomplex and in a rRNA free state in the less dense fractions. This indicates that some ribosomal proteins are stable when unassembled, while others are not. Whether free mitribosomal proteins might contribute to mitochondrial dysfunction is not known.

Next, we investigated the distribution of mitochondrial transcripts within the ribosomal fractions by qRT-PCR of each fraction (Figure 6B) to determine the distribution of the unprocessed transcripts within the sucrose gradient in the control mice. The distribution of the 12S and 16S rRNAs follows a similar pattern as that observed by immunoblotting against markers of the small and large ribosomal subunits, respectively (Figure 6B). However, we observe a redistribution of these transcripts in the Mrpp3 knockout mice to less dense fractions of the sucrose gradient (Figure 6B) that co-migrate with the unprocessed polycistronic transcripts found in the Mrpp3 knockout mice (Figure 6A).

Similarly, we observe that the mt-Co1 mRNA is present in the pool of free mt-mRNAs in the top fractions as well as with the ribosomal fractions in the control mice; however, it is re-distributed only to the earlier fractions in the Mrpp3 knockout mice because there are no correctly assembled mitribosomal subunits to recruit mt-mRNAs (Figure 6B). We find that the mt-rRNA 22S-28S is not present in its mature form, but instead it is also part of the large unprocessed transcript found in the Mrpp3 knockout mice (Figure S4C). We do not observe a polycistronic transcript in the control mice, indicating that the unprocessed transcript that co-migrates with the large subcomplex detected in the immunoblot Mrpp3 knockout mice (Figure 6A) is composed of ribosomal proteins that assemble on the unprocessed transcript containing the 12S and 16S rRNA. Rnase A treatment of mitochondrial lysates before sucrose gradient separation (Figure 6C) identified that the precursor transcript is lost upon Rnase treatment, as this transcript is not protected it is likely that it is not bound by a complete complement of mitribosomal proteins.

To identify the ribosomal proteins that are present in the subcomplex found in fraction 10 on the precursor transcript, we carried out proteomic analyses on fractions that contained the small (fraction 6) and large subunit (fraction 10) or the subcomplex (fraction 8 and 10) and the monosome (fractions 12 and 13) from both the control and Mrpp3 knockout mice (Figure 7). We confirmed that the monosome is not formed in the Mrpp3 knockout mice compared to the control mice where we find the highest abundance of both small and large ribosomal proteins in the monosome fractions (Figure 7). In control mice, the distribution of the small and large subunits is consistent with their migration in the gradient, where the small and large ribosomal subunit proteins are found in the fractions containing the monosome (fractions 12 and 13). In the Mrpp3 knockout mice, we find enrichment of large ribosomal proteins in fraction 10 indicating that these proteins assemble on the unprocessed transcript to form the subcomplex observed by immunoblotting (Figure 6A). Furthermore, as we detected an enrichment of the majority of large subunit proteins in fraction 10, this complex must be stalled close to the end of the large subunit assembly pathway. Taken together, we have shown that RNA processing is rate limiting for the assembly of the mitochondrial ribosomal subunits and the 5S monosome.

DISCUSSION

The compact organization of the mitochondrial genome in animals has eliminated the necessity for RNA splicing of non-coding inter- and intragenic sequences. Instead, the RNA genes separate the mtRNAs and RNAs in a way that permits the RNA endoribonuclease to cleave them from the large polycistronic transcripts produced by the mitochondrial RNA polymerase and release all individual mtRNAs (Culav et al., 1991). Although two enzymes involved in 5′ and 3′ RNA processing have been identified as the RNase P complex and ELAC2, the importance of RNA processing (Brzezniak et al., 2011; Holzmann et al., 2008; Sanchez et al., 2011) and the consequences of its loss have not been investigated in vivo. Here, we investigated the loss of RNA 5′ and 3′ processing upon conditional heart- and muscle-specific knockout of the Mrpp3 gene, encoding the endonuclease component of the mitochondrial RNase P complex. We show that RNA processing is essential for early embryo development, and its knockout in the heart and skeletal muscle leads to very early onset cardiomyopathy that causes the mice to die by 11 weeks. Clearly, RNA processing is essential for life since severely impaired biogenesis of mitochondrial ribosomes and also OXPHOS complexes leads to mitochondrial dysfunction. Our data further corroborate in vitro findings (Holzmann et al., 2008) and demonstrate that there are no alternative RNase P enzymes to compensate for the loss of the 5′ endonuclease activity of Mrp3. Impaired RNA processing has been shown recently due to mutation of the MrP3P, Mrp3, and ELAC2 genes in patients with mitochondrial disease and cardiomyopathy (Haack et al., 2013; Meldev et al., 2016; Viardo and Rosomarini, 2015). Interestingly, in all three cases although RNA processing was shown to be impaired there was not a significant effect on the mature mtRNA levels, which is particularly curious. These effects may be a consequence of growth of patient cells in vitro in highly anaerobic conditions or the tissue-specific response of fibroblasts compared to heart and muscle cells and tissues, both of which do not impose the same bioenergetic requirements as in vivo systems. The requirement for RNA processing for life results in the observation of very mild hypomorphic mutations in these genes, masking the deleterious effects on RNA metabolism and making these effects difficult to detect in these patients. Loss of Mrpp3 in vivo abolishes RNA 5′ end processing, trapping the individual mtmRNAs in precursor transcripts and thereby...
Figure 7. Differential Distribution of Mitochondrial Ribosomal Proteins upon Loss of MRPP3

Mass spectrometry was used to analyze the distribution of the mitochondrial ribosomal proteins in sucrose gradient fractions that contained the small (fraction 8) and large ribosomal subunit (fraction 10) and the monosome (fractions 12 and 13) in the control (L/L) mice and compared to the same fractions in the Mrpp3 knockout mice (L/L, cre), both at 11 weeks. The analyses were performed on fractions from three separate biological experiments from three control and three knockout mice (n = 3). The results of the ANOVA analysis were adjusted for multiple testing using the Benjamini and Hochberg procedure. Only proteins with adjusted p value of less than 0.01 (1% FDR) were termed significant and marked with asterisk (*).
leads to profound reduction of most mature mtRNAs. The posi-
tion of specific rRNAs or mtRNAs within or adjacent to clusters of
RNA genes suggested that perhaps the activity of ELAC2 could
rescue some of these transcripts by cleaving neighboring 5' ends
as we observed for mt-NiS. However, this was not the case for
the rRNAs where we found large unprocessed transcripts that
precluded the correct assembly of the small and large ribo-
subunits. Although the levels of ELAC2 were moderately
increased when MRPP3 was lost, they were not sufficient to
rescue the processing of specific 3' end sites that are also 5' end
sites of flanking RNAs. We conclude that RNA 5' end process-
ing is required as the initial step of RNA processing preceding 3' RNA
cleavage. It is not clear why 5' processing is required in
most cases for 3' processing to proceed, perhaps because the
MRPP3 enzyme is more agnostic to structure, while RNase Z
activity is inhibited by long 5' leaders both in rRNA processing
(Rashidinotem et al., 1999) and cleavage of tRNA-like RNAs (Wulz
et al., 2008).

Interestingly, we identify two different sites that are not
changed by the loss of MRPP3, one at the 5' end of mt-Cyt b and the other at the 5' end of mt-CoA. A gene trap mutant of
the Ptdcd gene in mice has shown that the 5' end of mt-Cyt b
is not processed in the absence of Ptdcd, suggesting that
this protein may have a role in its cleavage (Ku et al., 2008).
Ptdcd belongs to the PPR family of mitochondrial RNA-binding
proteins in animals; like MRPP3, however besides its PPRs it
lacks a canonical endoribonuclease domain, suggesting that
there may be a new fold within its tertiary structure that can carry
out this activity or that it recruits an unidentified endoribonu-
lease that can cleave this site. PPR proteins across different
kingdoms can interact and recruit other RNA-binding proteins
for their activity, such as the RNase P complex with the
MRPP1 and MRPP2 (Holzman et al., 2008; Li et al., 2015; Rein-
hard et al., 2015) and the LRRPRC/SLIP complex (Lagouge
et al., 2015; Ruzzenni et al., 2012; Sasaean et al., 2013) in
animals or PPR proteins and MORF proteins in plants (Takenaka
et al., 2015).

We annotated all the known MRPP3 cleavage sites in our
work, accounting for all in vivo processing events carried out by
the RNase P complex. Furthermore, we observed a dramatic
increase in 5' ends within internal regions of mtRNAs upon loss of
MRPP3 and 5' and processing. These are likely products of RNA
degradation, and the general loss of these transcripts at their 3' ends
suggests that degradation is in the 3' to 5' direction. It re-
mains to be determined which mitochondrial endonuclease is
responsible for this removal of unprocessed RNAs, and a recent
report suggests that LACTB2 may act as a mitochondrial endo-
nuclease (Lary et al., 2016).

Mitochondrial gene expression relies solely on the post-
transcriptional import of nuclear-encoded RNA-binding proteins
and other protein factors (Hilberg and Larson, 2014; Rashcham
et al., 2012). Knockout of mitochondrial RNA-binding proteins
has been shown to increase the levels of other nuclear-encoded
factors in an effort to compensate for the changes in gene
expression (Camar et al., 2011; Meltdav et al., 2009, 2014;
Park et al., 2007; Ruzzenni et al., 2012). We observed similar
increases in RNA-binding protein levels upon MRPP3 loss, but
most strikingly we observed a dramatic increase in POLRMT
levels and mitochondrial transcription. Furthermore, the varying
effects on the mitochondrial ribosomal proteins suggested that
the expression and stability of these proteins is differentially
regulated. Nevertheless the dramatic changes that result from
impaired RNA 5' end cleavage on transcription and protein
synthesis indicate that RNA processing couples these two
processes.

The cleavage of the 5' ends of rRNAs is essential for ribosome
assembly, which, in turn, is clearly a process that links transcrip-
tion and translation. In addition, it indicates that it is the initial
step of RNA processing, since ELAC2 and any other endoribonu-
clases cannot overcome the lack of 5' processing. The small
subunit is dependent on the mature 12S rRNA for assembly
because its assembly is severely impaired in the Mmp3
knockout mice, and consequently the 55S ribosome cannot
form. This may indicate that correctly processed 5' and 3' ends
are required to initiate the early assembly events in small subunit
formation. This mirrors bacterial ribosome assembly where the 5' and 3' ends of the 16S and 23S rRNAs must juxtapose for effi-
cient ribosome biogenesis (Young and Slizt, 1978). However,
it is interesting to note that a large subunit subcomplex can
form using an unprocessed 16S rRNA and contains a number
of large ribosomal subunit proteins, suggesting that some ribo-
osomal proteins can assemble without the requirement for 16S
rRNA processing. This is highly unusual, although the mitochon-
drial ribosome is unique in that it contains a much higher protein
content compared to rRNA, as well as additional ribosomal pro-
teins that are not found in cytoplasmic or prokaryotic ribosomes
(Rackham and Filipovska, 2014b; Suzuki et al., 2001a, 2001b),
suggesting that some of these could have unusual characteris-
tics that may allow them to initiate binding or their stabilization
by association with the unprocessed RNA. The MRPP3
knockout mouse has been useful to identify the ribosomal proteins
involved in the initial assembly of the large mitochondrial ribo-
somal subunit, as the rRNAs become available following tran-
scription. Taken together, we have shown that RNA processing
is rate limiting for the assembly of the mitochondrial ribosomal
subunits that occurs co-transcriptionally.

Here, we have shown the dire consequences of impaired mito-
chondrial RNA processing and that this is an essential step in
mitochondrial RNA metabolism required for biogenesis and
energy production. The next goal is to elucidate the precise
mode of ribosome assembly that is clearly dependent on the
generation of mature rRNAs. This will undoubtedly be a long
and interesting journey with additional complexities as this tiny
but mighty transcriptome has revealed so far.

EXPERIMENTAL PROCEDURES

Animals and Housing

Mmp3 heterozygous mice on a C57BL/6J background were generated by Ta-
cori. The puroycin cassette was removed by mating Mmp3+/− mice with transgenic mice ubiquitously expressing Cre recombinase. The resulting
Mmp3+/− mice were mated with mice ubiquitously expressing Cre recombi-
nase to generate heterozygous knockout Mmp3+/− mice that were bred with
each other to identify that the homozygous loss of Mmp3 was embryonic le-
thal. Heart- and skeletal muscle-specific knockout mice were generated by
crossing Mmp3+/− mice with transgenic mice expressing Cre under
the control of the muscle creatinin kinase promoter (Crem-cro). Double
heterozygous mice (B6g129C57Bl6J) were mated with Mippo3/3 mice to generate heart-specific knockout (B6g129C57Bl6J)−/− and control mice (B6g129C57Bl6J). Mice were housed in standard cages (45 cm x 29 cm x 12 cm) under a 12-hr light/dark schedule (lights on 7 a.m. to 7 p.m.) in controlled environmental conditions of 22 ± 1°C and 55 ± 10% relative humidity and fed a normal chow diet (Pepi and Mouse Chews, Specialty Foods) and water were provided ad libitum. The study was approved by the Animal Ethics Committee of the University of Western Australia (UWA) and performed in accordance with the NHMRC Australian code for the care and use of animals for scientific purposes, 8th Edition 2013 and Federation for Laboratory Animal Science Associations (FLASA), Germany.

Mitochondrial Isolation
Mitochondria were isolated from homogenized hearts and isolated by differential centrifugation as described previously (Lagouge et al., 2015; Mourier et al., 2014), with some modifications as described in the Supplemental information.

Sucrose Gradient Fractionation
Sucrose gradient fractionation was carried out as described previously (Lagouge et al., 2015; Melot et al., 2008; Ruzzoer et al., 2012) with some modifications as described in the Supplemental information.

Immunoblotting
Specific proteins were detected using rabbit polyclonal antibodies against POLDI (Lagouge et al., 2015; Metodiev et al., 2009; Ruzzenente et al., 2012) with specific proteins were detected using rabbit polyclonal antibodies against COXII, COXIV, and complex V subunit a (Abcam, diluted 1:1,000), MRPP1, MRPP2, ELAC2, MRPL44, MRPL37 (1:1,000), MRPL12 (1:500), MRPS16 (1:1,000), MRPS16 (Proteinscore), and MRPS16 (Sigma, diluted 1:1,000) and mouse monoclonal antibodies against a-actin, voltage-dependent anion channel (VDAC), NDUF, complex II, complex III, COX, COXIV, and complex V subunit a (Abcam, diluted 1:1,000), in Odyssey Blocking Buffer (LI-COR Biosciences). IR Dye 800CW goat anti-rabbit infrared-labeled antibody (LI-COR Biosciences) secondary antibodies were used, and the immunoblots were visualized using an Odyssey Infrared Imaging System (LI-COR Biosciences).

Mass Spectrometry: Data Analyses
Protein and peptide identification was performed using Speaks (Ma et al., 2005) v7.5. Raw data were searched against the canonical and synonym sequence database of the mouse reference proteome (protein ID UP0000000969, June 2016) from Uniprot. For data analysis, cytosine carbondimethylation was set as a “fixed modification” and methionine oxidation and protein N-terminal acetylation as “variable modification.” The digestion parameters were set to trypsin allowing for cleavage after lysine and arginine when not followed by proline with up to two missed cleavages. Database searching was performed using mass error tolerance of 1 ppm (parent) and 0.5 Da (fragments). Up to two variable modifications per peptide were allowed, and false discovery rate (FDR) estimation was enabled. Upon database searching, the peptide identification FDR was set at 1%; default criteria were used for protein identification. Peaks database search results were exported and used further for peptide and protein quantification. Label-free quantification was performed using Skyline (MacLean et al., 2010) using the MS1 filtering approach (Schilling et al., 2010). For the analyses, the peptides pep.xml Speaks output file and a custom fasta database containing the sequences of all identified mitoribosomal proteins were imported into Skyline. For the fasta database import, “Trypsin (PIR) PT” was set as a protease allowing for up to two missed cleavages. Regarding the full scan settings, the precursor charges were set to two, three, and four. MS1 filtering was performed using three peaks (M, M-1 and M-2). “Choplog” was selected as a mass analyzer with resolving power of 60,000 at 200 m/z. Only scans within 5 min of MS/MS Ions were considered. After data import, the extracted ion chromatograms of all peptides mapping to the selected mitoribosomal proteins were manually inspected. In some cases, the peak integration was adjusted by manual inspection. No peak was set as a peak. Data integration was performed over a retention time region, which is in agreement with the repeated analysis of the same and the remaining fractions. Upon manual evaluation, the peptide intensity data were exported and assembled into protein intensity data using an in-house script. Similarly to other label-free approaches (Nishizawa and Nishizawa, 2016), protein intensities in all four fractions were transformed into proportions of the total protein intensity in all four fractions, in order to derive the relative distribution of a protein in the gradient fractions. The proportion values in each fraction were used to perform a two-way ANOVA analysis using two factors, “Genotype,” “Fraction,” and their interaction. The ANOVA was performed in R (R Development Core Team, 2013, http://www.r-project.org/) using the analysis of variance function. The raw (National Health and Medical Research Council [NHMRC] Australian code for the care and use of animals for scientific purposes, 8th Edition 2013) and Federation for Laboratory Animal Science Associations (FLASA), Germany.

RNA Isolation, Northern Blotting, and qRT-PCR
RNA was isolated from total hearts or heart mitochondria using the miRNeasy Mini kit (Qiagen) incorporating an on-column mRNA-free Dnase digestion to remove all DNA. RNA-S (1 μg) was resuspended in 12% agarose formaldehyde gels and then transferred to 0.45 μm Nuclepore (introduces membrane (GE Healthcare Life Sciences)) and hybridized with biotinylated oligonucleotide probe specific to mouse mitoribosomal RNA (rRNA, and rRNA-R (Richman et al., 2009). Hybridizations were carried out overnight at 50°C in 5x sodium sodium citrate (SSC), 200 ng/mL H2O2, 7 SDS, and 100 μg/mL heparin, followed by washing. The signal was detected using either streptavidin-linked horseradish peroxidase or streptavidin-linked infrared-labeled antibody (diluted 1:1,000 in 3x SSC, 5% SDS, 25 mM NaHPO4, pH 7.0) by enhanced chemiluminescence (GE Healthcare Life Sciences) or an Odyssey infrared imaging system (LI-COR biosciences), respectively. 

RNA was prepared using the Quantitect Reverse Transcription Kit (Qiagen) and used as a template in the subsequent PCR that was performed using a Corbett RotorGene 3000 using SensiMix SYBR mix (BioLine) and normalized to 18S rRNA.

RNA-Seq, PARE, and Alignments
RNA sequencing was performed on total RNA or mitochondrial RNA from three control and three Mippo3 knockout mice using the Illumina HiSeq platform, according to the Illumina Tru-Seq protocol or PARE protocol (German et al., 2005; Meng et al., 2011; Richman and Nguyen, 2014). Library preparation, sequencing, and bioinformatics analyses are described in detail in the Supplemental information.

Transcription and Translation Assays
In organello transcription and translation assays were carried out in isolated heart mitochondria as described previously (Lagouge et al., 2015).

Blue Nase PAGE
BN-PAGE was carried out using isolated mitochondria from hearts as described previously (Mourier et al., 2014). BN-PAGE gels were analyzed by in-gel activity assays or by transferring to polyvinylidene fluoride (PVDF) and immunoblotting against the respiratory complexes.

Respiratory Chain Function and Complex Activity
The mitochondrial oxygen consumption flux was measured with an Oxygraph-2k (Oroboros Instruments), as previously described (Mourier et al., 2014; Lagouge et al., 2015).

Hotology
Mouse hearts were fixed with 10% neutral buffered formalin for 24 hr and stored in PBS or 70% ethanol. Tissues were embedded in paraffin, sectioned using a microtome, and transferred to positively charged slides. Slides were heated for 2 hr at 30°C and treated with xylene, xylen and ethanol (1:1), and decreasing concentrations of xylene (100%, 80%, 60%, 40%) before they were washed in distilled H2O (X3). The H&E staining was performed as described before (Richman et al., 2015). Skeletal muscle was frozen in OCT, and the tissue slices were analyzed by the tissue microarray by performing an in-situ hybridization using NADH and cytochrome c oxidase as substrates. Coverslips were attached using DPX mounting media and images were acquired using a Nikon Ti Eclipse inverted microscope using a Nikon 20x objective.
ACCESSION NUMBERS
The accession number for the RNA-seq and PAR-SE data reported in this paper is Gene Expression Omnibus (GEO): GSE83471.

SUPPLEMENTAL INFORMATION
Supplemental Information includes Supplemental Experimental Procedures, four figure, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.cepel.2016.07.001.

AUTHOR CONTRIBUTIONS
D.R., N.-G.L., and A.F. conceived the project and designed the experiments. D.R., J.D.B., S.M., S.S., I.K., I.A., J.E., A.-M.J.S., T.R.R., J.B.S., A.M., D.M., and A.F. conducted and analyzed the experiments. D.R. and A.F. wrote the manuscript, and the other authors approved the manuscript.

AKNOWLEDGMENTS
This project was supported by fellowships and project grants from the National Health and Medical Research Council (APP1035443, APP1040677, APP1041859, APP1024480, APP1002630, and APP1003976) to A.F. and C.R.), Australian Research Council (to A.F. and D.R.) the Cancer Council of Western Australia (to D.R.). Swedish Research Council (2015-04181 to N.-G.L.), Knut and Alice Wallenberg Foundation to N.-G.L.), and the Alexander von Humboldt Foundation (to A.F., D.R., and N.-G.L). S.M. was supported by an EMBO Short-Term Scholarship. We thank Larisa Schmitz, Julia Preuschoff, and Malte Stentenbach for technical assistance. Dr. Min-Jung for advice on the embryo isolation and Dr. Jorge Boucas for help with the bioinformatics pipelines for analyses of the transcriptome-wide differential expression.

Received: May 13, 2016
Revised: June 20, 2016
Accepted: July 13, 2016
Published: August 4, 2016

REFERENCES


R Development Core Team (2013). The R project for statistical computing (R Foundation for Statistical Computing).


Supplemental Information

Hierarchical RNA Processing Is Required for Mitochondrial Ribosome Assembly

Figure S1. Related to Figure 1 | The effects of MRPP3 loss on skeletal muscle.

(A) Mrpp3 transcript levels were determined in isolated total skeletal muscle RNA from control (L/L) and knockout (L/L, cre) mice by qRT-PCR (n=5). The data are expressed relative to control mice and normalized to 18S rRNA.

(B) Hematoxylin and eosin (H&E), NADH and COX staining of skeletal muscle from control (L/L) and knockout (L/L, cre) mice.

(C) Levels of Atf4, Atf5, and Chop mRNAs were determined in isolated total skeletal muscle RNA from control (L/L) and knockout (L/L, cre) mice by qRT-PCR (n=5). The data are expressed relative to control mice and normalized to 18S rRNA.
Figure S2. Related to Figure 2 and Figure 3 | Loss of the MRPP3 impairs mitochondrial protein synthesis and increases transcription

(A) *In organello* transcription was measured in heart mitochondria in the presence of $^{32}$P labeled UTP and the radiolabelled RNA was isolated and resolved on a 0.9% formaldehyde gel that was visualized by autoradiography. Porin was used as a loading control. The data shown are a darker exposure of the autoradiograph shown in Figure 2A.

(B) Northern blot detection of *mt-Co1*, *mt-tRNA* and *12S* rRNA on the cold blot shown in (A).

(C) Immunoblotting against all five electron transport chain complexes following BN-PAGE.
(D) Mitochondrial RNA junctions were measured in total skeletal muscle RNA from control (L/L) and knockout (L/L, cre) mice by qRT-PCR and normalized to 18S rRNA. Error bars indicate SEM; **p<0.01; ***p<0.001, Student’s t-test.
Figure S3. Related to Figure 4 | Transcriptome-wide map of mitochondrial RNA 5' ends.
(A) A transcriptome-wide map of the mean normalized 5′ end abundance (log_{10} mean normalized count; scale: 100 – 10,000,000) from three control (1, 5; L/L) and three knockout (2, 4; L/L, cre) mice in heavy (1, 2) and light (4, 5) strands. The mitochondrial genome is displayed in the central track (3), with the nucleotide position in base pairs displayed across the exterior; rRNAs are displayed in orange, mRNAs in green, tRNAs in blue and the non-coding region (NCR) in grey.

(B) Genome browser view of the 5′ end abundance from three control (L/L) and three knockout (L/L, cre) mice (top; mean normalized count), and relative changes in mean 5′ end abundance (middle; log_{2} fold change [KOmean/Ctrlmean]) and mean RNA-Seq coverage (bottom; log_{2} fold change [KOmean/Ctrlmean]) showing the 5′ cleavage sites of multiple tRNAs by MRPP3 and the downstream effect on the 3′ end cleavage in the absence of MRPP3. Regions of interests are shown in grey (for 5′ ends) and green (for 3′ ends) boxes. The browser view of the region containing tRNA^{Ile} and tRNA^{Met} on the heavy strand and tRNA^{Gln} on the light strand indicates decreased cleavage at the 5′ ends of tRNA^{Gln}, tRNA^{Met} and mt-Nd2 mRNA.

(C) Genome browser view of the 5′ end abundance from three control (L/L) and three knockout (L/L, cre) mice (top; mean normalized count), and relative changes in mean 5′ end abundance (middle; log_{2}[KOmean/Ctrlmean]), and mean RNA-Seq coverage (bottom; log_{2}[KOmean/Ctrlmean]) showing the cluster of five tRNAs (WANCY) encoded by both the heavy and light strand of the mitochondrial genome. Regions of interests are shown in grey (for 5′ ends) and green (for 3′ ends) boxes.
Figure S4. Related to Figure 4 and Figure 6 | Loss of MRPP3 causes non-specific mt-RNA degradation and accumulation of precursor RNAs.

(A) Genome browser view of the 5′ end abundance from three control (L/L) and three knockout (L/L, cre) mice (top; mean normalized count), and relative changes in mean 5′ end abundance (middle; log2 fold change [KOmean/Ctrlmean]) and mean RNA-Seq coverage (bottom; log2 fold change [KOmean/Ctrlmean]) showing the 5′ cleavage site of “mirror” mt-tRNA<sup>Tyr</sup> that acts as a 5′ end for the mt-Co1 mRNA in the absence of MRPP3 (shown in a red box).
(B) A logo of the degradation specificity of mt-RNAs in the absence of 5′ end
processing. A side logo is shown of the nucleotide distribution of the heavy and light
strands of the mouse mitochondrial genome.

(C) A continuous 10-30% sucrose gradient was used to determine the distribution of
the small and large ribosomal subunit and the monosome in the control (L/L) and
knockout (L/L, cre) mice. The distribution of the mt-tRNA<sub>Leu(UUR)</sub> mRNA were
analyzed by qRT-PCR. The data are expressed as % of total RNA abundance and they
are representative of results from at least four independent biological experiments.
Supplemental experimental procedures

Mitochondrial isolation
Hearts were cut and washed three times with ice cold PBS, and once with mitochondrial isolation buffer (MIB) containing 310 mM sucrose, 10 mM Tris-HCl and 0.05 % BSA (w/v) by centrifugation at 4,500 g for 1 min at 4°C. Heart pieces were homogenized in 5 ml of fresh MIB using a Potter S pestle (Sartorius). The homogenate was centrifuged at 1,000 g for 10 min at 4°C and the supernatant was centrifuged at 4,500 g for 15 min at 4°C to isolate mitochondria. Crude mitochondrial pellets were suspended in MIB supplemented with 1x Complete EDTA-free protease inhibitor cocktail (Roche). Protein concentration was determined by the Bradford or BCA method using BSA as a standard.

Sucrose gradient fractionation
Isolated mitochondria at 10 mg/ml were lysed in 10 mM Tris-HCl, pH 7.5, 260 mM sucrose, 100 mM KCl, 20 mM MgCl₂ and 2 % digitonin in the presence of 40 U/ml RNase inhibitor and 1x Complete EDTA-free Protease inhibitor cocktail for 20 min. The lysate was centrifuged at 9,200 g for 45 min at 4°C, the clarified lysate was loaded on a continuous 10-30% sucrose gradient (in 10mM Tris-HCl, pH 7.5, 100 mM KCl, 20 mM MgCl₂ in the presence of RNAse and protease inhibitors) and centrifuged at 71,000 g in an Optima Beckman Coulter preparative ultracentrifuge. Fractions were collected and one third of each fraction was precipitated with 0.02% sodium deoxycholate and 12% trichloroacetic acid (final concentration), washed twice with acetone, and the entire precipitate was resolved by SDS-PAGE. Protein markers of the mitochondrial ribosomal subunits were detected by immunoblotting, as described in the immunoblotting section.
Mass spectrometry based proteomic analysis of sucrose density gradient fractions

The remaining two thirds of each fraction obtained from the sucrose density gradients were subjected to chloroform/methanol precipitation (Wessel and Flugge, 1984) with some modifications. Two thirds of the gradient fractions were mixed with 1000 µl methanol. After addition of 400 µl chloroform and 200 µl water, samples were mixed and centrifuged at 4°C for 5 min at 20,817 g. The aqueous phase was discarded and the remaining interphase and lower phase were mixed with 1.5 ml of fresh methanol and centrifuged again. The supernatant was discarded and the protein pellets were dried for five to ten minutes at room temperature. The protein pellets were resuspended in 10 µl lysis buffer (6 M guanidinium chloride, 10 mM Tris(2-carboxyethyl)phosphine hydrochloride, 40 mM chloroacetamide and 100 mM Tris-HCl) as previously described by (Kulak et al., 2014) with some modifications. After lysis, samples were diluted ten times with 20 mM Tris pH 8.3 and digested overnight at 37°C with 300 ng of Trypsin Gold. Peptides were desalted using home-made StageTips (Empore Octadecyl C18, 3M; (Rappsilber et al., 2003; 2006)) and eluted with 80 µl of 60% acetonitrile/0.1% formic acid buffer. Samples were dried with a vacuum concentrator plus (Eppendorf) and resuspended in 4 µl of 0.1% formic acid for mass spectrometry.

For mass spectrometric analysis, peptides were separated online on a 25 cm 75 µm ID PicoFrit analytical column (New Objective) packed with 1.9 µm ReproSil-Pur media (Dr. Maisch) using an EASY-nLC 1000 (Thermo Fisher Scientific). The column was maintained at 50°C. Buffer A and B were 0.1% formic acid in water and 0.1% formic acid in acetonitrile respectively. Peptides were separated on a segmented gradient.
from 5% to 20% buffer B for 40 min and from 20% to 40% buffer B for 20 min at 200 nl/min. Eluting peptides were analyzed on an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific). The instrument was operated in a data dependent, “Top Speed” mode, with cycle times of 2 s. The “Universal Method” template was used with some modifications. Peptide precursor mass to charge ratio (m/z) measurements (MS1) were carried out at 60000 resolution in the 300 to 1500 m/z range. The MS1 AGC target was set to 1e6 and maximum injection time to 300 ms. Precursor priority was set to “Most intense” and precursors with charge state 2 to 7 only were selected for HCD fragmentation. Fragmentation was carried out using 27% collision energy. The m/z of the peptide fragments were measured in the ion trap using an AGC target of 2e3 and 300 ms maximum injection time. The option “Inject ions for All Available Parallelizable Time” was enabled. This option varies the maximum injection time and ion targets on the fly based on the available parallelizable time.

RNA-Seq, PARE and alignments

RNA sequencing was performed on total RNA from three control and three Mrpp3 knockout mice by the Cologne Genomics Centre (Cologne, Germany) on the Illumina HiSeq platform, according to the Illumina Tru-Seq protocol. We used random hexamer primers for cDNA library generation and carried out cytoplasmic rRNA depletion using the Ribo-Zero rRNA removal kit. Quality and adapter trimming was performed with Flexbar v.2.5. Reads were aligned to the mouse genome (GRCm38) using HISAT2 v2.0.1-beta (Kim et al., 2015). Guided transcriptome assembly (GRCm38.81) was done using StringTie v.1.1.0. Cufflinks v.2.2.1 was then used for analysis of differential expression (Trapnell et al., 2010). Briefly, cuffmerge was used
to merge all assemblies with the reference assembly as template, and cuffquant followed by cuffdiff (--dispersion-method per-condition) for final isoform quantification and differential analysis. The cummerbund package was used to process the data (Goff et al., 2012). The GSEA software was used with the pre-ranked gene list option, which requires generation of two files (.gmt and .rnk formats) (Mootha et al., 2003). The gene matrix transposed [.gmt] file format was produced based on specific GO terms and an in-house customized gene list for mitochondrial ribosomal proteins. GSEA rank file is a list of detected genes and a rank metric score such that top of the list are genes with the “strongest” up-regulation, at the bottom of the list are the genes with the “strongest” down-regulation and the genes not changing are in the middle. The metric used is the sign of the fold change multiplied by the -log10 (Q-value).

To analyse the mitochondrial transcriptome sequenced reads were initially aligned to the mouse genome (mm10) with HISAT2 v2.0.1-beta (Kim et al., 2015) (--very-sensitive --fr --rna-strandness RF --phred33 --no-mixed --no-discordant). The read IDs of all mitochondrial primary alignments were extracted with SAMtools v1.3 (Li et al., 2009), and used to filter the original alignments into nuclear and mitochondrial alignment files with Picard v1.23 (http://broadinstitute.github.io/picard/). The mitochondrial alignments were converted back to their original read sequence in fastq format with Picard, and realigned to the mitochondrial genome as before with the additional parameter --no-spliced-alignment. Primary mitochondrial alignments were extracted and separated according to their template strand-of-origin with SAMtools and converted to template strand-specific fragment BED files with BEDtools (Quinlan and Hall, 2010) and ad hoc scripts. Per-base depth of strand-specific fragment BED files were generated with BEDtools normalized by the total number of
properly mapped pairs per million across the whole genome (nuclear and mitochondrial) and converted to bedGraph format for visualization with Integrative Genomics Viewer v2.3.9 (Robinson et al., 2011; Thorvaldsdóttir et al., 2013). PARE library preparation and sequencing was performed on RNA isolated from heart mitochondria, as previously described (Mercer et al., 2011; Rackham and Filipovska, 2014) and sequencing was performed by Vertis Biotechnologie (Freising, Germany). The 5 bp adapter sequence was trimmed and paired-end reads were merged with CLC Genomics Workbench (Qiagen) if the read overlap exceeded 95% identity. Merged reads were aligned to the mouse genome (mm10) with Bowtie2 using default parameters (Ben Langmead and Salzberg, 2012), and the strand-specific depth of read coverage at the 5' terminal position of aligned reads was calculated with BEDTools (-d -5 -strand [+-]) (Quinlan and Hall, 2010). The read depth was normalized using the trimmed mean of M values (TMM) method (Robinson and Oshlack, 2010) from the edgeR package (Robinson et al., 2010) and subsequently converted to BedGraph format for visualization on the Integrative Genomics Viewer (IGV) (Robinson et al., 2011). Relative changes were calculated as log2 fold changes of the mean normalized 5' end coverage for three control (L/L) or knockout (L/L, cre) mice, plus a pseudocount of one. To generate the frequency plots, the mean per-base normalized coverage across a 31 bp window centered on each tRNA 5' end was calculated and plotted with R (R software R Core Team, 2013, http://www.R-project.org). For the sequence logo we used normalized coverage data between biological replicates, selected the top 50% of reads and searched the sequences that contain 10 nt before and after the 5' end for motifs. The sequence logo plots were produced using the web based online application Web Logo 3 (Crooks et al., 2004).
Transcription and translation assays

In organello transcription and translation assays were carried out in isolated heart mitochondria as described before (Lagouge et al., 2015). To measure transcription 800 µg of mitochondria were washed in 1 ml of transcription buffer (10 mM Tris pH 7.4, 25 mM sucrose, 75 mM sorbitol, 100 mM KCl, 10 mM K2HPO4, 50 µM EDTA, 5 mM MgCl2, 10 mM glutamate, 2.5 mM malate, 1 mg/ml BSA and 1 mM ADP) by centrifugation at 10,000 g, for 3 min at 4˚C. The mitochondrial pellet was suspended in 750 µl of transcription buffer supplemented with 40 µCi of 32P-UTP (PerkinElmer) and incubated for 40 min at 37˚C. After the incubation, mitochondria were washed once and suspended in 750 µl of fresh transcription buffer in the presence of 0.2 mM of cold UTP. A short chase was performed for 5 min at 37˚C to decrease the background and mitochondria were washed three times in 10 mM Tris pH 6.8, 0.15 mM MgCl2 and 10% glycerol. The mitochondrial pellet was suspended in 0.5 ml TRIzol (Invitrogen) for RNA extraction according to the manufacturer’s instructions. The isolated RNA was resolved on a 0.9% formaldehyde gel and transferred to a membrane that was visualized by autoradiography.

To measure mitochondrial translation, 500 µg mitochondria were incubated in 750 µl translation buffer (100 mM mannitol, 10 mM sodium succinate, 80 mM KCl, 5 mM MgCl2, 1 mM KPi, 25 mM HEPES pH 7.4, 5 mM ATP, 20 µM GTP, 6 mM creatine phosphate, 60 µg/ml creatine kinase and 60 µg/ml of all amino acids except methionine). Mitochondria were supplemented with 150 µCi of 35S methionine (PerkinElmer) for 60 min at 37˚C. After labelling, mitochondria were washed in translation buffer and suspended in RIPA lysis buffer. Protein concentration was measured and 50 µg of mitochondrial protein was resolved by SDS-PAGE and visualized by autoradiography.
Respiratory chain function and complex activity

The mitochondrial oxygen consumption flux was measured with an Oxygraph-2k (Oroboros Instruments), as previously described (Mourier et al., 2014) and (Lagouge et al., 2015), at 37°C by using 65 to 125 μg of crude mitochondria diluted in 2.1 ml of mitochondrial respiration buffer (120 mM sucrose, 50 mM KCl, 20 mM Tris-HCl, 4 mM KH₂PO₄, 2 mM MgCl₂, 1 mM EGTA, pH 7.2). The oxygen consumption rate was measured using either 10 mM pyruvate, 5 mM glutamate and 5 mM malate, or 10 mM succinate and 10 nM rotenone. Oxygen consumption was assessed in the phosphorylating state with 1 mM ADP or non-phosphorylating state by adding 2.5 μg/ml oligomycin. In the control mitochondria, the respiratory control ratio (RCR) values were >10 with pyruvate/glutamate/malate and >5 with succinate/rotenone. Respiration was uncoupled by successive addition of carbonyl cyanide m-chlorophenyl hydrazone (CCCP) up to 3 μM to reach maximal respiration. Mitochondria, 15–50 μg, were diluted in phosphate buffer (50 mM KH₂PO₄, pH 7.4) and spectrophotometric analyses of isolated respiratory chain complex activities were performed at 37 °C by using a Hitachi UV-3600 spectrophotometer. The citrate synthase activity was measured at 412 nm (ε = 13,600 M⁻¹ cm⁻¹) after the addition of 0.1 mM acetyl-CoA, 0.5 mM oxaloacetate and 0.1 mM 5,5'-dithiobis-2-nitrobenzoic acid (DTNB). The SDH activity was measured at 600 nm (ε = 21,000 M⁻¹ cm⁻¹) after the addition of 10 mM succinate, 35 μM dichlorophenolindophenol (DCPIP) and 1 mM KCN. The NADH dehydrogenase activity was determined at 340 nm (ε = 6, 220 M⁻¹ cm⁻¹) after addition of 0.25 mM NADH, 0.25 mM decylubiquinone and 1 mM KCN, controlling for rotenone sensitivity. The COX activity was measured by
standard N,N′,N,N′-tetramethylphenylene-1,4-diamine (TMPD) ascorbate assays, controlling for KCN sensitivity. All chemicals were obtained from Sigma-Aldrich.

References:


Rackham, O., and Filipovska, A. (2014). Analysis of the human mitochondrial...


Appendix C
Mitochondrial disorders are clinically and genetically diverse, with mutations in mitochondrial or nuclear genes able to cause defects in mitochondrial gene expression. Recently, mutations in several genes encoding factors involved in mt-RNA processing have been identified as causes of mitochondrial disease. Using whole-exome sequencing, we identified mutations in TRMT10C encoding the mitochondrial RNA polymerase 1 (MRPP1) in two unrelated individuals who presented at birth with lactic acidosis, hypotonia, feeding difficulties, and deafness. Both individuals died at 5 months after respiratory failure. MRPP1, along with MRPP2 and MRPP3, forms the mitochondrial ribonuclease P (mt-RNase P) complex that cleaves the 3′ ends of mt-tRNAs from polycistronic precursor transcripts. Additionally, a stable complex of MRPP1 and MRPP2 has m^1R^9 methyltransferase activity, which methylates mt-tRNAs at position 9 and is vital for folding mt-tRNAs into their correct tertiary structures. Analyses of fibroblasts from affected individuals harboring TRMT10C missense variants revealed decreased protein levels of MRPP1 and an increase in mt-RNA precursors indicative of impaired mt-RNA processing and defective mitochondrial protein synthesis. The pathogenicity of the detected variants—compound heterozygous c.542G>T (p.Arg181Leu) and c.814A>G (p.Val272Met) changes in subject 1 and a homozygous c.142G>T (p.Arg47Lys) variant in subject 2—was validated by the functional rescue of mt-RNA processing and mitochondrial protein synthesis defects after lentiviral transduction of wild-type TRMT10C. Our study suggests that these variants affect MRPP1 protein stability and mt-RNA processing without affecting m^1R^9 methyltransferase activity, identifying mutations in TRMT10C as a cause of mitochondrial disease and highlighting the importance of RNA processing for correct mitochondrial function.

Rajiv Bhakoo,1,* and Robert W. Taylor 2,*

Mitochondrial respiratory chain deficiencies lead to insufficient ATP production from oxidative phosphorylation (OXPHOS), resulting in a wide range of clinical presentations broadly recognized as “mitochondrial disorders.” Mitochondrial diseases are genetically diverse, owing to the necessary expression, co-ordination, and activity of factors encoded by both the mitochondrial and nuclear genomes for proper mitochondrial function. The 16.6 kb human mitochondrial DNA (mtDNA) encodes only 22 tRNAs, 2 rRNAs, and 13 polypeptides that are essential components of four of the five OXPHOS complexes.1 The remaining subunits of the respiratory complexes and all of the factors involved in mt-tRNA expression and maintenance are encoded by the nuclear genome, synthesized in the cytosol, and imported into mitochondria. Thus, there are a large number of potential genetic causes of mitochondrial disease, which has often complicated attempts to identify the correct genetic diagnosis. The advent of next generation sequencing has greatly expanded the list of known gene mutations associated with mitochondrial disease,2 including several genes involved in mitochondrial (mt)-tRNA processing and maturation.3-4 In mammalian mitochondria, all mt-tRNAs required for mitochondrial protein synthesis are encoded by the mitochondrial genome. Transcription of mtDNA produces long polycistronic transcripts that require further processing. Most mitochondrial open reading frames are separated by at least one mt-tRNA gene, with the structure of mt-tRNAs acting as “punctuation” marks in the transcripts prior to mt-tRNAs being excised at the 5′ end by the RNase P complex and at the 3′ end by the RNase Z enzyme. The mitochondrial RNase P is a ribonucleoprotein composed of three proteins, MRPP1, MRPP2, and MRPP3 (encoded by TRMT10C [MIM: 615423], HSDB7180 [MIM: 300256], and KIAA0391 [MIM: 609947]), respectively, whereas mitochondrial RNase Z is encoded by a single gene, ELAC2 (MIM: 603567).5-10 In addition to cleavage from the polycistronic transcripts, mt-tRNAs undergo many further modifications, with at least 30 different modified residues reported.5,11 One crucial modification is m^1R^9 methylation, which is involved in correct folding of mt-tRNAs and the formation of the mt-RNase P complex, ensuring proper mt-tRNA processing and mitochondrial protein synthesis.

Metodi D. Metodiev,1,10 Kylie Thompson,2,10 Charlotte L. Alston,3 Andrew A.M. Morris,4,4 Langping He,1 Zarah Assouline,5 Marlene Ro,1 Nadia Babu-Ruszson,1 Angela Pyle,5 Helen Griffin,6 Stefan Siria,7 Aleksandra Filipovska,7 Arnold Munnich,1,5 Patrick F. Chinnery,4,6,8 Robert McFarland,2 Agnes Rötig,1,6,7 and Robert W. Taylor.4,8

Mitochondrial disorders are clinically and genetically diverse, with mutations in mitochondrial or nuclear genes able to cause defects in mitochondrial gene expression. Recently, mutations in several genes encoding factors involved in mt-RNA processing have been identified as causes of mitochondrial disease. Using whole-exome sequencing, we identified mutations in TRMT10C encoding the mitochondrial RNA polymerase 1 (MRPP1) in two unrelated individuals who presented at birth with lactic acidosis, hypotonia, feeding difficulties, and deafness. Both individuals died at 5 months after respiratory failure. MRPP1, along with MRPP2 and MRPP3, forms the mitochondrial ribonuclease P (mt-RNase P) complex that cleaves the 3′ ends of mt-tRNAs from polycistronic precursor transcripts. Additionally, a stable complex of MRPP1 and MRPP2 has m^1R^9 methyltransferase activity, which methylates mt-tRNAs at position 9 and is vital for folding mt-tRNAs into their correct tertiary structures. Analyses of fibroblasts from affected individuals harboring TRMT10C missense variants revealed decreased protein levels of MRPP1 and an increase in mt-RNA precursors indicative of impaired mt-RNA processing and defective mitochondrial protein synthesis. The pathogenicity of the detected variants—compound heterozygous c.542G>T (p.Arg181Leu) and c.814A>G (p.Val272Met) changes in subject 1 and a homozygous c.142G>T (p.Arg47Lys) variant in subject 2—was validated by the functional rescue of mt-RNA processing and mitochondrial protein synthesis defects after lentiviral transduction of wild-type TRMT10C. Our study suggests that these variants affect MRPP1 protein stability and mt-RNA processing without affecting m^1R^9 methyltransferase activity, identifying mutations in TRMT10C as a cause of mitochondrial disease and highlighting the importance of RNA processing for correct mitochondrial function.

Metodi D. Metodiev,1,10 Kylie Thompson,2,10 Charlotte L. Alston,3 Andrew A.M. Morris,4,4 Langping He,1 Zarah Assouline,5 Marlene Ro,1 Nadia Babu-Ruszson,1 Angela Pyle,5 Helen Griffin,6 Stefan Siria,7 Aleksandra Filipovska,7 Arnold Munnich,1,5 Patrick F. Chinnery,4,6,8 Robert McFarland,2 Agnes Rötig,1,6,7 and Robert W. Taylor.4,8

Mitochondrial disorders are clinically and genetically diverse, with mutations in mitochondrial or nuclear genes able to cause defects in mitochondrial gene expression. Recently, mutations in several genes encoding factors involved in mt-RNA processing have been identified as causes of mitochondrial disease. Using whole-exome sequencing, we identified mutations in TRMT10C encoding the mitochondrial RNA polymerase 1 (MRPP1) in two unrelated individuals who presented at birth with lactic acidosis, hypotonia, feeding difficulties, and deafness. Both individuals died at 5 months after respiratory failure. MRPP1, along with MRPP2 and MRPP3, forms the mitochondrial ribonuclease P (mt-RNase P) complex that cleaves the 3′ ends of mt-tRNAs from polycistronic precursor transcripts. Additionally, a stable complex of MRPP1 and MRPP2 has m^1R^9 methyltransferase activity, which methylates mt-tRNAs at position 9 and is vital for folding mt-tRNAs into their correct tertiary structures. Analyses of fibroblasts from affected individuals harboring TRMT10C missense variants revealed decreased protein levels of MRPP1 and an increase in mt-RNA precursors indicative of impaired mt-RNA processing and defective mitochondrial protein synthesis. The pathogenicity of the detected variants—compound heterozygous c.542G>T (p.Arg181Leu) and c.814A>G (p.Val272Met) changes in subject 1 and a homozygous c.142G>T (p.Arg47Lys) variant in subject 2—was validated by the functional rescue of mt-RNA processing and mitochondrial protein synthesis defects after lentiviral transduction of wild-type TRMT10C. Our study suggests that these variants affect MRPP1 protein stability and mt-RNA processing without affecting m^1R^9 methyltransferase activity, identifying mutations in TRMT10C as a cause of mitochondrial disease and highlighting the importance of RNA processing for correct mitochondrial function.
which is probably important for the correct folding of most mt-tRNAs. In the case of mt-tRNA<sup>114</sup>, the unmodified in vitro transcript folds into an extended bulged hairpin,<sup>3,7</sup> but with the sole modification of N1 methyl-ation of adenine 9 (m<sup>1</sup>A9), the RNA adopts the classic cloverleaf structure.<sup>3,7</sup> It has been demonstrated that MRPP1 and MRPP2 can form a stable sub-complex that is active as a methyltransferase and is uniquely able to methylate both adenine and guanine nucleotides at position 9.<sup>7,11,14</sup> 19 of the 22 mt-tRNAs contain either A or G at position 9 and it is likely that all of these are subject to m<sup>1</sup>R9 methylation.<sup>7,11,14</sup>

We studied two children with suspected mitochondrial disease from unrelated families. Subject 1 (male) was the second child of healthy, non-consanguineous, white British parents, with a healthy older sister. He was born at term by a normal vaginal delivery after a normal pregnancy with a birth weight of 3.7 kg. He did not require resuscitation but was noted to be hypotonic and weak soon after birth. He fed poorly and gained weight slowly, partly due to gastro-esophageal reflux. Neonatal screening revealed significant hearing impairment subsequently confirmed to be sensorineural deafness. He was also found to have a raised plasma alanine transaminase of 439 U/L (normal range 15–60; ALT; normal range 7–40; gGT: 262 UI/L, normal range 0.7–2.1 mmol/L) with a high lactate to pyruvate ratio (170; normal range 30–50) was recorded at 1 month. She also had significantly impaired liver function (AST: 84 U/L, normal range 15–60; ALT: 52 U/L, normal range 7–40; gGT: 262 U/L, normal range 0.90–1.70 mmol/L) with a high lactate to pyruvate ratio (220; normal range 30–50; lactic acidosis, pH 7.4, high lactate to pyruvate ratio). Alcohol and plasma lactate were normal, plasma amino acid analysis was normal with the exception of a raised alanine concentration. He deteriorated rapidly, requiring tube feeding, and at the age of 4 months he suffered rhinovirus bronchiolitis requiring ventilatory support with CPAP. It proved impossible to wean him off ventilatory support with CPAP. It proved impossible to wean him off ventilatory support and he died at 5.5 months of age after withdrawal of this support.

Subject 2 (female) was the second child of unrelated par-
ts of Kurdish origin with a healthy older brother. She was born at term by a normal vaginal delivery after a normal pregnancy, weighing 3.05 kg. Hypotonia, poor sucking, and rapid weight loss, requiring tube feeding, and at the age of 4 months she suffered rhinovirus bronchiolitis requiring ventilatory support with CPAP. It proved impossible to wean him off ventilatory support with CPAP. It proved impossible to wean him off ventilatory support and she died at 5 months of age after withdrawal of this support.

Table 1. Biochemical and Clinical Findings in Individuals with TRMT10C Variants and Relevant Family History

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>TRMT10C Variants</th>
<th>OXPHOS Activities in Skeletal Muscle</th>
<th>Clinical Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>-</td>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>Subject 1 (male)</td>
<td>male</td>
<td>c.[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T]; [542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T];[542G&gt;T]</td>
<td>COX-deficient, ragged-red fibers</td>
<td>Lactate 6.2 mmol/L, pyruvate 0.145 mmol/L, lactate/pyruvate 4.28</td>
</tr>
</tbody>
</table>

**Other Clinical Features**

- Hypotonia, poor sucking, rapid weight loss, requiring tube feeding
- Lactate 6.2 mmol/L, pyruvate 0.145 mmol/L, lactate/pyruvate 4.28
- COX-deficient, ragged-red fibers

**Muscle Biopsy Findings**

- Age at Onset: 5 months
- Clinical Course: Death at 5 months

**Respiratory Chain Complex (RCC) Activities**

- Respiratory chain complex (RCC) activities are expressed as nmols NADH oxidised.min<sup>-1</sup>/unit citrate synthase<sup>103</sup> for complexes III and IV
- Respiratory chain complex (RCC) activities are expressed as nmol/min/mg protein

**Control Mean (Reference Range)**

- COX: 30–50
- Succinate dehydrogenase: 25–75
- Complex I: 30–45
- Complex II: 30–45
- Complex III: 30–45
- Complex IV: 30–45
- Citrate synthase: 70–170

**Clinical Features**

- Death at 5 months
- Hypotonia, poor sucking, rapid weight loss
- Lactate 6.2 mmol/L, pyruvate 0.145 mmol/L, lactate/pyruvate 4.28
- COX-deficient, ragged-red fibers

**Biochemical and Clinical Findings**

- Lactate concentration, and urine organic acid analysis were normal with the exception of a raised alanine concentration
- Blood spot acylcarnitine analysis was normal
- Plasma amino acid analysis was normal

**Further Investigations**

- Ultrasound of the kidneys was normal. Blood spot acylcarnitine analysis and plasma biotinidase were normal, plasma amino acid analysis was normal apart from a raised lactate concentration.
- He deteriorated rapidly, requiring tube feeding, and at the age of 4 months he suffered rhinovirus bronchiolitis requiring ventilatory support with CPAP. It proved impossible to wean him off ventilatory support and he died at 5.5 months of age after withdrawal of this support.

**References**

Brain MRI was undertaken at 2 months of age and was of poor quality but showed findings suggestive of bifrontal polymicrogyria. Acoustic oto-emissions were abnormal at 4 months, suggesting deafness. Unfortunately, auditory evoked potential was not done. She died at 5 months of age from respiratory distress.

Informed consent for diagnostic and research studies was obtained for both subjects in accordance with the Declaration of Helsinki protocols and approved by local Institutional Review Boards in Newcastle upon Tyne, UK, and Paris, France. Biochemical analysis of skeletal muscle samples identified clear mitochondrial enzyme defects involving both complex I and IV in both subjects, whereas complex III activity was normal in subject 1 but decreased in subject 2; both cases showed sparing of complex II activity (Table 1). Histopathological analysis of muscle from subject 1 revealed evidence of subsarcolemmal mitochondrial accumulation (ragged red fibers) and a mosaic pattern of cytochrome c oxidase (COX) deficiency (Figures 1A–1F).

Analysis of muscle DNA from both subjects excluded mtDNA abnormalities (mtDNA rearrangements and point mutations) and mtDNA copy number was shown to be normal in each case (data not shown). Whole-exome sequencing via previously described methodologies and bioinformatic filtering pipelines identified biallelic variants in TRMT10C (MIM: 615423; GenBank: NM_017819.3; also known as MRPP1 and RG9MTD1). Compound heterozygous c.542G>T (p.Arg181Leu) (ClinVar: SCV000264779.0) and c.814A>G (p.Thr272Ala) (ClinVar: SCV000264780.0) variants were identified in subject 1, whereas subject 2 was homozygous for the c.542G>T (p.Arg181Leu) variant, also identified in subject 1. Sanger sequencing was undertaken to validate the variants and confirm that these segregated with disease in each family (Figure 1G). Both identified TRMT10C variants are predicted to result in amino acid substitutions affecting evolutionarily conserved residues (Figure 1H) and are rare: the c.542G>T (p.Arg181Leu) variant is present on the ExAC database (10/120,324 alleles) and ESP6500 (1/11,824 alleles) whereas the c.814A>G (p.Thr272Ala) TRMT10C variant is absent on ExAC, ESP6500, and COSMIC. In silico predictions via SIFT, PolyPhen-2, and aGVGD suggest that the biophysico impact of the p.Arg181Leu substitution are relatively benign but that the proximity of the Arg181 residue to the TRM10-type domain (predicted from Met191) could hint at a crucial structural role that only an arginine residue can perform. In silico modeling of the TRMT10C variants via RaptorX and Phyre2 produced disparate predictions of MRPP1 protein structure and thus could not be used to indicate any potential misfolding as a consequence of the variants.

To investigate the functional effects of the identified TRMT10C variants, Western blot and mitochondrial protein synthesis assays were performed in fibroblast cell lines derived from both affected individuals and age-matched control subjects. These data showed that the steady-state expression of MRPP1 is reduced in fibroblasts derived from both subjects, with a more pronounced effect in subject 1 where a lower level of expression was observed.
levels of MRPP1 were markedly decreased in the subject cell lines, suggesting that the variants affect the stability of the protein (Figure 2A). On the other hand, levels of MRPP2 and MRPP3, the other two subunits of RNase P, were unchanged in fibroblasts from affected individuals (Figure 2A). The loss of MRPP1 protein level correlates with decreased steady-state levels of subunits of complex I (NDUFB8) and complex IV (COXI and COXII), in agreement with the multiple respiratory chain defects observed in muscle. We used blue native PAGE to determine the effects on the assembly and stability of the respiratory chain complexes and show a marked decrease of fully assembled complex I and complex IV, with a slight decrease in complex III levels (Figure 2C). The low steady-state levels of mtDNA-encoded proteins was due to impaired mitochondrial protein synthesis in subject fibroblasts as demonstrated by reduced incorporation of 35S-labeled methionine and cysteine (Figure 2D).

Because MRPP1 is known to be an essential subunit of the mitochondrial RNase P, we investigated whether fibroblasts from affected individuals showed evidence of impaired mitochondrial RNA processing. Northern blot analyses showed an increase in RNA precursor RNA19 when detected with either an MT-ND1 or MT-RNR2 probe (Figure 3A). However, the steady-state levels of the mature mRNAs were not significantly affected. No increase in precursors of MT-CO2 or MT-CO3 were observed, although the steady-state levels of mature MT-CO3 appeared to be slightly decreased in subject 2 (Figure 3A).

Processing of mt-tRNAs at the 3' end is carried out by ELAC2. Because both subjects had functional copies of ELAC2, it would be expected that the mt-tRNAs would be processed at the 3' end, but not at the 5' end, resulting in mt-tRNAs with an uncleaved mt-tRNA at the 3' end. The resolution of the Northern blots for mt-mRNAs was not sufficient to distinguish between mature mRNA and these pre-processed transcripts, thus, high-resolution Northern blot experiments were performed to assess the levels of mature mt-tRNAs (Figure 3B). Surprisingly, the steady-state levels of mt-tRNAs were not significantly altered in the affected individuals relative to controls, suggesting that the severe mitochondrial translation defect was not due to absence of cleaved mt-tRNAs. However, mt-tRNA19aa and mt-tRNA126aa appeared to have slightly
lower steady-state levels in subject fibroblasts relative to controls.

To further investigate precursor processing, we carried out RNA-seq analysis of mitochondrial RNA isolated from control and affected individuals. Differential analyses of mt-mRNA and mt-tRNA gene expression in TruSeq library datasets and small RNA library datasets, respectively, revealed no significant differences in mitochondrially encoded mt-mRNA, mt-rRNA, and mt-tRNA levels between the samples (not shown). However, when we investigated the changes in the abundance of reads across the entire mitochondrial transcriptome, we found an increase in the regions that span gene boundaries, where RNA processing is required to release individual mitochondrial RNAs from the precursor transcripts (Figure S1). Together, these data confirm an impairment of mt-tRNA processing efficiency without severe effects on mature mt-mRNA or mt-tRNA steady-state levels. It is possible that the cleavage of mt-tRNAs by mt-RNase P is less efficient in cells harboring TRMT10C/MBPP1 variants, but that the mt-tRNAs that are cleaved are very stable, thus retaining steady-state mt-tRNA levels to approximately wild-type levels.

All tRNAs undergo post-translational modification at numerous sites to promote their correct function. The mt-tRNAs are not exceptions and cleavage from the polycistronic mt-RNA transcripts is just one step in their maturation. In addition to their role in RNAse P activity, MRPP1 and MRPP2 act as an m^1R^9 methyltransferase. Methyl- ation of either G or A at position 9 is vital for the correct structure and function of mt-tRNAs. Thus, we sought to investigate the impact of the TRMT10C variants on m^1R^9 methyltransferase activity in subject fibroblasts. To this aim, we utilized two experimental approaches: (1) primer extension analysis of individual mt-tRNAs during which the reverse transcriptase-mediated extension of a radiolabelled primer is inhibited by the presence of methylated G9 in mt-tRNA leading to the accumulation of a single-base extension product (labeled m^1G) that is detectable in both control and in case subject RNA at similar levels. (2) Sequencing error rates at position 9 in mt-tRNA, mt-tRNA, and mt-tRNA determined by RNA-seq analysis of mitochondrial RNAs extracted from control and subjects’ fibroblasts. The relative abundance of individual nucleotides and indels generated by the presence of m^1G was analyzed as described previously.

Figure 3. TRMT10C Variants Lead to Altered RNA Processing but Do Not Affect Methylation at Position 9

(A) Northern blot analyses of mt-mRNA steady-state levels using radiolabelled probes specific for MT-ND1, MT-RNR2 (16 s rRNA), MT-CO2, and MT-CO3. Asterisk (*) denotes putative mRNA precursors.

(B) High-resolution Northern blot analyses of mt-tRNAs using radiolabelled oligonucleotide probes against mt-tRNAs with 5S rRNA as a loading control (performed as described previously).

(C) Primer extension analysis of m^1G in tRNA^Leu(UUR)^ using a radiolabeled primer (5’-TTATGCGATTACGGGCTCTGC-3’) annealing 1 base downstream of the modified residue. Primer and 3 μg RNA were denatured at 95°C for 1 min and cooled on ice. Primer extensions were carried out using AMV reverse transcriptase (Thermo Fisher Scientific) at 45°C for 1 hr and stopped by heating at 85°C for 15 min. After ethanol precipitation, the samples were analyzed by fractionation through a 12% polyacrylamide-agarose gel and autoradiography.

(D) Sequencing error rates at position 9 in mt-tRNA^Phe, mt-tRNA^Val, and mt-tRNA^Glu determined by RNA-seq analysis of mitochondrial RNAs extracted from control and subjects’ fibroblasts. The relative abundance of individual nucleotides and indels generated by the presence of m^1G was analyzed as described previously.

The American Journal of Human Genetics 98, 993–1000, May 5, 2016 997
We have demonstrated that tRNA splicing is affected in fibroblasts from affected individuals with mutant MRPP1, which is consistent with current knowledge of the function of MRPP1 as a component of mt-RNase P. However, to definitively prove that the mitochondrial OXPHOS defect is a consequence of the TRMT10C variants, lentiviral rescue experiments were performed to complement the respiratory phenotype expressed in cultured cells. Fibroblast cell lines from affected individuals were transfected with a lentiviral vector carrying a copy of the wild-type TRMT10C gene encoding MRPP1. The complemented cell lines displayed increased expression of MRPP1 protein level (Figure 4A), leading to a restoration of mitochondrial translation (Figures 4A and 4B) and normal levels of fully assembled respiratory chain complexes (Figure 4C). Furthermore, the level of mt-RNA precursors, elevated in subject fibroblasts, normalized after lentiviral transduction with wild-type TRMT10C (Figure 4D). These data verify the pathogenicity of the c.542G>T (p.Arg181Leu) and c.814A>G (p.Thr272Ala) TRMT10C variants, establishing these variants as causative of mitochondrial disease associated with multiple respiratory chain abnormalities.

Prenatal diagnosis has subsequently been offered to both families in subsequent pregnancies after the identification and validation of pathogenic TRMT10C variants. In the first family harboring compound heterozygous changes, both TRMT10C variants were identified in a later pregnancy after chorionic villus biopsy, leading to termination. In family 2 (homozygous variant), the fetus was heterozygous for the c.542G>T (p.Arg181Leu) variant and mitochondrial respiratory chain activities were normal in the chorionic villus biopsy sample (data not shown), supportive of an unaffected clinical status.

Of interest, we noted that both affected individuals showed decreased MRPP1 steady-state protein levels in fibroblasts, although the levels in subject 1 (compound heterozygote variants) had lower levels than subject 2 (homozygous variant), suggesting that the p.Thr272Ala mutant MRPP1 protein is less stable than the p.Arg181Leu mutant. However, despite having higher residual levels of MRPP1, cells from subject 2 exhibited a more severe impairment of mitochondrial protein synthesis resulting in lower steady-state levels of respiratory chain complexes I and IV, implying the p.Arg181Leu mutant protein is more stable but less active than the p.Thr272Ala mutant protein.

The impairment of mt-RNA processing observed in fibroblasts from the affected individuals was not as severe as we anticipated, with steady-state levels of mature mt-mRNAs and mt-tRNAs largely unaffected. However, these data fit well with reports of mutations in other proteins involved with RNA processing, given that mutations in ELAC2 and HSD17B10 have both been shown to lead to an accumulation of mt-RNA precursors without effects on the levels of mature mt-mRNA and mt-tRNAs. Furthermore, it has been shown that loss of MRPP2 levels leads to a reduction in steady-state levels of MRPP1.21,22 Given that the increase in RNA precursors (RNA19) we report was similar to those seen in subjects with deleterious HSD17B10 (MRPP2) variants (MIM: 300438),21 it is possible that decreased MRPP1 protein levels are particularly important for RNA processing because MRPP2 levels...
were not shown to be diminished in either of the cases documented here (Figure 2B). Conversely, we did not find any evidence of altered m^9R methylintransferase activity in fibroblasts from either affected individual, implying that the observed defect in mitochondrial protein synthesis is due to a decrease in the efficiency of mt-RNA processing or modification. The m^9R methylintransferase activity is carried out by a stable protein complex of MRPP1 and MRPP2, whereas the m^1R9 methyltransferase activity is carried out by a stable protein complex of MRPP1 and MRPP2, which contains the active site of the nuclease activity of G^iNase P. Therefore, one attractive hypothesis is to suggest that the p.Arg181Leu and p.Thr272Ala TRM10C variants disrupt the interaction between MRPP1 and MRPP5 without affecting the complex with MRPP2, although this requires future investigation.

Knowledge of defects in mt-tRNA processing or modification leading to disease has expanded in recent years including the aforementioned variants in HSD17B10 (MRPP2) and ELAC2, as well as mutations in numerous mt-tRNA modifier proteins including PUS7 (MIM: 608109), TRTL2, TRMT5 (MIM: 610230), TRMT10A (MIM: 613023), MTO1 (MIM: 614667), and GTPBP3 (MIM: 608536). Mutaions in TRMT10C (MRPP1) can now be added to this growing list as we show that the introduction of wild-type MRPP1 into fibroblasts from affected individuals is sufficient to rescue their mitochondrial defects, confirming these TRMT10C variants as pathogenic in mitochondrial disease associated with impaired mitochondrial translation.

Accession Numbers

The data reported in this paper have been deposited in GEO at GSE79120 and in Clívar at SCV000264779.0 and SCV000264780.0 for c.542G>T and c.814A>G, respectively.

Supplemental Data

Supplemental Data include one figure and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2016.03.010.

Acknowledgments

This work was supported by a Wellcome Trust Strategic Award (099019/Z/11/Z) (R.W.T., R.M., and K.T.). C.L.A. is supported by a UK National Health and Medical Research Council (NHMRC) Senior Research Fellow (APP1088442) and is supported by grants from the NHMRC (APP1078273 and APP1043182) and the Association Alexander von Humboldt Fellowship for Experienced Researchers. A.R. is supported by the Association Française contre les Myopathies and the Rare project (GENOMIT [05GM020]). RNA library construction and sequencing were performed at the Cologne Center for Genomics, Cologne, Germany.

Received: January 26, 2016
Accepted: March 14, 2016
Published: April 28, 2016; corrected online: June 21, 2016

Web Resources

OHIM, http://www.ohim.org/
Phyre2, http://www.sbg.bio.ic.ac.uk/phyre2/submit.html?ss=1
PolyPhun2, http://genetics.bwh.harvard.edu/pph2/
RaptorX, http://raptorx.uchicago.edu

References


The American Journal of Human Genetics 98, 993–1000, May 5, 2016 999


Supplemental Data

Recessive Mutations in *TRMT10C* Cause Defects in Mitochondrial RNA Processing and Multiple Respiratory Chain Deficiencies

Supplemental information

Figure S1: Mutations in MRPP1 cause accumulation of mitochondrial precursor RNA. Genome browser illustration of the average normalised read depth of the affected relative to control samples across the mitochondrial heavy and light strands (log2 fold change). The increase in mean coverage across tRNA-encoding regions is attributed to accumulation of unprocessed RNAs. Strand-specific coverage profiles were generated in bedGraph format with bedtools genomecov (Quinlan et al. 2010), normalised to library size (RPM; reads per million mapped). The data represented in this figure have been deposited in NCBI’s Gene Expression Omnibus (Edgar et al., 2002) and are accessible through accession number GEO: GSE79120.

Reference:
Appendix D
Mapping of Mitochondrial RNA-Protein Interactions by Digital RNase Footprinting

Gangqiang Liu,1,4 Timothy R. Mercer,1,3 Anne-Marie J. Shearwood,2 Stefan J. Silara,5 Moira E. Hibbs,5 John S. Mattick,1,4 Oliver Rackham,1 and Aleksandra Filipovska3,*

1Garvan Institute of Medical Research, Sydney NSW 2010, Australia
2Institute for Molecular Bioscience, The University of Queensland, Brisbane QLD 4072, Australia
3Western Australian Institute for Medical Research, Centre for Medical Research and School of Chemistry and Biochemistry, The University of Western Australia, Perth WA 6000, Australia
4St Vincent’s Clinical School and the School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney NSW 2052, Australia
5Correspondence: aleksandra.filipovska@uwa.edu.au

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

SUMMARY

Human mitochondrial DNA is transcribed as long polycistronic transcripts that encompass each strand of the genome and are processed subsequently into mature mRNAs, tRNAs, and rRNAs, necessitating widespread posttranscriptional regulation. Here, we establish methods for massively parallel sequencing and analyses of RNase-accessible regions of human mitochondrial RNA and thereby identify specific regions within mitochondrial transcripts that are bound by proteins. This approach provides a range of insights into the contribution of RNA-binding proteins to the regulation of mitochondrial gene expression.

INTRODUCTION

Mitochondria appear to possess very few transcription factors, but rather have a wide range of RNA-binding proteins that control gene expression at the posttranscriptional level (Åkänen et al., 2007; Rackham et al., 2012). The mitochondrial genome is transcribed as long polycistronic transcripts that are subsequently cleaved and processed into individual mRNAs (Brosz-Rain et al., 2011; Holzmann et al., 2008; Sanchez et al., 2011; Ojala et al., 1981). Recent sequencing of the human mitochondrial transcriptome has revealed previously unexpected complexities of mitochondrial RNAs and identified new transcripts (Mercer et al., 2011; Rackham et al., 2012). The mitochondrial genome has provided evidence that regulation of RNA processing, maturation, stability, translation, and degradation orchestrated by mitochondrial RNA-binding proteins play a major role in mitochondrial gene expression. This indicates that posttranscriptional mechanisms have assumed a prominent role in mammalian mitochondrial gene regulation and expression (Falkenberg et al., 2009; Rackham et al., 2012) and that achieving a profile of their interactions with mitochondrial RNAs will provide regulatory insights akin to transcription factor binding in the nuclear genome (Hesselberth et al., 2009). Here, we develop a digital RNase footprinting method that provides a global profile of RNA-protein interactions in human mitochondria and permits the investigation of posttranscriptional regulation.

RESULTS

Identification of Footprints on Mitochondrial RNAs

To identify footprints of proteins bound to mitochondrial RNAs, we treated isolated mitochondrial preparations with three endonucleases with different cleavage specificities (Figure 1A): RNase A, which cleaves single-stranded RNA after pyrimidine nucleotides; RNase T1, which cleaves single-stranded RNA after guanine residues; and RNase V1, which is a sequence-independent endonuclease that cleaves double-stranded RNAs. A mock-digested mitochondrial preparation was used as an accompanying control. Following the alignment and normalization of data sets to the mitochondrial genome, the site and frequency of RNase A, T1, and V1 cleavage in each data set were assigned to the position immediately upstream of the 5′ nucleotide of each read. A normalized cleavage score (C score) was ascribed to each nucleotide across the mitochondrial genome on all data sets (Experimental Procedures). By comparison to matched control digests of purified mitochondrial RNA, a RNA footprint detection algorithm identified significant contiguous sites with a lower C score than their flanking regions. These represent footprints within mitochondrial transcripts protected from endoribonuclease cleavage by mitochondrial RNA-binding proteins (Figure S1A).

In total, we identified 88 distinct footprints, 33 within mRNAs, 8 from rRNAs, and 7 within tRNAs. The majority of footprints are in specific regions of mitochondrial mRNAs (Figure 1B). We found that the remaining 40 footprints were in transcription regulatory sites within the D loop and in noncoding transcripts encoded by the light strand of the mitochondrial genome. The positioning of footprints within regions of noncoding transcripts indicates a role for proteins in their biogenesis or degradation (Rackham et al., 2011).
Figure 1. Mapping of Mitochondrial RNA-Protein Interactions and Small RNAs

(A) Schematic illustrating the method for mapping mitochondrial RNA-protein interactions by digital Ribo-seq footprinting.

(B) The identified mitochondrial RNA-binding protein footprints are shown in the mitochondrial transcriptome, including those of the TRMT10C protein. The outside and inner circles show mitochondrial genes on the heavy and light strand, respectively.
The digital RNase footprinting provides a global profile of protein interactions with the mitochondrial transcriptome. To provide insight into the identity of specific proteins interacting with the transcriptome, we performed small interfering RNA (siRNA) knockdown of the mitochondrial RNase P protein 1 (MRPP1), recently renamed RNA methyltransferase 10C (TRMT10C), one of the components of the RNase P complex required for cleavage of the 5' ends of mitochondrial mRNAs (Holzmann et al., 2008), and determined its depletion in cells to be over 95% by quantitative RT-PCR and immunoblotting (Figure S1B).

To identify TRMT10C footprints, we compared the data sets where TRMT10C was knocked down to data sets where mitochondria were treated with control siRNAs. We identified 17 unique footprints (Table S1) that were specifically absent from TRMT10C-knockdown mitochondrial data sets. These footprints predominantly reside within mitochondrial mRNAs, such as that found at the 5' end of RNAttaG (MT-5′), which is the 5' end of RNAttaG (MT-TL2) (Figure 1C). Notably, this footprint encompasses the THCP loop of RNAttaG and this region of RNAs has been shown recently to be protected by the mitochondrial RNase P in Arabidopsis, suggesting that there is an evolutionarily conserved binding mode for RNA 5' processing enzymes (Olsbert et al., 2013).

Recently, the presence of a noncoding small RNA (sRNA) population was identified in human mitochondria (Mercer et al., 2011). In our mock-digested mitochondrial lysates, we identified sRNAs ranging from 15 to 32 nt in size and representing 95.8% (Table S2) of the total reads aligned to the mitochondrial genome. Most identified sRNAs align to mitochondrial RNA genes and belong to distinct size classes (Figure 1D). The majority of the 19 nt class of sRNAs aligned to RNAttaG (MT-TP), positions 612–631 of the heavy strand (Figure 1E). This region lies immediately downstream of the major heavy strand promoter, and these sRNAs likely represent the products of abortive transcription initiation by the mitochondrial RNA polymerase (Falkenberg et al., 2007). The remaining sRNAs consist of two distinct species of sRNAs associated with mitochondrial RNA genes, one immediate downstream of the 5' cleavage site of RNAs and a second less abundant species immediately downstream of the 5' cleavage site of RNAs (Figure 1F), that vary in size (Figure S2). The expression of the two mitochondrial RNA-derived sRNA species was not correlated to each other, suggesting that they are generated by independent mechanisms, likely by RNA processing enzymes. For example, we can detect the presence of sRNAs that align to the 5’ end of RNAttaG in cells by northern blotting and show that knockdown of TRMT10C in cells decreased their abundance and that of the mature RNAttaG (Figure 1G), suggesting that the mitochondrial RNase P plays a role in the generation of sRNAs in mitochondria.

Mining of next-generation data sets from Arabidopsis and rice chloroplast genomes has identified sRNAs as potential footprints of RNA-binding proteins (Ryu and Schmitz-Linneweber, 2012; Zhelyazkova et al., 2012). The majority of the identified sRNA footprints are produced from noncoding regions in the chloroplast genome as well as from 5' and 3' ends of mRNAs thought to be regulated by the sequence-specific family of PPR RNA-binding proteins (Ryu and Schmitz-Linneweber, 2012; Zhelyazkova et al., 2012). We integrated our previous small RNA sequencing data set (Mercer et al., 2011) with the data sets from the endonuclease-treated samples, finding only a small number of the identified sRNAs produced from the human mitochondrial transcriptome overlap with RNA-binding protein footprints (Figure 1H). This suggests that most sRNAs are free to take part in intermolecular RNA-RNA interactions and may have regulatory roles in mitochondrial gene expression.

Identification of PPR Protein Footprints

RNA-binding proteins of the PPR family have emerged as important regulators of organellar gene expression and consequently cell health (Rachak and Filipovska, 2012; Schmitz-Linneweber and Smal, 2008). These proteins are similar to the PUF and TALE proteins in that they are predicted to have a sequence-specific mode of nucleic acid recognition (Boch et al., 2009; Filipovska and Rackham, 2012; Missiou and Bogdanove, 2009; Wang et al., 2002). Therefore, there is considerable interest in identifying their RNA targets to understand better their role in RNA metabolism or to use them in biotechnology for targeting specific RNAs of interest. Here, we used RNase digital footprinting to identify binding sites of the mammalian PPR protein, pentatricopeptide repeat domain protein 1 (PTCD1), previously found to affect mitochondrial RNA metabolism (Sanchez et al., 2011; Rachak et al., 2009). We identified five PTCD1-specific RNA footprints in samples from cells where this protein was knocked down using siRNAs, relative to controls. Examples of PTCD1 footprints in RNAttaG and RNAttaG and RNAttaG are shown in Figures 2A and 2B. We investigated the interaction between the identified RNA targets with PTCD1 in vitro using an RNA electrophoretic mobility shift assay (EMSA) (Figure 2C). The binding of the identified RNA targets to PTCD1 compared to a scrambled RNA control validated the specificity of this protein for these targets. This suggests that the RNase digital footprinting method combined with specific siRNA knockdown can be used to effectively predict the targets for RNA-binding proteins of interest.

Protein Footprints upon Stalled Mitochondrial Translation

Protein translation within mitochondria has diverged significantly from translation in the cytoplasm and prokaryotes (Suzuki et al., 2001a, 2001b) and is still poorly understood. To examine the
translational regulation of mRNAs, we treated cells with chloramphenicol to stall mitochondrial translation and performed RNA footprinting analyses. We compared data sets from mitochondria treated in the presence and absence of chloramphenicol to identify 270 footprints that were protected from endonuclease cleavage as a result of inhibiting translation. We identified 124 protein footprints (from the total of 270) that were within mitochondrial mRNAs, of which 22 were 25–35 nt in length and were recognized as mitochondrial ribosome stalling sites or mitoribosome footprints (Figure 3A). An example of a mitochondrial ribosomal footprint in MT-CO1 mRNA during translational stalling is shown in Figure 3B. The ribosome footprints were not enriched at start codons of mitochondrial mRNAs, suggesting that ribosome stalling induced by chloramphenicol may be

Figure 2. Identification of PTCD1 Footprints on Mitochondrial RNAs

(A and B) Specific PTCD1 protein footprints are shown at the 3' ends of tRNA^His (A) and tRNA^Ile (B).

(C) An RNA EMSA shows specific in vitro binding of purified PTCD1 protein with its RNA targets identified in (A) and (B) compared to control RNA.
attributed to the downstream secondary structure. The other 102 protected sites in mitochondrial mRNAs are likely RNA-binding protein footprints that are transiently associated with mRNAs between rounds of translation. The remaining 146 protein footprints from the total 270 identified were in noncoding rRNAs: 36 in rRNAs, 10 in tRNAs, and 7 in the D loop, as well as 86 in transcripts encoded by the light strand and 8 in noncoding regions of the heavy strand. Interestingly, we identified 9 ribosomal footprints in transcripts encoded by the light strand, suggesting that mitochondrial ribosomes may indiscriminately scan or initiate translation of RNAs that have cryptic start codons.

To understand the cause for the ribosomal stalling at specific regions in mitochondrial mRNAs, we investigated the correlation between the 22 identified mitoribosome footprints and RNA secondary structure. Regions within mitochondrial transcripts that are immediately upstream and downstream of the ribosomal footprints have low R scores (see Experimental Procedures), suggesting that these regions are less structured, compared to the regions further away from the footprints that are more structured and consequently have higher R scores (Figure 3C). We confirmed that the single-strand RNA structure upstream and downstream of ribosomal footprints is significant using the Wilcoxon test (see Experimental Procedures), suggesting the ribosome prevents folding of RNA immediately upstream and downstream during translation. We also considered whether the 22 identified ribosomal footprints correlate with codon usage by the mitochondrial ribosome (Jia and Higgs, 2005). We selected the lowest frequency codons in each ribosomal footprint with an average frequency of ~0.008 (Figure 3D) and compared these to the frequency of rare codons in 22 randomly selected mRNAs (Experimental Procedures). The median p value is 0.308 (Figure 3E), which suggests that ribosome stalling is not strongly influenced by the frequency of rare codons.

Digital Mapping of RNA Secondary Structures within Mitochondria

RNAseq cleavage coupled with deep sequencing can be used as a high-throughput method for RNA structural mapping (Frohm and Yoshizawa, 2012). Because RNAseq V1 preferentially cleaves double-stranded RNAs, we used it to analyze local secondary structures of mitochondrial transcripts in our control data sets. We found that the majority of mapped reads aligned to mitochondrial RNAs and that mitochondrial rRNAs, 10 in tRNAs, and 7 in the D loop, as well as 86 in transcripts encoded by the light strand and 8 in noncoding regions of the heavy strand. Interestingly, we identified 9 ribosomal footprints in transcripts encoded by the light strand, suggesting that mitochondrial ribosomes may indiscriminately scan or initiate translation of RNAs that have cryptic start codons.

To understand the cause for the ribosomal stalling at specific regions in mitochondrial mRNAs, we investigated the correlation between the 22 identified mitoribosome footprints and RNA secondary structure. Regions within mitochondrial transcripts that are immediately upstream and downstream of the ribosomal footprints have low R scores (see Experimental Procedures), suggesting that these regions are less structured, compared to the regions further away from the footprints that are more structured and consequently have higher R scores (Figure 3C). We confirmed that the single-strand RNA structure upstream and downstream of ribosomal footprints is significant using the Wilcoxon test (see Experimental Procedures), suggesting the ribosome prevents folding of RNA immediately upstream and downstream during translation. We also considered whether the 22 identified ribosomal footprints correlate with codon usage by the mitochondrial ribosome (Jia and Higgs, 2005). We selected the lowest frequency codons in each ribosomal footprint with an average frequency of ~0.008 (Figure 3D) and compared these to the frequency of rare codons in 22 randomly selected mRNAs (Experimental Procedures). The median p value is 0.308 (Figure 3E), which suggests that ribosome stalling is not strongly influenced by the frequency of rare codons.

Digital Mapping of RNA Secondary Structures within Mitochondria

RNAseq cleavage coupled with deep sequencing can be used as a high-throughput method for RNA structural mapping (Frohm and Yoshizawa, 2012). Because RNAseq V1 preferentially cleaves double-stranded RNAs, we used it to analyze local secondary structures of mitochondrial transcripts in our control data sets. We found that the majority of mapped reads aligned to mitochondrial RNAs and that mitochondrial rRNAs, 10 in tRNAs, and 7 in the D loop, as well as 86 in transcripts encoded by the light strand and 8 in noncoding regions of the heavy strand. Interestingly, we identified 9 ribosomal footprints in transcripts encoded by the light strand, suggesting that mitochondrial ribosomes may indiscriminately scan or initiate translation of RNAs that have cryptic start codons.

To understand the cause for the ribosomal stalling at specific regions in mitochondrial mRNAs, we investigated the correlation between the 22 identified mitoribosome footprints and RNA secondary structure. Regions within mitochondrial transcripts that are immediately upstream and downstream of the ribosomal footprints have low R scores (see Experimental Procedures), suggesting that these regions are less structured, compared to the regions further away from the footprints that are more structured and consequently have higher R scores (Figure 3C). We confirmed that the single-strand RNA structure upstream and downstream of ribosomal footprints is significant using the Wilcoxon test (see Experimental Procedures), suggesting the ribosome prevents folding of RNA immediately upstream and downstream during translation. We also considered whether the 22 identified ribosomal footprints correlate with codon usage by the mitochondrial ribosome (Jia and Higgs, 2005). We selected the lowest frequency codons in each ribosomal footprint with an average frequency of ~0.008 (Figure 3D) and compared these to the frequency of rare codons in 22 randomly selected mRNAs (Experimental Procedures). The median p value is 0.308 (Figure 3E), which suggests that ribosome stalling is not strongly influenced by the frequency of rare codons.

Digital Mapping of RNA Secondary Structures within Mitochondria

RNAseq cleavage coupled with deep sequencing can be used as a high-throughput method for RNA structural mapping (Frohm and Yoshizawa, 2012). Because RNAseq V1 preferentially cleaves double-stranded RNAs, we used it to analyze local secondary structures of mitochondrial transcripts in our control data sets. We found that the majority of mapped reads aligned to mitochondrial RNAs and that mitochondrial rRNAs, 10 in tRNAs, and 7 in the D loop, as well as 86 in transcripts encoded by the light strand and 8 in noncoding regions of the heavy strand. Interestingly, we identified 9 ribosomal footprints in transcripts encoded by the light strand, suggesting that mitochondrial ribosomes may indiscriminately scan or initiate translation of RNAs that have cryptic start codons.

To understand the cause for the ribosomal stalling at specific regions in mitochondrial mRNAs, we investigated the correlation between the 22 identified mitoribosome footprints and RNA secondary structure. Regions within mitochondrial transcripts that are immediately upstream and downstream of the ribosomal footprints have low R scores (see Experimental Procedures), suggesting that these regions are less structured, compared to the regions further away from the footprints that are more structured and consequently have higher R scores (Figure 3C). We confirmed that the single-strand RNA structure upstream and downstream of ribosomal footprints is significant using the Wilcoxon test (see Experimental Procedures), suggesting the ribosome prevents folding of RNA immediately upstream and downstream during translation. We also considered whether the 22 identified ribosomal footprints correlate with codon usage by the mitochondrial ribosome (Jia and Higgs, 2005). We selected the lowest frequency codons in each ribosomal footprint with an average frequency of ~0.008 (Figure 3D) and compared these to the frequency of rare codons in 22 randomly selected mRNAs (Experimental Procedures). The median p value is 0.308 (Figure 3E), which suggests that ribosome stalling is not strongly influenced by the frequency of rare codons.
Identification of Footprints of Mitochondrial RNA-Binding Proteins

We transformed the mapped reads to wig format, which represent the value of the F score between the experimental and control conditions. To estimate if the F score fold change is significant for a footprint, we built an empirical null distribution of F score (control); second, we divided the F score (control) by the F score (experiment) and the F score for the equivalent control data at the region, and shuffled them 1,000 times. For each upstream or downstream nucleotide, we employed the Wilcoxon test (95% confidence level of the interval) to compare the ribosomal footprint R score to random sites and calculated an average p value from the 1,000 tests.

Northern Blotting

Northern blotting of tRNA was carried out as described before (Pul and Hamilton, 2008), and the tRNAs were detected with biotinylated probes as described before (Rochman et al., 2009).

Purification of PTCD1 Protein

Human PTCD1 lacking its mitochondrial targeting sequence but including all Ribosome footprints were used as controls. At the end of the incubation, the reactions were analyzed by 15% PAGE in Tris-acetate-EDTA and fluorescence was detected using a Typhoon FLA 9500 biomolecular image (GE).

Correlation of R Score and rRNA Secondary Structure

Unpaired predictive value = \( \frac{TP}{TP + FP} \)

Paired predictive value = \( \frac{TN}{TN + FN} \)
Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}

MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}

The MCC is used here to measure the quality of R score classification. A coefficient of +1 represents a perfect prediction, 0 indicates no better than random prediction, and -1 indicates disagreement between prediction and observation.

ACCESSION NUMBERS

The Sequence Read Archive (SRA) accession number for the data reported in this paper is SRA098892.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.09.036.

Figure 4. In Vivo Analysis of Mitochondrial RNA Secondary Structure

(A) Histogram of the R score of every nucleotide across the whole mitochondrial transcriptome from control mitochondria and its matched purified RNA data sets. From center, the outer two tracks show the R score of each nucleotide of transcripts on the mitochondrial heavy strand, while the inner two tracks correspond to the light strand. This figure was constructed using Circos (Krzywinski et al., 2009).

(B) R score across the coding sequences (CDS) of mitochondrial mRNA transcripts. Transcripts were aligned by their translational start and stop sites; the horizontal line denotes the average R score of mitochondrial CDS.

(C and D) R score across every nucleotide of tRNAVal from control mitochondria (C) and its matched purified RNA data sets and their heat maps (D) showing the specific loops and stems of this tRNA.

846 Cell Reports 5, 839–848, November 14, 2013 ©2013 The Authors


Supplementary Figure Legends

Figure S-1. (A) The footprint detection algorithm. (B) The expression levels of TRMT10C and PTCD1 were significantly decreased in 143B cells treated with TRMT10C or PTCD1 siRNAs compared to control siRNA treated cells. Porin was used as a loading control.

Figure S-2. A histogram showing three examples of the proportion of sRNAs that align to tRNA$^{\text{Leu(UUR)}}$ (A), tRNA$^{\text{Ser(AGY)}}$ (B) and tRNA$^{\text{Gly}}$ (C) relative to total mapped reads.

Figure S-3. Kernel density plotting of R score of nucleotides in control mitochondria and matched purified mitochondrial RNA datasets.

Figure S-4. Correlation of R score with the 12S rRNA secondary structure model.

Figure S-5. Correlation of R score with the 16S rRNA secondary structure model.
A

chrM: $1.0 \times 10^4$

sRNA frequency (raw count)

3,325

16S rRNA, tRNA-Leu

B

chrM: $8.0 \times 10^4$

sRNA frequency (raw count)

12,180

tRNA-His, tRNA-Ser, tRNA-Leu

C

chrM: $3.0 \times 10^3$

sRNA frequency (raw count)

10,100

CO3, tRNA-Gly, ND3
Appendix E
RESEARCH ARTICLE

SLIRP Regulates the Rate of Mitochondrial Protein Synthesis and Protects LRPPRC from Degradation

Marie Lagouge1, Arnaud Mourier1, Hyun-Ju Lee1, Henrik Spåhr1, Timothy Wei3, Christian Kukat3, Eduardo Silva Ramos1, Elisa Motori1, Jakob D. Busch1, Stefan Siira4, German Mouse Clinic Consortium5, Elisabeth Kremmer6, Aleksandra Filipovska4*, Nils-Göran Larsson1,7

1 Department of Mitochondrial Biology, Max Planck Institute for Biology of Ageing, Cologne, Germany, 2 CECAD Research Center, University of Cologne, Cologne, Germany, 3 FACS and Imaging facility, Max Planck Institute for Biology of Ageing, Cologne, Germany, 4 Harry Perkins Institute of Medical Research, Centre for Medical Research and School of Chemistry and Biochemistry, The University of Western Australia, Perth, Australia, 5 German Mouse Clinic, Helmholtz Zentrum Münchner, German Research Center for Environmental Health GmbH, Neuherberg, Germany, 6 Helmholtz Zentrum München, Institute of Molecular Immunology, Munich, Germany, 7 Department of Laboratory Medicine, Karolinska Institute, Stockholm, Sweden

* Current address: Inserm U1138, Centre de Recherche des Cordeliers, Paris, France

Abstract
We have studied the in vivo role of SLIRP in regulation of mitochondrial DNA (mtDNA) gene expression and show here that it stabilizes its interacting partner protein LRPPRC by protecting it from degradation. Although SLIRP is completely dependent on LRPPRC for its stability, reduced levels of LRPPRC persist in the absence of SLIRP in vivo. Surprisingly, Slirp knockout mice are apparently healthy and only display a minor weight loss, despite a 50–70% reduction in the steady-state levels of mtDNA-encoded mRNAs. In contrast to LRPPRC, SLIRP is dispensable for polyadenylation of mtDNA-encoded mRNAs. Instead, deep RNA sequencing (RNAseq) of mitochondrial ribosomal fractions and additional molecular analyses show that SLIRP is required for proper association of mRNAs to the mitochondrial ribosome and efficient translation. Our findings thus establish distinct functions for SLIRP and LRPPRC within the LRPPRC-SLIRP complex, with a novel role for SLIRP in mitochondrial translation. Very surprisingly, our results also demonstrate that mammalian mitochondria have a great excess of transcripts under basal physiological conditions in vivo.

Author Summary
Mitochondria provide most of the energy required for key metabolic and cellular processes that are essential for life. The biogenesis of the mitochondrial oxidative phosphorylation system, the site of energy conversion, is dependent on the coordinated expression of the

PLOS Genetics | DOI:10.1371/journal.pgen.1005423 August 6, 2015 1/1 9

OPEN ACCESS


Editor: Gregory S. Barsh, Stanford University School of Medicine, UNITED STATES

Received: March 17, 2015
Accepted: July 4, 2015
Published: August 6, 2015

Copyright: © 2015 Lagouge et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The study was supported by a post-doctoral grant from the AXA Research Fund (ARF) to ML, by a European Research Council advanced investigator grant (268897) and by grants from the Deutsche Forschungsgemeinschaft (DFG) to NGL. AF is a National Health and Medical Research Council Senior Research Fellow (APP 1058442) and supported by an Alexander von Humboldt Fellowship for experienced researchers. Members of the GMC were funded by the German Federal Ministry of
mitochondrial and nuclear genomes. Mitochondrial gene expression is largely regulated at the post-transcriptional level by RNA-binding proteins, including the LRPPRC-SLIRP complex. It is still unclear how the proteins within this complex regulate mitochondrial RNA metabolism. Here, we have knocked out the Slirp gene in mice to dissect the individual roles of LRPPRC and SLIRP and provide further insights into the mechanisms governing post-transcriptional regulation of mitochondrial gene expression. LRPPRC is required for the maintenance of mitochondrial mRNA polyadenylation whereas SLIRP, by facilitating the presentation (or association) of mRNAs to the mitochondrial ribosome, regulates the rate of translation. In addition, we demonstrate that mitochondrial mRNAs in mammals are present in quantities that far exceed those needed to maintain normal physiology under basal conditions.

Introduction
Mitochondria are double-membrane bound organelles that have fundamental roles in energy metabolism, cell health and death, making them essential for life. The oxidative phosphorylation (OXPHOS) system is the major site of ATP production in mitochondria and is composed of proteins encoded by two genomes, the nuclear genome and mitochondrial DNA (mtDNA). Consequently coordinated regulation of nuclear and mitochondrial gene expression is required to meet the changing energy demands of the cell. The compact size and organization of mtDNA in animals has necessitated the evolution of unique mechanisms to regulate the expression of the 13 subunits of the OXPHOS system that are mitochondrially encoded. Mitochondrial gene expression is complex and predominantly regulated at the post-transcriptional level [1,2] by nuclear-encoded mitochondrial RNA-binding proteins that control the processing, maturation, translation, stabilization and degradation of mitochondrial RNAs [3]. The mitochondrial RNA polymerase (POLRMT) stimulated by mitochondrial transcription factor A (TFAM) and B2 (TFB2M) produces near-genome length polycistronic transcripts [3]. Because animal mtDNA lacks introns, the 22 mitochondrial tRNA genes that are arranged between the 2 rRNA and 11 mRNA coding genes act as punctuation marks to signal the processing of the polycistronic transcripts [4] by mitochondrial tRNA (mt-tRNA) processing enzymes [5–7]. The processed transcripts undergo extensive maturation, including polyadenylation at the 3' end of the mitochondrial mRNAs (mt-mRNAs) [8,9], and mitochondrial rRNAs (mt-rRNAs) and mt-tRNAs are modified enzymatically at specific nucleosides to enable proper folding and biogenesis of the translation machinery [2,3]. The matured mt-mRNAs are translated on mitochondrial ribosomes (mitoribosomes) [10], although it is not clear how they are recognized as they lack conventional 5' and 3' untranslated regions (UTRs), Shine-Dalgarno sequences and 5' 7- methylguanosine caps [11].

The mammalian family of RNA-binding pentatricopeptide repeat domain (PPR) proteins consists of seven nuclear-encoded mitochondrial proteins, each of which has a specific role in regulating mitochondrial gene expression from transcription and processing to maturation and translation [12]. The PPR protein LRPPRC first came to attention when a mutation of the LRPPRC gene was shown to cause a rare French-Canadian variant of Leigh syndrome characterized by cytochrome c oxidase deficiency [13]. In cultured cells, LRPPRC knockdown (KD) causes a reduction of mt-mRNA levels [14–16] and impaired mitochondrial translation [16]. LRPPRC physically interacts with SLIRP, which has an RNA recognition motif (RRM), consistent with a role in mitochondrial RNA metabolism [16]. LRPPRC and SLIRP form a complex that mediates mt-mRNA stability [15–17] and both proteins are co-stabilized within this...
complex because reduction of LRPPRC levels leads to concomitant reduction of SLIRP [14–18]. In mice, the LRPPRC-SLIRP complex regulates mt-mRNA stability, polyadenylation and coordinated mitochondrial translation [17]. We have also demonstrated that the bicoid stability factor (BSF, renamed DmLRPPRC1), one of the two Drosophila melanogaster orthologues of mammalian LRPPRC [19,20], has a very similar function as the mammalian one [21]. Furthermore, DmLRPPRC1 associates with one of the two fly orthologues of SLIRP [19,21], suggesting that the interaction between PPR-motif- and RRM-containing proteins is important for mitochondrial RNA metabolism and has been conserved through evolution.

To address the unclear in vivo role of SLIRP and its function within the LRPPRC-SLIRP complex, we generated Slirp knockout mice. Molecular analyses revealed that SLIRP is required to stabilize LRPPRC. In addition, our findings show that LRPPRC and SLIRP have distinct roles within the mt-mRNA-stabilizing complex they form, i.e. LRPPRC is required for maintenance of polyadenylation whereas SLIRP regulates the rate of translation. Very surprisingly, we also report that mice lacking SLIRP are apparently healthy despite a very drastic (50–70%) depletion of mt-mRNAs. These findings show that mt-mRNAs in mammalian mitochondria are present in quantities that far exceed those needed to maintain normal physiology.

**Results**

**Generation of Slirp knockout mice**

In mammals SLIRP forms a complex with the mitochondrial protein LRPPRC [16–18] and the complex is required for the stability of mt-mRNAs, polyadenylation and coordinated mitochondrial translation [15–17]. SLIRP is predicted to localize to mitochondria with a probability of 94.4% using the MitoProtII software [22] and we confirmed this prediction by using immunocytochemistry to show that endogenous SLIRP co-localizes with the mitochondrial ATPase complex (S1A Fig) in 143B cells.

To investigate the specific role of SLIRP within mitochondria in vivo, we generated a germ-line Slirp knockout (KO) mouse model (Slirp−/−) via excision of the floxed exon 2 of Slirp by expressing the Cre-recombinase under the control of the β-actin promoter (S1B Fig). The resulting Slirp−/− mice (S1C Fig) were inter-crossed to generate Slirp−/− mice and all expected genotypes were obtained at Mendelian ratios, thus showing that SLIRP, in contrast to LRPPRC [17], is not required for embryonic development. Mice lacking SLIRP were apparently healthy with no obvious phenotype, except a slight reduction in body weight (S1D Fig). In contrast to a previous report [23], we also found that lack of SLIRP does not impair fertility as crosses between Slirp−/− males or females and wild-type mice produced normal litter sizes. These findings show that the in vivo roles of SLIRP and LRPPRC are at least partly divergent.

**Slirp is necessary for the stability of mitochondrial mRNAs in vivo**

Steady-state SLIRP levels have been shown to correlate with those of LRPPRC [14–16] and conditional KO of Lrpprc causes complete loss of SLIRP [17]. Therefore, we investigated LRPPRC levels by immunoblotting of mitochondria isolated from Slirp−/− mice and found that ~25% of the LRPPRC protein remained in heart, liver and kidney of these mice (Fig 1A). The reduction of LRPPRC protein levels upon Slirp deletion was also confirmed by immunocytochemistry on mouse embryonic fibroblasts (MEFs) derived from Slirp−/− mice (Fig 1B). Our findings thus show that low levels of LRPPRC can be maintained without forming a complex with SLIRP, whereas in wild-type mice LRPPRC and SLIRP form a stable complex in both heart and liver, as confirmed by co-immunoprecipitation (S1E Fig). We found that LRPPRC is reduced by ~50% in heart, liver and kidney mitochondria in Lrpprc−/− mice, consistent with our previous results [24], and these levels were further reduced to ~25% in the Slirp−/− mice.
Fig 1. Loss of SLIRP compromises the stability of mitochondrial mRNAs and LRPPRC. (A) Immunoblot of SLIRP and LRPPRC protein levels in heart, liver, and kidney mitochondria from 12-week old wild-type (Slirp +/+), Slirp homozygous knockout (KO, Slirp -/-) and Lrpprc heterozygous KO (Lrpprc +/-) mice.
SLIRP Fine-Tunes Mitochondrial Translation

VDAC was used as a loading control. (B) Representative confocal microscopy images of Slirp+/+ (left) and Slirp−/− (right) MEFs, stained for LRPPRC and TOM20 as a mitochondrial marker. Magnifications of the dashed boxed areas show merged channels of LRPPRC and TOM20 (bottom, left) and single channels of LRPPRC (bottom, right). Scale bars presented, 10 μm. (C) In organo-de novo transcription assays performed on heart and liver mitochondria isolated from 12-week old Slirp+/+ and Slirp−/− mice. VDAC was used as a loading control. (D) Mitochondrial mRNA steady-state levels assessed by northern blotting in hearts of 12-week old Slirp+/+ and Slirp−/− mice. (E) Mitochondrial transcript steady-state levels assessed by qRT-PCR in hearts from 12-week old Slirp+/+ (white bars) and Slirp−/− (black bars) mice, as well as in hearts from 12-week old Lrpprc control (Lrpprc p/p), light grey bars) and conditional KO (Lrpprc p/cre, dark grey bars), n = 5. Error bars represent SEM. * p value < 0.05, ** p value < 0.01, *** p value < 0.001. (F) Measurement of mitochondrial transcript poly(A) tail length from heart mitochondria of Slirp+/+ (white bars), Slirp−/− (black bars), Lrpprc−/− (yellow bars) and Lrpprc p/c (cre grey bars) mice. Error bars represent SEM. n.s. means not significant. * p value < 0.05.

SLIRP does not affect the polyadenylation status of mitochondrial mRNAs

In mammalian mitochondria, mRNAs, with the exception of Nd6, contain short poly(A) tails [4,8], which are necessary to complete the termination codon for seven of the total 11 mt-mRNAs. The LRPPRC-SLIRP complex, which is involved in mt-mRNA stability, has been found to maintain polyadenylation [15,17], but the role for poly(A) tails in the regulation of mt-mRNA stability is unclear. However, polyadenylation appears to have roles in mitochondrial translation that are distinct from termination codon formation [25,26], which is consistent with the requirement of LRPPRC for coordinated translation in mammalian mitochondria [16,17]. Interestingly, we found that poly(A) tail length was intact in the absence of SLIRP (Figs 1F and S2D), which shows that SLIRP is not required for the in vivo maintenance of mitochondrial polyadenylation. This finding also demonstrates that the presence of poly(A) tails is not sufficient to ensure mt-mRNA stability in vivo when SLIRP is lost (Fig 1E and 1F). LRPPRC

DOI:10.1371/journal.pgen.1005423.g001
has been shown to promote polyadenylation by the mitochondrial poly(A) polymerase [15,27] and our findings show that normal poly(A) tail length can be maintained even if the levels of LRPPRC are low, as it is the case in the Lypprec−/− and Slirp−/− mice (Fig 1A and 1H). A corollary of this is that SLIRP may have an additional function besides maintaining mt-mRNA stability as part of the LRPPRC-SLIRP complex.

Loss of SLIRP affects the engagement of mitochondrial mRNAs with the mitochondrial ribosome

Next we investigated how decreased levels of mt-mRNAs affected the protein synthesis machinery in mitochondria from Slirp−/− mice. We measured the steady-state levels of the mitochondrial 12S and 16S rRNAs and found that they were increased in Slirp−/− relative to control mice (Figs 1E and S2B and S2C). The increase in 16S rRNA correlated with an increased amount of MRPL37, a mitochondrial ribosomal protein (MRP) of the large subunit, in Slirp−/− heart and liver mitochondria (S3A Fig). This apparent increase in mitoribosome biogenesis is presumably a compensatory response to the reduced mt-mRNA stability observed upon SLIRP loss.

To assess the state of association of mt-mRNAs with the mitoribosome we performed sucrose sedimentation gradient analyses of mitochondrial extracts. We used qRT-PCR to determine the sedimentation profile of the small (28S) and large (39S) ribosomal subunits and the fully assembled mitoribosome (55S) (Figs 2A and S3C). The 12S rRNA co-migrated with the MRPS35 protein of the 28S subunit and the 16S rRNA co-migrated with MRPL37 (Figs 2A and S3C, left panels), which shows that the 28S subunit was mainly present in fractions 6–7, the 39S subunit in fraction 9 and the 55S mitoribosome in fractions 11–12 in control mice. Strikingly, in Slirp−/− mitochondria of liver (Fig 2A, right panel) and heart (S3C Fig, right panel), the ribosome profiles were altered as shown by the continuous distribution of MRPL37 and 16S rRNA between fractions 9 and 12. This continuous distribution may occur as a consequence of the increased steady-state levels of mt-mRNAs and MRPs (S3A Fig). We also measured the abundance of mt-mRNAs in the different fractions of the gradient by qRT-PCR and could identify two distinct pools, one translationally inactive in fractions 4–5 and a second one, translationally active, that co-migrates with the assembled mitoribosome (Figs 2A and S3C).

To investigate the proportion of mt-mRNAs engaged with the mitoribosome without being misled by the global decrease of mt-mRNA levels in Slirp−/− mice, we normalized Slirp−/− mt-mRNA levels to those of the control samples (Fig 2B). Interestingly, after normalization, we found that mt-mRNAs were less engaged with the assembled mitoribosome in the Slirp−/− liver mitochondria in comparison with controls (Fig 2B). Strikingly, the profile was the opposite in heart where we found increased engagement of mt-mRNAs with the 55S mitoribosome in the Slirp−/− heart mitochondria in comparison with controls (S3D Fig).

Next we investigated the association of mt-mRNAs with the mitoribosome by performing RNA sequencing (RNAseq) of fractions from liver mitochondria that corresponded to the 28S and 39S subunits and to the 55S mitoribosome. In addition we carried out RNAseq of the fractions between the 39S subunit and the 55S mitoribosome, as we observed a continuous distribution of large subunit proteins and rRNA in this region of the gradient in Slirp−/− mitochondria. Differential expression analyses of the mt-mRNAs indicate a global and dramatic decrease of their abundance across the ribosomal fractions in liver mitochondria where SLIRP is lost (Fig 2C), which is in line with the reduced mt-mRNA steady-state levels previously assessed (S2C Fig). The levels of the Nd6 mt-mRNA associated with the mitoribosome were not affected by the loss of SLIRP (Fig 2C), suggesting that the association of Nd6 with the mitoribosome is possibly independent from the LRPPRC-SLIRP complex. Furthermore, both
SLIRP Fine-Tunes Mitochondrial Translation

Fig 2. SLIRP loss affects the engagement of mitochondrial mRNAs with the mitochondrial ribosome. (A) Sedimentation profile in sucrose density gradient of transcripts and ribosomal proteins from liver mitochondria from 12-week old wild-type (Slirp+/+, left panel) and Slirp homozygous knockout (Slirp−/−, right panel) animals. Individual mitochondrial transcripts were detected by qRT-PCR. The abundance of a given RNA in each fraction is shown as percentage of the total level in the control. The migration of the small (28S) and large (36S) mitochondrial ribosomal subunits, and of the assembled mitochondrial ribosome (55S) is assessed by the profiles of the 12S (blue line) and 16S (red line) rRNAs as well as by the migration of subunit-specific ribosomal proteins (MRPL37 and MRPS35) detected by immunoblotting. (B) Individual mRNA sedimentation profiles from the gradient described in (A). Slirp+/+ profiles are depicted in purple and Slirp−/− profiles are depicted in orange. Individual mitochondrial mRNAs were detected by qRT-PCR and the mRNA distribution profile is shown after normalization to controls, i.e. the quantity of a given mt-mRNA, named RNAx, in each fraction of the Slirp−/− gradient was normalized to the total RNAxquantity,total RNAxquantity ratio, where total RNAs is the sum of the RNAx quantity detected across all the fractions. (C) Hierarchical clustered expression levels of mitochondrial transcripts (log10FPKM) across all fractions from both Slirp+/+ and Slirp−/− mitochondria. (D) Hierarchical clustered log2 fold changes in transcript expression for each fraction, showing the overall decrease in mRNA levels of Slirp−/− compared to Slirp+/+ mitochondria.

doi:10.1371/journal.pgen.1005423.g002

SLIRP Fine-tunes the rate of mitochondrial protein synthesis

We proceeded to measure mitochondrial protein synthesis to assess the biological significance of the altered engagement of mt-mRNAs with the mitochondrial ribosome in the absence of SLIRP. We determined the rate of translation by following 35S-methionine incorporation into newly synthetized mitochondrial polypeptides over time in MEFs (Fig 3A) and in isolated heart, liver (Fig 3B) and kidney (S4A Fig) mitochondria. Interestingly, we found that the translation rate was impaired in the Slirp−/− MEFs as well as liver and kidney mitochondria (Figs 3A and S4A), which is in line with the observed reduced engagement of the mt-mRNAs with the mitochondrion (55S) in Slirp−/− liver mitochondria (Fig 2B and 2C). In contrast, in Slirp−/− heart mitochondria the incorporation of 35S-methionine was comparable to that of control heart mitochondria (Fig 3B), with the exception of Ndu2 and Cox1/Nd4 whose translation seemed to be affected by the loss of SLIRP. The maintenance of a comparable translation rate despite Slirp knockout is consistent with the observed increased engagement of mt-mRNAs with the 55S mitochondrion in Slirp−/− hearts (S3B Fig). These findings suggest that SLIRP is involved in preventing mature mRNAs to the mitochondrion in order to promote mitochondrial translation, but that its loss can be compensated for in certain tissues such as the heart. We found that the steady state levels of the mitochondrial translation initiation factor 3 (mtIF3) were increased in Slirp−/− mitochondria in comparison with controls, especially in the liver (S3B Fig), which likely constitutes a compensatory response to the impaired rate of translation. The tissue-specific mitochondrial translation defect, which is very minor in the heart and more apparent in the liver (Fig 3B) and kidney (S4A Fig) mitochondria. The rate of translation per mt-mRNA is presumably higher in the liver and kidney mitochondria, allowing a more pronounced compensatory response to the impaired translation rate in the heart. An intriguing question is whether the Slirp−/− heart mitochondria translate the mt-mRNAs more efficiently after the translation rate is impaired, analogous to the slight reduction in the size of the ribosomal subunits that we observed in Slirp−/− heart mitochondria (Fig 2D) compared to controls. Moreover, the substantial upregulation of the mitochondrial translation initiation factor 3 (mtIF3) in Slirp−/− heart mitochondria suggests that this factor, which regulates the presentation of mature mt-mRNA to the mitochondrial ribosome, is increased to compensate for the impaired translation rate in Slirp−/− heart mitochondria.
liver and kidney, does not seem to impact the assembly of the RC subunits, as their steady-state levels are similar in liver and heart mitochondria from control and Slirp-/- mice (Fig 3C). Furthermore the respiration was normal under phosphorylating, non-phosphorylating, and uncoupled conditions in Slirp-/- mitochondria from liver (Fig 3D) and heart (S4C Fig). Consistently, the RC enzyme activities of complexes I, II and IV were comparable in mitochondria from liver (Fig 3E) and heart (S4D Fig) in Slirp-/- and control mice.

Together, these data argue that SLIRP can act as a general activator of mitochondrial translation, whose loss (i) can be overcome by an unknown mechanism in tissues such as the heart and (ii) is not sufficient to induce OXPHOS dysfunction in tissues where translation rate is affected such as the liver and kidney. We hypothesize that despite their reduced stability and impaired loading onto the mitoribosome, mt-mRNAs in Slirp-/- mice are still translated at rates sufficient to preserve normal OXPHOS activity, which likely explains the absence of a pathophysiology in the Slirp-/- mice under basal conditions. In addition, the stability of the mitochondria-encoded RC subunits as observed by 35S-methionine pulse-chase assay in the Slirp-/- MEFs (S4B Fig) is likely contributing to the maintenance of normal OXPHOS function despite the decrease in translation.

Stabilization of LRPPRC cannot rescue mitochondrial mRNA stability in the absence of SLIRP

In plants, PPR proteins have been found to associate via protein-protein interactions with additional RNA-binding proteins, including RRM proteins to regulate gene expression [29]. The co-stabilization of SLIRP and LRPPRC as a complex [16–18] has provided a challenge to specifically decipher their individual roles in mitochondria. Our data suggest that the decreased levels of mt-mRNA in Slirp-/- mitochondria could possibly be a consequence of the decreased LRPPRC protein levels, where SLIRP could act as a stabilizing partner for LRPPRC without directly affecting mt-mRNA stability. To further investigate this hypothesis, we generated mice ubiquitously overexpressing Lrpprc on a Slirp KO background (Fig 4A), in an attempt to overcome the co-stability dependence of the two proteins. Interestingly, we found that LRPPRC protein levels could not be restored in the absence of SLIRP thus confirming that SLIRP is essential for LRPPRC protein stabilization. Mitochondrial protein turnover is regulated by several proteases [30], among which LONP1 has been shown to target components of the mitochondrial gene expression machinery [31,32]. By knocking down the expression of Lonp1 in Slirp-/- MEFs we could partially restore LRPPRC protein levels, demonstrating that LRPPRC is targeted for degradation by this matrix protease in the absence of SLIRP (Fig 4B). We used this rescue model to determine if the increased steady-state levels of LRPPRC would restore mt-mRNA levels in the absence of SLIRP (Fig 4C). However, the significant rescue of LRPPRC steady-state levels induced by the Lonp1 KD (Fig 4B) did not significantly increase mt-mRNA levels (Fig 4C), yet in the absence of any adverse effect on the mt-RNA degradation machinery (Fig 4B) [33]. In previous work we have shown that Lrpprc+/- mice, that have a ~50% reduction liver and kidney, does not seem to impact the assembly of the RC subunits, as their steady-state levels are similar in liver and heart mitochondria from control and Slirp-/- mice (Fig 3C). Furthermore the respiration was normal under phosphorylating, non-phosphorylating, and uncoupled conditions in Slirp-/- mitochondria from liver (Fig 3D) and heart (S4C Fig). Consistently, the RC enzyme activities of complexes I, II and IV were comparable in mitochondria from liver (Fig 3E) and heart (S4D Fig) in Slirp-/- and control mice.

Together, these data argue that SLIRP can act as a general activator of mitochondrial translation, whose loss (i) can be overcome by an unknown mechanism in tissues such as the heart and (ii) is not sufficient to induce OXPHOS dysfunction in tissues where translation rate is affected such as the liver and kidney. We hypothesize that despite their reduced stability and impaired loading onto the mitoribosome, mt-mRNAs in Slirp-/- mice are still translated at rates sufficient to preserve normal OXPHOS activity, which likely explains the absence of a pathophysiology in the Slirp-/- mice under basal conditions. In addition, the stability of the mitochondria-encoded RC subunits as observed by 35S-methionine pulse-chase assay in the Slirp-/- MEFs (S4B Fig) is likely contributing to the maintenance of normal OXPHOS function despite the decrease in translation.

Stabilization of LRPPRC cannot rescue mitochondrial mRNA stability in the absence of SLIRP

In plants, PPR proteins have been found to associate via protein-protein interactions with additional RNA-binding proteins, including RRM proteins to regulate gene expression [29]. The co-stabilization of SLIRP and LRPPRC as a complex [16–18] has provided a challenge to specifically decipher their individual roles in mitochondria. Our data suggest that the decreased levels of mt-mRNA in Slirp-/- mitochondria could possibly be a consequence of the decreased LRPPRC protein levels, where SLIRP could act as a stabilizing partner for LRPPRC without directly affecting mt-mRNA stability. To further investigate this hypothesis, we generated mice ubiquitously overexpressing Lrpprc on a Slirp KO background (Fig 4A), in an attempt to overcome the co-stability dependence of the two proteins. Interestingly, we found that LRPPRC protein levels could not be restored in the absence of SLIRP thus confirming that SLIRP is essential for LRPPRC protein stabilization. Mitochondrial protein turnover is regulated by several proteases [30], among which LONP1 has been shown to target components of the mitochondrial gene expression machinery [31,32]. By knocking down the expression of Lonp1 in Slirp-/- MEFs we could partially restore LRPPRC protein levels, demonstrating that LRPPRC is targeted for degradation by this matrix protease in the absence of SLIRP (Fig 4B). We used this rescue model to determine if the increased steady-state levels of LRPPRC would restore mt-mRNA levels in the absence of SLIRP (Fig 4C). However, the significant rescue of LRPPRC steady-state levels induced by the Lonp1 KD (Fig 4B) did not significantly increase mt-mRNA levels (Fig 4C), yet in the absence of any adverse effect on the mt-RNA degradation machinery (Fig 4B) [33]. In previous work we have shown that Lrpprc+/- mice, that have a ~50% reduction
of LRPPRC protein levels, have normal mitochondrial transcript stability [24]. In contrast, we show here that a similar LRPPRC level reduction combined with loss of SLIRP, as seen in

Fig 4. LRPPRC is degraded by LONP1 in the absence of SLIRP and LRPPRC alone cannot preserve mitochondrial transcript stability. (A) Immunoblotting of SLIRP and LRPPRC protein levels in liver and kidney mitochondria from 12-week old wild-type (Slirp+/-, Lrpprc+/-) and Slirp homozygous knockout (Slirp-/-, Lrpprc+/-) mice and mice overexpressing Lrpprc on a Slirp homozygous knockout background (Slirp-/-, Lrpprc+/+) and Slirp-/-, Lrpprc+/+ primary MEFs after transfection with a siRNA directed against the expression of the LONP1 protease (siLonp1) or with a scrambled siRNA (siCtrl). LONP1 was detected to assess the efficiency of the knockdown (KD), SUV3L1 and PNPT1 were detected to assess the steady-state level of the mitochondrial RNA degradosome and SDHA was used as a loading control (left panel). The right panel represents the quantification of three independent experiments. Error bars represent SEM. * p value < 0.05, ** p value < 0.01, *** p value < 0.001. n.s. means not significant. doi:10.1371/journal.pgen.1005423.g004
Slirp-/- MEFs upon Lonp1 KD (Fig 4B), induced a significant reduction in the steady-state levels of mitochondrial transcripts (Fig 4C). Taken together, these results show that SLIRP has a role in mt-mRNA stability that can be disconnected from its function in stabilizing LRPPRC. We thus conclude that both LRPPRC and SLIRP are required for maintaining mt-mRNA steady-state levels independent of their roles in stabilizing the other partner of the LRPPRC-SLIRP complex.

Discussion

It has previously been shown that the stability of SLIRP is absolutely dependent on the presence of LRPPRC [14–18]. Here, we show that a small fraction of LRPPRC can be maintained even if SLIRP is absent. It should be noted that SLIRP is necessary for maintaining normal levels of LRPPRC, which can otherwise be degraded by mitochondrial matrix proteases such as LONP1. Beyond their roles in co-stabilization, SLIRP and LRPPRC share a common direct role on mitochondrial transcript stability as we have shown that both proteins are required to maintain mt-mRNA steady-state levels. Indeed, mt-mRNA stability could not be restored by the sole rescue of LRPPRC levels in the absence of SLIRP. An alternative hypothesis is that the rescued LRPPRC is not fully functional in the absence of SLIRP and can therefore not fulfill its mt-mRNA stabilizing function. This is however very unlikely as we have shown that low levels of LRPPRC, independent of the LRPPRC-SLIRP complex, are sufficient, and therefore functional, for mt-mRNA poly(A) tail maintenance.

Interestingly we show that in contrast to its partner LRPPRC, SLIRP is not involved in the maintenance of the poly(A) tails of mt-mRNAs in vivo. This result was surprising given the reduced poly(A) tail abundance and subsequent accumulation of mt-mRNA oligo(A) tails reported upon SLIRP KD in cells [15], but was in line with the observation by the same authors that LRPPRC alone could stimulate mt-mRNA polyadenylation in vitro and that SLIRP only had a supportive role in this assay through the stabilization of LRPPRC [15]. However this last observation contrasts with an in vitro study showing that the extension of the poly(A) tail was enhanced when LRPPRC was complexed with SLIRP, compared to LRPPRC alone [27]. This is likely because recombinant PPR proteins can be unstable and prone to precipitation [34], and LRPPRC would require SLIRP for its in vitro stability and thereby would enhance its intrinsic activity required for poly(A) tail maintenance.

We find that SLIRP is not involved in poly(A) tail maintenance in vivo, but instead has a role in fine-tuning the rate of mitochondrial protein synthesis. Indeed, we have shown using RNAseq that SLIRP can globally facilitate the ordered association of mature mt-mRNAs with mitoribosome components, thereby affecting the rate of translation. The only exception was Nduf mRNA, whose stability required the presence of LRPPRC and SLIRP but whose engagement with the mitoribosome seemed in contrast to be unaffected by the loss of SLIRP. Whether this independence is conferred by the absence of a poly(A) tail is not clear. Furthermore, as mentioned above, the residual levels of LRPPRC are sufficient to stabilize the poly(A) tails of mt-mRNAs, enabling normal protein synthesis in the heart but not in the liver and kidney. This effect is independent of the interaction between LRPPRC and SLIRP as we have confirmed that this complex is present both in mouse heart and liver mitochondria. The apparent tissue-specific effect of SLIRP loss on mitochondrial translation could be explained by an unknown mechanism compensating for the absence of SLIRP in the heart. We found that the translation rate was also impaired in the Slirp-/- MEFs and we therefore hypothesize that the consequence of SLIRP loss could be linked to the proliferative status of the tissue. Indeed, faster rates of mitochondrial translation may be required in proliferating cells such as hepatocytes and MEFs.

The role of SLIRP in maintaining the translation rate would therefore be better illustrated in...
proliferative tissues, where its absence would confer a more obvious disadvantage. However, irrespective of the effects on protein synthesis, polypeptides are made in sufficient amounts for proper assembly of the OXPHOS complexes and SLIRP loss does not compromise mitochondrial respiration. The observation that a moderate decrease in mitochondrial translation does not lead to a reduced abundance of the steady state levels of mitochondria-encoded RC subunits is likely due to the stability of the RC subunits.

Notably, mice lacking SLIRP are apparently healthy, with the exception of a slight weight loss, despite having a profound (50–70%) depletion of mt-mRNAs. The levels of the mtDNA-encoded mt-mRNAs are thus present in great excess under normal physiological conditions. It is interesting to speculate that the excess of transcripts would enable robust and rapid activation of mitochondrial protein synthesis in response to sudden changes in metabolic demand. Our findings also indicate that excess mt-mRNAs could provide a buffer that can cope with dramatic reduction of transcription of mtDNA-encoded genes, as might occur when mtDNA undergoes replication in rapidly dividing cells.

Materials and Methods

Generation of Slirp knockout mice

The targeting vector for disruption of Slirp in ES cells (derived from C57BL/6N mice) was generated by using BAC clones from the C57BL/6J RPCI-23 BAC library by Taconic Artemis. To generate a conditional Slirp knockout (KO) allele, exon 2 of the Slirp locus was flanked by loxP sites. The puromycin resistance marker (PuroR) was flanked by F3 sites and inserted into intron 2. The puromycin resistance cassette was removed by mating Slirp+/loxP-Puro mice with transgenic mice ubiquitously expressing Flp recombinase. The resulting Slirp+/loxP mice were then backcrossed to C57BL/6N mice for seven generations and intercrosses were used to generate wild-type (WT, Slirp+/+) and Slirp homozygous KO (Slirp−/−) animals.

Ethics statement

This study was performed in strict accordance with the recommendations and guidelines of the Federation of European Laboratory Animal Science Associations (FELASA). The protocol was approved by the “Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen”.

The mice were housed in specific pathogen-free conditions with a 12 hr light-dark cycle and had free access to water and food. Phenotypical characterization was performed at the German Mouse Clinic on 15 Slirp+/+ and 15 Slirp−/− mice of each gender aged between 9 and 21 weeks. At the German Mouse Clinic, the mice were maintained according to the GMC housing conditions and German laws and the tests were performed as outlined in the standard operating procedures (SOP) linked to the EMPReSS website http://empress.bar.mrc.ac.uk.

Mitochondrial transcript poly(A) tail length measurement (MPAT assay)

The MPAT assay was adapted from previous protocols [15,35]. RNA was extracted with TRIzol Reagent (Invitrogen) from heart mitochondria. An adaptor DNA oligonucleotide (sequence in S1 Table) was phosphorylated by the T4 PNK (New England Biolabs) and 2.5 pmol of this phosphorylated adaptor DNA oligonucleotide was then ligated to the 3' end of 0.3 μg of total mitochondrial RNA, for each RNA species to be tested. The ligation reaction was performed using the T4 RNA ligase (New England Biolabs) for 2 hrs at 37°C. The ligated RNA was
extracted with the TRIzol Reagent and reverse transcribed using the High capacity cDNA reverse transcription kit (Applied Biosystems) and a primer specific of the adaptor DNA oligonucleotide sequence (anti-adaptor, sequence in S1 Table). A first round of PCR was carried out for 29 cycles using a gene-specific upper primer and the anti-adaptor primer. The PCR products were purified on G-50 micro columns (GE Healthcare) in order to remove the primers. A nested PCR was then carried out for 12 cycles using a gene-specific lower primer and an inner anti-adaptor primer (sequence in S1 Table), in order to improve specificity. The PCR products were then cloned into the pCR4-TUPO vector and transformed into chemically-competent bacteria. Finally, DNA from selected colonies was extracted and sequenced in order to assess the length of the poly(A) stretch on the 3′ end of each mt-mRNA species.

**Transcription and translation assays**

In *in organello* transcription and translation assays were performed on mitochondria isolated from mouse tissues by differential centrifugation as detailed in the S1 Text. Mitochondria, 800 μg, were collected for each *in organello* transcription assay and washed in 1 ml of transcription buffer (10 mM Tris pH 7.4, 25 mM sucrose, 75 mM sorbitol, 100 mM KCl, 10 mM K2HPO4, 50 μM EDTA, 5 mM MgCl2, 10 mM glutamate, 2.5 mM malate, 1 mg/ml RSA and 1 mM ADP). An aliquot of mitochondria was collected for immunoblotting with the VDAC antibody (Millipore) to ensure equal loading. The remaining mitochondria were pelleted by centrifugation at 10,000 g for 3 min at 4°C, suspended in 750 μl of transcription buffer supplemented with 30 μCi of [32P]-UTP (PerkinElmer) and incubated for 20 min at 37°C. After the incubation, mitochondria were washed once and suspended in 750 μl of fresh transcription buffer in the presence of 0.2 mM of cold UTP. A short chase was performed for 5 min at 37°C in order to decrease the background and mitochondria were washed three times in 10 mM Tris pH 6.8, 0.15 mM MgCl2 and 10% glycerol. The mitochondrial pellet was suspended in 1 ml TRIzol (Invitrogen) for RNA extraction according to the manufacturer’s instructions. The isolated RNAs were analyzed by northern blotting and the radiolabeled transcripts were visualized by autoradiography.

Mitochondria, 500 μg, were collected for *in organello* translation assays and incubated in 750 μl translation buffer (100 mM mannitol, 10 mM sodium succinate, 80 mM KCl, 5 mM MgCl2, 1 mM KPi, 25 mM HEPES pH 7.4, 5 mM ATP, 20 μM GTP, 6 mM creatine phosphate, 66 μg/ml creatine kinase and 60 μg/ml of all amino acids except methionine). An aliquot of the mitochondrial preparation was set aside for immunoblotting to ensure equal loading as described above. Mitochondria were supplemented with 150 μCi of [35S]methionine (PerkinElmer) for 10, 30 or 60 min at 37°C. After labeling, mitochondria were washed in translation buffer and suspended in a SDS-PAGE loading buffer. Translation products were resolved by SDS-PAGE and analyzed by autoradiography.

The mitochondrial translation rate was also assessed in cultured primary MEFs following a previously described method [36]. The translation products were labeled for 10, 30 and 60 min with 80 μCi/ml of a mixture of [35S]-methionine or [35S]-cysteine (Perkin Elmer) in DMEM lacking methionine and cysteine and in the presence of 100 μg/ml of the cytoplasmic translation inhibitor emetine (Sigma). After pulse labeling, a short chase was performed for 5 min at 37°C to decrease the background. The cells were then washed and lysed with RIPA lysing buffer. Protein concentration was measured and 50 μg of total cell extracts were resolved by SDS-PAGE and analyzed by autoradiography.
Cell culture and RNAi

Knockdown of the mitochondrial LONP1 protease was performed on Slirp<sup>+/-</sup> and Slirp<sup>-/-</sup> primary MEFs plated on 10 cm diameter dishes. MEFs at 80% confluence were transfected with 1.4 μg of either scrambled or Lonp1 siRNA (Stealth siRNA negative control, Med. GC and Stealth siRNA Lonp1 respectively, Life technologies) in 12 μl of Lipofectamine RNAi Max (Invitrogen) per dish. Cells were harvested after 72 hrs either in TRIzol Reagent (Invitrogen) for RNA extraction or in RIPA lysis buffer [37] for total cell protein extraction.

Northern blotting and qRT-PCR

For detection of mitochondrial RNAs, northern blotting and qRT-PCR were performed as described in the S1 Text.

Immunoblotting

Protein steady-state levels were assessed by immunoblotting as described in the S1 Text.

Sucrose density gradients

Heart and liver mitochondria, 1.2 mg, were lysed in the presence of 1% n-Dodecyl β-D-maltoside (Sigma). Lysates were loaded on 10–30% sucrose gradients and separated by centrifugation overnight as previously described [17,38]. Gradient fractions were collected as 750 μl aliquots.

RNA was extracted from one third of each fraction by using the TRIzol LS Reagent (Invitrogen) according to the manufacturer’s recommendations. The samples were subsequently treated with DNase I and used for cDNA synthesis. The transcript abundance in each fraction was assessed by qRT-PCR analysis using the Taqman probes listed in S1 Table. The remaining two thirds of each fraction (500 μl) were precipitated with trichloracetic acid, resolved by SDS-PAGE and ribosome-containing fractions were detected by immunoblotting using antibodies specific for individual proteins from the 28S (MRPS35, Proteintech) and 39S (MRPL37, Sigma) ribosomal subunits.

RNAseq

RNA from mitochondrial sucrose gradient fractions 6 and 9–12 was isolated and the concentration, purity, and integrity were confirmed using a BioAnalyzer. The libraries were constructed using the Illumina TruSeq Sample Prep Kit and deep sequencing of the mitochondrial RNAs was performed by the Cologne Center for Genomics at the University of Cologne on an Illumina MiSeq according to the manufacturer’s instructions. Raw sequencing reads were aligned to the mouse mitochondrial genome (chrM; mm10) with Bowtie2 v2.2.4 (-p 20—very-sensitive) [39]. Gene-specific read counts were summarised with featureCounts [40] from the Subread package v1.3.5-p4 (-b-T 20-s 2) using the Ensembl 78 (GENCODE VM4) annotation, modified to merge Nd4L/Nd4 and Atp8/Atp6 annotations into bicistronic transcripts. Raw fragment counts were normalised as fragments per kilobase per million mapped reads (FPKM) and expression changes calculated as log2 fold changes of FPKM values. Heat maps were made with gplots v2.15.0 and transcripts were hierarchically clustered by complete linkage of their euclidean distances with the hclust and dist functions of the base stats package in R v3.1.2.

Respiratory chain function and complex activity

The mitochondrial oxygen consumption flux and the respiratory chain complex activities were measured as described in the S1 Text and in previous works [41].
Supporting Information

S1 Fig. Whole body homozygous knockout of Slirp. (A) Subcellular localization of the endogenous SLIRP protein in 143B cells. (B) Targeting strategy for the disruption of the Slirp gene in mice. A set of primers (orange) was designed to detect the wild-type (WT, +/+ allele and another set of primers (blue) was designed to detect the knockout (KO, −/− allele generated after Cre recombination. (C) PCR analysis of a tail biopsy from 3-week old mice. WT pups showed a lack of DNA amplification with the KO allele primer set (blue) whereas Slirp homozygous KO pups showed a lack of DNA amplification with the WT allele primer set (orange). In the Slirp heterozygous KO mice (+/−), both the WT and the KO allele primer pairs generated bands at 383 bp and ~700bp, respectively: (D) Body weight of WT (Slirp+/+, white bars) and homozygous Slirp KO (Slirp−/−, black bars) mice measured in females and males at 11 weeks of age, n = 15. Error bars represent SEM. * p value < 0.05. (E) LRPPRC-Flag co-immunoprecipitation from heart and liver mitochondria of mice expressing the recombinant LRPPRC-Flag on a Slirp WT (Slirp+/+) or KO (Slirp−/−) background, followed by immunoblots against LRPPRC and SLIRP.

S2 Fig. SLIRP absence affects mitochondrial mRNA stability. (A) MiDNA quantification by qPCR in heart, liver and kidney samples from 12-week old wild-type (WT, Slirp+/+) and Slirp homozygous knockout (KO, Slirp−/−) mice. n = 5, error bars represent SEM. (B) Mitochondrial transcript steady-state levels assessed by northern blotting in hearts from 12-week old WT (+/+) and homozygous Slirp KO (−/−) mice, as well as in hearts from 12-week old Lrpprc control (p/p) and conditional KO (p/p, cre) mice. (C) Mitochondrial transcript steady-state levels assessed by qRT-PCR in liver samples from 12-week old Slirp+/+ and Slirp−/− mice. n = 5, error bars represent SEM. * p value < 0.05. ** p value < 0.01. *** p value < 0.001. (D) Distribution of the length of the poly(A) tails expressed in % of the total number of clones, with oligo(A) tail ≤ 10nt and poly(A) > 10nt, in heart mitochondria from Slirp−/−, Slirp−/−, Lrpprc heterozygous KO (Lrpprc +/−) and Lrpprc conditional KO (Lrpprc p/p, cre) mice.

S3 Fig. Loss of SLIRP affects the mitochondrial ribosome profile. (A) Steady-state levels of the mitochondrial ribosomal proteins MRPL37 and MRPS35 were assessed by immunoblotting of protein extracts from heart and liver mitochondria from 12-week old Slirp+/+ and Slirp−/− mice. VDAC was used as a loading control. (B) Steady-state levels of the mitochondrial translation initiation factors mif2 and mif3 as assessed by immunoblotting on heart and liver mitochondria from 12-week old Slirp+/+ and Slirp−/− mice. VDAC was used as loading control. (C) Sedimentation profiles of transcripts and ribosomal proteins in sucrose density gradients of extracts from heart mitochondria from 12-week old Slirp+/+ and Slirp−/− mice. Individual mitochondrial transcripts were detected by qRT-PCR. The RNA abundance is expressed as a percentage of the levels in the control. The migration of the small mitochondrial ribosomal subunit (28S) the large mitochondrial ribosomal subunit (39S) and the assembled mitochondrial ribosome (55S) was determined by assessing the profiles of the 12S and 16S rRNAs as well as the migration of MRPL17 protein of the large ribosomal subunit. (D) Individual mRNA sedimentation profiles from the gradient described in (C). Slirp+/+ profiles are depicted in purple and Slirp−/− profiles are depicted in orange. Individual mitochondrial mRNAs were detected by qRT-PCR and the mRNA distribution profile is shown after normalisation to controls.

S4 Fig. The absence of SLIRP does not affect respiratory chain activity. (A) Mitochondrial translation rate assessed by in organello 35S-methionine pulse labelling for 10, 30 and 60 minutes.
in isolated kidney mitochondria from 12-week old wild-type (Slirp+/+) and Slirp homozygous knockout (Slirp−/−) mice. (B) Mitochondria-encoded respiratory chain subunit stability assessed by a 60 minutes in cellulo 35S-methionine/cysteine pulse labelling followed by 7 and 22 hrs chase in Slirp+/+ and Slirp−/− primary MEFs. (C) Oxygen consumption rates of isolated heart mitochondria from 12-week old Slirp+/+ and Slirp−/− mice. Isolated mitochondria were incubated with complex I or complex II substrates. Each set of substrates was successively combined with ADP (to assess the phosphorylating respiration), oligomycin (to assess the non-phosphorylating respiration) and CCCP (to assess uncoupled respiration). n = 3. Error bars represent the SEM. (D) The activity of the respiratory chain complexes I (CI), II (CII) and IV (CIV) in heart mitochondria from 12-week old Slirp+/+ and Slirp−/− mice. The citrate synthase activity (CS) was used as a control. n = 3. Error bars represent the SEM.

TIFF

S1 Table. List of primers and probes used in the study.

PDF

S1 Text. Supporting information. This section contains the supporting methods and the appendix.

DOCX

Acknowledgments

The authors would like to thank Prof. Zofia Chrzanowska-Lightowlers for technical suggestions and critical input. RNA library construction and sequencing were performed at the Cologne Center for Genomics. The German Mouse Clinic Consortium members performed the phenotyping of the mice and analyzed the corresponding data.

Author Contributions

Conceived and designed the experiments: ML AF NGL. Performed the experiments: ML AM HJL HS TW CK ESR EM JDB SS EK AF. Analyzed the data: ML AM HJL HS TW CK ESR EM JDB SS AF. Wrote the paper: ML AF NGL.

References


SLIRP regulates the rate of mitochondrial protein synthesis and protects LRPPRC from degradation

Marie Lagouge, Arnaud Mourier, Hyun Ju Lee, Henrik Spåhr, Timothy Wai, Christian Kukat, Eduardo Silva Ramos, Elisa Motori, Jakob D. Busch, Stefan Siira, German Mouse Clinic Consortium, Elisabeth Kremmer, Aleksandra Filipovska\textsuperscript{a} and Nils-Göran Larsson\textsuperscript{a}.

Supporting Information
Supporting Methods

Bioinformatic prediction of SLIRP subcellular localization

Subcellular localization of SLIRP was predicted with the MitoProt II 1.0a4 software (http://ihg.gsf.de/ihg/mitoprot.html).

Immunocytochemistry

For the detection of endogenous SLIRP, 143B cells were seeded on coverslips, fixed with 8% paraformaldehyde, incubated with an anti-SLIRP antibody (ab51523, Abcam) and visualized using a secondary Alexa Fluor 488 goat anti-rabbit antibody (LifeTech). The mitochondrial staining was achieved by anti-ATPase (Abcam) visualized by a secondary Alexa Fluor 594 goat anti-mouse antibody (LifeTech). The image acquisition was performed with a Leica TCS SP5-X inverted confocal microscope (Leica Microsystems) using a HCX PL APO CS 100x/1.46 oil objective.

For detection of endogenous LRPPRC, primary MEFs were fixed with 4% paraformaldehyde and incubated with the following primary antibodies: rabbit anti-LRPPRC (LRP-130, sc-66844 from Santa Cruz, 1:250) and mouse anti-TOM20 (WH009804M1 from Sigma, 1:500). Visualization was performed by using the Alexa Fluor 594 goat anti-mouse antibody (Invitrogen) as a secondary antibody. The image acquisition was performed as described above.

Mitochondria isolation from mouse tissues

Mitochondria from heart, liver and kidney were isolated by differential centrifugation (2 x 10 min at 1,000 g, 4°C and 1 x 10 min at 10,000 g, 4°C) in mitochondrial isolation buffer (10 mM Tris pH 7.4, 320 mM Sucrose and 1 mM EDTA).

mtDNA quantification
DNA was extracted using the ‘DNeasy Tissue and Blood’ kit (Qiagen) according to the manufacturer’s instructions. Genotyping PCRs were performed using the primers listed in S1 Table. 5ng of total DNA from heart, liver and kidney were used for detection of mtDNA and genomic DNA by qPCR with the Taqman probes listed in S1 Table.

Northern blotting and qRT-PCR

RNA for northern blotting and qRT-PCR was isolated with TRizol (Invitrogen) and suspended in nuclease-free water (Ambion). For detection of the mitochondrial RNAs by northern blotting, 2 μg of heart total RNA was denatured in NorthernMax-Gly Sample Loading Dye (Ambion) for 30 min at 50°C, separated on a agarose gel containing formaldehyde and transferred to Hybond-N+ nylon membrane (GE Healthcare). Oligonucleotides for the detection of the mtRNAs were end-labeled with 40 μCi of [32P]-ATP by the T4 Polynucleotide Kinase (New England Biolabs). Probes for the detection of mt-mRNAs and mt-tRNAs were radiolabeled with 50 μCi of α-[32P]-dCTP using the Prime-It II random primer labeling kit (Stratagene). The radiolabeled probes were then incubated for 1 hr at 42°C (tRNAs) or at 65°C (mRNAs and tRNAs) with the nylon membrane and the radioactive signal was detected by autoradiography.

For detection of the mt-mRNAs and nuclear-encoded mitochondrial-related mRNAs by qRT-PCR, total RNA was treated with DNase I using the TURBO DNA-free kit (Ambion) according to the manufacturer’s instructions. DNase-treated RNAs, 2 μg, were reverse transcribed using the high capacity cDNA reverse transcription kit (Applied Biosystems) and qRT-PCR was performed on cDNAs using the Taqman Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems) and the Taqman probes listed in S1 Table.

Immunoblotting

Immunoblotting was performed by using 50 μg of mitochondrial proteins isolated from mouse heart, liver and kidney or from total cell extracts that were resolved by SDS-PAGE and transferred to a nitrocellulose membrane. LRPPRC was detected using a polyclonal antiserum
directed against human LRPPRC. The monoclonal antibody (15C4; rat IgG2a) directed against mouse SLIRP was generated using N-6xHis-SLIRP recombinant protein produced in E.coli using standard procedures. MRPL37 and MRPS35 antibodies were purchased from Sigma and Proteintech respectively. VDAC (Millipore) was used as a loading control. The anti-LONP1 antibody was a gift from the laboratory of Thomas Langer. SUV3L1 and PNPT1 antibodies were purchased from Abcam and Proteintech respectively and an antibody against SDHA (Invitrogen) was used as a loading control.

Respiratory chain function and complex activity

The mitochondrial oxygen consumption flux was measured with an Oxygraph-2k (Oroboros Instruments) as previously described (Mourier et al., 2014) at 37°C by using 65 to 125 µg of crude mitochondria diluted in 2.1 ml of mitochondrial respiration buffer (120 mM sucrose, 50 mM KCl, 20 mM Tris-HCl, 4 mM KH₂PO₄, 2 mM MgCl₂, 1 mM EGTA, pH 7.2). The oxygen consumption rate was measured using either 10 mM pyruvate, 5 mM glutamate and 5 mM malate, or 10 mM succinate and 10 mM rotenone. Oxygen consumption was assessed in the phosphorylating state with 1 mM ADP or non-phosphorylating state by adding 2.5 µM/ml oligomycin. In the control mitochondria, the respiratory control ratio (RCR) values were >10 with pyruvate/glutamate/malate and >5 with succinate/rotenone. Respiration was uncoupled by successive addition of carbonyl cyanide m-chlorophenyl hydrazone (CCCP) up to 3 µM to reach maximal respiration.

Mitochondria, 15–50 µg, were diluted in phosphate buffer (50 mM KH₂PO₄, pH 7.4) and spectrophotometric analyses of isolated respiratory chain complex activities were performed at 37 °C by using a Hitachi UV-3600 spectrophotometer. The citrate synthase activity was measured at 412 nm (ε = 13,600 M⁻¹ cm⁻¹) after the addition of 0.1 mM acetyl-CoA, 0.5 mM oxaloacetate and 0.1 mM 5,5'-dithiobis-2-nitrobenzoic acid (DTNB). The SDH activity was measured at 600 nm (ε = 21,000 M⁻¹ cm⁻¹) after the addition of 10 mM succinate, 35 µM
dichlorophenolindophenol (DCPIP) and 1 mM KCN. The NADH dehydrogenase activity was
determined at 340 nm (E = 6, 220 M⁻¹ cm⁻¹) after addition of 0.25 mM NADH, 0.25 mM
decylubiquinone and 1 mM KCN, controlling for rotenone sensitivity. The COX activity was
measured by standard N,N,N’,N’-tetramethylethylenylene-1,4-diamine (TMPD) ascorbate assays,
controlling for KCN sensitivity. All chemicals were obtained from Sigma-Aldrich.

**Statistical analyses**

The following statistical tests were performed using GraphPad Prism software:

Fig. 1D: Unpaired t-test to compare the *Slirp⁺⁺* versus *Slirp⁻⁻* conditions and *Lrpcre p/p* versus
*Lrpcre p/p, Cre conditions. Fig. 1E: Unpaired t-test with Welch’s correction to compare the *Slirp⁺⁺* and *Slirp⁻⁻* conditions. Figs. 3D and 3E: Unpaired t-test to compare the *Slirp⁺⁺* versus *Slirp⁻⁻* conditions. Fig. 4B: Unpaired t-test to compare the *Slirp⁺⁺* versus *Slirp⁻⁻* conditions and to
compare the siCtrl and siLonp1 conditions in *Slirp⁺⁺* MEFs. Fig. 4C: One way ANOVA.
## S1 Table

### Genotyping primers

| Strp WT Fw | TGTATCTGGCAGGACACCTGG |
| Strp WT Rv | ATACTAGGAAGCCTCGCAC |
| Strp KO Fw | GAAGGAGATGACGTCACGCT |
| Strp KO Rv | CCC TCC CTG TGG GAGAOG |

### Taqman probes for mtDNA quantification

| 16S      | Mm03975671_s1 |
| Atp6     | Mm03649417_g1 |
| 18S      | Hs99999901_s1 |

### Taqman probes for mRNA level measurement

| ls2ugkb   | Mm00437762_m1 |
| Strp      | Mm01266845_m1 |
| Lpprc     | Mm0051512_m1 |
| 12S       | A689669P |
| 16S       | Mm03975671_s1 |
| Cox1      | Mm04225243_g1 |
| Cox2      | Mm03294538_g1 |
| Cox3      | Mm04225261_g1 |
| Cyb       | Mm04225271_g1 |
| Atp6      | Mm03649417_g1 |
| Ns1       | Mm04225274_s1 |
Appendix

5 German Mouse Clinic, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Neuherberg, Germany

Thure Adler1,2, Dirk H. Busch3, Antonio Aguilar-Pimentel4,5, Markus Ollert6,7, Oana Amarie8,9, Tobias Stoege10, Ali Önder Yildirim11, Oliver Eickelberg12, Lore Becker13, Alexandra Vernaeken1
5, Thomas Köpstock5,16,20,21, Marion Horsch1, Johannes Beckers1,8,19, Kristin Moreth1, Raffi Bekeredjian6, Hugo Katus6, Lillian Garrett14, Sabine M. Höfter15, Annemarie Zimpich1,16, Wolfgang Wurst8,14,15,16,17,21, Oliver Puk14, Jochen Graw8, Wolfgang Hans1, Jan Rozman1,19, Martin Klingenspor4,10, Laura Brachthäuser17, Julia Calzada-Wack17, Dirk Janik17, Tanja Klein-Rodewald1,2, Frauke Neff17, Lidikó Rácz17, Andreas Zimmer12, Birgit Rathkolb1,13,19, Eckhard Wolf12, Manuela Östereicher1, Ralph Steinkamp1, Christoph Lengger1, Holger Maier1, Claudia Stoege1, Stefanie Leuchtenberger1, Valérie Gailus-Durner1, Helmut Fuchs1, Martin Hrabě de Angelis1,8,19

1 German Mouse Clinic, Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany

2 Institute for Medical Microbiology, Immunology and Hygiene, Technical University of Munich, Trogerstrasse 9, 81675 Munich, Germany

3 Department of Dermatology and Allergy, Biederstein, Klinikum rechts der Isar, Technische Universität München (TUM), Biedersteiner Str. 29, 80802 Munich,

4 Comprehensive Pneumology Center, Institute of Lung Biology and Disease, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany and Member of the German Center for Lung Research
5 Department of Neurology, Friedrich-Baur-Institut, Ludwig-Maximilians-Universität München, Ziemssenstrasse 1a, 80336 Munich, Germany
6 Department of Cardiology, University of Heidelberg, Im Neuenheimer Feld 410, 69120 Heidelberg, Germany
7 Institute of Pathology, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany
8 Institute of Developmental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany
9 Chair for Molecular Nutritional Medicine, Technische Universität München, Else Kröner-Fresenius Center for Nutritional Medicine, 85350 Freising, Germany
10 ZIEL – Center for Nutrition and Food Sciences, Technische Universität München, 85350 Freising, Germany
11 Clinical Research Group Molecular Allergology, Center of Allergy and Environment Munich (ZAUM), Technische Universität München (TUM), and Institute for Allergy Research, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
12 Institute of Molecular Psychiatry, University of Bonn, Sigmund-Freud-Straße 25, 53127 Bonn, Germany
13 Ludwig-Maximilians-Universität München, Gene Center, Institute of Molecular Animal Breeding and Biotechnology, Feodor-Lynen Strasse 25, 81377 Munich, Germany
14 Chair of Developmental Genetics, Center of Life and Food Sciences Weihenstephan, Technische Universität München, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany
15 Max Planck Institute of Psychiatry, Kraepelinstr. 2-10, 80804 Munich, Germany
16 Deutsches Institut für Neurodegenerative Erkrankungen (DZNE) Site Munich, Schillerstrasse 44, 80336 Munich, Germany
17 Munich Cluster for Systems Neurology (SyNergy), Adolf-Butenandt-Institut, Ludwig-Maximilians-Universität München, Schillerstrasse 44, 80336 Munich, Germany
18 Chair of Experimental Genetics, Center of Life and Food Sciences Weihenstephan, Technische Universität München, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany
19 Member of German Center for Diabetes Research (DZD), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany
20 German Network for Mitochondrial Disorders (mitoNET)
21 German Center for Vertigo and Balance Disorders, Munich, Germany