Characterization of the small mitochondrial RNAs, and their role in the regulation of mito-nuclear interactions.

(Andrea Pozzi)

(M.Sc.)

A thesis submitted for the degree of *Doctor of Philosophy* at
Monash University in 2020
School of Biological Science

Copyright notice

Table of Contents

ABSTRACT

Animal cells host two different genomes, mitochondrial and nuclear, providing the possibility for inter-genomic interaction. The nuclear genome regulates most cellular function, while the central role of the mitochondrial genome is the regulation of energy production to its host. Yet, a growing number of studies have implicated the mitochondria in many other biological functions over and above cellular metabolism *per se*. Indeed, mutations within the mitochondrial DNA (mtDNA) sequence have been found to influence many phenotypes, spanning thermal tolerance capacity to the aggressiveness of certain types of cancer. Although the mechanisms used by the mitochondria to exert these phenotypic effects remain largely unknown; emerging studies have provided new insights. For example, recently, researchers identified multiple new mitochondrial proteins affecting animal phenotypes, spanning from human health to sex determination. Among these new mitochondrial products, a new type of RNAs have been identified, the "small mitochondrial RNAs".

Currently, the only known features of the small mitochondrial RNAs are their length (~30nt), and that their sequence aligns to the mitochondrial genome. These RNAs have been found in several species, including several vertebrates and one invertebrate (*Ruditapes philippinarum*), but their function is still unknown. I am interested in characterizing these RNAs, including uncovering their function. In 2017, during my master's degree thesis, I hypothesized that the function of the small mitochondrial RNAs is to regulate the expression of nuclear genes through a process known as RNA interference. In this Ph.D. thesis, I have characterized the small mitochondrial RNAs and investigated if my 2017 hypothesis was correct. The work has proceeded through three research chapters, in which I have pursued this research together with my Ph.D. supervisor, Associate Professor Damian Dowling.

Through these three years of research, we have uncovered several key insights into the small mitochondrial RNAs, and we have provided multiple lines of evidence that supports the hypothesis first formulated in 2017. The key aspects that we investigated are presented in the three chapters of

this thesis. The first key aspect we investigated is the origin of the small mitochondrial RNAs, by demonstrating that these RNAs are encoded in the mitochondrial genome (Chapter 2). The second key aspect is whether and how the existence of these small mitochondrial RNAs is likely to affect interpretations of previous research into mitochondrial genetics. We contend that the existence of these RNAs provides a unique opportunity to re-interpret the research done to investigate how mitochondrial mutations affects human health. The general public and governments of several countries are now focusing more than ever on the effect of mitochondrial mutations on health, for topics such as the mitochondrial replacement therapy. Therefore, we decided to provide an overview of how the small mitochondrial RNAs might exert effects on human health (Chapter 3). Lastly, we investigated the function and conservation of the small mitochondrial RNAs. We found that the small mitochondrial RNAs can bind Argonaute 2, a key protein of the RNA interference mechanism. Furthermore, we found that some of these small RNAs are conserved in sequence and expression across multiple species spanning the entire *Chordata* phylum. Their ability to bind Argonaute 2 and their conservation across species are features resembling the microRNAs (miRNA); suggesting that the small mitochondrial RNAs are indeed functional RNAs that act like nuclear miRNAs. For this reason, we named them “mt-miRNAs”. In conclusion we fulfilled the initial goal of this thesis, by characterizing the mt-miRNAs and providing support of their functionality. This will pave the way for future studies to explore the extent to which mt-miRNAs affect organismal function via mitochondrial-nuclear interactions.

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes 1 original paper published in a peer reviewed journal and 2 articles submitted to peer reviewed journals. The core theme of the thesis is Mitonuclear communication. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the School of Biological Sciences under the supervision of A/Prof. Damian Dowling.

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution*	Co-author(s), Monash student
Second Chapter	The Genomic Origins of Small Mitochondrial RNAs: Are They Transcribed by the Mitochondrial DNA or by Mitochondrial Pseudogenes within the Nucleus (NUMTs)?	Published: *Genome Biology and Evolution*	70%. Concept, data analysis and writing	1) Damian Dowling, input into manuscript 30%	No
Third Chapter	Mitochondria Interference: a role for small mitochondrial RNAs as mediators of mitochondrial diseases?	In Review: *Trends in Molecular Medicine*	70%. Concept, data analysis and writing	1) Damian Dowling, input into manuscript 30%	No
Fourth Chapter	A new member in the Argonaute crew: the mt-miRNAs	In Review: *Current Biology*	75%. Concept, data analysis and writing	1) Damian Dowling, input into manuscript 25%	No

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student name: *Andrea Pozzi*
Student signature: **Date:** *24/04/2020*

I hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor name: *Damian Dowling*
Main Supervisor signature: **Date:** ***24/04/2020***

ACKNOWLEDGEMENTS

I have never written any acknowledgements in my thesis, as I believe that the people involved in my research and the one supporting me in my private life already know how I feel about them.

However, I decided to make a small exception in this case.

I made this exception to thanks my PhD Advisor, Damian Dowling, as without him this thesis wouldn’t exist. In fact, working to establish a new field as a PhD student has been incredibly hard, and without its support at every step, I am not sure if I would have been able to finish this work. We faced multiple challenges, sometimes within Monash and sometimes outside, but I have been always confident to have its support while facing these challenges. For example, during the last weeks of my PhD the COVID19 pandemic started, creating many problems in both my working and private life. Nonetheless, thanks to Damian support, we still managed to submit this PhD thesis with only ~1 week of delay. For your support during these three years, I am really grateful.

The research in this thesis was supported by an Australian Government Research Training Program (RTP) Scholarship.

CHAPTER ONE

Introduction

The tiny organelle within us: the mitochondria

Mitochondria are organelles found in almost all eukaryotic organisms, which are responsible for the production of Adenosine-Triphosphate (ATP). Indeed, to date, there has only been one eukaryote, Monocercomonoides, identified that has no mitochondria, but it is known to have evolved from ancestors that had mitochondria (Karnkowska et al. 2016). Thus, the presence of mitochondria could validly be argued to be a defining feature of the eukaryote domain of life (Zimmer 2009). Mitochondria are the evolutionary outcome of a symbiosis between an alpha-proteobacteria (ancestral mitochondria) and an ancient archaeon (Zimmer 2009). Since the origin of this endosymbiosis, the mitochondrial genome has undergone a massive process of reductive genome evolution (Khachane, Timmis, and Martins dos Santos 2007; Andersson et al. 1998), becoming a genetic system specialized in producing energy.

One of the most unique features of the mitochondrial genome is its strict maternal inheritance (Gyllensten, Wharton, and Wilson 1985). Indeed, the mitochondrial genome is inherited only from mother to offspring in most animal species, including vertebrates, arthropods and molluscs (Kuijper, Lane, and Pomiankowski 2015). Although each taxon uses different ways to remove the paternal mitochondrial genome from the zygote, strict maternal inheritance is well conserved across clades. Strict maternal inheritance is evolutionary conserved even among plants. Indeed, most plant species share this feature with animals, and only a few species are known for their ability to transmit the paternal mitochondrial genome (Greiner, Sobanski, and Bock 2015). However, since the structure and evolution of mitochondrial genomes strongly differ across taxa (Saccone et al. 1999), below we will focus on the vertebrate mitochondrial genome, which has a highly conserved structure, and for which strict maternal inheritance is present in all species.

In vertebrates, the typical mitochondrial genome contains 13 protein-coding genes, and was long thought to contain a total of 37 genes, 13 of which are protein-coding, 22 tRNA and 2 rRNA (Boore 1999). However, recent years have seen the discovery of new genes that lay cryptically within the mitochondrial DNA (mtDNA) sequence. One exceptional example is the discovery of the

Humanin protein (Lee, Yen, and Cohen 2013; Capt, Passamonti, and Breton 2016). The gene coding for this protein has been found inside the human mitochondrial genome, precisely within the gene encoding the rRNA 16S, and it has been linked to the onset of several diseases, such as Alzheimer's disease (Matsuoka 2009). Another example is the protein MOTS-c, encoded in the rRNA 12S gene, which has been associated with obesity and diabetes (Lee 2015). These cryptic mitochondrial products are not limited to the human mtDNA. In fact, recent studies found that the mtDNA of a few bivalve species encodes for sex-specific proteins which might be involved in the sex determination of these species (Breton et al. 2011; Milani et al. 2014). As recently observed, the number of newly identified cryptic mitochondrial products keeps increasing, thus propelling the interest in mitochondrial biology. Furthermore, as a growing number of studies link mitochondrial variability and human health, the medical community increases its focus on mitochondrial biology (Picard, Wallace, and Burelle 2016).

Healthy mitochondria, healthy people

While in many instances mitochondrial mutations have no obvious effect on phenotypic expression, in other cases they have been consistently correlated with diseases (Majamaa et al. 1998; Gorman et al. 2016). This is probably due to the synergistic relationship of the mitochondrial proteins in the production of ATP. Indeed, any mutation within a protein-coding gene that leads to a defect in at least one mitochondrial protein, might as a consequence cause a deficiency in energy production (Gorman et al. 2016). Similarly, mutations impairing the function of ribosomal and tRNAs can affect the synthesis of the mitochondrial proteins, thus, impairing the production of ATP (Fine et al. 2019; McFarland et al. 2004; Majamaa et al. 1998). These mitochondrial mutations involved with the dysfunction of the mitochondria characterize a group of genetic disorders known as mitochondrial diseases (Ozawa et al. 1991; Wallace 1994; Gorman et al. 2016). Among the mitochondrial diseases, the Myoclonic Epilepsy and Ragged-Red Fiber Disease (MERRF) and Leber's Hereditary Optic Neuropathy (LHON) are among the most studied diseases (Brown et al. 1992; Wallace et al. 1988). The most common effect of MERRF is general epilepsy, and the mitochondrial mutation mt-tRNA Lys A8344G is present in ~90% of these cases (Shoffner and Wallace 1992). Research on the mutation A8344G showed that it produces deficiencies in the enzyme complexes of the respiratory chain, mostly involving Complex I and complex IV, consistent with the effect of a defect in translation of all proteins encoded in the mitochondrial genome (Bindoff et al. 1991; Wallace et al. 1988). The most prominent symptom of LHON is vision loss, and this disease is associated in over 90% of the cases with the mutations found in positions 11778, 3460, and 14484 of the mtDNA sequence (Riordan-Eva and Harding 1995). In particular, research on mutations in position 3460 of the mitochondrial genome revealed that this mutation is correlated with a deficiency in Complex I activity. Furthermore, a study found that rotenone sensitive NADH:ubiquinone oxidoreductase activity in mitochondria is reduced ~80% in presence of this mutation (Majander et al. 1991). Thus, demonstrating a key role of the mitochondria in human health. However, the role of the mitochondria in human health is not limited to these examples. Indeed, it has recently been discovered that other common diseases such as Type 2 diabetes are linked to mitochondrial mutations (Hudson et al. 2014).

Since the discovery of the mitochondrial genome, researchers have reported that mitochondrial mutations are correlated with a surprising number of diseases, such as type 2 diabetes and Parkinson’s disease (Feder et al. 2009; Mohlke et al. 2005; Dölle et al. 2016; Abeliovich and Gitler 2016). Multiple studies have found that different sets of mitochondrial mutations, known as haplotypes, are either over- or underrepresented among individuals affected by type 2 diabetes, such as the haplotype J2 in Finnish populations (Marom, Friger, and Mishmar 2017; Feder et al. 2009; Mohlke et al. 2005). Similar effects have been observed in studies of Parkinson's disease. Authors found that neurons from patients with Parkinson's disease are not able to maintain a stable number of wild-type mtDNA copies, thus the number of wild-type mtDNA copies decreases, leading to the malfunction of the mitochondria (Dölle et al. 2016). These studies highlight a key role of the mitochondria in dictating trajectories of human health, over and above the role of mtDNA mutations in causing mitochondrial diseases. Indeed, a groundswell of research attention investigating the role of the mitochondria in regulating the maintenance of human health (Levin et al. 2013) has in turn helped motivate further research into novel treatments to prevent the progression of mitochondrial diseases, leading to the development of an early stage therapy known as mitochondrial replacement therapy.

Mitochondrial replacement therapy is a technique to replace the mtDNA in an oocyte of a female affected by a mitochondrial disease, caused by a mitochondrial mutation, with the mtDNA from the oocyte of a healthy donor (Craven et al. 2010; Tachibana et al. 2009). This is a relatively new technique, developed in the last 15 years, which is now legal in the United Kingdom and probably soon will be legal in other countries such as Singapore, USA, and Australia (Newson et al. 2019). However, pairing the parental nuclear DNA with the mtDNA of a third-party donor is not without challenges. Indeed, the technique invariably results in the carryover (typically around 2%) of a small amount of the mother's mutant mtDNA, and sometimes, this mutated mtDNA can increase in frequency and be expected to become dominant again post therapy (Craven et al. 2010). The factors that drive changes in the frequency of donor to mother haplotype frequency across the life course are not well understood. Furthermore, other researchers have reported evidence that the application of this technique to human cells may have a negative effect on their fitness (Dobler et al. 2018). Based on a risk analysis reconciling the outcome of previous studies, and using the most conservative of assumptions possible, the study reported that negative effects on the offspring's health would be expected to occur at the very least in one of 130 new-born children that are born to the therapy. Despite these difficulties, the interest for such therapy is high, leading several governments to legislate in their favour. The interest of the scientific community and the general public on the mitochondrial replacement therapy provides an example on how mitochondrial biology is becoming more relevant for human health.

The role of mitochondrial genomics in animal evolution

Research leveraging the study of the mtDNA has held a very prominent place in the field of evolutionary biology for over four decades, spanning studies of phylogeny to animal adaptation. In phylogenetic studies, the mtDNA has been used as a ‘neutral marker’ for the estimation of numerous evolutionary inferences (Cann, Stoneking, and Wilson 1987), and it is still used to investigate evolutionary relationships across close species or populations (Rubinoff and Holland 2005). There are multiple reasons why the mtDNA was embraced as the marker of choice in phylogenetic studies (Galtier et al. 2009). The first reason is that the mtDNA supposedly has no recombination, thus is inherited from only one parent allowing an easy reconstruction of one parental lineage. Another reason is that mitochondrial genes are very conserved across species, having only few duplications and introns (Gissi, Iannelli, and Pesole 2008). Furthermore, the mtDNA has multiple polymorphisms in natural populations due to its high mutation rate, which can be used to track recent population history (Galtier et al. 2009). A well-known example of the use of the mtDNA in phylogenetics is its application to investigating human (*Homo sapiens*) origins and evolution. The most famous case is the “mitochondrial eve”, referring to the study that dated the origin of the modern mtDNA haplotypes to ~200kya (Ingman et al. 2000; Cann, Stoneking, and Wilson 1987), sparking interest in both the scientific community and the general public.

Similarly, by studying the relationship between mtDNA variation and animal adaptation, mitochondrial biology greatly affected other branches of science, such as evolutionary biology. In fact, it was about two decades ago when researchers first found that natural selection influenced the mtDNA variation in human populations (Mishmar et al. 2003). In this study, the authors found that specific genes in the human mtDNA had different rates of synonymous/non-synonymous mutations depending on the environment considered, indicating that these genes might have been adapting to that specific environment. This study inspired many others to investigate if mtDNA variability might be shaped by the environment. Arguably, one of the most relevant types of adaptation to the environment is thermal tolerance. Thermal tolerance has been found to be affected by mitochondrial variation both in the wild, and in the lab, in at least one species, *Drosophila melanogaster*, suggesting this phenomenon might be conserved across more species (Camus et al. 2017; Lajbner et al. 2018). Indeed, researchers tested multiple combinations of mtDNA and nuclear genome to discover that the mtDNA is able to affect thermal tolerance regardless of the nuclear genome (Harada, Healy, and Burton 2019; Lajbner et al. 2018). As we approach a climate crisis ("Act Now and Avert a Climate Crisis" 2019), discovering more about the evolutionary capacity for organisms to respond to rapid climatic changes, such as increasing temperatures, is relevant for society, despite not leading to immediate technological advancement. In fact, through studies that probe the evolutionary capacity of organisms and ecosystems in response to climatic changes, we come closer to being able to predict whether specific taxa and ecosystems will be able to sustain and responds to the fast-pace changes they are currently experiencing, or whether they will need human support. Another adaptive trait affected by mtDNA variation, arguably with more immediate effect, is fertility. Multiple studies on multiple organisms have found that mtDNA variation has effects on fertility, spanning from insects to humans, suggesting a very conserved mechanism underlying this phenomenon (Nakada et al. 2006; Ruiz-Pesini et al. 2000; Yee et al. 2015; Camus et al. 2015). For example, by studying a Spanish cohort of men, researchers reported that the mt-haplotype T is very common across infertile men, suggesting a role for the mutations present in this haplotype in male infertility (Ruiz-Pesini et al. 2000). Another example comes from a recent study where the authors found that sterile mice harbouring mutated mtDNA in heteroplasmy with a wild type mtDNA molecule, recover their fertility when the absolute number of unmutated mtDNA copies rise above a specific threshold, highlighting the importance of mitochondrial mutations in male fertility (Jiang et al. 2017). However, while mtDNA mutations have been consistently shown to influence phenotypes such as male fertility, the type and magnitude of these effects are variable, as the effects are dependent on the interactions between mitochondrial and nuclear genomes.

Mitonuclear communication

Often the effect of mitochondrial mutations depends on the interaction of the mutated mitochondrial product (i.e. RNA or protein) with partner nuclear products. This interaction between mitochondrial and nuclear products is part of a phenomenon called mitonuclear communication, which describes how mitochondria and nucleus interact through proteins, RNAs, and metabolites (Ballard 2005; Monaghan and Whitmarsh 2015). Negative fitness effects have been reported when communication between mitochondria and nucleus is disrupted (Wallace et al. 1988; Ballard 2005; Ballard and Pichaud 2014). A recent example is a study of the effects of mitochondrial-nuclear interactions affecting lifespan in populations of fruit fly, *Drosophila melanogaster* (Vaught et al. 2020). The authors used a breeding scheme to create genetic strains of fruit flies (*D. melanogaster*) harbouring different combinations of mitochondrial and nuclear genotypes sourced from different global populations of flies. They reported that substitution of population-specific combinations of mitochondrial and nuclear genotypes, with a novel combination, drastically modified the lifespan of the flies, with the effects of the mitonuclear substitution contingent on the sex of the flies. Another example is a study of the marine copepod *Tigriopus californicus,* where authors found that individuals sourced from distinct populations of this species exhibit signatures of reproductive isolation when mating with each other. Indeed, the researchers found that while individuals from disjunct populations were able to generate fertile offspring (F1), the second generation (F2) exhibited greatly reduced fitness (Burton, Ellison, and Harrison 2006). The researchers found that this effect was led by the disruption of mitonuclear communication due to the mismatch across the two genomes. Indeed, the reduced fitness of the F2 was recovered by backcrossing F2 females with males of the F0 maternal line, which increased the proportion of the maternally derived nuclear genome, thus restoring the communication between mitochondrial and nuclear genome (Harrison and Burton 2006). However, while it is clear that modification of the organismal mitonuclear genotype manifests in effects on key life-history phenotypes, we still do not fully grasp the mechanisms that underpin these interactions.

The mechanism behind mitonuclear communication

The study of mitonuclear communication is in its early stages, however, studies investigating the molecular mechanisms of interaction between the two genomes are quickly growing. The best-known example of mitonuclear communication are the proteins that compose the Electron transport chain (ETC), a series of proteins complexes formed by a few mitochondrial proteins and multiple nuclear proteins. The ETC is necessary for complex eukaryotic life, as it provides energy for the cell. Thus, any disruption in the ETC would be fatal, and both genomes must keep "collaborating" for the survival of the cell. The mitonuclear communication requires synergic action between the two genomes, which affect their evolution (Blier, Dufresne, and Burton 2001). A hypothesis about the effects of a synergistic action between mitochondrial and nuclear genomes involves nuclear proteins and mitochondrial rRNAs. This hypothesis proposes that nuclear proteins interacting with mitochondrial products, such as rRNAs, should evolve faster to match the faster evolution rate present in the mtDNA (Blier, Dufresne, and Burton 2001). This hypothesis has been supported by several recent studies. For example, a one study found that nuclear genes encoding proteins that form part of the mitochondrial ribosome exhibit an enhanced mutation rate when compared to ribosomal proteins part of the nuclear ribosome, supposedly to match the higher mutation rate present in the mt-rRNAs (Barreto and Burton 2013). The study reported a similarly enhanced mutation rate across different arthropods, suggesting that this mechanism of interaction between the two genomes is conserved across species. Nonetheless, this is not the only proposed mechanism for mitonuclear communication.

In 2017, during my Master degree, I proposed an alternative mechanism for mitonuclear communication, which has the small mitochondrial RNAs at its core (Pozzi et al. 2017). These small mitochondrial RNAs are RNAs between 20 and 35 nucleotides in length that are encoded by the mitochondrial genome, which have been discovered in 2011 as a side finding during the comprehensive study on the human mitochondrial transcriptome (Mercer et al. 2011). However, it was in my first published work, where it was first proposed that these RNAs may form a key role in mitonuclear communication through RNA interference (Pozzi et al. 2017). RNA interference is an endogenous mechanism of post-transcriptional regulation that reduces the expression levels of the target gene. This down-regulation is possible thanks to a protein complex known as RISC (RNA-Induced Silencing Complex), which uses specific small non-coding RNA molecules to recognize the untranslated region (UTR) of a target messenger RNA (mRNA), and then block the translation of the mRNA (Ambros 2004; Cloonan 2015). The best studied group of small non-coding RNA able to use this mechanism of gene silencing are the microRNAs, miRNA; (Ambros 2001). The biological role of RNA interference, and the small non-coding RNA involved in achieving gene silencing, cannot be overstated. This is reflected by more than 150,000 papers published on this subject. Indeed, while the first studies of miRNA came through the discovery of their fundamental role in development in *C. elegans* (Fire et al. 1998), they have now been studied in the context of virtually every human disease (Picard, Wallace, and Burelle 2016); generating a wealth of publicly available data of small RNA-sequencing datasets . My aim is to leverage the data accumulated via the study of these small RNAs in RNA interference, to investigate whether the small mitochondrial RNAs can interact with nuclear mRNAs using this process, thus probing a putatively new pathway of communication between the two genomes.

Thesis aim

The broad goal of my PhD thesis is to understand whether previously identified small mitochondrial RNAs are functional, or mere residues of longer RNAs. Evidence for functionality and a role in RNA interference could change the way that biologists understand mtDNA variation. In this Ph.D. thesis, we characterize the small mitochondrial RNAs to the extent allowed within the framework of a 3-year Ph.D. study.

First Study Aim – Investigate the origin of small mitochondrial RNAs

Following the first discovery of the small mitochondrial RNAs (Mercer et al. 2011), several studies reported the presence of these RNAs across multiple species. In fact, up to the start of my Ph.D. in 2017, small mitochondrial RNAs had been identified in human, mouse, chicken, rainbow trout and killifish (Ro et al. 2013; Riggs et al. 2018; Bottje et al. 2017; Ma et al. 2016; Mercer et al. 2011; Pozzi et al. 2017). Nonetheless, none of these studies had performed a comprehensive investigation of the origin of these small RNAs. Indeed, although the small mitochondrial RNAs align to the mtDNA, parts of the mtDNA can be found inside the nuclear genome. These parts of mtDNA incorporated in the nuclear genome are known as NUMTs (Mishmar et al. 2004). Thus, there is the possibility that the small mitochondrial RNAs, might be expressed by NUMTs instead of by the mtDNA. This possibility undermines the very core of the hypothesis, which sees the small mitochondrial RNAs as an important regulator of mitonuclear communication. Therefore, investigating the origin of these RNAs is a priority, and the first study of my Ph.D. thesis.

In this study, we aimed to test the hypothesis that the small mitochondrial RNAs are transcribed directly by the mtDNA rather than by NUMTs resident in the nuclear genome. In particular, we focused on testing two predictions associated with the hypothesis. The first prediction we tested is that the expression of the small mitochondrial RNAs would not be correlated with the amount of NUMTs. To investigate this prediction, we compared the expression of these RNAs across six species of vertebrates, each characterised by different amounts of NUMT sequence, and across five tissues. Through this comparative analysis, we found that the expression of small mitochondrial RNAs is not correlated with the amount of NUMTs. In fact, the expression levels of small mitochondrial RNAs in species harbouring only small amount of NUMTs in their nuclear genome, such as *Gallus gallus* with ~1kb of NUMTs, is similar to species with much greater amounts of NUMTs, such as *Monodelphis domestica* that has 2000kb of NUMTs. Furthermore, to verify the presence of tissue-specific transcription from the NUMTs, we analysed tissues from different organs in each species, finding a consistent lack of correlation between levels of expression of small mitochondrial RNAs expression and NUMTs across these tissues.

The second prediction we tested is that the expression of the small mitochondrial RNAs and the mtDNA copy number would be correlated with each other. To verify this prediction, we compared tissues known to exhibit striking differences in the levels of mtDNA copy number, such as brain tissues (high copy number) and epithelium (low copy number), finding much higher expression of small mitochondrial RNAs in the brain. Thus, this comparison supports a correlation between the expression of small mitochondrial RNAs and mtDNA copy number. To further validate this prediction, we compared levels of the small mitochondrial RNAs expression across samples of the same human tissue (bladder cells), but across two different states (healthy versus cancer). Indeed, as we know that cancerous bladder cells samples are enriched in mtDNA, we predicted that these cells would be associated with higher expression levels of small mitochondrial RNAs. This is indeed what we found.

In conclusion, by compiling evidence that the expression of small mitochondrial RNAs is uncorrelated to levels of NUMT sequence, but closely correlated to tissues-specific levels of mtDNA copy number, we provided the first strong evidence for the hypothesis that the small mitochondrial RNAs are directly transcribed from the mtDNA sequence. The full article is already published in *Genome Biology and Evolution* as a Research Article.

Second Study Aim – A new hypothesis: Mitochondrial Interference

Interest in mitochondrial biology is rising in medicine, and advances in understanding the biology of, and the treatment, of mitochondrial diseases are happening at increasing speed (Picard, Wallace, and Burelle 2016). Most of this research is focused squarely on the effects of mtDNA mutations on cellular metabolism and energy production. However, due my focus on small RNAs, I developed a different perspective, which I describe in this study.

The aim of this study is to inspire researchers working in the field of mitochondrial biology, by providing a new and broader framework for the interpretation of how mitochondrial mutations may exert their effects on disease phenotypes in humans; a framework that considers the existence of the small mitochondrial RNAs.

This study integrates current knowledge on the underpinning mechanisms linking mtDNA mutations to mitochondrial disease, highlighting intriguing discrepancies between the predicted links and results of many studies. We then leverage these discrepancies to outline a new framework for studies seeking to validate links between mitochondrial mutations and effects on human health. We introduce the *mitochondrial interference hypothesis* and outline its supporting evidence. Briefly, mitochondrial interference describes a mechanism in which small mitochondrial RNAs may inhibit nuclear gene expression through RNA interference. In this second study of my Ph.D., we describe several studies supporting the presence of this mechanism and argue how this mechanism better fits the results of several studies, both historical and recent, on mitochondrial diseases. This work is now under consideration in *Trends in Molecular Medicine* as an Opinion Article.

Third Study Aim – Investigate the function of small mitochondrial RNAs

As described previously, mitochondrial mutations have been linked with a manifold number of diverse phenotypes, and in the work preceding my Ph.D. I have proposed the small mitochondrial RNAs as a mediator of these effects (Pozzi et al. 2017). The proposed mechanism is RNA interference, however, to my knowledge, there is no direct evidence demonstrating that small mitochondrial RNAs are part of this molecular mechanism. Nonetheless, the small mitochondrial RNAs have similar length to other small RNAs involved in RNA interference (about 30 nucleotides) and many proteins related to RNA interference are co-localized with the mitochondria, thus supporting the hypothesis that the small mitochondrial RNAs might be involved in this process (Ha and Kim 2014; Nishimasu et al. 2012; Mercer et al. 2011). Among these proteins, Argonaute 2 (Ago2) is the best candidate to interact with these RNAs. Ago2 is the best known protein in RNA interference, and its role is to use a miRNA to bind an mRNA and block its translation (Ha and Kim 2014). Furthermore, this protein has been found to co-localize with the mitochondria, thus providing a connection between Ago2 and the mitochondria. Although the reasons for colocalization of Ago2 to the mitochondria remain unknown, the ability of Ago2 to directly bind small RNAs and its connection to the mitochondria makes it a likely partner for the small mitochondrial RNAs.

Investigating the ability of the small mitochondrial RNAs to bind Ago2 is facilitated by the wealth of data accumulated in the last two decades on the nuclear miRNAs (Fire et al. 1998). Indeed, Ago2 is one of the first proteins discovered in the RNA interference mechanism and multiple studies on its role in cell regulation make for a wealth of publicly available datasets of small RNA sequencing libraries that are freely accessible (Ambros 2004; Ha and Kim 2014). The accessible data mostly consists of studies where Ago2 has been isolated while bound to the RNAs interacting with this protein, including the small mitochondrial RNAs. These studies have been performed on multiple species and tissues, however, the authors discarded the sequences aligning to the mitochondria in their original works. We therefore aim to re-purpose the original sequencing data from these studies and investigate the ability of Ago2 to bind these RNAs. Furthermore, we aim to identify how much the small mitochondrial RNAs have in common with nuclear RNAs, thus validating their role in cell regulation.

This chapter can be divided into three main sections. In the first section of the chapter, we investigate whether the small mitochondrial RNAs can bind Ago2, the main protein of the RNA interference. We leveraged several published datasets online, where the RNAs sequenced were obtained through the immunoprecipitation with Ago2. We were able to find multiple studies in mice and humans where the small mitochondrial RNAs can bind Ago2. Our results are quite reliable, as the small mitochondrial RNAs expression was consistently higher in Ago2-immunoprecipitation treatments compared the mock-immunoprecipitation (only IgG) treatments across several cell lines and multiple biological replicates, thus providing the first evidence of that the small mitochondrial RNAs may play a hitherto unrecognised role in RNA interference.

In the second section of the chapter, we investigate the conservation of expression and sequence of the small mitochondrial RNAs across several key model organisms. We compared the expression of the small mitochondrial RNA encoded within one particular RNA that we identified (our previous analysis had verified its ability to bind Ago2 across multiple cell lines), which we name the mt-tRNA Met because it sits within the mitochondrial tRNA for Methionine. By analysing the expression pattern of this small RNA across several key vertebrates, from *Homo sapiens* to *Danio rerio*, we confirmed its presence in all the *Chordata* phylum. Furthermore, by comparing the sequence of this small RNA to the sequence covering the rest of the mt-tRNA Met gene, we found that the region encoding the small RNA is more conserved. Therefore, these analyses identified at least one *bone fide* small mitochondrial RNA with miRNA-like features.

In the last section of the chapter, we perform several analyses to identify targets of this small mitochondrial RNA that could be linked to human health. Through a combination of computational and experimental techniques, we present strong evidence that the mitochondrial small RNA encoded in the mt-tRNA Met can bind a gene called CASP8 And FADD Like Apoptosis Regulator (CFLAR). The small RNA binds a region at the edge of the 3' untranslated region (UTR) of CFLAR. Interestingly, the coding region of this mRNA is only ~2Kb, while the UTR is ~12Kb. The CFLAR isoform having this ~12Kb UTR is a promoter of apoptosis, thus suggesting that this small RNA is involved in apoptosis regulation. Moreover, we found that this ~12Kb UTR is present but not expressed in other primates. In fact, while the ~12Kb UTR can be found in the same genomic region, the other primates analysed have the UTR at least 1Kb distant from the CFLAR gene, indicating that this region does not serve as a binding site for this small mitochondrial RNA in non-human primates. Thus, these analyses suggest the presence of a human-specific pathway of regulation led by the small mitochondrial RNAs. This work is now under review in *Current Biology* as a Research Article.

References

Abeliovich, Asa, and Aaron D. Gitler. 2016. “Defects in Trafficking Bridge Parkinson’s Disease Pathology and Genetics.” *Nature* 539 (7628): 207–16.

“Act Now and Avert a Climate Crisis.” 2019. *Nature* 573 (7774): 309.

Ambros, V. 2001. “microRNAs: Tiny Regulators with Great Potential.” *Cell* 107 (7): 823–26.

Ambros, Victor. 2004. “The Functions of Animal microRNAs.” *Nature* 431 (7006): 350–55.

Andersson, S. G., A. Zomorodipour, J. O. Andersson, T. Sicheritz-Pontén, U. C. Alsmark, R. M. Podowski, A. K. Näslund, A. S. Eriksson, H. H. Winkler, and C. G. Kurland. 1998. “The Genome Sequence of Rickettsia Prowazekii and the Origin of Mitochondria.” *Nature* 396 (6707): 133–40.

Ballard, J. William O. 2005. “Drosophila Simulans as a Novel Model for Studying Mitochondrial Metabolism and Aging.” *Experimental Gerontology* 40 (10): 763–73.

Ballard, J. William O., and Nicolas Pichaud. 2014. “Mitochondrial DNA: More than an Evolutionary Bystander.” Edited by Charles Fox. *Functional Ecology* 28 (1): 218–31.

Barreto, Felipe S., and Ronald S. Burton. 2013. “Evidence for Compensatory Evolution of Ribosomal Proteins in Response to Rapid Divergence of Mitochondrial rRNA.” *Molecular Biology and Evolution* 30 (2): 310–14.

Bindoff, L. A., C. Desnuelle, M. A. Birch-Machin, J. F. Pellissier, G. Serratrice, C. Dravet, M. Bureau, N. Howell, and D. M. Turnbull. 1991. “Multiple Defects of the Mitochondrial Respiratory Chain in a Mitochondrial Encephalopathy (MERRF): A Clinical, Biochemical and Molecular Study.” *Journal of the Neurological Sciences* 102 (1): 17–24.

Blier, P. U., F. Dufresne, and R. S. Burton. 2001. “Natural Selection and the Evolution of mtDNA-Encoded Peptides: Evidence for Intergenomic Co-Adaptation.” *Trends in Genetics: TIG* 17 (7): 400–406.

Boore, J. L. 1999. “Animal Mitochondrial Genomes.” *Nucleic Acids Research* 27 (8): 1767–80.

Bottje, Walter G., Bhuwan Khatri, Stephanie A. Shouse, Dongwon Seo, Barbara Mallmann, Sara K. Orlowski, Jeonghoon Pan, et al. 2017. “Identification and Differential Abundance of Mitochondrial Genome Encoding Small RNAs (mitosRNA) in Breast Muscles of Modern Broilers and Unselected Chicken Breed.” *Frontiers in Physiology* 8 (October): 816.

Breton, Sophie, Donald T. Stewart, Sally Shepardson, Richard J. Trdan, Arthur E. Bogan, Eric G. Chapman, Andrew J. Ruminas, Helen Piontkivska, and Walter R. Hoeh. 2011. “Novel Protein Genes in Animal mtDNA: A New Sex Determination System in Freshwater Mussels (Bivalvia: Unionoida)?” *Molecular Biology and Evolution* 28 (5): 1645–59.

Brown, M. D., A. S. Voljavec, M. T. Lott, I. MacDonald, and D. C. Wallace. 1992. “Leber’s Hereditary Optic Neuropathy: A Model for Mitochondrial Neurodegenerative Diseases.” *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 6 (10): 2791–99.

Burton, R. S., C. K. Ellison, and J. S. Harrison. 2006. “The Sorry State of F2 Hybrids: Consequences of Rapid Mitochondrial DNA Evolution in Allopatric Populations.” *The American Naturalist* 168 Suppl 6 (December): S14–24.

Camus, M. Florencia, Jonci N. Wolff, Carla M. Sgrò, and Damian K. Dowling. 2017. "Experimental Support That Natural Selection Has Shaped the Latitudinal Distribution of Mitochondrial Haplotypes in Australian Drosophila Melanogaster." *Molecular Biology and Evolution* 34 (10): 2600–2612.

Cann, R. L., M. Stoneking, and A. C. Wilson. 1987. "Mitochondrial DNA and Human Evolution." *Nature* 325 (6099): 31–36.

Capt, Charlotte, Marco Passamonti, and Sophie Breton. 2016. "The Human Mitochondrial Genome May Code for More than 13 Proteins." *Mitochondrial DNA. Part A, DNA Mapping, Sequencing, and Analysis* 27 (5): 3098–3101.

Cloonan, Nicole. 2015. "Re-Thinking miRNA-mRNA Interactions: Intertwining Issues Confound Target Discovery." *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 37 (4): 379–88.

Craven, Lyndsey, Helen A. Tuppen, Gareth D. Greggains, Stephen J. Harbottle, Julie L. Murphy, Lynsey M. Cree, Alison P. Murdoch, et al. 2010. "Pronuclear Transfer in Human Embryos to Prevent Transmission of Mitochondrial DNA Disease." *Nature* 465 (7294): 82–85.

Dobler, Ralph, Damian K. Dowling, Edward H. Morrow, and Klaus Reinhardt. 2018. "A Systematic Review and Meta-Analysis Reveals Pervasive Effects of Germline Mitochondrial Replacement on Components of Health." *Human Reproduction Update* 24 (5): 519–34.

Dölle, Christian, Irene Flønes, Gonzalo S. Nido, Hrvoje Miletic, Nelson Osuagwu, Stine Kristoffersen, Peer K. Lilleng, et al. 2016. "Defective Mitochondrial DNA Homeostasis in the Substantia Nigra in Parkinson Disease." *Nature Communications* 7 (November): 13548.

Feder, Jeanette, Ofer Ovadia, Ilana Blech, Josef Cohen, Julio Wainstein, Ilana Harman-Boehm, Benjamin Glaser, and Dan Mishmar. 2009. "Parental Diabetes Status Reveals Association of Mitochondrial DNA Haplogroup J1 with Type 2 Diabetes." *BMC Medical Genetics* 10 (June): 60.

Fine, Amena Smith, Christina L. Nemeth, Miriam L. Kaufman, and Ali Fatemi. 2019. "Mitochondrial Aminoacyl-tRNA Synthetase Disorders: An Emerging Group of Developmental Disorders of Myelination." *Journal of Neurodevelopmental Disorders* 11 (1): 29.

Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello. 1998. "Potent and Specific Genetic Interference by Double-Stranded RNA in Caenorhabditis Elegans." *Nature* 391 (6669): 806–11.

Florencia Camus, M., Jochen B. W. Wolf, Edward H. Morrow, and Damian K. Dowling. 2015. "Single Nucleotides in the mtDNA Sequence Modify Mitochondrial Molecular Function and Are Associated with Sex-Specific Effects on Fertility and Aging." *Current Biology: CB* 25 (20): 2717–22.

Galtier, N., B. Nabholz, S. Glémin, and G. D. D. Hurst. 2009. "Mitochondrial DNA as a Marker of Molecular Diversity: A Reappraisal." *Molecular Ecology* 18 (22): 4541–50.

Gissi, C., F. Iannelli, and G. Pesole. 2008. "Evolution of the Mitochondrial Genome of Metazoa as Exemplified by Comparison of Congeneric Species." *Heredity* 101 (4): 301–20.

Greiner, Stephan, Johanna Sobanski, and Ralph Bock. 2015. "Why Are Most Organelle Genomes Transmitted Maternally?" *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 37 (1): 80–94

Gorman, Gráinne S., Patrick F. Chinnery, Salvatore DiMauro, Michio Hirano, Yasutoshi Koga, Robert McFarland, Anu Suomalainen, David R. Thorburn, Massimo Zeviani, and Douglass M.

Turnbull. 2016. “Mitochondrial Diseases.” *Nature Reviews. Disease Primers* 2 (October): 16080.

Gyllensten, U., D. Wharton, and A. C. Wilson. 1985. “Maternal Inheritance of Mitochondrial DNA during Backcrossing of Two Species of Mice.” *The Journal of Heredity* 76 (5): 321–24.

Ha, Minju, and V. Narry Kim. 2014. “Regulation of microRNA Biogenesis.” *Nature Reviews. Molecular Cell Biology* 15 (8): 509–24.

Harada, Alice E., Timothy M. Healy, and Ronald S. Burton. 2019. “Variation in Thermal Tolerance and Its Relationship to Mitochondrial Function Across Populations of Tigriopus Californicus.” *Frontiers in Physiology* 10 (March): 213.

Harrison, J. Scott, and Ronald S. Burton. 2006. “Tracing Hybrid Incompatibilities to Single Amino Acid Substitutions.” *Molecular Biology and Evolution* 23 (3): 559–64.

Hudson, Gavin, Aurora Gomez-Duran, Ian J. Wilson, and Patrick F. Chinnery. 2014. “Recent Mitochondrial DNA Mutations Increase the Risk of Developing Common Late-Onset Human Diseases.” *PLoS Genetics* 10 (5): e1004369.

Ingman, M., H. Kaessmann, S. Pääbo, and U. Gyllensten. 2000. “Mitochondrial Genome Variation and the Origin of Modern Humans.” *Nature* 408 (6813): 708–13.

Jiang, Min, Timo Eino Sakari Kauppila, Elisa Motori, Xinping Li, Ilian Atanassov, Kat Folz-Donahue, Nina Anna Bonekamp, Sara Albarran-Gutierrez, James Bruce Stewart, and Nils-Göran Larsson. 2017. “Increased Total mtDNA Copy Number Cures Male Infertility Despite Unaltered mtDNA Mutation Load.” *Cell Metabolism* 26 (2): 429–36.e4.

Karnkowska, Anna, Vojtěch Vacek, Zuzana Zubáčová, Sebastian C. Treitli, Romana Petrželková, Laura Eme, Lukáš Novák, et al. 2016. “A Eukaryote without a Mitochondrial Organelle.” *Current Biology: CB* 26 (10): 1274–84.

Khachane, Amit N., Kenneth N. Timmis, and Vítor A. P. Martins dos Santos. 2007. “Dynamics of Reductive Genome Evolution in Mitochondria and Obligate Intracellular Microbes.” *Molecular Biology and Evolution* 24 (2): 449–56.

Kuijper, B., N. Lane, and A. Pomiankowski. 2015. “Can Paternal Leakage Maintain Sexually Antagonistic Polymorphism in the Cytoplasm?” *Journal of Evolutionary Biology* 28 (2): 468–80.

Lajbner, Zdeněk, Reuven Pnini, M. Florencia Camus, Jonathan Miller, and Damian K. Dowling. 2018. “Experimental Evidence That Thermal Selection Shapes Mitochondrial Genome Evolution.” *Scientific Reports* 8 (1): 9500.

Lee, Changhan, Kelvin Yen, and Pinchas Cohen. 2013. “Humanin: A Harbinger of Mitochondrial-Derived Peptides?” *Trends in Endocrinology and Metabolism: TEM* 24 (5): 222–28.

Lee, Yuh Chwen G. 2015. “The Role of piRNA-Mediated Epigenetic Silencing in the Population Dynamics of Transposable Elements in Drosophila Melanogaster.” *PLoS Genetics* 11 (6): e1005269.

Levin, Liron, Ilia Zhidkov, Yotam Gurman, Hadas Hawlena, and Dan Mishmar. 2013. “Functional Recurrent Mutations in the Human Mitochondrial Phylogeny: Dual Roles in Evolution and Disease.” *Genome Biology and Evolution* 5 (5): 876–90.

Ma, Hao, Gregory M. Weber, Hairong Wei, and Jianbo Yao. 2016. “Identification of Mitochondrial Genome-Encoded Small RNAs Related to Egg Deterioration Caused by Postovulatory Aging in Rainbow Trout.” *Marine Biotechnology* 18 (5): 584–97.

Majamaa, K., J. S. Moilanen, S. Uimonen, A. M. Remes, P. I. Salmela, M. Kärppä, K. A. Majamaa-Voltti, et al. 1998. “Epidemiology of A3243G, the Mutation for Mitochondrial Encephalomyopathy, Lactic Acidosis, and Strokelike Episodes: Prevalence of the Mutation in

an Adult Population." *American Journal of Human Genetics* 63 (2): 447–54.
Majander, A., K. Huoponen, M. L. Savontaus, E. Nikoskelainen, and M. Wikström. 1991. "Electron Transfer Properties of NADH:ubiquinone Reductase in the ND1/3460 and the ND4/11778 Mutations of the Leber Hereditary Optic Neuroretinopathy (LHON)." *FEBS Letters* 292 (1-2)
Marom, Shani, Michael Friger, and Dan Mishmar. 2017. "MtDNA Meta-Analysis Reveals Both Phenotype Specificity and Allele Heterogeneity: A Model for Differential Association." *Scientific Reports* 7 (February): 43449.
Matsuoka, Masaaki. 2009. "Humanin; a Defender against Alzheimer's Disease?" *Recent Patents on CNS Drug Discovery* 4 (1): 37–42.
McFarland, Robert, Andrew M. Schaefer, Julie L. Gardner, Stephen Lynn, Christine M. Hayes, Martin J. Barron, Mark Walker, Patrick F. Chinnery, Robert W. Taylor, and Douglass M. Turnbull. 2004. "Familial Myopathy: New Insights into the T14709C Mitochondrial tRNA Mutation." *Annals of Neurology* 55 (4): 478–84.
Mercer, Tim R., Shane Neph, Marcel E. Dinger, Joanna Crawford, Martin A. Smith, Anne-Marie J. Shearwood, Eric Haugen, et al. 2011. "The Human Mitochondrial Transcriptome." *Cell* 146 (4): 645–58.
Milani, Liliana, Fabrizio Ghiselli, Maria Gabriella Maurizii, Sergey V. Nuzhdin, and Marco Passamonti. 2014. "Paternally Transmitted Mitochondria Express a New Gene of Potential Viral Origin." *Genome Biology and Evolution* 6 (2): 391–405.
Mishmar, Dan, Eduardo Ruiz-Pesini, Martin Brandon, and Douglas C. Wallace. 2004. "Mitochondrial DNA-like Sequences in the Nucleus (NUMTs): Insights into Our African Origins and the Mechanism of Foreign DNA Integration." *Human Mutation* 23 (2): 125–33.
Mishmar, Dan, Eduardo Ruiz-Pesini, Pawel Golik, Vincent Macaulay, Andrew G. Clark, Seyed Hosseini, Martin Brandon, et al. 2003. "Natural Selection Shaped Regional mtDNA Variation in Humans." *Proceedings of the National Academy of Sciences of the United States of America* 100 (1): 171–76.
Mohlke, Karen L., Anne U. Jackson, Laura J. Scott, Erin C. Peck, Yong D. Suh, Peter S. Chines, Richard M. Watanabe, et al. 2005. "Mitochondrial Polymorphisms and Susceptibility to Type 2 Diabetes-Related Traits in Finns." *Human Genetics* 118 (2): 245–54.
Monaghan, Richard M., and Alan J. Whitmarsh. 2015. "Mitochondrial Proteins Moonlighting in the Nucleus." *Trends in Biochemical Sciences* 40 (12): 728–35.
Nakada, Kazuto, Akitsugu Sato, Kayo Yoshida, Takashi Morita, Hiromitsu Tanaka, Shin-Ichi Inoue, Hiromichi Yonekawa, and Jun-Ichi Hayashi. 2006. "Mitochondria-Related Male Infertility." *Proceedings of the National Academy of Sciences of the United States of America* 103 (41): 15148–53.
Newson, A. J., S. de Lacey, D. K. Dowling, S. Murray, C. M. Sue, D. R. Thorburn, L. Gillam, and C. Degeling. 2019. "Public Attitudes towards Novel Reproductive Technologies: A Citizens' Jury on Mitochondrial Donation." *Human Reproduction* 34 (4): 751–57.
Nishimasu, Hiroshi, Hirotsugu Ishizu, Kuniaki Saito, Satoshi Fukuhara, Miharu K. Kamatani, Luc Bonnefond, Naoki Matsumoto, et al. 2012. "Structure and Function of Zucchini Endoribonuclease in piRNA Biogenesis." *Nature* 491 (7423): 284–87.
Ozawa, Takayuki, Satoru Sugiyama, Masashi Tanaka, and Kazuki Hattori. 1991. "Mitochondrial DNA Mutations and Disturbances of Energy Metabolism in Myocardium." *Japanese Circulation Journal*. https://doi.org/10.1253/jcj.55.1158.
Picard, Martin, Douglas C. Wallace, and Yan Burelle. 2016. "The Rise of Mitochondria in

Medicine.” *Mitochondrion* 30 (September): 105–16.

Pozzi, Andrea, Federico Plazzi, Liliana Milani, Fabrizio Ghiselli, and Marco Passamonti. 2017. “SmithRNAs: Could Mitochondria ‘Bend’ Nuclear Regulation?” *Molecular Biology and Evolution*, April. https://doi.org/10.1093/molbev/msx140.

Riggs, Claire L., Amanda Summers, Daniel E. Warren, Göran E. Nilsson, Sjannie Lefevre, W. W. Dowd, Sarah Milton, and Jason E. Podrabsky. 2018. “Small Non-Coding RNA Expression and Vertebrate Anoxia Tolerance.” *Frontiers in Genetics* 9 (July): 230.

Riordan-Eva, P., and A. E. Harding. 1995. “Leber’s Hereditary Optic Neuropathy: The Clinical Relevance of Different Mitochondrial DNA Mutations.” *Journal of Medical Genetics* 32 (2):

Ro, Seungil, Hsiu-Yen Ma, Chanjae Park, Nicole Ortogero, Rui Song, Grant W. Hennig, Huili Zheng, et al. 2013. “The Mitochondrial Genome Encodes Abundant Small Noncoding RNAs.” *Cell Research* 23 (6): 759–74.

Rubinoff, Daniel, and Brenden S. Holland. 2005. “Between Two Extremes: Mitochondrial DNA Is Neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference.” *Systematic Biology* 54 (6): 952–61.

Ruiz-Pesini, Eduardo, Ana-Cristina Lapeña, Carmen Díez-Sánchez, Acisclo Pérez-Martos, Julio Montoya, Enrique Alvarez, Miguel Díaz, et al. 2000. “Human mtDNA Haplogroups Associated with High or Reduced Spermatozoa Motility.” *American Journal of Human Genetics* 67 (3): 682–96.

Saccone, C., C. De Giorgi, C. Gissi, G. Pesole, and A. Reyes. 1999. “Evolutionary Genomics in Metazoa: The Mitochondrial DNA as a Model System.” *Gene* 238 (1): 195–209.

Shoffner, J. M., and D. C. Wallace. 1992. “Mitochondrial Genetics: Principles and Practice.” *American Journal of Human Genetics* 51 (6): 1179–86.

Tachibana, Masahito, Michelle Sparman, Hathaitip Sritanaudomchai, Hong Ma, Lisa Clepper, Joy Woodward, Ying Li, Cathy Ramsey, Olena Kolotushkina, and Shoukhrat Mitalipov. 2009. “Mitochondrial Gene Replacement in Primate Offspring and Embryonic Stem Cells.” *Nature* 461 (7262): 367–72.

Vaught, Rebecca C., Susanne Voigt, Ralph Dobler, David J. Clancy, Klaus Reinhardt, and Damian K. Dowling. 2020. “Interactions between Cytoplasmic and Nuclear Genomes Confer Sex-Specific Effects on Lifespan in Drosophila Melanogaster.” *Journal of Evolutionary Biology*, February. https://doi.org/10.1111/jeb.13605.

Wallace, D. C. 1994. “Mitochondrial DNA Mutations in Diseases of Energy Metabolism.” *Journal of Bioenergetics and Biomembranes* 26 (3): 241–50.

Wallace, D. C., X. X. Zheng, M. T. Lott, J. M. Shoffner, J. A. Hodge, R. I. Kelley, C. M. Epstein, and L. C. Hopkins. 1988. “Familial Mitochondrial Encephalomyopathy (MERRF): Genetic, Pathophysiological, and Biochemical Characterization of a Mitochondrial DNA Disease.” *Cell* 55 (4): 601–10.

Yee, Winston K. W., Björn Rogell, Bernardo Lemos, and Damian K. Dowling. 2015. “Intergenomic Interactions between Mitochondrial and Y-Linked Genes Shape Male Mating Patterns and Fertility in Drosophila Melanogaster.” *Evolution; International Journal of Organic Evolution* 69 (11): 2876–90.

Zimmer, Carl. 2009. “Origins. On the Origin of Eukaryotes.” *Science* 325 (5941): 666–68.

CHAPTER TWO

The genomic origins of small mitochondrial RNAs: are they transcribed by the mitochondrial DNA or by mitochondrial pseudogenes within the nucleus (NUMTs)?

Title: The genomic origins of small mitochondrial RNAs: are they transcribed by the mitochondrial DNA or by mitochondrial pseudogenes within the nucleus (NUMTs)?

Author 1: Andrea Pozzi, Monash University, School of Biological Sciences, andrea.pozzi@monash.edu

Author 2: Damian K. Dowling, Monash University, School of Biological Sciences, damian.dowling@monash.edu

Status: Published – Genome Biology and Evolution

https://doi.org/10.1093/gbe/evz132

Abstract

Several studies have linked mitochondrial genetic variation to phenotypic modifications; albeit the identity of the mitochondrial polymorphisms involved remains elusive. The search for these polymorphisms led to the discovery of small non-coding RNAs, which appear to be transcribed by the mitochondrial DNA ("small mitochondrial RNAs"). This contention is, however, controversial because the nuclear genome of most animals harbours mitochondrial pseudogenes (NUMTs) of identical sequence to regions of mtDNA, which could alternatively represent the source of these RNAs. To discern the likely contributions of the mitochondrial and nuclear genome to transcribing these small mitochondrial RNAs, we leverage data from six vertebrate species exhibiting markedly different levels of NUMT sequence. We explore whether abundances of small mitochondrial RNAs are associated with levels of NUMT sequence across species, or differences in tissue-specific mtDNA content within species. Evidence for the former would support the hypothesis these RNAs are primarily transcribed by NUMT sequence, while evidence for the latter would provide strong evidence for the counter hypothesis that these RNAs are transcribed directly by the mtDNA. No association exists between the abundance of small mitochondrial RNAs and NUMT levels across species. Moreover, a sizable proportion of transcripts map exclusively to the mtDNA sequence, even in species with highest NUMT levels. Conversely, tissue-specific abundances of small mitochondrial RNAs are strongly associated with the mtDNA content. These results support the hypothesis that small mitochondrial RNAs are primarily transcribed by the mitochondrial genome and that this capacity is conserved across *Amniota* and, most likely, across most metazoan lineages.

Introduction

For the past four decades, biologists have routinely harnessed the genetic variability found within the mitochondrial DNA (mtDNA) sequence as an evolutionary marker, to help in resolving the population genetic processes shaping patterns of evolutionary divergence across natural populations and divergent taxa (Avise 2000; Avise et al. 1987; Bandelt et al. 1995; van Oven 2010). These inferences were typically made under the assumption that this genetic variation is associated with no phenotypic alterations to the organism, and therefore has evolved under a neutral equilibrium model (Ballard and Kreitman 1995). However, the results of numerous studies to emerge over the past two decades have challenged the assumption of selective neutrality of the mitochondrial genome, by showing that mitochondrial genetic variation is often tied to phenotypic expression, or by revealing adaptive signatures of mutational variation within the mtDNA sequence (Dowling 2014; Ballard and Pichaud 2014; Vaught and Dowling 2018; Ballard and Whitlock 2004; Bazin, Glémin, and Galtier 2006; James, Piganeau, and Eyre-Walker 2016).

For example, several studies on vertebrate and invertebrate study systems have leveraged experimental designs able to partition mitochondrial from nuclear genetic effects on organismal function (Camus et al. 2017; Tourmente et al. 2017; Morales et al. 2018; Innocenti, Morrow, and Dowling 2011; Estopinal et al. 2014; Mossman et al. 2016; Harrison and Burton 2006). These uncovered a role for mitochondrial genetic variation in the regulation of physiological, life-history, and health-related phenotypes. These studies were further augmented by studies reporting intriguing associations between particular mitochondrial haplogroups and patterns of disease penetrance in humans (Marom, Friger, and Mishmar 2017; Chalkia et al. 2018; Sloan, Fields, and Havird 2015; Singh and Kulawiec 2009; Hudson et al. 2014). For example, Marom *et al*. (2017) observed that certain diseases, such as Type 2 diabetes and Parkinson's disease, are enriched in humans carrying particular mtDNA haplogroups (haplogroups H, J, and K), while Hudson *et al.* (2014) noted associations between mtDNA haplogroups and the risk of developing several complex diseases, including ischemic stroke, Parkinson's disease, multiple sclerosis, and schizophrenia. Yet, while it is clear that mitochondrial genetic variation regularly affects phenotypic expression, uncovering the candidate mtDNA mutations involved in the regulation of these effects has proven difficult, given that the mutations that lie in the mtDNA sequence generally appear in tight linkage disequilibrium and are thus inherited as a block.

Most studies that have sought to uncover the candidate polymorphisms that underpin the mitochondrial genotype-phenotype linkage, have focused on genetic variants within the protein-coding region of the mtDNA (Camus et al. 2015; Harrison and Burton 2006; Singh and Kulawiec 2009; Weihe et al. 2009). Some of these studies identified candidate polymorphisms that confer a modification to the protein-coding sequence (Dowling, Tompkins, and Gemmell 2015; Patel et al. 2016; Camus et al. 2015). However, in several other cases, the polymorphisms underpinning the association between particular mtDNA haplogroups and patterns of phenotypic expression remained elusive, but do not appear to involve substitutions to the amino acid sequence (Camus et al. 2017; Hopkins et al. 2017; Takasaki 2009). For example, Hopkins *et al.* (2017) reported that specific mutations in the mitochondrial origin of replication (ORI) in humans are associated with the progression to aggressive grades of prostate cancer. Other studies have proposed that mitochondrial regulation of the phenotype may proceed via mechanisms other than direct changes to the protein coding sequence. Indeed, recent studies have reported the discovery of previously uncharacterized RNAs, which are putatively transcribed by the mitochondria, and that may serve important functions (Larriba, Rial, and Del Mazo 2018; Riggs et al. 2018; Mercer et al. 2011; Pozzi et al. 2017; Ro et al. 2013).

The importance of mitochondrial RNAs (mtRNAs), such as ribosomal RNAs (mt-rRNA) and transfer RNAs (mt-tRNA), which are necessary for translation of messenger RNA (mt-mRNA), has long been recognized (Storz 2002). However, in recent years, several independent studies have reported the presence of small non-coding RNAs that appear to be encoded by the mtDNA in fish, mammals, and clams, and the authors of these studies have suggested a role for these RNAs in driving phenotypic changes such as improved anoxia tolerance and involvement in sex determination (Pozzi et al. 2017; Riggs et al. 2018; Sanchez et al. 2011; Larriba, Rial, and Del Mazo 2018). The presence of these "small mitochondrial RNAs", across several species, suggests they might play a role in cell regulation.

This suggestion, however, is controversial, and it remains debated whether these small RNAs are non-functional, representing degraded residues of longer, well-known RNAs (Houseley and Tollervey 2009), or functional in their own right, representing a nucleic part of a ribonucleoprotein (Hogan et al. 2008). Ribonucleoproteins usually use small RNA to recognize specific DNA or RNA targets with which to bind, to exert their function (Hogan et al. 2008). A classic example of such a mechanism is RNA interference (RNAi), a process in which microRNAs (miRNA) of approximately 22 nucleotides in length, each with a specific sequence, lead the RNA-induced silencing complex (RISC) to the target mRNA where the protein can inhibit the translation of the mRNA (Ambros 2004; Cloonan 2015; Ha and Kim 2014).

In 2017, Pozzi and colleagues proposed that some of these mitochondrial small RNAs could use RNAi to shape patterns of nuclear gene regulation, thus invoking a hitherto unrecognized mechanism by which the mitochondria could broadly shape patterns of phenotypic expression (Pozzi et al. 2017). Furthermore, this prediction coincides with recent studies of Innocenti *et al.* (2011) and Baris *et al.* (2017), both of which identified an abundance of nuclear genes that are seemingly involved in the regulation of mitochondrial-nuclear interactions, but which had no previously known involvement in mitochondrial function. The proteins encoded by these nuclear genes are generally not transported into the mitochondrion, and thus it is unclear how their regulation could be affected by genetic polymorphisms within the mtDNA. The discovery of small mitochondrial RNAs may help to address these outstanding questions, if these RNAs are able to target a broad range of nuclear genes that have no previous known association with mitochondrial function. However, currently, even the origin of these small RNAs remains controversial. This is because the nuclear genome of eukaryotes is littered with mtDNA sequence in the form of nuclear mitochondrial pseudogenes (<u>Nu</u>clear <u>Mit</u>ochondrial DNA Segment, **NUMT**), which in theory could represent the source of these small RNAs.

The presence of NUMTs makes it difficult to definitively identify the origin of small mitochondrial RNA, since in many metazoans, much of the mtDNA sequence is duplicated via shorter inserts throughout the nuclear genome, with the NUMT sequence being almost identical to that of the mitochondrial DNA of origin (Lopez et al. 1994; Gaziev and Shaikhaev 2010). However, despite their widespread presence among most eukaryotes, the function of NUMTs remains largely unknown, and for the most part, they are considered non-functional elements (Hazkani-Covo, Zeller, and Martin 2010). Indeed, because the genetic code differs between mitochondrial and nuclear genomes, the mitochondrial sequences would not make the same protein using the nuclear code, and therefore these NUMTs are widely regarded as pseudogenes (Lopez et al. 1994). Furthermore, these sequences are "tuned" to be transcribed and translated using proteins specific to the mitochondria, suggesting the inability of the NUMTs to produce functional proteins even when entire genes are translocated to the nuclear genome (Smits, Smeitink, and van den Heuvel 2010). Yet, although NUMTs may not express proteins, they may still express small RNAs. Because NUMTs are mitochondrial in origin, any small RNAs they encode would be nearly identical in sequence to mRNAs encoded by mtDNA, rendering it difficult to determine whether small RNAs that have previously been identified are transcribed by the mitochondrial or nuclear genome.

Fortunately, NUMTs have several unique features relative to the mtDNA, and these features can be used to formulate two sets of testable predictions that allow the putative source of these small mitochondrial RNAs to be discerned. One of the features is that the amount of NUMT sequence differs greatly across metazoan species (Rogers and Griffiths-Jones 2012; Hazkani-Covo, Zeller, and Martin 2010). For example, the mouse harbors around 25 times more NUMT sequence than the chicken (~37 Kb or 0.15% of the mouse nuclear genome, ~1.52Kb or 0.01% of the chicken nuclear genome (Calabrese, Simone, and Attimonelli 2012; Bensasson, Feldman, and Petrov 2003). Another feature is that the NUMT amount is stable across tissues, while the mtDNA copy number changes greatly across different tissues. Here, we take advantage of these species differences in NUMT content, and tissue-specific differences mtDNA copy number, to home in on the probable origin of the small mitochondrial RNAs. Our inferences hinge on the assumption that the abundance of the small mitochondrial RNAs is correlated with the abundance of their source molecules, be that the amount of NUMT sequence, or copies of mtDNA. We contend that this assumption is plausible given the large differences (greater than a thousandfold) in both the amount of NUMT sequence present across distinct species, and in the number of mitochondria found across distinct tissues within a species. Based on this assumption, we formulate two predictions. Firstly, if the small mitochondrial RNAs are encoded primarily by NUMTs, then we predict that the level of transcription of these RNAs across species will be correlated to the amount of species-specific NUMT sequence. Species with large amounts of NUMT sequence should possess smaller mitochondrial RNAs than species with low amounts of NUMTs. On the contrary, if the mitochondrial DNA encodes the small mitochondrial RNAs, then it is predicted that there will be no correlation between the amount of NUMT sequence and the abundance of small mitochondrial RNAs across species. The second prediction leverages the observation that the copy number of NUMT pseudogenes will be stable across tissue types within a species, but the copy number of mtDNA molecules will differ greatly. Thus, if generally transcribed by NUMT sequence, we predict that the small mitochondrial RNAs will not exhibit predictable differences in abundance across tissue-types. In contrast, if transcribed by the mtDNA, then we predict the abundance of these small RNAs will be higher in tissues that are rich in mitochondria relative to tissues that are poor in mitochondrial content. In particular, tissues that exhibit naturally-enriched levels of mitochondria, such as the cancerous cells, would be predicted to show increased abundances of small mitochondrial RNAs when compared to their healthy counterparts that have less mitochondria (Williams et al. 2015; Vyas, Zaganjor, and Haigis 2016; Zhu et al. 2017).

Here, we tested these predictions, using publicly available data stored in the Short Read Archive (SRA) in the NCBI. Many miRNA-sequencing datasets are available for model organisms, such as chicken and mice, which can be repurposed to examine the profiles of small mitochondrial RNAs across species and tissues. To clarify, these datasets consist of RNA sequencing data in which the transcripts have been selected using a size-selection approach that selects transcripts smaller than 50nt. Due to this method, the small mitochondrial RNAs, the targets of our study, are represented in these libraries. Using these data, we first tested whether the amount of NUMTs present in the nuclear genome of a given species is associated with the amount of small mitochondrial RNA transcribed, across several species of the *Amniota* clade. We then focused on expression patterns of small mitochondrial RNAs across three tissues of mice and chicken, two species with very different amounts of NUMT sequence. To address the second prediction, we compared abundances of small mitochondrial RNA from a mitochondria-rich tissue and a mitochondria-poor tissue, in human datasets. Then, to further probe this prediction, we compared small mitochondrial RNAs abundance in a tissue naturally enriched in mitochondria (cancerous tissue) to its healthy counterpart with physiological levels of mtDNA copies.

Results

The small mitochondrial RNAs are widely expressed across tissues in *Amniota*

We first tested whether the abundance of small mitochondrial RNAs changes with the amount of NUMT sequence across species. We used data from Meunier *et al.*, (2013), who analyzed tissue-specific RNA libraries from each of six species of the *Amniota* clade. Each of these species is characterized by differing degrees of NUMT sequence insertions. We quantified the percentage of small RNAs that map to the mtDNA in each of these samples and compared this to the percentage of reads mapping to NUMT sequence (**Fig.1A)**. Accordingly, we can define three different type of reads: the ***mtDNA reads***, which include the total pool of reads that map to the mtDNA (those that map exclusively to the mtDNA as well as those that map both to mtDNA and NUMT sequence); ***mtDNA-only reads***, which map unambiguously and exclusively to the mtDNA and exclude the sequences present in the NUMTs; and ***NUMT reads***, which represent the reads that map both to the mtDNA and NUMT sequence. Generally, species with high levels of NUMTs have high proportions of reads that map jointly to both mtDNA and NUMTs (**Fig.1B**). However, even in species that have very high levels of NUMTS, such as the opossum, at least 25% of the mtDNA reads are nonetheless comprised of mtDNA-only reads. Notably, the chicken is characterized by very low levels of NUMT sequence. In this species, the percentage of mtDNA reads that map exclusively to the mtDNA sequence is about 95% across all sampled tissues.

We next investigated whether associations exist between the amount of NUMT sequence and the percentage of mtDNA reads across the different tissues of each species (**Fig.2**). Although we note there are insufficient samples per tissue and species to warrant formal statistical analysis (n = 1 data point per tissue per species), the analysis highlights an absence of any clear relationship between the NUMT content and total abundance of mtDNA reads across the six species.

Small mitochondrial RNA abundance is independent of the NUMT content

To further probe associations between NUMT content and abundances of mtDNA reads, we conducted a further set of analyses that focused on two of the six species surveyed above (chicken and mouse). These species exhibit very different NUMT contents (chicken 1.52 Kb, mouse 37.67 Kb), with many supporting small RNA-seq datasets available for each. Our analyses focused on three different tissue types, for which there are numerous samples available in public databases: the brain, heart, and liver (Reznik et al. 2016; Rooney et al. 2015).

Our analyses of tissue-specific abundances of mtDNA reads reinforce the large difference in levels of NUMT reads between the two species. In the chicken datasets, there are consistently lower levels of NUMT reads per tissue than in the mouse datasets (Mann Whitney, $U = 0$, $n = 80$ $p = 0.0001$, **Fig.3**). The mouse datasets show moderate levels of NUMT reads, which nonetheless rarely surpass the amount of mtDNA-only reads. Specifically, while the percentage of NUMT reads in the brain (median 1.18 ± 0.73%) is similar to the percentage of mtDNA-only reads (median 1.87 ± 0.4%), the heart and liver have a greater representation of mtDNA-only reads (medians of 3.75 ± 0.53% and 1.44 ± 0.1% respectively) than NUMT reads (medians of 2.65 ± 0.27% and 1.00 ± 0.1% respectively, Supplementary Material Table S2).

Overall, when comparing mtDNA-only to NUMT reads across all of the chicken samples, by pooling together the data from all tissues, the percentage of NUMT reads usually comprises less than 0.1% of the small RNA library (**Fig.4**), confirming observations derived from the dataset of Meunier *et al.* (2013) that we report in Fig. 1. Furthermore, we observed high abundances of mtDNA-only reads in the chicken samples, represented as a proportion of the total small RNA reads per library, ranging from 0.8 to 12% across the three tissues. Patterns observed across mouse samples similarly reinforced the results of our analysis of the Meunier *et al.* (2013) dataset. We detected a high percentage of NUMT reads in all tissues, which reflects the high amount of NUMT sequence in the mouse genome (~37Kb).

This pattern suggests that, at least in the chicken, the majority of the mtDNA reads must be transcribed directly from the mtDNA. The source of the mtDNA reads in mice remains open to question, and technically, it is possible that the NUMTs play a secondary role in their transcription. Notwithstanding, it must be noted that all NUMT reads also map to the mtDNA, and thus the mtDNA would remain the most likely candidate site for their transcription. Indeed, if the NUMTs play a sizable role in the transcription of mtDNA reads, we would expect that the total pool of mtDNA reads would be significantly lower in the chicken than in the mouse, given the much lower NUMT content of the chicken genome of NUMT sequence in the chicken. While there are no differences in the proportion of mtDNA-only reads between the species ($U = 621$, $n = 79$, $p = 0.12$), the chicken samples do exhibit a meaningful reduction in the abundance of mtDNA reads when compared to the mouse samples ($U = 493$, $n = 79$, $p = 0.003$). While at first, this result would suggest a role for the NUMTs in the transcription of small mitochondrial RNA, closer scrutiny of the tissue-specific samples suggests otherwise. While both the chicken liver and brain samples show statistically significant reductions in the abundance of mtDNA reads relative to the equivalent tissues in mice (liver $U = 37$, $n = 27$, $p = 0.004$; brain $U = 38$, $n = 26$, $p = 0.009$), the heart samples of the two species do not differ ($U = 68.5$, $n = 26$, $p = 0.43$). See Supplementary Material Table S1 for full statistical analyses.

The abundance of small mitochondrial RNAs is due to a few transcriptional hotspots

We next assessed whether differences in the abundances of NUMT and mtDNA-only reads across the two species are attributable to changes in the expression of small mitochondrial RNA molecules transcribed across many different regions of the mtDNA sequence, or by high levels of transcription at just a few specific sequence regions. In the chicken, almost all the mtDNA reads align to the rRNA16S, aligning just after the 2Kb mark (**Fig.5A-B**). This peak aligns to mtDNA-only reads, consistent with expectation given the high percentage of mtDNA-only reads reported above for this species. Indeed, NUMT reads are almost entirely absent, and only minor peaks of mtDNA-only reads can be seen in the brain and heart samples.

In the mouse, hotspots of expression of mtDNA reads were observed in the heart and liver. This was not the case for expression patterns in the brain, in which many different transcripts across the coding, tRNA and rRNA regions of the mtDNA sequence were observed (**Fig.5C-D**). In particular, the abundance of mtDNA reads present in the mouse heart and liver is attributable to one hotspot of transcription (close to position 12Kb in the mtDNA) aligning to mt-tRNA Ser2 gene. Due to the similarity between mtDNA and NUMT sequence, this peak could technically be either mtDNA or NUMT encoded (they are NUMT reads). The second large peak, comprising a mix of NUMT and mtDNA-only reads, is present in all samples, aligning on mt-rRNA16s and mt-tRNA Val genes, suggesting very high conservation of these mtDNA reads across tissues in the mouse. The abundance of mtDNA reads in the brain samples show a diffuse pattern of expression of thousands of transcripts across the entire coding sequence of mtDNA, lacking specific hotspots of transcription, which is difficult to explain, based on our current knowledge.

Abundance of small mitochondrial RNAs is closely associated with mtDNA content

We extended our investigation, turning to RNA libraries from humans, to examine the correlation between the abundance of mtDNA reads and mtDNA content, across tissues that are likely to exhibit strong differences in their mitochondrial content. We collated small RNA sequencing data from three studies, two from the brain (a putatively mitochondria-rich tissue) and one from the bladder epithelium (a putatively mitochondria-poor tissue). We calculated the percentage of small RNA reads mapping to the mtDNA sequence across each of these tissues (**Fig.6**). The mitochondria-rich brain samples have higher levels of transcription of mtDNA reads ($U = 1$, $n = 66$, $p = 0.0001$). In fact, the mean abundance of mtDNA reads in the brain (3.46 ± 0.37%) is more than thirty-fold more than the mean abundance in the bladder epithelium (0.15 ± 0.01%).

Finally, to further probe the relationship between expression profiles of mtDNA reads and mtDNA content, we leveraged data from a recent study examining the small RNA profiles of bladder samples taken from patients with bladder cancer relative to healthy patients (Itesako et al. 2014). It has previously been shown that cancerous bladder cells show higher levels of mtDNA than healthy bladder cells (Williams et al. 2015). The samples from cancerous tissue, with enriched mtDNA content, contained over fivefold more mtDNA reads than the samples from healthy tissue ($U = 0$, $n = 10$, $p = 0.008$) (**Fig.7**). This result reinforces the observation that levels of small mitochondrial RNAs are tightly linked to the total content of mtDNA within a tissue type, thus supporting the contention that these RNAs are transcribed by the mtDNA, and not by NUMT sequence within the nuclear genome.

Discussion

Here, we present two lines of support for the hypothesis that the mtDNA is the primary source of small mitochondrial RNA transcription. Furthermore, our study demonstrates that the high levels of small mitochondrial RNA transcripts observed across the sampled small RNA datasets are largely attributable to a smaller set of tissue-specific transcripts; an observation that supports the contention that these transcripts are likely to confer a functional role.

The first line of support for the hypothesis comes from our observation that abundances of small mitochondrial RNA were not correlated to levels of NUMT sequence across species. One striking example of this point comes from the case of the opossum, a species with almost 10 times more NUMT sequence (~2093Kb) than the platypus, monkey, and human. Despite this order of magnitude difference in the level of NUMT reads, the proportion of mtDNA reads mapping to NUMT sequence was not tangibly higher across major tissues of the opossum relative to those of these other species. Furthermore, the presence of small mitochondrial RNAs in the chicken, a species with a distinct paucity of NUMT sequence, represents definitive evidence of the ability of the mtDNA to transcribe small mitochondrial RNAs. In this species, the transcripts mapping to the mtDNA do not map to any other part of the genome and their presence can thus only be explained by transcription from the mtDNA sequence. Indeed, across six species examined, we observed high proportions of reads mapping exclusively to the mtDNA sequence, confirming that transcription of small RNAs from the mitochondria occurs generally across vertebrates. Furthermore, we were able to confirm these general patterns by focusing on patterns in two species for which we had high levels of inferential power, given that RNA libraries were available from a number of independent studies for each of these species, and at high levels of replication.

The second line of support comes from analyses that show a tight link between expression levels of the small mitochondrial RNAs, and the mtDNA content of particular tissues. Our analyses are based on the assumption that a tissue with the high-energy demands, such as the brain (Picard and McEwen 2014; Davey, Peuchen, and Clark 1998; Veltri, Espiritu, and Singh 1990), will have more mitochondria than a tissue with putatively low demands, such as the epithelium (Gibson et al. 1996). On this, while there is evidence that brain tissue is rich in mitochondria (Picard and McEwen 2014; Davey, Peuchen, and Clark 1998), there is less direct evidence that epithelium tissue is low in mitochondria. However, despite not knowing the number of mitochondria, we argue that the difference in energy consumption between these two tissues should make our assumption reasonable. Under this assumption, if the mitochondria are the main source of mtDNA reads, we predict the mitochondria-poor tissue to express lower levels of these RNAs than their mitochondria-rich counterpart. Our predictions were supported. In fact, the abundance of mtDNA reads is strikingly higher in brain samples than in epithelium samples. Moreover, this result is supported by the analysis of cancerous cells naturally enriched in mtDNA (Williams et al. 2015; Vyas, Zaganjor, and Haigis 2016). The cancerous cells showed a fivefold increase in their abundance of mtDNA reads, supporting the hypothesis that the mtDNA is a main player in the transcription of mtDNA reads.

We note that our lines of evidence support previous experiments performed in cell lines. In fact, Ro et al. (2013) investigated the origin of the small mitochondrial RNAs by analysing the expression of cells with depleted mitochondria (Rho0). The authors found that Rho0 cells have strong downregulation of small mitochondrial RNAs transcription, thus adding a complementary layer of evidence that support our findings, and showing that NUMTs have very little if any role in transcribing the small mitochondrial RNAs.

Further investigations into the origins of small mitochondrial RNAs should seek to leverage allelic differences between NUMT sequences and mtDNA. Over time, NUMT sequences are expected to diverge from their ancestral mtDNA sequence, accumulating SNPs that are unique to NUMT sequences and not found in the mtDNA. In theory, this divergence should provide an opportunity for future studies to investigate whether small mitochondrial RNAs that map exclusively to NUMT sequence exist. Currently, limitations in the available data preclude such an analysis. Given that different individuals are expected to harbour different SNPs in both mtDNA and NUMTs (i.e., natural allelic variation in these regions of sequence), such analysis would require both DNA and RNA data originating from the same individual (or from the same sets of individuals). Furthermore, small mitochondrial RNAs exist in very short sequences (~30nt), presenting small mutational targets in which is unlikely multiple mutations will occur. Therefore, such mutations would rarely lead to tangible molecular divergence between mtDNA and NUMT sequence. This will make future analyses based on molecular divergence challenging.

Considering the evidence presented here, we propose that the small mitochondrial RNA reads that map jointly to both mtDNA and NUMTs should be assumed to originate from the mtDNA. Moreover, we argue that we can safely assume that all small RNAs aligning to mtDNA, including in species not tested in our study, are likely to be transcribed by the mtDNA, and not by NUMTs. In fact, the lack of evidence that the NUMTs are transcribed, or play functional roles in any metazoan, aligns well with our hypothesis (Hazkani-Covo, Zeller, and Martin 2010). Previous studies have identified and explored the putative function of small mitochondrial RNAs in other species (Riggs et al. 2018; Pozzi et al. 2017; Mercer et al. 2011). Our results provide strong support that their analyses and interpretations are unlikely to have been confounded by transcripts encoded by NUMTs present within the nuclear genome.

We have provided evidence to support the ability of the mtDNA to transcribe small RNAs in chickens, mice and humans. Thus from an evolutionary perspective, this allows us to date the evolution of this ability to at least the birth of the *Amniota* clade. However, while our study is the first to deeply investigate the origin of the small mitochondrial RNAs, it is not the first to have identified the presence of these molecules. In fact, prior to this study, previous studies had identified small mitochondrial RNAs in humans, chicken, mice, fish, and clams (Pozzi et al. 2017; Riggs et al. 2018; Mercer et al. 2011; Larriba, Rial, and Del Mazo 2018; Bottje et al. 2017; Ro et al. 2013). This degree of conservation in the presence of these RNAs suggests that the ability to transcribe small RNAs from the mtDNA is conserved well beyond the Amniotes, extending to the origin of *Protostomia* around 600 Mya. Furthermore, we note a recent report of small mitochondrial RNAs in plants (Stone et al. 2015). Thus, we speculate that the ability of the mtDNA to transcribe small RNAs may date back to the early origins of eukaryogenesis, when the mitochondrial-eukaryote endosymbiosis was still in the incipient stages of evolution. While almost without evidence, such a possibility is intriguing, and worthy of further investigation.

Characteristics that are evolutionary conserved across many species tend to be linked to specific functions, since patterns of sequence conservation are shaped by strong selection (Margoliash 1963; Zuckerkandl and Pauling 1965; Kimura 1968). The fact that numerous metazoan species are able to transcribe small mitochondrial RNAs therefore provides hints of likely functionality. tRNAs are known for being one of the source for microRNA-like molecules named tRNA-derived RNA fragments (tRFs). These tRFs seem to be produced by both nuclear and organelle tRNAs. We hypothesize that such functionality would likely come through the involvement of these small RNAs in the RNAi mechanism. RNAi represents one of the most conserved molecular mechanisms among eukaryotes, the function of which is regulated by small non-coding nuclear RNAs of the same length of the small mitochondrial RNAs (Ha and Kim 2014). Importantly, our study shows that most small mitochondrial RNAs originate from the mt-tRNAs, aligning well with a previous study of human RNA that found evidence of a mt-tRNA binding to *Ago2* (Ha and Kim 2014). *Ago2* is a key protein involved in the regulation of RNAi, and which is present outside of the mitochondria and within the cytoplasm (Maniataki and Mourelatos 2005). Further connection to the tRNAs is represented by their general involvement with RNAi (Lee et al. 2009; Cloonan 2015). In fact, tRNAs encoded by the nucleus are a known source of small RNAs (~30nt) involved in RNAi, thus suggesting that the small mitochondrial RNAs might be generated in similar manner. Supporting this hypothesis, a recent study showed that organelles in plants also encode tRNA-derived small RNAs (Cognat et al. 2017). These results match the predictions made by Pozzi *et al.* (2017), who upon analysis of levels of expression, sequence lengths, and complementarity to target mRNAs, proposed a putative role for a subset of highly transcribed small mitochondrial RNAs in RNAi. Therefore, we argue that the presence of polymorphisms in sequences harbouring small RNAs may alter their function through interfering with the complementarity between small RNA and mRNA target.

Conclusion

We have presented several lines of support for the hypothesis that the mitochondrial genome is able to consistently transcribe small RNAs, across species, and in a tissue-specific manner. The next frontier is now to home in on the question of whether these small RNAs play a functional role in the regulation of the organismal phenotype via mitochondrial-nuclear sequence interactions.

Material and Methods

Data

The data for the first analysis, which includes samples from the brain, cerebellum, heart, kidney, and testis of several species (*Monodelphis domestica*, *Homo sapiens*, *Macaca mulatta*, *Ornithorhynchus anatinus*, *Mus musculus*, and *Gallus gallus*) are published in (Meunier et al. 2013), and are accessible in the Short Reads Archive in NCBI with the ID PRJNA174234. All individuals analysed were males, and in all mammals the brain samples originated from the prefrontal cortex. Full information and origin of all samples are described in Meunier et al. 2013.

To compare the correlation between the NUMT content and small mitochondrial RNAs abundance, we used the data from over 10 samples for three different tissues in two different species: *Mus musculus* and *Gallus gallus*. The data from *Mus musculus*, which includes samples from the brain, liver, and heart, comes from several studies. The 13 brain samples include four from PRJNA232648 (Hu et al. 2014), five from PRJNA283972 (Malmevik et al. 2016), and four from PRJNA326768 (Woldemichael et al. 2016). As described by the articles just cited, the samples originated from hippocampal neurons and neuroblastoma cell cultures. The 14 liver samples include one from PRJNA203543 (Yamtich et al. 2015), one from PRJNA401785, and 12 from PRJNA407374 (Hao and Waxman 2018). The original articles did not specify any specific section of the liver used for the sequencing. The 12 heart samples include three from PRJNA219640 (Crippa et al. 2016), seven from PRJNA314812 (Ooi et al. 2017), and two from PRJNA421014 (Huang et al. 2018). As described in the original articles, all samples were obtained from mouse heart ventricles. The data from *Gallus gallus*, which includes samples from brain, liver, and heart, comes from several studies. The 13 brain samples are from PRJNA396511 (Warnefors et al. 2017). The 13 liver samples include 10 from PRJNA396511 (Warnefors et al. 2017) and three from PRJNA434773. The 14 heart samples include 12 from PRJNA396511 (Warnefors et al. 2017), and two from PRJNA204941. Details on the sections from which these samples were extracted are unavailable.

To study the correlation between the mtDNA content and the abundance of small mitochondrial RNAs, we used data from *Homo sapiens*, which includes samples from brain and bladder from several studies. The 36 brain samples comprised 11 from PRJNA272617 (Hoss et al. 2015), and 25 from PRJNA394722 (Pantazatos et al. 2017). All the tissues were samples from the prefrontal cortex. The 10 bladder samples (5 controls and 5 cancer) are from PRJDB2583 (Itesako et al. 2014).

NUMT sequence estimation

The amount of NUMT sequence per genome of each species was extracted from a previously published article (Hazkani-Covo, Zeller, and Martin 2010). These authors had documented NUMT content by aligning the mitochondrial genome of each species on the respective nuclear genome using *blastn* with a cut-off value of 0.0001 (e-score). For our study, we selected only high quality genomes, thus eliminating the chance of mischaracterization of NUMTs due to assembly problems. The amount of NUMTs for each species (absolute and relative to the nuclear genome) is: chicken (1.52Kb, 0.0001%), mouse (37.67Kb, 0.0015%), platypus (244.198Kb, 0.0081%), monkey (261.622Kb, 0.0087%), human (266.478Kb, 0.0087%), opossum (2093.63Kb, 0.0698%).

The amount of NUMTs in each species is not always stable. The NUMT amount might differ slightly across individuals analysed within a given species, however this possibility cannot be tested without specific genomic data from the individuals analysed. Nonetheless, we expect that variability in NUMT amount across individuals within a species will be far lower than levels of variability across species. The differences between chicken and other high-NUMT amount species is between ~15 (mouse) and ~700 fold (opossum). Thus, we argue that such variability should not affect our results.

The genome sequences of the individuals included in this study are likely to be slightly different from the reference genomes used. In fact, both NUMTs and mtDNA show slight variation across individuals. However, we estimate that any such variation is likely to have had at most a negligible effect on our study. Indeed, we accounted for these differences by using non-stringent criteria during the alignment: a relatively small seed and allowing for a mismatch of one nucleotide in the sequence. Thus, unless multiple SNPs would arise in the short sequences encoding the small mitochondrial RNAs under study, the effects would likely be negligible.

Library Analysis

To obtain the percentage of reads mapping exclusively to the mtDNA (denoted ***mtDNA-only reads***) versus those that map both to NUMTS and mtDNA (denoted ***NUMT reads***), we applied a custom pipeline for each of the samples. At first, we aligned the RNA library to the reference mitochondrial genome of the species of interest, obtaining the ***mtDNA reads*** (these are all reads that map to mtDNA, both mtDNA-only and the NUMT reads). Then, we aligned the mtDNA reads onto the nuclear genome to obtain the reads mapping both to the nuclear and mitochondrial genome; which we defined as the NUMT reads (i.e. these reads mapped to both the mtDNA and to NUMT sequence). The reads of each library were aligned to their NCBI reference genomes using Bowtie2 (Langmead and Salzberg 2012), using the default settings of the --*local* function and a seed of 18 nucleotides with one mismatch allowed (-*n* 1). From the output of Bowtie2, we retained the overall percentage of reads mapping to the mitochondrial and nuclear (NUMTs) genomes. All the format conversion of the datasets during the analysis was done using standard approaches with *samtools* (Li et al. 2009) and *bedtools* (Quinlan and Hall 2010). The results are plotted using MATLAB and Statistics Toolbox Release 2018b, The MathWorks, Inc., Natick, Massachusetts, United States. To identify hotspots of transcription, the alignment files (mtDNA reads) for each sample were converted to a BedGraph format using *bedtools* (Quinlan and Hall 2010), and all samples from each tissue were merged using the *cat* function in the terminal. These data were then plotted using the package *Circlize* in R (Gu et al. 2014).

Statistical Analysis

The percentages of aligned small mitochondrial RNAs (mtDNA reads) extracted from (Meunier et al. 2013) were grouped first by species, creating six groups in which each data point represented the value of a different tissue from the same species (brain, cerebellum, heart, kidney, and testis). A formal statistical test of these data was not appropriate, because there is no independent replication of data points for each tissue in each species. The mtDNA content changes drastically among tissues, thus pooling the tissues in each species to perform a formal statistical test would result in an unreliable result.

The analyses of the chicken and mouse samples (heart, brain, liver), as well as the human samples (brain and bladder), were performed using the same tools. We conducted the analysis using standard functions in R and Past3 (https://folk.uio.no/ohammer/past/). The percentage of mtDNA reads aligned from each sample was clustered by tissue type. The distributions obtained were tested for normality using the Shapiro-Wilk test (Shapiro and Wilk 1965), and in every case, the null hypothesis was rejected. Therefore, to test the difference between each distribution we used a two way Mann-Whitney U test (Mann and Whitney 1947).

Funding

Australian Research Council, through grants DP170100165 and FT160100022, to D.K.D.

Author's contribution

A.P and D.K.D. conceived the study, and A.P. performed all the computational analysis and prepared the figures. A.P and D.K.D. interpreted the results and wrote the manuscript.

Acknowledgements

We wish to acknowledge all researchers that made their data freely available, thus giving us the chance to perform this work. Furthermore, we would like to thank Paul Sunnucks for the insightful discussions that improved this work.

References

Ambros, Victor. 2004. "The Functions of Animal microRNAs." *Nature* 431 (7006): 350–55.

Avise, John C., Jonathan Arnold, R. Martin Ball, Eldredge Bermingham, Trip Lamb, Joseph E. Neigel, Carol A. Reeb, and Nancy C. Saunders. 1987. "Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics." *Annual Review of Ecology and Systematics* 18: 489–522.

Avise, John C. 2000. *Phylogeography: The History and Formation of Species*. Harvard University Press.

Ballard, J. William O., and Nicolas Pichaud. 2014. "Mitochondrial DNA: More than an Evolutionary Bystander." Edited by Charles Fox. *Functional Ecology* 28 (1): 218–31.

Ballard, J. William O., and Michael C. Whitlock. 2004. “The Incomplete Natural History of Mitochondria.” *Molecular Ecology* 13 (4): 729–44.

Ballard, J. W., and M. Kreitman. 1995. “Is Mitochondrial DNA a Strictly Neutral Marker?” *Trends in Ecology & Evolution* 10 (12): 485–88.

Bandelt, H. J., P. Forster, B. C. Sykes, and M. B. Richards. 1995. “Mitochondrial Portraits of Human Populations Using Median Networks.” *Genetics* 141 (2): 743–53.

Bazin, Eric, Sylvain Glémin, and Nicolas Galtier. 2006. “Population Size Does Not Influence Mitochondrial Genetic Diversity in Animals.” *Science* 312 (5773): 570–72.

Bensasson, Douda, Marcus W. Feldman, and Dmitri A. Petrov. 2003. “Rates of DNA Duplication and Mitochondrial DNA Insertion in the Human Genome.” *Journal of Molecular Evolution* 57 (3): 343–54.

Bottje, Walter G., Bhuwan Khatri, Stephanie A. Shouse, Dongwon Seo, Barbara Mallmann, Sara K. Orlowski, Jeonghoon Pan, et al. 2017. “Identification and Differential Abundance of Mitochondrial Genome Encoding Small RNAs (mitosRNA) in Breast Muscles of Modern Broilers and Unselected Chicken Breed.” *Frontiers in Physiology* 8 (October): 816.

Calabrese, Francesco Maria, Domenico Simone, and Marcella Attimonelli. 2012. “Primates and Mouse NumtS in the UCSC Genome Browser.” *BMC Bioinformatics* 13 Suppl 4 (March): S15.

Camus, M. Florencia, Jonci N. Wolff, Carla M. Sgrò, and Damian K. Dowling. 2017. “Experimental Support That Natural Selection Has Shaped the Latitudinal Distribution of Mitochondrial Haplotypes in Australian Drosophila Melanogaster.” *Molecular Biology and Evolution* 34 (10): 2600–2612.

Chalkia, Dimitra, Yi-Cheng Chang, Olga Derbeneva, Maria Lvova, Ping Wang, Dan Mishmar, Xiaogang Liu, Larry N. Singh, Lee-Ming Chuang, and Douglas C. Wallace. 2018. “Mitochondrial DNA Associations with East Asian Metabolic Syndrome.” *Biochimica et Biophysica Acta* 1859 (9): 878–92.

Cloonan, Nicole. 2015. “Re-Thinking miRNA-mRNA Interactions: Intertwining Issues Confound Target Discovery.” *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 37 (4): 379–88.

Cognat, Valérie, Geoffrey Morelle, Cyrille Megel, Stéphanie Lalande, Jean Molinier, Timothée Vincent, Ian Small, Anne-Marie Duchêne, and Laurence Maréchal-Drouard. 2017. “The Nuclear and Organellar tRNA-Derived RNA Fragment Population in Arabidopsis Thaliana Is Highly Dynamic.” *Nucleic Acids Research* 45 (6): 3460–72.

Crippa, Stefania, Mohamed Nemir, Samir Ounzain, Mark Ibberson, Corinne Berthonneche, Alexandre Sarre, Gaëlle Boisset, et al. 2016. “Comparative Transcriptome Profiling of the Injured Zebrafish and Mouse Hearts Identifies miRNA-Dependent Repair Pathways.” *Cardiovascular Research* 110 (1): 73–84.

Davey, Gavin P., Stefan Peuchen, and John B. Clark. 1998. “Energy Thresholds in Brain Mitochondria: POTENTIAL INVOLVEMENT IN NEURODEGENERATION.” *The Journal of Biological Chemistry* 273 (21): 12753–57.

Dowling, Damian K. 2014. “Evolutionary Perspectives on the Links between Mitochondrial Genotype and Disease Phenotype.” *Biochimica et Biophysica Acta* 1840 (4): 1393–1403.

Dowling, Damian K., Daniel M. Tompkins, and Neil J. Gemmell. 2015. “The Trojan Female Technique for Pest Control: A Candidate Mitochondrial Mutation Confers Low Male Fertility across Diverse Nuclear Backgrounds in Drosophila Melanogaster.” *Evolutionary Applications* 8 (9): 871–80.

Estopinal, Christopher B., Isaac M. Chocron, Megan B. Parks, Emily A. Wade, Rachel M. Roberson, L. Goodwin Burgess, Milam A. Brantley Jr, and David C. Samuels. 2014. “Mitochondrial Haplogroups Are Associated with Severity of Diabetic Retinopathy.” *Investigative Ophthalmology & Visual Science* 55 (9): 5589–95.

Florencia Camus, M., Jochen B. W. Wolf, Edward H. Morrow, and Damian K. Dowling. 2015. “Single Nucleotides in the mtDNA Sequence Modify Mitochondrial Molecular Function and Are Associated with Sex-Specific Effects on Fertility and Aging.” *Current Biology: CB* 25 (20): 2717–22.

Gaziev, A. I., and G. O. Shaikhaev. 2010. “Nuclear Mitochondrial Pseudogenes.” *Molecular Biology* 44 (3): 358–68.

Gibson, P. R., R. P. Anderson, J. M. Mariadason, and A. J. Wilson. 1996. “Protective Role of the Epithelium of the Small Intestine and Colon.” *Inflammatory Bowel Diseases* 2 (4): 279–302.

Gu, Zuguang, Lei Gu, Roland Eils, Matthias Schlesner, and Benedikt Brors. 2014. “Circlize Implements and Enhances Circular Visualization in R.” *Bioinformatics* 30 (19): 2811–12.

Ha, Minju, and V. Narry Kim. 2014. “Regulation of microRNA Biogenesis.” *Nature Reviews. Molecular Cell Biology* 15 (8): 509–24.

Hao, Pengying, and David J. Waxman. 2018. “Functional Roles of Sex-Biased, Growth Hormone-Regulated MicroRNAs miR-1948 and miR-802 in Young Adult Mouse Liver.” *Endocrinology* 159 (3): 1377–92.

Harrison, J. Scott, and Ronald S. Burton. 2006. “Tracing Hybrid Incompatibilities to Single Amino Acid Substitutions.” *Molecular Biology and Evolution* 23 (3): 559–64.

Hazkani-Covo, Einat, Raymond M. Zeller, and William Martin. 2010. “Molecular Poltergeists: Mitochondrial DNA Copies (numts) in Sequenced Nuclear Genomes.” *PLoS Genetics* 6 (2): e1000834.

Hogan, Daniel J., Daniel P. Riordan, André P. Gerber, Daniel Herschlag, and Patrick O. Brown. 2008. “Diverse RNA-Binding Proteins Interact with Functionally Related Sets of RNAs, Suggesting an Extensive Regulatory System.” *PLoS Biology* 6 (10): e255.

Hopkins, Julia F., Veronica Y. Sabelnykova, Joachim Weischenfeldt, Ronald Simon, Jennifer A. Aguiar, Rached Alkallas, Lawrence E. Heisler, et al. 2017. “Mitochondrial Mutations Drive Prostate Cancer Aggression.” *Nature Communications* 8 (1): 656.

Hoss, Andrew G., Adam Labadorf, Jeanne C. Latourelle, Vinay K. Kartha, Tiffany C. Hadzi, James F. Gusella, Marcy E. MacDonald, et al. 2015. “miR-10b-5p Expression in Huntington’s Disease Brain Relates to Age of Onset and the Extent of Striatal Involvement.” *BMC Medical Genomics* 8 (March): 10.

Houseley, Jonathan, and David Tollervey. 2009. “The Many Pathways of RNA Degradation.” *Cell* 136 (4): 763–76.

Huang, Wei, Yuliang Feng, Jialiang Liang, Hao Yu, Cheng Wang, Boyu Wang, Mingyang Wang, et al. 2018. “Loss of microRNA-128 Promotes Cardiomyocyte Proliferation and Heart Regeneration.” *Nature Communications* 9 (1): 700.

Hudson, Gavin, Aurora Gomez-Duran, Ian J. Wilson, and Patrick F. Chinnery. 2014. “Recent Mitochondrial DNA Mutations Increase the Risk of Developing Common Late-Onset Human Diseases.” *PLoS Genetics* 10 (5): e1004369.

Hu, Zhonghua, Danni Yu, Qin-Hua Gu, Yanqin Yang, Kang Tu, Jun Zhu, and Zheng Li. 2014. “miR-191 and miR-135 Are Required for Long-Lasting Spine Remodelling Associated with Synaptic Long-Term Depression.” *Nature Communications* 5: 3263.

Innocenti, Paolo, Edward H. Morrow, and Damian K. Dowling. 2011. "Experimental Evidence Supports a Sex-Specific Selective Sieve in Mitochondrial Genome Evolution." *Science* 332 (6031): 845–48.

Itesako, Toshihiko, Naohiko Seki, Hirofumi Yoshino, Takeshi Chiyomaru, Takeshi Yamasaki, Hideo Hidaka, Tomokazu Yonezawa, et al. 2014. "The microRNA Expression Signature of Bladder Cancer by Deep Sequencing: The Functional Significance of the miR-195/497 Cluster." *PloS One* 9 (2): e84311.

James, Jennifer E., Gwenael Piganeau, and Adam Eyre-Walker. 2016. "The Rate of Adaptive Evolution in Animal Mitochondria." *Molecular Ecology* 25 (1): 67–78.

Kimura, M. 1968. "Evolutionary Rate at the Molecular Level." *Nature* 217 (5129): 624–26.

Langmead, Ben, and Steven L. Salzberg. 2012. "Fast Gapped-Read Alignment with Bowtie 2." *Nature Methods* 9 (4): 357–59.

Larriba, Eduardo, Eduardo Rial, and Jesús Del Mazo. 2018. "The Landscape of Mitochondrial Small Non-Coding RNAs in the PGCs of Male Mice, Spermatogonia, Gametes and in Zygotes." *BMC Genomics* 19 (1): 634.

Lee, Yong Sun, Yoshiyuki Shibata, Ankit Malhotra, and Anindya Dutta. 2009. "A Novel Class of Small RNAs: tRNA-Derived RNA Fragments (tRFs)." *Genes & Development* 23 (22): 2639–49.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25 (16): 2078–79.

Lopez, J. V., N. Yuhki, R. Masuda, W. Modi, and S. J. O'Brien. 1994. "Numt, a Recent Transfer and Tandem Amplification of Mitochondrial DNA to the Nuclear Genome of the Domestic Cat." *Journal of Molecular Evolution* 39 (2): 174–90.

Malmevik, Josephine, Rebecca Petri, Pina Knauff, Per Ludvik Brattås, Malin Åkerblom, and Johan Jakobsson. 2016. "Distinct Cognitive Effects and Underlying Transcriptome Changes upon Inhibition of Individual miRNAs in Hippocampal Neurons." *Scientific Reports* 6 (January): 19879.

Maniataki, Elisavet, and Zissimos Mourelatos. 2005. "Human Mitochondrial tRNAMet Is Exported to the Cytoplasm and Associates with the Argonaute 2 Protein." *RNA* 11 (6): 849–52.

Mann, H. B., and D. R. Whitney. 1947. "On a Test of Whether One of Two Random Variables Is Stochastically Larger than the Other." *Annals of Mathematical Statistics* 18 (1): 50–60.

Margoliash, E. 1963. "PRIMARY STRUCTURE AND EVOLUTION OF CYTOCHROME C." *Proceedings of the National Academy of Sciences of the United States of America* 50 (October): 672–79.

Marom, Shani, Michael Friger, and Dan Mishmar. 2017. "MtDNA Meta-Analysis Reveals Both Phenotype Specificity and Allele Heterogeneity: A Model for Differential Association." *Scientific Reports* 7 (February): 43449.

Mercer, Tim R., Shane Neph, Marcel E. Dinger, Joanna Crawford, Martin A. Smith, Anne-Marie J. Shearwood, Eric Haugen, et al. 2011. "The Human Mitochondrial Transcriptome." *Cell* 146 (4): 645–58.

Meunier, Julien, Frédéric Lemoine, Magali Soumillon, Angélica Liechti, Manuela Weier, Katerina Guschanski, Haiyang Hu, Philipp Khaitovich, and Henrik Kaessmann. 2013. "Birth and Expression Evolution of Mammalian microRNA Genes." *Genome Research* 23 (1): 34–45.

Morales, Hernán E., Alexandra Pavlova, Nevil Amos, Richard Major, Andrzej Kilian, Chris Greening, and Paul Sunnucks. 2018. "Concordant Divergence of Mitogenomes and a Mitonuclear Gene Cluster in Bird Lineages Inhabiting Different Climates." *Nature Ecology & Evolution* 2 (8): 1258–67.

Mossman, Jim A., Leann M. Biancani, Chen-Tseh Zhu, and David M. Rand. 2016. "Mitonuclear Epistasis for Development Time and Its Modification by Diet in Drosophila." *Genetics* 203 (1): 463–84.

Ooi, Jenny Y. Y., Bianca C. Bernardo, Saloni Singla, Natalie L. Patterson, Ruby C. Y. Lin, and Julie R. McMullen. 2017. "Identification of miR-34 Regulatory Networks in Settings of Disease and antimiR-Therapy: Implications for Treating Cardiac Pathology and Other Diseases." *RNA Biology* 14 (5): 500–513.

Oven, Mannis van. 2010. "Revision of the mtDNA Tree and Corresponding Haplogroup Nomenclature." *Proceedings of the National Academy of Sciences of the United States of America* 107 (11): E38–39.

Pantazatos, S. P., Y-Y Huang, G. B. Rosoklija, A. J. Dwork, V. Arango, and J. J. Mann. 2017. "Whole-Transcriptome Brain Expression and Exon-Usage Profiling in Major Depression and Suicide: Evidence for Altered Glial, Endothelial and ATPase Activity." *Molecular Psychiatry* 22 (5): 760–73.

Patel, Maulik R., Ganesh K. Miriyala, Aimee J. Littleton, Heiko Yang, Kien Trinh, Janet M. Young, Scott R. Kennedy, Yukiko M. Yamashita, Leo J. Pallanck, and Harmit S. Malik. 2016. "A Mitochondrial DNA Hypomorph of Cytochrome Oxidase Specifically Impairs Male Fertility in Drosophila Melanogaster." *eLife* 5 (August): e16923.

Picard, Martin, and Bruce S. McEwen. 2014. "Mitochondria Impact Brain Function and Cognition." *Proceedings of the National Academy of Sciences of the United States of America* 111 (1): 7–8.

Pozzi, Andrea, Federico Plazzi, Liliana Milani, Fabrizio Ghiselli, and Marco Passamonti. 2017. "SmithRNAs: Could Mitochondria 'Bend' Nuclear Regulation?" *Molecular Biology and Evolution*, April. https://doi.org/10.1093/molbev/msx140.

Quinlan, Aaron R., and Ira M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics* 26 (6): 841–42.

Reznik, Ed, Martin L. Miller, Yasin Şenbabaoğlu, Nadeem Riaz, Judy Sarungbam, Satish K. Tickoo, Hikmat A. Al-Ahmadie, et al. 2016. "Mitochondrial DNA Copy Number Variation across Human Cancers." *eLife* 5 (February). https://doi.org/10.7554/eLife.10769.

Riggs, Claire L., Amanda Summers, Daniel E. Warren, Göran E. Nilsson, Sjannie Lefevre, W. W. Dowd, Sarah Milton, and Jason E. Podrabsky. 2018. "Small Non-Coding RNA Expression and Vertebrate Anoxia Tolerance." *Frontiers in Genetics* 9 (July): 230.

Roach, Jared C., Gustavo Glusman, Arian F. A. Smit, Chad D. Huff, Robert Hubley, Paul T. Shannon, Lee Rowen, et al. 2010. "Analysis of Genetic Inheritance in a Family Quartet by Whole-Genome Sequencing." *Science* 328 (5978): 636–39.

Rogers, Hubert H., and Sam Griffiths-Jones. 2012. "Mitochondrial Pseudogenes in the Nuclear Genomes of Drosophila." *PloS One* 7 (3): e32593.

Rooney, John P., Ian T. Ryde, Laurie H. Sanders, Evan H. Howlett, Meryl D. Colton, Kaylyn E. Germ, Greg D. Mayer, J. Timothy Greenamyre, and Joel N. Meyer. 2015. "PCR Based Determination of Mitochondrial DNA Copy Number in Multiple Species." *Methods in Molecular Biology* 1241: 23–38.

Ro, Seungil, Hsiu-Yen Ma, Chanjae Park, Nicole Ortogero, Rui Song, Grant W. Hennig, Huili

Zheng, et al. 2013. "The Mitochondrial Genome Encodes Abundant Small Noncoding RNAs." *Cell Research* 23 (6): 759–74.
Sanchez, Maria I. G. Lopez, Tim R. Mercer, Stefan M. K. Davies, Anne-Marie J. Shearwood, Karoline K. A. Nygård, Tara R. Richman, John S. Mattick, Oliver Rackham, and Aleksandra Filipovska. 2011. "RNA Processing in Human Mitochondria." *Cell Cycle* 10 (17): 2904–16.
Shapiro, S. S., and M. B. Wilk. 1965. "An Analysis of Variance Test for Normality (Complete Samples)." *Biometrika* 52 (3/4): 591–611.
Singh, Keshav K., and Mariola Kulawiec. 2009. "Mitochondrial DNA Polymorphism and Risk of Cancer." *Methods in Molecular Biology* 471: 291–303.
Sloan, Daniel B., Peter D. Fields, and Justin C. Havird. 2015. "Mitonuclear Linkage Disequilibrium in Human Populations." *Proceedings. Biological Sciences / The Royal Society* 282 (1815). https://doi.org/10.1098/rspb.2015.1704.
Smits, Paulien, Jan Smeitink, and Lambert van den Heuvel. 2010. "Mitochondrial Translation and beyond: Processes Implicated in Combined Oxidative Phosphorylation Deficiencies." *Journal of Biomedicine & Biotechnology* 2010 (April): 737385.
Storz, Gisela. 2002. "An Expanding Universe of Noncoding RNAs." *Science* 296 (5571): 1260–63.
Takasaki, Shigeru. 2009. "Mitochondrial Haplogroups Associated with Japanese Alzheimer's Patients." *Journal of Bioenergetics and Biomembranes* 41 (5): 407–10.
Tourmente, Maximiliano, Misa Hirose, Saleh Ibrahim, Damian K. Dowling, Daniel M. Tompkins, Eduardo R. S. Roldan, and Neil J. Gemmell. 2017. "mtDNA Polymorphism and Metabolic Inhibition Affect Sperm Performance in Conplastic Mice." *Reproduction* 154 (4): 341–54.
Vaught, R. C., and D. K. Dowling. 2018. "Maternal Inheritance of Mitochondria: Implications for Male Fertility?" *Reproduction* 155 (4): R159–68.
Veltri, K. L., M. Espiritu, and G. Singh. 1990. "Distinct Genomic Copy Number in Mitochondria of Different Mammalian Organs." *Journal of Cellular Physiology* 143 (1): 160–64.
Vyas, Sejal, Elma Zaganjor, and Marcia C. Haigis. 2016. "Mitochondria and Cancer." *Cell* 166 (3): 555–66.
Warnefors, Maria, Katharina Mössinger, Jean Halbert, Tania Studer, John L. VandeBerg, Isa Lindgren, Amir Fallahshahroudi, Per Jensen, and Henrik Kaessmann. 2017. "Sex-Biased microRNA Expression in Mammals and Birds Reveals Underlying Regulatory Mechanisms and a Role in Dosage Compensation." *Genome Research* 27 (12): 1961–73.
Weihe, Andreas, Janina Apitz, Frank Pohlheim, Annabel Salinas-Hartwig, and Thomas Börner. 2009. "Biparental Inheritance of Plastidial and Mitochondrial DNA and Hybrid Variegation in Pelargonium." *Molecular Genetics and Genomics: MGG* 282 (6): 587–93.
Williams, Stephen B., Yuanqing Ye, Maosheng Huang, David W. Chang, Ashish M. Kamat, Xia Pu, Colin P. Dinney, and Xifeng Wu. 2015. "Mitochondrial DNA Content as Risk Factor for Bladder Cancer and Its Association with Mitochondrial DNA Polymorphisms." *Cancer Prevention Research* 8 (7): 607–13.
Woldemichael, Bisrat T., Ali Jawaid, Eloïse A. Kremer, Niharika Gaur, Jacek Krol, Antonin Marchais, and Isabelle M. Mansuy. 2016. "The microRNA Cluster miR-183/96/182 Contributes to Long-Term Memory in a Protein Phosphatase 1-Dependent Manner." *Nature Communications* 7 (August): 12594.
Yamtich, Jennifer, Seok-Jin Heo, Joseph Dhahbi, David I. K. Martin, and Dario Boffelli. 2015. "piRNA-like Small RNAs Mark Extended 3'UTRs Present in Germ and Somatic Cells." *BMC Genomics* 16 (June): 462.

Zhu, Xun, Yingying Mao, Tongtong Huang, Caiwang Yan, Fei Yu, Jiangbo Du, Juncheng Dai, Hongxia Ma, and Guangfu Jin. 2017. “High Mitochondrial DNA Copy Number Was Associated with an Increased Gastric Cancer Risk in a Chinese Population.” *Molecular Carcinogenesis* 56 (12): 2593–2600.

Zuckerkandl, Emile, and Linus Pauling. 1965. “Evolutionary Divergence and Convergence in Proteins.” In *Evolving Genes and Proteins*, edited by Vernon Bryson and Henry J. Vogel, 97–166. Academic Press.

Zuckerkandl, Emile, and Linus Pauling. 1965. “Evolutionary Divergence and Convergence in Proteins.” In *Evolving Genes and Proteins*, edited by Vernon Bryson and Henry J. Vogel, 97–166. Academic Press.

Figures

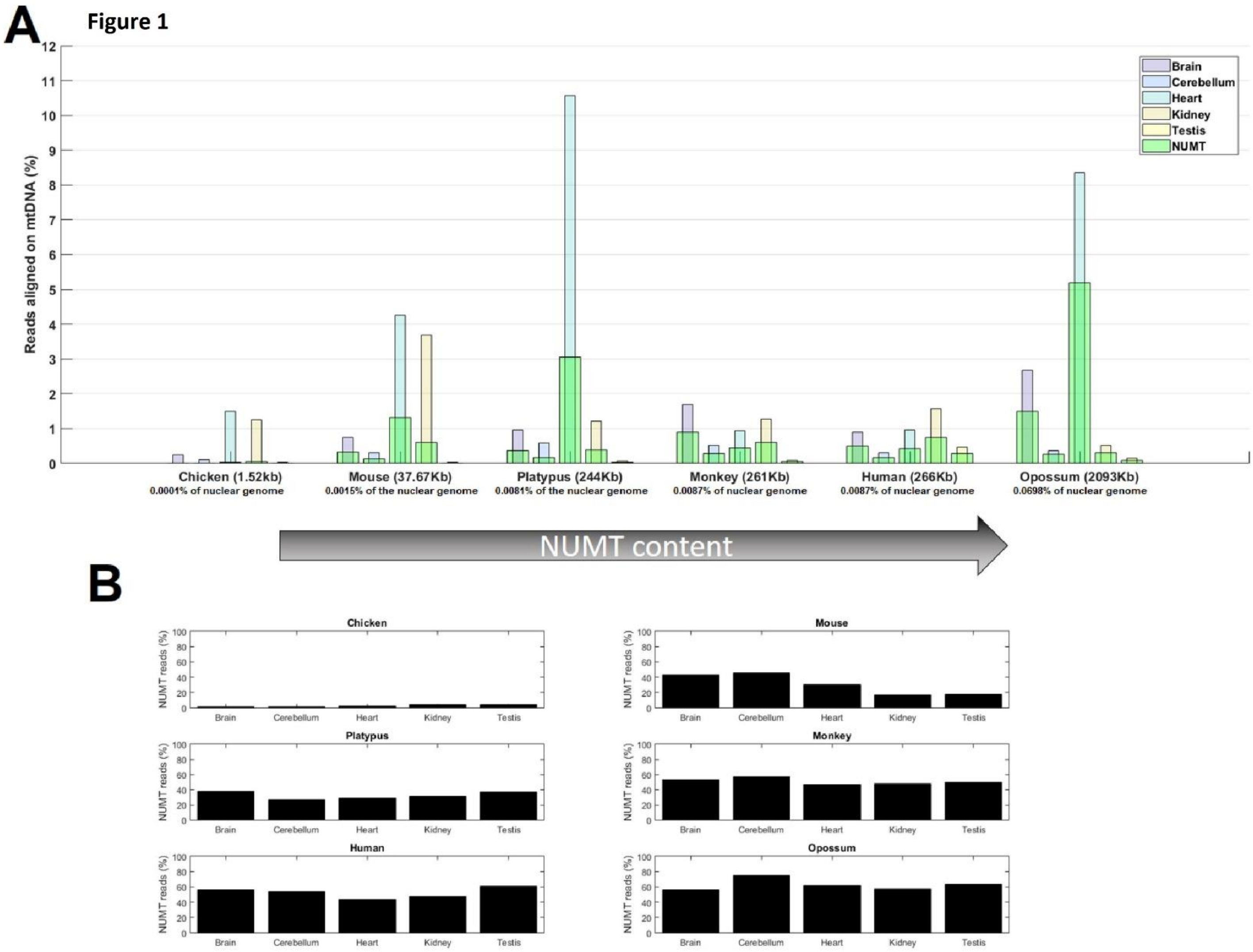


Fig.1 The abundance of small mitochondrial RNAs across species exhibiting different NUMT content. In (A), the percentage of small RNAs, relative to the total RNAs in the sample, that align exclusively to the mtDNA in five distinct tissues of six separate species. The NUMT content of each species is denoted in parentheses beside the species name, and the species are ordered by ascending NUMT content. Furthermore, the percentage of NUMTs present in their nuclear genome is denoted below the name of each species. Different tissues are represented by distinct colours: brain (violet), cerebellum (cyan), heart (light blue), kidney (light brown), and testis (yellow). The percentage of reads mapping to the mtDNA, both mtDNA-only and NUMT reads together, is shown on the Y-axis. On the same axis, the percentage of reads in the RNA libraries that map both to the mtDNA and to the nuclear genome (NUMTs reads) is highlighted in light green. For example, in the case of O. anatinus (platypus), the heart samples show ~3% of NUMT reads and ~7.5% of mtDNA-only reads, for a total of 10.5% mtDNA reads. In (B), the percentage of NUMT reads as a proportion of the overall mtDNA reads, for each sample. While in G. gallus, less than 10% of mtDNA reads map the NUMTs, in most samples of the other species almost half of the mtDNA reads map jointly to the NUMT sequence. Despite this, even in species with the highest level of NUMT sequence, around one-third of the mtDNA reads map uniquely to the mtDNA.

Figure 2

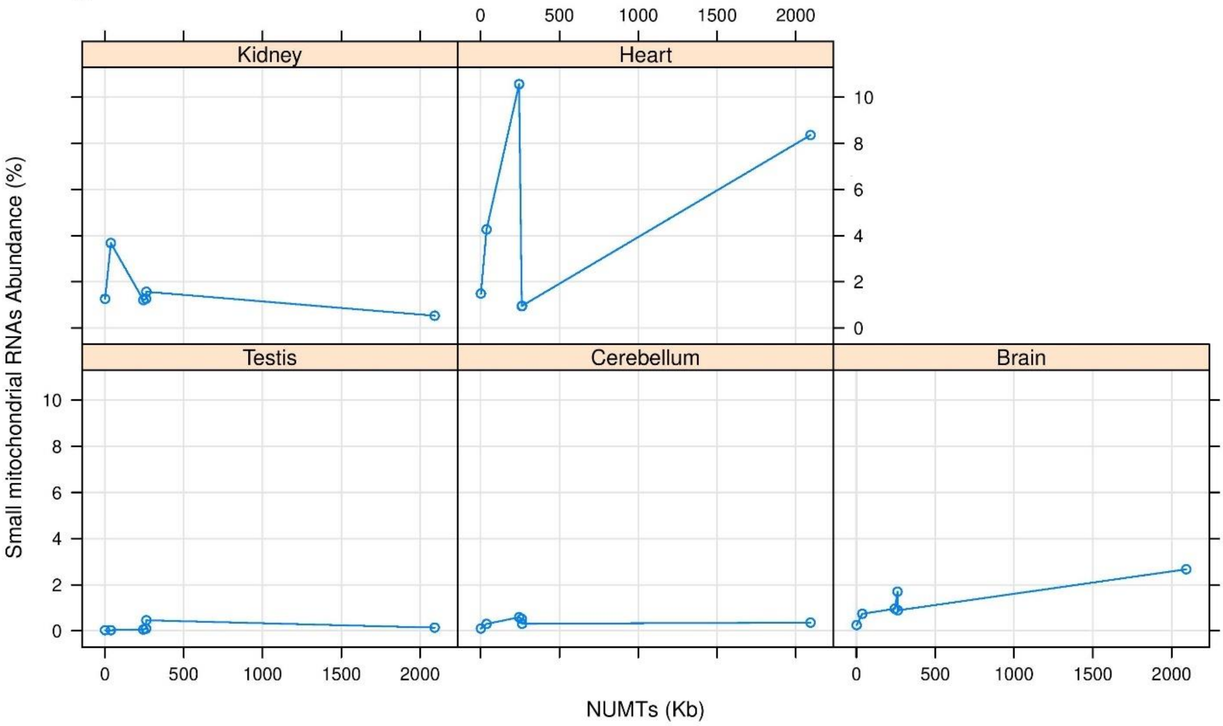


Fig.2 The relationship between the percentage of small mitochondrial RNAs in each tissue and NUMT content (Kb), across six species. There are six samples for each combination species/tissue, and each data point represents a sample of a single tissue from a single species. For example, in the top left plot we can see a data point at the 2000 mark on the X-axis, which indicates that small mitochondrial RNAs comprised less than 2% of the RNA library extracted from the kidney of a species with over 2000Kb NUMTs length (Opossum).

Figure 3

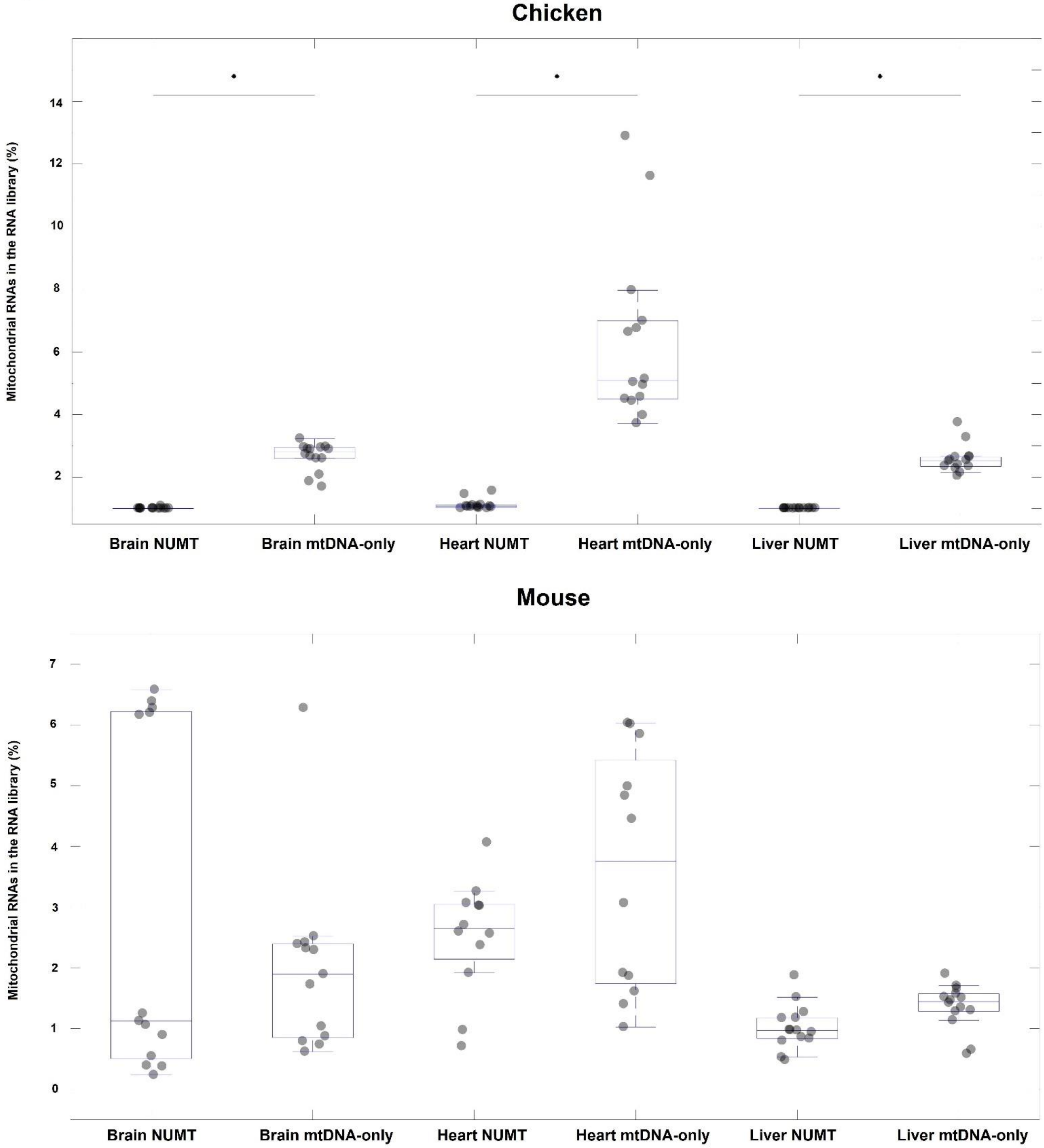


Fig.3 Boxplots of percentages of NUMT reads (those mapping both to the mtDNA and the nuclear genome) and mtDNA-only reads in the brain, heart, and liver, across samples taken from chicken and mouse. The percentage of reads in the RNA library mapping to the mtDNA is shown on the vertical axis. The NUMT boxes show the percentage of reads mapping jointly to both mtDNA and NUMT sequence (NUMT reads). The mtDNA-only boxes show reads mapping exclusively to the mtDNA and not to NUMT sequences (mtDNA-only reads). The horizontal line in each box indicates the median of the distribution, and the light grey circles indicate individual data points. The distributions were tested using Mann-Whitney U test and significance ($p < 0.05$) is indicated with an asterisk.

Figure 4

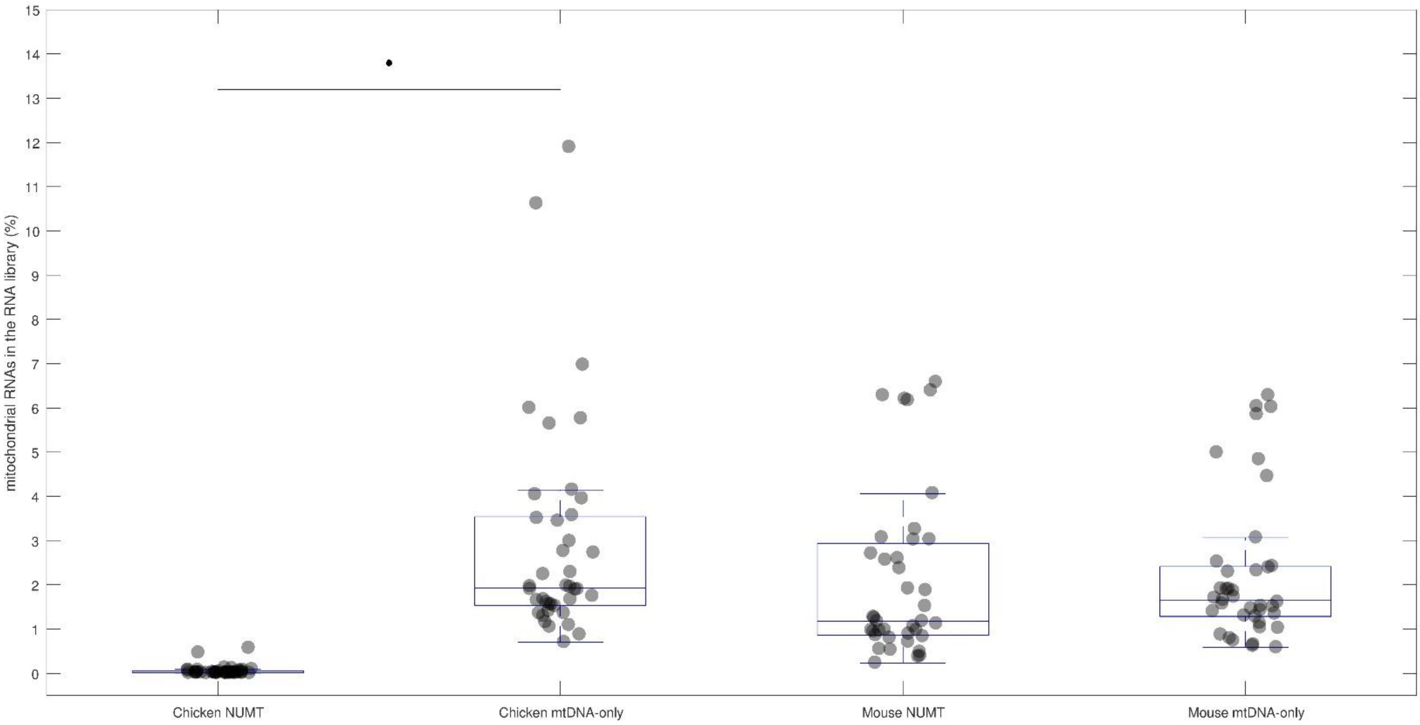


Fig.4 Boxplots showing the percentages of NUMT reads and mtDNA-only reads across all tissues pooled (heart, brain, and liver) of chicken and mouse. The percentage of reads in the RNA library mapping to the mtDNA is shown on the Y-axis. The NUMT boxes show the percentage of reads mapping to both mtDNA and NUMT sequence (NUMT reads). The mtDNA-only boxes show the reads mapping only the mtDNA and not the NUMTs (mtDNA-only reads). The horizontal line in each box indicates the median of the distribution, and the light grey circles indicate individual data points. The distributions were tested using Mann-Whitney U test and significance ($p < 0.05$) is indicated with an asterisk.

Figure 5

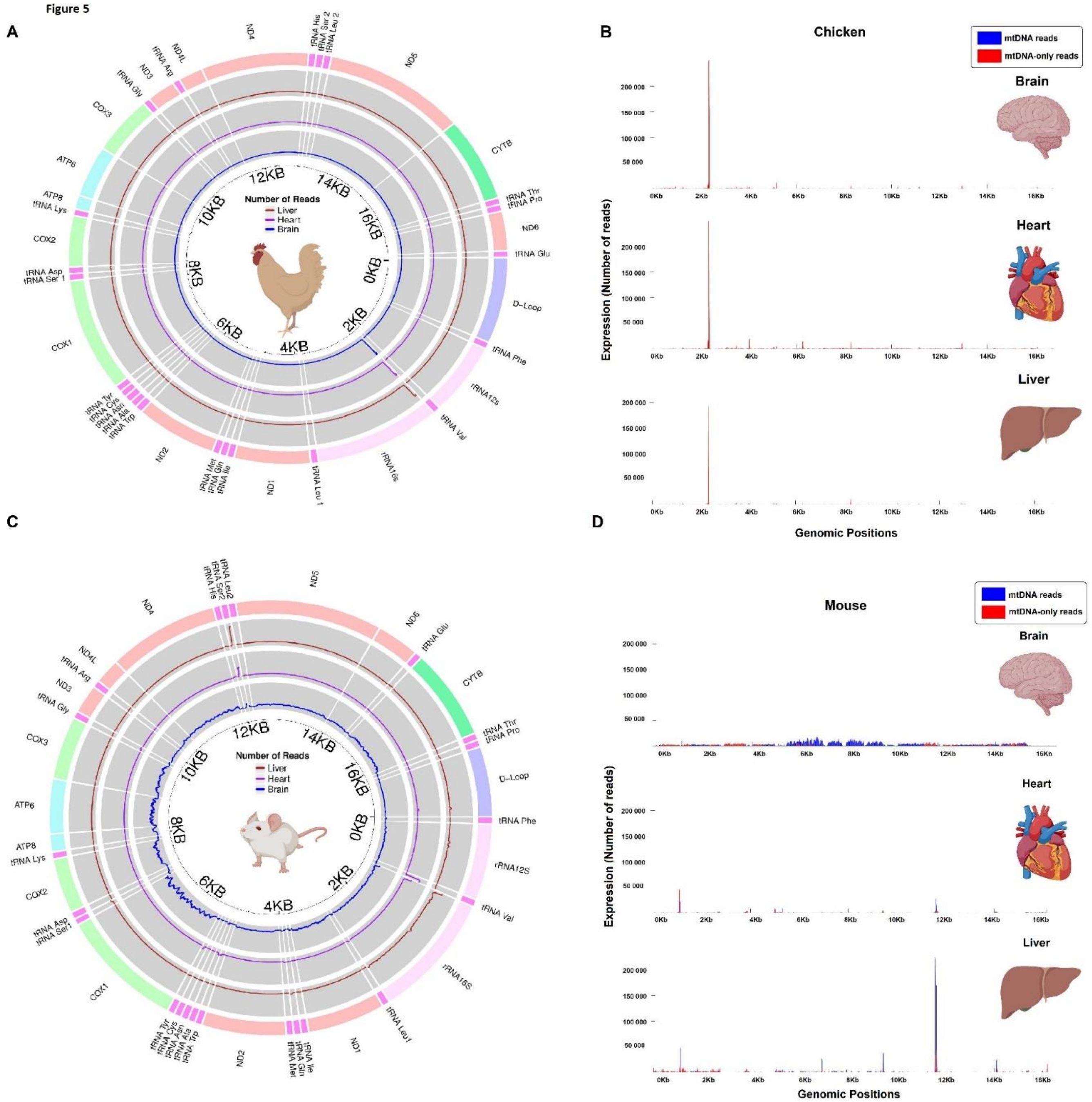


Fig.5 Expression profiles of small mitochondrial RNAs across three tissues of chicken (A & B) and mouse (C & D). Circular representations of (A) chicken mtDNA and (C) mouse mtDNA, in which three concentric circles show the number of reads aligning to each portion of the genome across the three tissue types. From the outer to the inner circle, we can see the expression of mtDNA reads in the liver (brown), heart (magenta) and brain (blue). The scale of each circle is not constant but changed accordingly to highlight the differences in expression among samples within a given tissue. However, within each tissue and species, the scale is constant across different regions of the mtDNA. The makers placed every 2kb show the approximate location of the genes in the genome. These markers are placed to facilitate the comparison with the linear representations. Linear representations of the chicken (B) and mouse (D) mtDNA, in which small mitochondrial RNAs are mapped to the sequence position, with both mtDNA reads (blue) and mtDNA-only reads (red). The portion of the mtDNA reads not covered by the mtDNA-only reads represent the sequences that map both the mtDNA and the nucleus (NUMT reads). The median number of reads mapping to a specific portion of the mtDNA represents the level of expression in this plot.

Figure 6

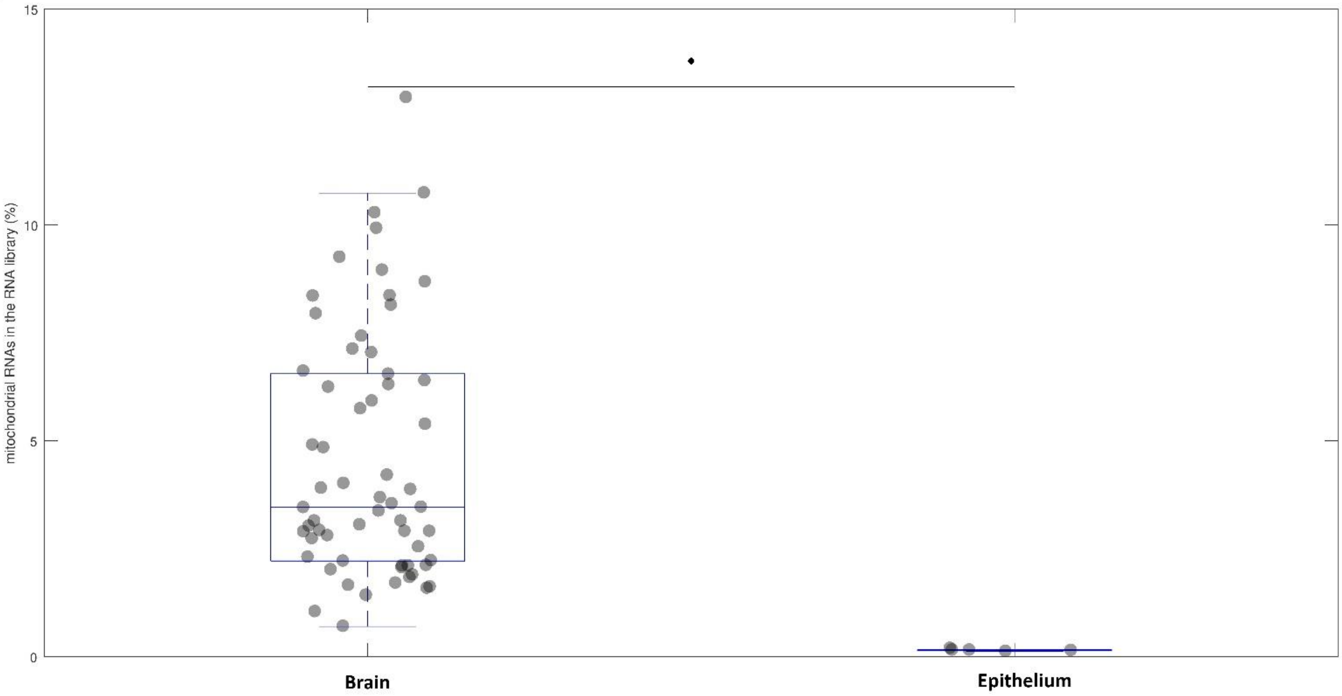


Fig.6 The percentage of small mitochondrial RNAs mapping to the mtDNA in the brain (putatively mitochondria-rich) and bladder epithelium (mitochondria-poor) tissue of humans. The percentage of mtDNA reads in the bladder epithelium can barely be seen because the amount is so low when compared to the brain mtDNA reads (less than 0.5% of all aligned reads). The horizontal line in each box indicates the median of the distribution, and the light grey circles indicate individual data points. The distributions were tested using Mann-Whitney U test and significance ($p < 0.05$) is indicated with an asterisk.

Figure 7

Fig.7 The percentage of small mitochondrial RNAs mapping to the mtDNA in two different types of cells: healthy (control) and cancerous (cancer). The horizontal line in each box indicates the median of the distribution, and the light grey circles indicate individual data points. The distributions were tested using Mann-Whitney U test and significance ($p < 0.05$) is indicated with an asterisk.

CHAPTER THREE

Mitochondria Interference: a role for small mitochondrial RNAs as mediators of mitochondrial diseases?

Title: Mitochondria Interference: a role for small mitochondrial RNAs as mediators of mitochondrial diseases?

Author 1: Andrea Pozzi, Monash University, School of

Biological Sciences, andrea.pozzi@monash.edu

Author 2: Damian K. Dowling, Monash University, School of

Biological Sciences, damian.dowling@monash.edu

Status: In Review – *Trends in Molecular Medicine*

Abstract

Investigation into the causes of mitochondrial diseases has centred on identifying mutations in the mtDNA sequence affecting mitochondrial respiratory capacity. We contend that pathogenic mitochondrial mutations will not always exert their negative effects through impairment of energy production via oxidative phosphorylation. Rather, we hypothesise that specific mutations could affect the function of a set of small RNAs transcribed within the mitochondrial DNA, perturbing communication between mitochondria and nucleus, leading to disease. Our hypothesis relies on two lines of evidence. Firstly, mutations underlying mitochondrial diseases often cannot be directly linked to negative effects on energy production or protein synthesis. Secondly, several studies have described the existence of functional small RNAs encoded by the mitochondria and speculated these RNAs may play influence the organismal phenotype through RNA interference. We synthesise these lines of evidence, presenting a roadmap to investigate the capacity of these small RNAs to act as mediators of mitochondrial diseases.

The Mitochondrial diseases

The mitochondria are intracellular organelles that have retained their own genome, encoding 37 known functional products. The mitochondrial genome encodes 13 proteins involved in oxidative phosphorylation (OXPHOS), and 24 functional RNAs involved in the synthesis of such proteins. These proteins act synergistically to produce ATP, and defects in one protein are sufficient to cause a deficiency in OXPHOS [1]. Likewise, mutations impairing the function of ribosomal and tRNAs affect the synthesis of these 13 proteins, leading to the impairment of OXPHOS function. Dysfunctional OXPHOS characterizes a group of genetic disorders defined as mitochondrial diseases [1,2]. Importantly, this definition assumes that in most cases the diseases will be underpinned by defective OXPHOS, thus leading to the hypothesis that ultimately all pathogenic mtDNA mutations linked to mitochondrial disease will affect the ability of the mitochondria to produce energy [1]. This hypothesis is grounded upon several empirical studies, which have demonstrated that several mitochondrial diseases (eg. <u>M</u>itochondrial <u>e</u>ncephalomyopathy, <u>l</u>actic acidosis, and stroke-like episodes, *MELAS*) are indeed associated with dysfunctional OXPHOS [3–5]. MELAS is arguably the best studied of mitochondrial diseases, a pathology where over 80% of the cases feature a mitochondrial mutation thought to impair mitochondrial protein translation in the mt-tRNA Leucine [5,6]. Indeed, authors using cell cultures demonstrated that the overexpression of nuclear proteins involved in protein translation partially suppress the MELAS phenotype [5]. Another example is the study on cerebellar ataxia, where the authors found that patients with specific mitochondrial mutations have dysfunctional OXPHOS [7]. Notwithstanding, while it is clear that several mtDNA mutations do negatively affect OXPHOS, and ultimately result in mitochondrial disease, there are several other cases for which striking discrepancies between theoretical predictions and empirical results remain unanswered. In this paper, we document and explore these discrepancies, and discuss an alternative mechanism by which mtDNA mutations might confer disease; a mechanism that is independent of direct alterations to OXPHOS function.

The first discrepancy: mitochondrial mutations and energy

The first discrepancy between theory and practice is that disease-associated mutations impairing energy production are usually located in mitochondrial genes that are not directly associated with the OXPHOS deficiency. We return to the case study of cerebellar ataxia described above for an example [7]. By analysing the functionality of OXPHOS in cultured cells from patients with cerebellar ataxia, the authors found that the protein complex most affected by pathogenic mutations, T14709C and A8344G, is Complex I. However, these mitochondrial mutations are in the mt-tRNAs Gln and Lys, both without any specific relation to proteins of Complex I. This finding is puzzling, especially because individuals carrying the mutation T14709C have been found to be healthy in several cases [8]. Indeed, although these mutations have been linked to several mitochondrial diseases, nothing is known as to the underpinning mechanism linking the mutated tRNA to mitochondrial disease [9]. Likewise, other studies present similar discrepancies when it comes to alignment of the mitochondrial genes harbouring the candidate pathogenic mutations and the OXPHOS complexes that are ultimately affected. For instance, a recent study investigated the mutation A3243G, arguably one of the most common mutations in mitochondrial diseases and, probably, the main cause of MELAs [10,11]. In this study on pigmentary retinopathy, the authors discovered that cells with this mutation produced as much energy as the healthy ones, albeit while still presenting pathogenic features [10]; thus providing a practical case whereby the A3243G mutation does not cause detectable decreases in energy production but are nonetheless still presumed to cause mitochondrial disease.

We contend that this discrepancy in the hypothesized universal association between disease-associated mtDNA mutations and OXPHOS function may provide insights into why the symptoms associated with disease-causing mtDNA mutations are so heterogeneous in their manifestation. Mitochondrial diseases such as MELAS include many symptoms, from stroke-like episodes, to diabetes mellitus and pigmentary retinopathy. Among the symptoms, heart-related symptoms are quite common and, due to the high energy demands of the heart, can easily be linked to a lack of energy [12]. According to the current paradigm, it is thought that the effects of pathogenic mtDNA mutations are heterogeneous due to tissue-specific differences in loads of mutant to wild type mtDNA, under heteroplasmy, meaning that the biochemical threshold, between functional and dysfunctional, is surpassed in some tissues but not others across a lifespan within individuals [1]. However, some mutations do not fit this paradigm. In fact, symptoms such as pigmentary retinopathy, or diabetes, affect cells not known for their high-energetic demands, and which are therefore likely to be less sensitive to mutations affecting OXPHOS, making any connection between the mitochondria and their pathogenesis less easily reconciled. In fact, the studies of cerebellar ataxia and pigmentary retinopathy discussed above indicate that disease-associated mutations in mitochondrial genes that appear to cause dysfunctional OXPHOS are often not associated with OXPHOS proteins, suggesting that an alternative explanation is required for how these mutations result in mitochondrial diseases.

The second discrepancy: mitochondrial tRNAs mutations

The mitochondrial genome harbours 22 tRNAs, which play a fundamental role in the translation of proteins. Mutations in mitochondrial tRNA genes have been linked to impairment in mitochondrial protein synthesis [13], by interfering with the binding between tRNA and aminoacyl-tRNA synthetases responsible for protein translation in the mitochondria [14]. A discrepancy between theory and practice, however, arises when considering the highly conserved function and structure of different tRNAs. Such strong conservation implies that the pathogenic potential of a mutation should be similar among tRNAs, thus a mutation associated with impaired protein synthesis at one position of the nucleotide sequence in one tRNA should be associated with similar impairment when found at the identical position of other tRNAs. Contrary to expectations, however, the position of pathogenic mutations is specific to each tRNA. Indeed, many mutations found within the mtDNA sequence are not pathogenic and widespread across populations. Mutations that are shared by many people in the same population and that can be traced back to a common ancestor are classified in groups named mtDNA haplogroups [15]. Haplogroup mutations are generally used for phylogenetic purposes (to infer evolutionary relationships between related taxonomic units)[16], and traditionally are not considered to incur negative effects (neutral *sensu lato [17]*). Noticeably, positions of neutral and pathogenic mutations are similar, highlighting the tRNA-specific nature of pathogenic mutations (**Fig.1**). An interesting example is represented by mutations found near nucleotide positions 27,36 on mt-tRNA Histidine and 31,42 respectively mt-tRNAs Leucine. These mutations are in similar positions in the tRNA structure, however, the mutations in the tRNA for the amino acid Leucine cause MELAS, while those on the tRNA for histidine do not have any known negative effects, and are indeed widespread among populations with haplogroup L1 (position 27) and L3 (position 36)[15]. Furthermore, some tRNAs (e.g. tRNA Leu1) have many more pathogenic mutations compared to other tRNAs, suggesting that tRNA conservation is not as strict for all tRNAs despite them having similar function and structure.

Indeed, when it comes to the role of tRNA mutations in mitochondrial disease, the discrepancy between theory and practice extends to the function of these RNAs. Indeed, if pathogenic mutations on the tRNAs exerted their negative effects through impairment of their ability to translate mRNAs into proteins, we would then predict that all mtDNA-encoded proteins possessing the particular amino acid translated by the mutant tRNA variant would be affected. This prediction is not upheld. For example, biochemical studies comparing mitochondrial enzyme functionality in tissue samples of patients suffering from MELAS, relative to control samples, brought about by the A3243G mutation in tRNA Leucine 1 reported that only two (I and IV) of five mitochondrial protein complexes are affected by this mutation; despite each of the complexes comprising of similar leucine content (13 to 20%)[18]. Furthermore, another study investigating the effect of mutations in tRNA Leucine found that the mutation A3243G, common in MELAS patients, downregulates translation of MT-ND4 and MT-ND5 only, leaving all the other mitochondrial proteins unaffected [5]. Interestingly, these two genes are known to vary in expression depending on the mtDNA haplotype [19]. In fact, a similar downregulation of ND4 and ND5, linked to mtDNA mutations generally, has been observed in other animals. Studies in fruit flies (*Drosophila melanogaster*) have reported mtDNA haplotype-mediated effects on the expression of ND5 and ND4, with the candidate mutations proposed to lie within the affected protein-coding genes themselves [20], suggesting that the expression of these proteins is sensitive to mechanisms that are unrelated to dysfunctional tRNAs. Taken together, these data highlight the presence of a discrepancy between theoretical and practical knowledge, where researchers have by tradition assumed that mitochondrial mutations in tRNAs will lead to impaired protein synthesis. Had this assumption been supported, then the associated impairment would be expected to be general across multiple protein complexes exhibiting high representation of the relevant amino acids. Instead, we argue that mutations in mitochondrial tRNAs, which have previously been associated with mitochondrial diseases, do not consistently lead to impaired OXPHOS, and thus the reasons for the pathogenicity of these mutations are yet to be adequately explained, and warrant consideration of an alternative hypothesis.

Population-specific effects

Mitochondrial-encoded proteins must interact with nuclear proteins to properly synthesize ATP, and any mismatching that might occur between mtDNA- and nuclear-encoded proteins could be pathogenic [21]. In fact, several mitochondrial mutations seem to be pathogenic only when in the presence of population-specific nuclear alleles [22,23]. For example, a recent study found that certain diseases, such as Type 2 diabetes and Parkinson's disease, are enriched in Caucasians carrying particular mitochondrial haplotypes [24]. The same mitochondrial mutations have no effect on Asian populations, suggesting that these mutations have to interact with population-specific nuclear alleles to be pathogenic. Likewise, another study found that a specific haplotype (N9a) confers resistance to Type 2 diabetes in people of Asian descent [25]. Indeed, there is emerging evidence that the biochemical threshold of mutant to wild type mtDNA molecule per cell, required to cause Type 2 Diabetes, differs across populations and background mtDNA haplotype. In Asians carrying the N9a haplotype, the frequency of mutated mitochondria must exceed 88% to confer a pathological phenotype, while in Europeans carrying other haplotypes (U,J,T) a frequency of 20% is sufficient to confer the same phenotype [22,26–30].

The mechanism underpinning the context-dependent pathogenicity of mtDNA mutations, across different nuclear genomic contexts, remains unknown. The current consensus hypothesis advocates that the proteins encoded by specific nuclear alleles could bind with more, or less, affinity to the interacting proteins encoded by specific mitochondrial alleles [31], with a lower affinity between the two alleles resulting in a deficiency in ATP production. Such a hypothesis hinges on the assumption that all alleles involved will affect the interactions between mitochondrial and nuclear proteins. However, most mutations do not affect protein structure (i.e., they are synonymous mutations), albeit emerging evidence suggests that synonymous mutations can nonetheless be functional, including those in the mtDNA sequence [20,32,33]. Furthermore, even mutations that affect protein structure may not necessarily cause defective OXPHOS [22]. For example, a recent study found that protein-protein binding domains are not as evolutionary conserved as previously thought [34], thus suggesting that non-synonymous mutations might not always be associated with negative phenotypic effects. Other evidence of possible communication between the mitochondria and the nucleus without the use of proteins comes from a recent meta-analysis of the mtDNA in humans [24]. In this work, the authors found that specific mitochondrial haplotypes are positively associated with longevity in Caucasians [24]. The polymorphisms that delineate these haplotypes are present in both coding and non-coding regions, with no specific mutation appearing to act as the main mediator underpinning the association with longevity. Nonetheless, mutations in the non-coding region seem to play a role, despite not being able to affect protein structure, reinforcing the idea of protein-independent mito-nuclear interactions. Considering this evidence, we argue that defective OXPHOS alone cannot explain all phenotypic effects associated with mitonuclear interactions in humans, and, in the next section, we propose an alternative framework to interpret these studies.

An alternative explanation

We propose the existence of a mechanism we call 'mitochondrial interference' to explain the discrepancies discussed above (**Fig.2**). Under this hypothesis, we contend that some occurrences of mitochondrial disease may originate from mutations within small RNAs encoded in the mitochondrial genome. In particular, we hypothesize the existence of functional small RNAs encoded in the mitochondria that are able to interfere with nuclear regulation. We predict that these small RNAs can modify nuclear protein expression through RNA interference. RNA interference is a process in which a small RNA leads a protein complex to a target mRNA, blocking its translation [35]. The small RNA can recognize the target mRNA thanks to the partial complementarity of their sequence, and mutations of either sequence, small RNA or target, could disrupt RNAi function. Disrupting the interaction between small RNAs and targets is often pathogenic [36]. Therefore, mutations in the mtDNA sequence that do not affect either protein structure or synthesis, may affect other products in the genome, including the expression and function of small non-coding RNAs. According to our hypothesis, and the data presented in the next paragraph, we propose that these mitochondrial small RNAs are encoded within other known mtDNA genes, such as tRNAs and protein-coding genes. The occurrence of these small RNA mutations, nested within known tRNA and coding genes, can explain the discrepancy of how these mutations can consistently confer mitochondrial disease without causing an associated effect on protein structure or synthesis (**Fig.3**). The mitochondrial interference hypothesis provides an alternative explanation for the observed pathogenicity of several mitochondrial mutations. Indeed, the wide range of targets that small RNAs could affect, unrelated to ATP production, could explain many of the various symptoms observed in humans. Indeed, most nuclear mRNAs have tissue-specific expression, and specific small RNAs likely affect different mRNAs depending on the tissue considered [37,38]. Therefore, the disruption of small RNA mediated regulation should be expected to cause tissue-specific effects, and be pathogenic only in specific tissues, consistent with what is consistently observed in mitochondrial diseases [1]. Importantly, the effects of mutations in these small RNAs will not only be expected to reduce the affinity with their target, but may also enhance their affinity with other targets, thus creating a wide range of possible, and complex, effects; which may explain the wide range of disease symptoms linked to mitochondrial mutations. Moreover, our hypothesis provides an alternative explanation for the observation of pathogenic mitonuclear interactions. In fact, while mutations that disrupt fundamental cellular function, such as energy production, cannot be expected to be widespread in a population due to the strong purifying selection against them [39,40], mutations affecting small RNAs can. The small mitochondrial RNAs are regulators acting on mRNAs, and thus unless the target mRNAs are necessary for fundamental cellular functions, polymorphisms in these small RNAs are unlikely to be lethal. While this alternative hypothesis provides a theoretical explanation for the discrepancies that we outlined above, it ultimately hinges on solid evidence of the existence of functional small mitochondrial RNAs. Such evidence is starting to emerge, and we discuss it in the next section.

Recent discoveries supporting the existence of mitochondrial interference

Recently, a new class of mitochondrial small non-coding RNAs of unknown function was identified in humans [41]. These small RNAs are between 20 and 30 nucleotides long and usually encoded within mt-tRNAs. These mitochondrial RNAs are similar to other well-known nuclear small RNAs such as microRNAs and piwi-interacting RNAs, having similar size, being transcribed in small clusters, and having longer RNA precursors [41–44]. Thus, we contend it is probable that they interact with the same proteins as the nuclear small RNAs [45]. Several lines of evidence already support this contention. We note that most of the proteins that are associated with regulation by piwi-interacting RNAs are localized on the mitochondrial membrane [46–48]. Furthermore, mt-tRNAs interacting with Argonaute 2, a key protein in the RNA interference mechanism, were recently identified in humans [45]. Additionally, a recent study found evidence that changes in cellular mitochondrial content have profound effects on global variability in protein expression [49], indicating a pervasive effect of the mitochondria on patterns of gene expression across the nuclear genome. In fact, the amount of small RNAs is correlated with the amount of mtDNA [42], and thus the amount of small mitochondrial RNAs present within a cell might directly underlie changes in protein expression. Moreover, several studies have now reported the widespread presence of small mitochondrial RNAs across vertebrates, including humans, and that the sequences encoding these small RNAs lie within other mitochondrial genes [42–44,50–52]. Taken together, these emerging studies indicate that mitochondrial small RNAs exist and are widespread, and that they might even provide the means to manipulating nuclear expression on behalf of the mitochondria. Moreover, these studies provide the backbone of preliminary evidence to suggest that mitochondrial interference may be a previously-overlooked genetic mediator of mitochondrial effects on human health.

Solving the discrepancies

The mitochondrial interference hypothesis can resolve the discrepancies between theory and practice noted above. The A3243G mutation, which is located within the mt-tRNA Leu1, has been consistently associated with many different symptoms, spanning from retinal neurodegeneration to stroke-like episodes [7,10]. The diversity of symptoms associated with this mutation is difficult to explain based on the current paradigm, which assumes that variation in levels of tissue-specific heteroplasmy will confer differences in the degree of impaired OXPHOS across tissues. However, this diversity could be explained if the mutation A3243G was located within a small mitochondrial RNA. Indeed, contrary to what is predicted under the current paradigm, the association of mutation A3243G to OXPHOS dysfunction is very weak, and thus heterogeneity in tissue-specific loads of A3243G should not expected to affect cell energy production. On the other hand, both mitochondrial and nuclear genomes exhibit tissue-specific expression [42], thus the interactions between the RNAs of both genomes will be predicted to lead to tissue-specific effects. Furthermore, mutations in nuclear miRNAs, or on their targets, are known to affect their function, thus we expect a similar effect involving the mitochondrial small RNAs and their mRNA targets. In this scenario, the mutation A3243G could affect different nuclear mRNA transcripts in each tissue, depending on which transcript is expressed. Thus, this prediction aligns with studies finding a correlation between the amount of mutated mtDNA heteroplasmy across multiple tissues and severity of the symptoms [1]. Indeed, a study found that the amount of mtDNA is positively correlated with the expression of small mitochondrial RNAs [42], thus if the load of non-mutated (healthy) mtDNA decreases, the amount of functional small mitochondrial RNAs will decrease, leaving only non-functional small mitochondrial RNAs encoded by the mtDNA carrying the mutation A3243G. This hypothetical scenario is supported by one, previously ignored, published finding. The first study reporting the presence of small mitochondrial RNA encoded in the mitochondria found that mt-tRNA Leu1 encodes a small mitochondrial RNA in the region where the mutation A3243G is located [41]. This small mitochondrial RNA is differentially expressed across multiple cell lines, supporting the presence of tissue-specific interactions[41]. Furthermore, the authors found no correlation between the expression of mt-tRNAs and the expression of small mitochondrial RNAs, suggesting that these small mitochondrial RNAs were not simply bioproducts of RNA turnover from the mt-tRNAs.

Therefore, we know that patients harbouring the mutation A3243G are expressing a mutated small mitochondrial RNA with unknown function, that might explain the pathogenicity of the mutation A3243G.

The mitochondrial interference hypothesis can resolve the discrepancies found while investigating the effect of mtDNA mutations across populations, and a good example would be the previously-mentioned study on Type 2 Diabetes and Parkinson's disease [24]. In that study, the same mitochondrial haplotypes conferred differences in effects between Caucasian and Asian populations. Under our hypothesis, we would predict that the Caucasian population is the only population possessing the nuclear allelic target of the pathogenic small mitochondrial RNA. Therefore, if this were the case, then such a nuclear allele would not be pathogenic by itself; that is, because the hypothetical small RNA binds a non-coding region of the mRNA, the effect of this mutation would appear only if paired to the mitochondria harbouring the pathogenic small RNA.

Concluding Remarks

Demonstrating both the existence and pervasiveness of mitochondrial interference is complex given that any such interference would involve interactions between two genomes – mitochondrial and nuclear. However, our hypothesis provides three key predictions that can be tested (**Fig.4**). Our first prediction is that the small mitochondrial RNAs can interfere with mRNA translation. This prediction can be tested by leveraging bioinformatic analyses to identify the targets of these small mitochondrial RNAs, then injecting small mitochondrial RNAs whose sequences specifically target candidate mRNAs of interest into cell cultures, and, finally, by performing a western blot to verify the downregulation of the translated proteins. However, this experiment carries several risks. Reliable identification of the mRNAs targeted by the small mitochondrial RNAs targets might be difficult; as previously seen with miRNAs, modern methods have strong confirmation bias and do not perform well with novel RNAs [53]. Nonetheless, this risk might be compensated by using high-throughput approaches, such as sequencing the entire proteome of multiple injected cell cultures, thus avoiding the necessity of a target identification [54]. By testing this prediction experimentally, it would be possible to verify the action of the small mitochondrial RNAs as inhibitors of mRNA translation. Our second prediction is that the small mitochondrial RNAs can bind to key proteins of RNAi. This prediction can be tested through RNA immunoprecipitation sequencing (RIP-seq) experiments, with nuclear proteins key to RNAi – such as Ago2, or Piwi. RIP-seq technique involves an RNA and RNA-binding protein co-immunoprecipitation, followed by RNA sequencing, thereby revealing all the RNAs able to bind the RNA-binding protein [55–57]. Testing the small mitochondrial RNAs ability to bind to Ago2, or Piwi, would have power to resolve the question of whether small mitochondrial RNAs may regulate patterns of nuclear gene and protein expression through RNA interference. The third prediction is that the function of cells that harbour pathogenesis-inducing mtDNA mutations (located within candidate small mitochondrial RNAs) can be restored through injection of the small mitochondrial RNA counterparts that lack the mutation. While the implementation of these approaches is not without its challenges, they offer promising avenues to experimentally test the key predictions of the mitochondrial interference hypothesis and therefore deserve immediate research attention.

In sum, the mitochondrial interference hypothesis proposes a new framework for the investigation of mitochondrial diseases. Current studies of the genetics underpinning the expression of mitochondrial diseases focus mostly on effects on mitochondrial energy production. Instead, we propose a role for the mitochondria beyond energy production, which involves regulation of the nuclear genome through mitochondrial small RNAs. We contend that recently discovered small mitochondrial RNAs may interfere with nuclear regulation and may therefore confer mitochondrial disease in a hitherto undescribed biological context; a context independent of OXPHOS functionality. Here, we have discussed small mitochondrial RNAs and their potential biomedical implications, highlighting their similarity with nuclear RNAs involved in the RNA interference mechanism. However, while focusing our attention on the potential role of these small RNAs in mitochondrial diseases, the potential implications extend beyond mitochondrial diseases *per se*, influencing common diseases such as Type 2 Diabetes. In conclusion, we hope that the mitochondrial interference hypothesis will inspire clinicians, medical scientists, and the scientific community in general to pursue new pathways in the study of mitochondrial biology in general, and mitochondrial diseases in particular.

References

1. Gorman GS, Chinnery PF, DiMauro S, Hirano M, Koga Y, McFarland R, et al. Mitochondrial diseases. Nat Rev Dis Primers. 2016;2: 16080. doi:10.1038/nrdp.2016.80

2. DiMauro S, Moraes CT. Mitochondrial encephalomyopathies. Arch Neurol. 1993;50: 1197–1208. Available: https://www.ncbi.nlm.nih.gov/pubmed/8215979

3. Pieczenik SR, Neustadt J. Mitochondrial dysfunction and molecular pathways of disease. Exp Mol Pathol. 2007;83: 84–92. doi:10.1016/j.yexmp.2006.09.008

4. Bratic I, Trifunovic A. Mitochondrial energy metabolism and ageing. Biochim Biophys Acta. 2010;1797: 961–967. doi:10.1016/j.bbabio.2010.01.004

5. Sasarman F, Antonicka H, Shoubridge EA. The A3243G tRNALeu(UUR) MELAS mutation causes amino acid misincorporation and a combined respiratory chain assembly defect partially suppressed by overexpression of EFTu and EFG2. Hum Mol Genet. 2008;17: 3697–3707. doi:10.1093/hmg/ddn265

6. Geffroy G, Benyahia R, Frey S, Desquiret-Dumas V, Gueguen N, Bris C, et al. The accumulation of assembly intermediates of the mitochondrial complex I matrix arm is reduced by limiting glucose uptake in a neuronal-like model of MELAS syndrome. Biochim Biophys Acta Mol Basis Dis. 2018;1864: 1596–1608. doi:10.1016/j.bbadis.2018.02.005

7. Lax NZ, Hepplewhite PD, Reeve AK, Nesbitt V, McFarland R, Jaros E, et al. Cerebellar ataxia in patients with mitochondrial DNA disease: a molecular clinicopathological study. J Neuropathol Exp Neurol. 2012;71: 148–161. doi:10.1097/NEN.0b013e318244477d

8. McFarland R, Schaefer AM, Gardner JL, Lynn S, Hayes CM, Barron MJ, et al. Familial myopathy: new insights into the T14709C mitochondrial tRNA mutation. Ann Neurol. 2004;55: 478–484. doi:10.1002/ana.20004

9. Chinnery PF. Mitochondrial Disorders Overview. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, et al., editors. GeneReviews®. Seattle (WA): University of Washington, Seattle; 2000. Available: https://www.ncbi.nlm.nih.gov/pubmed/20301403

10. Chichagova V, Hallam D, Collin J, Buskin A, Saretzki G, Armstrong L, et al. Human iPSC disease modelling reveals functional and structural defects in retinal pigment epithelial cells harbouring the m.3243A > G mitochondrial DNA mutation. Sci Rep. 2017;7: 12320. doi:10.1038/s41598-017-12396-2

11. Majamaa K, Moilanen JS, Uimonen S, Remes AM, Salmela PI, Kärppä M, et al.

Epidemiology of A3243G, the mutation for mitochondrial encephalomyopathy, lactic acidosis, and strokelike episodes: prevalence of the mutation in an adult population. Am J Hum Genet. 1998;63: 447–454. doi:10.1086/301959

12. Veltri KL, Espiritu M, Singh G. Distinct genomic copy number in mitochondria of different mammalian organs. J Cell Physiol. 1990;143: 160–164. doi:10.1002/jcp.1041430122

13. Yarham JW, Al-Dosary M, Blakely EL, Alston CL, Taylor RW, Elson JL, et al. A comparative analysis approach to determining the pathogenicity of mitochondrial tRNA mutations. Hum Mutat. 2011;32: 1319–1325. doi:10.1002/humu.21575

14. Sissler M, González-Serrano LE, Westhof E. Recent Advances in Mitochondrial Aminoacyl-tRNA Synthetases and Disease. Trends Mol Med. 2017;23: 693–708. doi:10.1016/j.molmed.2017.06.002

15. van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat. 2009;30: E386–94. doi:10.1002/humu.20921

16. Rubinoff D, Holland BS. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst Biol. 2005;54: 952–961. doi:10.1080/10635150500234674

17. Ballard JW, Kreitman M. Is mitochondrial DNA a strictly neutral marker? Trends Ecol Evol. 1995;10: 485–488. Available: https://www.ncbi.nlm.nih.gov/pubmed/21237113

18. Lin D-S, Kao S-H, Ho C-S, Wei Y-H, Hung P-L, Hsu M-H, et al. Inflexibility of AMPK-mediated metabolic reprogramming in mitochondrial disease. Oncotarget. 2017;8: 73627–73639. doi:10.18632/oncotarget.20617

19. Cohen T, Levin L, Mishmar D. Ancient Out-of-Africa Mitochondrial DNA Variants Associate with Distinct Mitochondrial Gene Expression Patterns. PLoS Genet. 2016;12: e1006407. doi:10.1371/journal.pgen.1006407

20. Camus MF, Wolff JN, Sgrò CM, Dowling DK. Experimental Support That Natural Selection Has Shaped the Latitudinal Distribution of Mitochondrial Haplotypes in Australian Drosophila melanogaster. Mol Biol Evol. 2017;34: 2600–2612. doi:10.1093/molbev/msx184

21. Wolff JN, Ladoukakis ED, Enríquez JA, Dowling DK. Mitonuclear interactions: evolutionary consequences over multiple biological scales. Philos Trans R Soc Lond B Biol Sci. 2014;369: 20130443. doi:10.1098/rstb.2013.0443

22. Chalkia D, Chang Y-C, Derbeneva O, Lvova M, Wang P, Mishmar D, et al. Mitochondrial DNA associations with East Asian metabolic syndrome. Biochim Biophys Acta. 2018;1859: 878–892. doi:10.1016/j.bbabio.2018.07.002

23. Hudson G, Gomez-Duran A, Wilson IJ, Chinnery PF. Recent Mitochondrial DNA Mutations Increase the Risk of Developing Common Late-Onset Human Diseases. PLoS Genet. 2014;10: e1004369. doi:10.1371/journal.pgen.1004369

24. Marom S, Friger M, Mishmar D. MtDNA meta-analysis reveals both phenotype specificity and allele heterogeneity: a model for differential association. Sci Rep. 2017;7: 43449.

doi:10.1038/srep43449

25. Fuku N, Park KS, Yamada Y, Nishigaki Y, Cho YM, Matsuo H, et al. Mitochondrial haplogroup N9a confers resistance against type 2 diabetes in Asians. Am J Hum Genet. 2007;80: 407–415. doi:10.1086/512202

26. Martikainen MH, Rönnemaa T, Majamaa K. Prevalence of mitochondrial diabetes in southwestern Finland: a molecular epidemiological study. Acta Diabetol. 2013;50: 737–741. doi:10.1007/s00592-012-0393-2

27. Martikainen MH, Rönnemaa T, Majamaa K. Association of mitochondrial DNA haplogroups and vascular complications of diabetes mellitus: A population-based study. Diab Vasc Dis Res. 2015;12: 302–304. doi:10.1177/1479164115579007

28. Crispim D, Canani LH, Gross JL, Tschiedel B, Souto KEP, Roisenberg I. The European-Specific Mitochondrial Cluster J/T Could Confer an Increased Risk of Insulin-Resistance and Type 2 Diabetes: An Analysis of the m.4216T > C and m.4917A > G Variants. Annals of Human Genetics. 2006. pp. 488–495. doi:10.1111/j.1469-1809.2005.00249.x

29. Feder J, Ovadia O, Blech I, Cohen J, Wainstein J, Harman-Boehm I, et al. Parental diabetes status reveals association of mitochondrial DNA haplogroup J1 with type 2 diabetes. BMC Med Genet. 2009;10: 60. doi:10.1186/1471-2350-10-60

30. Mohlke KL, Jackson AU, Scott LJ, Peck EC, Suh YD, Chines PS, et al. Mitochondrial polymorphisms and susceptibility to type 2 diabetes-related traits in Finns. Hum Genet. 2005;118: 245–254. doi:10.1007/s00439-005-0046-4

31. Barreto FS, Burton RS. Evidence for compensatory evolution of ribosomal proteins in response to rapid divergence of mitochondrial rRNA. Mol Biol Evol. 2013;30: 310–314. doi:10.1093/molbev/mss228

32. Hurst LD. Molecular genetics: The sound of silence. Nature. 2011;471: 582–583. doi:10.1038/471582a

33. Kimchi-Sarfaty C, Oh JM, Kim I-W, Sauna ZE, Calcagno AM, Ambudkar SV, et al. A “silent” polymorphism in the MDR1 gene changes substrate specificity. Science. 2007;315: 525–528. doi:10.1126/science.1135308

34. Jack BR, Meyer AG, Echave J, Wilke CO. Functional Sites Induce Long-Range Evolutionary Constraints in Enzymes. PLoS Biol. 2016;14: e1002452. doi:10.1371/journal.pbio.1002452

35. Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol. 2014;15: 509–524. doi:10.1038/nrm3838

36. Ardekani AM, Naeini MM. The Role of MicroRNAs in Human Diseases. Avicenna J Med Biotechnol. 2010;2: 161–179. Available: https://www.ncbi.nlm.nih.gov/pubmed/23407304

37. Ovando-Vázquez C, Lepe-Soltero D, Abreu-Goodger C. Improving microRNA target prediction with gene expression profiles. BMC Genomics. 2016;17: 364. doi:10.1186/s12864-016-2695-1

38. Alberti C, Manzenreither RA, Sowemimo I, Burkard TR, Wang J, Mahofsky K, et al. Cell-type specific sequencing of microRNAs from complex animal tissues. Nat Methods. 2018;15: 283–289. doi:10.1038/nmeth.4610

39. Stewart JB, Freyer C, Elson JL, Larsson N-G. Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease. Nat Rev Genet. 2008;9: 657–662. doi:10.1038/nrg2396

40. Cooper BS, Burrus CR, Ji C, Hahn MW, Montooth KL. Similar Efficacies of Selection Shape Mitochondrial and Nuclear Genes in Both Drosophila melanogaster and Homo sapiens. G3 . 2015;5: 2165–2176. doi:10.1534/g3.114.016493

41. Mercer TR, Neph S, Dinger ME, Crawford J, Smith MA, Shearwood A-MJ, et al. The human mitochondrial transcriptome. Cell. 2011;146: 645–658. doi:10.1016/j.cell.2011.06.051

42. Pozzi A, Dowling DK. The genomic origins of small mitochondrial RNAs: are they transcribed by the mitochondrial DNA or by mitochondrial pseudogenes within the nucleus (NUMTs)? Genome Biol Evol. 2019. doi:10.1093/gbe/evz132

43. Pozzi A, Plazzi F, Milani L, Ghiselli F, Passamonti M. SmithRNAs: could mitochondria “bend” nuclear regulation? Mol Biol Evol. 2017. doi:10.1093/molbev/msx140

44. Ro S, Ma H-Y, Park C, Ortogero N, Song R, Hennig GW, et al. The mitochondrial genome encodes abundant small noncoding RNAs. Cell Res. 2013;23: 759–774. doi:10.1038/cr.2013.37

45. Maniataki E, Mourelatos Z. Human mitochondrial tRNAMet is exported to the cytoplasm and associates with the Argonaute 2 protein. RNA. 2005;11: 849–852. doi:10.1261/rna.2210805

46. Shiromoto Y, Kuramochi-Miyagawa S, Daiba A, Chuma S, Katanaya A, Katsumata A, et al. GPAT2, a mitochondrial outer membrane protein, in piRNA biogenesis in germline stem cells. RNA. 2013;19: 803–810. doi:10.1261/rna.038521.113

47. Ding D, Liu J, Dong K, Melnick AF, Latham KE, Chen C. Mitochondrial membrane-based initial separation of MIWI and MILI functions during pachytene piRNA biogenesis. Nucleic Acids Res. 2019;47: 2594–2608. doi:10.1093/nar/gky1281

48. Nishimasu H, Ishizu H, Saito K, Fukuhara S, Kamatani MK, Bonnefond L, et al. Structure and function of Zucchini endoribonuclease in piRNA biogenesis. Nature. 2012;491: 284–287. doi:10.1038/nature11509

49. Guantes R, Rastrojo A, Neves R, Lima A, Aguado B, Iborra FJ. Global variability in gene expression and alternative splicing is modulated by mitochondrial content. Genome Res. 2015;25: 633–644. doi:10.1101/gr.178426.114

50. Larriba E, Rial E, Del Mazo J. The landscape of mitochondrial small non-coding RNAs in the PGCs of male mice, spermatogonia, gametes and in zygotes. BMC Genomics. 2018;19: 634. doi:10.1186/s12864-018-5020-3

51. Riggs CL, Summers A, Warren DE, Nilsson GE, Lefevre S, Dowd WW, et al. Small Non-coding RNA Expression and Vertebrate Anoxia Tolerance. Front Genet. 2018;9: 230. doi:10.3389/fgene.2018.00230

52. Bottje WG, Khatri B, Shouse SA, Seo D, Mallmann B, Orlowski SK, et al. Identification and Differential Abundance of Mitochondrial Genome Encoding Small RNAs (mitosRNA) in Breast Muscles of Modern Broilers and Unselected Chicken Breed. Front Physiol. 2017;8: 816. doi:10.3389/fphys.2017.00816

53. Cloonan N. Re-thinking miRNA-mRNA interactions: intertwining issues confound target discovery. Bioessays. 2015;37: 379–388. doi:10.1002/bies.201400191

54. Aslam B, Basit M, Nisar MA, Khurshid M, Rasool MH. Proteomics: Technologies and Their Applications. J Chromatogr Sci. 2017;55: 182–196. doi:10.1093/chromsci/bmw167

55. Lu Z, Guan X, Schmidt CA, Matera AG. RIP-seq analysis of eukaryotic Sm proteins identifies three major categories of Sm-containing ribonucleoproteins. Genome Biol. 2014;15: R7. doi:10.1186/gb-2014-15-1-r7

56. Clark PM, Loher P, Quann K, Brody J, Londin ER, Rigoutsos I. Argonaute CLIP-Seq reveals miRNA targetome diversity across tissue types. Sci Rep. 2014;4: 5947. doi:10.1038/srep05947

57. Zhang X, Zuo X, Yang B, Li Z, Xue Y, Zhou Y, et al. MicroRNA directly enhances mitochondrial translation during muscle differentiation. Cell. 2014;158: 607–619. doi:10.1016/j.cell.2014.05.047

Figures

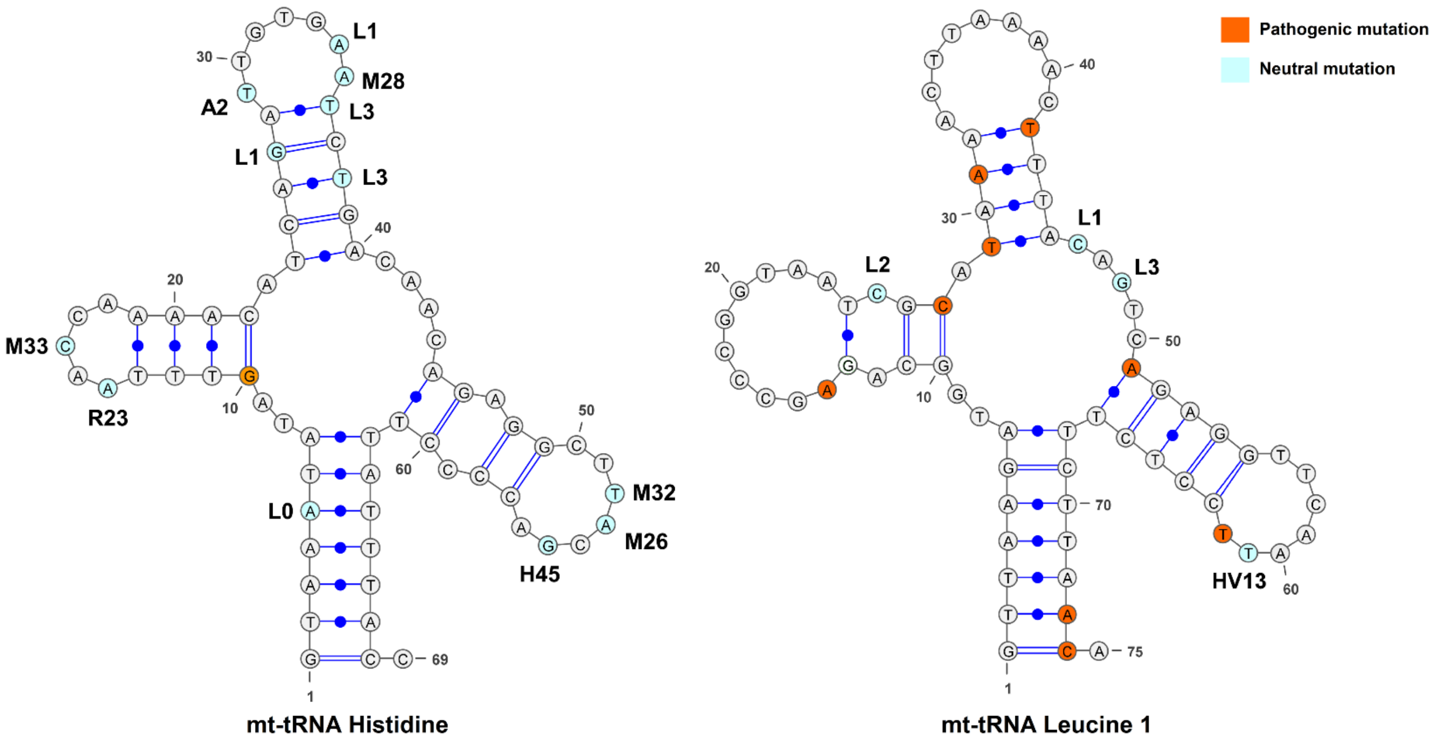


Fig.1 Variability in the position of mtDNA mutations across mt-tRNAs. Two human mt-tRNAs are shown (mt-tRNA Histidine on left, and Leucine 1 on right), in which nucleotides with different colours represent neutral (light blue) and pathogenic (orange) mutations. Neutral mutations are mutations without any known negative effects on fitness. The labels next to the neutral mutations indicate the haplogroup in which that mutation is common. These haplogroups have different frequencies depending on the population considered. For example, the haplogroup L is the most common in African populations.

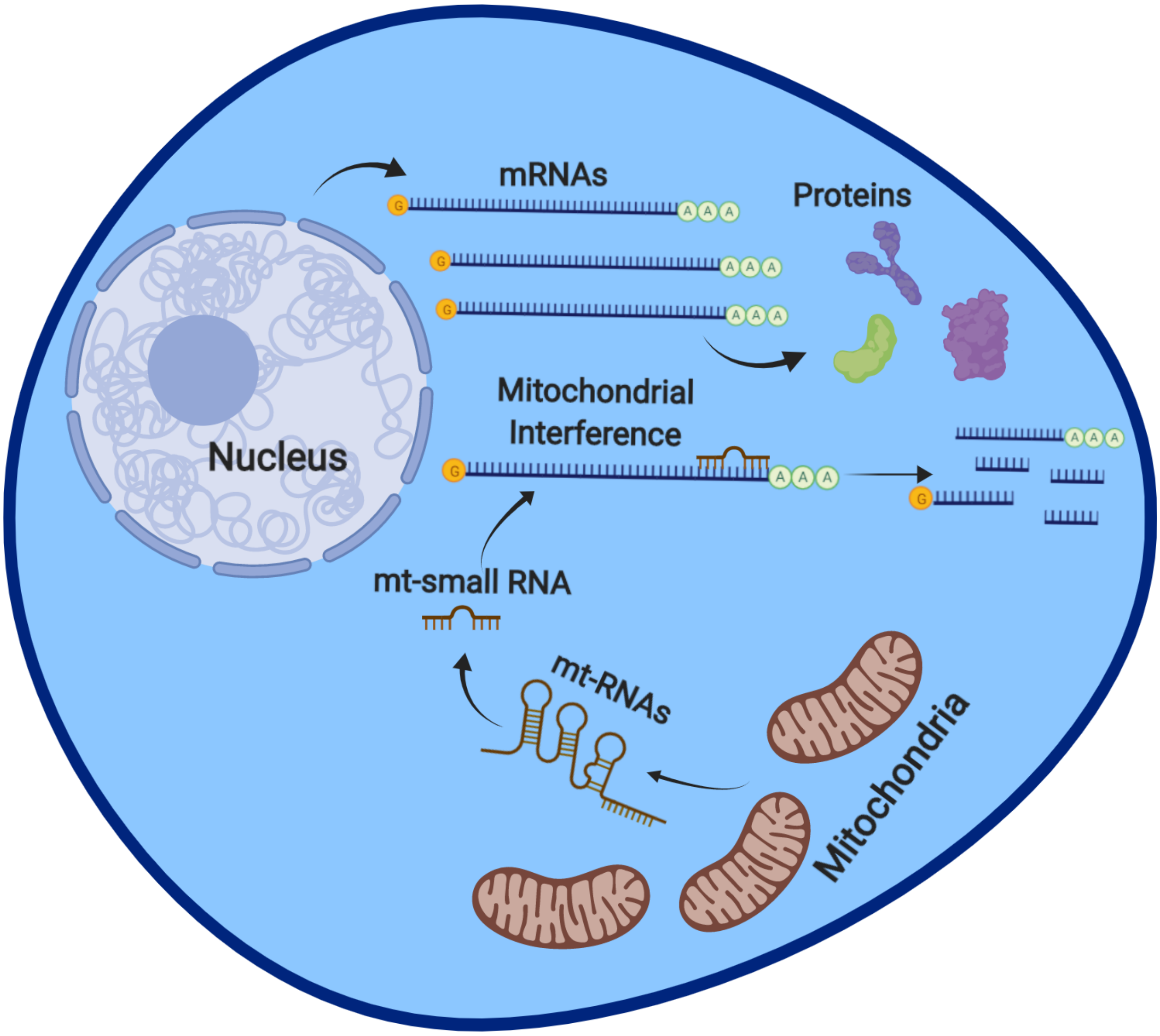


Fig.2. A schematic representation of the mitochondrial interference hypothesis. The mitochondrial genome transcribes small mitochondrial RNAs (mt-small RNA) that can bind a complementary mRNA made by the nucleus. Then the mt-small RNA binds the target mRNA in a region with similar sequence. This binding stops the translation of the nuclear mRNA, thus changing the expression of a nuclear gene. This process could in theory target virtually any nuclear gene, and thus could result in dysfunction of innumerable phenotypes.

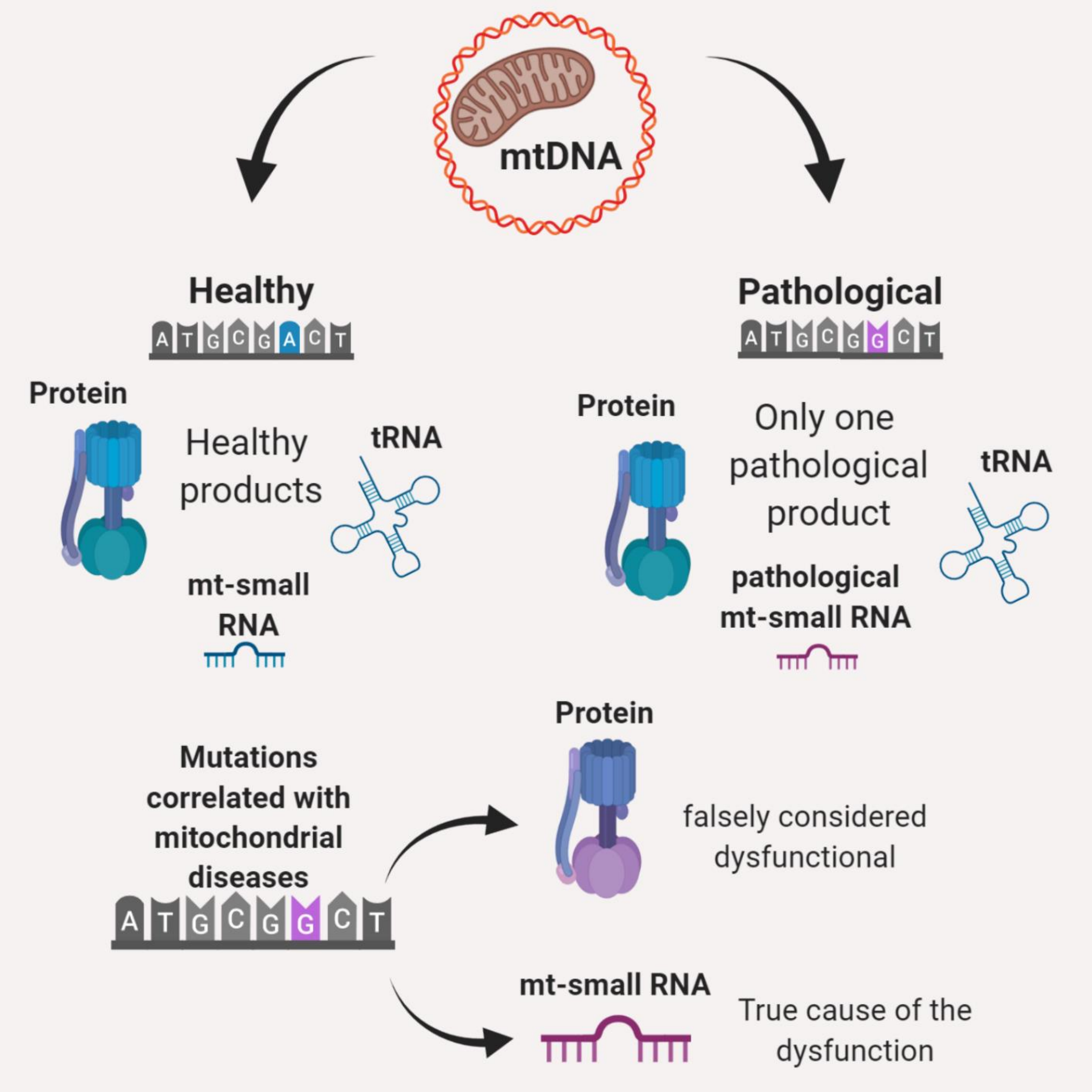


Fig.3 A representation of the misclassification of pathological mitochondrial mutations. Pathogenic mutations found in genes encoding proteins and tRNAs will be usually linked to dysfunction in these genes. However, the mitochondrial interference hypothesis suggests that such mutations might instead affect the transcription of mitochondrial small RNAs (mt-small RNAs) encoded within the protein coding sequence or the tRNAs. If so, these mutations affecting mt-small RNAs will have been previously overlooked and misclassified as mutations affecting either protein or tRNA function.

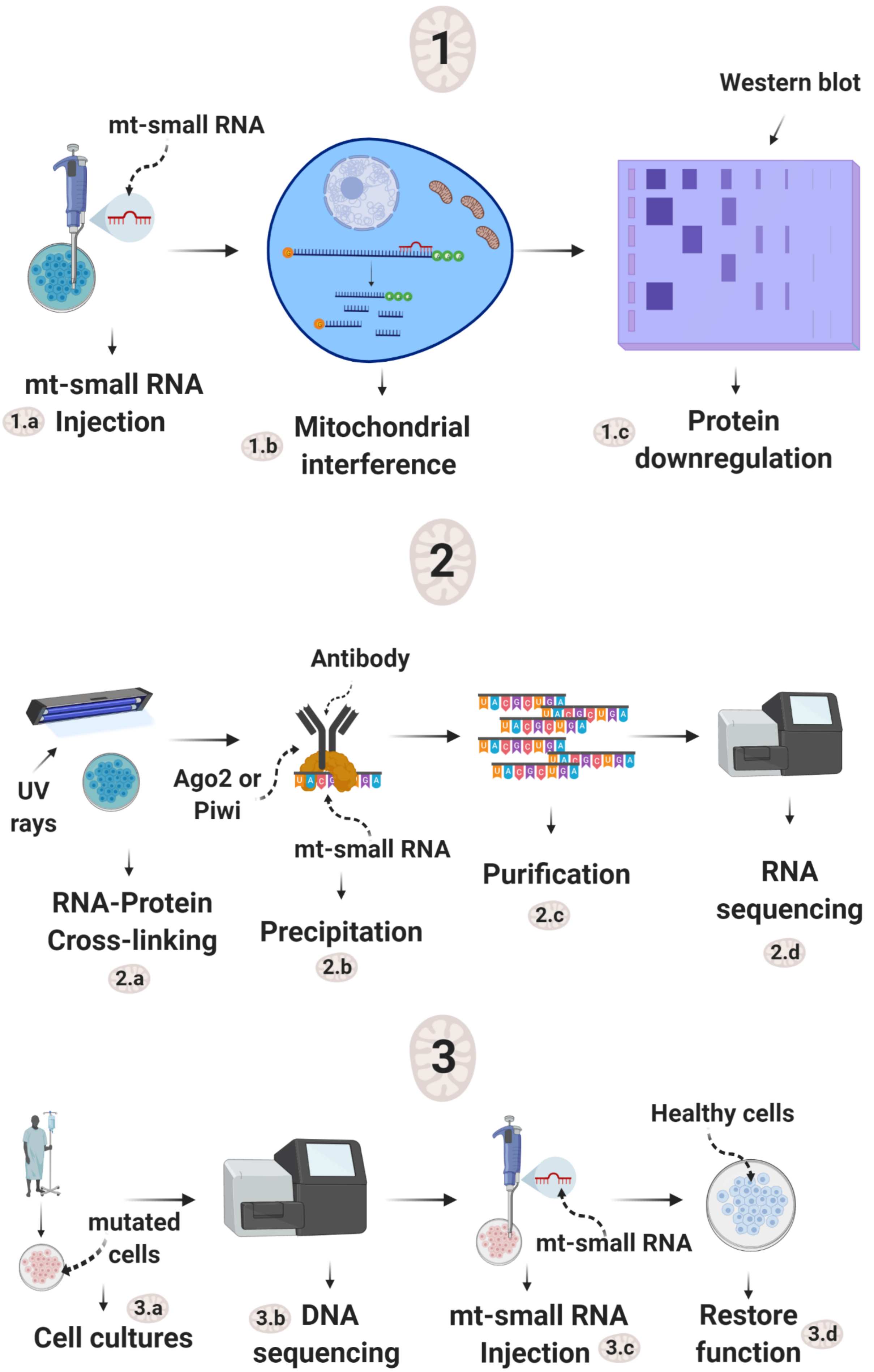

1
mt-small RNA
Western blot
mt-small RNA
1.a Injection
1.b Mitochondrial interference
1.c Protein downregulation
2
Antibody
UV rays
Ago2 or Piwi
mt-small RNA
RNA-Protein Cross-linking
2.a
Precipitation
2.b
Purification
2.c
RNA sequencing
2.d
3
Healthy cells
mutated cells
mt-small RNA
3.a
Cell cultures
3.b DNA sequencing
mt-small RNA Injection 3.c
Restore function 3.d

Fig.4 A schematic representation of the experiments necessary to verify the three key predictions of the mitochondrial interference hypothesis. Prediction 1 is that the small mitochondrial RNAs can interfere with mRNA translation. It can be tested by measuring expression of target proteins after injection of candidate mt-small RNAs, proving that their action is to inhibit protein expression. This involves injecting small mitochondrial RNAs (1.a) targeting specific mRNAs, which will cause mitochondrial interference (1.b), with protein downregulation then verified through western blot analysis (1.c). Prediction 2 is that the small mitochondrial RNAs can bind to key proteins of RNAi. It can be tested by determining whether small mitochondrial RNAs are bound to Argonaute 2 *in vivo*. It involves cross-linking Ago2 with the small mitochondrial RNAs (2.a), precipitated them using antibodies (2.b), purify the RNAs by removing the proteins (2.c), and, ultimately, sequencing the small mitochondrial RNAs (2.d). Prediction 3 is that the function of cells that harbour pathogenesis-inducing mtDNA mutations can be restored through injection of the small mitochondrial RNA counterparts that lack the mutation. This experiment involves two stages. In the first, the pathogenic mutations in the patient are identified though cell cultures (3.a) and sequencing (3.a). In the second, unmutated small mitochondrial RNAs corresponding to the mutated mitochondrial regions will be injected in the cell cultures, then the function of the healthy cells will be verified through histological and physiological analyses depending on the pathology investigated.

CHAPTER FOUR

A new member in the Argonaute crew: the mt-miRNAs

Title: A new member in the Argonaute crew: the mt-miRNAs

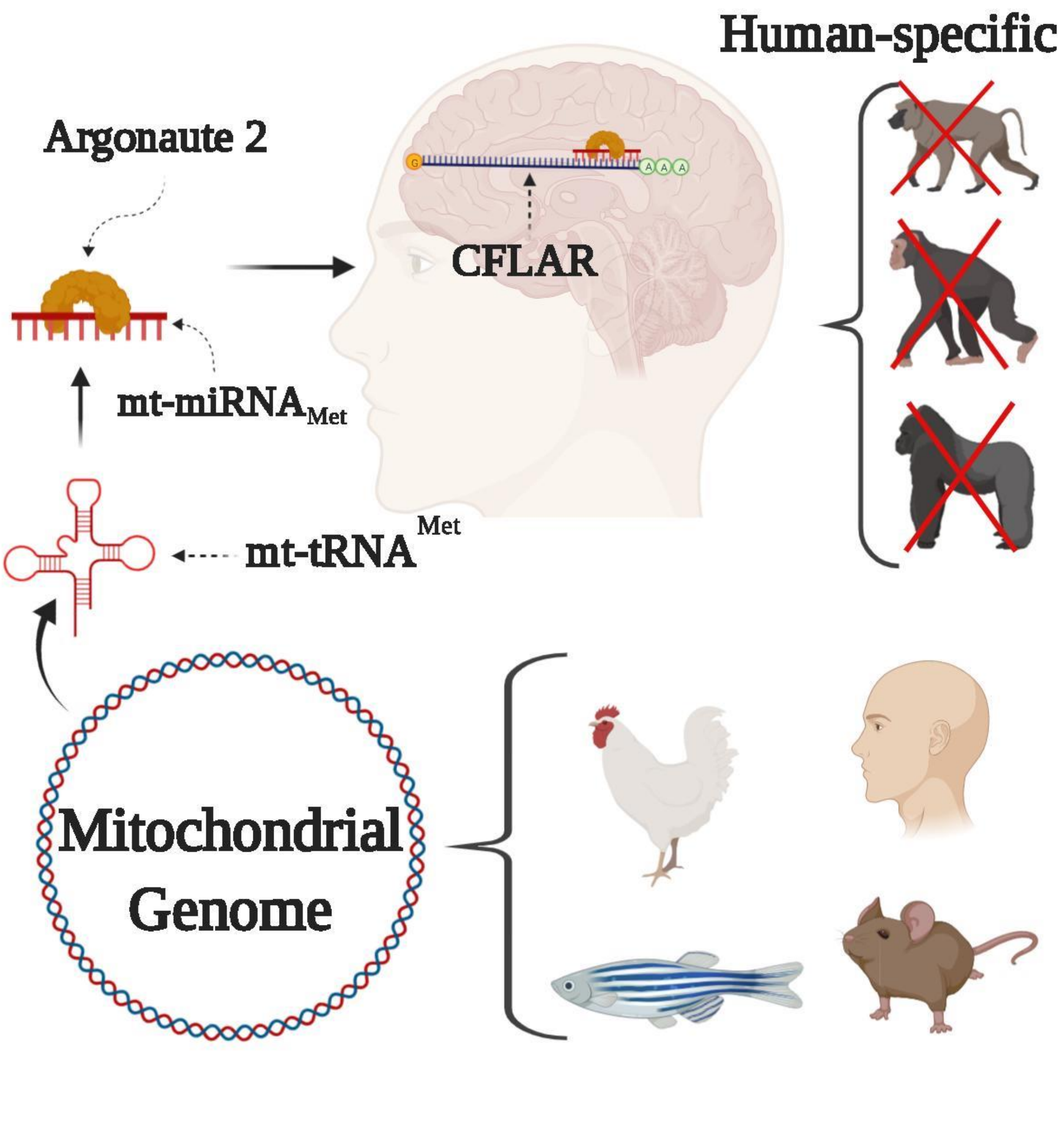
Author 1: Andrea Pozzi, Monash University, School of Biological Sciences, andrea.pozzi@monash.edu

Author 2: Damian K. Dowling, Monash University, School of Biological Sciences, damian.dowling@monash.edu

Status: In Review – *Current Biology* / Available on *BioArchivx*

In Brief & Graphical Abstract

We show that the mitochondrial genome encodes miRNAs able to bind Ago2. In particular, the mt-miRNA encoded within the gene mt-tRNA Met is conserved across multiple vertebrates and has a unique target in humans, CFLAR. The regulatory region of CFLAR is present only in humans, suggesting that the mt-miRNA Met might be part of a human specific regulatory pathway.



Highlights

- The mt-miRNAs can bind Ago2 in multiple species
- The biogenesis of mt-miRNAs is different from that of nuclear miRNAs
- The mt-miRNA encoded in mt-$tRNA_{Met}$ has been conserved for over 300 Mya
- The mt-$miRNA_{Met}$ targets CFLAR, a transcript involved in apoptosis
- The target region in CFLAR is present only in humans

Summary

Mutations within the mitochondrial genome have been linked to many diverse phenotypes. The mechanisms that explain the manifold array of mitochondrial genotypic effects on organismal function, and their context-dependency, have however remained a mystery. Here, we demonstrate that mitochondria are involved in nuclear gene regulation via RNA interference; transcribing mitochondrial (mt-)miRNAs able to bind nuclear mRNAs that previously had no known involvement in mitochondrial function. We identify one mt-miRNA that exhibits particularly high conservation in sequence and expression across vertebrates. We then screened for candidate targets of this mt-miRNA involved in insulin regulation, finding that this mt-miRNA represses CLFAR, an enhancer of apoptosis in the human brain. Through comparative analysis of CFLAR across primates, we show that the UTR bound by this mt-miRNA is present only in a human-specific CFLAR isoform. Our findings thus uncover a novel mechanism via which the mitochondria can regulate patterns of nuclear gene expression.

Introduction

Interest in mitochondrial biology is on the rise, with a growing number of studies highlighting the complex role of the mitochondria in cell regulation (Picard, Wallace, and Burelle 2016; Sloan et al. 2018; Sprenger and Langer 2019). Among these, numerous studies have found that sequence variation in the mitochondrial DNA (mtDNA) can affect the expression of a range of life-history and health related traits, from fertility, to longevity and thermal tolerance (Lajbner et al. 2018; Camus et al. 2017; Rand, Fry, and Sheldahl 2006; Song and Lewis 2008; Yee, Sutton, and Dowling 2013; James and Ballard 2003). Furthermore, while it is well known that the mtDNA can harbour loss-of-function mutations conferring mitochondrial disease in humans (Wallace 2018), emerging studies implicate mtDNA mutations in a range of other late-onset diseases not previously linked to mitochondrial function (Hudson et al. 2014). For example, Hopkins et al. (2017) recently reported an association between the frequency and type of mtDNA mutations and aggressiveness of prostate cancer. Furthermore, the phenotypic effects of these mtDNA mutations appear to be routinely moderated by the nuclear genetic background alongside which the mtDNA mutations are co-expressed (Hill et al. 2018), suggesting a broad role for intergenomic regulation ("mitonuclear communication") involving exchange of proteins, metabolites and genetic products between genomes (Wu et al. 2019; Moriyama, Koshiba, and Ichinohe 2019; Zhu, Ingelmo, and Rand 2014). The mechanistic basis of the molecular interactions that underpin mitonuclear regulation of cellular and organismal function, however, remains elusive.

In 2019, Kopinski et al. provided new insights into how the mitochondria communicate with the nucleus, reporting a key role for mitochondrial metabolites and subcellular redox levels. The authors found that variation in the level of intracellular mtDNA heteroplasmy (i.e. the frequency of normal to mutant mtDNA molecules) modulates mitochondrial metabolites, influencing the abundance of substrate necessary for methylation and acetylation of specific histones, thus affecting patterns of nuclear expression (Kopinski et al. 2019). Their findings were to some degree consistent with those of a previous study by Guantes et al. (2015), who reported strong correlations between mtDNA content and changes to the epigenetic and transcriptional profile of the cell. However, Guantes et al. (2015) found that some of the mtDNA-content mediated changes to cellular regulation were unlikely related to the presence of mitochondrial metabolites. Furthermore, they found that mtDNA content also shapes patterns of post-transcriptional regulation; a type of regulation that is not usually affected directly by histone methylation. Mitochondrial-mediated modification to levels of protein translation, affected by post-transcriptional modifications to gene expression, would require a mechanism that is able to interfere with mRNA translation. To date, such a mechanism is unknown to exist.

Mitochondrial RNAs (mt-RNAs) represent a possible mediator of patterns of mitochondrial-mediated post-transcriptional regulation. Functional mt-RNAs are well-known in the form of the 22 tRNAs and 2 rRNAs encoded by the typical vertebrate mitochondrial genome. However, novel types of RNAs of mitochondrial origin have recently been identified. For example, Dhir et al. (2018) described a new class of double-stranded RNAs encoded in the mitochondria that are able to trigger antiviral signaling in humans (Dhir et al. 2018). Although to date these double-stranded RNAs have only been identified in humans, different types of novel small mitochondrial RNAs have been described in multiple species across two metazoan phyla, Chordata and Mollusca (Ro et al. 2013; Riggs et al. 2018; Larriba, Rial, and Del Mazo 2018; Bottje et al. 2017; Mercer et al. 2011; Pozzi et al. 2017; Pozzi and Dowling 2019). Yet, despite increasing interest in the putative role these small mitochondrial RNAs may play in the regulation of cellular function, clear evidence of their functionality remains absent.

Given similarities in their length and sequence to microRNAs, Pozzi et al. (2017) hypothesized involvement of small mitochondrial RNAs in the regulation of mRNA translation through RNA interference (RNAi) (Pozzi et al. 2017). RNAi is a process in which a microRNA (miRNA) leads a protein complex to block translation of a target mRNA (Ambros 2004; Ha and Kim 2014). miRNAs are partially complementary to a regulatory region of the mRNAs, and due this close miRNA-mRNA affinity, the protein complex is able to precisely bind its target mRNA and hinder its binding to the ribosome (Ambros 2004; Cloonan 2015). Within this protein complex, the main protein binding the miRNAs is Ago2, an endonuclease shared across multiple species necessary for RNAi (Ha and Kim 2014; Cloonan 2015). Interestingly, Ago2 has previously been reported to co-localize with mitochondria (Bandiera et al. 2011; Zhang et al. 2014) and, moreover, to associate with mitochondrial tRNA (mt-tRNA$_{Met}$) in the cytoplasm (Maniataki and Mourelatos 2005). Accordingly, Ago2 is an excellent candidate to further probe the hypothesis that the small mitochondrial RNAs serve a similar role as nuclear encoded miRNAs in RNAi, and therefore constitute mitochondrial miRNAs (hereafter, “mt-miRNAs”).

Here, we investigate mitochondrial involvement in RNAi by verifying the presence of multiple features of miRNAs in the mitochondrial small RNAs. Firstly, we sought to verify binding between the mitochondrial small RNAs and Ago2. To this end, we leveraged published datasets of RNA sequencing (RNA-seq) and RNA-binding-protein co-immunoprecipitation sequencing (RIP-seq) (Townley-Tilson, Pendergrass, and Marzluff 2006). These datasets come from pre-published studies that reported novel mechanistic insights into RNAi. We re-purposed these datasets to investigate the capacity for mtDNA-mediated involvement in RNAi. Secondly, we tested whether the mitochondrial small RNAs are generated by pre-mitochondrial small RNAs, similarly to the case of miRNAs. To achieve this, we screened for the presence of matching transcriptional profiles between datasets of long and small RNAs extracted from the same individuals, which enabled us to determine if the mt-miRNAs are transcribed from transcripts of ~70nt length, as commonly happens in the miRNAs (Ha and Kim 2014). Thirdly, given that miRNAs have been shown to be conserved across multiple clades, we explored levels of sequence conservation in the mitochondrial small RNAs. We compared the presence of mitochondrial small RNAs that exhibit features like the miRNAs across multiple model organisms by using small RNAs-seq datasets from multiple independent and taxonomically diverse studies. Finally, we investigated the presence of mRNA targets for the most conserved of the small mitochondrial RNAs. We used a mix of computational and experimental methods to determine the presence of a target for the mt-miRNA. By combining multiple datasets and approaches, our study provides the first evidence of a functional relationship between the mitochondrial small RNAs and RNAi; thus, supporting the hypothesis that these RNAs are mitochondrially transcribed miRNAs (mt-miRNAs).

Results

Mt-miRNAs bind Ago2

To investigate the ability of the small mitochondrial RNAs to bind Ago2, we analysed a RIP-seq dataset including two different human cell lines; neural progenitor and teratoma-derived fibroblast. This dataset includes all the small RNAs that bind to Ago2, however, our focus was only on the mt-miRNAs, which have been previously ignored. Accordingly, we identified mt-miRNAs binding Ago2 in each of the two cell lines (**Fig.1**). Through this analysis, we found that the mt-miRNAs are present only in the Ago2 immunoprecipitations (IPs), mostly in the regions coding for mt-tRNAs, while the control IP samples have almost no mt-miRNA present. The expression of the mt-miRNAs differs across the different cell line types, which are reflective of different tissues, confirming previous studies demonstrating that mitochondrial transcription is usually tissue-specific (Mercer et al. 2011; Pozzi and Dowling 2019; Scheibye-Alsing et al. 2007). Aside from Ago2 RIP-seq, the authors of the original dataset performed other treatments on their samples to verify genuine binding of their focal miRNAs with Ago2. One of these treatments is particularly significant for our study: RNase I treatment. RNase I is an endonuclease able to digest RNAs that remain unbound to proteins, thus higher concentrations of RNAase I will be more effective in eliminating RNA contamination. We analysed samples treated with different RNase I concentrations and found that higher concentrations of RNase I have no effect on transcription levels of mt-miRNAs (See SI Fig.S1). Furthermore, we performed a gene-by-gene analysis on both Ago2-IP and control IP samples to verify whether the transcriptional signatures of the mt-miRNAs are consistent with them representing functional RNAs or noise (Pozzi et al. 2017; Mercer et al. 2011). These analyses showed that the mt-miRNAs with the highest level of transcription are encoded within mt-RNAs, confirming what found by previous studies in other human cell lines (Mercer et al. 2011; Ro et al. 2013). Due to the large number of genes analysed, the results of the gene-by-gene analysis are in the SI Fig.S2. Our analysis of the Ago2-IP samples provides the first evidence that the mt-miRNAs bind to Ago2, demonstrating their involvement in RNAi.

Mt-miRNAs are encoded in mt-tRNAs and protein-coding genes.

To further investigate the transcriptional signature of mt-miRNAs across tissues and species, we expanded the number of datasets analysed, by including more cell lines from humans and mice from other independent studies. This analysis confirmed the presence of mt-miRNAs binding to Ago2 in human (HeLa) and mouse embryonic cell lines (Scherer, Syverton, and Gey 1953) (**Fig.2**). To verify the enrichment of mt-miRNAs in Ago2-IP samples, we calculated the fold change between samples of the same cell line in which one samples had undergone the IP process, and the other had not. Through this experiment, we identified an enrichment in mt-miRNAs across Ago2 IP samples in both species. In HeLa cells, a mix of tRNAs and protein-coding genes are enriched for these mt-miRNAs, providing the first evidence the mt-miRNAs can be encoded within both mt-tRNAs and protein coding genes. Nonetheless, the most enriched mt-miRNA (5fold higher expression) is encoded in the mt-tRNA Met. This finding supports a previous study which found that mt-tRNA Met binds to Ago2 outside the mitochondria (Maniataki and Mourelatos 2005). The high expression of the mt-miRNA$_{Met}$ in these samples (HeLa cells), and in the previous analysis (neural progenitor and teratoma-derived fibroblast cell lines, Fig 1) suggests that although the mt-miRNAs have tissue-specific expression, some mt-miRNAs are conserved across tissues. The analysis of the mouse samples provides similar results. We found the mouse embryonic stem cells are enriched for mt-miRNAs encoded across multiple genes. Notably, the genes enriched for mt-miRNAs in the mouse samples differed from those in the human HeLa cell lines, with mt-ATP8 the only mt-miRNA exhibiting high expression in the Ago2-IP samples of both species.

To better understand the differences and similarities in the results between human and mouse samples, we performed a gene-by-by gene analysis of the transcriptional signature of the mt-miRNAs across the mouse embryonic stem cells and HeLa cell lines (See SI Fig.S3-4). Surprisingly, closer scrutiny of the putative mt-miRNA at mt-ATP8 in both species showed that the mt-miRNA is only likely to exist in humans (**Fig.3A**). In fact, while the human mt-ATP8 encodes an mt-miRNA (32nt long) with clear start and end, the same gene in mouse has only noise, without any clear transcriptional signature presence. This suggests that the enrichment patterns of the mt-miRNAs are accurate in predicting the presence of mt-miRNAs only when paired with detailed analysis of the mt-miRNAs transcriptional signature. Furthermore, we identified unusual transcriptional signatures that might shed light on some aspects of the biogenesis of the mt-miRNAs. Indeed, some mt-miRNAs are not fully encoded within a gene but overlap across two different genes (**Fig.3B**). We were able to identify this phenomenon only in the mouse samples, where both mt-miRNA Phe and mt-miRNA Thr partially overlap with the neighbour gene (an overlap of up to 6bp). Moreover, both of these mt-miRNAs have isoforms, “mt-isomiRs”, similar to observations in previous studies of nuclear miRNAs (Budak et al. 2016; Tan et al. 2014; Desvignes et al. 2015). These mt-miRNAs feature two different isoforms of different lengths: the short mt-isomiR ends where the first gene ends, while the longer mt-isomiR overlaps on the second gene by several nucleotides. Interestingly, these isoforms are quite long compared to other small RNAs such as miRNA (Ha and Kim 2014). In fact, the mt-isomiRs Phe are 32nt and 37nt long respectively, while the mt-isomiRs Thr are 37nt and 43nt long. Due to the length of these mt-isomiRs, it is possible that the longer isoform may represent a transitional stage for the shorter mature form. This phenomenon would be similar to what happens in piwi-interacting RNAs (piRNAs), in which proteins located on the surface of the mitochondria edit the length of piRNAs during their maturation process (Kim 2006; Nishimura et al. 2018; Ding et al. 2017; Bronkhorst and Ketting 2018). While the mechanism underpinning the generation of these mt-isomiRs remains unknown, our analysis provides the first report of mt-miRNAs encoded across the boundaries of genes, and, in general, the presence of mitochondrial products encoded across genes.

Mt-miRNAs and nuclear miRNAs have a different maturation process

Due to the presence of mt-isomiRs of different length, we investigated the possibility of these RNAs representing transition stages of the mt-miRNA maturation process. To address this, we re-purposed an RNA-seq dataset used for a study of the renal disease, Autosomal polycystic kidney disease, which includes several kidney samples from transgenic mice (Woo et al. 2017). The analysed dataset contains RNA samples in which the researchers had extracted both small and long RNAs, separately from the same individual mice (Woo et al. 2017). By comparing the transcriptional signature of small and long RNAs in the same samples, we can screen for the presence of matching transcriptional signatures, to determine whether either the start or end of each mt-miRNA matches with the expression of a mt-long RNA. However, we did not identify any matching transcriptional signatures between small and long RNAs (**Fig.4** SI Fig.S5). The small RNA samples show high levels of expression of mt-miRNAs in the mt-tRNAs Met and Ser 1, while the long RNAs samples have no distinct transcriptional hotspots. In fact, although most biological replicates of the long RNAs samples have a similar pattern of expression, we were not able to identify any signal suggesting the presence of functional RNAs. Thus, we argue that the consistent noisy pattern present in the long RNAs samples is most likely explained by the different chemical properties of the sequences, which would lead to some parts of the sequence being slightly overrepresented than others (Ross et al. 2013). On the contrary, the small RNA dataset exhibited a pattern concordant with the expression of mt-miRNAs: ~30nt sequences with clear cut start and end positions, consistent across multiple biological replicates (SI Fig S5). Thus our results suggest that the mt-miRNAs are not processed from pre-mt-miRNAs, as previously suggested, but are more likely to be matured directly from the polycistronic mt-RNA (Mercer et al. 2011; Rorbach and Minczuk 2012) from which mt-tRNAs and mt-rRNAs are matured.

Mt-miRNA$_{Met}$ is conserved across *Chordata*.

Nuclear miRNAs are usually conserved across species (Kenny et al. 2015; Lee, Risom, and Strauss 2007), and we investigated whether mt-miRNA$_{Met}$, which is conserved in both human and mouse (Fig. 1, 4), would be conserved in other model organisms. Thus, we analysed the expression and primary sequence of mt-miRNA$_{Met}$ across two tissues of four model organisms and verified its conservation in *Chordata* (**Fig 5**). The four model organisms (human, mouse, chicken, and zebrafish) show a very consistent transcriptional signature for the mt-miRNA$_{Met}$ across both brain and liver samples (**Fig 5A**). In fact, in all species the mt-miRNA$_{Met}$ 3' end is at the beginning of the mt-tRNA Met gene, while the 5' end is ~32nt after. However, one of the tissues has a different signature. Samples from the human liver exhibit a noisier transcriptional signature, suggesting that despite its broad conservation, this mt-miRNA might not be ubiquitously expressed across all tissues. Indeed, this result aligns with one of our previous findings in mouse, as we did not find the mt-miRNA$_{Met}$ in mouse embryonic stem cells, either transcribed or bound to Ago2. Similarly, to the transcriptional signature, we found that the primary sequence of the mt-miRNA$_{Met}$ was highly conserved across the four model organisms (**Fig 5B**). The mt-miRNA$_{Met}$ sequence is almost identical in all species, having only one polymorphism in amniotes, and four in zebrafish. When comparing the number of polymorphisms in the mt-miRNA$_{Met}$ to the polymorphisms in the rest of the mt-tRNA Met, we found that the polymorphisms are underrepresented in the mt-miRNA$_{Met}$. In fact, in human, mouse and chicken there are around 4 times more mutations in the rest of the mt-tRNA$_{Met}$ compared to the mt-miRNA$_{Met}$ (1/30 against 5/39), while in zebrafish there are around 2 times more (4/30 against 10/39). This suggests that the region harbouring the mt-miRNAMet might be under stronger purifying selection than its counterpart, which might be explained by the presence of overlapping selection due to the dual role of these regions in encoding both mt-miRNA$_{Met}$ and mt-tRNA Met.

The mt-miRNA$_{Met}$ targets CFLAR in human temporal lobe

We investigated the function of the very conserved mt-miRNA, mt-miRNA$_{Met}$, by verifying the presence of target mRNAs in the most well-characterized species, humans. To investigate the targets of mt-miRNA$_{Met}$, we screened all 67087 human transcripts using a computational target predictor, which found 8709 potential targets (**Fig 6A**). These targets have different scores, as some of them are more likely to be a genuine target compared to the others. However, since it is impossible to establish a fully objective threshold, we filtered these potential targets for a specific function. Although probably this mt-miRNA target many transcripts, we decided to focus on genes involved in insulin regulation, because this pathway is at the intersection of many diseases in which mitochondrial mutations and nuclear miRNAs appear to be involved (C. Lee et al. 2015; Mohlke et al. 2005; Chalkia et al. 2018; Duarte, Palmeira, and Rolo 2015; Heni et al. 2015). By filtering for insulin regulation, we identified 74 transcripts. These transcripts were then validated by using brain Ago2-IP samples enriched in miRNAs targets (**Fig 6B**). These brain samples are an ideal tissue to study the presence of these targets, given that mitochondrial mutations are related to many neurodegenerative diseases (Takasaki 2009; Dölle et al. 2016). Through this analysis we were able to validate one of the transcripts of the gene CASP8 and FADD-like apoptosis regulator (CFLAR), which, as the name suggests, is a key protein in the regulation of a caspase (CASP8) involved in the apoptosis pathway. However, as the mt-miRNA$_{Met}$ might not be the only miRNA binding this region, we verified how many nuclear miRNAs are able to bind the same region. By using programs for *in silico* prediction of miRNA targets in humans, we found that 13 nuclear miRNAs can bind the same region (**Fig 6C**). Nonetheless, by analysing other Ago2-IP samples from a similar part of the brain we found that none of these nuclear miRNAs are expressed in brain tissue, while the mt-miRNA$_{Met}$ is present (**Fig 6D**). To verify that the absence of these miRNAs was not due to technical mistakes, or bias of the library, we analysed three other common miRNAs (let-7, mir9, and mir100) finding that they are expressed in the brain, and in some instances, at a similar expression level to the mt-miRNA$_{Met}$. This analysis thus confirms that mt-miRNA$_{Met}$ binds to the CFLAR UTR.

To better understand the function of mt-miRNA$_{Met}$, we investigated the function and evolution of its target, CFLAR. By comparing the CFLAR genomic region across six primates with high-level nuclear genome sequencing, we found that the human CFLAR has a unique structure that might be related to the presence of the binding site of mt-miRNA$_{Met}$ (**Fig7A**). Indeed, the binding site of mt-miRNA$_{Met}$ is near the end of the 12kb 3' UTR of CFLAR. This is puzzling given the human CFLAR transcript consists of ~2kb of protein-coding sequence and its UTR is 6 times longer than the coding region. Furthermore, this UTR is not present in the transcript of any other species, and it is absent from the CFLAR gene of other primates. However, a sequence of 12kb length is unlikely to be absent from closely related species, thus we expanded our investigation to the flanking regions of the CFLAR gene. Through this analysis we found that other primates have the UTR vaguely conserved in the flanking regions of the CFLAR gene (**Fig7B,** See SI Fig.S6). This evidence suggests that the 12kb UTR of CFLAR is conserved and transcribed only in humans, thus its regulation by mt-miRNA$_{Met}$ is probably possible only in humans.

Discussion

The mtDNA is involved in RNA interference.

Our study confirmed the presence of two shared features of nuclear miRNAs and mt-miRNAs, suggesting a role for the mtDNA in RNAi. The first shared feature is the ability to bind to Ago2. We provided definitive evidence that the mt-miRNAs can bind Ago2, the key protein in gene regulation through RNAi. Although we found some noise in the Ago2-IP samples, the genes having clearly defined mt-miRNAs were strongly upregulated in the comparison between Ago2-IP and mock-IP, thus supporting a genuine binding of these RNAs to Ago2. We verified the consistent presence of mt-miRNAs binding to Ago2 across multiple independent studies, suggesting that technical differences in RNA preparation, or sequencing, did not significantly affect the mt-miRNAs. Furthermore, our analyses uncovered clear transcriptional signatures across multiple genes and samples, confirming similar patterns observed in previous studies that analysed the expression patterns of small mitochondrial RNAs not bound to Ago2 (Pozzi et al. 2017; Pozzi and Dowling 2019; Mercer et al. 2011). The second shared feature is the conservation of an mt-miRNA across multiple species. This feature is present in many miRNAs (Lee, Risom, and Strauss 2007; Ambros 2004), and has been used before for phylogenetic purposes (Lee, Risom, and Strauss 2007; Sempere et al. 2006; Kenny et al. 2015). However, ours is the first study to demonstrate that some mt-miRNAs are conserved across multiple diverged species within Chordata. Although the conservation of miRNAs usually relies on verifying the presence of the miRNA sequence in the genome of the target species, this method was not possible for the mt-miRNAs because their sequence lies cryptic within the sequence of other host genes – in the case of mt-miRNA$_{Met}$, this small RNA lies within the first half of the mt-tRNA Met gene. Thus, the presence of this sequence across multiple species only proves that the host gene is conserved. Nonetheless, by using small RNA expression data, we showed that the mt-miRNA$_{Met}$ is expressed across multiple tissues and species with a conserved transcriptional signature. Arguably, this level of evidence is more reliable than the benchmark generally used for the identification of miRNAs, because we not only demonstrated the presence of the sequences within the genome, but also showed a clear and conserved transcriptional signature. Indeed, the conservation of these RNAs, in both sequence and expression, across species suggests a function, and the only known function of Ago2 is inhibition of mRNA translation (Ha and Kim 2014; Cloonan 2015). However, definitive proof that the mt-miRNAs act in RNAi in the same way as the miRNAs is still warranted. Indeed, to verify if the effect is similar, it would be necessary to know the target of specific mt-miRNAs and demonstrate their downregulation in protein expression through western blot analysis (Taylor et al. 2013). However, as a preliminary investigation of this phenomena, we used a mix of computational and experimental data to identify *bone fide* targets of the mt-miRNA that are most likely to be functional.

The mt-miRNAMet - the selfish miRNA?

By investigating whether the mt-miRNA$_{Met}$ targets any mRNA in the human brain, we uncovered evidence that mt-miRNA might be involved in the protection of mitochondria from apoptosis. We identified strong evidence that CFLAR is a target of mt-miRNA$_{Met}$. CFLAR is a pseudo-caspase with multiple alternative transcripts known for having anti-apoptotic effect (S. A. Sarkar et al. 2009; He and He 2013). However, the mt-miRNA$_{Met}$ targets only a specific isoform of CFLAR, named CFLAR$_L$, since this is the only transcript transcribing the 12kb UTR hosting the Ago2 binding site. CFLAR$_L$ is different from the other transcripts (Yu, Jeffrey, and Shi 2009; Tsuchiya, Nakabayashi, and Nakano 2015), and it is one of the most efficient activators of procaspase-8, a key factor in apoptosis (Chang et al. 2003). The activation of procaspase-8 is a known trigger of apoptosis, a cellular mechanism in which the mitochondria are destroyed, along with the hosting cell (Chang et al. 2003). This process is fundamental for eukaryotic organisms, however, it means extinction for the genetic material within the cell, thus this process could be negatively selected at cellular level, even if it was under positive selection at the organismal level (increased fitness of the individual). Intriguingly, a mitochondrion expressing an RNA such as mt-miRNA$_{Met}$, would likely have an advantage relative to other mitochondria within a cellular environment if it was associated with a decreased possibility its host cell would die (Rand 2001).

This is supported by the evidence that the mitochondrial genome is under strong selection (Ferguson and von Borstel 1992; MacAlpine, Perlman, and Butow 2000). Therefore, we hypothesize that this mt-miRNA might be protecting mitochondrial fitness at cellular level, thus being the first “selfish miRNA”, a miRNA expressed from an organelle which “protects” its own genome from a cellular mechanism.

The mt-miRNAMet is involved in a human-specific regulation pathway

Our results suggest that the regulation of CFLAR by mt-miRNA$_{Met}$ is probably unique to humans. Indeed, mt-miRNA$_{Met}$ binds a UTR not transcribed in any other species. UTRs are not translated into proteins, and their role is usually to harbour regulatory sequences, thus, the presence of such a long UTR is puzzling (Yoon et al. 2008). Its presence, a long UTR sequence might be explained if this sequence harbours important regulatory sequences for this gene, such as the binding region for mt-miRNAMet. Our study supports this hypothesis, as a shortening or loss of this UTR, as seen in the other primates, would remove the binding site necessary for mt-miRNA$_{Met}$ to regulate CFLAR expression, thus disrupting transcript regulation. Indeed, losing this regulatory region might cause an increase of CFLAR$_L$ and potentially cell death through apoptosis (Tsuchiya, Nakabayashi, and Nakano 2015). Furthermore, our finding of a human-specific regulatory pathway for CFLAR aligns with the conflicting results found in experiments on CFLAR function (Tsuchiya, Nakabayashi, and Nakano 2015). In fact, most experiments have been performed in humans and mice, which although express the same protein, do not share the same regulatory region (Irmler et al. 1997; Shu, Halpin, and Goeddel 1997; Micheau et al. 2002). Indeed, mouse, like all other species, do not possess the 12kb UTR. Therefore, experiments on CFLAR function have conflicting results across humans and mice most likely due to the presence of species-specific regulation, such mt-miRNA$_{Met}$ as in humans, supporting the finding that mt-miRNA$_{Met}$ is part of a human-specific pathway.

Are the mt-miRNAs acting outside the mitochondria?

Notwithstanding the ability of the mt-miRNAs to bind Ago2, their function is expected to be exerted outside the mitochondria. However, the mechanism that would lead the mt-miRNAs out of the mitochondria and into the cytoplasm remains unknown. The ability to bind Ago2 suggests the ability of mt-miRNAs to translocate outside to reach the protein, most likely using the same, yet not fully characterized, transport mechanism used by other nucleic acids. Indeed, proteins like the PNPase have been demonstrated to be involved in RNA translocation through the mitochondrial membrane (Wang et al. 2012). Furthermore, a recent study showed that small mtDNA fragments travel outside of the mitochondria through mitochondrial pores that might be used by small RNAs too, since they have very similar characteristics (Moriyama, Koshiba, and Ichinohe 2019). Our study is not the first supporting the ability of mt-miRNAs to move outside the mitochondria. Indeed, Maniataki and Mourelatos (2005) previously reported that the mitochondrial mRNA mt-tRNA Met (that we have shown here encodes mt-miRNA$_{Met}$), is able to bind Ago2 outside the mitochondria. Furthermore, two other studies have previously provided evidence that Ago2 might be able to move inside of the mitochondria or to co-localize with it (Bandiera et al. 2011; Zhang et al. 2014). These studies performed western blots targeting Ago2 in samples with mitochondria-isolates, finding a positive signal for this protein. However, we believe the results of these studies better reflect co-localization of Ago2 and mitochondria than transport of Ago2 into the mitochondria. In fact, in the analyses of both Bandera et al. (2011) and Zhang et al. (2014), the Ago2 band was incredibly faint in the mitochondrial fractions, while other co-localization studies have revealed that Ago2 is very often co-localized with mitochondria (Olivieri et al. 2010; Vagin et al. 2013; Rogers et al. 2017). In sum, this indicates that Ago2 is being localized on the surface of the mitochondria, similarly to other proteins involved in RNAi (Vagin et al. 2013; Rogers et al. 2017). In this case, the evidence of mt-miRNA binding Ago2 suggests the presence of a mechanism enabling small RNAs transport through the mitochondrial membrane, potentially through PNPase (Dhir et al. 2018; D. Sarkar and Fisher 2006). Furthermore, we are aware of many types of small RNAs moving between mitochondria and cytoplasm, such as MitomiRs (Duarte, Palmeira, and

Rolo 2015), nuclear miRNAs localized within the mitochondria; and double-stranded RNAs, acting in the cytoplasm to trigger an immune response (Dhir et al. 2018). The mt-miRNAs might use the same, as-yet uncharacterized transport mechanism used by these other RNAs.

A new class of RNA without a new name

Small mitochondrial RNAs have been previously reported and given multiple new names across a series of studies. The first study reporting their existence did not explicitly name these RNAs, simply mentioning the presence of highly expressed small mitochondrial RNAs (Mercer et al. 2011). In 2013, Ro *et al.* showed that small mitochondrial RNAs were encoded in both mouse and humans, potentially having a function in mitochondrial regulation (Ro et al. 2013). This study named these RNAs as 'mitosRNAs', however, they annotated thousands of RNAs in this new class, in which the described RNAs had very different sizes (from ~15nt to ~120nt), were barely expressed, and had no associated evidence of function. In 2017, another group tried to name these RNAs using other standards (based on small RNAs of high expression, they called them ***sm***all ***h***ighly ***t***ranscribed small RNAs; smithRNAs), but introducing further confusion as to the nomenclature of these RNAs (Pozzi et al. 2017). This confusion persists due to the lack of precise standards when classifying new functional RNAs. Therefore, we have decided to keep the nomenclature simple and reasonable, adhering to the historical precedent, and following the example set by the authors discovering one of the most famous classes of small non-coding RNAs: the piRNAs (Grivna et al. 2006; Kim 2006). The discoverer of the piRNAs used their ability to bind the protein Piwi as the criterion to define them. Thus, we classify only those small mitochondrial RNAs able to bind Ago2 as mt-miRNAs. Furthermore, as mentioned above, we believe that further increasing the RNA nomenclature would not benefit the scientific community (Ro et al. 2013; Pozzi et al. 2017; Srinivasan and Das 2015), hence we propose adding a simple prefix (mt-) to define this class of small RNAs (Budak et al. 2016). Along with tRNAs and rRNAs, which receive the prefix mt- when referring to those encoded by mtDNA, we contend that miRNAs that bind Ago2 should similarly receive the same prefix when encoded in the mtDNA.

Some mt-miRNAs are inexplicably long

Some mt-miRNAs are almost twice the length of nuclear miRNAs but still able to bind Ago2, a case never seen before. Given that the protein complex necessary for the RNAi usually binds RNAs that are 20-30nt long (Ha and Kim 2014; Cloonan 2015), it remains unclear if longer mt-miRNAs will have similar functions once bound to Ago2. Some small RNAs in *C. elegans* are 34nt long, longer than most mt-miRNAs (~32nt), and are able to bind another protein involved in RNAi, Piwi (Ha and Kim 2014; Cloonan 2015)4). However, no previous miRNAs have been reported as long as the long isoform of mt-miRNA Thr (41nt), which usually would suggest this mt-miRNA is an intermediate precursor of the mt-miRNA. However, our analysis in paired long and short mt-RNAs found no support for this precursor hypothesis, due to the lack of longer RNAs in the mt-tRNA Thr. We suggest this RNA might be associated with a different function from canonical miRNAs. Indeed, in studies using Ago2-IP samples, the RNAs of ~35 length are usually considered fragments of miRNA targets, such as mRNAs, and not miRNAs (Boudreau et al. 2014). The hypothesis that the mt-tRNA Thr might be a target of Ago2 is supported by our analysis of mt-miRNAs in RNA-seq samples without Ago2-IP treatment in mouse. In these samples, we found clearly defined transcriptional signatures in several mt-tRNAs, such as mt-tRNA Ser 1, but not in the mt-tRNA Thr, suggesting that the mt-tRNA Thr does not encode for any small RNA and what was found in the Ago2-IP samples is a fragment of a longer RNA bound to Ago2. However, the samples analysed in both studies come from different tissues, thus it is possible that lack of mt-miRNAs from mt-tRNA Thr results from tissue-specific expression. The presence of long mt-miRNAs or the binding of Ago2 to these mt-tRNAs is puzzling, and without any obvious explanation. Nonetheless, this result shows that the mt-miRNAs and the miRNAs do not share a similar biogenesis, as the mt-miRNAs are not matured from ~70nt pre-miRNAs as are the nuclear miRNAs.

Multiple gene layers in the mitochondria.

Our study shows that the mtDNA harbours multiple gene layers, and that the products of overlapping genes are selected during the primary mt-RNA maturation. We demonstrated that the mt-miRNAs are encoded within other genes, specifically protein-coding genes or mt-tRNAs, aligning to observations of previous studies (Pozzi et al. 2017; Pozzi and Dowling 2019; Ro et al. 2013; Riggs et al. 2018; Bottje et al. 2017; Mercer et al. 2011; Larriba, Rial, and Del Mazo 2018). Likewise, other mitochondrial products are encoded within multiple genes: double-stranded RNAs (Dhir et al. 2018), long non-coding RNAs (Rackham et al. 2011), and proteins (C. Lee et al. 2015; C. Lee, Yen, and Cohen 2013; K. H. Kim et al. 2018). However, some of these mitochondrial products have been better described than others. The double-stranded RNAs have been discovered only very recently, and function in the immune response (Dhir et al. 2018). The function of long non-coding RNAs is still unknown, but their nuclear counterparts are known for having extensive roles in gene expression regulation (Mattick 2003). There are also several newly discovered mitochondrial proteins, with 'humanin' the first and most intensely studied (Hashimoto et al. 2001). This protein is encoded within rRNA 16S and, although its function is not fully understood, it seems to somehow have protective effects against Alzheimer's disease (Matsuoka 2009). Thus, the number of functional mitochondrial products identified in recent years is quickly increasing, providing strong support for the presence of multiple new regulatory layers within the mitochondrial genome. These findings are intriguing because the presence of these novel mitochondrial products suggests a reinterpretation of the candidate mechanisms by which pathogenic mutations in the mtDNA sequence exert their effects on organismal health and function may be required in several cases.

The mt-miRNAs change our perspective on mitonuclear interactions.

The mt-miRNAs add another level of complexity to the dynamics of mitonuclear communications. Indeed, because the mt-miRNAs share features with the miRNAs, we expect them to affect cell biology in a similar manner. The miRNAs play pervasive roles in cell regulation (Friedman et al. 2009), and the presence of mt-miRNAs indicates that the mtDNA might broadly affect cell regulation as well. The mt-miRNAs might act as a vector to affect nuclear regulation in many ways. Indeed, by using sequence complementarity, mt-miRNAs could lead Ago2 to interfere with the gene expression of virtually any nuclear mRNA that exhibits partial sequence complementarity (Cloonan 2015; Ambros 2004). This could help explain many of the diverse phenotypes linked to mtDNA mutations observed in recent years (Hopkins et al. 2017; Hudson et al. 2014; Dobler et al. 2014). Recent studies, for example, have demonstrated clear associations of mtDNA mutations on a range of phenotypes, ranging from thermal tolerance, to cognitive function, to fertility (Lajbner et al. 2018; Camus et al. 2017; Yee, Sutton, and Dowling 2013; Dowling, Abiega, and Arnqvist 2007; Roubertoux et al. 2003), as well as a range of human diseases not previously associated with mitochondrial genetics (Hopkins et al. 2017; Hudson et al. 2014). However, the mechanisms underpinning these diverse effects associated with the mitochondrial genome are yet to be understood. We believe that RNAi, mediated through mt-miRNAs, might well provide the explanation for the diversity of phenotypic effects associated with mitochondrial sequence variation.

The unusual maturation of the mt-miRNAs.

The mitochondria generally transcribe long mt-RNA precursors encoding multiple genes (D'Souza and Minczuk 2018), potentially including the mt-miRNAs. Indeed, the transcriptional signature of the mt-miRNAs does not match the transcriptional signature of the long mitochondrial RNAs (~70nt), suggesting that the mt-miRNAs are matured directly from the pre-mt-RNAs. Mitochondrial genes are usually transcribed in few long polycistronic mt-RNA, RNAs including multiple genes, and these RNAs are called pre-mt-RNAs (Van Haute et al. 2015). Although previously the presence of intermediate RNAs were proposed for the mt-miRNAs (Pozzi et al. 2017), our results suggest that the mt-miRNAs originate from the pre-mt-RNAs. Indeed, we identified multiple cases of specific isoforms of mt-miRNAs whose sequence spanned two genes. In these cases, it is clear that miRNAs cannot originate from a mature tRNA, since they would then lack part of the sequence. Furthermore, mt-tRNAs undergo a vast number of nucleotide modifications through their maturation (Richter et al. 2018; Pan 2018), thus mt-miRNAs originating from mature tRNAs will have these RNA modifications, thus escaping normal sequencing methods. That is, it would be impossible for the mt-miRNAs presented in our analyses to exhibit these modified nucleotides, given the methods used for the RNA sequencing would not have captured these modified RNAs (Zheng et al. 2015). Therefore, according to our results, the mt-miRNAs encoded in mt-tRNAs are either matured from them before the posttranscriptional modifications, or directly from the mt-RNA precursor. Our study provides the first clues as to the biogenesis of the mt-miRNAs and, in general, supports the hypothesis that the mitochondria regulates its products mostly from the pre-mt-RNA (Sloan and Wu 2016; Lavrov et al. 2016).

Overlapping selection.

Mitochondrial genes are known to be under constant strong purifying selection (J. W. Ballard and Kreitman 1995; Ballard and Whitlock 2004), however, the existence of multiple gene layers suggests the presence of overlapping selection pressures that could alter the strength or direction of selection on particular regions of mtDNA sequence. Indeed, theory suggests that multiple products encoded within the same region of a gene would affect the selection pressure on this region (Rogozin et al. 2002). We hypothesise that the presence of overlapping genes such as mt-miRNAs and dsRNAs (Dhir et al. 2018), on canonical genes such as mt-tRNAs, will increase the effect of purifying selection in order to preserve the function of these products. Our results support this hypothesis, showing that within the mt-tRNA Met, the region harbouring the mt-miRNAMet has fewer polymorphisms than did the rest of the tRNA gene. While more analysis across a broad number of species and mt-miRNAs will be necessary to fully test this hypothesis, this study provides the first evidence that this phenomenon might exist.

Author Contributions

A.P. and D.K.D. conceived the study. A.P. performed the analyses. A.P. and D.K.D. discussed the results and wrote the manuscript.

Acknowledgement

We wish to acknowledge all researchers who made their data freely available, thus giving us the chance to perform this work. The study was funded by the Australian Research Council, through grants DP170100165 and FT160100022 to D.K.D, and through Australian Government Research Training Program (RTP) Scholarship to A.P.

References

Ambros, Victor. 2004. "The Functions of Animal microRNAs." *Nature* 431 (7006): 350–55.

Ballard, J. William O., and Michael C. Whitlock. 2004. "The Incomplete Natural History of Mitochondria." *Molecular Ecology* 13 (4): 729–44.

Ballard, J. W., and M. Kreitman. 1995. "Is Mitochondrial DNA a Strictly Neutral Marker?" *Trends in Ecology & Evolution* 10 (12): 485–88.

Bandiera, Simonetta, Silvia Rüberg, Muriel Girard, Nicolas Cagnard, Sylvain Hanein, Dominique Chrétien, Arnold Munnich, Stanislas Lyonnet, and Alexandra Henrion-Caude. 2011. "Nuclear Outsourcing of RNA Interference Components to Human Mitochondria." *PloS One* 6

Bottje, Walter G., Bhuwan Khatri, Stephanie A. Shouse, Dongwon Seo, Barbara Mallmann, Sara K. Orlowski, Jeonghoon Pan, et al. 2017. "Identification and Differential Abundance of Mitochondrial Genome Encoding Small RNAs (mitosRNA) in Breast Muscles of Modern Broilers and Unselected Chicken Breed." *Frontiers in Physiology* 8 (October): 816.

Boudreau, Ryan L., Peng Jiang, Brian L. Gilmore, Ryan M. Spengler, Rebecca Tirabassi, Jay A. Nelson, Christopher A. Ross, Yi Xing, and Beverly L. Davidson. 2014. "Transcriptome-Wide Discovery of microRNA Binding Sites in Human Brain." *Neuron* 81 (2): 294–305.

Bronkhorst, Alfred W., and René F. Ketting. 2018. "Trimming It Short: PNLDC1 Is Required for piRNA Maturation during Mouse Spermatogenesis." *EMBO Reports*. https://doi.org/10.15252/embr.201845824.

Budak, Hikmet, Reyyan Bulut, Melda Kantar, and Burcu Alptekin. 2016. "MicroRNA Nomenclature and the Need for a Revised Naming Prescription." *Briefings in Functional Genomics* 15 (1): 65–71.

Camus, M. Florencia, Jonci N. Wolff, Carla M. Sgrò, and Damian K. Dowling. 2017. "Experimental Support That Natural Selection Has Shaped the Latitudinal Distribution of

Mitochondrial Haplotypes in Australian Drosophila Melanogaster." *Molecular Biology and Evolution* 34 (10): 2600–2612.

Chalkia, Dimitra, Yi-Cheng Chang, Olga Derbeneva, Maria Lvova, Ping Wang, Dan Mishmar, Xiaogang Liu, Larry N. Singh, Lee-Ming Chuang, and Douglas C. Wallace. 2018. "Mitochondrial DNA Associations with East Asian Metabolic Syndrome." *Biochimica et Biophysica Acta* 1859 (9): 878–92.

Chang, David W., Zheng Xing, Vanessa L. Capacio, Marcus E. Peter, and Xiaolu Yang. 2003. "Interdimer Processing Mechanism of Procaspase-8 Activation." *The EMBO Journal* 22 (16): 4132–42.

Cloonan, Nicole. 2015. "Re-Thinking miRNA-mRNA Interactions: Intertwining Issues Confound Target Discovery." *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 37 (4): 379–88.

Desvignes, T., P. Batzel, E. Berezikov, K. Eilbeck, J. T. Eppig, M. S. McAndrews, A. Singer, and J. H. Postlethwait. 2015. "miRNA Nomenclature: A View Incorporating Genetic Origins, Biosynthetic Pathways, and Sequence Variants." *Trends in Genetics: TIG* 31 (11): 613–26.

Dhir, Ashish, Somdutta Dhir, Lukasz S. Borowski, Laura Jimenez, Michael Teitell, Agnès Rötig, Yanick J. Crow, et al. 2018. "Mitochondrial Double-Stranded RNA Triggers Antiviral Signalling in Humans." *Nature* 560 (7717): 238–42.

Ding, Deqiang, Jiali Liu, Kunzhe Dong, Uros Midic, Rex A. Hess, Huirong Xie, Elena Y. Demireva, and Chen Chen. 2017. "PNLDC1 Is Essential for piRNA 3' End Trimming and Transposon Silencing during Spermatogenesis in Mice." *Nature Communications* 8 (1): 819.

Dobler, R., B. Rogell, F. Budar, and D. K. Dowling. 2014. "A Meta-Analysis of the Strength and Nature of Cytoplasmic Genetic Effects." *Journal of Evolutionary Biology* 27 (10): 2021–34.

Dölle, Christian, Irene Flønes, Gonzalo S. Nido, Hrvoje Miletic, Nelson Osuagwu, Stine Kristoffersen, Peer K. Lilleng, et al. 2016. "Defective Mitochondrial DNA Homeostasis in the Substantia Nigra in Parkinson Disease." *Nature Communications* 7 (November): 13548.

Dowling, Damian K., Katia Chávez Abiega, and Göran Arnqvist. 2007. "Temperature-Specific Outcomes of Cytoplasmic-Nuclear Interactions on Egg-to-Adult Development Time in Seed Beetles." *Evolution; International Journal of Organic Evolution* 61 (1): 194–201.

D'Souza, Aaron R., and Michal Minczuk. 2018. "Mitochondrial Transcription and Translation: Overview." *Essays in Biochemistry* 62 (3): 309–20.

Duarte, Filipe V., Carlos M. Palmeira, and Anabela P. Rolo. 2015. "The Emerging Role of MitomiRs in the Pathophysiology of Human Disease." *Advances in Experimental Medicine and Biology* 888: 123–54.

Ferguson, L. R., and R. C. von Borstel. 1992. "Induction of the Cytoplasmic 'Petite' Mutation by Chemical and Physical Agents in Saccharomyces Cerevisiae." *Mutation Research* 265 (1): 103–48.

Friedman, Robin C., Kyle Kai-How Farh, Christopher B. Burge, and David P. Bartel. 2009. "Most Mammalian mRNAs Are Conserved Targets of microRNAs." *Genome Research* 19 (1): 92–105.

Grivna, Shane T., Ergin Beyret, Zhong Wang, and Haifan Lin. 2006. "A Novel Class of Small RNAs in Mouse Spermatogenic Cells." *Genes & Development* 20 (13): 1709–14.

Ha, Minju, and V. Narry Kim. 2014. "Regulation of microRNA Biogenesis." *Nature Reviews. Molecular Cell Biology* 15 (8): 509–24.

Hashimoto, Y., T. Niikura, H. Tajima, T. Yasukawa, H. Sudo, Y. Ito, Y. Kita, et al. 2001. "A

Rescue Factor Abolishing Neuronal Cell Death by a Wide Spectrum of Familial Alzheimer's Disease Genes and Abeta." *Proceedings of the National Academy of Sciences of the United States of America* 98 (11): 6336–41.

He, Ming-Xiao, and You-Wen He. 2013. "CFLAR/c-FLIPL: A Star in the Autophagy, Apoptosis and Necroptosis Alliance." *Autophagy* 9 (5): 791–93.

Heni, Martin, Stephanie Kullmann, Hubert Preissl, Andreas Fritsche, and Hans-Ulrich Häring. 2015. "Impaired Insulin Action in the Human Brain: Causes and Metabolic Consequences." *Nature Reviews. Endocrinology* 11 (12): 701–11.

Hill, Geoffrey E., Justin C. Havird, Daniel B. Sloan, Ronald S. Burton, Chris Greening, and Damian K. Dowling. 2018. "Assessing the Fitness Consequences of Mitonuclear Interactions in Natural Populations." *Biological Reviews of the Cambridge Philosophical Society*, December. https://doi.org/10.1111/brv.12493.

Hopkins, Julia F., Veronica Y. Sabelnykova, Joachim Weischenfeldt, Ronald Simon, Jennifer A. Aguiar, Rached Alkallas, Lawrence E. Heisler, et al. 2017. "Mitochondrial Mutations Drive Prostate Cancer Aggression." *Nature Communications* 8 (1): 656.

Hudson, Gavin, Aurora Gomez-Duran, Ian J. Wilson, and Patrick F. Chinnery. 2014. "Recent Mitochondrial DNA Mutations Increase the Risk of Developing Common Late-Onset Human Diseases." *PLoS Genetics* 10 (5): e1004369.

Irmler, M., M. Thome, M. Hahne, P. Schneider, K. Hofmann, V. Steiner, J. L. Bodmer, et al. 1997. "Inhibition of Death Receptor Signals by Cellular FLIP." *Nature* 388 (6638): 190–95.

James, Avis C., and J. William O. Ballard. 2003. "Mitochondrial Genotype Affects Fitness in Drosophila Simulans." *Genetics* 164 (1): 187–94.

Kenny, Nathan J., Yung Wa Sin, Alexander Hayward, Jordi Paps, Ka Hou Chu, and Jerome H. L. Hui. 2015. "The Phylogenetic Utility and Functional Constraint of microRNA Flanking Sequences." *Proceedings. Biological Sciences / The Royal Society* 282 (1803): 20142983.

Kim, Kyung Hwa, Jyung Mean Son, Bérénice A. Benayoun, and Changhan Lee. 2018. "The Mitochondrial-Encoded Peptide MOTS-c Translocates to the Nucleus to Regulate Nuclear Gene Expression in Response to Metabolic Stress." *Cell Metabolism*, June. https://doi.org/10.1016/j.cmet.2018.06.008.

Kim, V. Narry. 2006. "Small RNAs Just Got Bigger: Piwi-Interacting RNAs (piRNAs) in Mammalian Testes." *Genes & Development* 20 (15): 1993–97.

Kopinski, Piotr K., Kevin A. Janssen, Patrick M. Schaefer, Sophie Trefely, Caroline E. Perry, Prasanth Potluri, Jesus A. Tintos-Hernandez, et al. 2019. "Regulation of Nuclear Epigenome by Mitochondrial DNA Heteroplasmy." *Proceedings of the National Academy of Sciences of the United States of America* 116 (32): 16028–35.

Lajbner, Zdeněk, Reuven Pnini, M. Florencia Camus, Jonathan Miller, and Damian K. Dowling. 2018. "Experimental Evidence That Thermal Selection Shapes Mitochondrial Genome Evolution." *Scientific Reports* 8 (1): 9500.

Larriba, Eduardo, Eduardo Rial, and Jesús Del Mazo. 2018. "The Landscape of Mitochondrial Small Non-Coding RNAs in the PGCs of Male Mice, Spermatogonia, Gametes and in Zygotes." *BMC Genomics* 19 (1): 634.

Lavrov, Dennis V., Marcin Adamski, Pierre Chevaldonné, and Maja Adamska. 2016. "Extensive Mitochondrial mRNA Editing and Unusual Mitochondrial Genome Organization in Calcaronean Sponges." *Current Biology: CB* 26 (1): 86–92.

Lee, Changhan, Kelvin Yen, and Pinchas Cohen. 2013. "Humanin: A Harbinger of Mitochondrial-

Derived Peptides?” *Trends in Endocrinology and Metabolism: TEM* 24 (5): 222–28.
Lee, Changhan, Jennifer Zeng, Brian G. Drew, Tamer Sallam, Alejandro Martin-Montalvo, Junxiang Wan, Su-Jeong Kim, et al. 2015. “The Mitochondrial-Derived Peptide MOTS-c Promotes Metabolic Homeostasis and Reduces Obesity and Insulin Resistance.” *Cell Metabolism* 21 (3): 443–54.
Lee, Chung-Tien, Tyler Risom, and William M. Strauss. 2007. “Evolutionary Conservation of microRNA Regulatory Circuits: An Examination of microRNA Gene Complexity and Conserved microRNA-Target Interactions through Metazoan Phylogeny.” *DNA and Cell Biology* 26 (4): 209–18.
MacAlpine, D. M., P. S. Perlman, and R. A. Butow. 2000. “The Numbers of Individual Mitochondrial DNA Molecules and Mitochondrial DNA Nucleoids in Yeast Are Co-Regulated by the General Amino Acid Control Pathway.” *The EMBO Journal* 19 (4): 767–75.
Maniataki, Elisavet, and Zissimos Mourelatos. 2005. “Human Mitochondrial tRNAMet Is Exported to the Cytoplasm and Associates with the Argonaute 2 Protein.” *RNA* 11 (6): 849–52.
Matsuoka, Masaaki. 2009. “Humanin; a Defender against Alzheimer’s Disease?” *Recent Patents on CNS Drug Discovery* 4 (1): 37–42.
Mattick, John S. 2003. “Challenging the Dogma: The Hidden Layer of Non-Protein-Coding RNAs in Complex Organisms.” *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 25 (10): 930–39.
Mercer, Tim R., Shane Neph, Marcel E. Dinger, Joanna Crawford, Martin A. Smith, Anne-Marie J. Shearwood, Eric Haugen, et al. 2011. “The Human Mitochondrial Transcriptome.” *Cell* 146 (4): 645–58.
Micheau, Olivier, Margot Thome, Pascal Schneider, Nils Holler, Jürg Tschopp, Donald W. Nicholson, Christophe Briand, and Markus G. Grütter. 2002. “The Long Form of FLIP Is an Activator of Caspase-8 at the Fas Death-Inducing Signaling Complex.” *The Journal of Biological Chemistry* 277 (47): 45162–71.
Mohlke, Karen L., Anne U. Jackson, Laura J. Scott, Erin C. Peck, Yong D. Suh, Peter S. Chines, Richard M. Watanabe, et al. 2005. “Mitochondrial Polymorphisms and Susceptibility to Type 2 Diabetes-Related Traits in Finns.” *Human Genetics* 118 (2): 245–54.
Moriyama, Miyu, Takumi Koshiba, and Takeshi Ichinohe. 2019. “Influenza A Virus M2 Protein Triggers Mitochondrial DNA-Mediated Antiviral Immune Responses.” *Nature Communications* 10 (1): 4624.
Nishimura, Toru, Ippei Nagamori, Tsunetoshi Nakatani, Natsuko Izumi, Yukihide Tomari, Satomi Kuramochi-Miyagawa, and Toru Nakano. 2018. “PNLDC1, Mouse Pre-piRNA Trimmer, Is Required for Meiotic and Post-Meiotic Male Germ Cell Development.” *EMBO Reports* 19 (3). https://doi.org/10.15252/embr.201744957.
Olivieri, Daniel, Martina M. Sykora, Ravi Sachidanandam, Karl Mechtler, and Julius Brennecke. 2010. “An in Vivo RNAi Assay Identifies Major Genetic and Cellular Requirements for Primary piRNA Biogenesis in Drosophila.” *The EMBO Journal* 29 (19): 3301–17.
Pan, Tao. 2018. “Modifications and Functional Genomics of Human Transfer RNA.” *Cell Research* 28 (4): 395–404.
Picard, Martin, Douglas C. Wallace, and Yan Burelle. 2016. “The Rise of Mitochondria in Medicine.” *Mitochondrion* 30 (September): 105–16.
Pozzi, Andrea, and Damian K. Dowling. 2019. “The Genomic Origins of Small Mitochondrial RNAs: Are They Transcribed by the Mitochondrial DNA or by Mitochondrial Pseudogenes

within the Nucleus (NUMTs)?” *Genome Biology and Evolution*, June. https://doi.org/10.1093/gbe/evz132.

Pozzi, Andrea, Federico Plazzi, Liliana Milani, Fabrizio Ghiselli, and Marco Passamonti. 2017. “SmithRNAs: Could Mitochondria ‘Bend’ Nuclear Regulation?” *Molecular Biology and Evolution*, April. https://doi.org/10.1093/molbev/msx140.

Rackham, Oliver, Anne-Marie J. Shearwood, Tim R. Mercer, Stefan M. K. Davies, John S. Mattick, and Aleksandra Filipovska. 2011. “Long Noncoding RNAs Are Generated from the Mitochondrial Genome and Regulated by Nuclear-Encoded Proteins.” *RNA* 17 (12): 2085–93.

Rand, David M. 2001. “The Units of Selection on Mitochondrial DNA.” *Annual Review of Ecology and Systematics* 32 (1): 415–48.

Rand, David M., Adam Fry, and Lea Sheldahl. 2006. “Nuclear–Mitochondrial Epistasis and Drosophila Aging: Introgression of Drosophila Simulans mtDNA Modifies Longevity in D. Melanogaster Nuclear Backgrounds.” *Genetics*. https://doi.org/10.1534/genetics.105.046698.

Richter, Uwe, Molly E. Evans, Wesley C. Clark, Paula Marttinen, Eric A. Shoubridge, Anu Suomalainen, Anna Wredenberg, Anna Wedell, Tao Pan, and Brendan J. Battersby. 2018. “RNA Modification Landscape of the Human Mitochondrial tRNALys Regulates Protein Synthesis.” *Nature Communications* 9 (1): 3966.

Riggs, Claire L., Amanda Summers, Daniel E. Warren, Göran E. Nilsson, Sjannie Lefevre, W. W. Dowd, Sarah Milton, and Jason E. Podrabsky. 2018. “Small Non-Coding RNA Expression and Vertebrate Anoxia Tolerance.” *Frontiers in Genetics* 9 (July): 230.

Rogers, Alicia K., Kathy Situ, Edward M. Perkins, and Katalin Fejes Toth. 2017. “Zucchini-Dependent piRNA Processing Is Triggered by Recruitment to the Cytoplasmic Processing Machinery.” *Genes & Development* 31 (18): 1858–69.

Rogozin, Igor B., Alexey N. Spiridonov, Alexander V. Sorokin, Yuri I. Wolf, I. King Jordan, Roman L. Tatusov, and Eugene V. Koonin. 2002. “Purifying and Directional Selection in Overlapping Prokaryotic Genes.” *Trends in Genetics: TIG* 18 (5): 228–32.

Rorbach, Joanna, and Michal Minczuk. 2012. “The Post-Transcriptional Life of Mammalian Mitochondrial RNA.” *Biochemical Journal* 444 (3): 357–73.

Ro, Seungil, Hsiu-Yen Ma, Chanjae Park, Nicole Ortogero, Rui Song, Grant W. Hennig, Huili Zheng, et al. 2013. “The Mitochondrial Genome Encodes Abundant Small Noncoding RNAs.” *Cell Research* 23 (6): 759–74.

Ross, Michael G., Carsten Russ, Maura Costello, Andrew Hollinger, Niall J. Lennon, Ryan Hegarty, Chad Nusbaum, and David B. Jaffe. 2013. “Characterizing and Measuring Bias in Sequence Data.” *Genome Biology* 14 (5): R51.

Roubertoux, Pierre L., Frans Sluyter, Michèle Carlier, Brice Marcet, Fatima Maarouf-Veray, Chabane Chérif, Charlotte Marican, et al. 2003. “Mitochondrial DNA Modifies Cognition in Interaction with the Nuclear Genome and Age in Mice.” *Nature Genetics* 35 (1): 65–69.

Sarkar, Devanand, and Paul B. Fisher. 2006. “Human Polynucleotide Phosphorylase (hPNPase Old-35): An RNA Degradation Enzyme with Pleiotrophic Biological Effects.” *Cell Cycle* 5 (10): 1080–84.

Sarkar, S. A., B. Kutlu, K. Velmurugan, S. Kizaka-Kondoh, C. E. Lee, R. Wong, A. Valentine, H. W. Davidson, J. C. Hutton, and S. Pugazhenthi. 2009. “Cytokine-Mediated Induction of Anti-Apoptotic Genes That Are Linked to Nuclear Factor Kappa-B (NF-kappaB) Signalling in Human Islets and in a Mouse Beta Cell Line.” *Diabetologia* 52 (6): 1092–1101.

Scheibye-Alsing, Karsten, Susanna Cirera, Michael J. Gilchrist, Merete Fredholm, and Jan

Gorodkin. 2007. “EST Analysis on Pig Mitochondria Reveal Novel Expression Differences between Developmental and Adult Tissues.” *BMC Genomics* 8 (October): 367.

Scherer, W. F., J. T. Syverton, and G. O. Gey. 1953. “… the Propagation in Vitro of Poliomyelitis Viruses: IV. Viral Multiplication in a Stable Strain of Human Malignant Epithelial Cells (strain HeLa) Derived from an Epidermoid ….” *Journal of Experimental*. http://jem.rupress.org/content/97/5/695.abstract.

Sempere, Lorenzo F., Charles N. Cole, Mark A. McPeek, and Kevin J. Peterson. 2006. “The Phylogenetic Distribution of Metazoan microRNAs: Insights into Evolutionary Complexity and Constraint.” *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution* 306 (6): 575–88.

Shu, H. B., D. R. Halpin, and D. V. Goeddel. 1997. “Casper Is a FADD- and Caspase-Related Inducer of Apoptosis.” *Immunity* 6 (6): 751–63.

Sloan, Daniel B., Jessica M. Warren, Alissa M. Williams, Zhiqiang Wu, Salah E. Abdel-Ghany, Adam J. Chicco, and Justin C. Havird. 2018. “Cytonuclear Integration and Co-Evolution.” *Nature Reviews. Genetics*, July. https://doi.org/10.1038/s41576-018-0035-9.

Sloan, Daniel B., and Zhiqiang Wu. 2016. “Molecular Evolution: The Perplexing Diversity of Mitochondrial RNA Editing Systems.” *Current Biology: CB*.

Song, Gyun Jee, and Vivian Lewis. 2008. “Mitochondrial DNA Integrity and Copy Number in Sperm from Infertile Men.” *Fertility and Sterility* 90 (6): 2238–44.

Sprenger, Hans-Georg, and Thomas Langer. 2019. “The Good and the Bad of Mitochondrial Breakups.” *Trends in Cell Biology* 0 (0). https://doi.org/10.1016/j.tcb.2019.08.003.

Srinivasan, Hemalatha, and Samarjit Das. 2015. “Mitochondrial miRNA (MitomiR): A New Player in Cardiovascular Health.” *Canadian Journal of Physiology and Pharmacology* 93 (10): 855–61.

Takasaki, Shigeru. 2009. “Mitochondrial Haplogroups Associated with Japanese Centenarians, Alzheimer’s Patients, Parkinson's Patients, Type 2 Diabetic Patients and Healthy Non-Obese Young Males.” *Journal of Genetics and Genomics = Yi Chuan Xue Bao* 36 (7): 425–34.

Tan, Geok Chin, Elcie Chan, Attila Molnar, Rupa Sarkar, Diana Alexieva, Ihsan Mad Isa, Sophie Robinson, et al. 2014. “5’ isomiR Variation Is of Functional and Evolutionary Importance.” *Nucleic Acids Research* 42 (14): 9424–35.

Taylor, Sean C., Thomas Berkelman, Geetha Yadav, and Matt Hammond. 2013. “A Defined Methodology for Reliable Quantification of Western Blot Data.” *Molecular Biotechnology* 55 (3): 217–26.

Townley-Tilson, W. H. D., S. A. Pendergrass, and W. F. Marzluff. 2006. “Genome-Wide Analysis of mRNAs Bound to the Histone Stem–loop Binding Protein.” *RNA* . http://rnajournal.cshlp.org/content/12/10/1853.short.

Tsuchiya, Yuichi, Osamu Nakabayashi, and Hiroyasu Nakano. 2015. “FLIP the Switch: Regulation of Apoptosis and Necroptosis by cFLIP.” *International Journal of Molecular Sciences* 16 (12): 30321–41.

Vagin, Vasily V., Yang Yu, Anna Jankowska, Yicheng Luo, Kaja A. Wasik, Colin D. Malone, Emily Harrison, et al. 2013. “Minotaur Is Critical for Primary piRNA Biogenesis.” *RNA* 19 (8): 1064–77.

Van Haute, Lindsey, Sarah F. Pearce, Christopher A. Powell, Aaron R. D’Souza, Thomas J. Nicholls, and Michal Minczuk. 2015. “Mitochondrial Transcript Maturation and Its Disorders.” *Journal of Inherited Metabolic Disease* 38 (4): 655–80.

Wallace, Douglas C. 2018. “Mitochondrial Genetic Medicine.” *Nature Genetics* 50 (12): 1642–49.
Wang, Geng, Eriko Shimada, Carla M. Koehler, and Michael A. Teitell. 2012. “PNPASE and RNA Trafficking into Mitochondria.” *Biochimica et Biophysica Acta* 1819 (9-10): 998–1007.
Woo, Yu Mi, Do Yeon Kim, Nam Jin Koo, Yong-Min Kim, Sunyoung Lee, Je Yeong Ko, Yubin Shin, et al. 2017. “Profiling of miRNAs and Target Genes Related to Cystogenesis in ADPKD Mouse Models.” *Scientific Reports* 7 (1): 14151.
Wu, Zheng, Sebastian Oeck, A. Phillip West, Kailash C. Mangalhara, Alva G. Sainz, Laura E. Newman, Xiao-Ou Zhang, et al. 2019. “Mitochondrial DNA Stress Signalling Protects the Nuclear Genome.” *Nature Metabolism*, December. https://doi.org/10.1038/s42255-019-0150-8.
Yee, Winston K. W., Katherine L. Sutton, and Damian K. Dowling. 2013. “In Vivo Male Fertility Is Affected by Naturally Occurring Mitochondrial Haplotypes.” *Current Biology: CB* 23 (2): R55–56.
Yoon, Kihoon, Daijin Ko, Mark Doderer, Carolina B. Livi, and Luiz O. F. Penalva. 2008. “Over-Represented Sequences Located on 3’ UTRs Are Potentially Involved in Regulatory Functions.” *RNA Biology* 5 (4): 255–62.
Yu, Jong W., Philip D. Jeffrey, and Yigong Shi. 2009. “Mechanism of Procaspase-8 Activation by c-FLIPL.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (20): 8169–74.
Zhang, Xiaorong, Xinxin Zuo, Bo Yang, Zongran Li, Yuanchao Xue, Yu Zhou, Jie Huang, et al. 2014. “MicroRNA Directly Enhances Mitochondrial Translation during Muscle Differentiation.” *Cell* 158 (3): 607–19.
Zheng, Guanqun, Yidan Qin, Wesley C. Clark, Qing Dai, Chengqi Yi, Chuan He, Alan M. Lambowitz, and Tao Pan. 2015. “Efficient and Quantitative High-Throughput tRNA Sequencing.” *Nature Methods* 12 (9): 835–37.
Zhu, Chen-Tseh, Paul Ingelmo, and David M. Rand. 2014. “G×G×E for Lifespan in Drosophila: Mitochondrial, Nuclear, and Dietary Interactions That Modify Longevity.” *PLoS Genetics* 10 (5): e1004354.

Figures

Figure 1

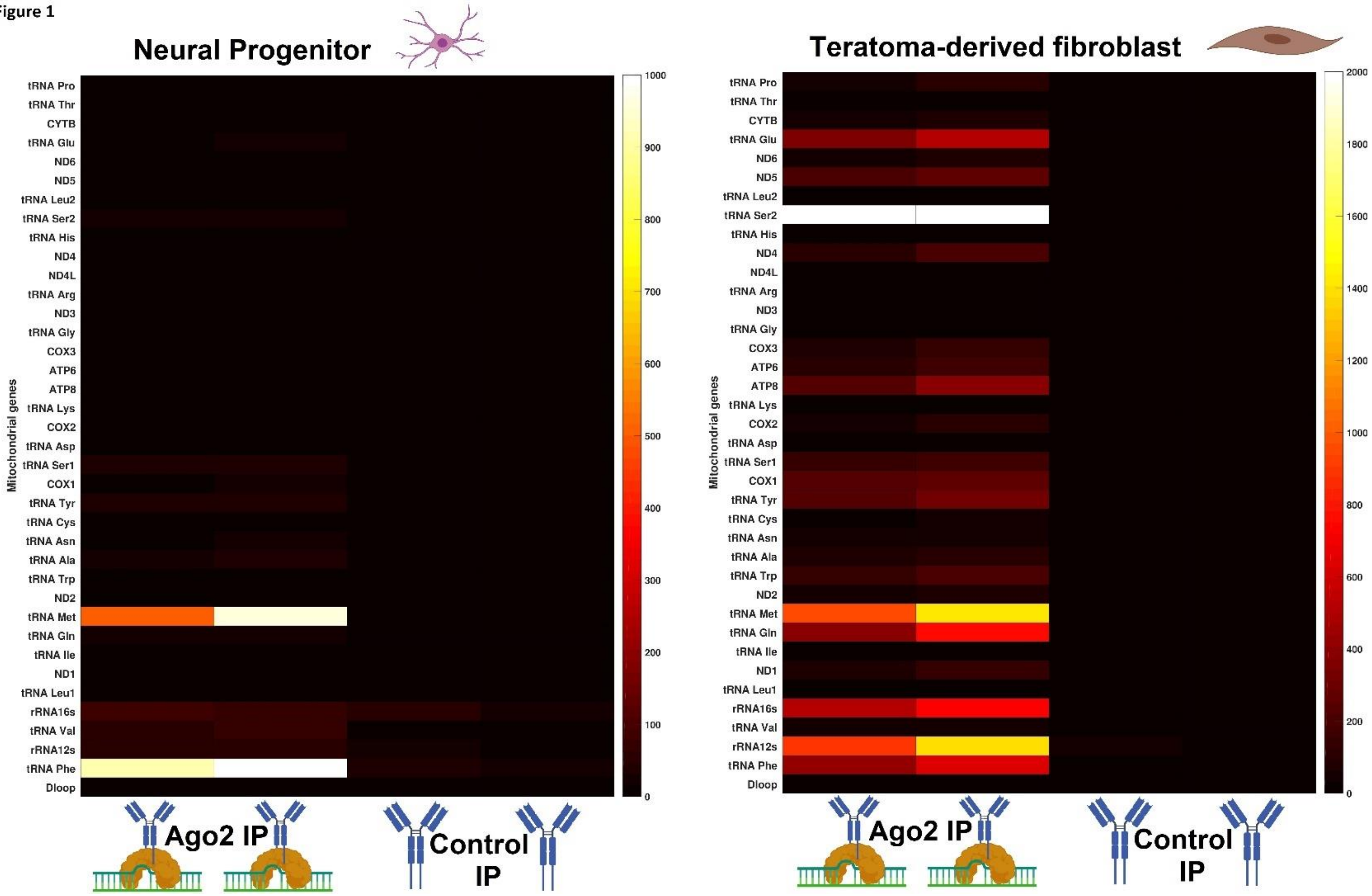


Fig.1 The mt-miRNAs bind Ago2. In two different cell lines, Ago2-IP samples are enriched in expression of specific mt-miRNAs when compared to a control IP, which contains only the IgG antibody. On the Y-axis, all the canonical mitochondrial genes are listed. On the X-axis, there are two biological replicates for each treatment. The two heatmaps are divided because they represent experiments from different cell lines. The sample type and treatments are illustrated through BioRender icons. The color scale is black-red-white, in which black and white indicate the lowest and highest levels of expression respectively. As the mt-miRNAs are an uncharacterized type of RNAs, we did not perform any normalization, thus the expression is measured in raw read counts. Gene-by-gene analyses of the transcriptional signatures are shown in Figure S2.

Figure 2

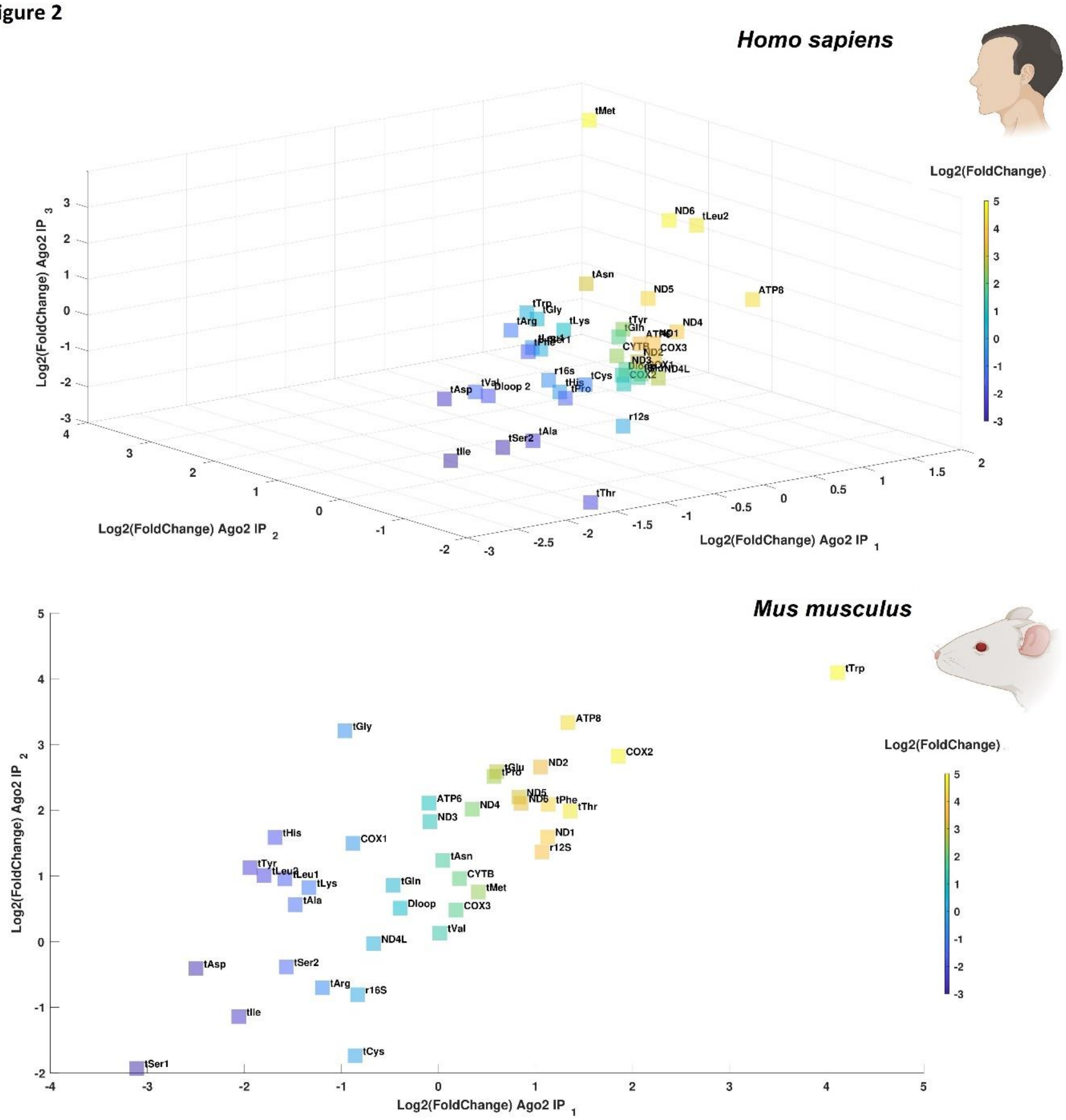


Fig.2 The mt-miRNAs are enriched in Ago2-IP samples. The plots show that Ago2-IP samples are enriched in expression for specific mt-miRNAs when compared to the same sample without Ago2-IP treatment. The enrichment is calculated though Log_2 fold change, thus a fold change of two means that the gene has four times higher expression when compared to its counterpart without Ago2-IP treatment. We used a color scale of blue-green-yellow, in which blue and yellow indicate the lowest and highest levels of expression respectively. As the mt-miRNAs are an uncharacterized type of RNAs, we did not perform any normalization, thus the expression is measured in raw read counts. The species used in each plot is indicated through small BioRender icons. **A)** The 3D-plot shows patterns of mt-miRNA expression in Ago2-IP treated samples of HeLa cells. The three axes represent three different biological replicates for each of the two treatments (IP and non-IP). **B)** The plot shows patterns of mt-miRNA expression the Ago2-IP treated samples derived from mouse embryonic stem cells. The two we represent two different biological replicates for each of the two treatments (IP and non-IP). Gene-by-gene analyses of the transcriptional signatures are shown in Figure S3 (human) and Figure S4 (mouse).

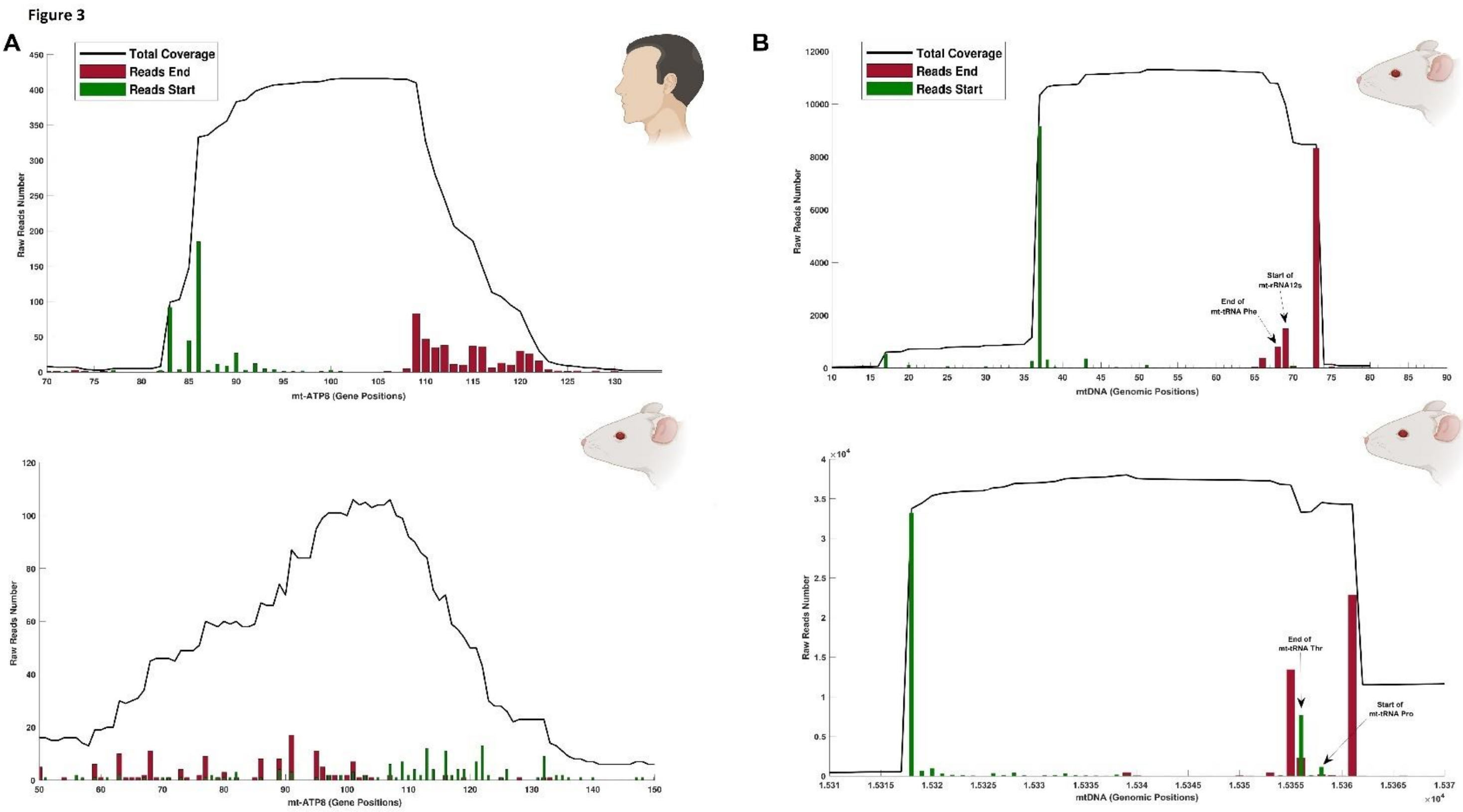


Fig.3 mt-miRNAs are encoded across genes. A) shows the coverage of small RNAs within the gene mt-ATP8 in human HeLa cell lines (upper left hand panel) and mice embryonic stem cells (lower left hand panel), and highlights the difference between a transcriptional profile showing a genuine mt-miRNA (human), and noise from the alignment of random small mitochondrial RNAs (mouse). The total coverage for each position in the gene is indicated with a black line, while the number of reads starting and ending at each position is indicated using green and red bars respectively. This type of representation provides the resolution necessary to verify the presence or absence of an mt-miRNA within a mitochondrial gene. As the mt-miRNAs are an uncharacterized type of RNAs, we did not perform any normalization, thus the expression shown on the Y-axis is measured in raw read counts. The X-axis shows the gene positions of each read. **B)** The panels represented on the right side of the figure highlight the transcriptional profile of two mt-miRNAs encoded across genes in mice embryonic stem cells. The first mt-miRNA (upper right-hand panel) is encoded mostly within the mt-tRNA Phenylalanine; however, a long isoform of this mt-miRNA includes five nucleotides of the gene mt-rRNA12s. The second mt-miRNA (lower right-hand panel) is encoded mostly within the mt-tRNA Threonine; however, a long isoform of this mt-miRNA includes six nucleotides of the gene mt-tRNA Proline. Small black arrows indicate either the end or start of mt-genes where relevant.

Figure 4

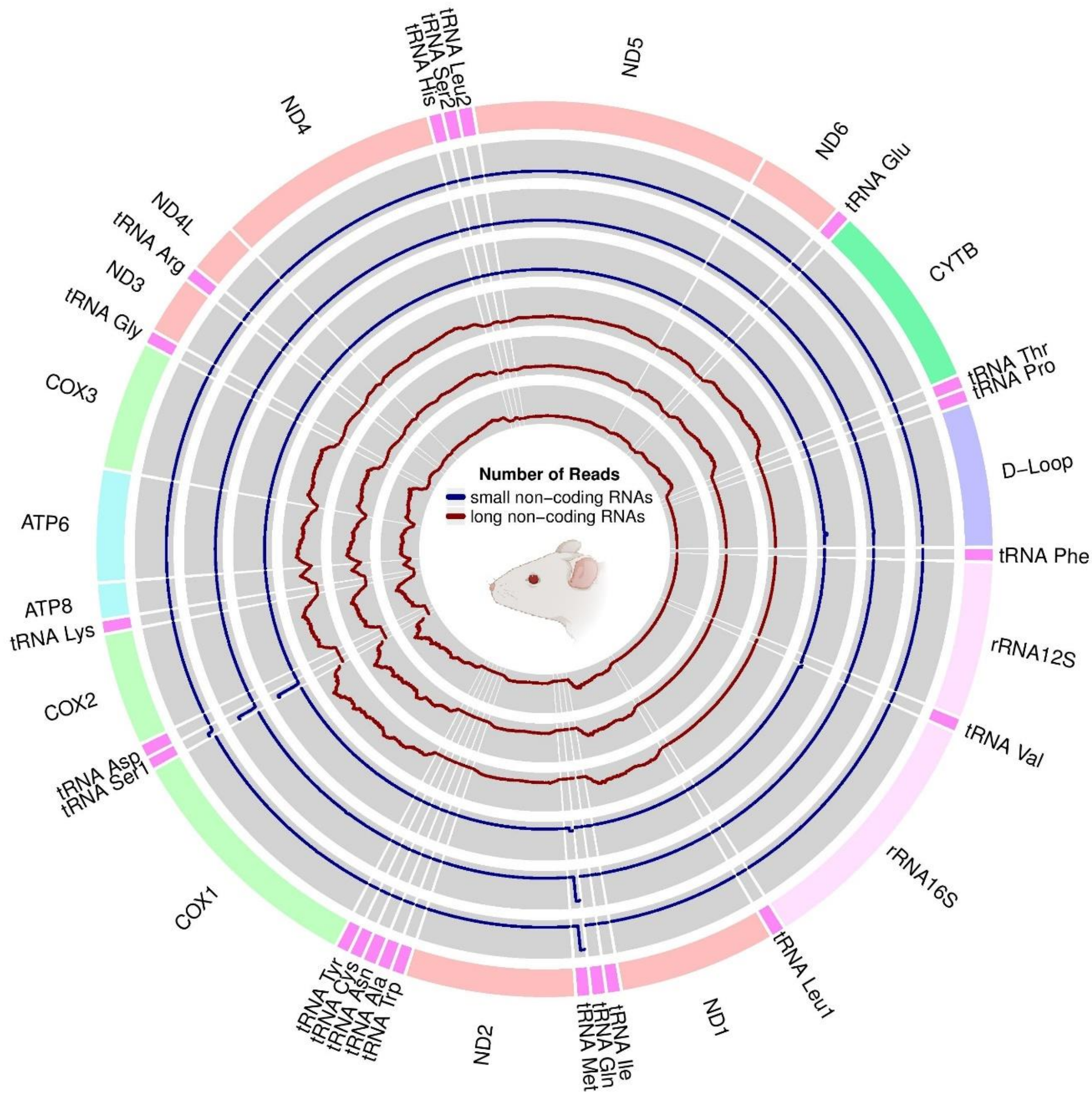


Fig.4 mt-miRNAs are not matured through pre-mt-miRNAs. The figure represents the transcriptional profile of all canonical mitochondrial genes for both small and long RNAs. The expression of small RNAs in each gene is indicated in dark blue, while the expression of long non-coding RNAs, both polyadenylated and not, is represented in dark red. As the mt-miRNAs are an uncharacterized type of RNAs, we did not perform any normalization, thus the expression is measured in raw read counts. To see more replicates, check Figure S5.

Figure 5

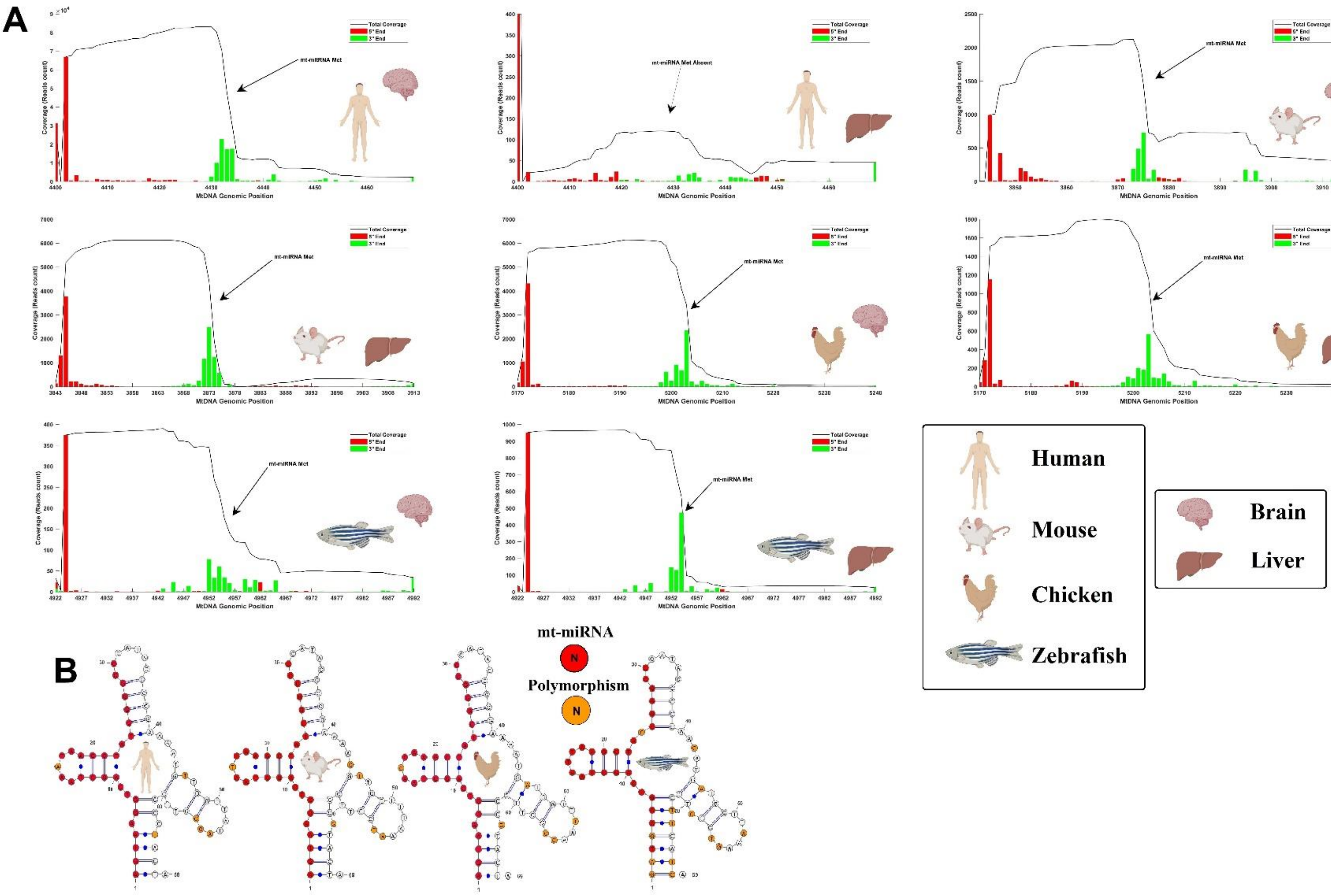


Fig.5 mt-miRNA$_{Met}$ is conserved across species. In the figure panels, the expression and sequence of the mt-miRNA$_{Met}$ are represented for several major model organisms: human (*Homo sapiens*), mouse (*Mus musculus*), chicken (*Gallus gallus*), and zebrafish (*Danio rerio*). For each species, tissues from two different organs are considered; brain and liver. The species and tissue of each sample are indicated through BioRender icons. **A**) The coverage (read count) in the mitochondrial genome, corresponding to the region of the mt-tRNA Met and mt-miRNA$_{Met}$, is depicted. With exception of human liver tissue, the mt-miRNA Met is present across all species and tissues. The Y-axis represents the expression in raw reads count, while the X-axis represents the genomic positions in the mitochondrial genome of reference for each species. **B**) We represented the mt-tRNA Met of each species while highlighting in red the sequences corresponding to the mt-miRNA$_{Met}$. We highlight in orange the presence of polymorphisms across the species, by comparing reference mitochondrial genomes of each species.

Figure 6

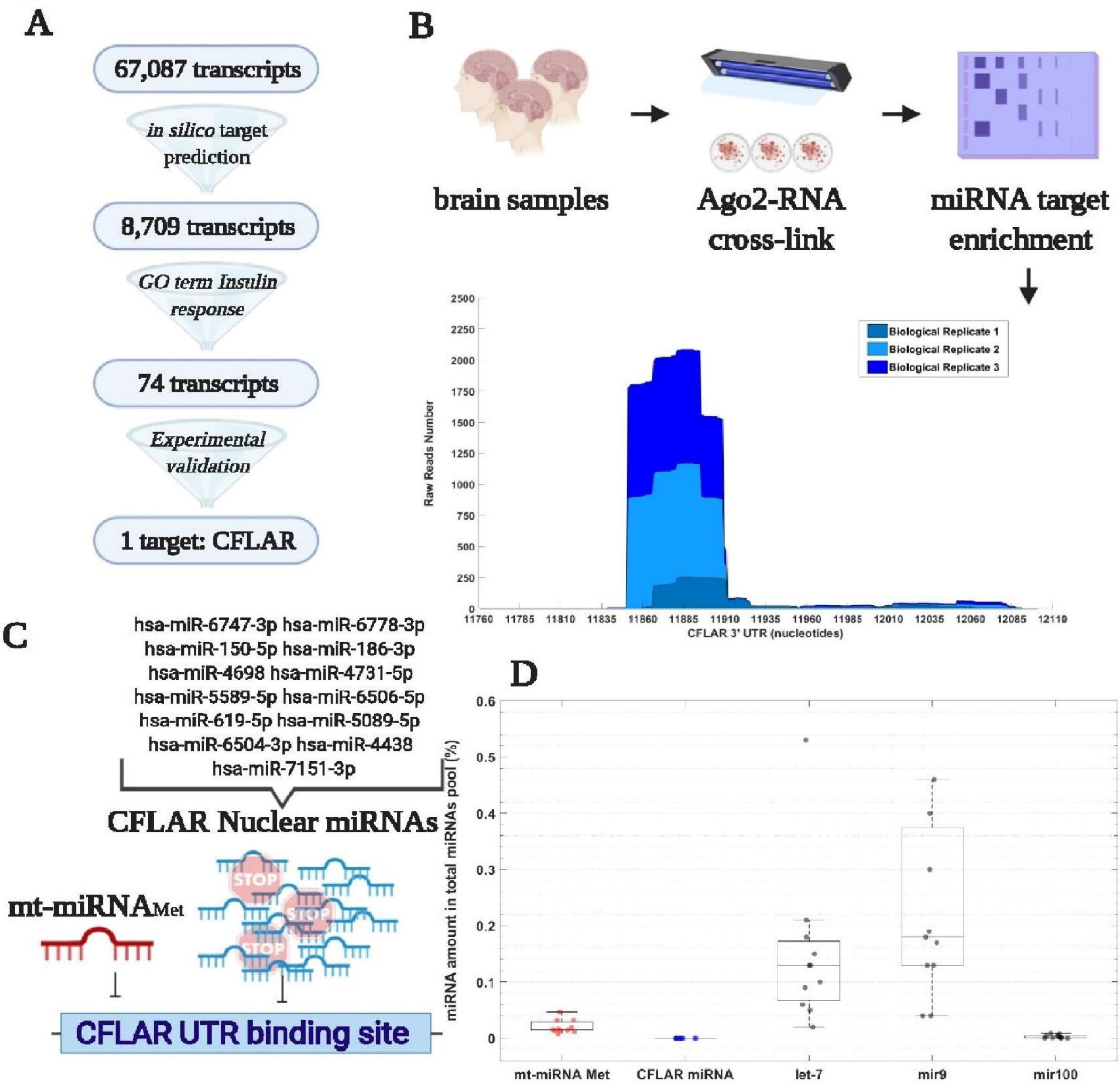


Fig.6 mt-miRNA$_{Met}$ targets the gene CFLAR in brain tissue. The figure represents the identification of an mRNA target for the mt-miRNA$_{Met.}$ **A**) A schematic representation of the target prediction used to identify the CFLAR gene. Briefly, we analysed the UTR of all human transcripts, identifying 8,709 possible targets through *in silico* prediction. Across these targets, we extracted only the ones involved in insulin response, a mechanism key in many mitochondria-related diseases. We then experimentally validated these targets by using an Ago2-IP dataset. **B**) We experimentally validated the binding between Ago2 and CFLAR by analysing a dataset composed

of three independent brain tissue samples. In this dataset Ago2 has been cross-linked to target mRNAs, and these RNAs have been isolated by size selection in gel. We show the high coverage present in the region predicted to bind mt-miRNA$_{Met}$ on the CFLAR 3' UTR across three biological replicates. The X- and Y- axes represent the positions within the 3' UTR of CFLAR and the abundance of reads. **C**) a schematic representation of the binding of mt-miRNA$_{Met}$ and the CFLAR nuclear miRNAs, a group of miRNAs predicted to bind this region of CFLAR in humans. A list of these miRNAs is present on the side of the figure, using the nomenclature present on miRBase. **D**) The boxplot represents the expression of four miRNAs, and a group of miRNAs named here CFLAR miRNAs. The expression is represented as the percentage representation of a given miRNA relative to the entire pool of miRNAs. Due to the absence of all the CFLAR miRNAs, we represented them as one group of pooled miRNAs (CFLAR miRNA). The presence of let-7, mir9, and mi100 to highlight that common miRNAs are expressed in a comparable way to mt-miRNA$_{Met}$ in the brain samples that we analysed.

Figure 7

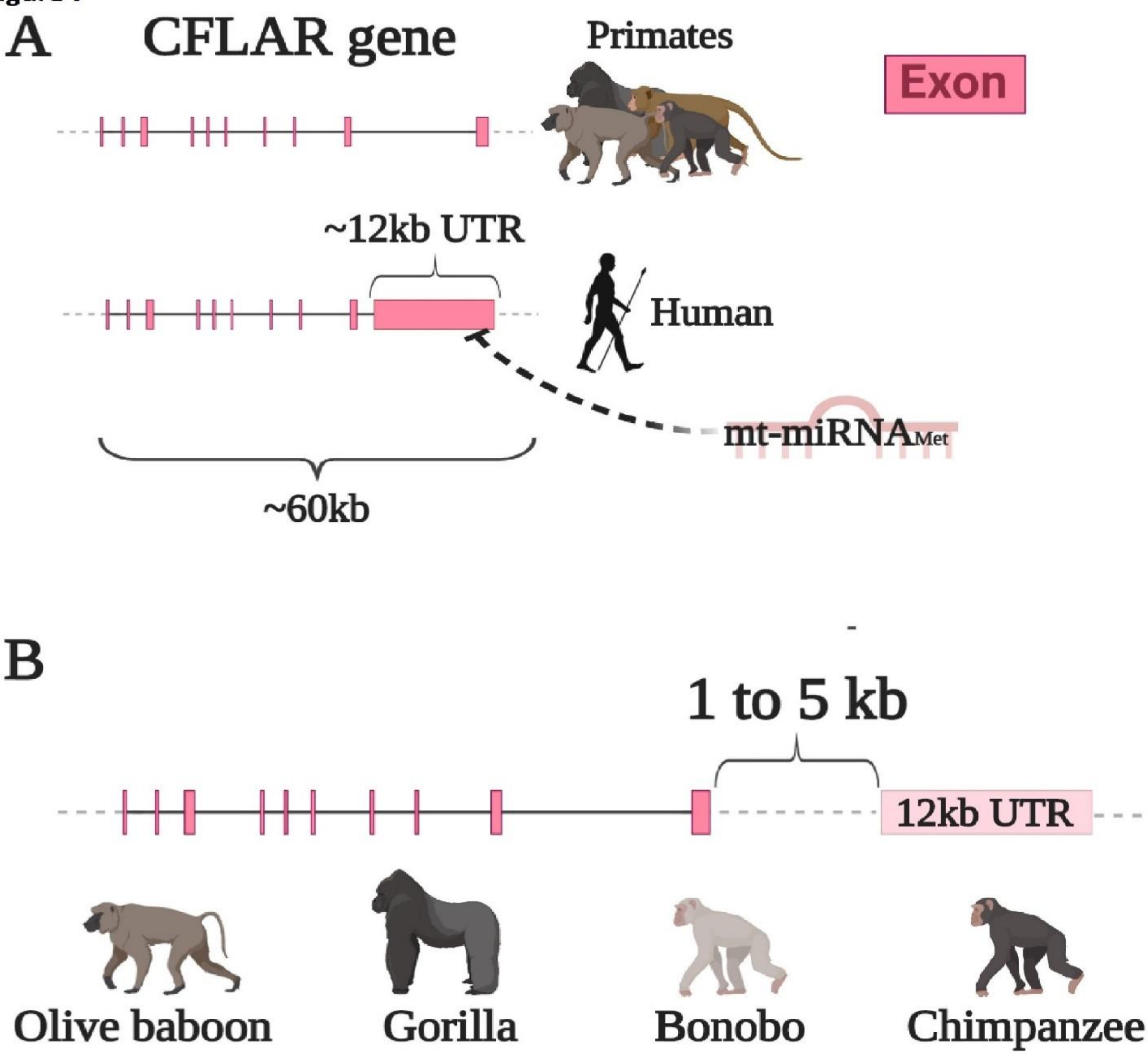


Fig.7 The mt-miRNA$_{Met}$ can regulate CFLAR in humans only. In this figure, the genomic region of CFLAR is represented across multiple species, showing its conservation across primates **A**) A schematic representation of the genomic region of the CFLAR across multiple primates. We highlight the presence of this ~12kb UTR only in humans and show that the region binding mt-miRNA$_{Met}$ is at the end of the UTR. **B**) We made a schematic representation showing that the ~12kb UTR present in humans is not translated anymore in primates but can be found at a short distance from the CFLAR gene (between 1 and 5 kb). No other miRNAs bind the CFLAR region bound by mt-miRNA$_{Met.}$ To see more details about the alignment of the genomic region of CFLAR between humans and other primates see Figure S6.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
Prefrontal cortex (BA9) from healthy individuals	Obtained from individuals in the original study	PRJNA272617
Liver samples from healthy individuals and patients suffering of chronic hepatitis C (CHC)	Obtained from individuals in the original study	PRJNA347838
Temporal lobe (TL) from patients undergone surgical temporal lobectomies	Obtained from individuals in the original study	PRJNA299324
Deposited Data		
Human brain tissues (TL)	Maragkakis M. *et al.* 2016	PRJNA299324
Human brain tissues (BA9)	Hoss AG. *et al.* 2015	PRJNA272617
Human liver tissues (CHC)	Butt AM. *et al.* 2016	PRJNA347838
HeLa cell culture	Zhang K. *et al.* 2018	PRJNA473925
Neuronal progenitor cell culture	Young-Soo Kwon	PRJNA438936
teratoma-derived fibroblast cell culture	Young-Soo Kwon	PRJNA438936
Mouse brain tissues	Viljetic B. *et al.* 2017	PRJNA309689
Mouse liver tissues	Jee D. *et al.* 2018	PRJNA395699
Mouse embryonic stem cell culture (mESC)	Zamudio JR. *et al.* 2014	PRJNA218026
Mouse kidney tissues	Woo YM. *et al.*2017	PRJNA342099
Chicken brain and liver tissues	Warnefors M. *et al.* 2017	PRJNA396511
Zebrafish brain and liver tissues	Vaz C. *et al.* 2015	PRJNA245824
Experimental Models: Cell Lines		
HeLa cell culture	Xiang-Dong Fu	GSM3168220
Neuronal progenitor cell culture	Young-Soo Kwon	GSM3052811
teratoma-derived fibroblast cell culture	Young-Soo Kwon	GSM3052815
Mouse embryonic stem cell culture (mESC)	Jesse R Zamudio	GSM1224442
Experimental Models: Organisms/Strains		
Mouse Brain tissues from mix of strains 129 and C57BL/6	The Jackson laboratory	No:002448 No:000664
Mouse liver tissues from strain B6.Cg-Tg(Vav1-icre)A2Kio/J vav1-cre	The Jackson laboratory	No:008610
kidney tissues from mice with floxed Pkd1	Stefan Somlo	NA
Chicken brain and liver tissues	Peter Jensen	NA
Zebrafish brain and liver tissues from Singapore strain	Obtained from individuals in the original study	NA
Software and Algorithms		
BowTie2	Langmead and Salzberg 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Bedtool	Quinlan and Hall 2010	https://bedtools.readthedocs.io/en/latest/
Samtool	Li H. *et al.* 2009	http://samtools.sourceforge.net/
R package *circlize*	Gu Z. *et al.* 2014	https://jokergoo.github.io/circlize_book/
MATLAB	The MathWorks, Inc.	https://au.mathworks.com/
MR-micro T	Maragkakis M. *et al.* 2011	http://diana.imis.athena-innovation.gr/DianaTools/
LastZ	Harris RS. 2007	https://help.rc.ufl.edu/doc/LASTZ

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and request about the reagents can be directed to the Lead Contact, Andrea Pozzi (andreapozzi.a@gmail.com). However, as the data are sourced from previously published articles, we can only redirect you to the appropriate source, and we cannot personally provide full details of each reagent.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

This study is based on published data deposited in the Sequence Read Archive (SRA) in NCBI, thus all information listed here are sourced from the original study where the data were obtained.

Cell lines

According to the original authors (Zhang et al., 2018), HeLa cells were grown on 15 cm plates using MEDM plus 10% FBS. The neuronal progenitor (NP) and teratoma-derived fibroblast (TDF) cell culture, were derived from human embryonic stem cells (hESC). The differentiation of hESC into NP was induced by replacing the original growing medium with DMEM/F12 supplemented with 2% B27, 100ng/ml FGF, 100ng/ml EGF and 5ng/ml heparin. Further details on the methods can be found at GSE115146 in the Gene Expression Omnibus (GEO) database and in the original article.

The differentiation of hESC into TDF was induced by injection of resuspended cells into mice homozygous for severe combined immune deficiency spontaneous mutation (SCID). The tumors grew in after 6 weeks were then removed and cultured in a medium of 10% FBS, nonessential amino acids, 2mM glutamine, 1% penicillin/streptomycin and 0.55μM β-mercaptoethanol. Further details on the methods, as described by the original authors, can be found at GSE112006 in the Gene Expression Omnibus (GEO) database.

According to the original authors (Zamudio et al., 2014), the embryonic stem cell culture (mESC) were grown on gelatinized tissue culture plates in Dulbecco's Modified Essential Media supplemented with multiple other nutrients. The full list of nutrients and details about the growing method can be found at GSE50595 in the GEO database and in the original article.

Model organisms

According to the original authors (Jee et al., 2018; Viljetic et al., 2017), the experiments on mice sourced from the Jackson laboratory were carried out in compliance with their institutional protocols, thus we believe that correct ethics for the experiments have been followed.

According to the original authors (Woo et al., 2017), the experiments on mice sourced by Stefan Somlo were carried out in compliance with the Animal Care and Use Committee (IACUC) rules at Sookmyung Women's University. Therefore, we believe that correct ethics were followed during the experiment.

According to the original authors, the samples from the chickens were sourced by Peter Jensen, however, there is not any further details about how the chickens were housed and maintained.

According to the original authors (Vaz et al., 2015), the zebrafish (Singapore strain) were maintained according to the Animal Care and Use Committee (IACUC) rules. Therefore, we believe that correct ethics were followed during the experiment.

Human samples

The BA9 samples were obtained by the original authors (Hoss et al., 2015), and the gathering of the samples were exempt from ethics approval because the study involves only tissue collected post-mortem, and consequently not classified as human subjects. This decision was taken by the Boston University School of Medicine Institutional Review Board (Protocol H-28974).

The CHC samples were obtained by the original authors (Butt et al., 2016), and the original authors declared that written informed consent for the use of biological samples and clinical records was given by all the participants. Furthermore, they declared that their work was done in accordance with the ethical guidelines of the 1975 Declaration of Helsinki and the International Conference on Harmonization Guidelines for Good Clinical Practice.

The TL samples were obtained by the original authors (Maragkakis et al., 2016; Nakaya et al., 2013), and the authors provided only a brief description of the samples. Indeed, they only mention that the samples come from temporal surgical lobectomy of three unrelated individuals. Although no statements on ethics is present in the original articles, we believe that is likely that the correct ethics approval processes are followed in the institution where the experiments where performed (University of Pennsylvania).

METHODS DETAILS

Ago2 co-immunoprecipitation sequencing

According to the original authors (Zhang et al., 2018), before the Ago2 immunoprecipitation (IP) of HeLa cells, the cultures were UV irradiated at 400mj. The IP was performed using Anti-Ago2 (Abnova, H00027161-M01), and following the IP the RNA libraries were made and sequenced using Hi-Seq 2500. Further details can be found in the original article and in the GEO database (GSE115146).

According to the original authors, before the Ago2-IP of the NP and TDF, the cultures were UV irradiated at 400mj. The IP was performed using Anti-Ago2 (ProteinTech) and following the IP the RNA libraries were made and sequenced using Hi-Seq 2000. Further details can be found in the GEO database (GSE112006).

According to the original authors (Zamudio et al., 2014), the mESC cells have been modified to leave only a modified Ago2 active in these cells. This modified Ago2 gene is known as FLAG-hemagglutinin (HA)-tagged hAgo2 (FHAgo2), thus has a specific epitope that can be targeted for IP. Furthermore, contrary to the experiments in other cell lines, in this case the cell cultures were lysed before the UV-cross linking and IP**.** Further details can be found in the original article and in the GEO database (GSE50595).

The Ago2-IP of the TL samples were performed using 2A8 Anti-Ago monoclonal antibody tethered to Protein A Dynabeads (Invitrogen), according to the original authors (Maragkakis et al., 2016). Further details can be found in the original article and in the BioProject NCBI database (PRJNA299324).

Ago2 immunoprecipitation miRNA targets enrichment

The original authors of the study performed the mRNA target enrichment (Maragkakis et al., 2016), and we describe here a short summary of their method. After the Ago2-IP of the three brain samples, the authors isolated longer RNAs (miRNA targets) using nitrocellulose filter, and 8% PAGE gels. Once the targets were purified, they have been re-amplified by PCR and sequenced. Further information about the methods can be found in previous works of the authors (Maragkakis et al., 2016; Nakaya et al., 2013).

Mt-miRNAs sequence alignment

The alignment of the RNA-seq libraries, both with and without IP, were performed using BowTie2 (Langmead & Salzberg, 2013). We did not remove the adapter for each library, and instead used the soft clipping provided by bowtie2 through the setting *--local*. Due to the possible differences between the reference genome used, and the sequence of the individual used for analysis, we used non-stringent parameters in the alignment. Thus, for the alignment we set a seed of 20 nucleotides (-*L 20*) and we allowed up to 1 mismatch (-*N 1*) between the small RNA and the reference genome. To gauge the overall amount of small RNAs within one gene, we used samtools *idxstats* function (Li et al., 2009), which outputs the reads aligning to each chromosome. Likewise, to obtain the coverage for each nucleotide, and thus identify the mitochondrial genes having the transcriptional signature of a mt-miRNA, we used the function *genomecov* of bedtools (Quinlan & Hall, 2010). To generate the total coverage for each gene, we used the setting *genomecov -d -i* while for the 5' and 3' coverage we added the setting *-5* and *-3* respectively. To generate the correct format to plot the gene coverage in the R package *circlize*, we used the parameter *-bg* instead of *-d* on bedtools *genomecov*. The alignment was done using the current reference genomes for human (NC_012920.1) and mouse (NC_005089.1).

Mt-miRNA$_{Met}$ polymorphism identification

The identification of polymorphism on the mt-miRNA$_{Met}$ across multiple species was performed by visual comparison of the sequences downloaded from the mt-tRNAs database mitotRNAdb (http://mttrna.bioinf.uni-leipzig.de/mtDataOutput/Welcome). We only used the reference sequences to establish the presence of polymorphisms, and we did not include any population data, because the quality of population data (i.e. the frequency of specific polymorphisms) are very different across the organisms considered. Therefore, we used the reference genome present in mitotRNAdb for *Danio rerio* (NC_002333.2), *Gallus gallus* (NC_001323.1), *Homo sapiens* (NC_012920.1), and *Mus musculus* (NC_005089.1).

Mt-miRNA$_{Met}$ target prediction

The *in silico* prediction of the mt-miRNA$_{Met}$ target was done using the web-server MR-microT (Kanellos et al., 2014; Reczko et al., 2012), a program able to use a custom miRNA sequence to identify potential target mRNAs in the human genome (using the database Ensembl v84). This program outputs both a score, indicating how likely is the predicted target to be real, and the regions of the sequence where miRNA and mRNAs are predicted to interact. As mentioned in the results, the list of potential targets predicted by this program is 8,709, which we filtered using the gene ontology (GO) term Insulin Response (GO:0032869) to obtain a list of 74 potential targets. To experimentally verify these targets, we extracted the coverage of each gene using bedtools (Quinlan & Hall, 2010) with settings *genomecov -d -i* , and then inspected by eye for the presence of sequences in the regions corresponding to the prediction of MR-micro T. Although this method might be missing several targets of this gene (false negative), we focused on finding targets that have the highest chance of being real (true positive). It is worth noting that the mRNA target identified using this method, CFLAR, has a MR-micro T score with high likelihood to be predictive (>0.8) according to the authors of the program (Kanellos et al., 2014; Reczko et al., 2012).

Comparative analysis mt-miRNA$_{Met}$ and nuclear miRNAs

The miRNAs were quantified using the output of samtools *flagstat* option, which provides the number of aligned reads, both the total amount and as a percentage of the RNA pool, for the specified reference sequence. We used mt-tRNA Met as the reference sequence to quantify the amount of mt-miRNA Met reads expressed, indeed, over 90% of the reads come from the region of the mt-miRNA$_{Met}$, thus using the coverage of the entire mt-tRNA Met provides a very good approximation of mt-miRNA$_{Met}$ without excluding any of its isoforms. To identify miRNAs that might be binding the same regions as mt-miRNA$_{Met}$, we used TargetscanHuman 7.2 (http://www.targetscan.org/) using CFLAR as a query, and then we annotated all the miRNAs predicted to bind the same, or similar, positions of mt-miRNA$_{Met}$. Using this approach, we annotated 13 miRNAs: hsa-miR-6747-3p, hsa-miR-6778-3p, hsa-miR-150-5p, hsa-miR-186-3p, hsa-miR-4698, hsa-miR-4731-5p, hsa-miR-5589-5p, hsa-miR-6506-5p, hsa-miR-619-5p, hsa-miR-5089-5p, hsa-miR-6504-3p, hsa-miR-4438, hsa-miR-5095, hsa-miR-7151-3p. To quantify the nuclear miRNAs, we used the sequences present in miRBase (http://www.mirbase.org/) as a reference sequence for *samtools*. However, none of these nuclear miRNAs was found in brain tissues. To verify the absence of artifacts, or errors in our analysis, we used the same method to quantify common, well-known miRNAs. The miRNAs we chose are hsa-Let7a-1, hsa-miR-9-5p, hsa-miR-100. The similar abundance of these miRNAs to the mt-miRNA$_{Met}$ confirmed the likely absence of artifacts or errors in the analysis, thus confirming the absence of other known miRNAs binding the same position of mt-miRNA$_{Met}$.

CFLAR Untranslated region analysis

We verified the presence of the 3' untranslated region (UTR) of 20kb length found in *Homo sapiens* across multiple species by searching on the online version of blastn using default settings (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch). After not finding any hit on blastn, we tested the presence of this UTR across several species in the Gene database of NCBI (https://www.ncbi.nlm.nih.gov/gene). Due to the lack of this UTR across any other species, we focused more detailed analysis on species closer to *Homo sapiens* as they are more likely to have some trace of this sequence. Thus, we selected five other high-quality genomes from different primates and compared the full genomic region of CFLAR. The species involved are *Homo sapiens* (ID: 8837), *Pan troglodytes* (ID: 459872), *Pan paniscus* (ID: 100972044), *Pongo abelii* (ID: 100172025), *Gorilla gorilla* (ID: 101130329), and *Papio anubis* (ID: 101006231). The alignment of the different regions was made using LastZ (http://www.bx.psu.edu/~rsharris/lastz/), with settings *--step=10 --nogapped --ambiguous=iupac --matchcount=20 --format=rdotplot* , which provided a matrix of the alignment that was then used for the dotplot in R.

SUPPLEMENTARY REFERENCES

Butt, A. M., Raja, A. J., Siddique, S., Khan, J. S., Shahid, M., Tayyab, G.-U.-N., Minhas, Z., Umar, M., Idrees, M., & Tong, Y. (2016). Parallel expression profiling of hepatic and serum microRNA-122 associated with clinical features and treatment responses in chronic hepatitis C patients. *Scientific Reports*, *6*, 21510. https://doi.org/10.1038/srep21510

Hoss, A. G., Labadorf, A., Latourelle, J. C., Kartha, V. K., Hadzi, T. C., Gusella, J. F., MacDonald, M. E., Chen, J.-F., Akbarian, S., Weng, Z., Vonsattel, J. P., & Myers, R. H. (2015). miR-10b-5p expression in Huntington's disease brain relates to age of onset and the extent of striatal involvement. *BMC Medical Genomics*, *8*, 10. https://doi.org/10.1186/s12920-015-0083-3

Jee, D., Yang, J.-S., Park, S.-M., Farmer, D. T., Wen, J., Chou, T., Chow, A., McManus, M. T., Kharas, M. G., & Lai, E. C. (2018). Dual Strategies for Argonaute2-Mediated Biogenesis of Erythroid miRNAs Underlie Conserved Requirements for Slicing in Mammals. *Molecular Cell*, *69*(2), 265–278.e6. https://doi.org/10.1016/j.molcel.2017.12.027

Kanellos, I., Vergoulis, T., Sacharidis, D., Dalamagas, T., Hatzigeorgiou, A., Sartzetakis, S., & Sellis, T. (2014). MR-microT: a MapReduce-based MicroRNA target prediction method. *Proceedings of the 26th International Conference on Scientific and Statistical Database Management*, 47. https://dl.acm.org/citation.cfm?id=2618289

Langmead, B., & Salzberg, S. L. (2013). Langmead. 2013. Bowtie2. *Nature Methods*, *9*, 357–359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* , *25*(16), 2078–2079.

https://doi.org/10.1093/bioinformatics/btp352
Maragkakis, M., Alexiou, P., Nakaya, T., & Mourelatos, Z. (2016). CLIPSeqTools--a novel bioinformatics CLIP-seq analysis suite. *RNA* , *22*(1), 1–9. https://doi.org/10.1261/rna.052167.115

Nakaya, T., Alexiou, P., Maragkakis, M., Chang, A., & Mourelatos, Z. (2013). FUS regulates genes coding for RNA-binding proteins in neurons by binding to their highly conserved introns. *RNA* , *19*(4), 498–509. https://doi.org/10.1261/rna.037804.112

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* , *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Reczko, M., Maragkakis, M., Alexiou, P., Grosse, I., & Hatzigeorgiou, A. G. (2012). Functional microRNA targets in protein coding sequences. *Bioinformatics* , *28*(6), 771–776. https://doi.org/10.1093/bioinformatics/bts043

Vaz, C., Wee, C. W., Lee, G. P. S., Ingham, P. W., Tanavde, V., & Mathavan, S. (2015). Deep sequencing of small RNA facilitates tissue and sex associated microRNA discovery in zebrafish. *BMC Genomics*, *16*, 950. https://doi.org/10.1186/s12864-015-2135-7

Viljetic, B., Diao, L., Liu, J., Krsnik, Z., Wijeratne, S. H. R., Kristopovich, R., Dutre-Clarke, M., Kraushar, M. L., Song, J., Xing, J., Chen, K. C., & Rasin, M.-R. (2017). Multiple roles of PIWIL1 in mouse neocorticogenesis. In *bioRxiv* (p. 106070). https://doi.org/10.1101/106070

Woo, Y. M., Kim, D. Y., Koo, N. J., Kim, Y.-M., Lee, S., Ko, J. Y., Shin, Y., Kim, B. H., Mun, H., Choi, S., Lee, E. J., Shin, J.-O., Park, E. Y., Bok, J., & Park, J. H. (2017). Profiling of miRNAs and target genes related to cystogenesis in ADPKD mouse models. *Scientific Reports*, *7*(1), 14151. https://doi.org/10.1038/s41598-017-14083-8

Zamudio, J. R., Kelly, T. J., & Sharp, P. A. (2014). Argonaute-bound small RNAs from promoter-proximal RNA polymerase II. *Cell*, *156*(5), 920–934. https://doi.org/10.1016/j.cell.2014.01.041

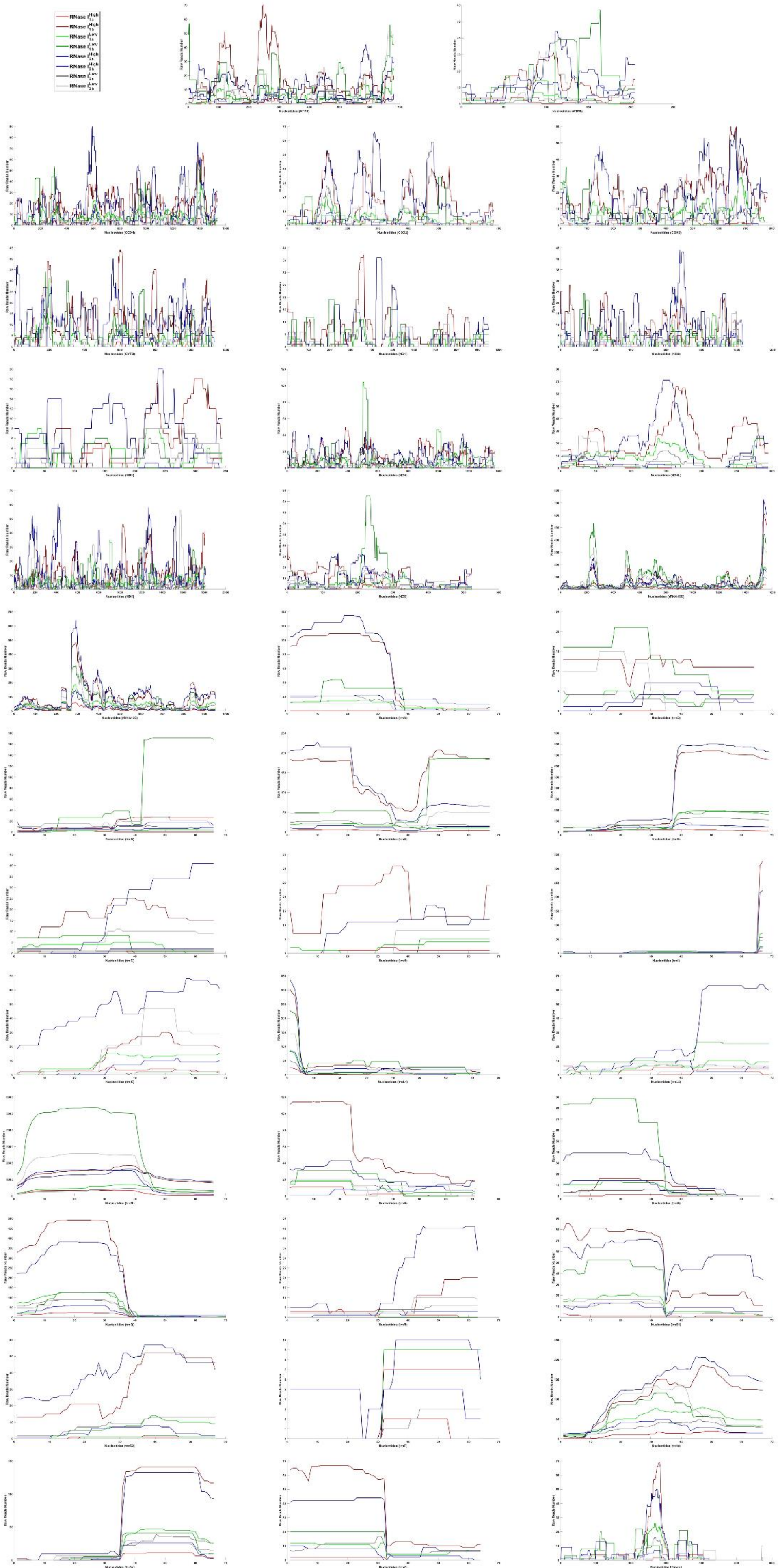
Zhang, K., Zhang, X., Cai, Z., Zhou, J., Cao, R., Zhao, Y., Chen, Z., Wang, D., Ruan, W., Zhao, Q., Liu, G., Xue, Y., Qin, Y., Zhou, B., Wu, L., Nilsen, T., Zhou, Y., & Fu, X.-D. (2018). A novel class of microRNA-recognition elements that function only within open reading frames. *Nature Structural & Molecular Biology*, *25*(11), 1019–1027. https://doi.org/10.1038/s41594-018-0136-3

SUPPLEMENTARY FIGURES

The layout of the figures is according to the CellPress policy and the figures resolution is low to avoid having a thesis of over 100 Mbs. Zoom in to see the pattern in a single plot.

The **figure S1** highlights the lack of effect from different concentrations of RNase I, across multiple mitochondrial genes.

In the following figure, we show the transcriptional signatures of putative small RNAs across all mitochondrial genes, gene by gene. The gene order is protein-coding, rRNAs, tRNAs, and within each group, the genes are in alphabetic order. In the following paragraph, we highlight the genes encoding mt-miRNAs, using the same prerequisite criteria as in the manuscript (i.e. transcriptional signatures with clear cut ends across multiple biological replicates indicate the presence of an mt-miRNA). Among the protein-coding genes, none of the genes has mt-miRNAs fitting the rationale for mt-miRNAs identification. Among the genes encoding for the two mt-rRNAs, only the mt-rRNA12s have an mt-miRNAs encoded at position 250 which is consistently transcribed across multiple biological replicates and have well-defined start and end. Among the mt-tRNA genes, only mt-tRNA Phe (trnF), mt-tRNA Met (trnM), mt-tRNA Gln (Q), and mt-tRNA Trp (trnW) are consistently transcribed across multiple biological replicates and have well-defined start and end.



RNase I High 1a
RNase I High 1b
RNase I Low 1a
RNase I Low 1b
RNase I High 2a
RNase I High 2b
RNase I Low 2a
RNase I Low 2b

F**igure S2** highlights the enrichment of Ago2-IP samples compared to IgG-IP samples, across multiple mitochondrial genes.

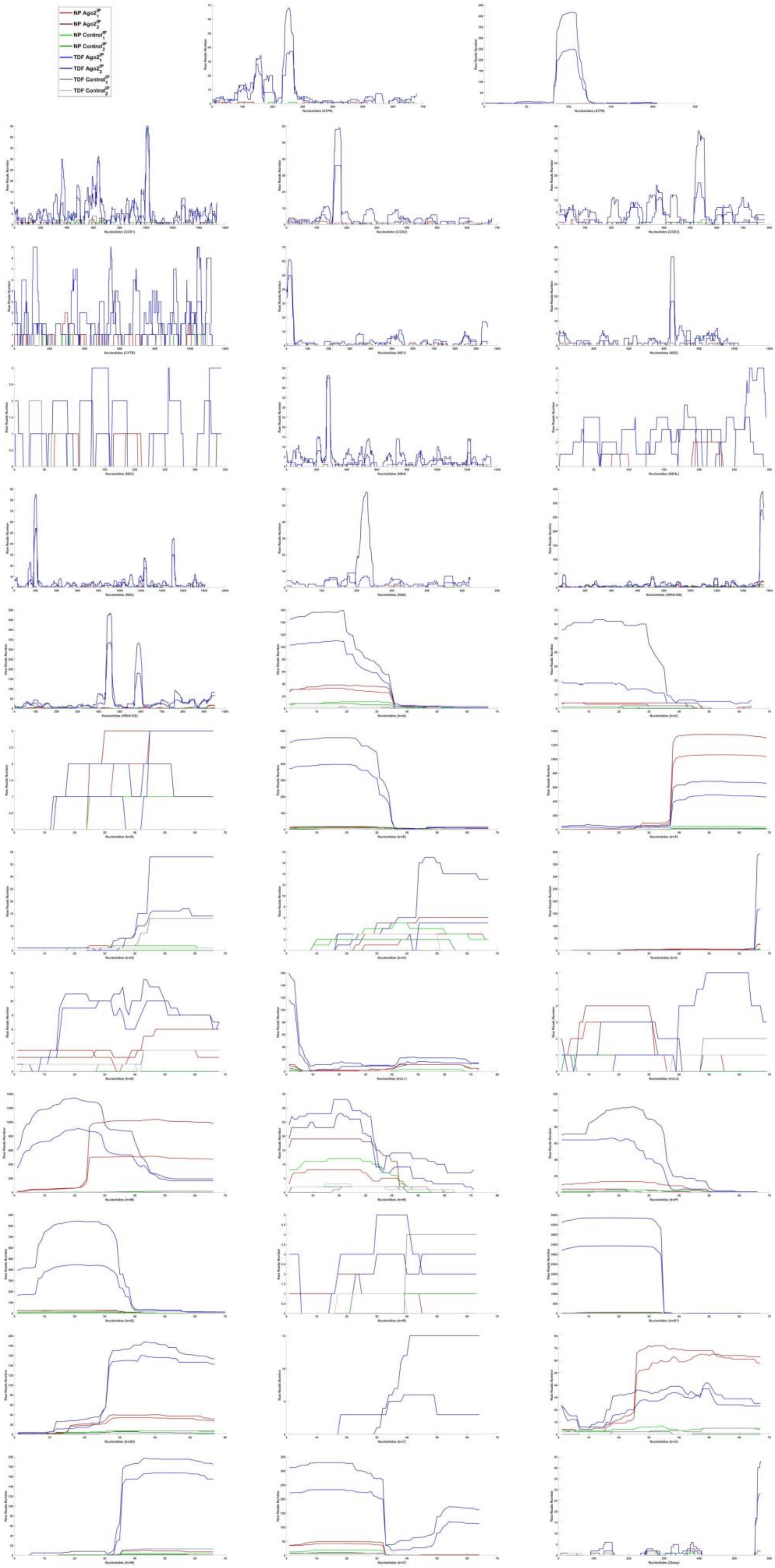
In the following figure, we show the transcriptional signature of small RNAs across all mitochondrial genes, divided by gene for neuronal progenitors (NP) and teratoma-derived fibroblast (TDF). The gene order is protein-coding, rRNAs, tRNAs, and within each group, the genes are in alphabetic order. In the following paragraph, we highlight the genes encoding for mt-miRNAs, following the same prerequisites as outlined in the manuscript (that transcriptional signatures with clear cut ends across multiple biological replicates indicate the presence of an mt-miRNA).

In neuronal progenitors (NP)

Among the protein-coding genes, none of the genes has mt-miRNAs fitting the prerequisites for mt-miRNA identification. Among the genes encoding for the two mt-rRNAs, none of the genes has an mt-miRNA consistently transcribed across multiple biological replicates and a well-defined start and end. Among the mt-tRNA genes, only mt-tRNA Phe (trnF), mt-tRNA Met (trnM), mt-tRNA Ser 2 (trnS2), and mt-tRNA Tyr (trnY) have mt-miRNAs consistently transcribed across multiple biological replicates and having well-defined start and end.

In teratoma-derived fibroblast (TDF)

Among the protein-coding genes, only ATP8 encodes for an mt-miRNAs fitting the rationale for mt-miRNAs identification. Among the genes encoding for the two mt-rRNAs, none of the genes has an mt-miRNA consistently transcribed across multiple biological replicates and has a well-defined start and end. The peaks are due to the overlap of many different small RNAs partially overlapping. Among the mt-tRNA genes, only mt-tRNA Glu (trnE), mt-tRNA Phe (trnF), mt-tRNA Met (trnM), mt-tRNA Gln (trQ), mt-tRNA Ser 1 (trnS1), mt-tRNA Ser 2 (trnS2), mt-tRNA Trp (trnW) and mt-tRNA Tyr (trnY) have mt-miRNAs consistently transcribed across multiple biological replicates and having well-defined start and end.



NP Ago2$_1^{IP}$
NP Ago2$_2^{IP}$
NP Control$_1^{IP}$
NP Control$_2^{IP}$
TDF Ago2$_1^{IP}$
TDF Ago2$_2^{IP}$
TDF Control$_1^{IP}$
TDF Control$_2^{IP}$

Figure S3 highlights the enrichment in mt-miRNAs of specific genes after Ago2-IP treatment, across multiple mitochondrial genes.

In the following figure, we show the small RNAs transcriptional signature of all the mitochondrial genes, divided by gene. The gene order is protein-coding, rRNAs, tRNAs, and within each group, the genes are in alphabetic order. In the following paragraph, we highlight the genes encoding for mt-miRNAs, following the prerequisite condition that transcriptional signatures with clear cut ends across multiple biological replicates indicate the presence of an mt-miRNA. The samples are paired, thus the samples of each treatment have to be compared to their corresponding number. For example, $Input_1$ should be compared to $Ago2^{IP}{}_1$ and not to $Ago2^{IP}{}_2$.

Among the protein-coding genes, none of the genes has mt-miRNAs fitting our criteria for mt-miRNA identification. Among the genes encoding for the two mt-rRNAs, none of the genes has an mt-miRNA consistently transcribed across multiple biological replicates and a well-defined start and end. Among the mt-tRNA genes, only mt-tRNA Ile (trnI), mt-tRNA Glu (trnQ), and mt-tRNA Ser1 (trnS1) are consistently transcribed across multiple biological replicates and have well-defined start and end.

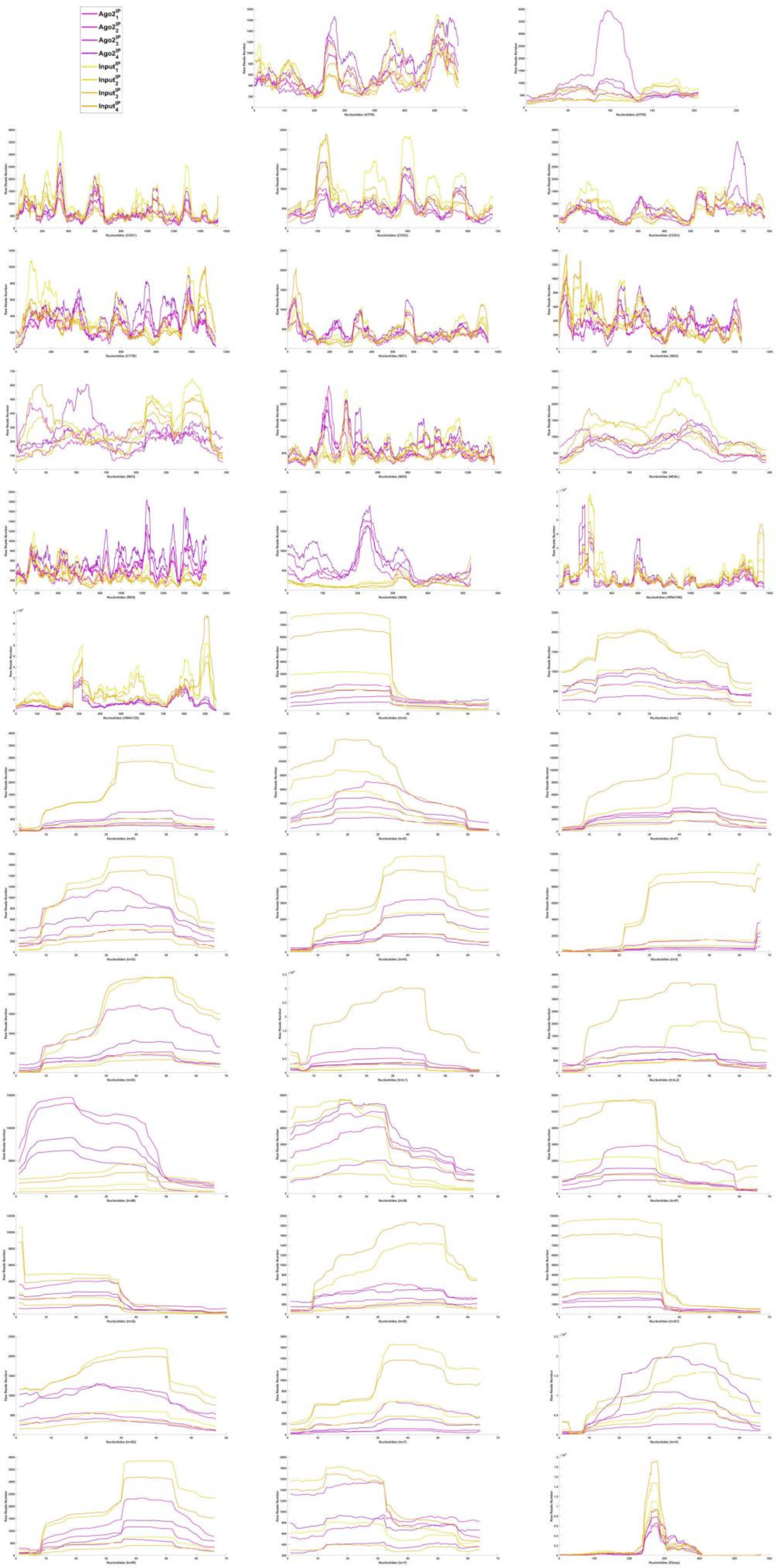


Figure S4 highlights the enrichment in mt-miRNAs of specific genes after Ago2-IP treatment, across multiple mitochondrial genes.

In the following figure, we show the small RNAs transcriptional signature of all the mitochondrial genes, divided by gene. The gene order is protein-coding, rRNAs, tRNAs, and within each group, the genes are in alphabetic order. In the following paragraph, we highlight the genes encoding for mt-miRNAs, following the prerequisite condition that transcriptional signatures with clear cut ends across multiple biological replicates indicate the presence of an mt-miRNA. The samples are paired, thus the samples of each treatment have to be compared to their corresponding number. For example, $InputIP_1$ should be compared to $Ago2^{IP}{}_1$ and not to $Ago2^{IP}{}_2$.

Among the protein-coding genes, only COX1, COX3, and ND4L have mt-miRNAs our criteria for mt-miRNA identification. Among the genes encoding for the two mt-rRNAs, none of the genes has an mt-miRNA consistently transcribed across multiple biological replicates and a well-defined start and end. Among the mt-tRNA genes, only mt-tRNA Glu (trnE), mt-tRNA Phe (trnF), mt-tRNA Leu2 (trnL2), mt-tRNA Met (trnM), and mt-tRNA Thr (trnT) have mt-miRNAs consistently transcribed across multiple biological replicates and showing well-defined start and end.

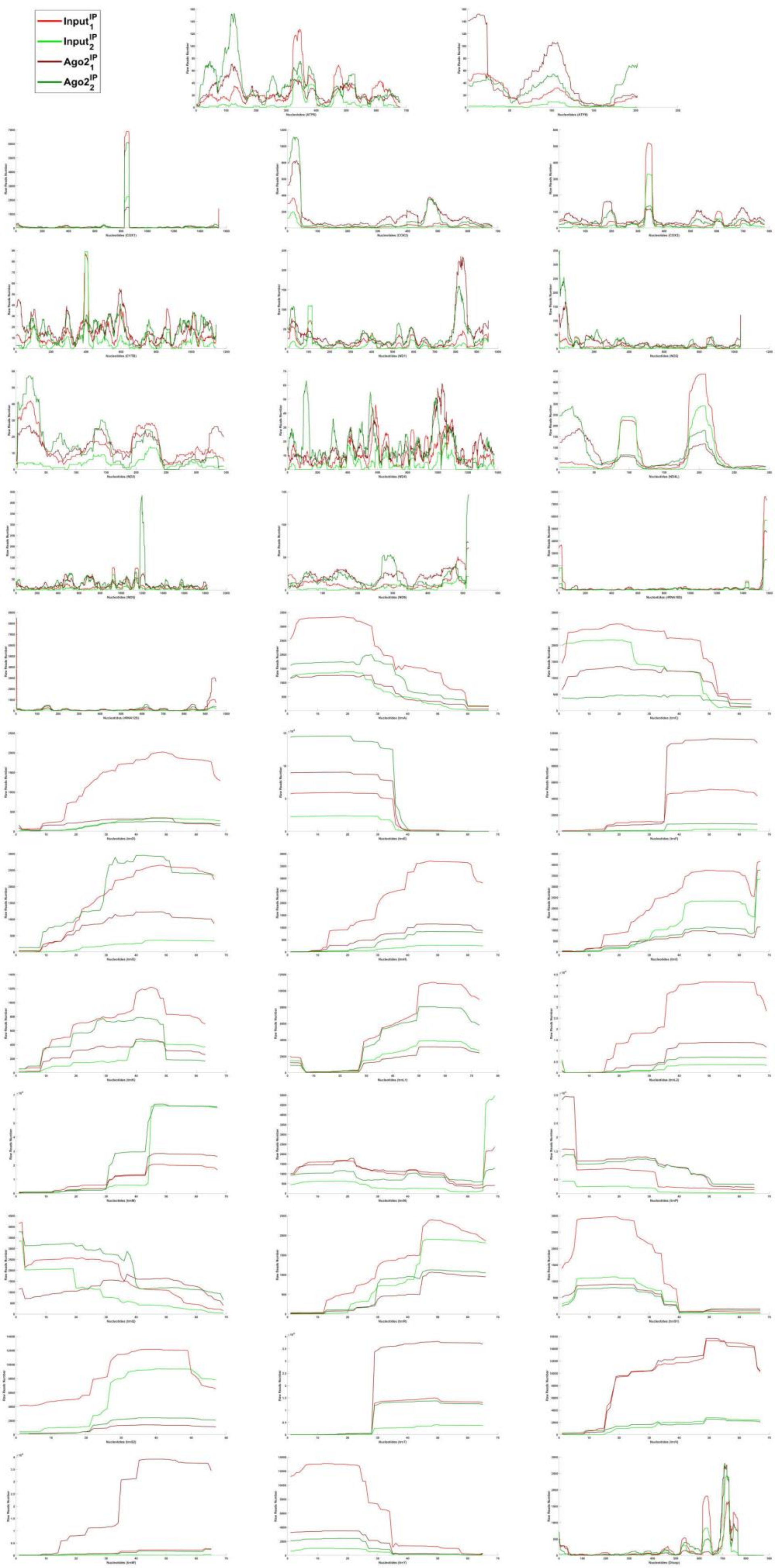
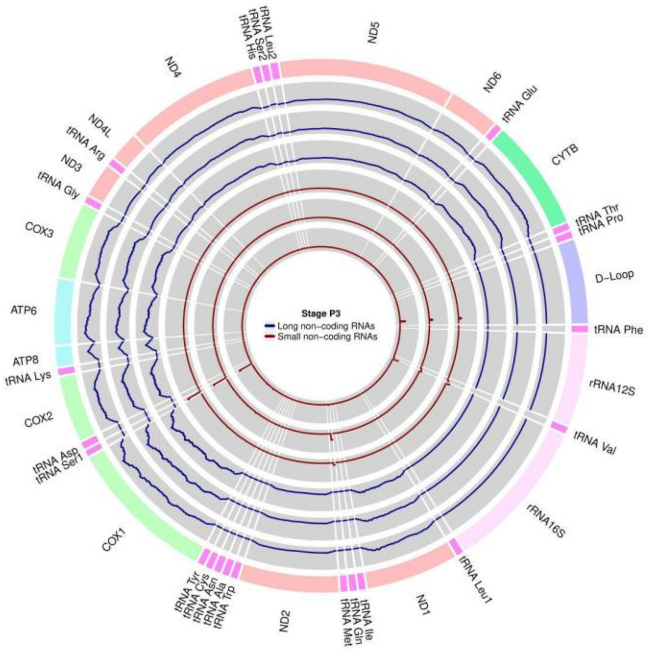
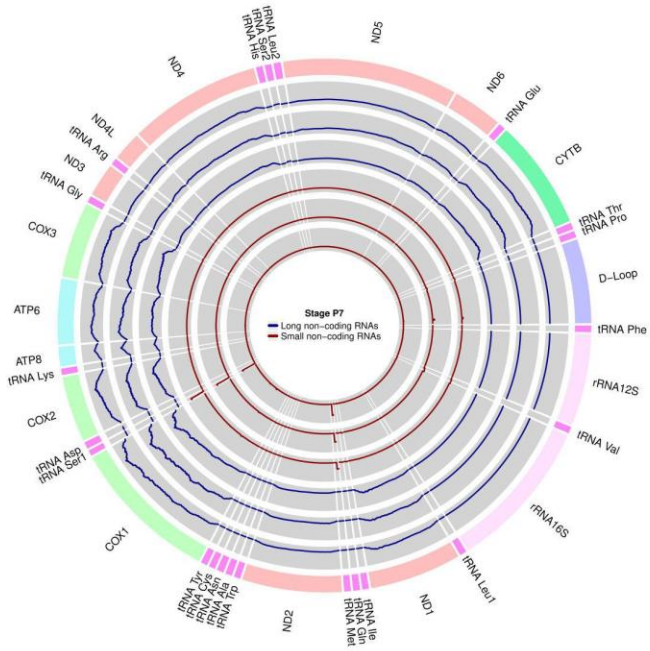
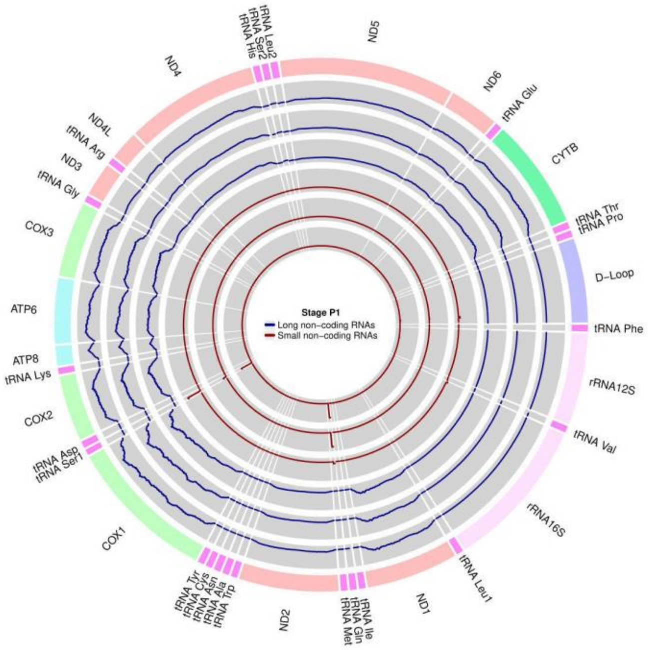

Input$_1^{IP}$
Input$_2^{IP}$
Ago2$_1^{IP}$
Ago2$_2^{IP}$

Figure S5 shows the expression of small and long mitochondrial RNAs across the entire mitochondrial genome in mouse.

In the following figure, we show that the small RNA transcriptional signatures at stages, day1 (P1), day3 (P3), and day7 (P7). There are three biological replicates for each stage, and for each biological replicate both long and small RNAs have been sequenced. Thus, there is one small and one long RNA sample for each biological replicate, providing the best comparison for matching transcriptional signatures. The samples are not specifically enriched for any small or long RNAs, thus they show a mix of coding and non-coding RNAs.

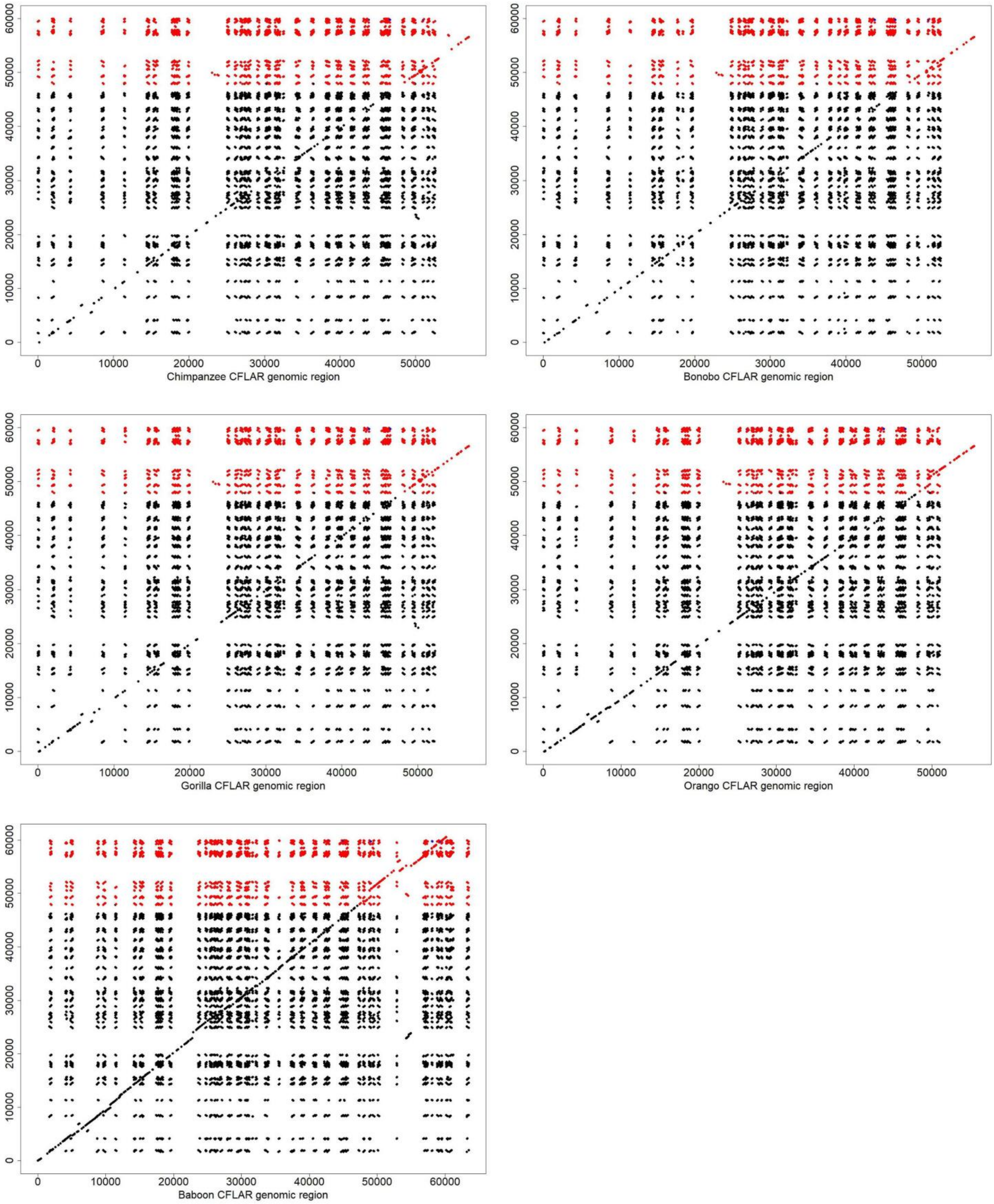

Stage P3
Long non−coding RNAs
Small non−coding RNAs


Stage P7
Long non−coding RNAs
Small non−coding RNAs


Stage P1
Long non−coding RNAs
Small non−coding RNAs

The **figure S6** shows the conservation of the CFLAR genomic region across five primates having a high-quality sequenced genome.

In the following figure, we show five dot-plots comparing the CFLAR genomic region in human (*Homo sapiens)* on the Y-axis to the respective region in chimpanzee (*Pan troglodytes*), bonobo (*Pan paniscus*), gorilla (*Gorilla gorilla*), orango (*Pongo abelii*), and olive baboon (*Papio anubis*) on the X-axis. In these plots each dot represents a sequence match between the two genomic regions, thus showing the conservation of the region on the diagonal of the plot. The region of the CFLAR 3’ UTR is highlighted in red, and the region of binding of mt-miRNAMet is highlighted in blue. The binding region is barely visible, as it has been misplaced to neighbor regions in all species, and it is not present on the diagonal.


Chimpanzee CFLAR genomic region
Bonobo CFLAR genomic region
Gorilla CFLAR genomic region
Orango CFLAR genomic region
Baboon CFLAR genomic region

CHAPTER FIVE

General Discussion

Discussion

In this thesis we investigated some of the fundamental questions about the small mitochondrial RNAs, thus laying the necessary groundwork to start a new branch of mitochondrial biology focused on the mt-miRNAs. First, we confirmed that the small mitochondrial RNAs are indeed transcribed by the mtDNA, thus fulfilling for us and others a fundamental prerequisite, and justifying the further study into the function of these RNAs. Second, we confirmed that these RNAs are conserved across vertebrates spanning ~500mya of divergence (Shu et al. 1999). Third, we confirmed that the small mitochondrial RNAs can bind Ago2, a key protein of RNA interference (RNAi), thus proving strong evidence that the mt-miRNAs are involved in the regulation of gene expression. This allowed us to use the name 'mt-miRNAs", since their ability to bind Ago2 suggests that they are the mitochondrial counterpart of the nuclear miRNAs. Finally, we characterized the most conserved mt-miRNA, mt-miRNA Met, identifying one of its nuclear targets, combining computation and experimental evidence. Through this investigation of fundamental questions about the mt-miRNA, we contributed new advances to the fields of molecular and evolutionary biology.

The evolutionary conservation of mt-miRNAs

Through the comparative analyses performed in the second and fourth chapters, we demonstrated that the mt-miRNAs are conserved across multiple taxa. In the second chapter, we confirmed the findings of previous studies about the conservation and expression of the mt-miRNAs, while providing new insight in untested species and tissues. Indeed, on one hand we confirmed the presence of the mt-miRNAs in *Homo sapiens, Mus musculus,* and *Gallus gallus,* initially discovered by other researchers (Mercer et al. 2011; Bottje et al. 2017; Larriba, Rial, and Del Mazo 2018; Ro et al. 2013). On the other hand, we presented for the first-time the presence of mt-miRNAs from non-placental mammals, and we found tissue-specific transcriptional signatures across multiple tissues for all the species studied. Interestingly, while the transcriptional signatures of the mt-miRNAs most expressed in each tissue were not conserved across chicken and mouse, in the fourth chapter we demonstrate that the expression of a specific mt-miRNA (mt-miRNA-Met) is conserved across these species. This means that although often the most expressed mt-miRNA differs across tissues, some of them are conserved with lower expression in multiple tissues. For example, in the fourth chapter we show that cell cultures from teratoma-derived fibroblast and neuronal progenitor differ in their most-expressed mt-miRNAs, mt-miRNA Ser and mt-miRNA Met respectively. However, the teratoma-derived fibroblast still expresses mt-miRNA Met, although in lower amount. This phenomenon was not clear at the time of the second study, where our focus was on the general expression of mt-miRNAs but became clear when we focused on gene-by-gene analyses in the fourth chapter. This phenomenon indicates that while the most expressed mt-miRNAs vary across species and tissues, specific mt-miRNAs can be conserved across species and tissues. Indeed, the mt-miRNA Met is highly conserved across species and tissues. In the fourth chapter, we analysed mt-miRNA Met conservation across four different key vertebrates (human, mouse, chicken, zebrafish), revealing that this mt-miRNA sequence has accumulated only a handful of mutations in the ~500mya since the origin of *Chordata* (Shu et al. 1999). The high level of sequence conservation of mt-miRNA Met suggests that other species outside the phylum *Chordata* could plausibly also have this mt-miRNA. In fact, in my previous research, we identified for the first time the mt-miRNAs outside the *Chordata* clade, finding a clam (*Ruditapes philippinarum*) expressing many putative mt-miRNAs (Pozzi et al. 2017). Unfortunately, the absence of immunoprecipitation data in this species prevented us from verifying the function of these RNAs, however, through the evidence reported in the fourth chapter, we can now say that the small RNAs mapping to the mitochondria identified in the *R. philippinarum* are likely mt-miRNAs (thus functional). *R. philippinarum* is part of a different phylum (*Mollusca*), which exhibits an ancient divergence from *Chordata*. Thus, considering that the mt-miRNAs are conserved across such distant phyla, it is plausible that the mt-miRNAs are similarly conserved in other metazoan phyla. This possibility is supported by the findings of other researchers, who discovered that even plant mtDNA expresses many small RNAs (Wu et al. 2015). By combining the results of all these studies, we can speculate that these mt-miRNAs might have originated in the ancestral eukaryote. In conclusion, while the discoveries made in this thesis provide solid evidence that the mt-miRNAs are conserved in the *Chordata*, our results lay the foundation for further research aiming to fully explore the mt-miRNAs origin and conservation across eukaryotes.

The function of mt-miRNAs

In the third and fourth chapter, we outline the path required to demonstrate the function of the mt-miRNAs (Chapter 3) and then perform the first step in understanding their function (Chapter 4). In the third chapter we describe several approaches to investigate the function of the mt-miRNAs, including outlining the experimental design that we chose to pursue in the fourth chapter. In the fourth chapter we used RNA-immunoprecipitation data to validate the mechanism used by the mt-miRNAs to exert their function. We chose this experimental design for two main reasons. First, by discovering the mechanism of action of these RNAs, we can use the knowledge of such mechanisms to understand how mitochondrial mutations affect organismal function. Second, we proposed a well-known mechanism, RNAi, so that we could leverage the many datasets available. These datasets originated from two decades of studies used to investigate how nuclear miRNAs regulated protein expression through RNAi, thus, providing optimal data to investigate the function of the mt-miRNAs. Therefore, by following the experimental design discussed in my third chapter, we were able to provide compelling evidence in Chapter 4 that the mechanism of action of the mt-miRNAs is through RNAi.

In this thesis we demonstrated that the mt-miRNAs are part of the RNAi mechanism, through their binding to Ago2. However, we did not fully explore all the possible pathways through which these RNAs could be involved. In fact, in the last chapter of this thesis we performed analyses on multiple datasets to verify the presence of bona fide mt-miRNAs able to bind Ago2, leaving less time to explore all the possible targets of the mt-miRNAs. In my thesis, we prioritized demonstrating that the mt-miRNAs can bind Ago2, a finding that we believe will provide crucial for the future development of the field. Indeed, the main question to emerge since first documented evidence of this new class of small mitochondrial RNAs is whether they are functional or degraded RNAs from the mitochondria. By providing a link to a well-known molecular mechanism, RNAi, we present the first solid evidence of function of the mt-miRNAs. In our opinion, providing a stable ground for the field to develop by demonstrating the origin and function of these RNAs must be done before starting to investigate all the possible phenotypes affected by these RNAs. Furthermore, the single target we were able to identify by using very strict criteria and multiple approaches (CFLAR) revealed a series of unexpected aspects of the mt-miRNAs.

The gene CFLAR is known as an oncosuppressor due to its ability to repress apoptosis (Sarkar et al. 2009). However, not all products from this gene are the same. Indeed, while the gene encoding the protein CFLAR is conserved across many mammals, its mRNA isoforms are not very conserved. One of these isoforms, $CFLAR_L$, is the longest transcript (~14Kb) and the only one known as an apoptosis activator (Yu, Jeffrey, and Shi 2009). Therefore, the $CFLAR_L$ isoform, by activating apoptosis, can increase the rate in which the mitochondria are destroyed. Coincidentally, this isoform is the only one targeted by mt-miRNA Met, as all the other isoforms in humans and other species lack the region binding mt-miRNA Met. The ability of this mt-miRNA to regulate a protein that would destroy its host (the mitochondria) resembles a "defence mechanism" protecting the mitochondria from the nucleus. This possible defence mechanism would involve the mt-miRNA Met blocking the translation of $CFLAR_L$ even though the nucleus expresses the gene to start the apoptosis. This phenomenon would reflect a conflict between the two genomes, as the cell tries to destroy itself while the mitochondria tries to survive. Although hard to verify, this eventuality is fascinating. In fact, while it is well known that virtually all eukaryotic cells harbour many selfish elements within their genome (i.e. transposons) this would be the first case of a miRNA at the centre of an intragenomic conflict between mitochondrial and nuclear genomes. The mt-miRNA Met interaction with $CFLAR_L$ played out through Ago2, could be interpreted as the first case of endogenous genomic conflict. While most researchers are focused on how either the mitochondria or the nucleus adapt to each other to keep collaborating (ie. ATP production), the presence of an endogenous genomic conflict suggests that despite billions of years of coevolution, the nuclear genome still has not fully domesticated its ancestral invader.

Another implication of the role of mt-miRNA Met in apoptosis is for human health. In fact, both mitochondria and apoptosis are known to be relevant in diseases with high mortality rate, such as Alzheimer’s disease (Shimohama 2000; Kukreja et al. 2014). Alzheimer’s disease is a neurodegenerative disease in which neurons in the centers for memory and cognition slowly die through apoptosis, leaving the person affected with multiple disabilities, such as frequent memory loss (Shimohama 2000; Alzheimer 1907). Due to the key role of apoptosis regulation in this disease, the mt-miRNA Met might be involved in this disease. Indeed, studies have demonstrated that increases in the mutation rate in the mtDNA increase the degeneration of neurons, and thus, mutations in the mt-miRNA Met might plausibly be contributing to this phenotype. Nonetheless, identifying specific mitochondrial mutations able to increase neurodegeneration has been difficult, indeed, as we still do not know which mechanism mitochondrial mutations would use. In conclusion, more research would be necessary to understand if the mt-miRNA Met, and other mt-miRNAs, are relevant for human health.

In the third chapter of this thesis, we described the potential role in human health of another mt-miRNA, which, however, is yet to be characterized. This is the mt-miRNA encoded in position 3243 of the mtDNA. We will call this mt-miRNA 3243 due to the relevance of the mutations appearing in position (Deschauer et al. 1999). In fact, it has been estimated that over 80% of the individuals with Mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes (MELAS) carry the A3243G mutation (Ikeda et al. 2018). The existence of mt-miRNA 3243 was documented when the first comprehensive study about mitochondrial transcription was published (Mercer et al. 2011). However, probably due to the amount of data discovered in that article, the finding of mt-miRNA 3243 has been largely ignored. Nonetheless, we hope that through this thesis we provided enough evidence for the mt-miRNAs existence and function to warrant more investigation of the the mt-miRNA 3243 and its potential role in MELAS. As I am writing this Discussion during the COVID19 pandemic, I am unsure on how research in the world will proceed, however, I hope that as soon as possible we will be able to test the function of this mt-miRNA. The ideal method to test if mt-miRNA 3243 has a role in MELAS would be by injecting mt-miRNAs with the nucleotide A in position 3243 into cultured cells of patients having high level of A3243G mtDNA. Then it would be important to compare key features such as mitochondrial functionality, transcriptome and proteome between healthy cells and cells with mutated mtDNA.

Future Directions

The goal of this thesis was to characterize the mt-miRNAs. However, due to the novelty of this field, we were able to only scratch the surface during my PhD studies. Therefore, I will continue investigating this field in my postdoc. Luckily, this year I have been awarded a JSPS Standard Postdoctoral Fellowship that will allow me to continue my research at Tohoku University, Japan. As I plan to continue my research on the mt-miRNAs during my postdoc, and I have already discussed at length how to investigate other aspects of the mt-miRNAs in my third chapter, I will now discuss mostly the research I plan to do in my postdoc.

In this thesis I have described several molecular aspects of the mt-miRNAs, however, in my postdoc I will focus on the most evolutionary aspect of these RNAs. In my postdoc, I plan to investigate if the mt-miRNAs have a role in animal adaptation. In particular, I will leverage published human genomes, both modern and ancient, to verify if there are signals of selection on the mt-miRNAs. Scanning mitochondrial genomes for signs of selection has been done before, however, not while considering the existence of “mitonuclear blocks (MBs)”.

An MB is a conceptual innovation which I defined as the pairing of a mt-miRNA that binds a partially complementary nuclear mRNA and regulates its function through RNAi (**Fig.1**). In an

MB only the mRNA has a direct effect on phenotypes and the mt-miRNA functions as a regulatory element of the mRNA. Therefore, the same mt-miRNA can be involved in multiple functions by affecting multiple mRNAs, forming multiple MBs. These multiple MBs regulate pathways within the cell, and mutations in genes encoding for either the mitochondrial or the nuclear RNA can lead to MB disruption and changes in the organism phenotype. Defining the MB is vital to adapt normal methods used to study natural selection to the study of the mt-miRNAs. In fact, usually these methods focus on only one genome (Bamshad and Wooding 2003), either mitochondrial or nuclear. However, mutations arising on either part of the MB can modify the phenotype, thus, analyses investigating only one genome (mtDNA) have lower power of detecting signals of adaptation (Sloan, Fields, and Havird 2015). Instead, by identifying the MBs present in a species, it should be possible to then verify the conservation of these MBs across populations exhibiting different phenotypes.

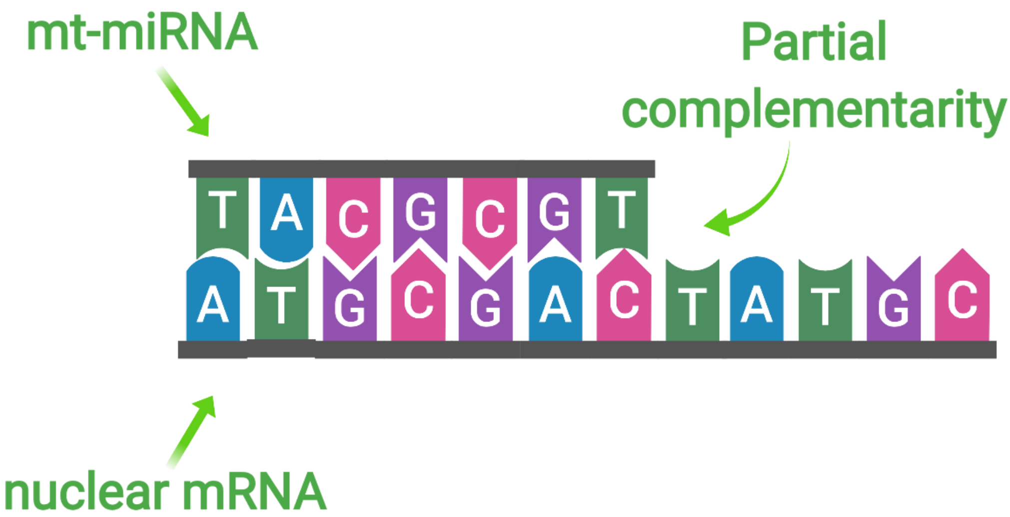


Fig.1 In this figure, we represented a mitonuclear block (MB) where a mt-miRNA binds the regulatory region of nuclear mRNA to regulate its translation.

The phenotype that I predict to be more likely influenced by MB disruption is thermal tolerance. This prediction is based on the research previously done by other researchers on the relationship between mtDNA mutations and phenotype changes (Mishmar et al. 2003; Camus et al. 2017). In fact, thermal tolerance is now one of the best-known cases where different mt-haplotypes can influence how fruit flies will react to changes in temperature. This research has been done in the Dowling lab and spans multiple studies. In these studies, the authors found that along the east coast of Australia, where a climatic cline is present, fruit flies have different frequencies of mt-haplotypes, which correlate with the changes in temperature along the cline. Furthermore, the researchers demonstrated that flies with the mt-haplotype more common in subtropical areas are more resistant to heat shock when compared to the flies carrying the mt-haplotype more common in temperate areas. This comparison holds when the flies are in an isogenic nuclear background (have the same nuclear genome), thus demonstrating that variability in the mtDNA can affect thermal tolerance. This convinced me that investigating whether MBs are conserved across populations that live in different temperatures might reveal signs of adaptation in the mt-miRNA. In fact, I expect populations living in warm and cold environments to have different MBs related to temperature regulation. Finding that MBs favourable for cold environments in populations living in cold environments and discovering the absence of such MBs in populations living in warm environments would suggest that the mt-miRNAs are indeed adaptive. Moreover, by combining ancient DNA (aDNA) data to data from modern populations would further improve the results. It is possible to use the aDNA to verify the presence of specific MBs in ancient populations, and then verify if these MBs persisted in more recent populations. By combining these genomics data to climatic data, we can verify if specific populations lost MBs when the environment changed. In conclusion, by using the findings of this thesis, in my postdoc I will be able to investigate if the mt-miRNAs are not only functional, but adaptive.

Conclusion

In this thesis we elucidated aspects of mitochondrial biology that were previously unknown. We characterized the mt-miRNAs by confirming their origin and function (in the second and fourth chapter) while at the same time providing suggestions for the medical community on how to further progress our findings in their own projects (in the third chapter). This thesis contributes to the development of the field of mitonuclear communication, but at the same time, provides new information for core fields of biology, such as cellular and molecular biology. This thesis focuses on a subject that has to date yielded less than 10 published works, albeit many of which have been at the forefront of mitochondrial biology (Mercer et al. 2011; Bottje et al. 2017; Ro et al. 2013; Larriba, Rial, and Del Mazo 2018; Ma et al. 2016; Riggs et al. 2018; Riggs, Woll, and Podrabsky 2019; Riggs and Podrabsky 2019). Thus, my work leverages a newly emerging field of research, and tries to open the way for further research into this field by providing necessary information for others to keep investigating how the mitochondria can affect so many diverse animal phenotypes, influencing health (Picard, Wallace, and Burelle 2016), adaptation (Rand and Mossman 2020) and speciation (Hill 2015).

References

Alzheimer, Alois. 1907. "About a Peculiar Disease of the Cerebral Cortex." *Centralblatt Für Nervenheilkunde Psychiatrie* 30 (1): 177–79.

Bamshad, Michael, and Stephen P. Wooding. 2003. "Signatures of Natural Selection in the Human Genome." *Nature Reviews. Genetics* 4 (2): 99–111.

Bottje, Walter G., Bhuwan Khatri, Stephanie A. Shouse, Dongwon Seo, Barbara Mallmann, Sara K. Orlowski, Jeonghoon Pan, et al. 2017. "Identification and Differential Abundance of Mitochondrial Genome Encoding Small RNAs (mitosRNA) in Breast Muscles of Modern Broilers and Unselected Chicken Breed." *Frontiers in Physiology* 8 (October): 816.

Camus, M. Florencia, Jonci N. Wolff, Carla M. Sgrò, and Damian K. Dowling. 2017. "Experimental Support That Natural Selection Has Shaped the Latitudinal Distribution of Mitochondrial Haplotypes in Australian Drosophila Melanogaster." *Molecular Biology and Evolution* 34 (10): 2600–2612.

Deschauer, M., T. Wieser, S. Neudecker, A. Lindner, and S. Zierz. 1999. "Mitochondrial 3243 A-->G Mutation (MELAS Mutation) Associated with Painful Muscle Stiffness." *Neuromuscular Disorders: NMD* 9 (5): 305–7.

Hill, Geoffrey E. 2015. "Mitonuclear Ecology." *Molecular Biology and Evolution* 32 (8): 1917–27.

Ikeda, Takahiro, Hitoshi Osaka, Hiroko Shimbo, Makiko Tajika, Masayo Yamazaki, Ayako Ueda, Kei Murayama, and Takanori Yamagata. 2018. "Mitochondrial DNA 3243A>T Mutation in a Patient with MELAS Syndrome." *Human Genome Variation* 5 (September): 25.

Kukreja, Lokesh, Gregory C. Kujoth, Tomas A. Prolla, Fred Van Leuven, and Robert Vassar. 2014. "Increased mtDNA Mutations with Aging Promotes Amyloid Accumulation and Brain Atrophy in the APP/Ld Transgenic Mouse Model of Alzheimer's Disease." *Molecular Neurodegeneration*. https://doi.org/10.1186/1750-1326-9-16.

Larriba, Eduardo, Eduardo Rial, and Jesús Del Mazo. 2018. "The Landscape of Mitochondrial Small Non-Coding RNAs in the PGCs of Male Mice, Spermatogonia, Gametes and in Zygotes." *BMC Genomics* 19 (1): 634.

Ma, Hao, Gregory M. Weber, Hairong Wei, and Jianbo Yao. 2016. "Identification of Mitochondrial

Genome-Encoded Small RNAs Related to Egg Deterioration Caused by Postovulatory Aging in Rainbow Trout." *Marine Biotechnology* 18 (5): 584–97.

Mercer, Tim R., Shane Neph, Marcel E. Dinger, Joanna Crawford, Martin A. Smith, Anne-Marie J. Shearwood, Eric Haugen, et al. 2011. "The Human Mitochondrial Transcriptome." *Cell* 146 (4): 645–58.

Mishmar, Dan, Eduardo Ruiz-Pesini, Pawel Golik, Vincent Macaulay, Andrew G. Clark, Seyed Hosseini, Martin Brandon, et al. 2003. "Natural Selection Shaped Regional mtDNA Variation in Humans." *Proceedings of the National Academy of Sciences of the United States of America* 100 (1): 171–76.

Picard, Martin, Douglas C. Wallace, and Yan Burelle. 2016. "The Rise of Mitochondria in Medicine." *Mitochondrion* 30 (September): 105–16.

Pozzi, Andrea, Federico Plazzi, Liliana Milani, Fabrizio Ghiselli, and Marco Passamonti. 2017. "SmithRNAs: Could Mitochondria 'Bend' Nuclear Regulation?" *Molecular Biology and Evolution*, April. https://doi.org/10.1093/molbev/msx140.

Rand, David M., and Jim A. Mossman. 2020. "Mitonuclear Conflict and Cooperation Govern the Integration of Genotypes, Phenotypes and Environments." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 375 (1790): 20190188.

Riggs, Claire L., and Jason E. Podrabsky. 2019. "An Anoxia-Tolerant Cell Line, AL4, Provides New Insight into Small Non-Coding RNA Biology." *The FASEB Journal* 33 (1_supplement): 714.4–714.4.

Riggs, Claire L., Amanda Summers, Daniel E. Warren, Göran E. Nilsson, Sjannie Lefevre, W. W. Dowd, Sarah Milton, and Jason E. Podrabsky. 2018. "Small Non-Coding RNA Expression and Vertebrate Anoxia Tolerance." *Frontiers in Genetics* 9 (July): 230.

Riggs, Claire L., Steven Cody Woll, and Jason E. Podrabsky. 2019. "MitosRNAs and Extreme Anoxia Tolerance in Embryos of the Annual Killifish Austrofundulus Limnaeus." *Scientific Reports* 9 (1): 19812.

Ro, Seungil, Hsiu-Yen Ma, Chanjae Park, Nicole Ortogero, Rui Song, Grant W. Hennig, Huili Zheng, et al. 2013. "The Mitochondrial Genome Encodes Abundant Small Noncoding RNAs." *Cell Research* 23 (6): 759–74.

Sarkar, S. A., B. Kutlu, K. Velmurugan, S. Kizaka-Kondoh, C. E. Lee, R. Wong, A. Valentine, H. W. Davidson, J. C. Hutton, and S. Pugazhenthi. 2009. "Cytokine-Mediated Induction of Anti-Apoptotic Genes That Are Linked to Nuclear Factor Kappa-B (NF-kappaB) Signalling in Human Islets and in a Mouse Beta Cell Line." *Diabetologia* 52 (6): 1092–1101.

Shimohama, S. 2000. "Apoptosis in Alzheimer's Disease--an Update." *Apoptosis: An International Journal on Programmed Cell Death* 5 (1): 9–16.

Shu, D-G, H-L Luo, S. Conway Morris, X-L Zhang, S-X Hu, L. Chen, J. Han, M. Zhu, Y. Li, and L-Z Chen. 1999. "Lower Cambrian Vertebrates from South China." *Nature* 402 (6757): 42–46.

Sloan, Daniel B., Peter D. Fields, and Justin C. Havird. 2015. "Mitonuclear Linkage Disequilibrium in Human Populations." *Proceedings. Biological Sciences / The Royal Society* 282 (1815). https://doi.org/10.1098/rspb.2015.1704.

Wu, Zhiqiang, James D. Stone, Helena Štorchová, and Daniel B. Sloan. 2015. "High Transcript Abundance, RNA Editing, and Small RNAs in Intergenic Regions within the Massive Mitochondrial Genome of the Angiosperm Silene Noctiflora." *BMC Genomics* 16 (November): 938.

Yu, Jong W., Philip D. Jeffrey, and Yigong Shi. 2009. "Mechanism of Procaspase-8 Activation by c-FLIPL." *Proceedings of the National Academy of Sciences of the United States of America* 106 (20): 8169–74.