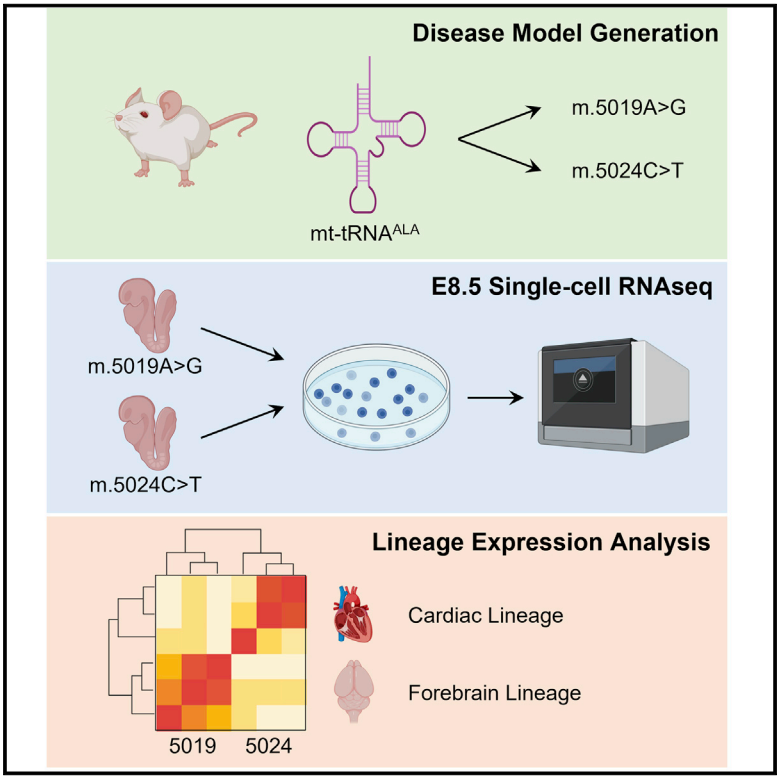
Cell

Cell lineage-specific mitochondrial resilience during mammalian organogenesis

Graphical abstract



Highlights

- Cell lineage-specific mitochondrial gene expression emerges early in development
- Disrupting mt-protein synthesis induces lineage-specific transcriptional responses
- These responses are both mutation and tissue specific
- Transcription factors promoting organelle resilience mediate the response

Authors

Stephen P. Burr, Florian Klimm, Angelos Glynos, ..., Michal Minczuk, James B. Stewart, Patrick F. Chinnery

Correspondence

pfc25@cam.ac.uk

In brief

Emergence of lineage-specific gene expression profiles during early mouse development provides fresh insights into organ-specific differences in mitochondrial function in health and disease in mammals.

Burr et al., 2023, Cell *186*, 1212–1229
March 16, 2023 © 2023 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.cell.2023.01.034

CellPress

CellPress
OPEN ACCESS

Cell

Article

Cell lineage-specific mitochondrial resilience during mammalian organogenesis

Stephen P. Burr,[1,2,12] Florian Klimm,[1,2,3,4,5,6,12] Angelos Glynos,[1,2,12] Malwina Prater,[1,2,12] Pamella Sendon,[7] Pavel Nash,[2] Christopher A. Powell,[2] Marie-Lune Simard,[7] Nina A. Bonekamp,[7,8] Julia Charl,[9] Hector Diaz,[10] Lyuba V. Bozhilova,[1,2] Yu Nie,[1,2] Haixin Zhang,[1,2] Michele Frison,[1,2] Maria Falkenberg,[10] Nick Jones,[3] Michal Minczuk,[2] James B. Stewart,[7,11] and Patrick F. Chinnery[1,2,13,*]

[1]Department of Clinical Neurosciences, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK
[2]Medical Research Council Mitochondrial Biology Unit, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK
[3]Department of Mathematics, Imperial College London, London, UK
[4]EPSRC Centre for Mathematics of Precision Healthcare, Imperial College, London, UK
[5]Department of Computational Molecular Biology, Max Planck Institute for Molecular Genetics, Ihnestraße 63-73, D-14195 Berlin, Germany
[6]Department of Computer Science, Freie Universität Berlin, Arnimallee 3, D-14195 Berlin, Germany
[7]Max Planck Institute for Biology of Ageing, Cologne, Germany
[8]Department of Neuroanatomy, Mannheim Centre for Translational Neuroscience (MCTN), Medical Faculty Mannheim/Heidelberg University, Heidelberg, Germany
[9]Institute of Biochemistry, University of Cologne, Otto-Fischer-Strasse 12-14, Cologne, Germany
[10]Department of Medical Biochemistry and Cell Biology, University of Gothenburg, PO Box 440, Gothenburg 405 30, Sweden
[11]Biosciences Institute, Faculty of Medical Sciences, Wellcome Centre for Mitochondrial Research, Newcastle University, Newcastle upon Tyne, UK
[12]These authors contributed equally
[13]Lead contact
*Correspondence: pfc25@cam.ac.uk
https://doi.org/10.1016/j.cell.2023.01.034

SUMMARY

Mitochondrial activity differs markedly between organs, but it is not known how and when this arises. Here we show that cell lineage-specific expression profiles involving essential mitochondrial genes emerge at an early stage in mouse development, including tissue-specific isoforms present before organ formation. However, the nuclear transcriptional signatures were not independent of organelle function. Genetically disrupting intra-mitochondrial protein synthesis with two different mtDNA mutations induced cell lineage-specific compensatory responses, including molecular pathways not previously implicated in organellar maintenance. We saw downregulation of genes whose expression is known to exacerbate the effects of exogenous mitochondrial toxins, indicating a transcriptional adaptation to mitochondrial dysfunction during embryonic development. The compensatory pathways were both tissue and mutation specific and under the control of transcription factors which promote organelle resilience. These are likely to contribute to the tissue specificity which characterizes human mitochondrial diseases and are potential targets for organ-directed treatments.

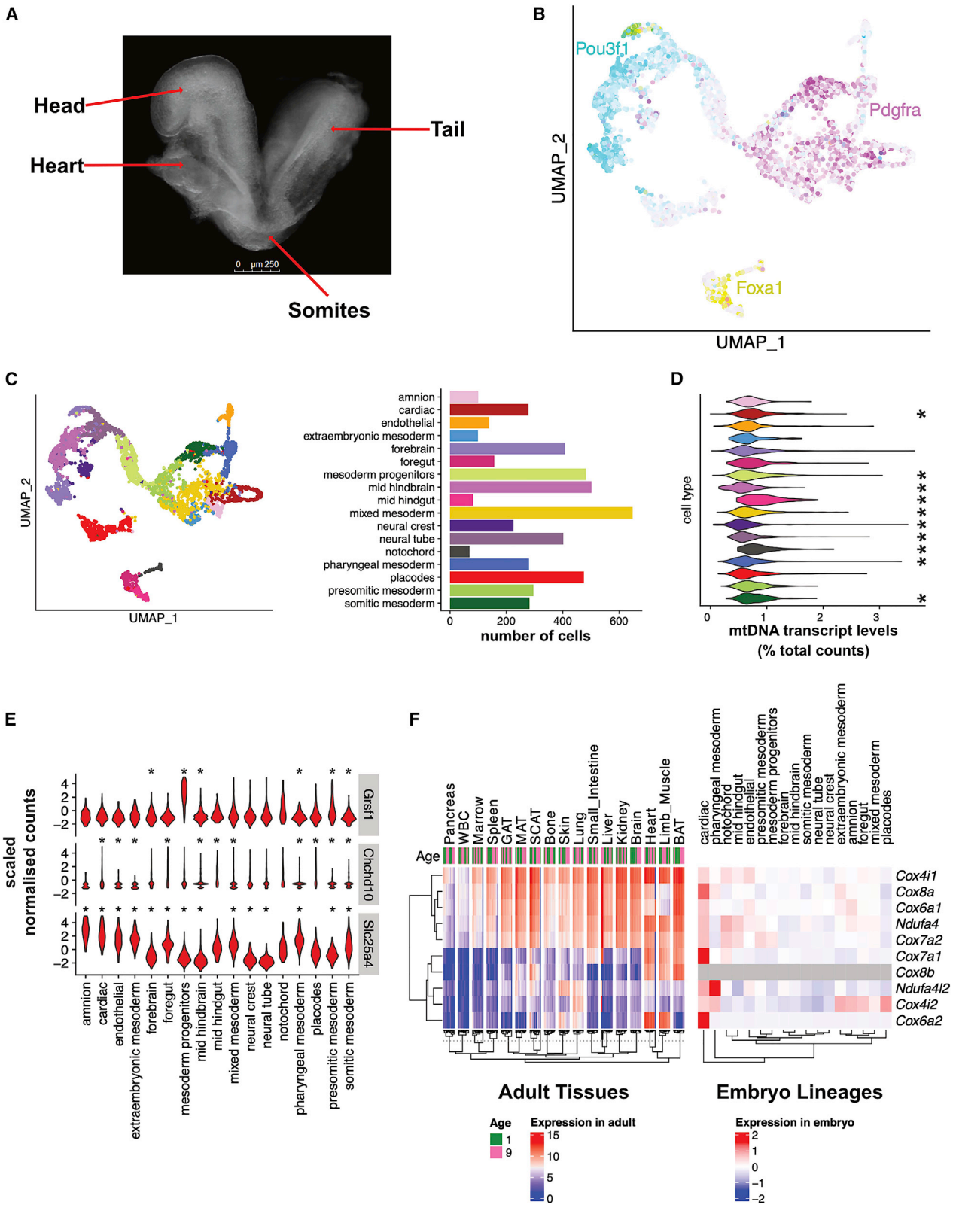
INTRODUCTION

Mitochondria are the principal cellular energy source under aerobic conditions. Reducing equivalents generated from intermediary metabolism donate electrons to complex I, III, and IV of the respiratory chain, pumping protons from the matrix to the intermembrane space. The electrochemical force is harnessed by complex V to regenerate adenosine triphosphate (ATP) through oxidative phosphorylation (OXPHOS). ATP is required for all active cellular processes, including protein synthesis, cell division, and the maintenance of ionic gradients across the membranes.[1,2]

Most of the ~1,100 proteins constituting the mitochondrial proteome are encoded by the nuclear genome that is inherited from both parents,[3,4] but thirteen essential respiratory chain proteins are encoded by multiple copies of the 16.5-kb mitochondrial genome (mtDNA), which is inherited independently down the maternal line.[5] mtDNA also codes for 24 rRNAs and tRNAs, which partner with the nuclear-encoded gene-expression apparatus to perform intra-mitochondrial protein synthesis. The assembly of the respiratory chain on the inner mitochondrial membrane also requires many nuclear-encoded assembly factors. Thus, the coordinated and close cooperation between the nuclear and mitochondrial genomes is essential for aerobic ATP production and fundamental for cell viability.

Different tissues and organs have a strikingly different capacity for oxidative ATP production reflecting their energy

 © 2023 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).



(legend on next page)

Table 1. Number of cells after alignment and pre-processing in both batches

Batch A: Wild-type	Batch A: 5024 mutation	Batch B: Wild-type	Batch B: 5019 mutation	Batch B: 5024 mutation
2,174	2,456	2,750	1,649	1,991

Batch A consists of wild-type and the 5024 mutation. Batch B consists of wild-type, 5019, and 5014 mutation. In total, 11,020 cells pass the quality control filtering.

dependence,[4,6–8] but it is not known how or when this emerges during development. Here we show that cell lineage-specific transcriptional profiles are present by embryonic day 8.5 (E8.5) in mice, including organ-specific mitochondrial protein isoforms present before organogenesis is complete. Disrupting intra-mitochondrial protein synthesis down-regulated nuclear genes known to buffer the effects of mitochondrial toxins when knocked-out *in vitro*,[9] thus maintaining cell lineage viability during *in vivo* development. Single cell transcriptomic analysis identified lineage and mtDNA mutation-specific mechanisms that compensate for mitochondrial dysfunction during key developmental stages in the early embryo, including transcription factors known to regulate gene networks. These findings show that nuclear-mitochondrial crosstalk is established during early mammalian development, long before organ maturation. This provides resilience to mitochondrial insults and likely contributes to the organ-specific vulnerability that characterizes rare genetic mitochondrial diseases and common human disorders where mitochondrial mechanisms play a role.

RESULTS

Differential expression of nuclear-encoded mitochondrial genes during organogenesis

Although mitochondrial function is known to vary between organs at birth,[7,10,11] it is not known when this arises. We determined whether the differences were apparent when organ-specific cell lineages are established during embryogenesis. We performed single cell transcript analysis on six whole C57Bl/6 wild-type (WT) mouse embryos following routine mating (Figure S1A). Embryos were collected at E8.5 post plugging, with the developmental stage confirmed morphologically (Figure 1A). 4,924 cells passed transcript quality control (Table 1). Uni-form Mani-fold Approximation and Projection (UMAP)[12] of the single cell transcript resolved the ectodermal, endodermal, and mesodermal germ layers based on *Foxa1*, *Pdgfra*, and *Pou3F1* expression (Figure 1B)[13,14] and allowed the identification of 17 distinct tissue lineages based on their transcriptomic signature (Figure 1C).[15] There were significant differences in the overall abundance of mtDNA-encoded transcripts across several of the identified tissue lineages (Figure 1D), and we also saw a heterogeneous expression pattern for the nuclear-encoded mitochondrial transcripts (Table S1, showing all genes significantly differentially expressed in at least one lineage). Some nuclear-mitochondrial genes (60%, 631/1043) showed equivalent expression levels across all tissues, including representative examples *Ndufb5*, *Timm44*, and *Cox10* (Figure S1B). However, the remaining 412 showed differential expression in at least one lineage including essential mitochondrial genes, with the three most significant being *Grsf1*, *Chchd10*, & *Slc25a4* (Figure 1E). This was also apparent when we specifically analyzed the transcripts of all known tissue-specific gene isoforms of electron transport chain subunits (Figure 1F). For example, in the cardiac lineage we saw expression of *Cox6a2* and *Cox7a1*, which encode heart-specific isoforms of complex IV subunits (Figure S1C),[16] recapitulating the expression pattern seen in adult mouse heart and muscle tissues (Figure 1F).[17] Furthermore, we saw a clear subpopulation of the cardiac lineage cells expressing both of these isoforms (Figure S1D), likely representing cardiomyocytes, which are known to be present by E8.5.[18,19] Like-wise, the oxygen-responsive *Cox4i2* isoform[20] was differentially expressed in several lineages (Figure 1F). Together, these findings indicate that the cell nucleus already modulates mitochondrial transcripts in a tissue-specific manner in early embryonic development, pre-empting organ maturation, as opposed to specific genes being expression in response to established organ physiological function.

Next, we determined whether the lineage-specific nuclear transcriptional signatures were established independently of the mitochondrion or were directly influenced by mtDNA gene expression. To do this *in vivo,* we studied the nuclear transcriptional response to two different point mutations in the mitochondrial $tRNA^{Ala}$ gene (*mt-Ta*) that disrupt intra-mitochondrial translation and OXPHOS:

Figure 1. Differential expression of nuclear-encoded mitochondrial genes during organogenesis

(A) Representative image of an E8.5 mouse embryo, with morphological features labeled. Scale bar, 250 μm.

(B) UMAP of single-cell transcripts defining ectodermal, endodermal, and mesodermal cell lineages based on *Pou3f1* (blue), *Foxa1* (yellow), and *Pdgfra* (magenta) expression, respectively.

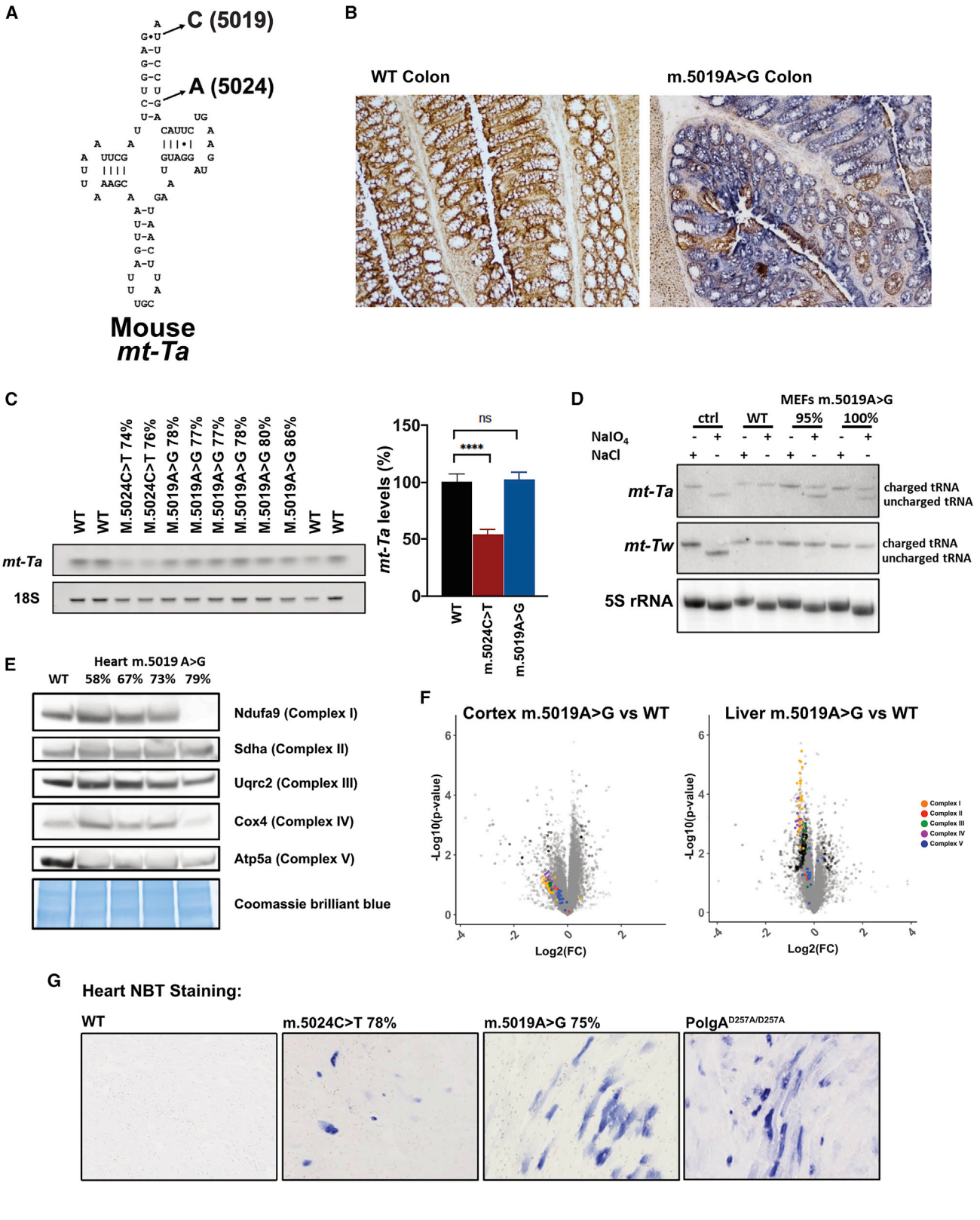
(C) Left, UMAP showing 17 different cell lineages, defined by unbiased analysis of known marker-gene transcript expression levels. Each dot represents a single cell. Right, number of cells in each lineage.

(D) Abundance of mtDNA-encoded transcripts across different cell types, showing significant differential gene expression in 10 of the 17 lineages. Colors match the lineages defined in (C). *$p < 0.05$.

(E) Violin plots showing transcript abundance of selected nuclear-encoded mitochondrial genes *Grsf1*, *Chchd10*, and *Slc25a4*, showing significant differential expression across the 17 cell lineages. *$p < 0.05$. Selected genes represent the top three most differentially regulated transcripts across at least two cell lineages. p values for all nuclear mitochondrial genes with significant differential expression in at least one cell lineage can be found in Table S1.

(F) Heatmap showing expression levels of known electron transport chain (ETC) subunit gene isoforms in bulk RNA-seq from adult mouse tissues (left)[17] and single-cell RNA-seq from E8.5 embryonic cell lineages (right). Strong expression of the muscle isoforms *Cox7a1* and *Cox6a2* is seen in embryonic cells of the cardiac lineage, revealing that the heart/muscle-specific expression pattern of these isoforms seen in adult tissues is already established by the early stages of organogenesis. UMAP, Uni-form Mani-fold Approximation and Projection.

See also Figure S1.



(legend on next page)

the previously reported m.5024C>T mutation,[21] and the m.5019A>G mutation (Figure 2A).

Characterization of the m.5019A>G *mt-Ta* mutation

The m.5019A>G mouse was identified in the offspring derived from a female heterozygous for the proof-reading deficient mtDNA polymerase allele *PolgA*^{D257A}.[22] Individual pups were screened for a high proportion of mitochondrial complex IV deficient colonic crypts as described.[21] Laser-capture microdissection of individual crypts revealed a single mtDNA molecule co-segregating with the mitochondrial deficiency that bore three mutations *in cis*: m.5019A>G, m.15200A>G, and m.16232A>T (Figures 2B and S2A). m.15200A>G is synonymous at amino acid Q352 in *Cytb* and thus does not disrupt mitochondrial function. m.16232A>T affects the binding site of the Tfam/Polrmt/Tefm binding complex within the light strand promoter (LSP) region of the mtDNA control region.[23] A competitive *in vitro* transcription assay showed that m.16232A>T led to a very mild decrease in transcription initiation, which was extenuated by increasing levels of Tfam to the transcribed template (Figures S2B–S2F). Consistent with this, steady-state levels of the mitochondrial transcripts including the tRNAs, mRNAs, and polycistronic pre-mRNAs of mitochondria, all sensitive indicators of transcription alterations,[24,25] remained unaffected compared to WT mice (Figures S2G–S2I). The LSP promoter also plays a role in primer formation for mtDNA replication, but we found no difference in mtDNA copy number (Figures S2J–S2K). Unlike in the m.5024C>T mutation,[21] we saw no evidence of an in organellar translation defect in m.5019A>G heart and liver, highlighting the different molecular mechanisms involved. Together, these results make it highly unlikely that m.16232A>T is the primary pathogenic mutation causing the complex IV deficiency, although a contributory role cannot be excluded. We therefore concluded that m.5019A>G was the primary pathogenic mtDNA mutation in this mouse line.

Despite being only 5 base pairs (bp) from the m.5024C>T mutation (Figure 2A),[21] the m.5019A>G mutation did not cause tRNA instability (Figure 2C), likely because the mutation causes the G-U wobble base-pairing to convert to a canonical G-C base-pairing in the mutant tRNA. The acylation state of *mt-Ta* was assessed by periodate oxidation, followed by northern blotting.[26] Uncharged *mt-Ta* were observed in embryonic fibroblasts derived from the mouse (MEFs) (Figure 2D), and in the small intestinal epithelium and heart of m.5019A>G mice (Figure S2L). The high-resolution northern blotting showed that the m.5019A>G mutated tRNA^{Ala} was the same length as WT tRNA^{Ala}, indicating that the post-transcriptional CCA extension was intact, which excludes this as a contributory pathogenic mechanism. Western blots for subunits of the mitochondrial respiratory complexes in the heart (Figure 2E) and small intestinal epithelium (Figures S2M and S3A) showed a marked reduction in complexes I and IV (cytochrome *c* oxidase, COX), and some reduction in complex III, consistent with a mitochondria-specific translational defect. In keeping with this, quantitative proteomic analysis of cerebral cortex and liver showed reduced abundance of both nuclear and mtDNA-encoded OXPHOS subunits in m.5019A>G mice compared to controls (Figure 2F). In the liver, this preferentially involved complexes I, III, and IV, consistent with a mitochondrial-specific translational defect (Figure S3B). The bulk proteomic analysis in adult tissues showed no consistent difference in the abundance of proteins involved in mitochondrial ribosome assembly nor in the ribosomal proteins themselves, making it unlikely that ribosomal stability contributed to the molecular mechanism.

Mitochondrial function was assessed in multiple tissues by comparing complex IV activity to nuclear encoded mitochondrial complex II (succinate dehydrogenase, SDH) *in situ*.[27] Mice of all ages assayed with >75% m.5019A>G showed COX deficient cells in many tissues, including the heart (Figures 2G, S3C, and S3D), colonic and intestinal epithelium, liver, skeletal muscle, kidney (Figure S3D), choroid plexus, and hippocampus

Figure 2. Characterization of the m.5019A>G *mt-Ta* mutation

(A) Position and effect of the m.5024C>T and m.5019A>G mutations in the *mt-Ta* gene.

(B) Representative COX-SDH dual histochemical staining of the colonic crypts from a WT mouse and an m.5019A>G mouse. Blue staining of the m.5019A>G tissue indicates COX-deficient crypts, which harbored higher levels of the m.5019A>G mutation than adjacent brown COX-normal crypts, quantified in Figure S2A.

(C) tRNA northern blots from heart tissue, showing levels of *mt-Ta* in mice bearing the m.5024C>T and m.5019A>G mutations at various heteroplasmy levels (%). Quantification of northern blot signal, shown in the bar graph, reveals significant reduction of *mt-Ta* in high heteroplasmy m.5024C>T mice compared to WT, whilst m.5019A>G mice of comparable heteroplasmy maintain normal *mt-Ta* levels. Tukey's multiple comparison test from a one-way ANOVA. 18S rRNA shown as loading control.

(D) High-resolution northern blot of MEFs from m.5019A>G or WT mice, showing aminoacylation states of *mt-Ta* and *mt-Tw* in the presence ("+") of $NaIO_4$ or NaCl. Incubation of freshly extracted RNA with $NaIO_4$ results in the removal of a single 3′ ribose on all tRNA not charged with an amino acid, leaving the charged tRNAs unaffected. Following the removal of amino acids from the charged tRNA, the initial ratio of charged to uncharged tRNA in the live cells can be visualized by high-resolution Urea PAGE, which can separate RNA molecules differing in molecular weight by a single ribose. A specific *mt-Ta* aminoacylation defect is seen in m.5019A>G MEFs when exposed to $NaIO_4$. 5S rRNA shown as loading control. Ctrl, RNA samples intentionally deacylated (alkaline treatment) and incubated with $NaIO_4$ or NaCl.

(E) Western blot of heart lysates from WT and m.5019A>G mice of varying relative mutation levels for selected respiratory chain complex subunits, showing increasing complex I and IV instability with increasing mutation load (%). Coomassie brilliant blue shown as loading control.

(F) Volcano plots showing differential protein levels of respiratory chain complex subunits in cerebral cortex and liver samples derived from m.5019A>G mice compared to WT controls. Black dots, significantly different based on a two-sample t test (FDR = 0.05). While respiratory chain subunit levels in cortex samples of m.5019A>G mice reveal a mild decrease compared to controls, liver samples exhibit a stronger response involving components of complex I, III, and IV.

(G) NTB staining of hearts from WT, m.5024C>T, m.5019A>G, and *PolgA* $^{D257A/D257A}$ mtDNA mutator mice, showing increased incidence of COX deficient cells in m.5019A>G heart compared to m.5024C>T & WT. Cells with normal COX activity exclude NTB, whilst cells with COX deficiency stain blue. *PolgA* $^{D257A/D257A}$ is included as a positive control for COX deficiency. COX, Cytochrome *c* Oxidase; FDR, false discovery rate. SDH, succinate dehydrogenase; MEF, mouse embryonic fibroblast; WT, wild-type; PAGE, polyacrylamide gel electrophoresis; NTB, nitrotetrazoleum blue; ns, not significant. ****$p < 0.0001$, error bars represent SD.

See also Figures S2 and S3.

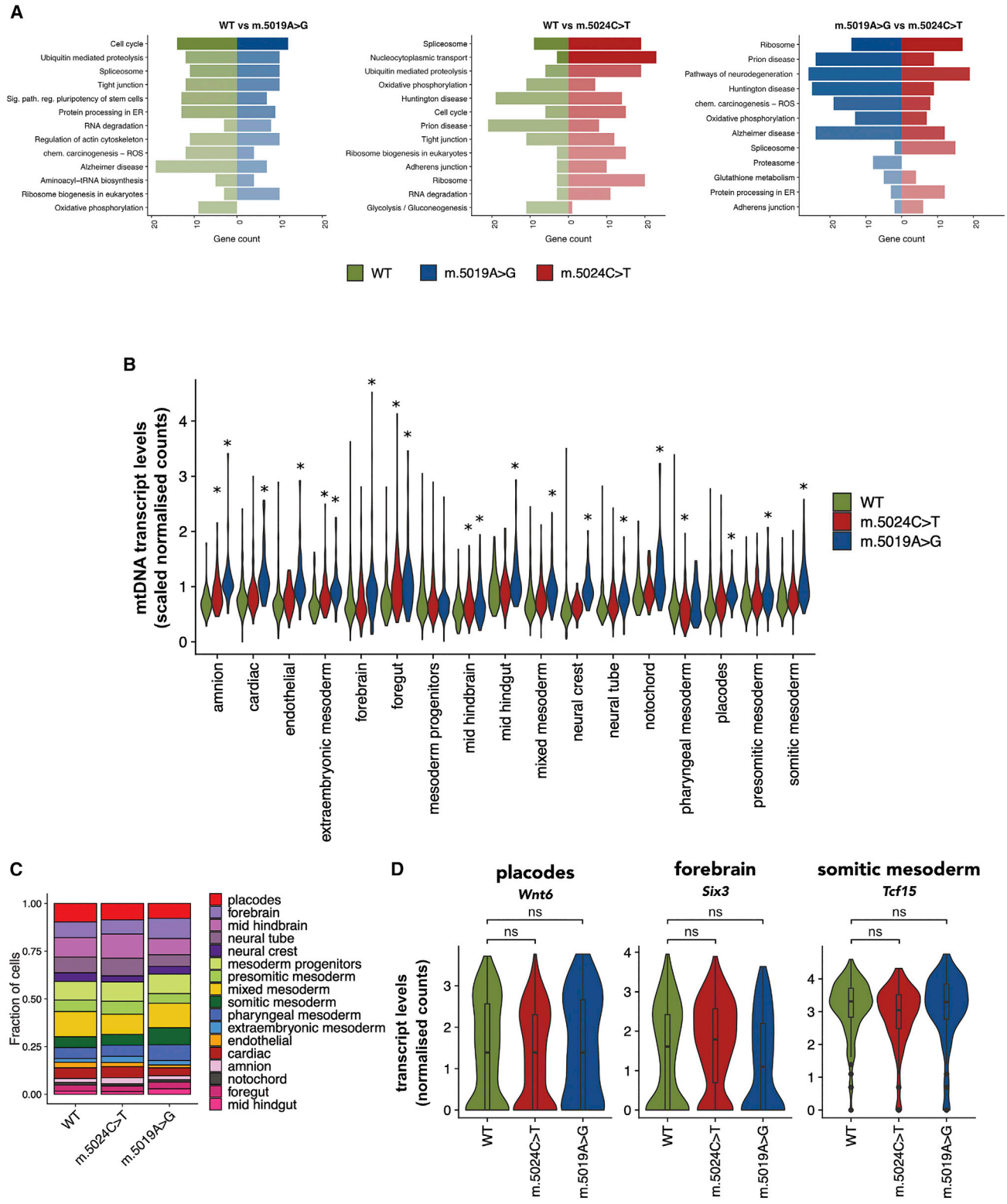


Figure 3. *mt-Ta* variants induce an embryo-wide transcriptional response

(A) Bar charts showing KEGG pathways enriched (adjusted $p < 0.05$) for transcripts with significant differential expression in the three pairwise comparisons of WT, m.5024C>T, and m.5019A>G. Bar length indicates the number of upregulated pathway genes for the corresponding strain in each comparison.

(legend continued on next page)

(Figure S3E) from ~2 years of age. Physiological studies showed that the m.5019A>G mutation, like the m.5024C>T mutation, results in less body fat than WT mice of the same nuclear genetic background (Figures S3F–S3G).

Mt-Ta variants induce an embryo-wide transcriptional response

E8.5 embryos with high heteroplasmy levels were isolated from m.5024C>T and m.5019A>G mothers. ~10,000 live cells were isolated from six m.5024C>T embryos (mean heteroplasmy level 69%, range 41%–93%, Figure S4A), with 4,447 cells passing QC for 10X sequencing (Table 1). Allele-specific transcript levels of the linked m.13715C>T mutation in m.5024C>T embryos[21] confirmed similar levels across the 17 tissue lineages (Figure S4B), in keeping with measurements in neonates where different tissues all have comparable heteroplasmy levels (Figure S3C). ~5,000 live cells were collected from three m.5019A>G embryos with a mean heteroplasmy level of 86% (range 72%–94%, Figure S4A), with 1,649 cells passing QC (Table 1). Heteroplasmy values were validated at the transcript level using the linked m.15200A>G *Cytb* mutation,[28] as for the m.13715C>T mutation in m.5024C>T embryo cells (Figure S4B), again showing similar levels across all lineages, as seen in neonatal m.5019A>G animals (Figure S4C). UMAPs generated using all 11,020 cells showed near identical clustering patterns for m.5019A>G, m.5024C>T, and WT cells (Figure S4D), and marker gene analysis identified the same 17 cell lineages in both mutant strains as seen in WT embryos (Figure S4E). Pseudobulk analysis of both m.5024C>T and m.5019A>G embryos showed dysregulation of genes involved in diverse intracellular pathways (Figure 3A), including changes in expression of nuclear-encoded mitochondrial genes[3] in both mice (Figure S4F; Table S2). These included the downregulation of transcripts involved in tRNA aminoacylation, the electron transport chain, and ATP synthesis when compared to WT embryos (Figure S4F), in keeping with the known defect of intra-mitochondrial translation.[21] We also saw an upregulation of mtDNA transcripts, which can be seen as a response to the mtDNA mutations in all cell lineages (Figure 3B).

Despite the observed OXPHOS defect defects in adult mice harboring the m.5024C>T[21] and m.5019A>G mutations, we saw only minor differences in the number of cells committed to each cell lineage at E8.5 (Figure 3C), no difference in the gene expression profiles that defined each lineage (Figure 3D), and no obvious developmental abnormalities originating during organogenesis (Figure S4G). We therefore sought evidence that the observed transcriptional changes were mediating an effective compensatory response.

Intra- and extra-organellar buffering in response to the *mt-Ta* mutations

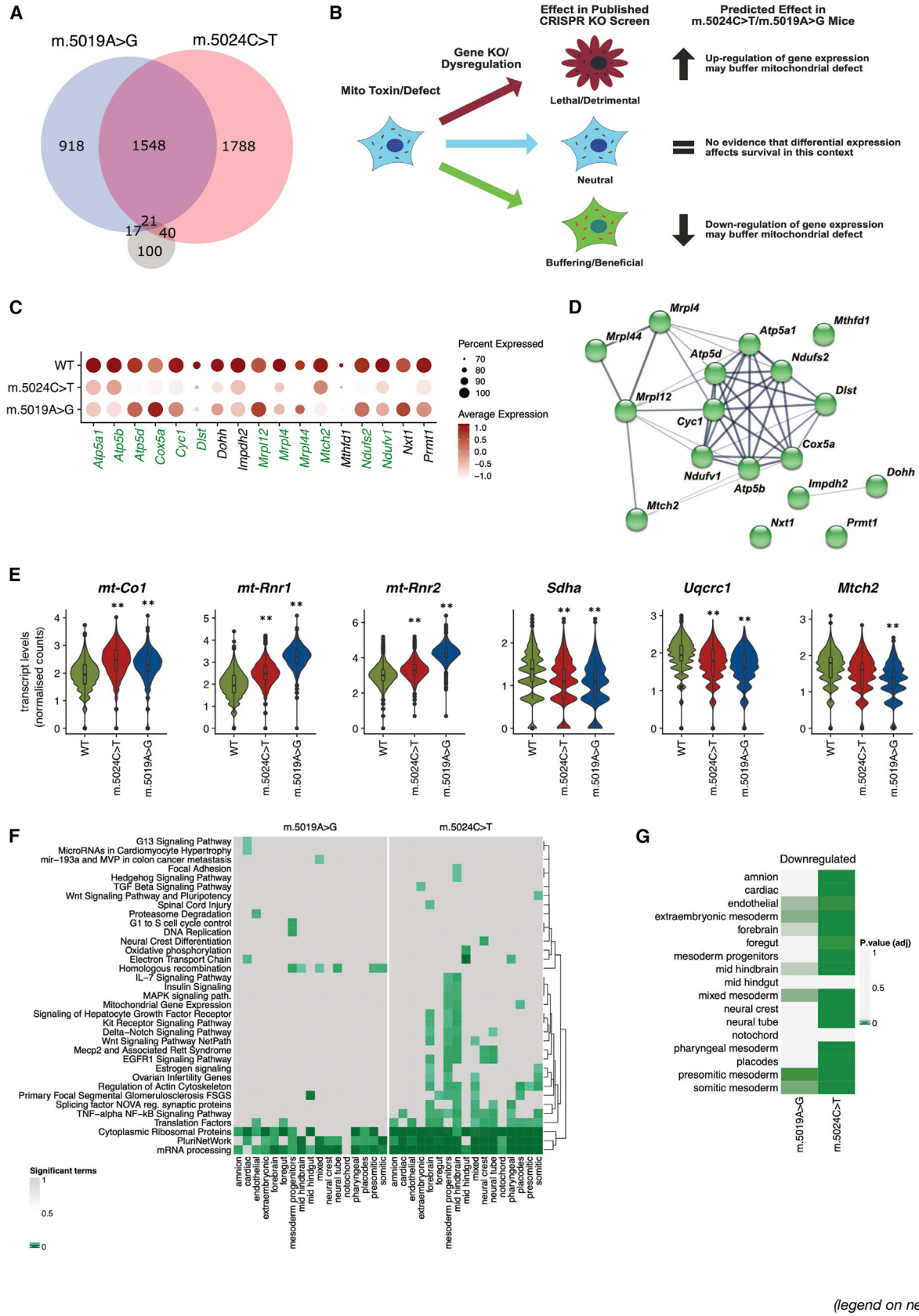
Genome-wide *in vitro* CRISPR screens using single molecule inhibitors of specific OXPHOS complexes identified 192 distinct genetic modifiers, including 38 knockouts that were synthetic sick/lethal, 91 that acted as epistatic buffers of the toxin, and 63 that suppressed the toxic effects of specific respiratory chain poisons.[9] We identified 178 mouse homologs from this list of 192 genetic modifiers, of which sixty-one (34%) were dysregulated in at least one cell lineage in m.5024C>T mice (hypergeometric p(adj) = 2.07 x 10^{-6}); and thirty-eight (21%) were dysregulated in at least one cell lineage in m.5019A>G mice (p(adj) = 4.61x10^{-2}). Twenty-one (12%) of the 178 genetic modifiers were dysregulated in both mutants (Figure 4A; Table S3).

The synthetic sick/lethal genes preferentially killed or suppressed cell proliferation when knocked out *in vitro* before exposing cells to a specific OXPHOS complex inhibitor.[9] By contrast, knocking out epistatic buffer genes prevented the effects of a specific OXPHOS inhibitor, and knockout of suppressor genes enhanced cell growth in the presence of the exogenous mitochondrial toxin. By analogy, we hypothesized that synthetic sick/lethal genes would either not be affected or would have increased expression in m.5024C>T and m.5019A>G mice, because if expression were reduced, this would impact cell lineage development. Conversely, epistatic buffer and suppressor genes would be under-expressed because knockout of these genes *in vitro* prevented the toxic effects of OXPHOS inhibitors (Figure 4B).

In line with this prediction, we saw significant enrichment for epistatic buffer/suppressor genes with reduced expression in *mt-Ta* mice compared to WT (hypergeometric *p(adj)* = 2.25 x 10^{-13} for m.5024C>T; 3.29 x 10^{-4} for m.5019A>G Figure 4C), predominantly involving predicted mitochondrial proteins (*p(adj)* = 2.38 x 10^{-30} (m.5024C>T); 3.41 x 10^{-9} (m.5019A>G)) (Figure 4C). These formed a coherent interaction network[29] involving intrinsic OXPHOS and related proteins centralized around *Atp5d* and *Mrpl12* and the down-regulation of nuclear-encoded OXPHOS complex I, III, and IV subunits, mitochondrial ribosomes, and ATP-synthase (Figure 4D). Several non-mitochondrial or "extrinsic" proteins were also involved in epigenetically regulating nuclear gene expression (*Prmt1*), nucleotide synthesis (*Impdh2*), and one-carbon metabolism (*Mthfd1*) (Figure 4C).

Our findings validate the importance of the candidate genetic modifiers identified through *in vitro* CRISPR screens by showing they have biological relevance during embryonic development and organogenesis *in vivo*. A further genome-wide *in vitro* CRISPR screen using IMT1, an inhibitor of mitochondrial transcription,[30] identified 526 genetic modifiers putatively involved in the cellular response to mitochondrial dysfunction.[31] Mouse homologs of many of these genetic modifiers were also dysregulated in m.5024C>T (25%, p(adj) = 3.55 x 10^{-15}) and m.5019A>G (20%, p(adj) = 1.06 x 10^{-13}) embryos (Figure S4H; Table S4), including numerous genes involved in the mTORC1 pathway (*Mtor*, *Rictor*, *Lamtor4*, *Rraga*, and *Rpl11*). Thus, the combined *in vitro* data indicates that differential regulation of these genetic modifiers in m.5024C>T and m.5019A>G embryos

(B) Violin plots showing mtDNA transcript expression levels in WT, m.5024C>T, and m.5019A>G mouse cells across the 17 identified cell lineages. *p < 0.05 (mutant vs. WT).
(C) Bar chart displaying the fraction of cells committed to each cell of the 17 identified cell lineages in WT, m.5019A>G, and m.5024C>T mouse embryos at E8.5.
(D) Violin plots displaying the transcriptional profile of selected lineage-specific marker genes *Wnt6*, *Six3*, and *Tcf15* in their respective cell types in WT, m.5024C>T, and m.5019A>G mice, showing no significant differences in expression of these lineage-defining genes. WT, wild-type.
See also Figure S4.



(legend on next page)

likely reflects a positive compensation, rather than a passive response to the reduced synthesis of mtDNA encoded subunits, although we cannot exclude other parallel mechanisms.

Mutation-specific and cell lineage-specific transcriptional signatures

Pseudobulk analysis also revealed mutation-specific differential expression of mitochondrial genes at the level of the whole embryo (Figure 4E) and in pooled cells from specific lineages (Figure S5A; Table S5). A corresponding WikiPathway analysis (Figure 4F) showed differential transcription of genes involved in mRNA processing, ribosomal pathways, and the pluripotency network across numerous cell lineages in both mutant strains, with multiple additional pathways being differentially regulated in just one or a few lineages. A similar analysis using GO Biological Processes produced comparable results (Figure S5B), highlighting the importance of both embryo-wide and lineage-specific transcriptional changes in response to the mitochondrial mutations. Next, we used our mouse models to investigate mutation-specific and lineage-specific expression patterns of the genetic modifiers identified in genome-wide CRISPR screens.[9] Compared to WT embryos, downregulated genes in both mutant lines were enriched for epistatic buffer/suppressor genes, although this signature was much stronger across multiple cell-lineages in m.5024C>T embryos than in m.5019A>G embryos, where fewer cell types showed strong enrichment (Figure 4G). This pattern was also seen in the global pathway analysis (Figure 4F) and in the initial marker gene analysis, where in addition to the downregulated ameliorating genes (Figure 4C), we saw a small but significant enrichment of epistatic buffer/suppressor genes that were upregulated (hypergeometric p(adj) = 0.0157) in m.5019A>G embryos, a finding that was not apparent in m.5024C>T embryos (hypergeometric p(adj) = 0.186). Taken together, these findings show evidence of a stronger compensatory down-regulation of epistatic buffer/suppressor genes in m.5024C>T embryos compared to m.5019A>G and could potentially explain the reduced mean litter size of 5 pups for m.5019A>G mice, compared to 6 pups per litter for m.5024C>T and WT mice (Figure S5C). For m.5019A>G mice, higher maternal heteroplasmy levels were associated with a decrease in litter size, which was not seen in m.5024C>T mice (Figure S5D). Although this could be explained by maternal factors influencing the pregnancy, we also saw a higher pre-weaning mortality for m.5019A>G pups, compared to m.5024C>T pups (Figure S5E), in keeping with a sub-optimal compensatory response to the m.5019A>G mutation within the pups themselves.

The scaled expression of individual epistatic buffers/suppressors across different cell types revealed lineage-specific expression patterns for most of the genes (Figures 5A and S6A; Table S6). For example, the buffering gene *Nr2f2*[9] was more strongly downregulated in mesodermal, neural tube, and neural crest lineages in m.5024C>T than in m.5019A>G mice. *Nr2f2,* also known as chicken ovalbumin upstream promoter transcription factor II (*COUP-TFII*), plays a key role in regulating the development of several organs including blood vessels derived from mesoderm[32] and specific neuronal types in the brain.[33,34] On the other hand, multiple OXPHOS subunit transcripts were preferentially upregulated inm.5019A>G mice, involving complexes I (*Ndufa6*, *Ndufa10*, *Ndufb3*, *Ndufb5*, *Ndufb7*, *Ndufb9*, *Ndufb10*, *Ndufb11*, *Nduf11*, *Ndufs5*, and *Ndufs8*), IV (*Cox5a* and *Cox14*), and 5 (*Atp5b*, *Atp5d*, and *Atp5e*) particularly in cardiac precursors. On first consideration, this upregulation could be seen as a beneficial, akin to the "ragged-red" or "sick mitochondrion" response to an intra-mitochondrial translation defect.[35] However, all were previously defined as epistatic buffer/suppressor genes,[9] so their up-regulation is likely to be disadvantageous in m.5019A>G mice (Figure 4B). Interestingly, we did not see coordinated regulation of genes involved in the integrated mitochondrial stress response (ISRmt) in either mutant strain (Figure S6B), despite the ISRmt being implicated in the cellular response to metabolic stress.[36] However, this is consistent with previous reports that transcription of Atf4, a key regulator of the ISRmt,[37] is not increased in heart tissue of adult m.5024C>T mice.[38] Taken together, these findings (Figure 5A) show mutation and lineage-specific differences in transcriptional compensatory responses that are likely to contribute to organ specific phenotypes.

Figure 4. Intra- and extra-organellar buffering in response to the *mt-Ta* mutations

(A) Venn diagram showing the intersection between published genetic modifiers of mitochondrial dysfunction identified through an *in vitro* CRISPR screen (gray)[9] and dysregulated genes in m.5019A>G (blue) and m.5024C>T (red) embryos. A full list of these intersecting genes with corresponding p values can be found in Table S3.

(B) Schematic showing the predicted concordance between CRISPR knockout screens using OXPHOS inhibitors and pathogenic mtDNA mutations in E8.5 embryonic cells, highlighting the potential buffering effect of downregulating genes whose knockout in the *in vitro* CRISPR screen was beneficial.

(C) Dot plot showing relative expression levels of 17 epistatic buffer/suppressor genes identified in the *in vitro* CRISPR screen that were significantly downregulated in at least one mutant strain when compared to WT. Genes labeled in green are annotated in the Mouse MitoCarta3.0 inventory.

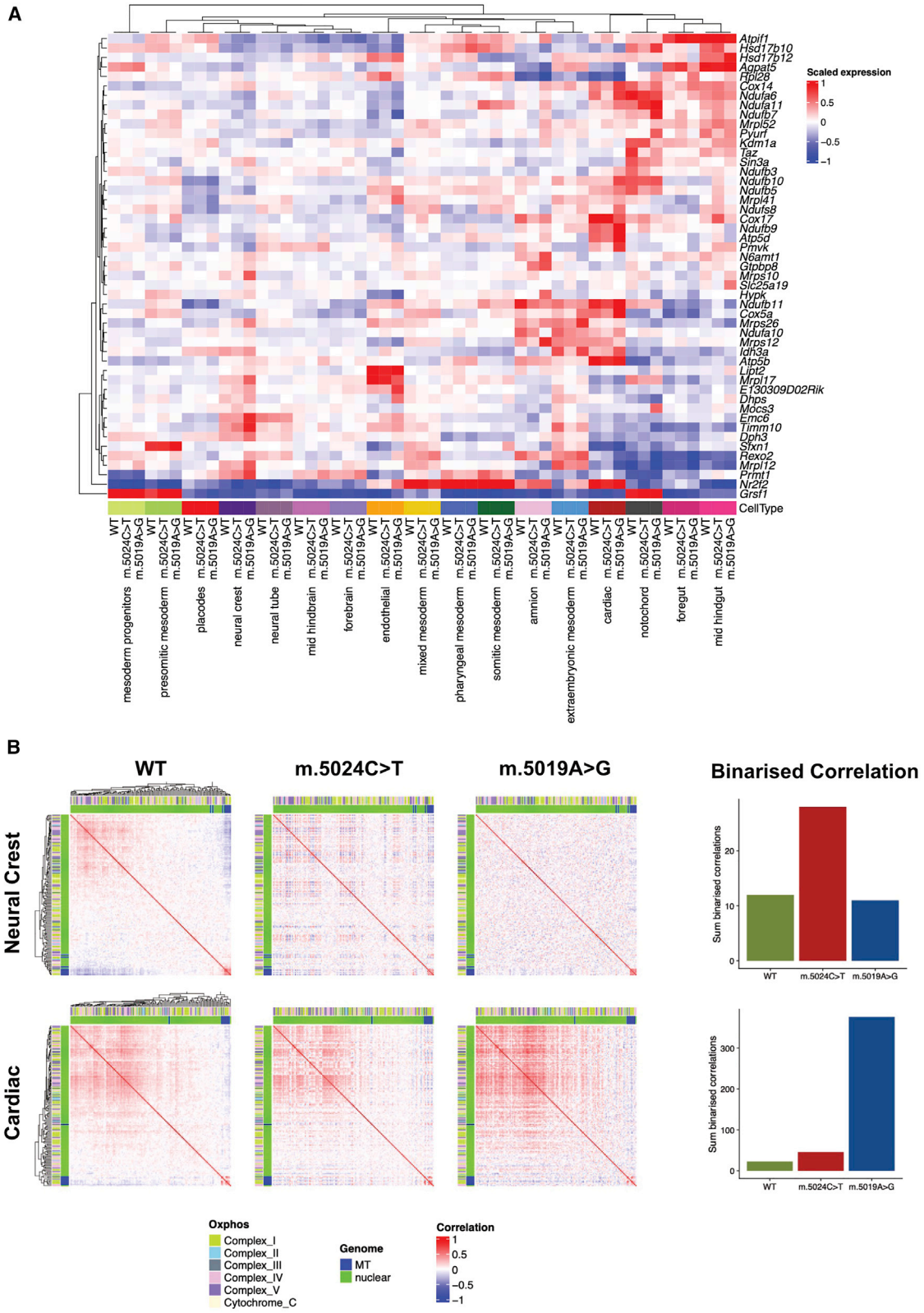
(D) STRING protein interaction network showing genes from (E); the core network comprises several nuclear-encoded OXPHOS subunits and components of the mitochondrial ribosome. Network generated by STRING v11. Minimum required interaction score = 0.4. Thickness of connecting lines indicates degree of confidence in prediction: thicker line corresponds to higher confidence of interaction.

(E) Violin plots showing pseudobulk expression profiles of illustrative mtDNA and nuclear gene transcripts encoding mitochondrial proteins in WT, m.5024C>T and m.5019A>G embryos. **p < 0.01. Full dataset shown in Table S2.

(F) Heatmap showing significantly enriched WikiPathways (2019 mouse annotation), based on differentially regulated genes in at least one cell lineage in m.5019A>G and m.5024C>T embryos compared to WT. mRNA processing, pluripotency, and ribosomal protein genes are enriched across multiple lineages in both mutants, whilst other pathways show lineage- and strain-specific enrichment. p < 0.05 for all green squares.

(G) Heatmap showing enrichment of identified epistatic buffers/suppressors amongst genes that are downregulated in m.5019A>G and m.5024C>T mutants compared to WT across all cell lineages. m.5024C>T embryos show a stronger downregulation signature than m.5019A>G embryos, perhaps representing a stronger compensatory response to mitochondrial dysfunction. p < 0.05 for all green bars. WT, wild-type; CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats; OXPHOS, oxidative phosphorylation.

See also Figures S4 and S5.



(legend on next page)

Finally, we looked for evidence of a coordinated response involving transcripts coding for the same OXPHOS complex. In WT mice, the co-expression of transcripts was most obvious in the more advanced tissue lineages, particularly the cardiac and neural crest (Figure 5B). In some cell types, the degree of co-expression between nuclear transcripts was stronger in the mutant mice, but this was accompanied by a weaker correlation between nuclear-encoded and mtDNA-encoded transcripts (Figure S6C), implying a disruption of nuclear-mtDNA communication. This was most marked in m.5019A>G mice.

Cell lineage-specific compensatory mechanisms are coordinated by gene regulatory networks

To determine the upstream mechanism responsible for the compensatory responses, we performed *cis*-regulatory analysis to derive gene "regulons" from the entire dataset using single-cell regulatory network inference and clustering (SCENIC)[39] and studied their expression in 17 cell lineages in m.5019A>G, m.5024C>T, and WT E8.5 mice. Each regulon defined a gene regulatory network co-regulated by a known transcription factor coding for a gene repressor or activator. We generated a regulon activity matrix through unbiased hierarchical clustering, which grouped the three mouse strains by cell lineage (Figure S7A). Relative scaling of regulon activity in each cell type revealed several regulons with mutation-specific activity patterns across multiple cell lineages (Figure 6A; Table S7), as well as lineage-specific activation of many regulons (Figure 6A; Table S7), often linked to genes involved in development of that tissue type. For example, we found increased activity of regulons linked to the GATA transcription factors, including *Gata4*, *Gata5*, and *Gata6* in developing cardiac cells and *Foxa3* in gut progenitor cells (Figures 6B and S7B), as seen previously.[40,41] To further validate this approach, we compared the genes in our SCENIC-derived regulons to previously published *in vitro* RNA-seq data from human and mouse neural crest-derived cell lines with siRNA knockdown of the transcription factor *Sox10*.[42,43] We saw significant overlap between genes perturbed by *Sox10* knockdown and several of the neural crest-specific regulons identified in our SCENIC analysis, including *Sox10*, *Sox9*, *Tfap2a*, and *Tfap2c* (Figure S7C). Several transcription factors were more active in WT than *mt-Ta* mice, including *Mecom* in the endothelial lineage, which is involved in arterial endothelial cell differentiation and proliferation.[44] Conversely, *Sox9* and *Sox10* showed higher activity in neural crest cells (NSCs) in the mutant strains than in WT mice (Figure 6C). *Sox9* and *Sox10* are essential for the migration and differentiation of neural crest cells in the developing embryo,[45] and were linked to the co-expression of *Ets1*, *Sox5*, *Tfapb2*, *Zeb2*, *Erbb3*, *Cdh6*, and *Cdh11* in both mutant strains (Figure 6D). These genes are involved in neural crest specification and the epithelium-to-mesenchyme transition, a key step in neural crest migration which commences at the 4–5 somite stage, corresponding to E8.0–8.5,[46,47] building on recent findings that mitochondrial function plays an important role in specification and differentiation of the neural crest lineage.[48]

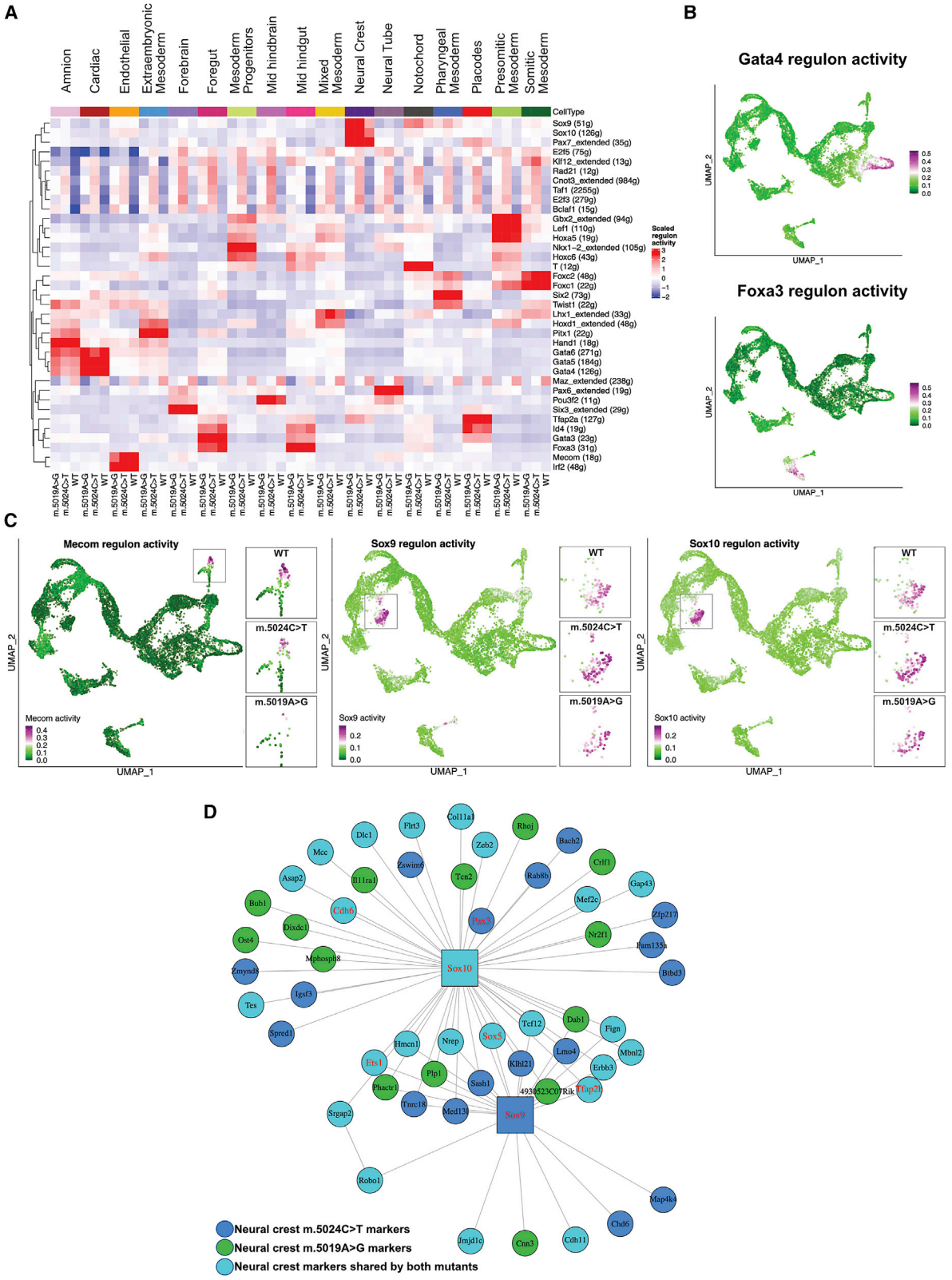
To provide independent experimental validation of our key findings, we sub-cloned mouse embryonic fibroblasts (MEFs) from m.5024C>T mice that had shown several regulons with increased activity across multiple cell-lineages *in vivo* (Figures 6A and S7A). By generating different clones with stable heteroplasmy levels from 20% to 90% (Figure S7D), we established an *in vitro* system allowing us to evaluate the impact of knocking down specific gene regulons on an isogenic nuclear background. To validate the *in vitro* model, we performed bulk RNA-seq on 10 clonal MEF lines. Principal component analysis (PCA) clustered MEF clones based on their heteroplasmy level (Figure 7A), and we saw evidence of widespread differential gene expression between low and high heteroplasmy MEF clones (Figure 7B), supporting our hypothesis that transcriptional changes play a key role in the adaptive response to mitochondrial dysfunction. Pathway enrichment analysis revealed differential regulation of genes in many of the same pathways observed in the single cell data (Figures 4F and 7C), despite the MEFs being a different tissue lineage established at a later developmental stage and being maintained *in vitro.*

To validate the role of the gene regulatory networks identified in m.5024C>T E8.5 embryos (Figure 6A), small interfering (si) RNAs were designed and validated using qPCR against 5 transcription factors, *E2f3*, *Taf1*, *Klf12*, *Maz*, and *Bclaf1*, controlling SCENIC regulons that were differentially regulated across multiple cell lineages in m.5024C>T vs. WT E8.5 embryos *in vivo* (Figure 6A). These were chosen over cell type-specific regulons, such as *Sox9*, *Sox 10*, and *Mecom*, on the basis that lineage-specific effects were less likely to be reproducible in the m.5024C>T MEF cells. Knockdown of *E2f3*, *Taf1*, *Maz*, and *Bclaf1* suppressed MEF cell proliferation, with the greatest impact on high heteroplasmy clones (Figure 7D). siRNA against *Klf12* had no impact on cell proliferation likely because, unlike *in vivo* cell lineages, this gene had very low expression levels in untreated MEFs. Given that cell proliferation rates are widely recognized as being a sensitive indicator of OXPHOS dysfunction[49,50] and that cell proliferation is essential for development, these findings support our main conclusion: that lineage-specific transcriptional responses act via coordinated gene networks controlled by specific transcription factors in response to mitochondrial dysfunction.

Figure 5. Mutation- and cell lineage-specific differential expression of epistatic buffer/suppressor genes and comparison between expression profiles

(A) Heatmap showing relative expression levels of 51 epistatic buffer/suppressor genes identified in CRISPR knockout screens, filtered to show genes that exhibit differential expression patterns between WT and m.5024C>T or m.5019A>G. Filters applied on scaled gene expression: minimum difference between mutant - WT < −0.2, and maximum for the scaled gene expression >0.5 for at least 1 cell type, across all cell lineages and mouse strains.

(B) Correlation between individual gene expression levels of nuclear and mtDNA-encoded mitochondrial transcripts (MitoCarta v3.0) for the neural crest and cardiac lineages. Left, correlation matrices for WT, m.5024C>T, and m.5019A>G mice. Outer bar on each axis shows the OXPHOS subunits corresponding to each gene. Inner bar on each graph axis whether the genes are expressed from nuclear DNA or mtDNA. Right, sum of the binarized correlations from each matrix where a correlation with $r^2 > 0.5$ is counted. Overall, the figure shows that more correlated gene expression in neural crest cells for m.5024C>T and more correlated gene expression in cardiac cells for m.5019A>G mice. Additional lineages are shown in Figure S6C for forebrain, mixed mesoderm, and foregut. WT, wild-type.
See also Figure S6.



(legend on next page)

DISCUSSION

Defects of intra-mitochondrial translation are the most common cause of rare inherited mitochondrial disorders, which affect ~1 in 8,000 humans.[51] Although these diseases share a common biochemical defect of OXPHOS, there is considerable clinical heterogeneity. Strikingly different phenotypes are seen in patients with closely related mutations affecting the mtDNA-encoded tRNA genes,[52,53] but the reasons for this are not known. The different properties of mitochondria found in different cell types has been proposed as an explanation for patterns of organ involvement in rare inherited mitochondrial diseases,[16,54] but when and how this occurs is not clear. Here we show that mice with pathogenic *mt-Ta* mutations only 5 bp apart induce distinct nuclear transcriptional responses in specific cell lineages by E8.5 during embryonic development. This raises the possibility that similar mechanisms contribute to tissue selectivity, which characterizes different human mitochondrial disorders and are at play during development *in utero*. Nuclear genetic factors required for mitochondrial biogenesis are regulated at the transcript level,[55] contributing to inter-tissue differences in mitochondrial activity.[56] Thus, the lineage-specific transcriptional signatures we saw in WT mice are likely reflected in distinct mitochondrial proteomes at the cellular level, in keeping with our observation of tissue-specific isoform transcripts which emerge before organ maturation.

Signals from the mitochondrion to the cell nucleus are poorly understood but are likely to occur at multiple levels. In addition to ATP, the cellular redox state (reflected in the NADH/NAD^+ ratio) and organellar calcium release likely play a role in different contexts.[57] Tricarboxylic acid intermediates are also implicated, including fumarate, succinate, and α-ketoglutarate, which can modulate nuclear gene expression through epigenetic modifications. Amino acid and one-carbon metabolism may also be involved, as may the mitochondrial unfolded protein response (UPRmt) and ISRmt.[36] However, it is not clear how different mtDNA mutations could induce different nuclear transcriptional response. As in humans, our findings in *mt-Ta* mice cannot be explained by a shared OXPHOS defect, emphasizing the importance of defining the underlying upstream molecular mechanisms: m.5024C>T compromises the stability of *mt-Ta*,[21] whereas this study reveals that m.5019A>G affects tRNA charging, likely due to disruption of G:U pairing in the anticodon stem.[58] Here we show they impact on pathways intrinsic and extrinsic to mitochondria in a cell lineage-specific manner.

It is striking how many of the genes downregulated in response to the *mt-Ta* mutations were also identified in *in vitro* screens as knockouts that buffered or rescued cell lines from mitochondrial toxins,[9] including an inhibitor of intra-mitochondrial transcription.[31] This independent corroboration implicates the same mechanisms in tissue-specific compensatory responses before birth, potentially resolving the long-standing paradox as to why severe OXPHOS diseases present in neonatal life or later. We show that many of the buffering mechanisms are under the control of specific transcriptional factors, raising the intriguing possibility of delivering systemic treatments against known druggable targets which specifically influence one tissue and not another. The entire dataset is available as a resource to accelerate mitochondrial target discovery (Table S7).

Given the emerging evidence implicating the UPRmt and related ISRmt in the pathogenesis of mitochondrial disorders,[36] we were surprised not to see strong upregulation of ISR transcripts in any of the cell lineages in E8.5 *mt-Ta* mice (Figure S6B). However, only a mild signature of the UPRmt/ISRmt was observed in mice with complex IV deficiency due to homozygous Surf1 mutations,[59] and there was no activation of the UPRmt in mice with cardiomyopathy caused by loss of CLPP.[60] Thus, the presence of an overt OXPHOS defect does not necessarily trigger the UPRmt/ISRmt in mice, despite clear biochemical defects in affected tissues. Recent evidence has cast light on the temporal dynamics of the ISRmt, defining three distinct phases involving local and systemic effects that take time to evolve and are partly dependent on FGF21.[61,62] Based on these findings, it possible that that our observations at E8.5 precede activation of the ISRmt. If correct, this makes the mechanisms that we have identified even more attractive targets to treat or perhaps prevent the onset of symptoms in mitochondrial diseases, including TFs linked to stress responses distinct from the classical ISR (Figure S7A). Mitochondria also play a role in common late onset diseases involving similar organ systems, leading to type 2 diabetes mellitus and Parkinson disease.[63] Our observations are therefore likely to have much broader therapeutic relevance than rare inherited disorders of mitochondrial metabolism.

Limitations of the study

Our principal findings relate to one embryonic stage (E8.5) in mice with two *mt-Ta* mutations. It will be important to extend this work to include earlier developmental stages to establish when the compensatory responses first emerge and to

Figure 6. Cell lineage-specific compensatory mechanisms are coordinated by gene regulatory networks

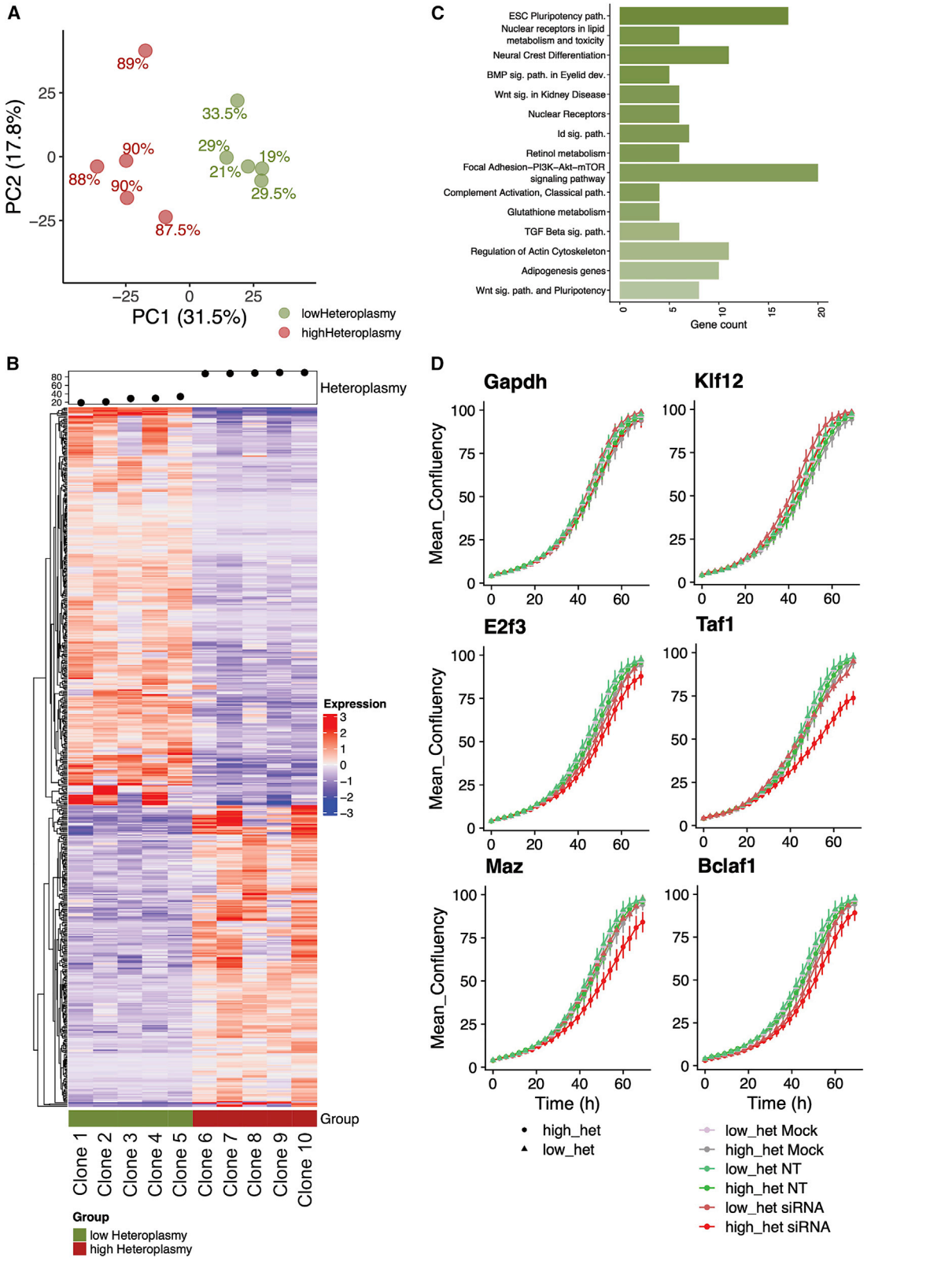
(A) Heatmap of SCENIC regulon activity across all cell lineages in WT, m.5024C>T, and m.5019A>G mice, showing selected lineage-specific and mutation-specific transcriptional activation patterns. Number in brackets is the number of genes (g) identified with high confidence in each regulon. Full regulon activity dataset can be found in Table S7.

(B) UMAPs showing localized activity of the heart-specific Gata4 transcription factor in the cardiac lineage and Foxa3 activity, required for liver development from gut endoderm, in the intestinal lineages. WT, m.5024C>T, and m.5019A>G cells are overlaid on the same UMAP. Gray/pink indicate increased regulon activity.

(C) UMAPs showing differential lineage-specific activity of the Mecom (endothelial, less active), Sox9 (neural crest, more active), and Sox 10 (neural crest, more active) regulons in m.5024C>T and m.5019A>G mice compared to WT. In the main UMAP, WT, m.5024C>T, and m.5019A>G cells are overlaid, with each mouse strain displayed separately in the expanded panels.

(D) Network displaying differentially expressed marker genes identified in the Sox9 and Sox 10 regulons in the neural crest lineage in m.5024C>T and m.5019A>G mice, with just under half of the identified marker genes (24/56) being shared by both mutant strains. Lines between circles and boxes indicate the regulon(s) associated with that gene. Lines between circles indicate predicted protein-protein interactions. Red text indicates genes involved in neural crest differentiation. WT, wild-type; UMAP, Uni-form Mani-fold Approximation and Projection.

See also Figure S7.



(legend on next page)

determine whether they persist or change in later prenatal stages and throughout life. Studying other mtDNA mutations and other species will establish whether our observations have more general relevance. This will be particularly challenging in humans for ethical reasons. Although transcript abundance plays an important role in regulating mitochondrial activity,[55,56] other compensatory mechanisms may be involved, including at the post-transcriptional level,[64] and these should be explored further. In addition, given the relationship between cell size, mtDNA content, and mtDNA-encoded transcript abundance, it would also be interesting to explore whether cell size is important during development and forms part of the compensatory response, although this could be confounded by cell cycle effects.

Although not directly relevant to our main conclusions, the precise molecular mechanism of the m.5019A>G mt-Ta mutation would benefit from further interrogation. Unlike the m.5024C>T mice, our preliminary analyses showed no evidence of similar defects in organelle translation of mitochondria isolated from m.5019A>G mice. Despite alterations in rRNA abundance, quantitative assays of proteins involved in mitochondrial ribosome structure and assembly showed no alterations in mitochondrial ribosomes from adult tissues, but a direct investigation of mitochondrial ribosome assembly was not performed.

Finally, although we provided independent validation for transcription factors that were down-regulated across multiple tissues *in vivo*, it will be important to carry out similar validation experiments for all of the compensatory transcription factors, and in specific cell types. This will enable a further dissection of downstream molecular mechanisms that are also likely to be tissue specific. Definitive proof will require *in vivo* validation. This could be achieved through targeted gene or protein disruption, but will be very challenging to perform during development and at the single cell level. Targeted small molecule approaches or single cell multi-omics provide a way forward.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animals
  - Cell lines
- METHOD DETAILS
  - Identification of the m.5019A>G strain
  - mtDNA heteroplasmy measurements
  - In vitro transcription assays
  - Measurement of aminoacylation status
  - Histology
  - RNA isolation and northern blot analysis
  - Western Blot analysis
  - Southern blots
  - Preparation of E8.5 embryos for scRNA-seq
  - Single-cell heteroplasmy measurements
  - 10X chromium scRNA-seq
  - scRNA-seq & gene expression analysis
  - Proteome profiling of m.5019A>G and WT adult mouse tissues
  - TMT labeling
  - High-pH first dimension reverse-phase fractionation
  - LC-MS/MS
  - Proteomic data analysis
  - Single-cell cloning of immortalized m.5024C>T MEF cells
  - Bulk RNA-seq of m.5024C>T MEF clones
  - siRNA knockdown of selected transcription factors in m.5024C>T MEF clones
  - Cell proliferation analysis
  - Additional datasets
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2023.01.034.

Figure 7. Knockdown of key *in vivo* regulon transcription factors impairs cell proliferation in a heteroplasmy-dependent manner *in vitro*

(A) PCA of bulk RNA-seq data from ten clonal populations of m.5024C>T MEF cells showing differential clustering of clones based on heteroplasmy level. Low heteroplasmy <35%, high heteroplasmy >80%. Clones were selected from a larger group shown in Figure S7D.

(B) Heatmap of genes showing significant differential expression between low and high heteroplasmy m.5024C>T MEF clones, highlighting the widespread transcriptional differences between the two groups.

(C) Bar chart showing WikiPathways (2019 mouse annotation) significantly enriched for genes differentially expressed between low and high heteroplasmy m.5024C>T MEF clones. Several of these, including various signaling and pluripotency pathways, overlap with those seen in m.5024C>T E8.5 embryonic cell lineages (Figure 4F).

(D) Growth curves showing proliferation rates for low and high heteroplasmy m.5024C>T MEF clones that were left untreated (mock), transfected with a non-targeting siRNA pool (NT), or transfected with an siRNA pool to the indicated gene (siRNA). Mock and NT treated clones showed no reduction in proliferation rate, as was the case for clones targeted with siRNA to *Gapdh* (control housekeeping gene) and *Klf12*. In clones transfected with siRNA to *E2f3*, *Taf1*, *Maz*, and *Bclaf1*, there was reduced proliferation that was more marked in high heteroplasmy clones. Plots show combined proliferation data from 5 low heteroplasmy and 5 high heteroplasmy clones gathered from two independent siRNA knockdowns per clone. Error bars indicate SEM.

PCA, principal component analysis; MEF, mouse embryonic fibroblast; siRNA, small interfering RNA.

See also Figure S7.

ACKNOWLEDGMENTS

We are grateful to the Cambridge Institute for Medical Research Flow Cytometry Core Facility for access to flow sorting facilities; Katarzyna Kania and the Cancer Research UK Cambridge Institute Genomics Core Facility for 10X single-cell RNA-seq; Keira Turner (Mitochondrial Biology Unit) and staff at the Anne McLaren Building for animal husbandry; The Wellcome-MRC Cambridge Stem Cell Institute NGS Library Facility for bulk RNA-seq library preparation; The Cambridge Center for Proteomics for quantitative proteomic analysis; David Ron for pBSSVD2005; The Max Planck for Biology of Aging FACS & Imaging Core Facility for support; graphical abstract created with BioRender with help from Brandon Lockey. PFC is a Wellcome Trust Principal Research Fellow (212219/Z/18/Z), and a UK NIHR Senior Investigator, who receives support from the Medical Research Council Mitochondrial Biology Unit (MC_UU_00028/7), the Medical Research Council (MRC) International Centre for Genomic Medicine in Neuromuscular Disease (MR/S005021/1), the Leverhulme Trust (RPG-2018-408), an MRC research grant (MR/S035699/1), an Alzheimer's Society Project Grant (AS-PG-18b-022). PS was supported by Marie Sklodowska-Curie ITN-REMIX [721757]. This research was supported by the NIHR Cambridge Biomedical Research Center (BRC-1215-20014) and the Knut and Alice Wallenberg Foundation (KAW 2017.0080 and Wallenberg Scholar to M. Falkenberg), and funding from the Max-Planck Gesennschaft & United Mitochondrial Disease Foundation grant/award number: 13-053R to J.B.S. This study was supported by the Swedish Research Council (2018-02439 to M. Falkenberg), the Swedish Cancer Foundation (22 2385 Pj to M. Falkenberg), the Knut and Alice Wallenberg Foundation (KAW 2017.0080 Wallenberg Scholar to M. Falkenberg), and grants from the Swedish state under the agreement between the Swedish government and the county councils, the ALF agreement (ALFGBG-727491 to M. Falkenberg). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

AUTHOR CONTRIBUTIONS

S.P.B. and A.G. performed all embryo and *in vitro* MEF experiments; F.K. and M.P. were responsible for data analysis and bioinformatics; P.S., P.N., M.-L.S., N.A.B. and H.D. performed characterization experiments for the m.5019A>G mouse model; L.V.B. performed analysis of litter size data; A.G., M. Frison, and Y.N. conducted adult mouse proteomic experiments; S.P.B., P.F.C., and J.B.S. wrote the paper; S.P.B., F.K., A.G., M.P., J.B.S. and P.F.C. planned experiments and discussed the data; F.K., M.P., M. Falkenberg, M.M. and N.J. assisted with manuscript preparation; M. Falkenberg, M.M., N.J., J.B.S. and P.F.C. supervised the project; J.B.S. and P.F.C. initiated and supported the research.

DECLARATION OF INTERESTS

The m.5024C>T mouse is available to lisence for commercial use from Max Plank Innovation. J.B.S. is a co-inventor on this license. All other authors declare no relevant conflicts of interest.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: April 19, 2022
Revised: October 28, 2022
Accepted: January 26, 2023
Published: February 23, 2023

REFERENCES

1. Vafai, S.B., and Mootha, V.K. (2012). Mitochondrial disorders as windows into an ancient organelle. Nature *491*, 374–383. https://10.1038/nature11707.
2. Wallace, D.C. (2018). Mitochondrial genetic medicine. Nat. Genet. *50*, 1642–1649. https://10.1038/s41588-018-0264-z.
3. Calvo, S.E., Clauser, K.R., and Mootha, V.K. (2016). MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res. *44*, D1251–D1257. https://10.1093/nar/gkv1003.
4. Pagliarini, D.J., Calvo, S.E., Chang, B., Sheth, S.A., Vafai, S.B., Ong, S.E., Walford, G.A., Sugiana, C., Boneh, A., Chen, W.K., et al. (2008). A mitochondrial protein compendium elucidates complex I disease biology. Cell *134*, 112–123. https://10.1016/j.cell.2008.06.016.
5. Stewart, J.B., and Chinnery, P.F. (2015). The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. Nat. Rev. Genet. *16*, 530–542. https://10.1038/nrg3966.
6. Mootha, V.K., Bunkenborg, J., Olsen, J.V., Hjerrild, M., Wisniewski, J.R., Stahl, E., Bolouri, M.S., Ray, H.N., Sihag, S., Kamal, M., et al. (2003). Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. Cell *115*, 629–640. https://10.1016/s0092-8674(03)00926-7.
7. Fernández-Vizarra, E., Enríquez, J.A., Pérez-Martos, A., Montoya, J., and Fernández-Silva, P. (2011). Tissue-specific differences in mitochondrial activity and biogenesis. Mitochondrion *11*, 207–213. https://10.1016/j.mito.2010.09.011.
8. Benard, G., Faustin, B., Passerieux, E., Galinier, A., Rocher, C., Bellance, N., Delage, J.P., Casteilla, L., Letellier, T., and Rossignol, R. (2006). Physiological diversity of mitochondrial oxidative phosphorylation. Am. J. Physiol. Cell Physiol. *291*, C1172–C1182. https://10.1152/ajpcell.00195.2006.
9. To, T.L., Cuadros, A.M., Shah, H., Hung, W.H.W., Li, Y., Kim, S.H., Rubin, D.H.F., Boe, R.H., Rath, S., Eaton, J.K., et al. (2019). A Compendium of Genetic Modifiers of Mitochondrial Dysfunction Reveals Intra-organelle Buffering. Cell *179*, 1222–1238.e17. https://10.1016/j.cell.2019.10.032.
10. Johnson, D.T., Harris, R.A., French, S., Blair, P.V., You, J., Bemis, K.G., Wang, M., and Balaban, R.S. (2007). Tissue heterogeneity of the mammalian mitochondrial proteome. Am. J. Physiol. Cell Physiol. *292*, C689–C697. https://10.1152/ajpcell.00108.2006.
11. Johnson, D.T., Harris, R.A., Blair, P.V., and Balaban, R.S. (2007). Functional consequences of mitochondrial proteome heterogeneity. Am. J. Physiol. Cell Physiol. *292*, C698–C707. https://10.1152/ajpcell.00109.2006.
12. Konopka, T. (2020). umap: Uniform Manifold Approximation and Projection. R package version. 0.2.7.0.
13. Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. Nature *566*, 490–495. https://10.1038/s41586-019-0933-9.
14. Argelaguet, R., Clark, S.J., Mohammed, H., Stapel, L.C., Krueger, C., Kapourani, C.A., Imaz-Rosshandler, I., Lohoff, T., Xiang, Y., Hanna, C.W., et al. (2019). Multi-omics profiling of mouse gastrulation at single-cell resolution. Nature *576*, 487–491. https://10.1038/s41586-019-1825-8.
15. Ibarra-Soria, X., Jawaid, W., Pijuan-Sala, B., Ladopoulos, V., Scialdone, A., Jörg, D.J., Tyser, R.C.V., Calero-Nieto, F.J., Mulas, C., Nichols, J., et al. (2018). Defining murine organogenesis at single-cell resolution reveals a role for the leukotriene pathway in regulating blood progenitor formation. Nat. Cell Biol. *20*, 127–134. https://10.1038/s41556-017-0013-z.
16. Kunz, W.S. (2003). Different metabolic properties of mitochondrial oxidative phosphorylation in different cell types–important implications for mitochondrial cytopathies. Exp. Physiol. *88*, 149–154. https://10.1113/eph8802512.
17. Tabula Muris Consortium (2020). A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. Nature *583*, 590–595. https://10.1038/s41586-020-2496-1.
18. Mittnenzweig, M., Mayshar, Y., Cheng, S., Ben-Yair, R., Hadas, R., Rais, Y., Chomsky, E., Reines, N., Uzonyi, A., Lumerman, L., et al. (2021). A single-embryo, single-cell time-resolved model for mouse gastrulation. Cell *184*, 2825–2842.e22. https://10.1016/j.cell.2021.04.004.

19. Tyser, R.C., Miranda, A.M., Chen, C.M., Davidson, S.M., Srinivas, S., and Riley, P.R. (2016). Calcium handling precedes cardiac differentiation to initiate the first heartbeat. Elife *5*, e17113. https://10.7554/eLife.17113.

20. Fukuda, R., Zhang, H., Kim, J.W., Shimoda, L., Dang, C.V., and Semenza, G.L. (2007). HIF-1 regulates cytochrome oxidase subunits to optimize efficiency of respiration in hypoxic cells. Cell *129*, 111–122. https://10.1016/j.cell.2007.01.047.

21. Kauppila, J.H.K., Baines, H.L., Bratic, A., Simard, M.L., Freyer, C., Mourier, A., Stamp, C., Filograna, R., Larsson, N.G., Greaves, L.C., and Stewart, J.B. (2016). A Phenotype-Driven Approach to Generate Mouse Models with Pathogenic mtDNA Mutations Causing Mitochondrial Disease. Cell Rep. *16*, 2980–2990. https://10.1016/j.celrep.2016.08.037.

22. Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J.N., Rovio, A.T., Bruder, C.E., Bohlooly-Y, M., Gidlöf, S., Oldfors, A., Wibom, R., et al. (2004). Premature ageing in mice expressing defective mitochondrial DNA polymerase. Nature *429*, 417–423. https://10.1038/nature02517.

23. Posse, V., Shahzad, S., Falkenberg, M., Hällberg, B.M., and Gustafsson, C.M. (2015). TEFM is a potent stimulator of mitochondrial transcription elongation in vitro. Nucleic Acids Res. *43*, 2615–2624. https://10.1093/nar/gkv105.

24. Park, C.B., Asin-Cayuela, J., Cámara, Y., Shi, Y., Pellegrini, M., Gaspari, M., Wibom, R., Hultenby, K., Erdjument-Bromage, H., Tempst, P., et al. (2007). MTERF3 is a negative regulator of mammalian mtDNA transcription. Cell *130*, 273–285. https://10.1016/j.cell.2007.05.046.

25. Harmel, J., Ruzzenente, B., Terzioglu, M., Spåhr, H., Falkenberg, M., and Larsson, N.G. (2013). The leucine-rich pentatricopeptide repeat-containing protein (LRPPRC) does not activate transcription in mammalian mitochondria. J. Biol. Chem. *288*, 15510–15519. https://10.1074/jbc.M113.471649.

26. Pearce, S.F., Rorbach, J., Van Haute, L., D'Souza, A.R., Rebelo-Guiomar, P., Powell, C.A., Brierley, I., Firth, A.E., and Minczuk, M. (2017). Maturation of selected human mitochondrial tRNAs requires deadenylation. Elife *6*, e27596. https://10.7554/eLife.27596.

27. Simard, M.L., Mourier, A., Greaves, L.C., Taylor, R.W., and Stewart, J.B. (2018). A novel histochemistry assay to assess and quantify focal cytochrome c oxidase deficiency. J. Pathol. *245*, 311–323. https://10.1002/path.5084.

28. Zhang, H., Esposito, M., Pezet, M.G., Aryaman, J., Wei, W., Klimm, F., Calabrese, C., Burr, S.P., Macabelli, C.H., Viscomi, C., et al. (2021). Mitochondrial DNA heteroplasmy is modulated during oocyte development propagating mutation transmission. Sci. Adv. *7*, eabi5657. https://10.1126/sciadv.abi5657.

29. Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. *47*, D607–D613. https://10.1093/nar/gky1131.

30. Bonekamp, N.A., Peter, B., Hillen, H.S., Felser, A., Bergbrede, T., Choidas, A., Horn, M., Unger, A., Di Lucrezia, R., Atanassov, I., et al. (2020). Small-molecule inhibitors of human mitochondrial DNA transcription. Nature *588*, 712–716. https://10.1038/s41586-020-03048-z.

31. Mennuni, M., Filograna, R., Felser, A., Bonekamp, N.A., Giavalisco, P., Lytovchenko, O., and Larsson, N.G. (2022). Metabolic resistance to the inhibition of mitochondrial transcription revealed by CRISPR-Cas9 screen. EMBO Rep. *23*, e53054. https://10.15252/embr.202153054.

32. Sissaoui, S., Yu, J., Yan, A., Li, R., Yukselen, O., Kucukural, A., Zhu, L.J., and Lawson, N.D. (2020). Genomic Characterization of Endothelial Enhancers Reveals a Multifunctional Role for NR2F2 in Regulation of Arteriovenous Gene Expression. Circ. Res. *126*, 875–888. https://10.1161/CIRCRESAHA.119.316075.

33. Alzu'bi, A., Lindsay, S.J., Harkin, L.F., McIntyre, J., Lisgo, S.N., and Clowry, G.J. (2017). The Transcription Factors COUP-TFI and COUP-TFII have Distinct Roles in Arealisation and GABAergic Interneuron Specification in the Early Human Fetal Telencephalon. Cereb. Cortex *27*, 4971–4987. https://10.1093/cercor/bhx185.

34. Hu, J.S., Vogt, D., Lindtner, S., Sandberg, M., Silberberg, S.N., and Rubenstein, J.L.R. (2017). Coup-TF1 and Coup-TF2 control subtype and laminar identity of MGE-derived neocortical interneurons. Development *144*, 2837–2851. https://10.1242/dev.150664.

35. Schon, E.A., Bonilla, E., and DiMauro, S. (1997). Mitochondrial DNA mutations and pathogenesis. J. Bioenerg. Biomembr. *29*, 131–149.

36. Suomalainen, A., and Battersby, B.J. (2018). Mitochondrial diseases: the contribution of organelle stress responses to pathology. Nat. Rev. Mol. Cell Biol. *19*, 77–92. https://10.1038/nrm.2017.66.

37. Quirós, P.M., Prado, M.A., Zamboni, N., D'Amico, D., Williams, R.W., Finley, D., Gygi, S.P., and Auwerx, J. (2017). Multi-omics analysis identifies ATF4 as a key regulator of the mitochondrial stress response in mammals. J. Cell Biol. *216*, 2027–2045. https://10.1083/jcb.201702058.

38. Filograna, R., Koolmeister, C., Upadhyay, M., Pajak, A., Clemente, P., Wibom, R., Simard, M.L., Wredenberg, A., Freyer, C., Stewart, J.B., and Larsson, N.G. (2019). Modulation of mtDNA copy number ameliorates the pathological consequences of a heteroplasmic mtDNA mutation in the mouse. Sci. Adv. *5*, eaav9824. https://10.1126/sciadv.aav9824.

39. Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat. Methods *14*, 1083–1086. https://10.1038/nmeth.4463.

40. Friedman, J.R., and Kaestner, K.H. (2006). The Foxa family of transcription factors in development and metabolism. Cell. Mol. Life Sci. *63*, 2317–2328. https://10.1007/s00018-006-6095-6.

41. Pikkarainen, S., Tokola, H., Kerkelä, R., and Ruskoaho, H. (2004). GATA transcription factors in the developing and adult heart. Cardiovasc. Res. *63*, 196–207. https://10.1016/j.cardiores.2004.03.025.

42. Lai, X., Liu, J., Zou, Z., Wang, Y., Wang, Y., Liu, X., Huang, W., Ma, Y., Chen, Q., Li, F., et al. (2021). SOX10 ablation severely impairs the generation of postmigratory neural crest from human pluripotent stem cells. Cell Death Dis. *12*, 814. https://10.1038/s41419-021-04099-4.

43. Saur, A.L., Fröb, F., Weider, M., and Wegner, M. (2021). Formation of the node of Ranvier by Schwann cells is under control of transcription factor Sox10. Glia *69*, 1464–1477. https://10.1002/glia.23973.

44. Hou, S., Li, Z., Dong, J., Gao, Y., Chang, Z., Ding, X., Li, S., Li, Y., Zeng, Y., Xin, Q., et al. (2022). Heterogeneity in endothelial cells and widespread venous arterialization during early vascular development in mammals. Cell Res. *32*, 333–348. https://10.1038/s41422-022-00615-z.

45. Haldin, C.E., and LaBonne, C. (2010). SoxE factors as multifunctional neural crest regulatory factors. Int. J. Biochem. Cell Biol. *42*, 441–444. https://10.1016/j.biocel.2009.11.014.

46. Trainor, P.A. (2005). Specification of neural crest cell formation and migration in mouse embryos. Semin. Cell Dev. Biol. *16*, 683–693. https://10.1016/j.semcdb.2005.06.007.

47. Simões-Costa, M., and Bronner, M.E. (2015). Establishing neural crest identity: a gene regulatory recipe. Development *142*, 242–257. https://10.1242/dev.105445.

48. Costa, R., Muccioli, S., Brillo, V., Bachmann, M., Szabò, I., and Leanza, L. (2021). Mitochondrial dysfunction interferes with neural crest specification through the FoxD3 transcription factor. Pharmacol. Res. *164*, 105385. https://10.1016/j.phrs.2020.105385.

49. Carelli, V., Vergani, L., Bernazzi, B., Zampieron, C., Bucchi, L., Valentino, M., Rengo, C., Torroni, A., and Martinuzzi, A. (2002). Respiratory function in cybrid cell lines carrying European mtDNA haplogroups: implications for Leber's hereditary optic neuropathy. Biochim. Biophys. Acta *1588*, 7–14.

50. Gómez-Durán, A., Pacheu-Grau, D., López-Gallardo, E., Díez-Sánchez, C., Montoya, J., López-Pérez, M.J., and Ruiz-Pesini, E. (2010). Unmasking the causes of multifactorial disorders: OXPHOS differences between mitochondrial haplogroups. Hum. Mol. Genet. *19*, 3343–3353. https://10.1093/hmg/ddq246.

51. Gorman, G.S., Schaefer, A.M., Ng, Y., Gomez, N., Blakely, E.L., Alston, C.L., Feeney, C., Horvath, R., Yu-Wai-Man, P., Chinnery, P.F., et al. (2015). Prevalence of nuclear and mitochondrial DNA mutations related to adult mitochondrial disease. Ann. Neurol. *77*, 753–759. https://10.1002/ana.24362.

52. Gorman, G.S., Chinnery, P.F., DiMauro, S., Hirano, M., Koga, Y., McFarland, R., Suomalainen, A., Thorburn, D.R., Zeviani, M., and Turnbull, D.M. (2016). Mitochondrial diseases. Nat. Rev. Dis. Primers *2*, 16080. https://10.1038/nrdp.2016.80.

53. Ciafaloni, E., Ricci, E., Shanske, S., Moraes, C.T., Silvestri, G., Hirano, M., Simonetti, S., Angelini, C., Donati, M.A., Garcia, C., et al. (1992). MELAS: clinical features, biochemistry, and molecular genetics. Ann. Neurol. *31*, 391–398. https://10.1002/ana.410310408.

54. Rossignol, R., Malgat, M., Mazat, J.P., and Letellier, T. (1999). Threshold effect and tissue specificity. Implication for mitochondrial cytopathies. J. Biol. Chem. *274*, 33426–33432. https://10.1074/jbc.274.47.33426.

55. Scarpulla, R.C. (2008). Transcriptional paradigms in mammalian mitochondrial biogenesis and function. Physiol. Rev. *88*, 611–638. https://10.1152/physrev.00025.2007.

56. Leary, S.C., Battersby, B.J., and Moyes, C.D. (1998). Inter-tissue differences in mitochondrial enzyme activity, RNA and DNA in rainbow trout (Oncorhynchus mykiss). J. Exp. Biol. *201* (Pt 24), 3377–3384.

57. Sharma, R., Reinstadler, B., Engelstad, K., Skinner, O.S., Stackowitz, E., Haller, R.G., Clish, C.B., Pierce, K., Walker, M.A., Fryer, R., et al. (2021). Circulating markers of NADH-reductive stress correlate with mitochondrial disease severity. J. Clin. Invest. *131*, e136055. https://10.1172/JCI136055.

58. Naganuma, M., Sekine, S.i., Chong, Y.E., Guo, M., Yang, X.L., Gamper, H., Hou, Y.M., Schimmel, P., and Yokoyama, S. (2014). The selective tRNA aminoacylation mechanism based on a single G*U pair. Nature *510*, 507–511. https://10.1038/nature13440.

59. Pulliam, D.A., Deepa, S.S., Liu, Y., Hill, S., Lin, A.L., Bhattacharya, A., Shi, Y., Sloane, L., Viscomi, C., Zeviani, M., and Van Remmen, H. (2014). Complex IV-deficient Surf1(-/-) mice initiate mitochondrial stress responses. Biochem. J. *462*, 359–371. https://10.1042/BJ20140291.

60. Seiferling, D., Szczepanowska, K., Becker, C., Senft, K., Hermans, S., Maiti, P., König, T., Kukat, A., and Trifunovic, A. (2016). Loss of CLPP alleviates mitochondrial cardiomyopathy without affecting the mammalian UPRmt. EMBO Rep. *17*, 953–964. https://10.15252/embr.201642077.

61. Khan, N.A., Nikkanen, J., Yatsuga, S., Jackson, C., Wang, L., Pradhan, S., Kivelä, R., Pessia, A., Velagapudi, V., and Suomalainen, A. (2017). mTORC1 Regulates Mitochondrial Integrated Stress Response and Mitochondrial Myopathy Progression. Cell Metab. *26*, 419–428.e5. https://10.1016/j.cmet.2017.07.007.

62. Forsström, S., Jackson, C.B., Carroll, C.J., Kuronen, M., Pirinen, E., Pradhan, S., Marmyleva, A., Auranen, M., Kleine, I.M., Khan, N.A., et al. (2019). Fibroblast Growth Factor 21 Drives Dynamics of Local and Systemic Stress Responses in Mitochondrial Myopathy with mtDNA Deletions. Cell Metab. *30*, 1040–1054.e7. https://10.1016/j.cmet.2019.08.019.

63. Stewart, J.B., and Chinnery, P.F. (2021). Extreme heterogeneity of human mitochondrial DNA from organelles to populations. Nat. Rev. Genet. *22*, 106–118. https://10.1038/s41576-020-00284-x.

64. Morita, M., Gravel, S.P., Chénard, V., Sikström, K., Zheng, L., Alain, T., Gandin, V., Avizonis, D., Arguello, M., Zakaria, C., et al. (2013). mTORC1 controls mitochondrial activity and biogenesis through 4E-BP-dependent translational regulation. Cell Metab. *18*, 698–711. https://10.1016/j.cmet.2013.10.001.

65. Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W., and Clayton, D.A. (1981). Sequence and gene organization of mouse mitochondrial DNA. Cell *26*, 167–180. https://10.1016/0092-8674(81)90300-7.

66. Norrander, J., Kempe, T., and Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene *26*, 101–106. https://10.1016/0378-1119(83)90040-9.

67. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://10.1093/bioinformatics/bts635.

68. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15. https://10.1186/s13059-017-1382-0.

69. Gatto, L., and Lilley, K.S. (2012). MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. Bioinformatics *28*, 288–289. https://10.1093/bioinformatics/btr645.

70. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740. https://10.1038/nmeth.3901.

71. Rath, S., Sharma, R., Gupta, R., Ast, T., Chan, C., Durham, T.J., Goodman, R.P., Grabarek, Z., Haas, M.E., Hung, W.H.W., et al. (2021). MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. Nucleic Acids Res. *49*, D1541–D1547. https://10.1093/nar/gkaa1011.

72. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287. https://10.1089/omi.2011.0118.

73. Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinf. *14*, 128. https://10.1186/1471-2105-14-128.

74. Gaspari, M., Falkenberg, M., Larsson, N.G., and Gustafsson, C.M. (2004). The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells. EMBO J. *23*, 4606–4614. https://10.1038/sj.emboj.7600465.

75. Posse, V., Hoberg, E., Dierckx, A., Shahzad, S., Koolmeister, C., Larsson, N.G., Wilhelmsson, L.M., Hällberg, B.M., and Gustafsson, C.M. (2014). The amino terminal extension of mammalian mitochondrial RNA polymerase ensures promoter specific transcription initiation. Nucleic Acids Res. *42*, 3638–3647. https://10.1093/nar/gkt1397.

76. Gammage, P.A., Viscomi, C., Simard, M.L., Costa, A.S.H., Gaude, E., Powell, C.A., Van Haute, L., McCann, B.J., Rebelo-Guiomar, P., Cerutti, R., et al. (2018). Genome editing in mitochondria corrects a pathogenic mtDNA mutation in vivo. Nat. Med. *24*, 1691–1695. https://10.1038/s41591-018-0165-9.

77. Davies, S.M.K., Lopez Sanchez, M.I.G., Narsai, R., Shearwood, A.M.J., Razif, M.F.M., Small, I.D., Whelan, J., Rackham, O., and Filipovska, A. (2012). MRPS27 is a pentatricopeptide repeat domain protein required for the translation of mitochondrially encoded proteins. FEBS Lett. *586*, 3555–3561. https://10.1016/j.febslet.2012.07.043.

78. Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J. Roy. Stat. Soc. B *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

79. Shevchenko, G., Musunuri, S., Wetterhall, M., and Bergquist, J. (2012). Comparison of extraction methods for the comprehensive analysis of mouse brain proteome using shotgun-based mass spectrometry. J. Proteome Res. *11*, 2441–2451. https://10.1021/pr201169q.

80. McAlister, G.C., Nusinow, D.P., Jedrychowski, M.P., Wühr, M., Huttlin, E.L., Erickson, B.K., Rad, R., Haas, W., and Gygi, S.P. (2014). MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. Anal. Chem. *86*, 7150–7158. https://10.1021/ac502040v.

81. Gatto, L., Gibb, S., and Rainer, J. (2021). MSnbase, efficient and elegant R-based processing and visualization of raw mass spectrometry data. J. Proteome Res. *20*, 1063–1069. https://10.1021/acs.jproteome.0c00313.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Total OXPHOS Blue Native WB Antibody Cocktail	Abcam	ab110412; RRID: AB_2847807
Goat Anti-mouse HRP	Promega	W4021; RRID: AB_430834
Mouse anti-OXPHOS Cocktail	Abcam	Ab110413; RRID: AB_2629281
ECL Sheep Anti-mouse IgG HRP-linked Whole Antibody	Biorad	AAC10P; RRID: AB_321929
Chemicals, peptides, and recombinant proteins		
Lipofectaimine 3000	Invitrogen	L3000015
Lipofectamine RNAiMAX Transfection Reagent	Invitrogen	13,778,030
DMEM-GlutaMAX 4.5 g/L Glucose	Gibco	31,966,021
DMEM 4.5 g/L Glucose	Gibco	41,966,029
Fetal Bovine Serum	Gibco	10,270,106
Pen/Strep	Gibco	15,140,122
NaOAc	Sigma	S2889
$NaIO_4$	Sigma	S1878
NaCl	Invitrogen	AM9670G
$MgCl_2$	Invitrogen	AM9530G
BSA	Sigma	A2153
DTT	Thermo Scientific	P2325
ATP	Invitrogen	AM8110G
GTP	Invitrogen	AM8130G
CTP	Invitrogen	AM8220G
UTP	Invitrogen	AM8140G
[α-P32]UTP	Hartmann Analytic	FP-210
RNase Inhibitor, murine	NEB	M0314
Lysine	Sigma	L5501
EDTA	Thermo Scientific	17,892
Formamide	Sigma	47,671
HindIII	NEB	R0104
PstI	NEB	R0140
BamHI	NEB	R0136
NdeI	NEB	R0111
SapI	NEB	R0569
Tris-HCl pH 8.0	Invitrogen	AM9855G
Trizma Pre-set crystals pH 8.5	Sigma	T8818
Urea	Sigma	U5128
Xylene cyanol	Thermo Scientific	422,690,050
Bromophenol blue	Sigma	B0126
Polyacrylamide (19:1 acrylamide/bis)	Sigma	A2917
Glycerol Tolerant Buffer	National Diagnostics	NAT1220
Gel Loading Buffer II	Invitrogen	AM8546G
TriZol	Ambion	15,596,018
Na_2HPO_4	Sigma	S9763
Ethanol	VWR	20,821
SDS	Sigma	L3771
Saline sodium citrate	Invitrogen	15,557,044

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RNA Secure	Ambion	AM7005
Heparin	Sigma	H0200000
IRDye 800CW Streptavidin	Li-cor	962-32230
Tween 20	Life Technologies	003,005
RIPA Buffer	Sigma	R0278
Coomassie Brilliant Blue	Thermo Scientific	20,279
NuPAGE 12% Bis-Tris Precast Gel	Thermo Scientific	NP0343BOX
ECL Western Blotting Detection Reagent	GE	28,980,926
Leibovitz's L-15 Medium	Sigma	L1518
Trypsin/EDTA	Gibco	15,400,054
DRAQ7	Invitrogen	D15106
Proteinase K	Ambion	AM2546
Recombinant mouse Polrmt	This manuscript	N/A
Recombinant mouse Tfb2m	This manuscript	N/A
Recombinant mouse Tfam	This manuscript	N/A
Protease Inhibitor Cocktail	Abcam	Ab271306
βOG	Sigma	850511P
Critical commercial assays		
Monarch Genomic DNA Purification Kit	NEB	T3010
PyroMark PCR Kit	QIAGEN	978,703
PyroMark Q48 Advanced Reagents Kit	QIAGEN	974,002
PyroMark Q24 Advanced Reagents Kit	QIAGEN	970,902
Gentra Puregene Tissue Kit	QIAGEN	158,063
RNeasy Plus Mini Kit	QIAGEN	74,134
Prime-IT II Random Primer Labeling Kit	Agilent	300,385
NEBNext Poly(A) mRNA Kit	NEB	E7490
NEBNext Ultra II Directional RNA Library Prep Kit	NEB	E7760
Chromium Single Cell 3′ Library + Gel Bead Kit v3	10X Genomics	1,000,075
Chromium Chip B Kit	10X Genomics	1,000,073
High-sensitivity D5000 ScreenTape & Reagents	Agilent	5067-5592/5593
High-sensitivity D1000 ScreenTape & Reagents	Agilent	5067-5584/5585
D5000 ScreenTape & Reagents	Agilent	5067-5588/5589
RNA ScreenTape & Reagents	Agilent	5067-5576/5577/5578
QuBit dsDNA HS Assay Kit	Invitrogen	Q32851
QuBit RNA HS Assay Kit	Invitrogen	Q32855
KAPA Library Quantification Kit	KAPA	KK4824
Pierce BCA Protein Assay Kit	Thermo Scientific	23,227
TMTpro 16plex Label Reagent Set	Thermo Scientific	A44522
Deposited data		
10X scRNA-seq data files	This manuscript, deposited in GEO	GEO: GSE200746
Msc80 Schwann cell RNA-seq data files	Saur et al.[43]	GEO: GSE152687
Neural Crest Stem Cell RNA-seq data files	Lai et al.[42]	GEO: GSE160312
Adult Mouse Tissue Transcriptome data files	Tabula Muris Compendium[17]	GEO: GSE132042
Mouse Embryonic Fibroblast Bulk RNA-seq data files	This manuscript, deposited in GEO	GEO: GSE214740
Experimental models: Cell lines		
Primary Mouse Embryonic Fibroblasts (m.5019A>G, m.5024C>T & WT)	This manuscript & J. Stewart	Kauppila et al.[21]

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
C57Bl/6 wild-type mice	Charles River Laboratories	Strain code 027; RRID: IMSR_CRL:027
m.5024C>T mice	J. Stewart	Kauppila et al.[21]; RRID: MGI:5902139
m.5019A>G mice	This manuscript	MGI: 6860509
Oligonucleotides		
Riboprobe *mt-Ta* forward: TAATACGACTCACTATAG GGAGACTAAGGACTGTA AGACTTCATC	IDT	N/A
Riboprobe *mt-Ta* reverse: GAGGTCTTAGCTTAATTAAAG	IDT	N/A
Riboprobe *mt-Tw* forward: TAATACGACTCACTATAGG GAGACCAGAAGTTAAACT TGTGTG	IDT	N/A
Riboprobe *mt-Tw* reverse: AGAAGTTTAGGATATACTAG	IDT	N/A
Riboprobe *mt-Tl1* forward: TAATACGACTCACTATAG GGAGACTATTAGGGAGA GGATTTGAAC	IDT	N/A
Riboprobe *mt-Tl1* reverse: ATTAGGGTGGCAGAGCCAGG	IDT	N/A
Riboprobe 5S rRNA end-labelled oligo: AGCCTACAGCACCCGGTATTCCCAG GCGGTCTCCCATCCAAGTACTAACCA	IDT	N/A
18S probes	IDT	N/A
Mouse specific biotin labeled oligos	IDT	N/A
Pyrosequencing m.5024C>T forward: ATACTAGTCCGCGAGCCTTCAAA	IDT	N/A
Pyrosequencing m.5024C>T reverse: /5Biosg/AATTGCAAATTCGAAGGTGTAGA	IDT	N/A
Pyrosequencing m.5024C>T sequencing: CACACAAGTTTAACTTCTGA	IDT	N/A
Pyrosequencing m.5019A>G forward: /5biosg/TTCCACCCTAGCTATCATAAGC	IDT	N/A
Pyrosequencing m.5019A>G reverse: CGTAGGTTTAATTCCTGCCAATCT	IDT	N/A
Pyrosequencing m.5019A>G sequencing: GATGTAGGATGAAGTCTTA	IDT	N/A
Mouse Non-targeting siRNA SMARTPool	Dharmacon	D-001810-10-05
Mouse *Gapdh* ON-TARGETplus siRNA SMARTPool	Dharmacon	D-001830-20-05
Mouse *Klf12* ON-TARGETplus siRNA SMARTPool	Dharmacon	L-065329-01-0005
Mouse *E2f3* ON-TARGETplus siRNA SMARTPool	Dharmacon	L-045247-00-0005
Mouse *Taf1* ON-TARGETplus siRNA SMARTPool	Dharmacon	L-058598-00-0005
Mouse *Maz* ON-TARGETplus siRNA SMARTPool	Dharmacon	L-059837-01-0005
Mouse *Bclaf1* ON-TARGETplus siRNA SMARTPool	Dharmacon	L-058999-00-0005
Recombinant DNA		
pBSSVD2005 – SV40 Large-T Antigen	D. Ron	RRID: Addgene_21826
pAM1 – NLL-MetRS	Bibb et al.[65]	RRID: Addgene_51401
pUC-19	Norrander et al.[66]	RRID: Addgene_50005
WT LSP Genestrand	Eurofins Genomics	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mutant LSP Genestrand	Eurofins Genomics	N/A
Software and algorithms		
PyroMark Analysis Software	QIAGEN	9,024,325
STARsolo (version 2.7.3a)	Dobin et al.[67]	https://github.com/alexdobin/STAR/blob/master/docs/STARsolo.md
Scanpy (version 1.7.2)	Wolf et al.[68]	https://scanpy.readthedocs.io/en/stable/
Prism (v9.0)	GraphPad	https://www.graphpad.com/scientific-software/prism/
STRING (v11)	Szklarczyk et al.[29]	https://string-db.org/
Incucyte Software (v2020C)	Sartorius	https://www.sartorius.com/en/products/live-cell-imaging-analysis/live-cell-analysis-software
Proteome Discoverer (v2.4)	Thermo Scientific	https://www.thermofisher.com/uk/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/multi-omics-data-analysis/proteome-discoverer-software.html
Mascot (v2.6)	Matrix Science	https://www.matrixscience.com/
MSnbase (v2.22.0)	Gatto et al.[69]	http://lgatto.github.io/MSnbase/
Perseus (v2.0.7.0)	Tyanova et al.[70]	http://coxdocs.org/doku.php?id=perseus:start
MitoCarta (v3.0)	Rath et al.[71]	https://www.broadinstitute.org/mitocarta/mitocarta30-inventory-mammalian-mitochondrial-proteins-and-pathways
clusterProfiler (v3.15)	Yu et al.[72]	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
Enrichr	Chen et al.[73]	https://maayanlab.cloud/Enrichr/
Other		
Ssniff M−H Low-Phytoestrogen mouse diet	Ssniff spezialdiaeten GmbH	V155
Ssniff M-Z Low-Phytoestrogen mouse diet	Ssniff spezialdiaeten GmbH	V115
SAFE105 Universal mouse diet	Safe Diets	SAFER105*

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Patrick Chinnery (pfc25@medschl.cam.ac.uk).

Materials availability

The m.5019A>G mouse line and immortalised MEF cell line derived from it are freely available to researchers upon completion of a non-profit Material Transfer Agreement (MTA). Please direct queries regarding availability of these materials to Dr. Jim Stewart (Jim.Stewart@newcastle.ac.uk).

Data and code availability

- Single-cell and bulk RNA-seq data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Original northern, Southern & western blot images have been deposited and are publicly available as of the date of publication. The DOI is listed in the key resources table.
- Code to reproduce the analysis and figures is available in Github repository: https://github.com/nmalwinka/2022_Burr_Klimm_Glynos_Prater__Cell_lineage_specific_mitochondrial_resilience_during_mammalian_organ.

- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

The mouse lines used in this study were m.5024C>T (Allele symbol: mt-Ta^{m1Jbst}, MGI ID: 5902095), m.5019A>G (Allele symbol: mt-Ta^{m2Jbst}, MGI ID: 6860509) and WT, all mice used in this study were bred on the C57BL/6 background.

The initial identification and characterization of the m.5019A>G mouse was carried out at the Max Planck Institute for Biology of Aging. All animal work was performed in accordance with recommendations and guidelines of the Federation of European Laboratory Animal Science Associations (FELASA). All experiments were approved and permitted by the Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen in accordance with German and European Union regulations.

The mice used were on the C57Bl/6NCrl background strain (Charles River Laboratories, Germany strain code 027). Mice were housed in a 12 h light/dark cycle at 21°C and fed *ad libitum* on a standard mouse food (ssniff M−H Low-Phytoestrogen) or an enhanced diet when breeding or newly weaned mice (ssniff M-Z Low-Phytoestrogen) by Ssniff Spezialdiaeten GmbH.

Mice were sacrificed by exposure of CO_2 and death confirmed via cervical dislocation, unless the brain was to be examined, in which case death was confirmed by exsanguination. Tissues were dissected, drained of blood, and washed in PBS. Tissues for histochemical analysis were frozen in isopentane that was previously cooled to −160°C in liquid nitrogen. Tissues for other analyses were snap frozen by emersion in an Eppendorf tube, be emersion in liquid nitrogen.

Breeding of the m.5024C>T and m.5019A>G heteroplasmic mouse lines and WT controls in Cambridge, UK, was performed according to the Animal (Scientific Procedures) Act 1986 under Home Office Project Licence P6C97520A. All experiments followed the ARRIVE guidelines and were approved by the University of Cambridge Animal Welfare Ethical Review Body (AWERB).

Mice were kept in individually-ventilated cages (Tecniplast) at 20-24°C, 45-65% humidity, with *ad libitum* access to SAFE 105 universal diet (Safe Diets) and water. Breeding females used for timed matings were aged between 8 and 12 weeks. To set up timed matings, stud males were first moved to an empty cage to acclimatise before introducing the female. Females were subsequently checked for vaginal plugs each morning, and noon on the day of a positive plug check was designated E0.5. At E8.5 following a positive vaginal plug check, pregnant females were culled by cervical dislocation, with death confirmed by exsanguination.

Mouse litter size data was collected within 48 h of birth, pre-wean deaths were counted as any pup losses between this initial litter-size count and weaning of pups at 2-3 weeks of age.

Cell lines

Primary m.5024C>T MEFs were isolated from individual E13.5 embryos and heteroplasmy levels of individual embryos were determined by pyrosequencing. Primary MEFs were immortalized by transfection with the SV40 large T antigen (pBSSVD2005). The plasmid was a gift from David Ron (Addgene plasmid # 21826).

Briefly, early passage primary MEFs (passage 0, 2x10^5 cells/well in a 6 well format) were transfected with 3 μg of SV40 plasmid using Lipofectamine 3000 (Invitrogen) according to the manufacturer's instructions. After 24 h, cells were trypsinized and serially diluted into 96 well plates to achieve single colonies. In the next weeks, 96 well plates were checked regularly for fast growing immortalized clones. Upon reaching confluency, immortalized MEFs of individual wells were trypsinized and expanded to larger formats. Heteroplasmy levels of individual immortalized MEF clones were re-analyzed by pyrosequencing. Primary and immortalized MEFs were maintained in DMEM-GlutaMAX (Gibco) containing 4.5 g/L of glucose supplemented with 10% fetal bovine serum (Gibco), 1% Pen/Strep (Gibco) and 50 μg/mL Uridine at 37°C and 5% CO2 in a humidified incubator.

METHOD DETAILS

Identification of the m.5019A>G strain

Mouse breeding to introduce mutant alleles, identification of the line through screening of complex IV deficiency in the colon, and laser-capture microdissection followed by mtDNA sequencing.[21]

mtDNA heteroplasmy measurements

At weaning, ear skin biopsies were taken from all heteroplasmic mice to allow each animal to be assigned a reference heteroplasmy value. DNA was extracted from biopsy samples using the Monarch Genomic DNA Purification Kit (NEB) according to the manufacturer's protocol and quantified using a NanoDrop spectrophotometer (Thermo Scientific). Heteroplasmy measurements were made by pyrosequencing using the QIAGEN Q48 Autoprep or Q24 vacuum workstation systems. Briefly, a section of the mitochondrial genome containing the mutation of interest was amplified from 50 to 100ng extracted DNA using the PyroMark PCR Kit (QIAGEN) and primers designed according to the manufacturer's instructions. After confirming successful amplification by gel electrophoresis, heteroplasmy levels were analyzed by running PCR product on either the Q48 Autoprep or Q24 vacuum workstation along with an

appropriate sequencing primer and following the appropriate manufacturer's protocol. Heteroplasmy was then call using PyroMark analysis software (QIAGEN) and reported as the percentage of mutant base present in the sample.

In vitro transcription assays

Recombinant mouse Polrmt, Tfb2m and Tfam were expressed and purified according to previously described protocols.[74,75] Mouse mitochondrial LSP regions 16,110 to 16,259 (for LSP "A" - WT) and 16,120 to 16,259 (for LSP "B" - WT or LSP "B" - mutant variants), followed by a random DNA sequence upstream of the promoter, were obtained from Eurofins Genomics in the form of linear dsDNA (GeneStrands). LSP "A" - WT and LSP "B" (WT or the corresponding mutant version) containing DNA fragments were cloned in pUC19 under HindIII/PstI and PstI/BamHI respectively. Dual promoter templates used in the *in vitro* transcription reactions were created by linearization of the previous plasmids with NdeI and SapI, generating run-off transcripts (R.O.) of about 321 nt for LSP "A" - WT and 302 nt LSP "B".

In vitro transcription reactions were performed in 25 μL reaction volume, in buffer containing 25 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 64 mM NaCl, 100 μg/mL BSA, 1 mM DTT, 400 μM ATP, 150 μM GTP, 150 μM CTP, 10 μM UTP, 0.02 μM α-32P UTP (3000 Ci/mmol, Hartmann Analytic GmbH) and 4 U RNase inhibitor Murine (New England Biolabs). 4 nM of DNA template, 30 nM Polrmt, 40 nM Tfb2m and 0,3, 0,6, 1,1 or 2.2 μM Tfam as indicated in the figures, were used in each reaction. All reactions were set up on ice, initiated by the addition of ribonucleotides and incubated at 32°C for 10 min, unless otherwise stated. Then, they were stopped by the addition of stop buffer (10 mM Tris-HCl pH 8, 200 mM NaCl, 1 mM EDTA and 100 μg/mL proteinase K) followed by incubation at 42°C for 45 min. Transcripts were purified by ethanol precipitation and pellets resuspended in loading buffer (98% formamide, 10 mM EDTA, 0.025% xylene cyanol and 0.025 bromophenol blue). Time-course experiments were performed in a total reaction volume of 125 μL per condition and at the indicated times, 25 μL were removed and the reaction ended with stop buffer. Reaction products were resolved on a denaturing 4% polyacrylamide gel followed by exposure on Phosphorimager. The intensity of the transcripts on the images generated from Phosphorimager were quantified using the program Multi Gauge.

Measurement of aminoacylation status

Total RNA was extracted from cells or heart tissue using TRIzol.[76] Recovered RNA pellets were resuspended in a solution of 100 mM NaOAc (pH 5) and kept on ice. Once RNA was fully resuspended, a $NaIO_4$ solution was added to 50mM and incubated at room temperature for 30 min. The reaction was quenched by addition of glucose to 100mM followed by a further 5-min incubation at room temperature. Samples were purified by filtration through MicroSpin G-25 spin columns (GE Healthcare, 27-5325-01) and the RNA precipitated with ethanol. The precipitated samples were centrifuged at 12,000 g at 4°C for 15 min and the resulting pellets were washed in 80% ethanol and centrifuged at 7500 g at 4°C for 5 min. The washed pellets were resuspended in 1 M Lysine pH 7.4 for 1 h at 45°C and then passed again through G-25 spin columns. Control samples were prepared alongside by treating samples with 100 mM NaCl instead of $NaIO_4$. A deacylated negative control sample was prepared in advance by incubating 20 μg total mouse RNA in 200 mM Tris-HCl at 75°C for 5 min then precipitating and resuspending the sample in 100 mM NaOAc (pH 5) with the experimental samples. All samples were subsequently resolved by high resolution denaturing Urea PAGE (8 M Urea, 10% polyacrylamide (19:1 acrylamide/bis), 1X Glycerol tolerant buffer (GTB)). Between 1.5 μg and 5 μg were loaded into the gel with Gel Loading Buffer II (Invitrogen, AM8546G) and run for approximately 5 h at 75 W. The gel was then blotted by capillary transfer onto Amersham Hybond-N+ membranes which were cross linked in a CL-1000 UV crosslinker at 0.120 J per cm^2. The resulting northern blots were then hybridised and visualised.[76] Riboprobe primers used for this method are listed in the Key Resources Table

Histology

COX-SDH dual-staining was performed as described in ref. 21 and 38 and NBTx staining was performed as described in ref. 27.

RNA isolation and northern blot analysis

Total RNA was isolated from snap frozen heart tissues with TRIzol (Ambion) following the manufacturer's standard protocol. The extracted RNA was quantified with the Nanodrop and resolved on formaldehyde-agarose gel followed by a transfer onto Amersham Hybond-N+ membrane (GE Healthcare). Transcripts of interest were detected with non-radioactive method.[77] Membranes were probed with mouse-specific biotin-labeled oligonucleotides by overnight incubation in hybridization buffer (5X saline sodium citrate [SSC], 20 mM Na_2HPO_4, 7% SDS, 0.5x RNA secure [Ambion], and 100 μg/mL heparin) at 50°C followed by washing and signal detection with IRDye 800CW dye-labeled streptavidin (dilution 1:5000 in TBS, 0.05% tween 20) in an LI-COR Biosciences imaging system.

Western Blot analysis

Protein was extracted by lysis in RIPA buffer. Samples were quantified by BCA assay and 40 μg was resolved by western blot.[76] Steady state levels of OXPHOS complexes were then determined using Total OXPHOS Blue Native WB Antibody Cocktail (1:250, ab110412) and goat anti-mouse conjugated to HRP (1:2000, Promega W4021). Equal gel loading was verified by staining with Coomassie Brilliant Blue (Life technologies).

30 μg of protein lysates from intestinal epithelial tissues were separated by SDS-PAGE in 4-12% Bis-Tris precast gel (Thermo Scientific) using Novex Midi-Cell (Thermo Scientific). Proteins were then transferred to polyvinylidene difluoride (PVDF) membrane (GE Healthcare) in a Criterion Blotter (Bio-Rad) The membrane was probed with mouse Anti-OXPHOS cocktail antibodies (Abcam) and

detected with ECL Anti-mouse IgG horseradish peroxidase linked whole antibody (from sheep) (Bio-Rad). Signal development was detected using ECL Western Blotting Detection Reagent (GE Healthcare) and scanned in LI-COR 43 imaging (LI-COR Biosciences).

Southern blots

DNA was extracted from tissues using the Gentra Puregene Tissue Kit (Qiagen). 5 μg of DNA was digested with *SacI* (New England Biolabs) for electrophoresis. DNA was transferred to Hybond-N+ membranes (Amersham) and UV crosslinked for blotting. mtDNA was detected using the pAM1 plasmid,[65] after radiolabeling with the Prime-IT II Random Primer Labeling kit (Agilent Technologies). 18S probes was used as the loading control for nuclear DNA. The membrane was sealed in a plastic bag and exposed to a phosphor screen by placing them both in a film cassette. The film was scanned using Typhoon FLA 7000 (GE Healthcare) to visualize the mtDNA.

Preparation of E8.5 embryos for scRNA-seq

At E8.5 following a positive vaginal plug check, pregnant females were culled, the uterus was dissected out and placed in ice-cold Leibovitz's L-15 medium (Sigma) supplemented with 2% FBS (Gibco). Embryos were isolated manually under a Leica M205 FA fluorescence dissecting microscope. Briefly, uterine tissues were removed to release individual embryos contained in decidual capsules. The decidua was carefully dissected from around each embryo and extra embryonic tissues, including Reichert's membrane and the allantois, were removed. Embryos were confirmed as being the correct stage based on their morphological appearance (Theiler stage 13).

Following dissection, embryos were pooled and washed twice in chilled PBS before adding 100 μL 1X Trypsin/EDTA (Gibco) and incubating at 37°C for 10 min. Pools were then pipetted up and down gently 20 times using a P200 pipet to create a single-cell suspension and the trypsin was inactivated by addition of 900 μL chilled L-15 medium +2% FBS. Suspensions were then passed through a 50 μm filter to remove any cell clumps/larger debris into 5 mL round-bottomed polypropylene tubes and stained with DRAQ7 viability dye (Invitrogen). Live cells (i.e. DRAQ7 negative) were sorted using a BD Melody flow sorter into a 1.5 mL microcentrifuge tube containing L-15 medium +2% FBS to a final volume of 47 μL ready for single-cell sequencing. Finally, 40 live single cells were sorted into 96-well PCR plates, immediately placed on dry ice, and then stored at −70°C.

Single-cell heteroplasmy measurements

96-well PCR plates containing single cells were centrifuged at 2150 rpm for 5 min at 4°C. 4 μL of lysis buffer containing 1% Tween 20, 50 mM Tris-HCl (pH 8.5) and 2 mg/mL Proteinase K was then added to each well. Following a quick centrifugation for 1 min at 1700 rpm, plates were incubated at 37°C for 30 min 8 μL of dH_2O was then added to each well and following a second centrifugation for 1 min at 1700 rpm, plates were further incubated at 80°C for 15 min to inactivate Proteinase K. 4 μL of the cell lysate was added to 21 μL of PyroMark PCR Kit reaction mix that included primers designed to amplify the region of mtDNA containing the mutation.

Measurement of heteroplasmy then proceeded using the QIAGEN Q48 Autoprep.[28]

10X chromium scRNA-seq

Single-cell RNA-seq libraries were prepared in the Cancer Research UK Cambridge Institute Genomics Core Facility using the following: Chromium Single Cell 3′ Library & Gel Bead Kit v3, Chromium Chip B Kit and Chromium Single Cell 3′ Reagent Kits v3 User Guide (Manual Part CG000183 Rev C; 10X Genomics). Cell suspensions were loaded onto the Chromium instrument with the expectation of collecting gel-bead emulsions containing single cells. RNA from the barcoded cells for each sample was subsequently reverse-transcribed on a C1000 Touch Thermal cycler (Bio-Rad) and all subsequent steps to generate single-cell libraries were performed according to the manufacturer's protocol with no modifications. cDNA quality and quantity were measured with an Agilent TapeStation 4200 (High Sensitivity 5000 ScreenTape) after which 25% of the material was used for gene expression library preparation.

Library quality was confirmed with an Agilent TapeStation 4200 (High Sensitivity D1000 ScreenTape) to evaluate library sizes, and Qubit 4.0 Fluorometer (Thermo Scientific Qubit dsDNA HS Assay Kit) to evaluate dsDNA quantity. Each sample was normalized and pooled at equimolar concentration. To confirm the final concentration, the pooled library was qPCRed using KAPA Library Quantification Kit on a QuantStudio 6 Flex before sequencing.

Initially a pool of 2 samples was sequenced on 1 lane of SP flowcell, which was then followed by a pool of 3 samples, also sequenced on 1 lane of SP flowcell. Sequencing was performed on an Illumina NovaSeq6000 with the following parameters: 28 bp, read 1; 8 bp, i7 index; and 91 bp, read 2.

scRNA-seq & gene expression analysis

Quality control and pre-processing

Raw FASTQ files from 10X sequencing were demultiplexed and aligned in a two-pass alignment to the *Mus musculus* reference genome (version Mus_musculus.GRCm38.102) using STARsolo (version 2.7.3a)[67] with default settings set according to the 10X chromium chemistry used. For the expression matrix analysis, we used Scanpy (version 1.7.2)[68] and perform standard pre-processing steps: removing cells with less than 200 genes, removing genes that have been detected in less than 3 cells, normalisation to 10,000 reads per cell, and log-transformation. After this quality control, we have 11,020 cells in total.

Clustering and cell-type annotation

To cluster the cells with Scanpy, we first construct a k-nearest-neighbour graph with the default k = 15 and, second, perform a modularity maximisation with resolution parameter 10. For each cluster, we then identify differentially expressed genes (DEGs) with a Wilcoxon rank-sum tests, for which we choose a significance threshold of 0.05 for and use the Benjamini–Hochberg procedure to obtain multiple-testing corrected p values. To annotate each cluster with a cell type, we compare the obtained top 100 DEGs with the top 100 marker genes as identified by ref. 15. Specifically, for each cluster, we compute the Jaccard index between the DEGs and each of the twenty cell types and then assign the cell type with the highest Jaccard index to this cluster.

For data processing for identification of marker genes, SCENIC regulons, and data visualisation on UMAP, raw counts were processed using Seurat (v.4.0.5) and SCTransform (sctransform R package, v.0.3.2) with regressing out batch effect, followed by Harmony (v.0.1.0). SCENIC analysis was performed on the integrated dataset using SCENIC R package (v.1.2.4) and python Arboreto module (v.0.1.5).Cell-type-specific, and cell-type-and-genotype-specific marker genes were calculated with Seurat function ''FindMarkers'', using a Wilcoxon rank-sum test with significance threshold of 0.05 and use of the Benjamini-Hochberg procedure[78] to obtain multiple-testing corrected p values.

Enrichment with buffering genes from ref. [9,31] was performed for each cell type & genotype specific markers (also for each regulon genes) using hypergeometric test with adjusting for multiple testing (Holm method). Heatmaps were produced using R package ComplexHeatmap (v.2.8.0) with hierarchical clustering using euclidean distance (for grouped by cell type- Figure 6, or individual cell type like in Figure 7A). Pathway analysis was performed using enrichR R package (v.3.0).

Pearson correlation of genes involved in oxidative phosphorylation (as annotated in mouse MitoCarta v3)[71] was calculated separately for WT, m.5019A>G and m.5024C>T embryos for selected cell types. To measure the number of correlated gene pairs (right) for each genotype, correlation data was binarised with a cut-off of >0.5 for correlation coefficient r^2, to count presence or absence of correlation between pairs of genes. Quantile-quantile (qq) plots were created to visually inspect changes in gene pair correlations between WT and mutant genotypes.

Proteome profiling of m.5019A>G and WT adult mouse tissues

Mouse tissues used for proteomics were collected within 10 min of cervical dislocation and directly snap frozen into liquid nitrogen. The tissues were then conserved at −70°C.

For liver samples, 100 mg of liver was homogenised using a Precelly's 24 (Bertin Instruments) in 300 μL ice-cold lysis buffer (1% SDS/50mM of HEPES buffer plus protease inhibitor (Abcam)). To remove debris, two successive centrifugation steps were carried out on the supernatant, each at 16,000 g for 10 min at 4°C. Protein concentration of the supernatant was estimated through a Pierce BCA Protein Assay Kit (Thermo Scientific) and 100 μg was submitted to the Cambridge Center for Proteomics (University of Cambridge) for TMT labeled Tandem Mass Spectrometry.

For brain samples a modified protocol from ref. 79 was followed. Briefly; 100 mg of cortex per sample was homogenised using a Precellys 24 (Bertin Instruments) in 300 μL ice-cold lysis buffer (10mM Tris-HCl (pH7.4), 0.15 mM NaCl, 1 mM EDTA, PBS containing 1% beta-OG (Sigma) (w/v), protease inhibitor). Subsequently, all samples were incubated at 4°C for 1h with mild agitation followed by centrifugation at 10,000 g for 30 min at 4°C. The supernatant was quantified using the Bradford assay and 100 μg was delipidated with a tri-*n*-butyl phosphate (TBP)/acetone/methanol mixture. The resulting pellet was washed sequentially with TBP, acetone, and methanol and air-dried. Samples were submitted to Cambridge Center for Proteomics (University of Cambridge) for TMT labeled Tandem Mass Spectrometry.

TMT labeling

TMT labelling was performed using a TMTpro 16plex Label Reagent Set (Thermo Scientific) according to the manufacturer's protocol. 100 μg of each digested protein sample was labeled individually with 12 TMTpro tags for the two experiments. Post-labelling, samples were combined and cleaned on a Sep-Pak C18 cartridge, dried and dissolved in 20 mM ammonium formate (pH 10). The solution was then pipetted into a sample vial and placed into the autosampler of a Waters Acquity UPLC pump (Waters Corporation, Milford MA).

High-pH first dimension reverse-phase fractionation

The following LC conditions were used for the fractionation of the TMT samples: desalted peptides were resuspended in 0.1 mL 20 mM ammonium formate (pH 10.0) + 4% (v/v) acetonitrile. Peptides were loaded onto an Acquity bridged ethyl hybrid C18 UPLC column (Waters; 2.1 mm i.d. x 150 mm, 1.7 μm particle size), and profiled with a linear gradient of 5-60% acetonitrile +20 mM ammonium formate (pH 10.0) over 60 min, at a flow-rate of 0.25 mL/min. Chromatographic performance was monitored by sampling eluate with a diode array detector (Acquity UPLC, Waters) scanning between wavelengths of 200 and 400 nm. Samples were collected in 1-min increments and reduced to dryness by vacuum centrifugation.

LC-MS/MS

Dried fractions from the high pH reverse-phase separations were resuspended in 40 μL of 0.1% formic acid and placed into a glass vial. 2 μL of each fraction was injected by the HPLC autosampler and separated by the LC method detailed below. Fractions were combined into pairs (i.e. the first fraction with the middle fraction etc) and were analyzed by LC-MS/MS.

LC-MS/MS experiments were performed using a Dionex Ultimate 3000 RSLC nanoUPLC (Thermo Scientific) system and a Lumos Orbitrap mass spectrometer (Thermo Scientific). Peptides were loaded onto a pre-column (Thermo Scientific PepMap 100C18, 5 mm particle size, 100 Å pore size, 300 mm i.d. x 5 mm length) from the Ultimate 3000 auto-sampler with 0.1% formic acid for 3 min at a flow rate of 15 μL/min. After this period, the column valve was switched to allow elution of peptides from the pre-column onto the analytical column. Separation of peptides was performed by C18 reverse-phase chromatography at a flow rate of 300 nL/min using a Thermo Scientific reverse-phase nano Easy-spray column (Thermo Scientific PepMap C18, 2mm particle size, 100 Å pore size, 75 mm i.d. x 50 cm length). Solvent A was water +0.1% formic acid and solvent B was 80% acetonitrile, 20% water +0.1% formic acid. The linear gradient employed was 2-40% B in 93 min. (Total LC run time was 120 min including a high organic wash step and column re-equilibration).

The eluted peptides from the C18 column LC eluant were sprayed into the mass spectrometer by means of an Easy-Spray source (Thermo Scientific). All *m/z* values of eluting peptide ions were measured in an Orbitrap mass analyzer, set at a resolution of 120,000 and were scanned between *m/z* 380-1500 Da. Data dependent MS/MS scans (Top Speed) were employed to automatically isolate and fragment precursor ions by collision-induced dissociation (CID, Normalised Collision Energy (NCE): 35%) which were analyzed in the linear ion trap. Singly charged ions and ions with unassigned charge states were excluded from being selected for MS/MS and a dynamic exclusion window of 70 s was employed. The top 10 most abundant fragment ions from each MS/MS event were then selected for a further stage of fragmentation by Synchronous Precursor Selection (SPS) MS3[80] in the HCD high energy collision cell using HCD (High energy Collisional Dissociation), (NCE: 65%). The *m/z* values and relative abundances of each reporter ion and all fragments (mass range from 100 to 500 Da) in each MS3 step were measured in the Orbitrap analyser, which was set at a resolution of 60,000. This was performed in cycles of 10 MS3 events before the Lumos instrument reverted to scanning the *m/z* ratios of the intact peptide ions and the cycle continued.

Proteomic data analysis

Proteome Discoverer v2.4 (Thermo Scientific) and Mascot (Matrix Science) v2.6 were used to process raw data files. Data were searched against the UniProt *M. musculus* database (63,550 entries) and with the common repository of contaminant proteins (cRAP, 125 sequences). Protein identification allowed an MS tolerance of ±10 ppm and an MS/MS tolerance of ±0.8 Da ppm along with permission of up to 2 missed tryptic cleavages. Quantification was achieved by calculating the sum of centroided reporter ions within a ±2 millimass unit (mmu) window.

Pre-processing and statistical analysis of the protein abundance were carried out using Perseus (v2.0.7.0)[70]: After the categorical annotation of groups, feature rows were filtered so that there are minimal three values in at least one group; Data were then normalised and logarithmised with missing values imputed from normal distribution; Two sample t-test was used to extract differential features applying the cut-off value of S0 = 1 and FDR = 0.05. Data correlation and principal component analysis were tested using MSnbase (v2.22.0).[81] The tissue specific expression pattern was checked against the Mouse Expression Atlas. Mitochondrial proteins and pathways were annotated referring to MitoCarta (v3.0).[71] Gene Set Enrichment Analysis was executed using clusterProfiler (v3.15),[72] with additional pathway enrichment analysis explored using Enrichr.[73]

Single-cell cloning of immortalized m.5024C>T MEF cells

Clonal populations were isolated by sorting single cells from a suspension of immortalized m.5024C>T MEF cells using a BD FACSMelody cell sorter. Cells were sorted into individual wells of a flat-bottomed 96-well plate containing 200 μL of culture medium. One week after sorting, wells were checked for presence of a single colony of cells. After two weeks, successful clones were expanded to 24-well plates and samples were taken for DNA extraction and heteroplasmy measurement. Selected low and high heteroplasmy clones were further expanded to T-75 flasks and used for RNA-seq and siRNA experiments.

Bulk RNA-seq of m.5024C>T MEF clones

$1x10^7$ cells per MEF clone were harvested following trypsinization and total RNA was extracted from the resulting cell pellets using the QIAGEN RNeasy Plus Mini Kit according to the manufacturer's instructions, including removal of genomic DNA via eliminator column. Purified RNA was eluted in RNAse-free water and RNA integrity was checked using Agilent RNA Tapestation and Qubit RNA HS assays. cDNA libraries were generated from 1000 ng of total RNA using the NEB PolyA Kit and NEB Ultra II Directional Library Kit. Library QC was performed using Agilent DNA 5000 Tapestation and Qubit dsDNA HS assays. Following QC, sample libraries were pooled and sequenced on a Novaseq 6000 SP flowcell with 50 bp paired end reads.

QC of the sequencing was assessed using FastQC (v0.11.9), fastq_screen (v0.14.1), Qualimap (v.2.2.1) and summarised with MultiQC (v1.11). Reads were aligned to the human genome (GRCh38) with STAR aligner (v2.7.8a), with a mean of 87.4% reads uniquely mapping and mean of 12.4 M paired reads/sample. Gene quantification was determined with STAR.

Counts extracted with STAR were used to perform differential gene analysis in R (version 4.2.1) using package DESeq2 (v.1.36.0). Since the samples were sequenced in pool on 2 separate lanes, counts were combined using DESeq2 function "collapseReplicates". Genes with no expression were filtered out. Read counts were normalised on estimated size factors. Fold changes were calculated with function lfcShrink using estimator "ashr". Principal component analysis (PCA) was performed on rlog-transformed count data for all genes. GO enrichment analysis was performed using enrichR package (v.3.1) and Kegg pathway analysis using clusterProfiler (v.4.4.4), on DEGs with absolute log2 fold change >1 and adjusted p value < 0.05.

siRNA knockdown of selected transcription factors in m.5024C>T MEF clones

Knockdowns were performed in 24-well plates using ON-TARGETplus siRNA SMARTPools from Dharmacon. Briefly, for each well, oligos were diluted to a concentration of 6 pmol in 100 μL of Opti-MEM I Reduced Serum Medium (Gibco) and 1 μL of Lipofectamine RNAiMAX reagent was added with gentle mixing, followed by incubation at room temperature for 10-20 min. Meanwhile, freshly trypsinised MEF cells were diluted to a concentration of 20,000 cells per mL in antibiotic-free culture medium. 500 μL of diluted cells were added to each well and mixed gently by rocking the plate back and forth. This resulted in a total volume of 600 μL per well, containing 10,000 cells and a final oligo concentration of 10 nM. Plates were then transferred to an IncuCyte S3 Live-Cell Analysis System (Sartorius) housed in an humidified incubator held at 37°C and 5% CO_2.

Cell proliferation analysis

Following siRNA knockdowns, cell proliferation rate was measured on an IncuCyte S3 Live-Cell Analysis System (Sartorius). Eight brightfield images (10× magnification) per well were recorded at 3-h intervals. A confluency mask was created using the IncuCyte software (v2020C) and applied to all images to calculate a confluency % for each image. For each timepoint, mean confluency and SEM were calculated for each well using the eight individual images.

Two independently conducted proliferation experiments were merged and means for each clone calculated. For the late time point (60h), Pearson correlation was used to assess correlation between heteroplasmy and mean confluency. Two statistical methods were used to assess change in confluency due to heteroplasmy levels: t-test at the late time point (60h) and area under curves (AUC).

Additional datasets

The Tabula Muris Consortium data was used to compare with expression of genes involved in oxidative phosphorylation in E8.5 WT mouse embryo. Bulk RNA-seq stages: 1 month and 9-month-old, were selected because there were highest number of replicates for these stages, and they were representing post-natal young and adult mouse. Raw counts were processed with DESeq2, and variance stabilizing transformation was used to visualise gene expression on the heatmap together with the mouse embryo.

In order to validate genes identified in SCENIC regulons, Sox10 regulon. Two separate Sox10 siRNA RNA-seq datasets (GSE160312 in human pluripotent stem cells,[42] and GSE152687 in mouse Schwann cells,[43] were used to identify genes regulated by the transcription factor (TF) Sox10, and were compared with the SCENIC Sox10 regulon genes that were predicted to be regulated by Sox10 TF. For each regulon identified by SCENIC (with minimum 30 genes), percentage of overlap with the siRNA datasets was calculated, and plotted as a line-plot with regulons ordered by the percent overlap.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data in Figures 2C and S2 were analyzed in GraphPad Prism v9. Details of the tests are found in the accompanying figure legend.

Supplemental figures

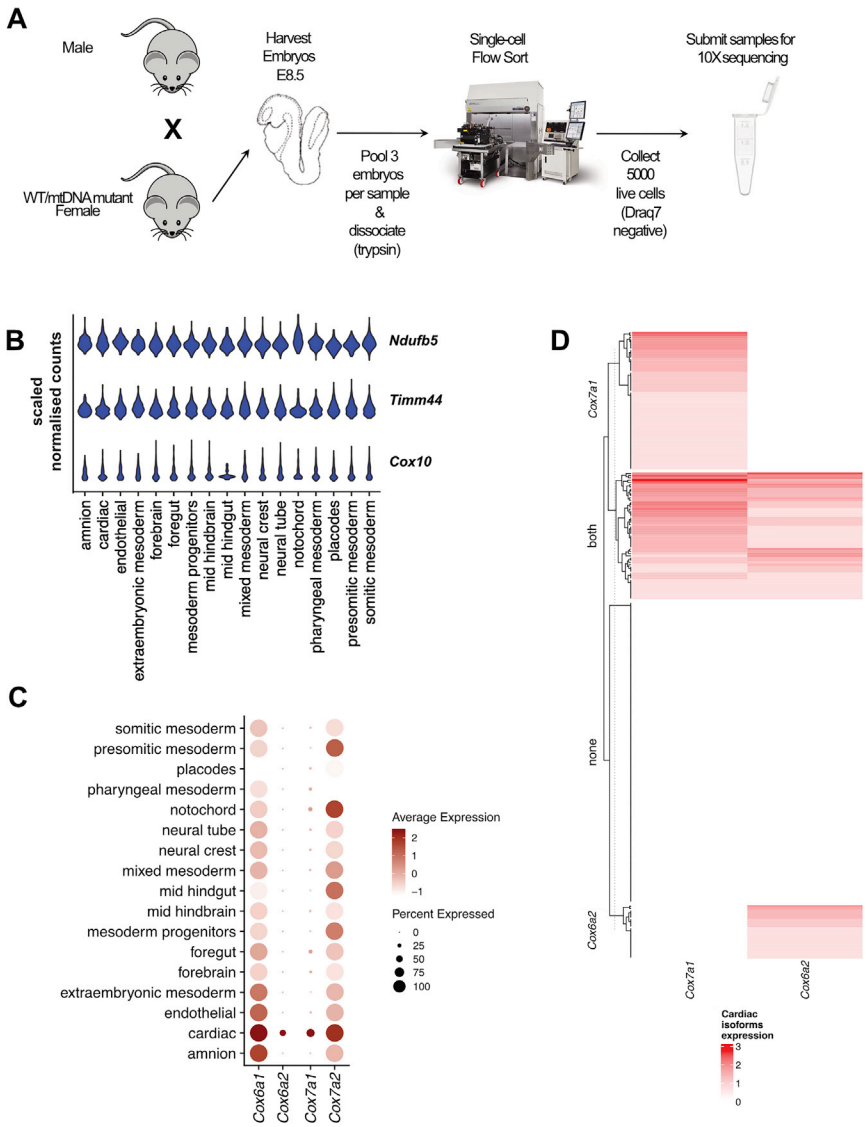


Figure S1. Differential expression of nuclear-encoded mitochondrial genes during organogenesis, related to Figure 1

(A) Overview of the experimental isolation and preparation of E8.5 embryonic cells for single-cell RNA-seq.

(B) Violin plots showing expression profiles of illustrative nuclear-encoded mitochondrial genes *Ndufb5*, *Timm44*, and *Cox10*, showing no significant differential expression across the 17 cell lineages. For the full dataset see Table S1.

(C) Dot plot showing relative expression of the *Cox6a1*, *Cox6a2*, *Cox7a1*, and *Cox7a2* gene isoforms in the 17 identified cell lineages of WT E8.5 embryos. Diameter of each dot is proportional to the percentage of cells expressing each isoform in the corresponding lineage, darker shading indicates higher average expression in that lineage.

(D) Heatmap showing relative expression of the heart/muscle specific Complex IV subunit isoforms *Cox7a1* and *Cox6a2* in WT E8.5 cardiac lineage cells. The majority of *Cox6a2* expressing cells also express *Cox7a1*, suggesting co-expression of the two isoforms in early cardiomyocytes. WT, wild-type

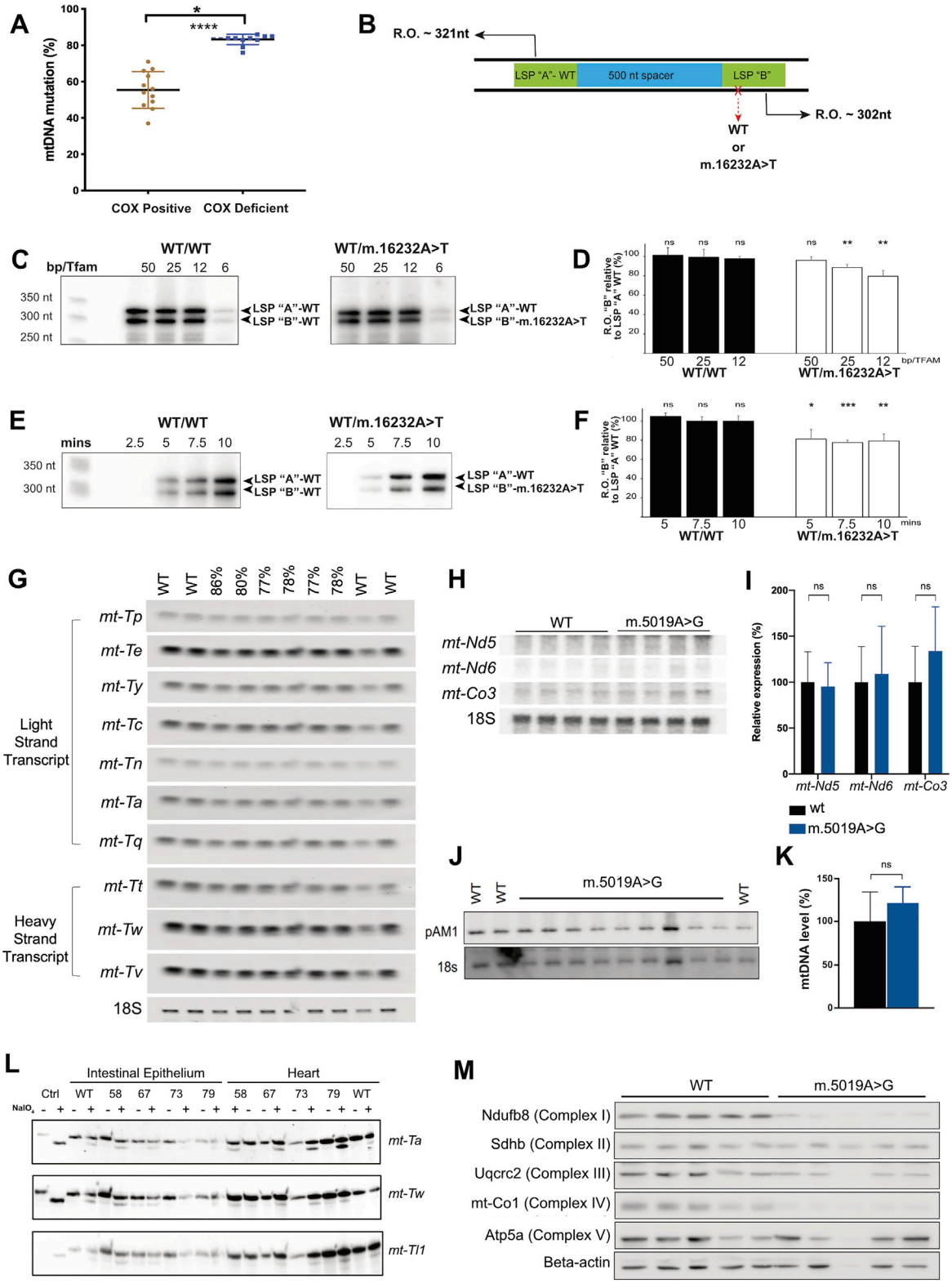


Figure S2. Characterization of the m.5019A>G *mt-Ta* mutation, related to Figure 2

(A) Heteroplasmy levels of COX-positive and COX-deficient colonic crypts isolated from m.5019A>G mice measured by pyrosequencing, showing a significant increase of mutation level in COX-deficient crypts. Mann-Whitney U Test.

(legend continued on next page)

(B) Schematic of the plasmid used for the m.16232A>T *in vitro* competitive transcription assay. The WT LSP sequence (LSP "A", left) leads to a 331nt transcript, and the m.16232A>T mutation leads to the smaller 314nt transcript (LSP "B", right).
(C) Autoradiograms of wild-type (WT) and m.16232A>T transcripts under varying DNA/Tfam concentrations. Left are two WT promoters, and right are the WT and m.116232A>T promoters.
(D) Bar graph showing quantification of radiography signals from Figure S2C. Preferential transcription of the WT sequence over m.16232A>T is seen as Tfam concentration increases. bp/Tfam = number of DNA base pairs per molecule of Tfam.
(E) Time course experiment of the competitive transcription assay. Left are two WT promoters, and right are the WT and m.16232A>T promoters.
(F) Bar graph showing quantification of radiography signals from Figure S2E. Higher relative transcription of WT vs m.16232A>T is seen at all experimental timepoints.
(G) Northern blots of steady-state mitochondrial tRNAs from heart tissue of WT and m.5019A>G mice with varying levels of mutation, showing no significant difference between WT and m.5019A>G for any of the assayed tRNAs. 18s rRNA shown as loading control.
(H) Northern blots of *mt-Nd5*, *mt-Nd6*, and *mt-Co3* transcripts from small intestinal epithelium of WT and m.5019A>G mice.
(I) Quantification of the northern blot signals from Figure S2H, with no significant differences in expression level seen for any of the transcripts. Mann-Whitney U test. 18S rRNA shown as loading control.
(J) Southern blot of relative mtDNA copy number from heart tissue of WT and m.5019A>G mice detected using the pAM1 plasmid.
(K) Quantification of the Southern blot signals from Figure S2J, revealing no significant difference in copy number between WT and m.5019A>G tissues. Mann-Whitney U test. 18S rRNA shown as loading control.
(L) High-resolution northern blot of MEFs from m.5019A>G or WT mice, showing aminoacylation states of *mt-Ta*, *mt-Tw*, and *mt-Tl1* in the presence ('+') or absence ('- ') of $NaIO_4$. Incubation of freshly extracted RNA with $NaIO_4$ results in the removal of a single 3' ribose on all tRNA not charged with an amino acid, leaving the charged tRNAs unaffected. Following the removal of amino acids from the charged tRNA, the initial ratio of charged to uncharged tRNA in the live cells can be visualized by high-resolution Urea PAGE, which can separate RNA molecules differing in molecular weight by a single ribose. A specific *mt-Ta* aminoacylation defect is seen in m.5019A>G heart tissue when exposed to $NaIO_4$. 5S rRNA shown as loading control. Ctrl, RNA samples intentionally deacylated (alkaline treatment) and incubated with $NaIO_4$ or NaCl.
(M) Western blots of selected mitochondrial complex subunits in small intestinal epithelium of m.5019A>G and WT mice.

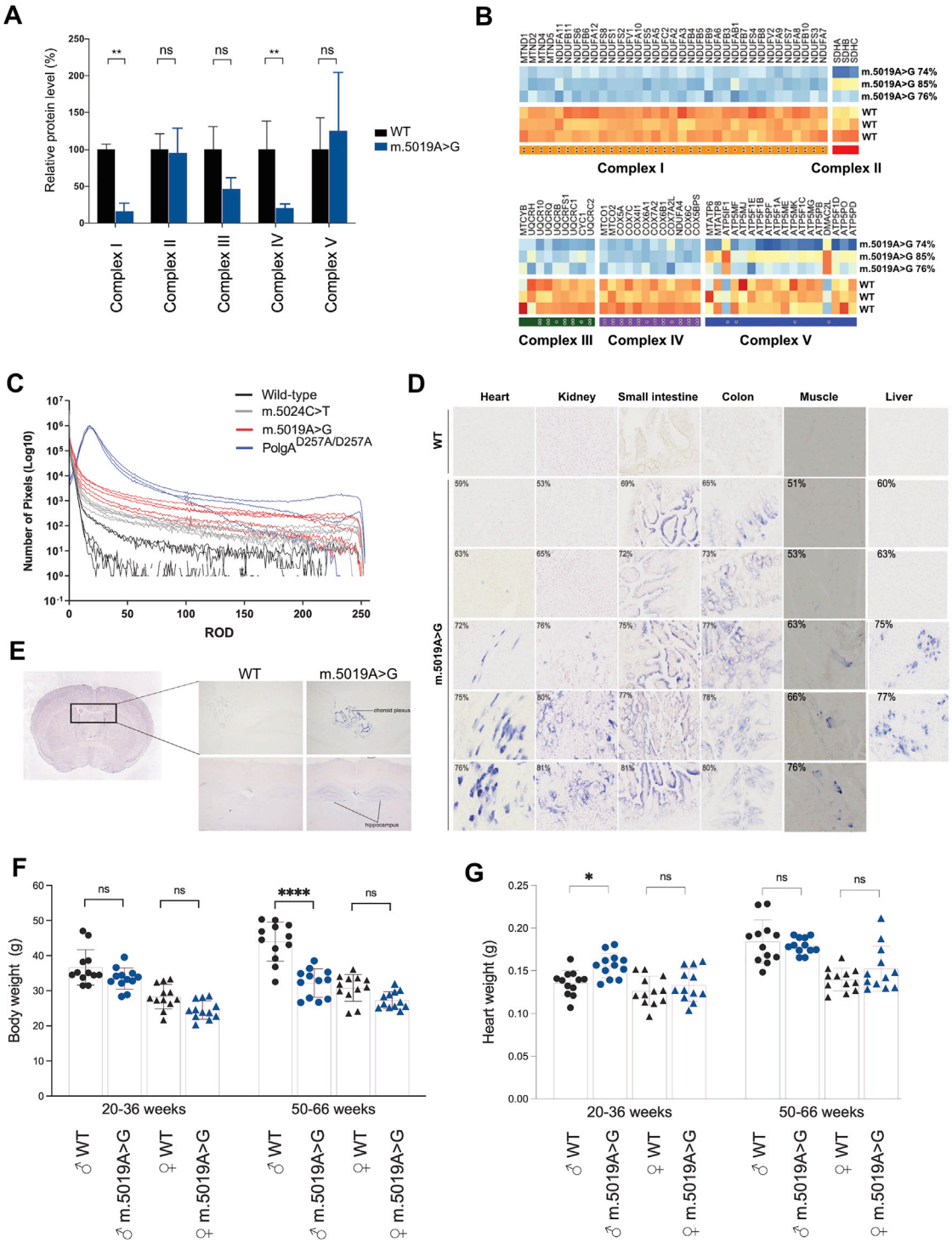


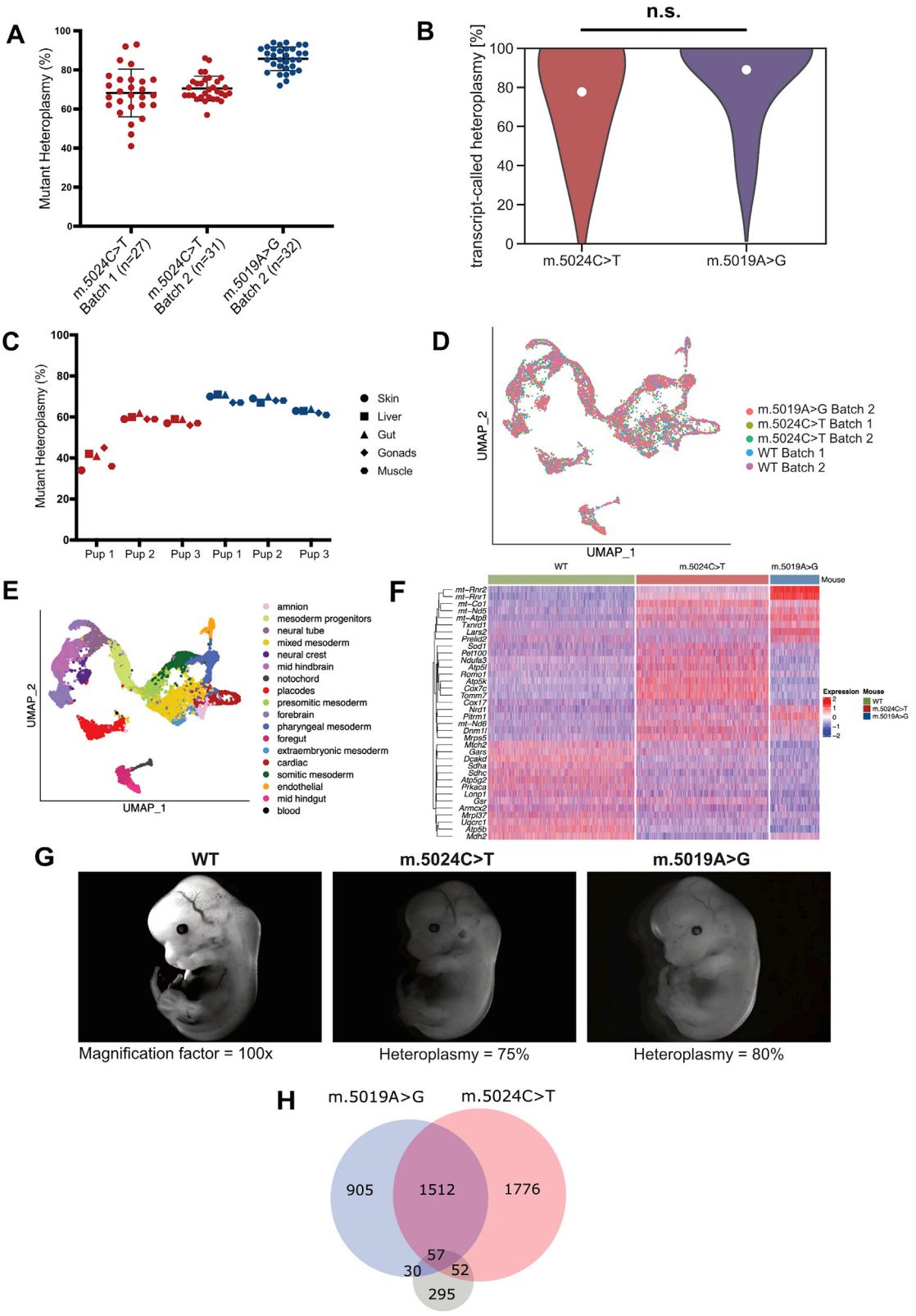
Figure S3. Additional characterization of the m.5019A>G *mt-Ta* mutation, related to Figure 2

(A) Quantification of western blots in Figure S2M, showing significant reduction of Ndufb8 (complex I) and mt-Co1 (complex IV) at the protein level in m.5019A>G mice. Sidak's multiple comparison test from a two-way ANOVA.

(B) Heat maps showing the protein expression ratio of respiratory complex subunits found in liver samples from m.5019A>G mice and WT controls. Comparing the m.5019A>G samples to their WT counterparts, reveals a significant decrease in the protein levels of most subunits belonging to respiratory complexes I, III, and IV.

(legend continued on next page)

(C) Pixel color analysis of heart sections quantifying relative amount of NBT staining of histology images from WT, m.5024C>T, m.5019A>G and *PolgA*$^{D257A/D257A}$ mice. x axis represents white to dark blue pixel colors (left to right) and y axis the $\log_{10}$ of the number of pixels with that color.
(D) Representative NBT histology sections of various tissues from WT and m.5019A>G mice of varying mutation levels. COX deficiency increases with m.5019A>G heteroplasmy across all tissues assayed, with the most striking effect seen in small intestine and colon. Panel showing m.5019A>G 75% Heart also appears in Figure 2F.
(E) Brain sections from 100-week-old WT and m.5019A>G mice stained with NBT. Blue-staining COX-deficient cells are visible in the choroid plexus and hippocampus of m.5019A>G mice but not in WT controls.
(F) Body weight measurements from male and female WT and m.5019A>G mice at 20–36 weeks and 49–68 weeks of age. Body mass of the mutant mice was decreased due to a lower accumulation of body fat compared to wild-type animals. n = 9 animals per group. Mann-Whitney U test.
(G) Heart weight measurements from male and female WT and m.5019A>G mice at 20–36 weeks and 49–68 weeks of age. Hearts from m.5019A>G mice did not show evidence of enlargement, unlike m.5024C>T mice at the same age, which showed signs of developing cardiomyopathy. n = 9 animals per group. Mann-Whitney U test. COX, Cytochrome *c* Oxidase, WT, wild-type, LSP, light strand promoter, ROD, relative optical density; NBT, nitrotetrazoleum blue; ns, not significant, *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, error bars indicate SD.



(legend on next page)

Figure S4. *mt-Ta* variants induce an embryo-wide transcriptional response, related to Figures 3 and 4

(A) Representative single-cell heteroplasmy measurements from cells isolated from the m.5024C>T and m.5019A>G embryo pools used for 10X single-cell RNA-seq. Heteroplasmy measured by pyrosequencing, horizontal bars represent population mean.

(B) Single-cell heteroplasmy calls from scRNA-seq data based on m.5024C>T and m.5019A>G transcript variants. Wilcoxon Rank-Sum test revealed no significant difference between the heteroplasmy distributions in the m.5024C>T & m.5019A>G datasets ($p = 0.308$).

(C) Mean bulk tissue heteroplasmy measurements from various tissues of three m.5024C>T (red) and three m.5019A>G (blue) littermates taken at P0 (day of birth). Similar heteroplasmy levels were seen across multiple tissue types, suggesting that tissue-specific differential segregation of heteroplasmy does not occur widely during embryogenesis in these mouse strains. Heteroplasmy measured by pyrosequencing.

(D) UMAP showing overlay of all cells passing QC from WT, m.5024C>T, and m.5019A>G embryos. Cells colored according to batch with no evidence of strong batch effects seen.

(E) UMAP showing overlay of all cells passing QC from WT, m.5024C>T, and m.5019A>G embryos. Cells colored according to cell type, confirming that cells cluster strongly according to lineage across multiple batches. NB: Blood cell lineage was only identified in m.5019A>G embryos.

(F) Heatmap showing relative expression levels of mtDNA and nuclear encoded mitochondrial gene transcripts in m.5019A>G, m.5024C>T, and WT embryonic cells.

(G) Representative images of WT, m.5024C>T, and m.5019A>G embryos, with corresponding heteroplasmy measurements, at E13.5, showing comparative morphology at the end of organogenesis. Embryos appear normal for this developmental stage across all mouse strains. Images taken with a combined magnification factor of 100× (10× objective, 10× eyepiece).

(H) Venn diagram showing the intersection between published genetic modifiers of mitochondrial dysfunction identified through an *in vitro* CRISPR screen (gray)[31] and dysregulated genes in m.5019A>G (blue) and m.5024C>T (red) embryos. WT, wild-type; UMAP, Uni-form Mani-fold Approximation and Projection; CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats.

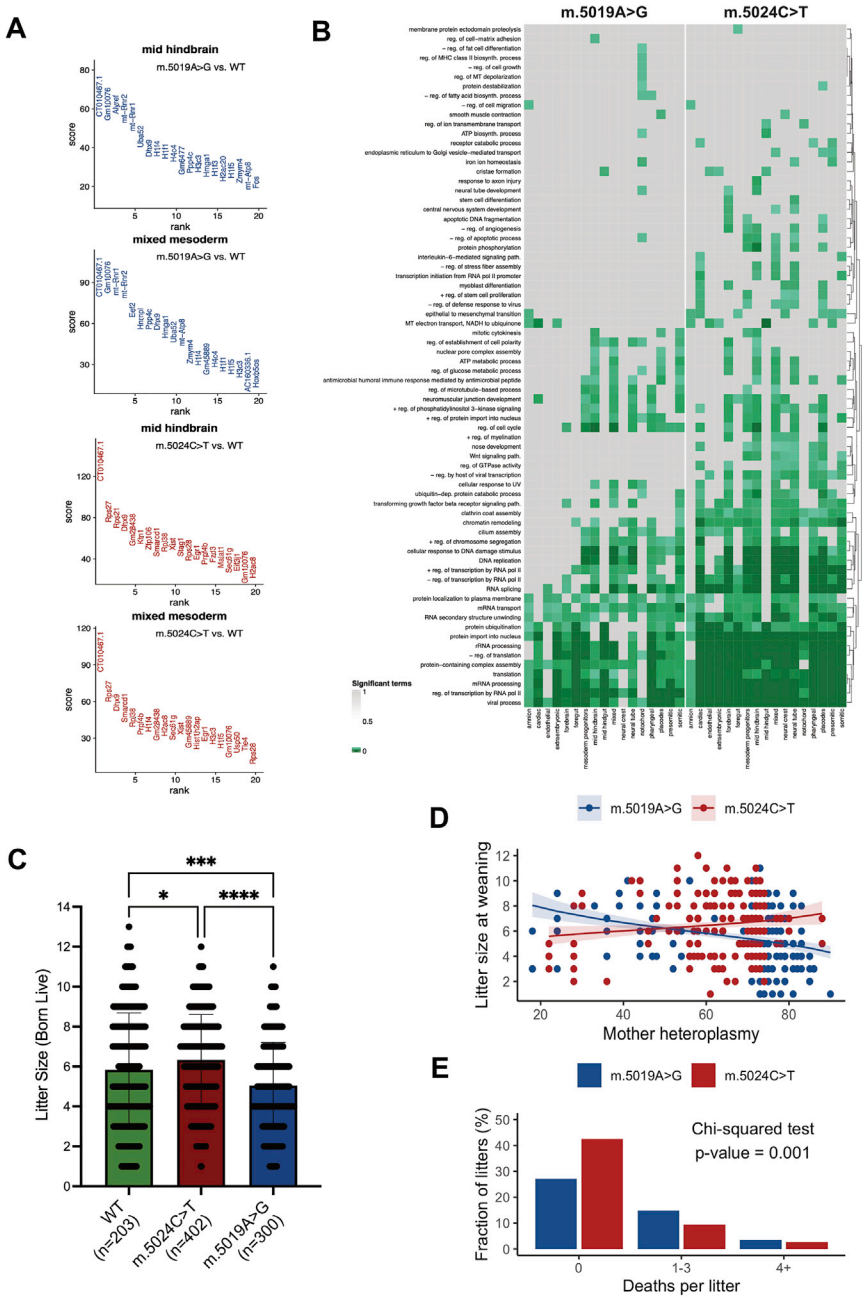


Figure S5. Mutation- and cell lineage-specific transcriptional signatures, related to Figure 4

(A) Top 20 most upregulated genes, ranked left to right by relative transcript expression score ($-\log_{10}$(p.adj) * abs(l2fc)) compared to WT, in the mid hindbrain and mixed mesoderm lineages for both m.5019A>G and m.5024C>T embryos. Ranked lists for other lineages can be found in Table S5.

(B) Heatmap showing significantly enriched GO Biological Processes (2018 annotation), based on differentially regulated genes for each cell lineage in m.5024C>T and m.5019A>G embryos compared to WT. $p < 0.05$ for all green squares.

(C) Litter sizes (weaned pups) for WT, m.5024C>T and m.5019A>G mice. Average litter size for the m.5019A>G strain was significantly smaller than both WT and m.5024C>T strains. p values calculated by one-way ANOVA with Tukey's correction for multiple comparisons.

(D) Scatterplot of litter size at weaning vs. mother heteroplasmy (skin biopsy) for m.5024C>T and m.5019A>G mouse colonies. A significant reduction in litter size was seen at high mother heteroplasmy for the m.5019A>G strain, but not for the m.5024C>T strain. Fit lines represent quasi-Poisson GLM with 95% confidence intervals.

(E) Bar chart showing number of pre-wean deaths recorded in m.5024C>T and m.5019A>G litters. Chi-squared test confirmed significantly higher incidence of pre-wean deaths in m.5019A>G litters compared to m.5024C>T. GO, Gene Ontology; WT, wild-type; ns, not significant; ****$p < 0.0001$, error bars represent SD.

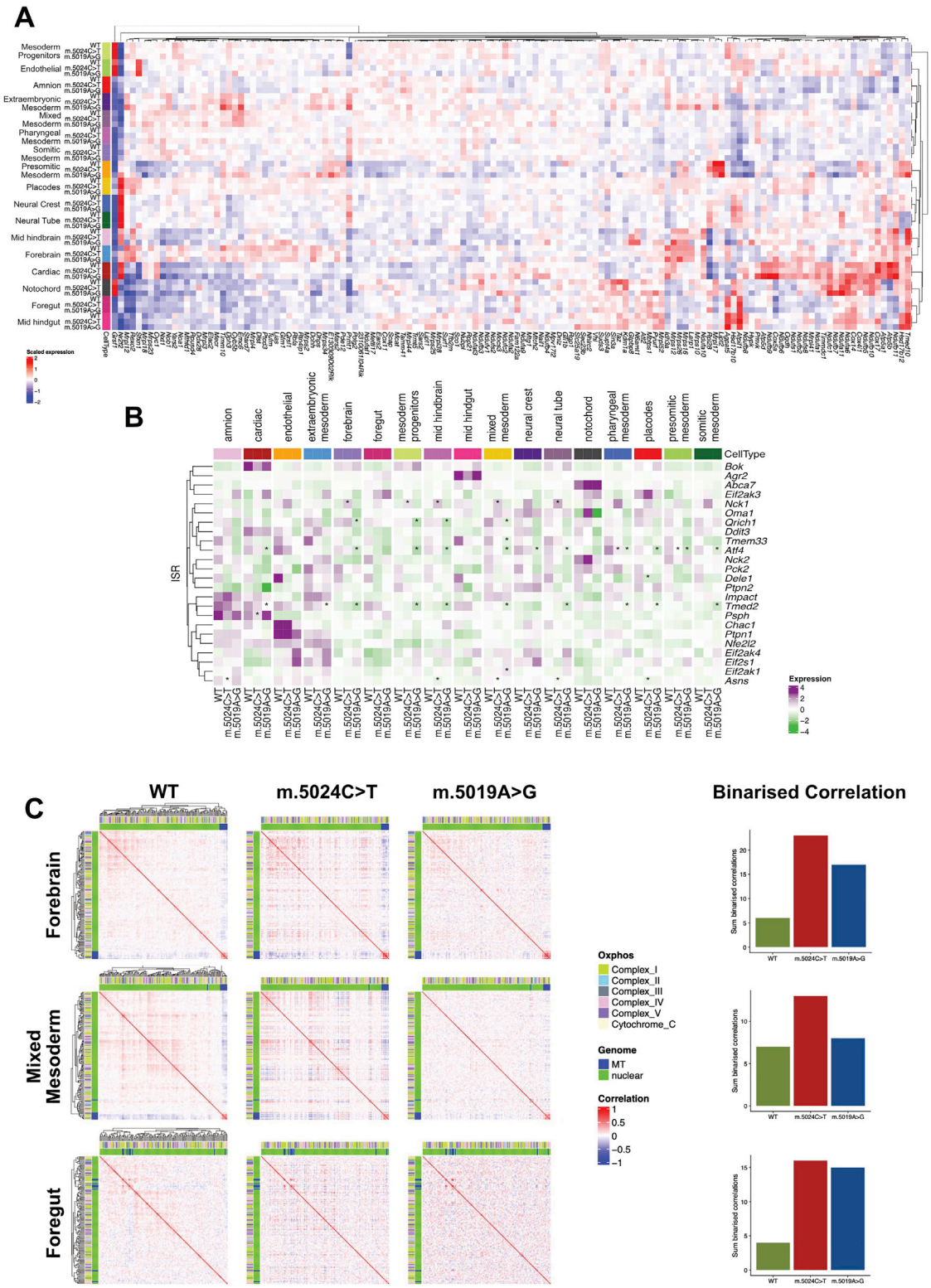


Figure S6. Mutation- and cell lineage-specific expression patterns of buffering genes in response to the *mt-Ta* mutations, related to Figure 5

(A) Unfiltered heatmap showing relative expression levels of epistatic buffer/suppressor genes identified in CRISPR knockout screens across all cell lineages and mouse strains. CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats.

(legend continued on next page)

(B) Heatmap showing relative expression levels of ISRmt genes in WT, m.5024C>T, and m.5019A>G cells across all cell lineages. Transcription of the key ISRmt regulator *Atf4* was downregulated in multiple lineages, particularly in m.5019A>G cells, suggesting that ISRmt is not activated in the mutant strains. ISRmt gene list consists of genes annotated under the Gene Ontology term ''Integrated Stress Response Signaling'' (GO ID: 0140467) plus known *Atf4* target genes.[37] *$p < 0.05$

(C) Correlation between individual gene expression levels of nuclear and mtDNA-encoded mitochondrial transcripts (MitoCarta v3.0) for the forebrain, mixed mesoderm, and foregut cell lineages. Left, correlation matrices for WT, m.5024C>T, and m.5019A>G mice. Outer bar on each axis shows the OXPHOS subunits corresponding to each gene. Inner bar on each axis shows whether the genes are expressed from nuclear DNA or mtDNA. Right, sum of the binarized correlations from each matrix, where a correlation with $r^2 > 0.5$ is counted.

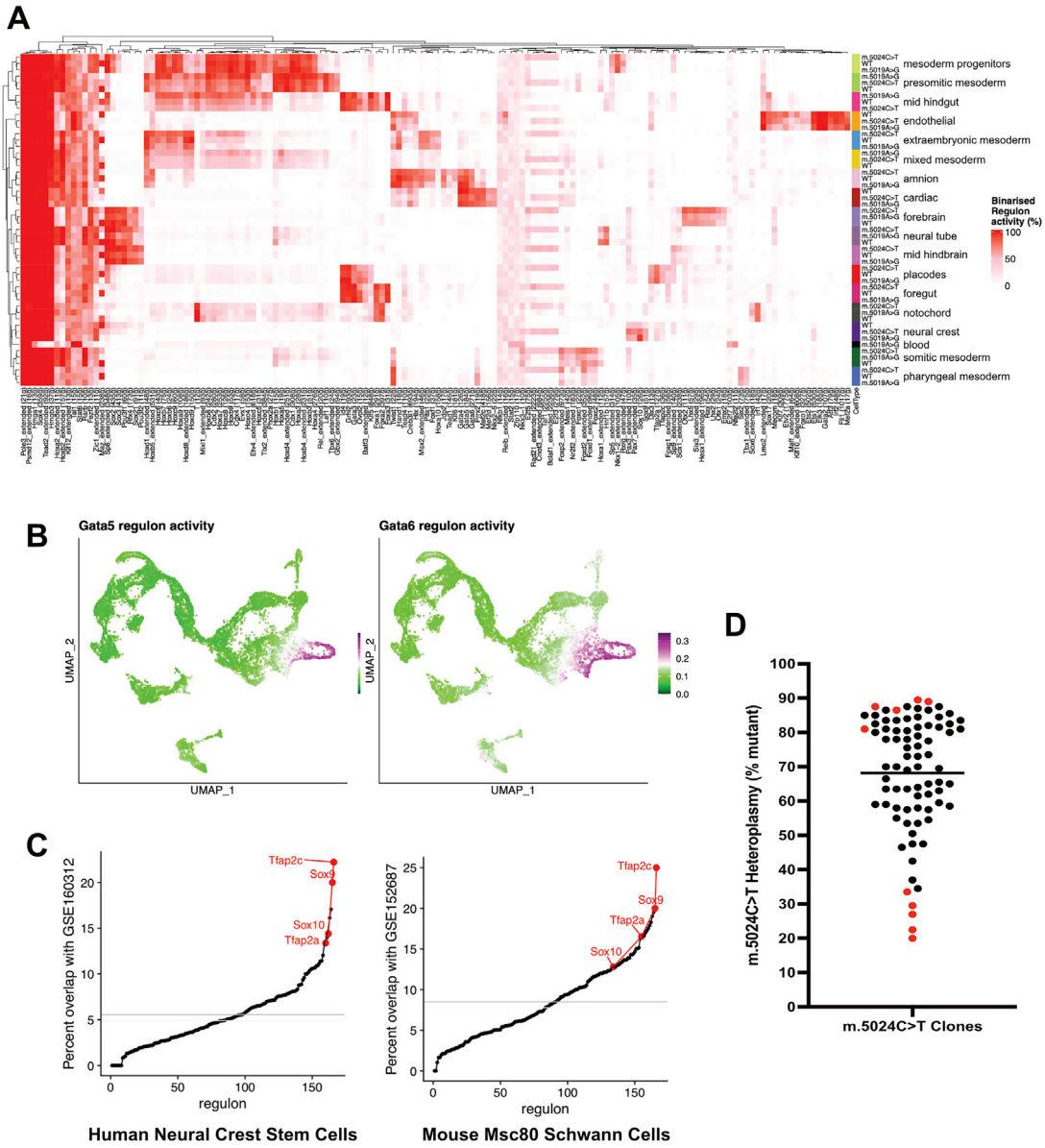


Figure S7. Cell lineage-specific compensatory mechanisms are coordinated by gene regulatory networks, related to Figures 6 and 7

(A) SCENIC heatmap showing binary activity of identified regulons. A regulon is defined as "active" by having a minimum of 50% cells showing regulon activity in at least one group. Both embryo-wide and lineage-specific patterns of regulon activity are seen in addition to mutation specific signatures for some regulons.

(B) UMAPs showing localized activity of the heart-specific Gata5 and Gata6 regulons, showing increased regulon activity in cardiac lineage cells. WT, m.5024C>T, and m.5019A>G cells are overlaid on the same UMAP. Red, increased regulon activity.

(C) Scatterplots showing the percentage overlap between marker genes in SCENIC regulons (Figure S7A) and differentially regulated genes identified in two published RNA-seq datasets comparing WT and Sox10 knockdown in human (GEO: GSE160312)[42] and mouse (GEO: GSE152687)[43] neural crest-derived cells. Neural crest-specific regulons identified by SCENIC, including Sox10, Sox9, Tfap2a, and Tfap2c, were amongst those with the highest overlap, confirming the validity of the SCENIC analysis.

(D) Heteroplasmy measurements of 80 single cell clones isolated from cultured immortalized MEF cells derived from m.5024C>T embryos. Measurements were made by pyrosequencing in technical triplicate from bulk-extracted DNA (mean heteroplasmy for each clone shown). Clones highlighted in red were selected as 'high' and 'low' heteroplasmy clones used for validatory experiments described in Figure 7. UMAP, Uni-form Mani-fold Approximation and Projection; WT, wild-type; MEF, mouse embryonic fibroblast.