Stem Cell Reports, Volume *16*

Supplemental Information

iPSC culture expansion selects against putatively actionable mutations in the mitochondrial genome

Maike Kosanke, Colin Davenport, Monika Szepes, Lutz Wiehlmann, Tim Kohrn, Marie Dorda, Jonas Gruber, Kaja Menge, Maike Sievert, Anna Melchert, Ina Gruh, Gudrun Göhring, and Ulrich Martin

Supplemental Information:

Supplemental Figures and Tables:

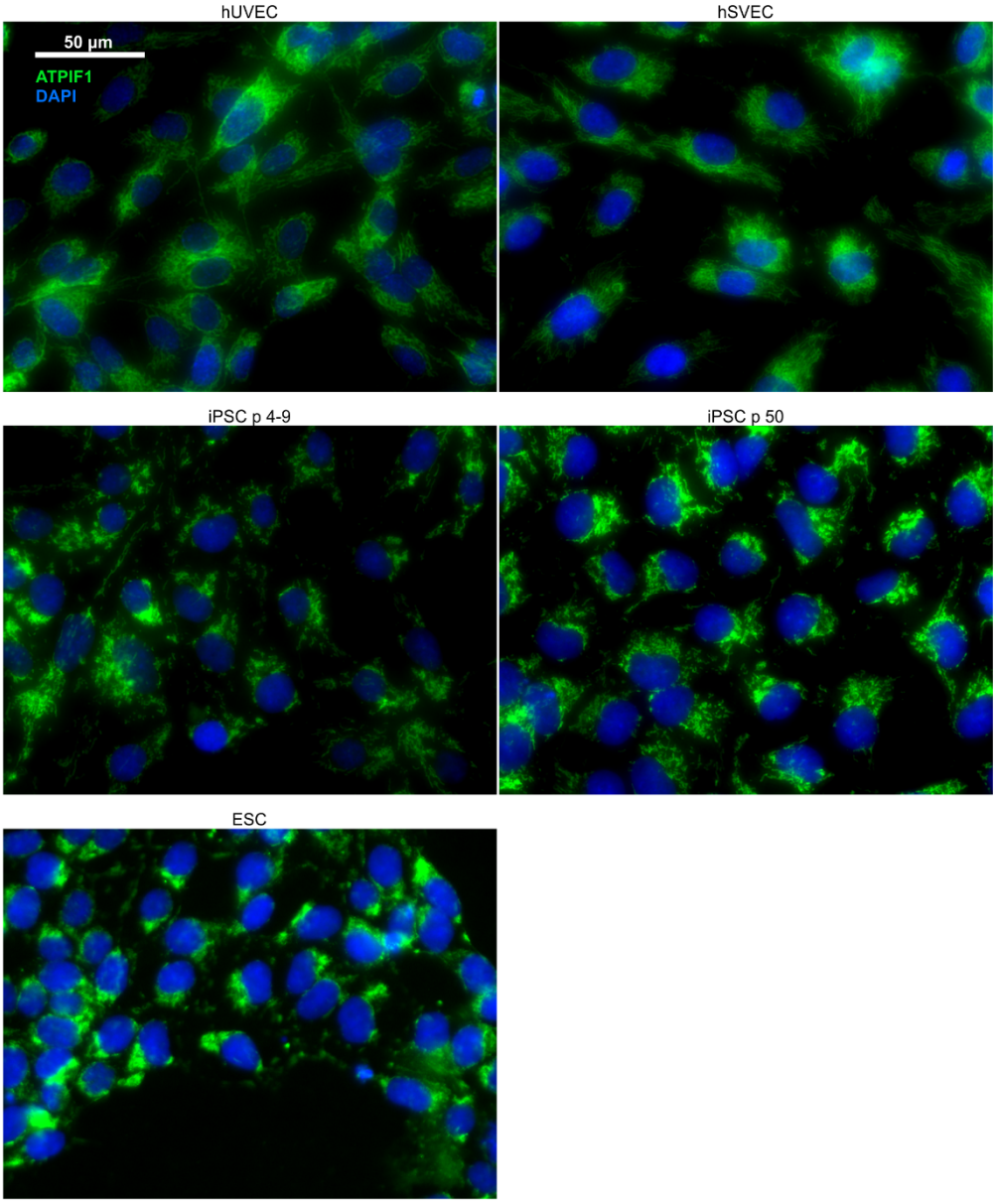

Figure S1: Mitochondrial network in ECs, iPSCs, and ESCs; Mentioned in first and second Result section. Adherent ECs (umbilical vein (hUVEC) and saphenous vein (hSVEC)), iPSCs of early and late passage, and ESCs were seeded on gelatine or Geltrex coated glass cover slides, cultured for 2-3 days in EMG2 or E8 medium, fixed with PFA, stained with anti-ATPIF1 (ATP synthase subunit IF1) antibody and DAPI nuclear staining, and analyzed by fluorescence microscopy. Pictures of different representative donors and clones are shown. Mitochondria are displayed in green and nuclei staining in blue. Scale bar 50µm.

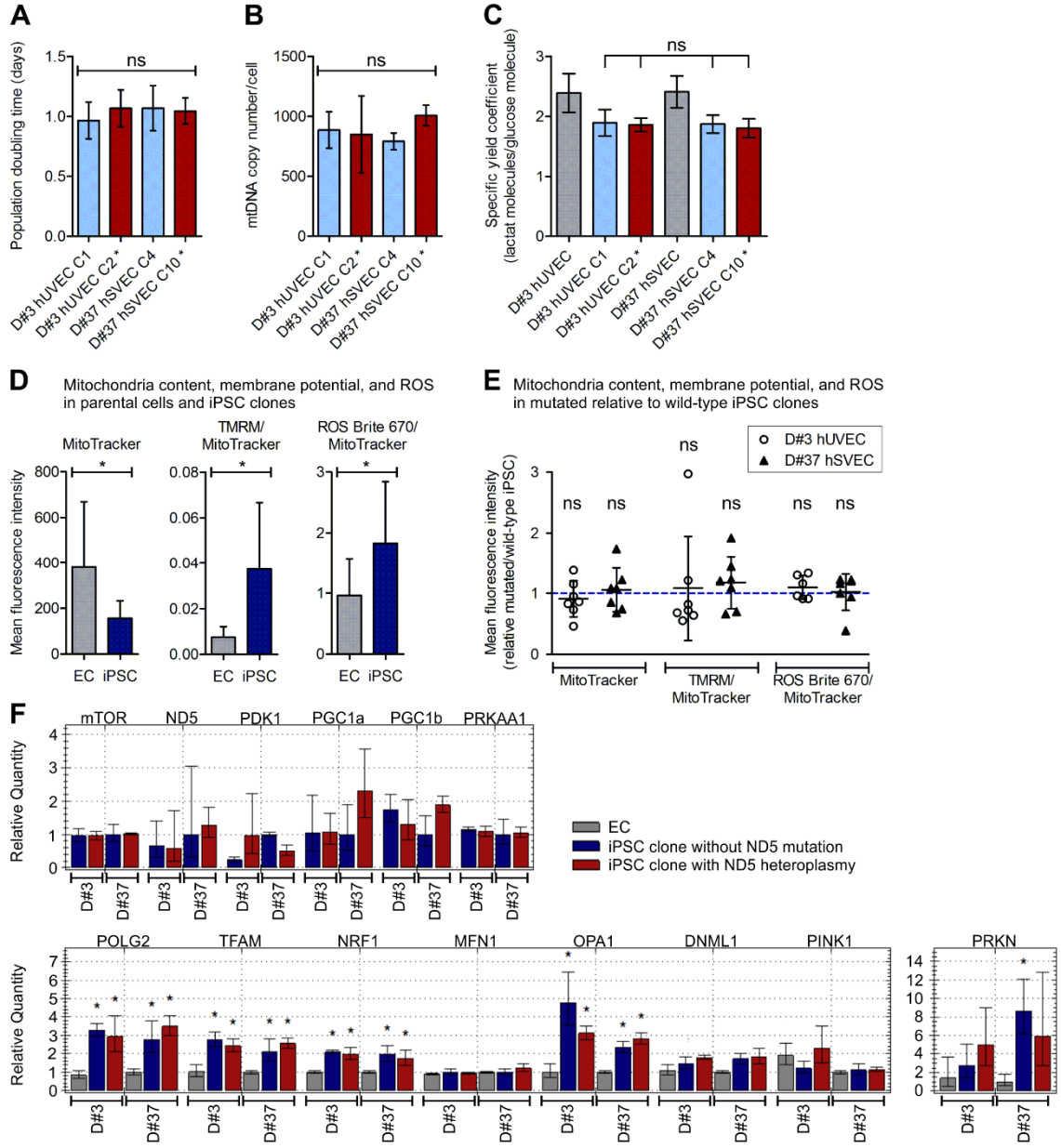

Figure S2: No detectable effect of specific *ND5* mutations on the functionality of undifferentiated iPSCs; Addressed in first Result section. iPSC clones with putatively actionable *ND5* mutations at an intermediate heteroplasmy level (D#37 hSVEC C10 with m.12686T>C at heteroplasmy level of ~23% and D#3 hUVEC C2 with m.13099G>A and ~38% heteroplasmy) and their sister iPSC clones without *ND5* mutations derived from the same parental cell population were analyzed. Mutated iPSC clones are marked with *. **A** Population doubling calculated as time in days needed for doubling of cell number measured by cell counting. Cell numbers were determined over 8-12 passages. N = 8-12 independent cell harvests each with 3 technical replicates. Mean with SD. Tested for significant differences by One-way Anova with Tukey's Multiple Comparison Test. **B** mtDNA copy number per iPSCs. Mean with SD. N = 2-4. No significant differences (One-way Anova). **C** Yield coefficient of lactate from glucose. Mean with SD. EC N = 2-3 and iPSC N = 3-4 biological replicates each with 2-3 technical replicates. Statistical analysis perfumed using One-way Anova. **D-E** Determination of mitochondrial content (MitoTracker), membrane potential (TMRM), and reactive oxygen species (ROS Brite 670) per cell via live cell staining and flow cytometric analysis. **D** Quantification calculated as geometric mean of fluorescence intensity of staining-positive populations. Mean with SD. EC N = 4; iPSC N = 8 including 4 independent batches of cells of mutated and 4 of wild-type clones. * indicates significant difference between groups (p-value < 0.05; Unpaired two-tailed t-test or nonparametric two-tailed Mann Whitney test). **E** Comparison of metabolic features of mutated against wild-type iPSC clones. Plot displays the deviation of values measured in mutated iPSC clones from the values in their wild-type counterparts quantified by geometric mean of fluorescence intensity and normalized to mitochondrial content (MitoTracker geometric mean). Mean

with SD. N = 7. Statistical analysis: Wilcoxon signed-rank test for derivation from hypothetical value of 1. **F** Gene expression analysis in iPSC clones and parental cells for genes involved in metabolism ((*mTOR*, *ND5*, *PDK1*, *PGC1a*, *PGC1b*, *PRKAA1*), mtDNA replication and transcription (*POLG2*, *TFAM*, *NRF1*), mitochondria dynamics (*MFN1*, *OPA1*, *DNML1*), and mitophagy (*PINK1*, *PRKN*) using qRT-PCR. Expression of all genes were normalized against housekeeping gene expression and is displayed relative to control iPSC clone D#37 hSVEC C4 (upper graph) or parental cells (D#37 hSVEC EC) (lower graph). Mean with SD; N = 3. Statistical analysis: One-way Anova with Tukey's Multiple Comparison Test for significant differences of gene expression in iPSC clones compared to parental cells (p-value < 0.05). No significant differences in gene expression between iPSC clones (One-way Anova with Tukey's Multiple Comparison).

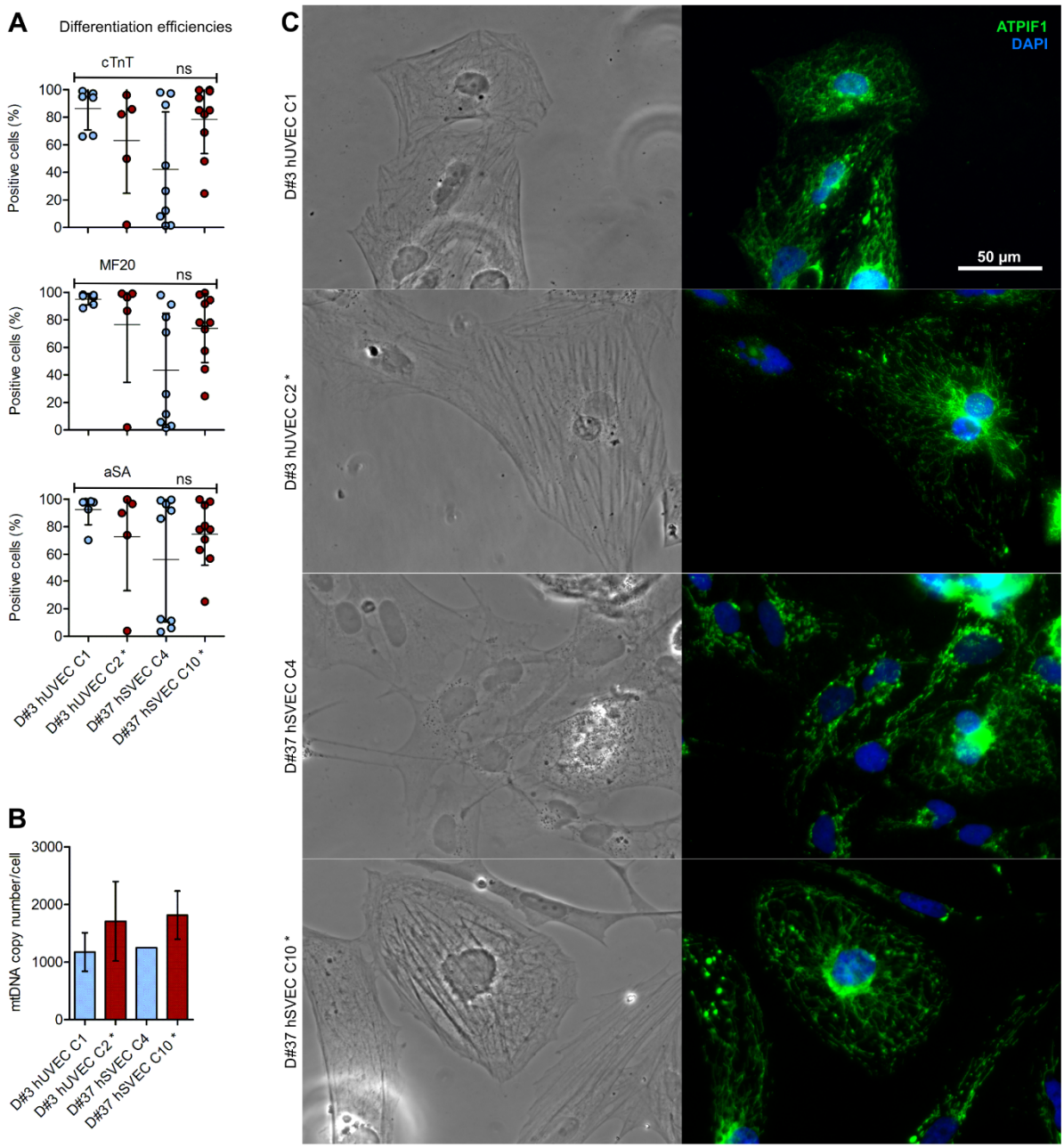

Figure S3: Intermediate heteroplasmy level of putatively actionable *ND5* mutation does not alter differentiation potential of iPSCs into CMs; Mentioned in third Result section. 2 mutated iPSC clones (marked with *; D#3 hUVEC C2 and D#37 hSVEC C10) each harbor a different putatively actionable mutation in *ND5* at intermediate heteroplasmy level and their isogenic sister iPSC clones were differentiated into CMs. **A** Differentiation efficiency quantified by positive staining for cTnT (Troponin-T), MF20 (MYH1E), and aSA (Sarcomeric α-actinin) and flow cytometry analysis at differentiation day (dd) 14. Mean with SD. N = 5-10 independent differentiations. Statistical analysis: Kruskal-Wallis test with Dunn's Multiple Comparison Test. **B** mtDNA copy number per CM determined at dd12-20. Mean with SD. N = 1 or 3 differentiations with differentiation efficiencies > 80% CM-marker positive cells. **C** CM at dd15 were seeded on fibronectin/gelatin (50µg/ml in 0.02% gelatin) coated glass cover slides, cultured for 3 days in differentiation medium, fixed with PFA, stained with anti-ATPIF1 (ATP synthase subunit IF1) antibody (green) and DAPI nuclear staining (blue), and analyzed by fluorescence microscopy. Scale bar 50µm.

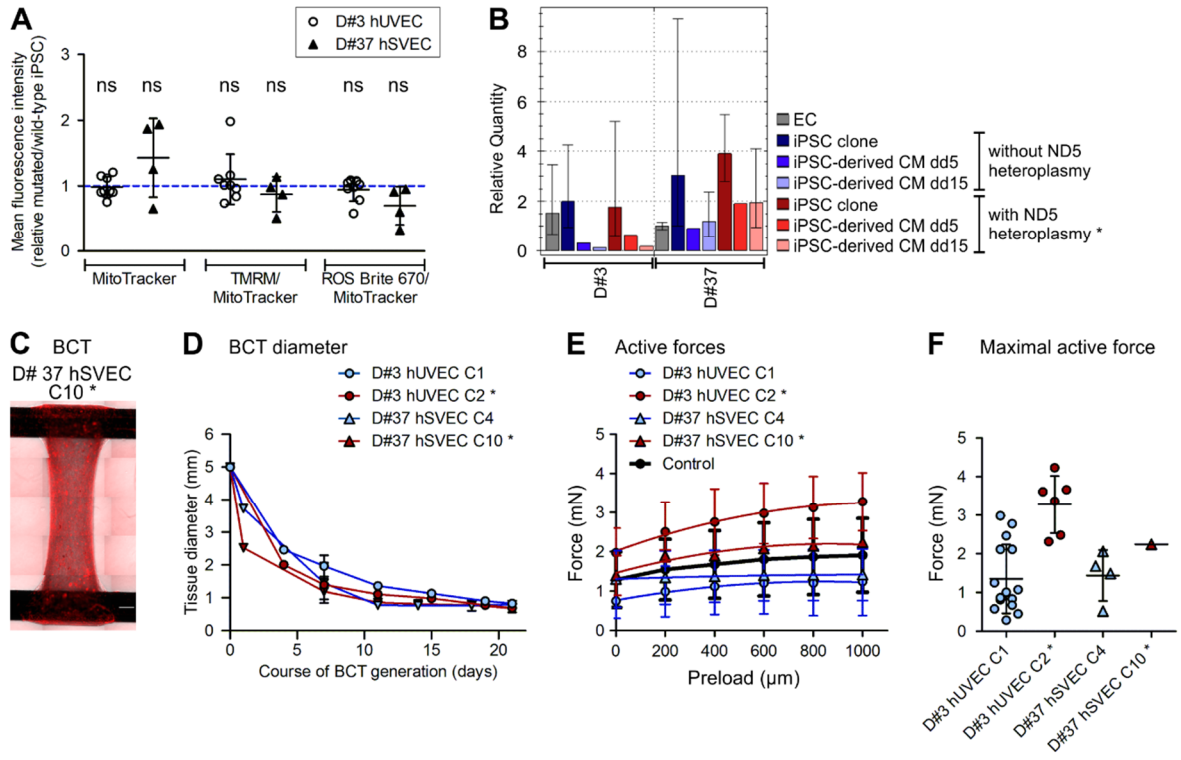

Figure S4: Impact of specific *ND5* mutations on CM functionality; Addressed in third Result section. iPSC clones with putatively actionable *ND5* mutations at an intermediate heteroplasmy level (marked with *; D#37 hSVEC C10 with m.12686T>C at heteroplasmy level of ~23% and D#3 hUVEC C2 with m.13099G>A and ~38% heteroplasmy) and their sister iPSC clones derived from the same parental cell population were differentiated into CMs. **A** Determination of mitochondrial content (MitoTracker), membrane potential (TMRM), and reactive oxygen species (ROS Brite 670) per cell via live cell staining and flow cytometric analysis. Comparison of metabolic features of mutated against wild-type iPSC-derived CMs. Plot displays the deviation of values measured in CMs derived of mutated iPSC clones from the values in their wild-type counterparts quantified by geometric mean of fluorescence intensity and normalized to mitochondrial content (MitoTracker geometric mean). Mean with SD. N = 4-8. Values do not deviate significantly from hypothetical value of 1 (Wilcoxon signed-rank test). **B** Relative quantification of *ND5* expression in parental endothelial cells (EC), iPSC clones and iPSC-derive CMs at differentiation day (dd) 5 and 15 using qRT-PCR. For each sample expression of *ND5* was normalized against housekeeping gene expression and is displayed relative to D#37 hSVEC EC. Mean with SD; iPSC N = 3; CM N = 1-2. **C-F** Bioartificial cardiac tissue (BCTs) were generated and active forces were measured. **C** Representative pictures of one BCT. Scale bar 500µm. **D** Diameter of BCT measured over course of BCT generation. Mean ± SD; N = 5-7 BCTs; 1-2 differentiations (biological replicates). **E** Active forces. Mean ± SD; N = 1-15 BCTs; 1-2 differentiations. Control N = 44 BCTs from 9 independent differentiations of 3 standard clonal iPSC lines. **F** Maximum active force. Mean ± SD; N = 1-15 BCTs; 1-2 differentiations.

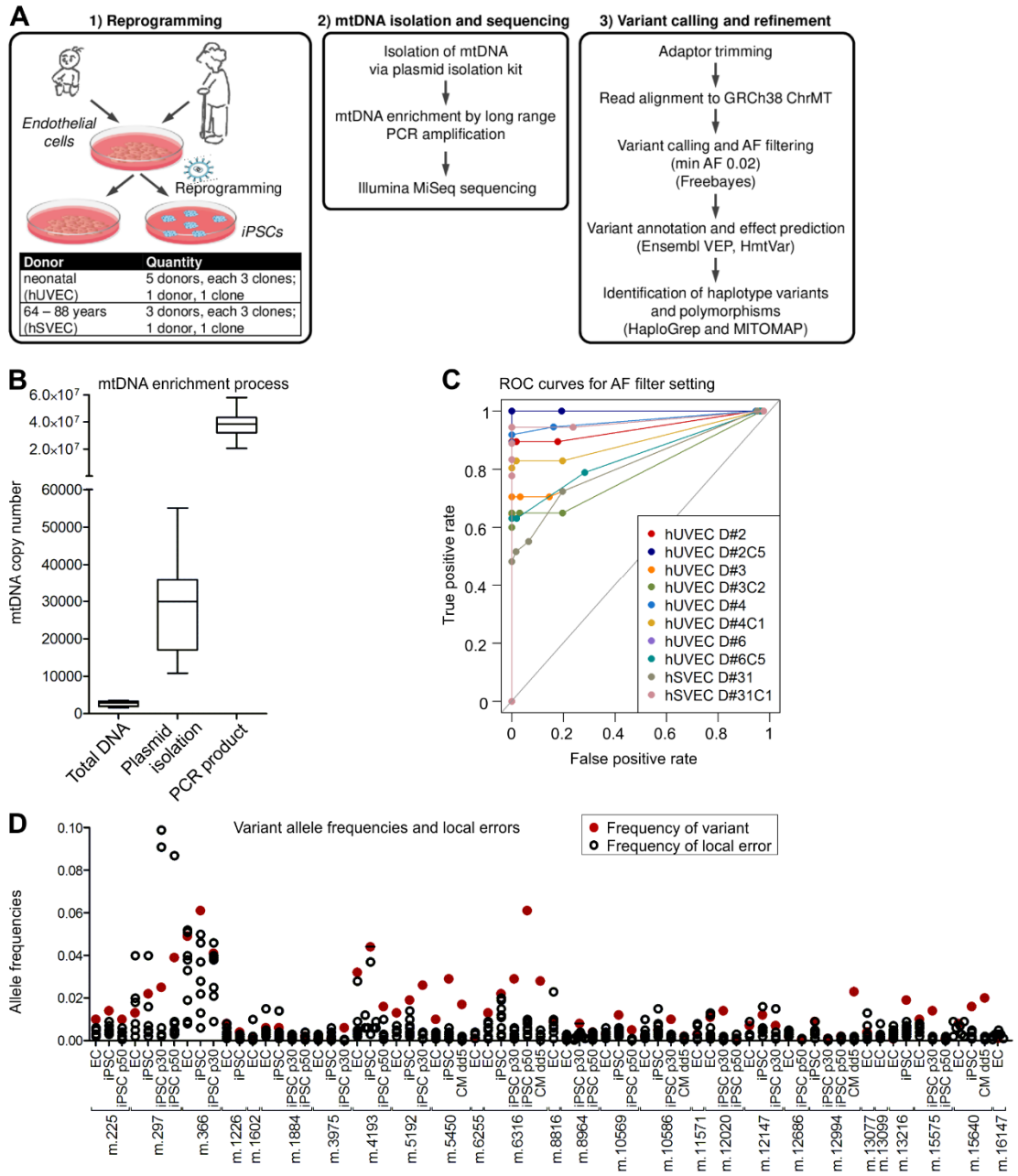

Figure S5: Establishment of our mtDNA sequencing approach; Related to Experimental Procedures section. A Workflow including cell reprogramming, sequencing and sequencing data analysis. **B** mtDNA enrichment by plasmid isolation and long-range PCR amplification for mtDNA sequencing. mtDNA copy number is measured relative to gDNA copy number after total DNA isolation, which represents normal cellular proportion of mtDNA and gDNA, after plasmid isolation, and after additional long range PCR amplification of mtDNA. N = 12. **C** Receiver operating characteristic (ROC) curve with AF as discrimination threshold (AF 0.005, AF 0.01, AF 0.02, AF 0.04, AF 0.06, AF 0.08, AF 0.1, AF 0.5, AF 1). **D** A choice of variants with very low variant allele frequencies in iPSC clones and/or parental cells were additionally investigated individually within their genetic context by inspecting and counting sequencing reads directly. The plot shows the AF of single nucleotide variants (SNVs), calculated from the number of reads with variant, in red against the background of reads with the non-reference and invariant nucleotide exemplarily for 27 loci.

Table S1: Variants detected by mtDNA sequencing; Related to Table 1. List of all variants detected by mtDNA sequencing of 26 iPSC clones at early passage (on average p6.5), of additional 7 of the iPSC clones at intermediate (p30) and high (p50) passage, of the 10 corresponding parental cell populations, and cardiomyocytes differentiated off 4 of the iPSC clones during differentiation process at differentiation day (dd) 0, dd5, and dd15. Variants were divided into groups and categories: Variants of group A, non-transmitted variants, were detected only in parental cell populations. Group B variants were homoplasmic in parental cell population and iPSCs derived thereof. Group C, heteroplasmic variants were detected in an iPSC clone with AF > 0.02 at any passage but were less frequent in the corresponding parental cell population. A final group comprises variants found heteroplasmic in differentiated cardiomyocytes. Those variants basically belong to the group of heteroplasmic variants that were transmitted to cardiomyocytes during differentiation. Variants of category 1 are haplotype variants defining a haplogroup. Category 2 variants are polymorphisms and defined by a population variant frequency (MITOMAP (NCBI GenBank)) ≥ 0.02. Category 3 variants are donor-specific variants which were unique to a donor and rare in the population context. heteroplasmic variants were further divided into the groups of amplified variants (which increase in heteroplasmy level of > 10 fold during iPSC culture expansion), stable variants (with fold change ~1), purified variants (decreased by > 10 fold), fluctuating variants (change of heteroplasmy level did not allow confident classification in one of the above mentioned groups), and "nd in p30-50" variants (iPSC clones not analyzed in high passages). Identification of haplotype variants and polymorphisms was based on HaploGrep, NCBI GenBank variant frequency (obtained from MITOMAP), and HmtVar. Variant effect prediction was performed, based on a consensus of *in silico* prediction algorithms obtained from snpEff impact, CADD, Condel, and HmtVar. The output of each tool for each variant is presented in the respective columns and variants classified as putatively actionable by a harmful designation by snpEff (high impact which indicates frame shift mutations) or at least by two of the other algorithms are marked in red in the consensus column. A choice of variants with very low variant allele frequencies in iPSC clones and/or parental cells were additionally investigated individually within their genetic context by inspecting and counting sequencing reads directly. Variant allele frequencies that were determined using this method are underlined in yellow. (); existence of variant in parental cell population or iPSC clone not confirmed with statistical confidence (p-value > 0.05).

Supplemental excel file

Table S2: Manual determination of very low variant allele frequencies; Related to Experimental Procedures section and mentioned in first Result section.
A Manual determination of very low variant allele frequencies of single nucleotide variants (SNVs). The table lists SNVs and their allele frequencies in iPSC clones and parental cell population determined by individual examination of variants within their genetic context. Results are displayed as the number of reads per respective nucleotide embedded in an 11bp sequence. The number for reference allele reads are given in bold, and for variant allele reads in red. Average local error rates calculation and results, statistical test results, as well as local detection limits are presented.
B Manual determination of very low variant allele frequencies of small insertions. The table lists the small insertions and their allele frequencies in iPSC clones and parental cell population determined by individual examination of variants within their genetic context. Results are displayed as the number of reads for the reference (bold) and variant sequence (red) embedded in an 11bp sequence.

Supplemental excel file

Table S3: Karyotype analysis of parental cells and iPSC clones; Mentioned in second Result section.

Sample	Karyotype	Sample	Karyotype	
Neonatal donors				
Parental cells		**iPSC clones**		
			Early passage (~p8)	**Late passage (~p52)**
D#1 hUVEC	46,XX[15]	**D#1 hUVEC C2**	46,XX[10]	Not analyzed
		D#1 hUVEC C5	46,XX[10]	Not analyzed
		D#1 hUVEC C6	46,XX[10]	Not analyzed
D#2 hUVEC	46,XX[15]	**D#2 hUVEC C4**	46,XX[10]	Not analyzed
D#3 hUVEC	46,XX[15]	**D#3 hUVEC C1**	46,XX[15]	46,XX[20]
		D#3 hUVEC C2	Not analyzed	46,XX[15]
D#4 hUVEC	Not analyzed	**D#4 hUVEC C1**	46,XY[15]	Not analyzed
D#5 hUVEC	Not analyzed	**D#5 hUVEC C1**	46,XX[12]	Not analyzed
D#6 hUVEC	Not analyzed	**D#6 hUVEC C2**	46,XY[12]	Not analyzed
Aged donors				
Parental cells		**iPSC clones**		
			Early passage (~p8)	**Late passage (~p52)**
D#31 hSVEC (64 years)	47,XY,+12[3] / 46,XY,del(13)(q21q31)[2] / 46,XY[10]	**D#31 hSVEC C1**	47,XY,+12[10] / 46,XY[3]	Not analyzed
		D#31 hSVEC C3	47,XY,+12[13] / 48,idem,+mar[2]	Not analyzed
		D#31 hSVEC C4	47,XY,+12[14]	Not analyzed
D#37 hSVEC (73 years)	46,XY[10]	**D#37 hSVEC C8**	46,XY[10]	Not analyzed
		D#37 hSVEC C10	Not analyzed	46,XY[26]
D#38 hSVEC (88 years)	46,XY[10]	**D#38 hSVEC C5**	46,XY[10]	Not analyzed
		D#38 hSVEC C6	46,XY[10]	Not analyzed
		D#38 hSVEC C9	46,XY[10]	Not analyzed
D#40 hSVEC (79 years)	46,XY[10]	**D#40 hSVEC C4**	46,XY[10]	Not analyzed

Table S4: Primer list; Related to Experimental Procedures section. Primer sequences for mtDNA copy number determination by qRT-PCR and whole mtDNA amplification for mtDNA sequencing.

Gene symbol	Primer Forward sequence	Primer Reverse sequence	Product length	Annealing (°C)	Application	Source
LYZL4	GGC CTT CTT ATT GCT GAG GTT C	TAG TCT TCT CCC TTC CTG TCT GC	123	66	Copy number	
PTBP2	TCT CCA TTC CCT ATG TTC ATG C	GTT CCC GCA GAA TGG TGA GGT G	114	65	Copy number	
GKN2	TGC TGA CCA TCT TTG GGA TAC	TTC AGG CTC TCT TGT ACC TTC C	112	60	Copy number	
mtDNA 8294	CCA CTG TAA AGC TAA CTT AGC ATT AAC C	GTG ATG AGG AAT AGT GTA AGG AGT ATG G	143	63	Copy number	Andréasson, BioTechniques. 2002
RNR2	CGT TCA AGC TCA ACA CCC ACT AC	GTT TTA ATC TGA CGC AGG CTT ATG	156	62	Copy number	
ND1	ATA CCC ATG GCC AAC CTC CTA C	TTC ATA GTA GAA GAG CGA TGG TGA GAG	259	62	Copy number	Wurmb-Schwark, Forensic Sci Int. 2002
F2480A/R 10858A	AAA TCT TAC CCC GCC TGT TT	AAT TAG GCT GTG GGT GGT TG	8379	54	Whole mtDNA amplification	McElhoe, Forensic Sci Int Genet. 2014
F10653B/ R2688B	GCC ATA CTA GTC TTT GCC GC	GGC AGG TCA ATT TCA CTG GT	8605	54	Whole mtDNA amplification	McElhoe, Forensic Sci Int Genet. 2014
PINK1	GGA CGC TGT TCC TCG TTA	ATC TGC GAT CAC CAG CCA	218	63	Gene expression	Gegg, PLOSone. 2009
PRKN	GCA TCT TCC AGC TCA AGG AG	CTT TTC TCC ACG GTC TCT GC	165	64	Gene expression	Li, J Clin Med. 2018
Mfn1	TGT TTT GGT CGC AAA CTC TG	CTG TCT GCG TAC GTC TTC CA	160	56	Gene expression	Cartoni, J Physiol. 2005
OPA1	GTG CTG CCC GCC TAG AAA	AGA CTG GCA GAC CTC ACA GG	282	65	Gene expression	
DNM1L	AGG AAG GAG GCG AAC TGT G	TGC TCT GCG TTC CCA CTA C	185	65	Gene expression	
POLG2	GCA GTC CTC ATT GGA ACA ACT	TTG TGG TGT CTC TGC TTC TCA	122	60	Gene expression	
TFAM	ATG GCG TTT CTC CGA AGC AT	CAG ATG AAA ACC ACC TCG GTA A	133	61	Gene expression	Prigione, Stem cell. 2010
NRF1	AAC AAA ATT GGG CCA CGT TAC A	TCT GGA CCA GGC CAT TAG CA	291	62	Gene expression	Prigione, Stem cell. 2010
PRKAA1	CAG CCG AGA AGC AGA AAC AC	TTT GCC AAC CTT CAC TTT GC	101	60	Gene expression	
PGC1A	TCA GTC CTC ACT GGT GGA CA	TGC TTC GTC GTC AAA AAC AG	351	61	Gene expression	Cartoni, J Physiol. 2005
ND5	TAT GTG CTC CGG GTC CAT C	CAG GGA GGT AGC GAT GAG AG	235	63	Gene expression	Amstrong, Stem cell. 2010
PINK1	GGA CGC TGT TCC TCG TTA TG	GGA TGT TGT CGG ATT TCA GG	171	65	Gene expression	adapted Gegg, PLOSone. 2009
PDK1	TTC GGA TCA GTG AAT GCT TG	ACC AAT TGA ACG GAT GGT GT	137	56	Gene expression	Vengellur, Physiol genomics. 2005
mTOR	AGG CCG CAT TGT CTC TAT CA	GCA GTA AAT GCA GGT AGT CAT CC	80	62	Gene expression	adapted Lv, Cell Physiol Biochem. 2017
PGC1B	CCT TGT GTT AAG GCG GAC AG	CCG AGG TGA GGT GCT TAT GTA G	157	62	Gene expression	
Mfn2	ATG CAT CCC CAC TTA AGC AC	CCA GAG GGC AGA ACT TTG TC	301	62	Gene expression	Cartoni, J Physiol. 2005
ß-MHC	GGC ACA GCC ATG GGA GAT TC	TCG AAC TTG GGT GGG TTC TGC	263	60	Gene expression	
α-MHC	CCG TGA AGG GAT AAC CAG GGG	ACT TGG GTG GGT TCT GCT GC	281	60	Gene expression	
PLN	GCT GCC AAG GCT ACC TAA AAG	GAC GTG CTT GTT GAG GCA TTT	182	60	Gene expression	
RYR2	AGA ACT TAC ACA CGC GAC CTG	CAT CTC TAA CCG GAC CAT ACT GC	199	60	Gene expression	
Troponin (TNNT2)	TTC ACC AAA GAT CTG CTC CTC GCT	TTA TTA CTG GTG TGG AGT GGG TGT GG	166	60	Gene expression	
CX43	Gap junction protein, alpha 1, 43kDa (Bio Rad; Cat. #10025636)		130	60	Gene expression	
GAPDH	CCA TCT TCC AGG AGC GAG ATC	GCA GAG ATG ATG ACC CTT TTG G	145	61	Gene expression	
RPL13A	CGC CCT ACG ACA AGA AAA AGC	CCT GGT ACT TCC AGC CAA C	123	61	Gene expression	

Table S5: Variants with minimum AF of 0.02 can be detected in sequencing replicates with repetitious accuracy; Related to Experimental Procedures section. Comparison of variant allele frequencies (AF) in early passages (p6) of three iPSC clones (D#3 hUVEC C1, D#3 hUVEC C2, and D#37 hSVEC C10) and in samples of differentiation day (dd) 0 from targeted cardiomyocyte (CM) differentiations of the respective clones (upper part). The dd0 samples of these CM differentiations for which differentiation was initiated at p6 represent basically independent replicates (including separate sample preparation and sequencing) of the sequencing of the corresponding iPSC clones. Absence of substantial variation in the heteroplasmy levels between dd0, dd5, and dd15 samples of CM differentiations of different iPSC clones (lower part) rebuts a role of amplification and sequencing artefacts.

Variants	Donor	iPSC clone	Variant allele frequency		Fold change	delta AF
			p6.5	dd0		
m.12020C>T	D#37 hSVEC	C10	0.026	0.027	1.04	0.00
m.12686T>C	D#37 hSVEC	C10	0.225	0.203	1.11	0.02
m.72T>C	D#3 hUVEC	C2	0.929	0.946	1.02	0.02
m.5192A>C	D#3 hUVEC	C1	0.019	0.027	1.41	0.01
m.5894_5895insC	D#37 hSVEC	C10	0.979	0.971	1.01	0.01
m.13099G>A	D#3 hUVEC	C2	0.374	0.363	1.03	0.01
m.4193T>C	D#3 hUVEC	C1	0.044	0.013	3.49	0.03
m.5450C>T	D#3 hUVEC	C2	0.029	0.020	1.41	0.01

Variants	Donor	iPSC clone	Variant allele frequency			Fold change	delta AF
			dd0	dd5	dd15		
m.72T>C	D#3 hUVEC	C2	0.946	0.948	0.967	0.99	0.01
m.297A>C	D#37 hSVEC	C4	0.071	0.077	0.073	0.98	0.00
m.4193T>C	D#3 hUVEC	C1	0.013	0.010	0.014	0.96	0.00
m.5192A>C	D#3 hUVEC	C1	0.027	0.024	0.022	1.13	0.00
m.5450C>T	D#3 hUVEC	C2	0.020	0.017	0.018	1.11	0.00
m.5894_5895insC	D#37 hSVEC	C10	0.971	0.972	0.967	1.00	0.00
m.6316A>C	D#37 hSVEC	C4	0.032	0.028	0.032	0.99	0.00
m.10569G>A	D#3 hUVEC	C2	0.014	0.016	0.012	1.09	0.00
m.12020C>T	D#37 hSVEC	C10	0.027	0.032	0.028	0.98	0.00
m.12686T>C	D#37 hSVEC	C10	0.203	0.185	0.205	1.00	0.01
m.13099G>A	D#3 hUVEC	C2	0.363	0.358	0.364	1.00	0.00

Table S6: Antibody and dye list; Related to Experimental Procedures section.

Antibody / Dye	Host / Isotype	Dilution	Company	Application
ATPIF1	Mouse IgG1	1:200	Thermo Fisher Scientific	Immunofluorescence
Anti-mouse IgG-Alexa Fluor 488	Donkey	1:200	Jackson ImmunoResearch	Immunofluorescence
DAPI		1:30000	Sigma-Aldrich	Immunofluorescence
MitoTracker		1µM	Invitrogen	Flow cytometry
ROS Brite 670		5µM	Biomol; AAT Bioquest	Flow cytometry
Tetramethylrhodamine (TMRM)		100nM	Thermo Fisher Scientific	Flow cytometry
Sarcomeric α-actinin	mouse IgG1	1:800	Sigma-Aldrich	Flow cytometry
MYH1E (MF20)	mouse IgG2a	1:50	Hybridoma Bank, University of Iowa, US	Flow cytometry
Troponin-T	mouse IgG1	1:100	Richard Allan Scientific	Flow cytometry
Anti-mouse IgG Alexa Fluor 647	Donkey	1:300	Jackson ImmunoResearch	Flow cytometry

Data S1: Galaxy workflows; Related to Experimental Procedures section.

Supplemental file

Supplemental Experimental Procedures:

Reprogramming and iPSC culture. Derivation and culture of endothelial cells, virus production, retroviral reprogramming, iPSC characterization, and culture was performed as previously described (Haase et al., 2009). In short, endothelial cells (ECs) were isolated from umbilical vein (hUVEC) from healthy newborns and saphenous vein (hSVEC) from aged patients (64-88 years) that underwent coronary bypass surgery. Human material was collected after approval by the local Ethics Committee and following the donor's or the newborn's parental written informed consent. ECs were cultivated in Endothelial Growth Medium (EGM-2) (Lonza) and $2*10^5$ cells of early passages (mean p4.5; range p3-7) were reprogrammed by ectopic expression of Oct4, Sox2, Nanog, and Lin28 (lentiviral transduction, monocistronic factors, multiplicity of infection (MOI) 20) with exception of clone D#40 hSVEC C5 which was reprogrammed via ectopic expression of Oct4, Sox2, Klf4, and c-Myc (lentiviral transduction, monocistronic factors, MOI 1). In total, 26 EC-derived iPSC clones were derived from 10 donors. 3 clones were generated for each donor D#1 hUVEC-D#5 hUVEC, D#31 hSVEC, D#37 hSVEC, and D#38 hSVEC. 1 clone was selected for the donors D#6 hUVEC and D#40 hSVEC. iPSC clones were cultivated as colonies on mouse embryonic fibroblasts (MEF) in Knock-out Dulbecco's Modified Eagle's medium (DMEM) (Gibco) supplemented with 20% KnockOut serum replacement, 1% Non-essential Amino acids, 1mM L-Glutamine, 0.1mM β-Mercaptoethanol (all obtained from Life Technologies), and 10ng/ml bFGF (Institute for Technical Chemistry, Leibniz University Hannover, Germany). Before analysis, iPSCs were transferred to monolayer culture on Geltrex (Thermo Fisher Scientific) in in-house Essential 8 (E8) medium (DMEM Nutrient Mixture F-12 (Gibco) supplemented with 543mg/L $NaHCO_3$, 2ng/ml TGFβ (PeproTech), 10.7µg/ml human recombinant transferrin, 14µh/l sodium selenite, 20µg/l insulin, 64µg/l ascorbic acid 2-phosphate (all obtained from Sigma-Aldrich), and 100ng/ml bFGF) and cultured for generally 3-5 passages.

Primer Design for the determination of mitochondrial DNA copy number using quantitative real-time PCR (qRT-PCR). A quantitative real-time PCR (qRT-PCR) based analysis system was developed which quantifies mitochondrial DNA (mtDNA) copy number relative to genomic DNA (gDNA) copy number. In contrast to most other studies which use only one primer pair, our system measures number of mtDNA and gDNA copies by each three primer pairs. Sequences for two primer pairs targeting mtDNA regions were assimilated from previous studies (Andreasson et al., 2002; von Wurmb-Schwark et al., 2002). The other primer pairs were designed employing Primer3 (v0.4.0). During primer design, target loci were chosen that were not within the common 4700bp mtDNA deletion site or genomic regions with high CNVs or mutation occurrence (manually inspected using variation data provided by Ensembl (release 90, August 2017) (Hunt et al., 2018).

mtDNA extraction and amplification. The workflow for mtDNA sequencing and data processing is displayed in Figure S5A. mtDNA was extracted via QIAprep Spin Miniprep kit (Qiagen) from early passage iPSC clones (mean p6.5), and from late passages (p30 and p50) of 7 clones. Furthermore, enriched mtDNA was isolated from the corresponding parental cell populations at a similar passage as subjected to reprogramming (mean p4.5). Isolation was essentially performed according to manufacturer's recommendation and yielded, on average, 12 fold increase of mtDNA compared to total DNA isolation (Figure S5B). Entire mtDNA was then amplified using 2 primer pairs (F2480A/R10858A (3:1) and F10653B/R2688B (1:1)) (Table S4) generating two overlapping PCR products of 8379bp and 8605bp fragment size (McElhoe et al., 2014). PCR reactions were performed with Herculase II Fusion DNA Polymerase (Agilent) in technical replicates with 50ng enriched mtDNA as template per reaction. Altogether, the input of enriched mtDNA accounted for 200ng which equals to ~$5*10^8$ mtDNA molecules ($3.2*10^4$ cells (theoretical input) * 12 (average fold enrichment) * 1300 (average mtDNA copy number per cell)), isolated from, on average, $5*10^6$ iPSCs or $2.5*10^6$ parental cells. PCR conditions were: 94°C for 5min as initiation step, then 23 cycles of 94°C for 40s, 54°C for 20s, and 54°C for 4.15min, followed by termination step of 72°C for 3min. Therefore, our protocol reduced PCR amplification cycle to 23 compared to 30 in most other published approaches of mtDNA amplification for sequencing purposes. mtDNA PCR amplification resulted in an enrichment of mtDNA of, on average, ~15000 fold over normal cellular content (Figure S5B). Each amplicon was individually inspected by gel electrophoresis and amplicons from the same sample were pooled,

Read trimming, quality assessment, and alignment. Quality and adapter trimming of reads was performed by Trim Galore (v0.4.3; Cutadapt v1.14) including removal of adapter, 10bp from the 3' end of read 1, 30bp from the 3' end of read 2, as well as low-quality ends from reads (Phred < 28). Reads were aligned to the human genome GRCh38 chromosome MT (the Cambridge Reference sequence (rCRS; NC_012920)) using BWA-MEM (Galaxy Version 0.7.17.1) (Li, 2013). Quality of trimmed and aligned reads were inspected employing FastQC Read Quality reports (v.0.11.6; Galaxy wrapper version v0.71) and Qualimap Multi-sample BAM QC analysis (v.2.2.1). Reads had an average length of 215bp, with Phred score > 28 over whole read length. 98% of all reads mapped to mitochondrial reference genome with equal quality across the reference genome (59.8 mean mapping quality). The coverage was 19000 on average.

Variant calling and refinement. Variants were called using FreeBayes (v1.0.2; Galaxy implementation v1.0.2.29-3) (Ewing and Green, 1998; Ewing et al., 1998) with filters set to ploidy 10, minimum coverage 100, and minimum AF (alternative frequency) 0.02. The evaluation of the AF filter threshold was performed based on knowledge of population frequency of variants and the high coverage was utilized to identify false and true variants. Variants that were detected at a low frequency in iPSC clones of different donors but were not describe as polymorphism were unlikely to be true-positive. Subsequently some selected variants were individually analyzed as described below in the section "Detection of variants at very low AF". A receiver operating characteristic (ROC) curve with AF as discrimination threshold (Figure S5C) showed that AF 0.02 retrieve very high specificity while maintaining most true positive variants. To exclude the few remaining false positive variants, remaining primer sequences that had evaded the trimming process were manually removed and variants that were detectable in different donors (with minimum AF 0.005) but not described as polymorphisms were excluded as probable false variant calls.

Determination of allele fraction of variants with very low AF in iPSC clones and parental cell population. The presence of a choice of variants (in total 128 variants at 38 different genetic regions) at very low AF (in general, ≤ 0.02) in iPSC clones or parental cell populations was examined individually within their genetic context. For this purpose, sequencing reads were inspected directly from fastq raw files. Both reads of the paired end sequencing were investigated and the number of reads with the variant enclosed by a 11bp long sequence ranging from position -5bp to +5bp was counted as well as the number of reads with the reference sequence. For calculation of average error for each genetic location, the numbers of reads that comprised not the variant or reference nucleotide were determined. In addition, the 11bp long sequences directly upstream and downstream of the variant were analyzed in the same manner. Hence, the average error of the amplicon sequencing was calculated for every variant as the mean number of reads with non-reference and non-variant nucleotide at variant location or in the middle of analyzed upstream and downstream sequence (in general, N = 8) and presence of variant was confirmed with p-value 0.05 against background of errors (Figure S5D, Table S2A). Frequencies of INDELs were determined by extracting and counting reads for both INDEL and reference sequence embedded by an 11bp sequence, directly from fastq files (Table S2B). Average error of the amplicon sequencing was calculated for every variant as the mean number of reads with non-reference and non-variant nucleotide at variant location or in the middle of analyzed upstream and downstream sequence (in general, N = 8). Distribution of errors for each variant was evaluated via the D'Agostino and Pearson omnibus normality test (where applicable, else the Shapiro-Wilk normality test was used). One sample one-tailed t-test and Wilcoxon one-tailed signed-rank test were performed, respectively, to compare number of reads with variant as hypothetical value against background of errors. The null hypothesis, assuming number of reads with variant is within local error range, was rejected with p values of 0.05.

Karyotype analysis. After treatment of hiPSCs with colcemid (Invitrogen) for 30 min, cells were detached with trypsin and metaphases were prepared according to standard procedures. Fluorescence R-banding using chromomycin A3 and methyl green was performed as previously described (Schlegelberger et al., 1999). At least 10 metaphases were analyzed per clone at a minimum of 300 bands. Karyotypes were described according to the international System for Human Cytogenetic Nomenclature (ISCN).

Differentiation of cardiomyocytes (CMs). For CM differentiation, $1*10^6$ undifferentiated cells/well were aggregated in 3ml E8 medium containing 10µM Rho-Kinase inhibitor Y-27632 (RI) on low attachment 6 well plates (Greiner Bio-one GmbH) on an orbital shaker (70rpm; Infors GmbH) at differentiation day (dd) -3). Differentiation was started (dd0) using 5µM CHIR99021 (LU Hannover) in CDM3 (Burridge et al., 2014; Halloin et al., 2019) to activate the WNT pathway. Exactly 24h later it was followed by a subsequent inhibition of the WNT pathway (dd1), where CDM3 medium including 2µM Wnt-C59 (LU Hannover) was added for 48h. On dd3 and dd5 fresh CDM3 medium was added. From dd7 on, the medium was changed to basic serum free medium (bSF; DMEM with 1% non-essential amino acids, 5.6mg/l transferrin, 37.2µg/l sodium-selenit, 1mM L-glutamine, and 10µg/ml insulin) and was changed every other day. The differentiation efficiency was addressed on dd14 by flow cytometry for the cardiac markers, cardiac troponin T (cTnT), alpha sarcomeric actinin (αSA), and pan-myosin heavy chain (MF20) as described earlier (Halloin et al., 2019) (Table S6). mtDNA of $5*10^6$ iPSC-derived cardiomyocytes at dd0, dd5, and dd15 was isolated, amplification, and subjected to sequencing as described above.

Bioartificial cardiac tissues (BCT) generation and cultivation. BCTs were generated and cultivated as described earlier (Kensah et al., 2011; Kensah et al., 2013). Briefly, differentiated CMs were dissociated using STEMdiff™ Cardiomyocyte Dissociation Kit (STEMCELL Technologies) according to manufacturer's instructions. 250µl cell-matrix mixture/BCT composed of 10% Matrigel (BD Biosciences), 1.35mg/ml rat collagen type I (Trevigen), 3.6% 0.4M NaOH, $1*10^6$ dissociated CMs, and $1*10^5$ irradiated human foreskin fibroblasts (ATCC) was poured into bottomless silicon molds containing two titanium rods (6mm initial slack length). Following 30min solidification, the tissues were cultivated in BCT medium with 30µM L-ascorbic acid

(Kensah et al., 2011) for 21 days, where spontaneously contracting BCTs formed around the rods. Starting on day 7, a growing static stretch protocol of 0.4mm/4days was applied. Tissue maturation and CM viability was monitored during cultivation using an AxioObserver Z1 fluorescence microscope and ZEN 3.1 software (Zeiss). To visualize the distribution of viable CMs, the mitochondria-specific dye tetramethylrhodamine methyl ester (TMRM, 25nM; Thermo Fisher Scientific) was added. Active contraction forces and spontaneous beating frequencies of BCTs were measured on day 21 of tissue cultivation as described earlier 53 using a custom-made bioreactor system (Central Research Devices Service Unit, MHH, Hannover). BCTs were placed into a culture vessel (BCT medium, 37°C) allowing for contraction force recording, and precise application of increasing preload. After every preload step (200µm increments), spontaneous and paced contraction (mean of 5 biphasic pulses, 10ms, 25V) values were measured.

Immunofluorescence staining and microscopic analysis. ECs, iPSCs, and CMs were seeded on cover slides coated with 1% gelatin, Geltrex, or 50µg/ml fibronectin in 0.02% gelatin, respectively. After a culture period of 1-2 days for ECs and iPSCs and 3 days for CMs, cells were fixed with 4% Paraformaldehyde (PFA) for 20min at 4°C, permeabilized with 0.1% Triton X 100 (Sigma-Aldrich) in PBS for 5min at room temperature, and unspecific binding sites were blocked with solution of 5% donkey serum (Chemikon) and 0.25% Triton X 100 in tris-buffered saline (TBS) for 30min at 4°C. Incubation with monoclonal anti-ATPIF1 (ATP synthase subunit IF1) antibody diluted in staining buffer (PBS with 1% bovine serum albumin (BSA) (Sigma-Aldrich)) was performed overnight at 4°C. List of antibodies is provided in Table S6. Secondary antibody staining was performed for 1h at 4°C and DAPI nuclear staining for 1min.

Supplemental References:

Andreasson, H., Gyllensten, U., and Allen, M. (2002). Real-time DNA quantification of nuclear and mitochondrial DNA in forensic analysis. Biotechniques *33*, 402-404, 407-411.

Burridge, P.W., Matsa, E., Shukla, P., Lin, Z.C., Churko, J.M., Ebert, A.D., Lan, F., Diecke, S., Huber, B., Mordwinkin, N.M., *et al.* (2014). Chemically defined generation of human cardiomyocytes. Nat Methods *11*, 855-860.

Ewing, B., and Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res *8*, 186-194.

Ewing, B., Hillier, L., Wendl, M.C., and Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res *8*, 175-185.

Haase, A., Olmer, R., Schwanke, K., Wunderlich, S., Merkert, S., Hess, C., Zweigerdt, R., Gruh, I., Meyer, J., Wagner, S., *et al.* (2009). Generation of induced pluripotent stem cells from human cord blood. Cell Stem Cell *5*, 434-441.

Halloin, C., Schwanke, K., Lobel, W., Franke, A., Szepes, M., Biswanath, S., Wunderlich, S., Merkert, S., Weber, N., Osten, F., *et al.* (2019). Continuous WNT Control Enables Advanced hPSC Cardiac Processing and Prognostic Surface Marker Identification in Chemically Defined Suspension Culture. Stem Cell Reports *13*, 366-379.

Hunt, S.E., McLaren, W., Gil, L., Thormann, A., Schuilenburg, H., Sheppard, D., Parton, A., Armean, I.M., Trevanion, S.J., Flicek, P., *et al.* (2018). Ensembl variation resources. Database *2018*.

Kensah, G., Gruh, I., Viering, J., Schumann, H., Dahlmann, J., Meyer, H., Skvorc, D., Bar, A., Akhyari, P., Heisterkamp, A., *et al.* (2011). A novel miniaturized multimodal bioreactor for continuous in situ assessment of bioartificial cardiac tissue during stimulation and maturation. Tissue Eng Part C Methods *17*, 463-473.

Kensah, G., Roa Lara, A., Dahlmann, J., Zweigerdt, R., Schwanke, K., Hegermann, J., Skvorc, D., Gawol, A., Azizian, A., Wagner, S., *et al.* (2013). Murine and human pluripotent stem cell-derived cardiac bodies form contractile myocardial tissue in vitro. European Heart Journal *34*, 1134-1146.

McElhoe, J.A., Holland, M.M., Makova, K.D., Su, M.S., Paul, I.M., Baker, C.H., Faith, S.A., and Young, B. (2014). Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq. Forensic Sci Int Genet *13*, 20-29.

Schlegelberger, B., Metzke, S., Harder, S., Zühlke-Jenisch, R., Zhang, Y., and Siebert, R. (1999). Classical and Molecular Cytogenetics of Tumor Cells. In Diagnostic Cytogenetics, R.-D. Wegner, ed. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 151-185.

von Wurmb-Schwark, N., Higuchi, R., Fenech, A.P., Elfstroem, C., Meissner, C., Oehmichen, M., and Cortopassi, G.A. (2002). Quantification of human mitochondrial DNA in a real time PCR. Forensic Sci Int *126*, 34-39.