Title

Genetically and metabolically corrected pluripotent stem cells from patients with mtDNA disease

Authors

Koski, Amy
Ma, Hong
Folmes, Clifford DL
et al.

Publication Date

2015-09-01

DOI

10.1016/j.mito.2015.07.106

Peer reviewed

LETTER

doi:10.1038/nature14546

Metabolic rescue in pluripotent cells from patients with mtDNA disease

Hong Ma[1,2], Clifford D. L. Folmes[3], Jun Wu[4], Robert Morey[5], Sergio Mora-Castilla[5], Alejandro Ocampo[4], Li Ma[4], Joanna Poulton[6], Xinjian Wang[7], Riffat Ahmed[1,2], Eunju Kang[1,2], Yeonmi Lee[1,2], Tomonari Hayama[1,2], Ying Li[1,2], Crystal Van Dyken[1,2], Nuria Marti Gutierrez[1,2], Rebecca Tippner-Hedges[1,2], Amy Koski[1,2], Nargiz Mitalipov[1,2], Paula Amato[8], Don P. Wolf[2], Taosheng Huang[7], Andre Terzic[3], Louise C. Laurent[5], Juan Carlos Izpisua Belmonte[4] & Shoukhrat Mitalipov[1,2]

Mitochondria have a major role in energy production via oxidative phosphorylation[1], which is dependent on the expression of critical genes encoded by mitochondrial (mt)DNA. Mutations in mtDNA can cause fatal or severely debilitating disorders with limited treatment options[2]. Clinical manifestations vary based on mutation type and heteroplasmy (that is, the relative levels of mutant and wild-type mtDNA within each cell)[3,4]. Here we generated genetically corrected pluripotent stem cells (PSCs) from patients with mtDNA disease. Multiple induced pluripotent stem (iPS) cell lines were derived from patients with common heteroplasmic mutations including 3243A>G, causing mitochondrial encephalomyopathy and stroke-like episodes (MELAS)[5], and 8993T>G and 13513G>A, implicated in Leigh syndrome. Isogenic MELAS and Leigh syndrome iPS cell lines were generated containing exclusively wild-type or mutant mtDNA through spontaneous segregation of heteroplasmic mtDNA in proliferating fibroblasts. Furthermore, somatic cell nuclear transfer (SCNT) enabled replacement of mutant mtDNA from homoplasmic 8993T>G fibroblasts to generate corrected Leigh-NT1 PSCs. Although Leigh-NT1 PSCs contained donor oocyte wild-type mtDNA (human haplotype D4a) that differed from Leigh syndrome patient haplotype (F1a) at a total of 47 nucleotide sites, Leigh-NT1 cells displayed transcriptomic profiles similar to those in embryo-derived PSCs carrying wild-type mtDNA, indicative of normal nuclear-to-mitochondrial interactions. Moreover, genetically rescued patient PSCs displayed normal metabolic function compared to impaired oxygen consumption and ATP production observed in mutant cells. We conclude that both reprogramming approaches offer complementary strategies for derivation of PSCs containing exclusively wild-type mtDNA, through spontaneous segregation of heteroplasmic mtDNA in individual iPS cell lines or mitochondrial replacement by SCNT in homoplasmic mtDNA-based disease.

Maternally inherited mtDNA encodes 13 proteins critical for oxidative phosphorylation, while the remaining protein subunits are encoded by nuclear DNA. Therefore, mitochondrial biogenesis requires coordinated interaction of protein subunits encoded by both genomes[2]. Mutations in mtDNA occur at a higher rate than in nuclear DNA, resulting in life-threatening conditions[3,6].

We have described a strategy to prevent transmission of mtDNA mutations to children involving mitochondrial replacement[7]. To explore the feasibility of generating genetically corrected autologous PSCs, herein, we focus on three of the most common pathogenic mtDNA mutations. Skin samples were donated by a MELAS patient carrying a 3243A>G heteroplasmic mutation in $tRNA^{Leu}$ (*MT-TL1*)[8] and by Leigh syndrome patients carrying heteroplasmic or homoplasmic 8993T>G mutations affecting the ATPase 6 gene (*MT-ATP6*)[3], and heteroplasmic 13513G>A mutation in the *MT-ND5* gene[9]. A panel of ten iPS cell lines from each mutation type was generated and quantitative mtDNA mutation analysis was carried out using amplification refractory mutation system-quantitative polymerase chain reaction (ARMs-qPCR), with a detection threshold of 0.5%. In MELAS iPS cell lines, the mutation was undetectable in five lines and varied from 33% to 100% in the remaining five lines, compared to 29% heteroplasmy in parental fibroblasts (Table 1 and Extended Data Fig. 1a). In iPS cell lines from the heteroplasmic 8993T>G mutation, the mutation was undetectable in one line and ranged from 29% to 87% in the remaining lines, compared to 52% heteroplasmy in parental fibroblasts (Table 1 and Extended Data Fig. 1b). Mutation segregation in individual iPS cell lines from 13513G>A fibroblasts also ranged from 0% to 100%, compared to 84% heteroplasmy in fibroblasts (Table 1 and Extended Data Fig. 1c). Previous studies suggested that segregation of heteroplasmic mtDNA is specific to iPS cells and may occur during or after reprogramming[10,11]. To explore mechanisms, parental fibroblasts carrying 3243A>G and 13513G>A mutations were subcloned and mutation loads in individual clones were analysed. Among ten randomly selected MELAS samples, five were homoplasmic containing either wild type (A) or mutant (G) at the 3243 position. The remaining five contained varying heteroplasmy levels similar to iPS cells (Table 1 and Extended Data Fig. 1d). Variable heteroplasmy levels were also observed in 13513G>A fibroblasts including homoplasmic mutant and wild-type clones (Table 1). Thus, segregation of heteroplasmic mtDNA mutations occurs in skin fibroblasts and may reflect a common phenomenon[12].

Isogeneic MELAS iPS cell lines carrying wild-type or mutant mtDNA maintained typical PSC morphology and formed teratomas containing cells and tissues from all three germ layers (Extended Data Fig. 2a, b). We next carried out whole mtDNA sequencing using the Illumina MiSeq platform and confirmed the 3243A>G mutation in parental MELAS-fib (46.8%), MELAS-iPS1 and MELAS-iPS3 (100%) cells while MELAS-iPS2 was homoplasmic for the wild-type allele (Supplementary Table 1). MELAS-fib also carried four additional heteroplasmic mutations with one variant carried to MELAS-iPS1 and MELAS-iPS2 (Supplementary Table 1).

The 3243A>G mutation perturbs $tRNA^{Leu}$ function and impairs mitochondrial protein synthesis as well as respiratory complex activity, with the homoplasmic mutation leading to prenatal lethality in humans[13]. Oxygen consumption rate (OCR) was employed as an indicator of mitochondrial respiration and energy production. Mutant MELAS-iPS1 and MELAS-iPS3 exhibited significantly lower OCR

[1]Center for Embryonic Cell and Gene Therapy, Oregon Health & Science University, 3303 S.W. Bond Avenue, Portland, Oregon 97239, USA. [2]Division of Reproductive & Developmental Sciences, Oregon National Primate Research Center, Oregon Health & Science University, 505 N.W. 185th Avenue, Beaverton, Oregon 97006, USA. [3]Center for Regenerative Medicine and Department of Medicine, Division of Cardiovascular Diseases, Mayo Clinic, Rochester, Minnesota 55905, USA. [4]Gene Expression Laboratory, Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. [5]Department of Reproductive Medicine, University of California, San Diego, Sanford Consortium for Regenerative Medicine, 2880 Torrey Pines Scenic Drive, La Jolla, California 92037, USA. [6]Department of Obstetrics and Gynaecology, John Radcliffe Hospital, University of Oxford, Headington, Oxford OX3 9DU, UK. [7]Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, USA. [8]Division of Reproductive Endocrinology, Department of Obstetrics and Gynecology, Oregon Health and Science University, 3181 Southwest Sam Jackson Park Road, Portland, Oregon 97239, USA.

Table 1 | Distribution of mtDNA variants in fibroblast and iPS clones derived from patients with heteroplasmic mutations

iPS clones						Fibroblast clones			
3243 A>G	Mutant G (%)	8993 T>G	Mutant G (%)	13513 G>A	Mutant A (%)	3243 A>G	Mutant G (%)	13513 G>A	Mutant A (%)
Parental fibroblasts	29	Parental fibroblasts	52	Parental fibroblasts	84	Parental fibroblasts	29	Parental fibroblasts	84
iPS1	100	iPS1	62	iPS1	100	fib1	100	fib1	0
iPS2	0	iPS2	72	iPS2	2	fib2	100	fib2	68
iPS3	100	iPS3	32	iPS3	4	fib3	0	fib3	24
iPS4	0	iPS4	52	iPS4	0	fib4	93	fib4	64
iPS5	0	iPS5	29	iPS5	80	fib5	8	fib5	58
iPS6	33	iPS6	66	iPS6	11	fib6	21	fib6	48
iPS7	0	iPS7	87	iPS7	19	fib7	3	fib7	69
iPS8	78	iPS8	72	iPS8	32	fib8	97	fib8	70
iPS9	88	iPS9	46	iPS9	100	fib9	100	fib9	63
iPS10	0	iPS10	0	iPS10	72	fib10	0	fib10	100

($P < 0.05$) when compared to the wild-type MELAS-iPS2 (Fig. 1a). Fibroblasts differentiated from MELAS-iPS1 and parental MELAS-fib also displayed low levels of mitochondrial function. In contrast, these respiratory defects were absent in MELAS-iPS2-derived fibroblasts. In general, mitochondrial respiration correlated with the heteroplasmy levels in cells (Fig. 1b and Table 1). The greater reliance on oxidative metabolism in wild-type MELAS-iPS2 was confirmed by the elevated OCR to ECAR (extracellular acidification rate) ratio, which provides a measure for the relative contribution of oxidative metabolism versus glycolysis (Extended Data Fig. 3a). Mutant MELAS iPS cells and their derivatives displayed significantly decreased OCR/ECAR ratios, indicating a greater reliance on glycolysis (Extended Data Fig. 3a, b). We next differentiated MELAS iPS cells into neuronal progenitor cells (NPCs, Extended Data Fig. 3c, d)[14,15]. Diminished metabolic profiles in mutant NPCs recapitulated those observed in undifferentiated iPS cells (Extended Data Fig. 3e). Cardiomyocyte differentiation[15] of mutant MELAS iPS cells was severely compromised due to massive cell death.

As expected, all iPS cell lines from homoplasmic 8993T>G fibroblasts carried mutant mtDNA (Extended Data Fig. 1e and Extended

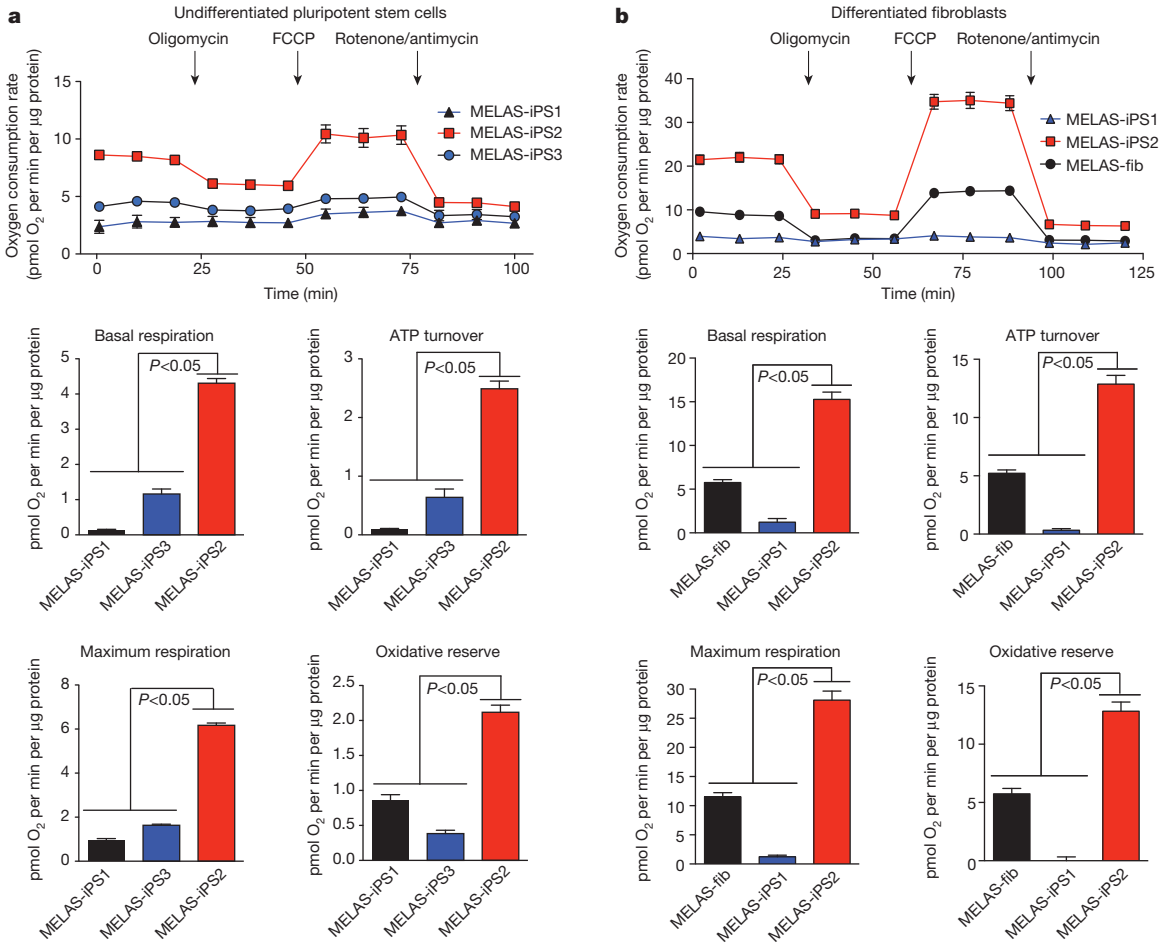


Figure 1 | Mitochondrial respiratory function in MELAS samples. **a**, Oxygen consumption rate (OCR) in undifferentiated MELAS-iPS1, MELAS-iPS2 and MELAS-iPS3 cells ($n = 9$ per cell line, biological replicates) in response to 0.5 µg ml^{-1} oligomycin, 1 µM fluorocarbonyl cyanide phenylhydrazone (FCCP), 0.5 µM rotenone and 1 µM antimycin. Wild-type MELAS-iPS2 displayed higher levels of oxygen consumption when compared to mutant MELAS-iPS1 and MELAS-iPS3. **b**, OCR in MELAS-iPS1 and MELAS-iPS2 derived fibroblasts and parental MELAS-fib ($n = 10$ per cell line, biological replicates). Error bars are mean ± s.e.m. and OCR data are representative of at least 2–3 independent experiments. Significance established with one-way analysis of variance (ANOVA) with Tukey's multiple comparison test.

Table 2 | Summary of 47 SNPs found in the mtDNA of Leigh-fib and Leigh-NT1 lines

Nucleotide position	Leigh-NT1	Leigh-fib	Locus	Effects	Nucleotide position	Leigh-NT1	Leigh-fib	Locus	Effects
152	C	T	Control region		10400	T	C	*MT-ND3*	Syn
248	A	Deletion	Control region	–	10410	C	T	*MT-TR*	–
489	C	T	Control region	–	10609	T	C	*MT-ND4L*	Non-syn
3010	A	G	*MT-RNR2*	–	10873	C	T	*MT-ND4*	Syn
3206	T	C	*MT-RNR2*	–	12406	G	A	*MT-ND5*	Non-syn
3970	C	T	*MT-ND1*	Syn	12418	Deletion	A	*MT-ND5*	Frame shift
4086	C	T	*MT-ND1*	Syn	12705	T	C	*MT-ND5*	Syn
4216	T	C	*MT-ND1*	Non-syn	12882	C	T	*MT-ND5*	Syn
4883	T	C	*MT-ND2*	Syn	13759	G	A	*MT-ND5*	Non-syn
5178	A	C	*MT-ND2*	Non-syn	13928	G	C	*MT-ND5*	Non-syn
6392	T	C	*MT-CO1*	Syn	14668	T	C	*MT-ND6*	Syn
6527	A	G	*MT-CO1*	Syn	14783	C	T	*MT-CYB*	Syn
6962	G	A	*MT-CO1*	Syn	14979	C	T	*MT-CYB*	Non-syn
7775	A	G	*MT-CO2*	Non-syn	15043	A	G	*MT-CYB*	Syn
8414	T	C	*MT-ATP8*	Non-syn	15301	A	G	*MT-CYB*	Syn
8473	C	T	*MT-ATP8*	Syn	15676	T	C	*MT-CYB*	Syn
8507	A	G	*MT-ATP8*	Non-syn	16148	C	T	Control region	–
8701	G	A	*MT-ATP6*	Non-syn	16162	A	G	Control region	–
8993	T	G	*MT-ATP6*	Non-syn	16172	T	C	Control region	–
9053	G	A	*MT-ATP6*	Non-syn	16223	T	C	Control region	–
9540	C	T	*MT-CO3*	Syn	16244	G	A	Control region	–
9548	G	A	*MT-CO3*	Syn	16304	T	C	Control region	–
10310	G	A	*MT-ND3*	Syn	16362	C	T	Control region	–
10398	G	A	*MT-ND3*	Non-syn					

Syn, synonymous; non-syn, non-synonymous.

Data Table 1). Therefore, we pursued mitochondrial replacement by SCNT with wild-type oocyte mitochondria. Following our reported protocol, two stable nuclear transfer–embryonic stem cell lines were established (Leigh-NT1 and Leigh-NT2)[16]. Genotyping confirmed that both lines contained predominantly oocyte wild-type mtDNA (Extended Data Fig. 1f) with limited low mutated mtDNA carryover (<1%) at passage 5 that became undetectable upon extended propagation (Extended Data Table 2).

Cytogenetic G-banding revealed that Leigh-iPS1 and Leigh-NT1 retained normal diploid karyotypes with no detectable numerical or structural chromosomal abnormalities (Extended Data Fig. 2c). However, Leigh-NT2 showed a XXXY tetraploid karyotype (Extended Data Fig. 2c). Fingerprinting by short tandem repeat analysis (STR) also revealed that Leigh-NT2 contained both oocyte and Leigh-fib alleles (Extended Data Table 3), consistent with failed enucleation. STR profiles for Leigh-NT1 and Leigh-iPS1 were identical to Leigh-fib (Extended Data Table 3). Both Leigh-iPS1 and Leigh-NT1 lines maintained typical PSC morphology, expressed pluripotency markers[7] and formed teratomas containing cells and tissues from all three germ layers (Extended Data Fig. 2d, e).

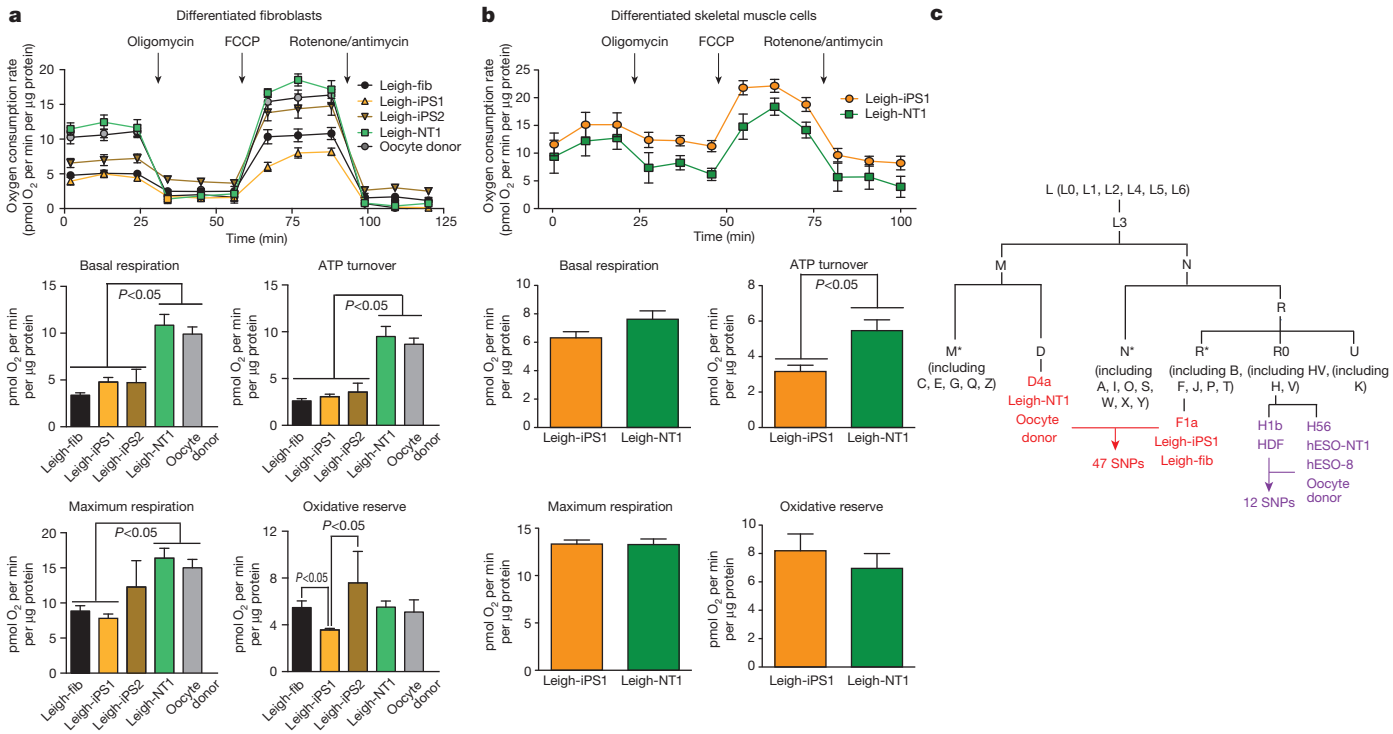


Figure 2 | Restoration of mitochondrial respiratory function in Leigh-NT1. **a**, OCR in Leigh-NT1, Leigh-iPS1 and Leigh-iPS2 derived fibroblasts, parental and oocyte donor fibroblasts ($n = 9$, 8, 10, 9 and 8 per cell line, respectively, biological replicates). **b**, OCR in Leigh-NT1and Leigh-iPS1 derived skeletal muscle cells ($n = 6$ biological replicates per cell line). **c**, mtDNA haplotype distances for oocyte and somatic cell donors based on mitochondrial phylogenetic tree from PhyloTree (http://phylotree.org/tree/main.htm)[27]. Asterisks indicate subgroups for mtDNA haplotypes. Error bars are mean ± s.e.m. and OCR data are representative of at least 2–3 independent experiments. Significance established using Kruskal–Wallis with Dunn's multiple comparison test or one-way ANOVA with Tukey's multiple comparison test (maximum respiration in **a**) and Student's *t*-test in **b**.

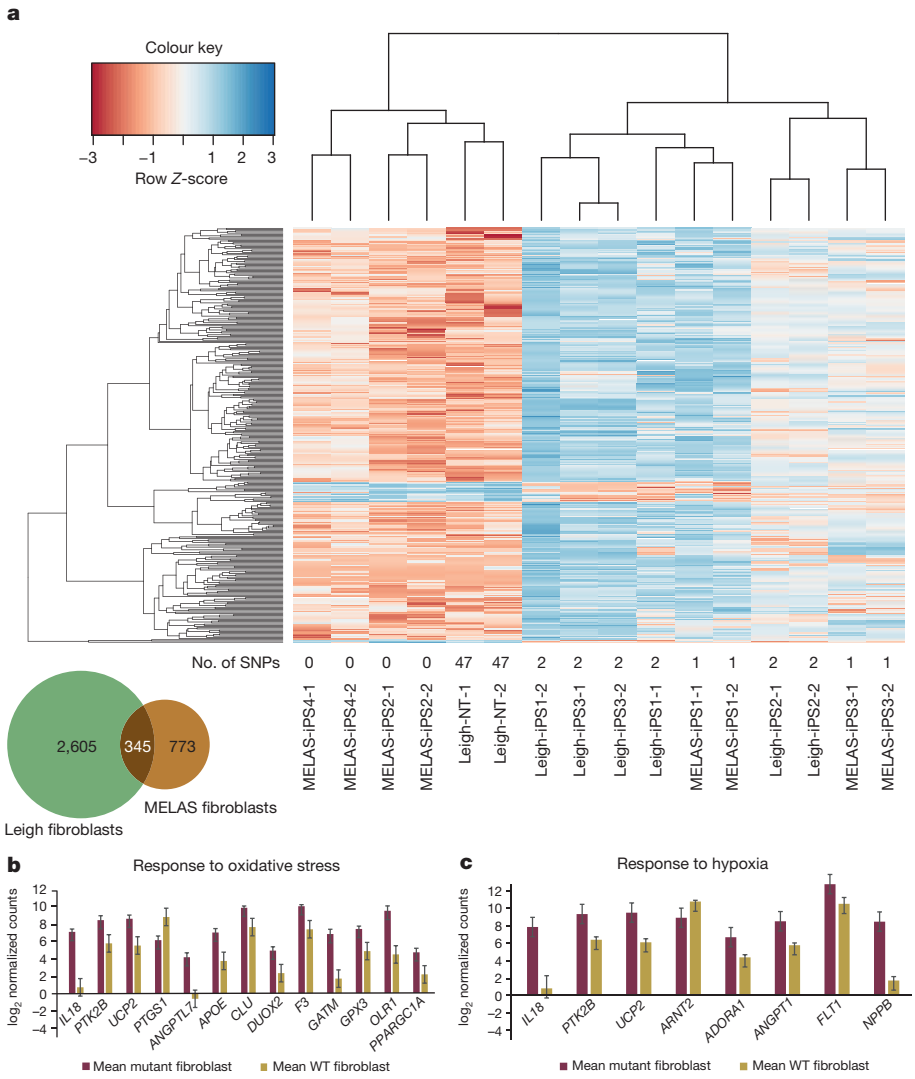


Figure 3 | Global gene expression analysis by RNA-seq. **a**, Heat map displaying 345 genes that are differentially expressed (adjusted *P* value <0.05) between wild-type PSCs (biological duplicates of 3 independent cell lines) versus mutant PSCs (biological duplicates of 5 independent cell lines). **b**, Functional enrichment analysis of genes displayed in the heat map that are known to be correlated with a response to oxidative stress. **c**, Functional enrichment analysis of genes displayed in the heat map that are known to be correlated with a response to hypoxia. Bar graphs are mean ± s.e.m., using all samples described in **a**.

Whole mtDNA sequencing confirmed the presence of homoplasmic 8993T>G mutation in Leigh-fib and Leigh-iPS1 and also identified a second homoplasmic 4216T>C mutation in the *MT-ND1* gene (Extended Data Fig. 1f and Supplementary Table 2). This non-synonymous mutation has been previously associated with Leber's hereditary optic neuropathy[15]. Leigh-NT1 mtDNA sequence differed from Leigh-fib at 47 nucleotide positions. In addition to the pathogenic 8993T>G and 4216T>C mutations, differences included 10 single nucleotide polymorphisms (SNPs) in the D-loop region, 2 in the 16S rRNA gene, 1 in the tRNA-R gene and 34 in protein genes (Table 2). We also detected two heteroplasmic variants in Leigh-NT1, four in Leigh-fib and three in Leigh-iPS1 (Supplementary Table 2). Clinical symptoms associated with these variants have not been reported.

We measured metabolic function in fibroblasts differentiated from Leigh-NT1, Leigh-iPS1 and Leigh-iPS2 and compared them to parental Leigh-fib and healthy skin fibroblasts from the oocyte donor. As expected, the homoplasmic 8993T>G mutation resulted in low mitochondrial oxidative capacity. In contrast, these respiration defects were absent in fibroblasts differentiated from Leigh-NT1 (Fig. 2a). Leigh-NT1 also displayed a metabolic profile and OCR/ECAR ratios similar to oocyte donor fibroblasts (Fig. 2a and Extended Data Fig. 4a). We observed varying levels of oxidative reserve for Leigh-iPS1 and Leigh-iPS2 compared to parental Leigh-fib (Fig. 2a), reflecting inherent variability within differentiated fibroblast populations. Both Leigh-NT1 and Leigh-iPS1 effectively generated skeletal muscle cells[17] (Extended Data Fig. 4b), with Leigh-iPS1 skeletal muscle cells displaying significantly lower ATP turnover ($P<0.05$) (Fig. 2b). Extensive cell death was observed in Leigh-iPS1 during directed cardiomyocyte differentiation (Extended Data Fig. 4c). These results demonstrate complete functional rescue of mitochondrial activity in Leigh-NT1 through restoration of the wild-type mtDNA.

Evolution of mtDNA has resulted in a series of neutral polymorphic variants within the human population often associated with regional migration and adaptation to climate[18]. The largest difference between distant human mtDNA haplotypes has been estimated at 95 SNPs[19]. In the present study, phylogenetic analysis assigned oocyte and Leigh-NT1 mtDNA to the D4a haplotype while the Leigh-iPS1 and Leigh-fib mtDNA haplotype was F1a (Supplementary Table 2). D4a is a descendant from the M while F1a comes from the N macro-haplo-group per the human mtDNA mutation tree (Fig. 2c). Safety evaluations of mitochondrial replacement therapy suggest possible harmful secondary outcomes reflecting nuclear–mitochondrial incompatibility[20,21]. Despite 'unmatched' donor mtDNA, Leigh-NT1 demonstrated lineage-specific differentiation and restoration of metabolic activity, implying normal nuclear–mitochondrial interaction. We further investigated a hESO-NT1 derived by SCNT from healthy fetal fibroblasts (human dermal fibroblast (HDF)) and IVF-derived hESO-8 carrying identical mtDNA[16]. hESO-NT1 mtDNA differed from HDF at 12 nucleotide positions (Fig. 2c and Supplementary Table 3). Metabolic profiles in NPCs and cardiomyocytes differentiated from hESO-NT1 and hESO-8 displayed similar metabolic profiles (Extended Data Fig. 5a–e). Next, we asked whether the 3243A>G

and 8993T>G mutations induced detectable changes in global gene expression and compared transcriptomes by RNA-seq for undifferentiated and differentiated PSCs. Undifferentiated PSCs containing wild-type or mutant mtDNA showed 154 differentially expressed transcripts (adjusted P value <0.05). This small number of differences is consistent with the predominantly glycolytic metabolism of pluripotent stem cells, which protects them from the deleterious effects of mtDNA mutations[22]. Global gene expression analysis of fibroblasts differentiated from isogenic MELAS lines identified 1,118 differentially expressed genes in mutant and wild-type cells (Extended Data Fig. 6a), whereas 2,950 genes were differentially expressed in fibroblasts differentiated from mutant Leigh-iPS1, Leigh-iPS2 and Leigh-iPS3 compared to the wild-type Leigh-NT1 cells (Extended Data Fig. 6b). Hierarchical clustering using a multiple bootstrap resampling algorithm showed that the Leigh-NT1 fibroblasts were similar to hESO-NT1, hESO-NT2, hESO-7 and hESO-8 fibroblasts (Extended Data Fig. 6c). These findings further support the notion that oocyte mtDNA in Leigh-NT1 interacts normally with nuclear DNA as long as the mtDNA sequence differences are neutral.

Next, we asked whether any of the differentially expressed genes were common to both 3243A>G and 8993T>G mutations, and found 345 genes that were shared, 96% of which were overexpressed in the mutant cells (Fig. 3a). Functional enrichment analysis identified genes associated with a response to hypoxia and oxidative stress[23,24] (Fig. 3b, c; P value <0.001). However, we did not observe an enrichment of genes associated with metabolism, stress response, epigenetic regulation, and cell signalling, which was reported in a recent MELAS study[13] (Extended Data Fig. 6d).

Finally, we addressed whether the 3243A>G and 8993T>G mutations specifically impact gene expression of mtDNA-encoded transcripts[25]. We found that transcripts expressed from mtDNA accounted for approximately 20% of the total cellular transcriptome, with similar expression levels across different mutations (Extended Data Fig. 7; adjusted P value >0.05).

We demonstrate complementary strategies for generating genetically and functionally corrected PSCs for patients with mtDNA disease. For the most common mtDNA syndromes caused by heteroplasmic mutations, generation of multiple iPS cell lines allows recovery of clones with exclusively wild-type mtDNA due to spontaneous segregation of heteroplasmic mtDNA. SCNT enables correction of homoplasmic mutations through replacement with donor mtDNA, and generation of PSCs with transcriptional and epigenetic profiles similar to embryo-derived embryonic stem cells[26]. Recovery of metabolic function despite haplotype differences between patient and donor mtDNA suggests that normal nuclear-to-mitochondrial interactions are highly conserved within species. Generation of genetically corrected PSCs from patients with mtDNA disease enables the transition from palliative care to therapeutic interventions based on regenerative medicine.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 September 2014; accepted 12 May 2015.
Published online 15 July 2015.

1. McBride, H. M., Neuspiel, M. & Wasiak, S. Mitochondria: more than just a powerhouse. *Curr. Biol.* **16,** 551–560 (2006).
2. Wallace, D. C. A mitochondrial bioenergetic etiology of disease. *J. Clin. Invest.* **123,** 1405–1412 (2013).
3. Taylor, R. W. & Turnbull, D. M. Mitochondrial DNA mutations in human disease. *Nature Rev. Genet.* **6,** 389–402 (2005).
4. Grossman, L. I. & Shoubridge, E. A. Mitochondrial genetics and human disease. *BioEssays* **18,** 983–991 (1996).
5. Goto, Y., Nonaka, I. & Horai, S. A mutation in the tRNA$^{Leu(UUR)}$ gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature* **348,** 651–653 (1990).
6. Tuppen, H. A., Blakely, E. L., Turnbull, D. M. & Taylor, R. W. Mitochondrial DNA mutations and human disease. *Biochim. Biophys. Acta* **1797,** 113–128 (2010).
7. Tachibana, M. *et al.* Towards germline gene therapy of inherited mitochondrial diseases. *Nature* **493,** 627–631 (2013).
8. Morten, K. J., Poulton, J. & Sykes, B. Multiple independent occurrence of the 3243 mutation in mitochondrial tRNA^{leuUUR} in patients with the MELAS phenotype. *Hum. Mol. Genet.* **4,** 1689–1691 (1995).
9. Chol, M. *et al.* The mitochondrial DNA G13513A MELAS mutation in the NADH dehydrogenase 5 gene is a frequent cause of Leigh-like syndrome with isolated complex I deficiency. *J. Med. Genet.* **40,** 188–191 (2003).
10. Fujikura, J. *et al.* Induced pluripotent stem cells generated from diabetic patients with mitochondrial DNA A3243G mutation. *Diabetologia* **55,** 1689–1698 (2012).
11. Hamalainen, R. H. *et al.* Tissue- and cell-type-specific manifestations of heteroplasmic mtDNA 3243A>G mutation in human induced pluripotent stem cell-derived disease model. *Proc. Natl Acad. Sci. USA* **110,** E3622–E3630 (2013).
12. Folmes, C. D. *et al.* Disease-causing mitochondrial heteroplasmy segregated within induced pluripotent stem cell clones derived from a patient with MELAS. *Stem Cells* **31,** 1298–1308 (2013).
13. Picard, M. *et al.* Progressive increase in mtDNA 3243A>G heteroplasmy causes abrupt transcriptional reprogramming. *Proc. Natl Acad. Sci. USA* **111,** E4033–E4042 (2014).
14. Liu, G. H. *et al.* Progressive degeneration of human neural stem cells caused by pathogenic LRRK2. *Nature* **491,** 603–607 (2012).
15. Lian, X. *et al.* Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. *Nature Protocols* **8,** 162–175 (2013).
16. Tachibana, M. *et al.* Human embryonic stem cells derived by somatic cell nuclear transfer. *Cell* **153,** 1228–1238 (2013).
17. Borchin, B., Chen, J. & Barberi, T. Derivation and FACS-mediated purification of PAX3$^+$/PAX7$^+$ skeletal muscle precursors from human pluripotent stem cells. *Stem Cell Rev.* **1,** 620–631 (2013).
18. Wallace, D. C., Brown, M. D. & Lott, M. T. Mitochondrial DNA variation in human evolution and disease. *Gene* **238,** 211–230 (1999).
19. Brandon, M. C. *et al.* MITOMAP: a human mitochondrial genome database–2004 update. *Nucleic Acids Res.* **33,** D611–D613 (2005).
20. Burgstaller, J. P. *et al.* MtDNA segregation in heteroplasmic tissues is common *in vivo* and modulated by haplotype differences and developmental stage. *Cell Rep.* **7,** 2031–2041 (2014).
21. Chinnery, P. F. *et al.* The challenges of mitochondrial replacement. *PLoS Genet.* **10,** e1004315 (2014).
22. Folmes, C. D., Dzeja, P. P., Nelson, T. J. & Terzic, A. Metabolic plasticity in stem cell homeostasis and differentiation. *Cell Stem Cell* **11,** 596–606 (2012).
23. Henchcliffe, C. & Beal, M. F. Mitochondrial biology and oxidative stress in Parkinson disease pathogenesis. *Nat. Clin. Pract. Neurol.* **4,** 600–609 (2008).
24. Boland, M. L., Chourasia, A. H. & Macleod, K. F. Mitochondrial dysfunction in cancer. *Front. Oncol.* **3,** 292 (2013).
25. Mercer, T. R. *et al.* The human mitochondrial transcriptome. *Cell* **146,** 645–658 (2011).
26. Ma, H. *et al.* Abnormalities in human pluripotent cells due to reprogramming mechanisms. *Nature* **511,** 177–183 (2014).
27. van Oven, M. & Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **30,** E386–E394 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors acknowledge the OHSU Embryonic Stem Cell Research Oversight Committee and the Institutional Review Board for providing oversight and guidance. We thank skin and oocyte donors and the Women's Health Research Unit staff at the Center for Women's Health, University Fertility Consultants and the Reproductive Endocrinology & Infertility Division in the Department of Obstetrics & Gynecology of Oregon Health & Science University for their support and procurement of gametes. We are grateful to M. Tachibana and A. Polat for help with derivation of PSCs and to M. Sparman for technical support. We are indebted to S. Gokhale for teratoma analysis and M. C. T. Penedo for microsatellite genotyping. We thank the staff at the Institute for Genomic Medicine Genomics Facility at UCSD for sequencing the RNA-seq libraries. Studies were supported by the Leducq Foundation, Mayo Clinic Center for Regenerative Medicine and OHSU and UCSD institutional funds. Work in the laboratory of J.C.I.B. was supported by the G. Harold and Leila Y. Mathers Charitable Foundation and the Leona M. and Harry B. Helmsley Charitable Trust (2012-PG-MED002).

Author Contributions H.M. and S.M. conceived the study and designed the experiments. S.M., P.A., H.M., R.A., E.K., Y.L., N.M.G. and R.T.-H. derived and cultured PSCs. J.P. derived MELAS fibroblasts. H.M., T.H., Y.L., C.V.D., A.K. and N.M. performed the DNA/RNA extractions and mtDNA ARMs-qPCR analyses. C.D.L.F. and A.T. performed Seahorse studies on differentiated fibroblasts and data analysis. X.W. and T.H. performed MiSeq studies. R.M., S.M.-C. and L.C.L. performed RNA-seq and bioinformatic analysis of the data. J.W., A.O., L.M. and J.C.I.B. performed NPC, skeletal muscle cell, and cardiomyocyte differentiations, Seahorse studies and data analysis. H.M., C.D.L.F., J.W., R.M., D.P.W., L.C.L., A.T., J.C.I.B. and S.M. analysed the data and wrote the paper.

Author Information Processed data sets can be downloaded from NCBI GEO under accession GSE61390 for RNA-seq. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.M. (mitalipo@ohsu.edu).

METHODS

The study protocols and informed consent for human subjects were approved by the OHSU Embryonic Stem Cell Research Oversight Committee and the Institutional Review Board. No statistical methods were used to predetermine sample size.

SCNT and iPS cell derivation and culture. Fibroblasts were acquired from Coriell Cell Repositories or donated by patients directly for our study. Fibroblasts were cultured in DMEM F12 medium supplemented with 10% fetal bovine serum (HyClone) and 50 μM uridine. SCNT procedures were performed as described previously[16]. Sendai virus-based reprogramming was carried out according to the manufacturer's protocol (CytoTune-iPS Reprogramming Kit, Life Technologies). Colonies with typical iPS cell morphology were isolated and manually propagated as described previously[26] in Knockout DMEM medium (Invitrogen) supplemented with 20% knockout serum replacement (Invitrogen), 0.1 mM nonessential amino acids (Invitrogen), 1 mM L-glutamine (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma), 1 × penicillin–streptomycin (Invitrogen) and 4 ng ml^{-1} basic fibroblast growth factor (Sigma). All cell cultures were free of mycoplasma contamination. Origin of all cell lines has been authenticated by STR and mtDNA genotyping.

Fibroblast differentiation. Differentiation of PSCs to fibroblasts was induced by culture in fibroblast medium (DMEM F12 with 10% FBS) for 2–3 weeks in absence of mouse embryonic fibroblast (mEF) feeder layers. Resulting differentiated cells were FACS sorted for TRA-1-60$^-$ (BD Biosciences), SSEA4$^-$ (Santa Cruz), CD56$^-$ (BD Biosciences) and CD13$^+$ (BD Biosciences) cells[28]. The CD13$^+$ cells were further expanded in the fibroblast medium.

NPC differentiation and culture. For NPC differentiation, a published protocol[14] was followed with minor modifications. PSCs were collected using collagenase IV (Life Technologies), washed twice with 1× DPBS without calcium and magnesium (Corning Cellgro), and cultured in Neural Induction Medium 1 (NIM-1: 50% Advanced DMEM/F12 (Invitrogen), 50% Neurobasal (Invitrogen), 1× B27 (Invitrogen), 1× N2 (Invitrogen), 2 mM GlutaMAX (Invitrogen) supplemented with 10 ng ml^{-1} hLIF (Peprotech), 4 μM CHIR99021 (Selleckchem), 3 μM SB431542 (Selleckchem), 2 μM dorsomorphin (Sigma), and 0.1 μM Compound E (EMD Chemicals Inc.)). Cells were cultured in NIM-1 medium for 2 days with daily medium change and then switched to Neural Induction Medium 2 (NIM-2: 50% Advanced DMEM/F12, 50% Neurobasal, 1× N2, 1× B27, 2 mM GlutaMAX and 10 ng ml^{-1} hLIF, 4 μM CHIR99021, 3 μM SB431542 and 0.1 μM Compound E). After 5 days culture in NIM-2 (daily medium change), cells were treated with 10 μM Y27632 (Selleckchem) for 1 h and 'dome'-shaped colonies were manually picked and treated with Accumax (Innovative Cell Technologies) for 10 min at 37 °C. Cells were then gently pipetted to obtain single cell suspension and replated onto Matrigel-coated 6-well plates at a density 3.5×10^4 per cm^2 in Neural Progenitor cell Maintenance Medium (NPMM: 50% Advanced DMEM/F12, 50% Neurobasal, 1× B27, 1× N2, 2 mM GlutaMAX, 10 ng ml^{-1} hLIF, 3 μM CHIR99021 and 2 μM SB431542) supplemented with 10 μM Y27632. NPCs were maintained on Matrigel-coated dishes in NPMM with daily medium change and passaged upon reaching 70% to 80% confluence using Accumax.

Skeletal muscle differentiation. Skeletal muscle differentiation was based on a previous report with minor modifications[17]. Briefly, PSCs plated on Matrigel-coated plates were grown to 40% confluence in mTeSR1 medium and then switched to Skeletal Muscle Induction Medium (SMIM, DMEM/F12, ITS, 3 μM CHIR99021). After 4 days culture in SMIM with daily medium change, cells were cultured in Skeletal Muscle Expansion Medium (SMEM: DMEM/F12, ITS and 20 ng ml^{-1} FGF2) for an additional 14 days with daily medium change. Cells were then cultured in Skeletal Muscle Differentiation Medium (SMDM: DMEM/F12 and ITS only) for an additional 18 days.

Cardiomyocyte differentiation. Cardiomyocyte differentiation was performed with adaptation based on the inhibition of GSK3 and Wnt pathways[15]. Briefly, PSCs were collected after Accutase (Life Technologies) treatment and cultured on Matrigel-coated plates in RPMI supplemented with B27 without insulin (Invitrogen) to 80–90% confluency. Cells were then incubated with 12 μM CHIR99021 (Selleckchem) for 16 h. At day 3, cells were incubated with 5 μM IWP2 (Tocris) for 48 h. At day 7, medium was replaced to RPMI supplemented with complete B27. Medium was replaced every 3 days. Contracting cardiomyocytes were observed on day 12 of differentiation.

Immunocytochemistry. Cultured cells were fixed with 4% paraformaldehyde for 15 min at room temperature and then permeabilized with 0.2% Triton X-100 in PBS for 10 min. Cells were washed 3× with PBST (PBS 1×, 0.02% Tween-20) and blocked with 10% goat or donkey serum (Sigma) for 1 h at room temperature. Cells were then incubated with primary antibodies diluted in PBST overnight at 4 °C, washed 3× with PBST and incubated with secondary antibodies (1:500, Molecular Probes) for 1 h at room temperature. Cells were washed 3× and mounted in Prolong Gold Antifade Mountant (Life Technologies). Image acquisition was performed on a Zeiss LSM 780 confocal microscope. Primary antibodies were: PAX6 (1:100, Convance), NESTIN (1:200, Millipore), MF20 (1:100, DSHB), OCT4 (1:100, Santa Cruz) and NANOG (1:40, R&D Systems).

Teratoma assay. Approximately 3–5 million undifferentiated PSCs were injected into the hindleg muscle of 8-week-old, SCID, beige male mice (Charles River) using an 18-gauge needle. Six to seven weeks after injection, mice were euthanized and tumours were dissected, sectioned and histologically characterized for the presence of representative tissues as described previously[7]. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.

mtDNA heteroplasmy analysis by ARMs-qPCR. The amplification refractory mutation system quantitative PCR assay (ARMs-qPCR) was used to measure mtDNA carryover in Leigh-NT1 and NT2 as previously described[16]. Primers and TaqMan MGB probes were designed to detect the 8993T>G mutation site. The nondiscriminative (ND) and discriminative (D) assays were mixed and measured with Rotor-Gene Multiplex PCR Kit (Qiagen). All reactions were run in duplicate with two different amounts of input DNA: 1–4 ng and 1:8 dilutions. The SDS software generated a standard curve using four eightfold dilutions plus a final fourfold dilution. The percentage of mtDNA carryover in relation to the total mtDNA content was calculated by the equation: heteroplasmy = 100×(quantity D/quantity ND). ARMs-qPCR was also applied to detect 8993T>G, 3243A>G and 13513G>A heteroplasmy levels in fibroblasts and iPS cells using primers and TaqMan MGB probes specifically targeting to mutation sites.

Whole mtDNA sequencing analysis by MiSeq. Single PCR amplification of entire human mtDNA was performed with primers mtDNA-F-2120, (GGAC ACTAGGAAAAAACCTTGTAGAGAGAG) and mtDNA-R-2119 (AAAGAGC TGTTCCTCTTTGGACTAACA) under the following conditions: 94 °C for 1 min followed 98 °C for 10 s and 68 °C for 16 min ×30 cycles and then 72 °C for 10 min. PCR amplifications were performed using TAKARA LA Taq polymerase (Takara Biotechnology) and the concentrations of PCR products were measured using a Qubit 2.0 Fluorometer. The Nextera XT DNA sample preparation kit (Illumina) was used to prepare the libraries. Sequencing was performed on an Illumina MiSeq instrument and the data were analysed using NextGENe software. Briefly, sequence reads ranging from 100 to 200 bp were quality filtered and processed using BLAT algorithm. Sequence error correction feature (condensation) was performed to reduce false-positive variants and produce sample consensus sequence and variant calls. Alignment without sequence condensation was used to calculate percentage of mitochondrial genome with depth of coverage of 1,000. Starting from quality FASTQ reads, the reads were quality filtered and converted to FASTA format. Filtered reads were then aligned to the revised Cambridge Reference Sequence (rCRS) of the human mtDNA (NC_012920.1) followed by variant calling. Variant heteroplasmy was calculated by NextGENe software as follows: Base heteroplasmy (mutant allele frequency %) = mutant allele (forward + reverse)/total coverage of all alleles C, G, T, A (forward + reverse) × 100. The clinical significance of the variants was then analysed with MitoMaster (http://www.mitomap.org/MITOMASTER/WebHome).

Live cell oxygen consumption. XF24 or XF96 extracellular flux analysers (Seahorse Biosciences) were used to measure oxygen consumption rates (OCR) as described[12]. In brief, stem-cell-derived fibroblasts were seeded at a density of 50,000 cells per well of a XF24 cell culture microplate and incubated for 24 h to ensure attachment. Before assay, cells were equilibrated for 1 h in unbuffered XF assay medium supplemented with 25 mM glucose, 1 mM sodium pyruvate, 2 mM glutamax, 1× nonessential amino acids and 1% FBS in a non-CO_2 incubator. Mitochondrial processes were examined through sequential injections of oligomycin (0.5 μg ml^{-1}), carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone (FCCP, 1 μM) and rotenone (0.5 μM)/antimycin A (1 μM). Indices of mitochondrial function were calculated as basal respiration rate (baseline OCR − rotenone/antimycin A OCR), ATP dependent (basal respiration rate − oligomycin OCAR), maximal respiration rate (FCCP OCR − rotenone/antimycin A OCR) and oxidative reserve (maximal respiration rate − basal respiration rate). For other cell types, an XF96 extracellular flux analyser was used with 20,000 cells seeded to each well of a XF96 cell culture microplate. After a 24-h attachment period, mitochondrial processes were examined using the same protocol as above. Each plotted value was normalized to total protein quantified using a Bradford protein assay (Bio-rad). Results were presented as mean ± s.e.m. One-way ANOVA was used for three group comparisons and Student's *t*-test was used for two group comparisons. A P value less than 0.05 was considered significant.

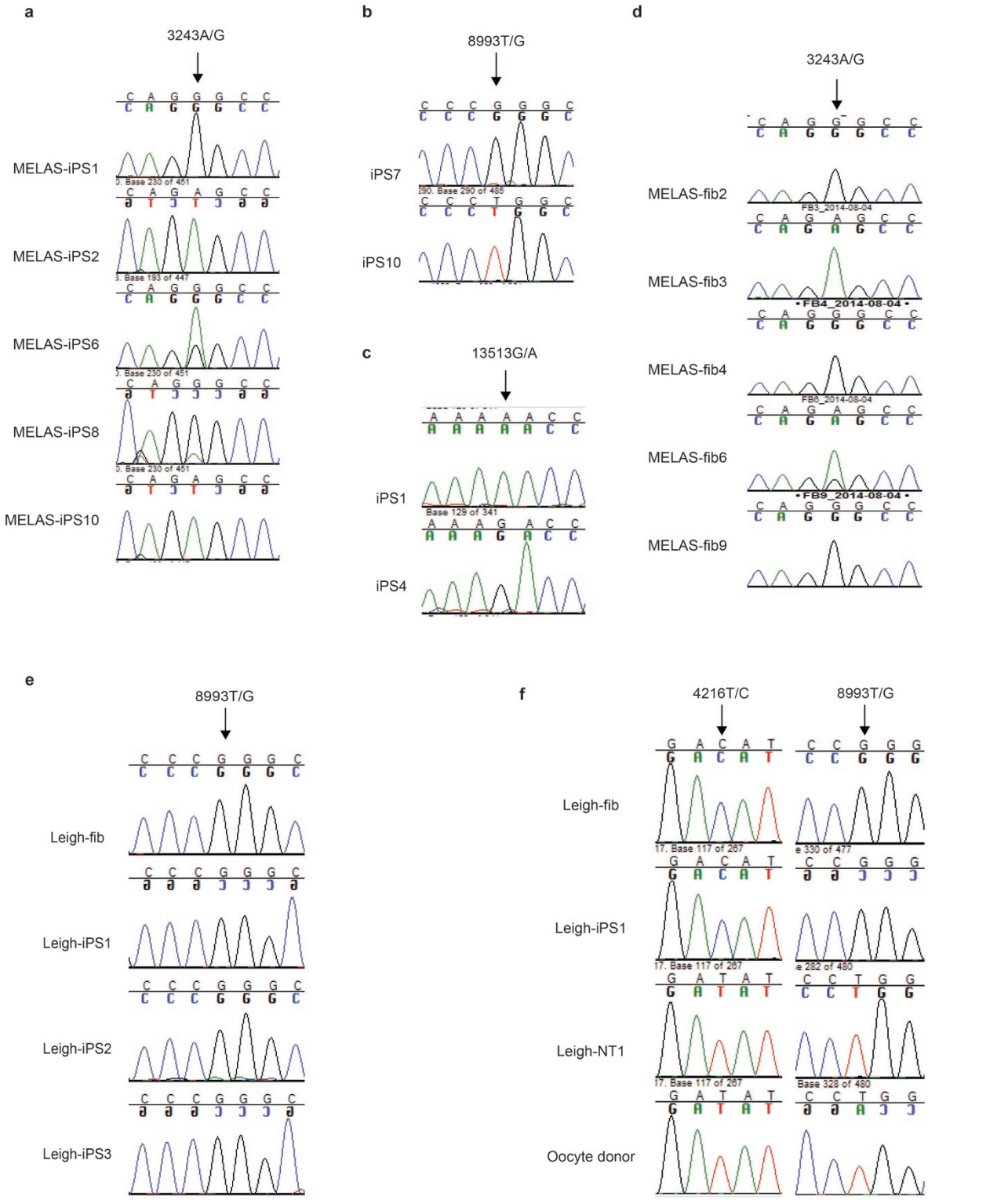
Flow cytometric analysis. The efficiency of differentiation protocols was assessed by FACS. For cardiomyocyte differentiation, 0.25 million cells were fixed in the presence of 1% (vol/vol) paraformaldehyde at room temperature for 20 min. Fixed cells were then incubated in 90% (vol/vol) cold methanol for 15 min at 4 °C, rinsed

two times and incubated overnight at 4 °C with a primary antibodies against GATA4 (Santa Cruz) and cTnT (Pierce). After staining, cells were rinsed two times and incubated in the presence of 1:1,000 secondary antibodies (donkey Alexa 488 and 567; Molecular Probes) for 30 min. After staining, cells were washed two times and re-suspended for analysis.

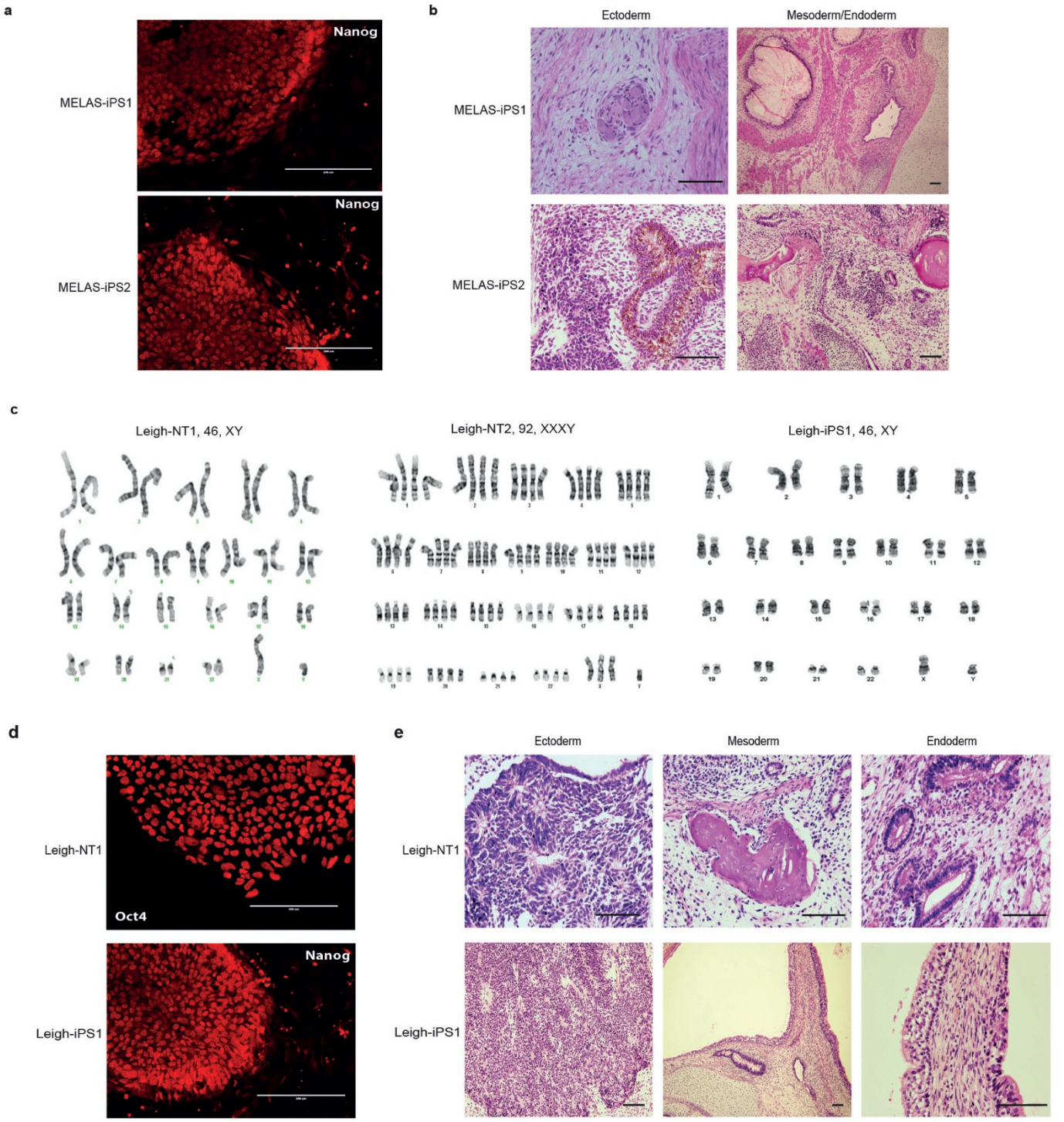
Real time RT–PCR. RNA was isolated using RNeasy kit (Qiagen) as per manufacturer's instructions. cDNA synthesis was performed using the iScriptTM cDNA synthesis kit for RT–PCR (BioRad). Real-time PCR was performed using the SYBR Green Supermix (BioRad). The levels of expression of respective genes were normalized to corresponding 18S values and are shown as fold change relative to the value of the control sample. All reactions were done in triplicate.

RNA-seq library construction and data analysis. RNA was isolated with Micro-to-Midi Total RNA Purification System (Life Technologies), quality evaluated (RNA6000 Nano Kit and BioAnalyzer 2100, Agilent) made into sequencing libraries, sequenced and mapped as previously described[26]. Libraries were constructed using 500 ng input RNA per sample. Approximately 27 million reads were generated per sample, and 73% of these reads were uniquely mapped. Counts for each gene were quantified using the python script rpkmforgenes and annotated using Ensembl GRCh37. Genes without at least one sample with at least five reads were removed from the analysis. The count data was normalized and differential expression was performed using the R (v.3.1.1) package DESeq2 (v.1.4.5). Briefly, DESeq2 uses negative binomial generalized linear models and shrinkage estimation for dispersions and fold changes to improve stability and interpretability of the estimates[29]. It reports a *P* value and an adjusted *P* value using the Benjamini–Hochberg procedure. Genes with an adjusted *P* value less than 0.05 were considered differentially expressed unless otherwise noted. Heat maps were constructed using the R (v.3.1.1) package gplots (v.2.14.2). Each variable was standardized by subtraction of its mean value and division by its standard deviation across all samples. All functional enrichment analyses were generated using the Genomic Regions Enrichment of Annotations Tool (v. 2.0.2)[30] with default settings. Hierarchical clustering was performed with the R package pvclust, with Euclidian distance and average linkage with 10,000 bootstraps.

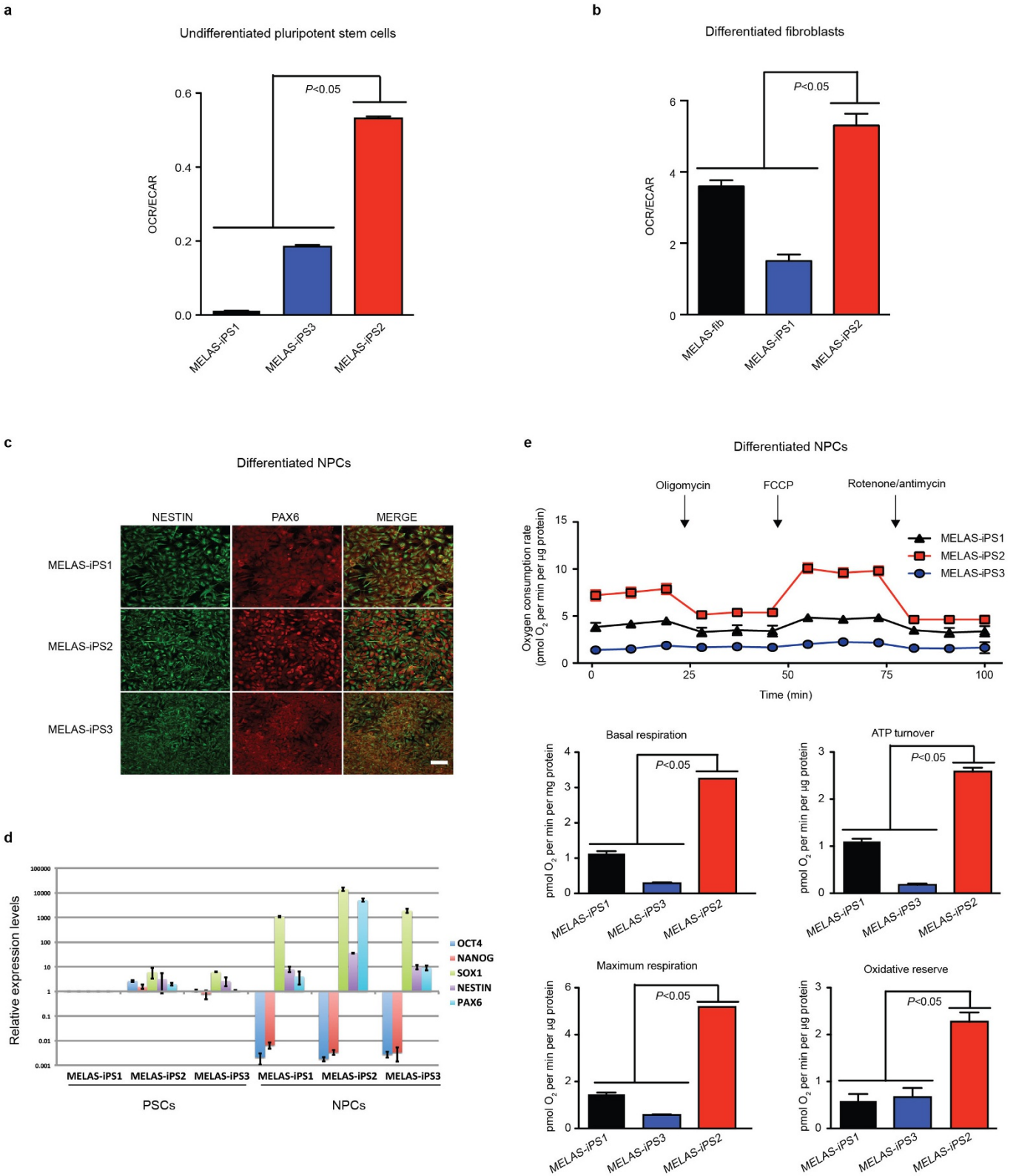
28. Paull, D. *et al.* Nuclear genome transfer in human oocytes eliminates mitochondrial DNA variants. *Nature* **493,** 632–637 (2013).
29. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** (2014).
30. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnol.* **28,** 495–501 (2010).



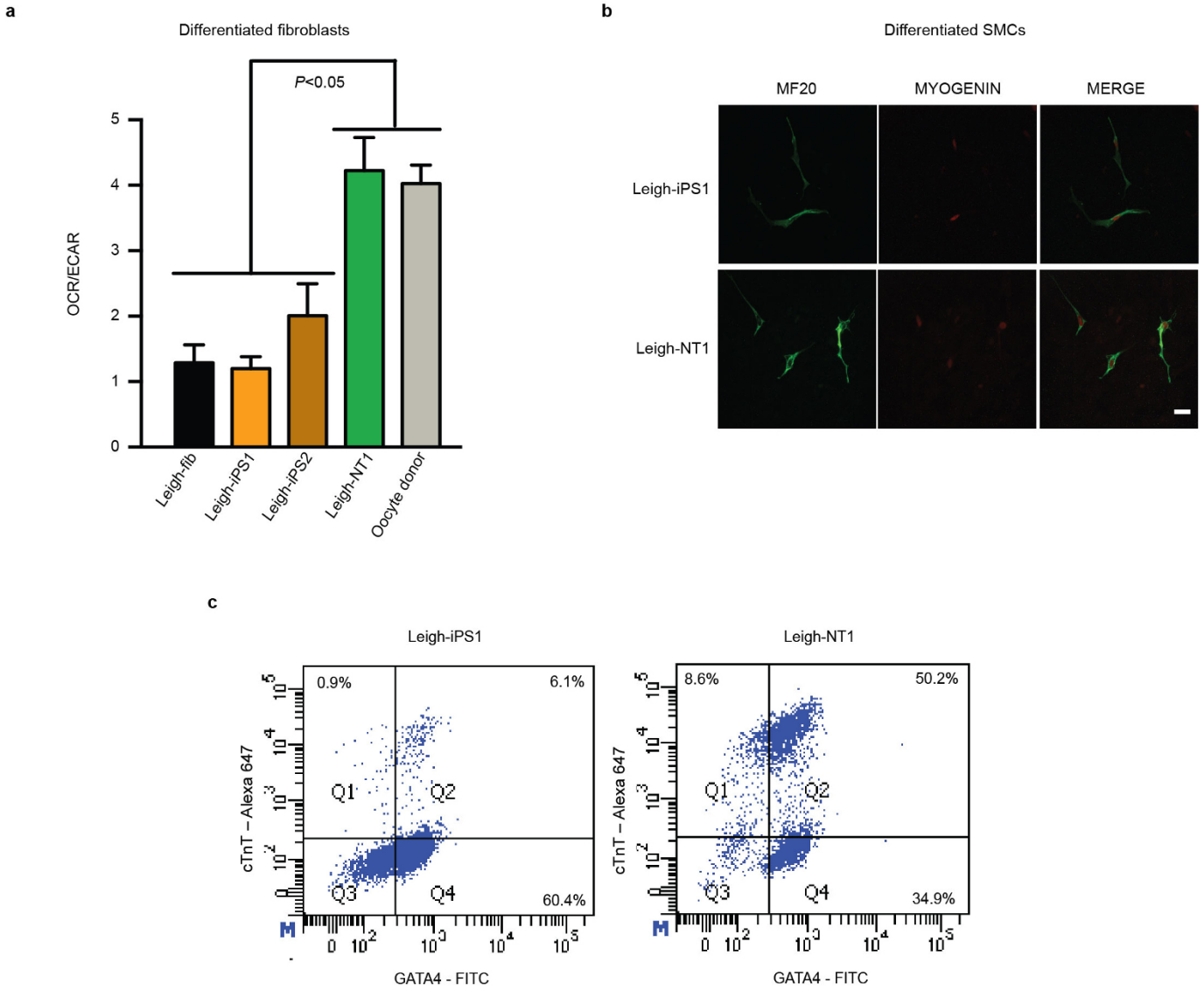
Extended Data Figure 1 | mtDNA genotyping by Sanger sequencing. **a**, Chromatographs showing mtDNA genotyping at 3243 position (arrow) in representative MELAS iPS cells. **b**, Chromatographs showing mtDNA genotyping at 8993 position (arrow). **c**, mtDNA at 13513 position (arrow) in representative iPS cells derived from Leigh syndrome patients. **d**, Chromatographs showing either wild-type A or mutant G allele at position 3243 in representative MELAS fibroblast clones. **e**, mtDNA genotyping demonstrated that all Leigh-iPS cell lines and Leigh-fib contain a G mutation allele at mtDNA position 8993. **f**, mtDNA genotyping demonstrated that Leigh-fib and Leigh-iPS1 cell lines contained a C mutant allele at position 4216 and a G mutant allele at position 8993, while Leigh-NT1 line carried oocyte mtDNA with a wild-type T allele at both positions.



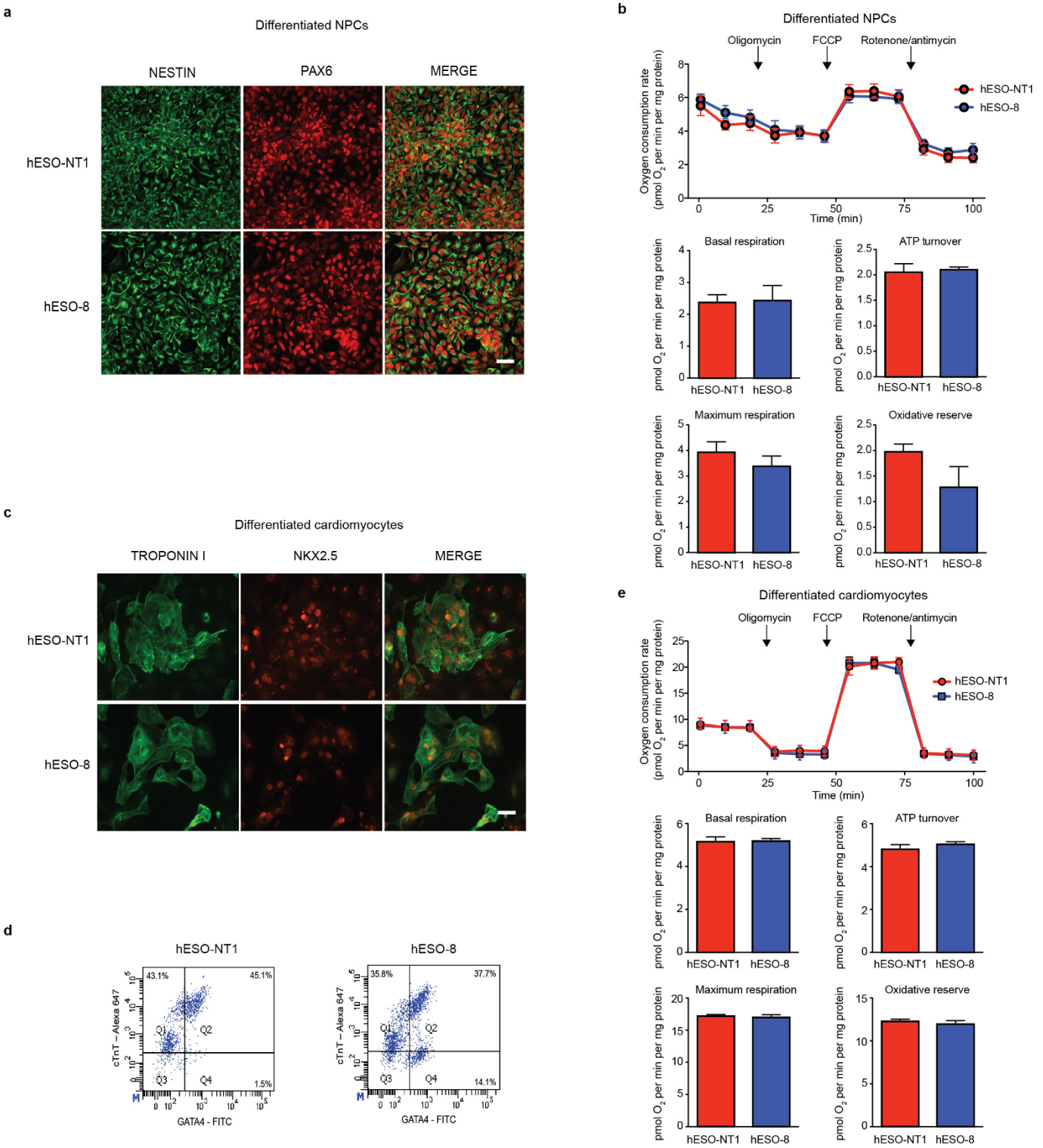
Extended Data Figure 2 | Cytogenetic, pluripotency and teratoma analyses. **a**, MELAS-iPS1 and MELAS-iPS2 expressing NANOG detected by immunocytochemistry. Scale bars, 200 μm. **b**, Histological analyses of teratoma tumours produced after injections of MELAS-iPS1 and MELAS-iPS2 cells into SCID mice. Scale bars, 200 μm. **c**, Cytogenetic G-banding analysis confirmed that Leigh-NT1 and Leigh-iPS1 exhibited normal 46XY karyotypes and Leigh-NT2 exhibited a XXXY tetraploid karyotype. **d**, Leigh-NT1 and Leigh-iPS1cells expressed OCT4 and NANOG. Scale bars, 200 μm. **e**, Histological analyses of teratoma tumours produced after injections of Leigh-NT1 and Leigh-iPS1 cells into SCID mice. Scale bars, 200 μm. Haematoxylin and eosin staining of teratoma sections identify derivatives of ectoderm, mesoderm and endoderm.



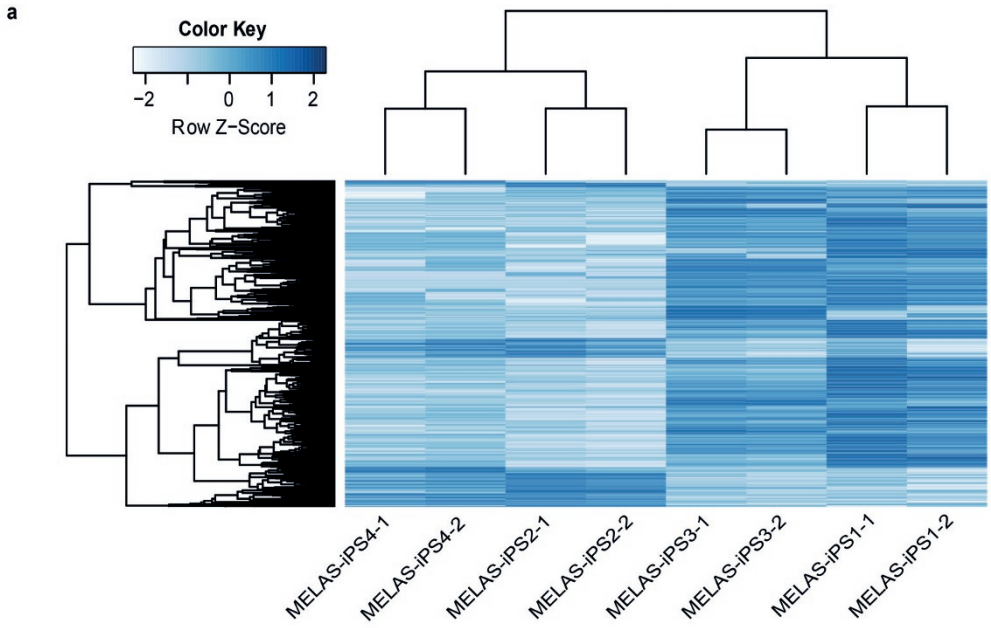
Extended Data Figure 3 | Metabolic function in differentiated cells from MELAS iPS cells. **a**, OCR/ECAR ratio in MELAS-iPS cells. Mutant MELAS-iPS1 and MELAS-iPS3 displayed significantly decreased OCR/ECAR ratios compared to wild-type MELAS-iPS2 ($P < 0.05$), indicating a greater reliance on glycolysis ($n = 9$ per cell line, biological replicates). **b**, OCR/ECAR ratio in MELAS-iPS cell derived fibroblasts ($n = 10$ per cell line, biological replicates). **c**, Immunofluorescence analysis for neural progenitor markers in MELAS-iPS derived NPCs. Scale bar, 100 µm. **d**, Quantitative analysis of PSC (OCT4 and NANOG) or NPC (SOX1, NESTIN and PAX6) marker expression in MELAS-iPS cells and NPCs ($n = 3$ per cell line, biological replicates). **e**, OCR of MELAS-iPS cell derived NPCs ($n = 6$ per cell line, biological replicates). Error bars are mean ± s.e.m. Significance established with one-way ANOVA with Tukey's multiple comparison test.

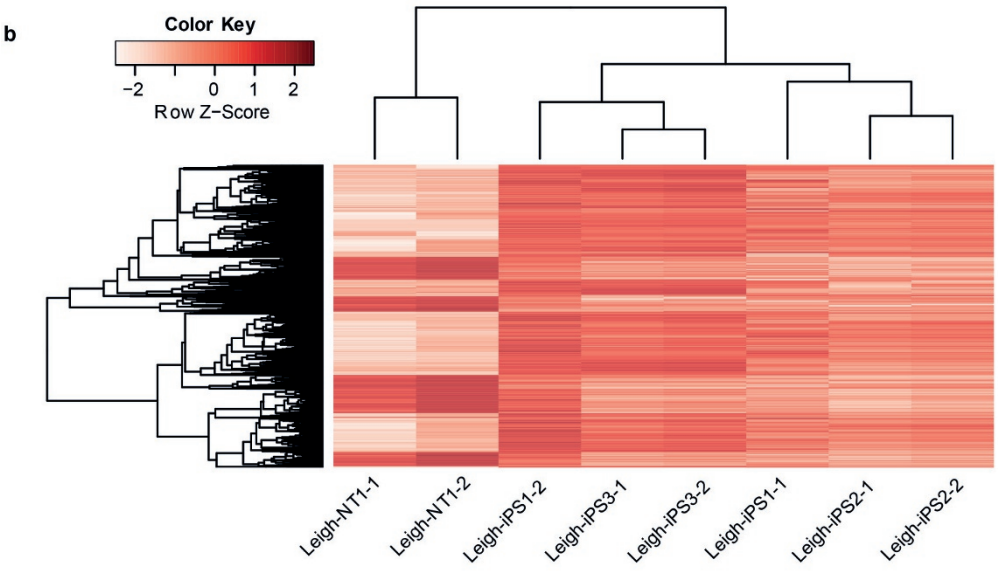


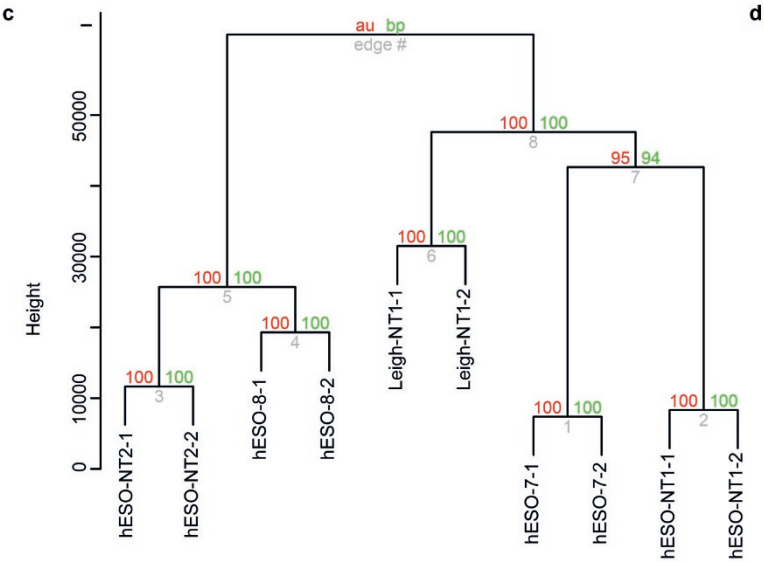
Extended Data Figure 4 | Metabolic function in differentiated cells from Leigh syndrome PSCs. **a**, OCR/ECAR ratio in Leigh-iPS1, Leigh-iPS2 and Leigh-NT1 derived fibroblasts, parental and oocyte donor fibroblasts ($n = 9, 8, 10, 9$ and 8 per cell line, respectively, biological replicates). **b**, Immunofluorescence analysis of Leigh-iPS1- and Leigh-NT1-derived skeletal muscle cells labelled with MF20 and myogenin antibodies. Scale bar, 100 μm. **c**, Cardiomyocyte differentiation efficiency in Leigh-iPS1 and Leigh-NT1 evaluated by FACS for CTnT-Alexa 647 and GATA4-FITC antibodies ($n = 3$ per cell line, biological replicates). Error bars are mean ± s.e.m. Significance established with one-way ANOVA with Tukey's multiple comparison test.

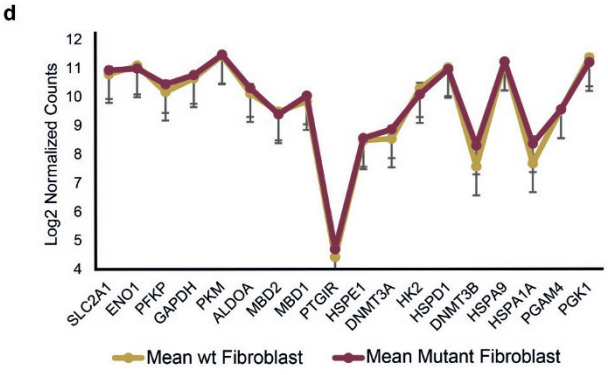


Extended Data Figure 5 | Metabolic function in hESO-NT1 and hESO-8. **a**, Immunofluorescence analysis of hESO-NT1 and hESO-8 derived NPCs with nestin and PAX6 antibodies. Scale bar, 100 µm. **b**, Metabolic profiles of NPCs differentiated from hESO-NT1 and hESO-8 ($n = 6$ per cell line, biological replicates). **c**, Immunofluorescence analysis of hESO-NT1 and hESO-8 derived cardiomyocytes with troponin I and NKX2.5 antibodies. Scale bar, 100 µm. **d**, Efficiency of cardiomyocyte differentiation in hESO-NT1 and hESO-8 evaluated by FACS analysis for CTnT-Alexa 647 and GATA4-FITC antibodies ($n = 3$ per cell line, biological replicates). **e**, OCR of hESO-NT1 and hESO-8 derived cardiomyocytes ($n = 6$ per cell line, biological replicates). Error bars are mean ± s.e.m. Significance established with Student's *t*-test.

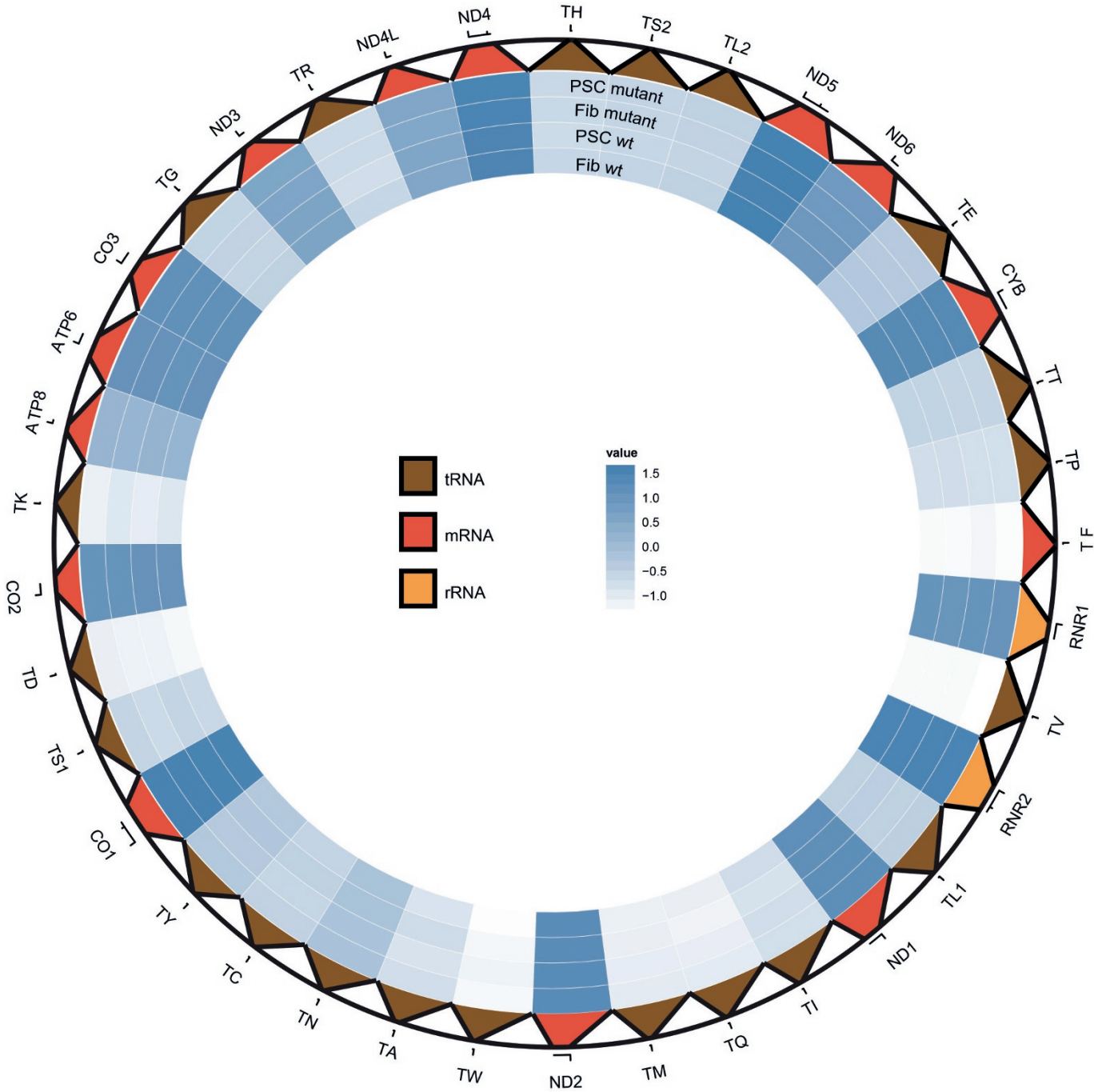








Extended Data Figure 6 | RNA-seq analyses of fibroblasts differentiated from MELAS and Leigh syndrome PSCs carrying wild-type and mutant mtDNA. **a**, Heat map showing all differentially expressed 1,118 genes (adjusted P value < 0.05) between fibroblasts differentiated from mutant MELAS iPS cells ($n = 4$ from biological duplicates of MELAS-iPS2 and MELAS-iPS4) and wild-type MELAS iPS cells ($n = 4$ from biological duplicates of MELAS-iPS1 and MELAS-iPS3). **b**, Heat map demonstrating differentially expressed 2,950 genes (adjusted P value < 0.05) between fibroblasts derived from wild-type Leigh-NT1 (biological duplicates) and mutant Leigh iPS cells ($n = 6$ from biological duplicates of Leigh-iPS1, Leigh-iPS2 and Leigh-iPS3). **c**, Hierarchical clustering using Euclidean distance and average linkage using pvclust, which employs a multiple bootstrap resampling algorithm to calculate the approximately unbiased (AU, red) and bootstrap probability (BU, green) values for cluster distinctions. Hierarchical clustering showed that the Leigh-NT1 fibroblasts were similar to hESO-NT1, hESO-NT2, hESO-7 and hESO-8 fibroblasts. **d**, Mean $\log_2$ normalized counts $\pm$ s.e.m. for genes previously reported to be differentially expressed in MELAS cytoplasmic hybrid clones and involved in metabolic and stress response, signalling pathways and epigenetic modifying processes (wild type fibroblast; $n = 14$ from biological duplicates of 7 independent cell lines; mutant fibroblast $n = 10$ from biological duplicates of 5 independent cell lines).



Extended Data Figure 7 | RNA-seq analysis of the mitochondrial transcriptome. Circular heat map displaying average expression levels for all mitochondrial genes grouped by sample differentiation status and presence or absence of a mutation in the mitochondrial genome (Fib mutant (including primary fibroblasts and PSC derived fibroblasts) with mutant mtDNA $n = 14$, biological duplicates of 7 independent cell lines; Fib wild type (including primary fibroblasts and PSC derived fibroblasts) with wild-type mtDNA $n = 14$, biological duplicates of 7 independent cell lines; PSC mutant (undifferentiated IVF-ESC, NT-ESC and iPS cells) with mutant mtDNA $n = 3$; PSC wild type (undifferentiated IVF-ESC, NT-ESC and iPS cells) with wild-type mtDNA $n = 12$). The expression of mtDNA-encoded genes was similar irrespective of 3243A>G or 8993T>G mutations (adjusted P value >0.05).

Extended Data Table 1 | Mutation loads in Leigh syndrome iPS cells with homoplasmic mutations

PSC lines	8993 T>G mutant G %	4216 T>C mutant C %
Leigh-iPS1	100	100
Leigh-iPS2	100	100
Leigh-iPS3	100	100
Leigh-iPS4	100	100
Leigh-iPS5	100	100
Leigh-NT1	0	0

Extended Data Table 2 | Quantitative mutant mtDNA carryover analysis in Leigh-NT1

Cell line	% mutant mtDNA P5 (±SD)	% mutant mtDNA P15 (±SD)	% mutant mtDNA P20 (±SD)	% mutant mtDNA P30 (±SD)	% mutant mtDNA P40 (±SD)
Leigh-NT1	0.14±0.06	undetectable	undetectable	undetectable	undetectable

Extended Data Table 3 | Short tandem repeat analysis of oocyte donors, Leigh-NT2 and iPS cells from the Leigh syndrome patient

Sample	Leigh group					
	Leigh-fib	Leigh-iPS1	Leigh-NT1	Oocyte donor1	Leigh-NT2	Oocyte donor2
Sex AME	M XY	M XY	M XY	F XX	F+M XXXY	F XX
D1S548	152/172	152/172	152/172	152/168	152/168/172/176	168/176
D2S1333	301/305	301/305	301/305	285/301	285/301/305	285/301
D3S1768	188/200	188/200	188/200	188/196	188/192/200	192/192
D4S2365	288/296	288/296	288/296	296/300	280/288/296	280/296
D4S413	123/123	123/123	123/123	123/123	123/155/157	155/157
D5S1457	119/123	119/123	119/123	123/123	119/123	119/123
D6S501	168/172	168/172	168/172	168/172	168/172/176	172/176
D7S513	179/197	179/197	179/197	187/201	179/187/189/197	187/189
D9S921	183/183	183/183	183/183	183/189	183/183	183/183
D10S1412	156/162	156/162	156/162	159/162	156/162/165	162/165
D11S2002	254/254	254/254	254/254	254/254	254/254	254/254
D11S925	303/305	303/305	303/305	303/305	282/295/303/305	282/295
D12S364	266/270	266/270	266/270	270/272	266/270/278	270/278
D12S67	260/268	260/268	260/268	252/260	252/256/260/268	252/256
D13S765	192/196	192/196	192/196	188/200	188/192/196	188/192
D16S403	137/141	137/141	137/141	145/145	135/137/141	135/141
D17S1300	257/273	257/273	257/273	257/269	257/265/269/273	265/269
D18S537	196/200	196/200	196/200	196/208	196/200/204	200/204
D18S72	305/305	305/305	305/305	301/305	305/305	305/305
DXS2506*	286*	286*	286*	282/286	282/286	282/282
D6S291	199/201	199/201	199/201	199/199	199/201	199/201
D6S276	239/239	239/239	239/239	227/235	227/239/251	227/251

* Male samples only show one mark of DXS2506 on the X chromosome.