# SUPPLEMENTAL MATERIAL

**This PDF file includes:**



**Supplementary methods:**

*Data and material sharing:* Cell and mouse lines generated in this work can be requested from the corresponding author and will be delivered through a material transfer agreement (MTA).

*Generation of heteroplasmic mice:* Heteroplasmic mice were generated by electro-fusing cytoplasts from conplastic BL/$6^{NZB}$ zygotes to recipient C57BL/6JOlaHsd (BL/$6^{C57}$) one-cell embryos; the fused cell pairs were then cultured overnight and transplanted as two-cell embryos into pseudo pregnant Hsd:ICR (CD-1®) females to complete development to term13. To our knowledge, there is no agreed consensus on the naming of heteroplasmic mouse strains. We propose the following nomenclature for heteroplasmic mouse strains: NUCLEAR GENOME-mtCYTOPLASMIC GENOME #1 + CYTOPLASMIC GENOME #2. Thus, the C57BL/6J-mtC57BL/6+NZB strain contains the C57BL/6J nuclear genome and the C57BL/6J and NZB cytoplasmic (mitochondrial) genomes. For simplicity, we abbreviate strain names as BL/6C57-NZB etc. Only offspring of the established heteroplasmic mice were used. To prevent accidental selection of private mutations and maintain a homogeneous nuclear background between BL/$6^{C57}$, BL/$6^{NZB}$ and BL/$6^{C57\text{-}NZB}$ strains, females were systematically outcrossed to BL/$6^{C57}$ males. Only females with an initial level of NZB heteroplasmy above 20%, measured in the tail, were used for colony maintenance.

*Molecular confirmation and quantification of heteroplasmy:* All mice were genotyped at postnatal day 21 using DNA extracted from the tail tip. Total genomic DNA was isolated from tails by REDExtract-N-Amp Tissue PCR (Sigma-Aldrich). To determine the heteroplasmy level in each tissue or cell culture, total genomic DNA was isolated using the DNeasy Blood &Tissue kit

(Qiagen). mtDNA in individual animals and tissues was genotyped by detection of the G4276A nucleotide polymorphism in C57 mtDNA. This polymorphism forms part of a BamH1 restriction site that is absent from NZB mtDNA. Total genomic DNA was PCR amplified with the 5Prime Master Mix (5Prime) kit using the following primers: 5'-AAGCTATCGGGCCCATACCCCG-3' (3862-3884) and 5'-GTTGAGTAGAGTGAGGGATGGG-3' (4503-4525). PCR conditions were as follows: 95ºC, 30s; 58ºC, 30s; 72ºC, 45s for 30 cycles. A 15 µl aliquot of the PCR product was digested with 20 units of BamH1 (New England Biolabs) at 37ºC for 2 hours. After agarose gel electrophoresis, DNA was visualized with a Gel Doc XR+System (Bio-Rad), and band intensity was quantified with Quantity One 1-D analysis software. The proportion of C57 mtDNA was calculated by summing the intensities of the 414bp and 250bp BamH1 fragments and dividing by the sum of the intensities of the undigested 664bp fragment and the 414bp and 250bp BamH1-digested fragments. To correct for heteroduplex formation, a standard curve was generated by mixing pure C57 and NZB mtDNAs. In addition, this technique was confirmed by solid-phase minisequencing, as previously described[48]. Mice with more than >30% heteroplasmy measured in the tail and in the specified tissues were used for RNAseq, proteomics, histology, and electron microscopy analyses.

*Heteroplasmy statistical analysis*: Heteroplasmy dynamics were modelled as $h(t) = 1 / (1 - (1-h_0)/h_0\ e^{-\beta t})$, where h is heteroplasmy, t is time after conception, $h_0$ is initial heteroplasmy at conception, and β is segregation rate. For each dataset, the values of $h_0$ for each sample and of β for each tissue were simultaneously inferred using a maximum likelihood approach implemented with Marcov Chain Monte Carlo (MCMC), assuming normally-distributed measurement noise and with confidence intervals derived using bootstrapping with the percentile method. To assess the influence of chemical and genetic perturbations, the model was expanded to allow β to be a function of time ($\beta_0$ before treatment and $\beta_1$ after treatment) or a function of a particular mouse type ($\beta_0$ for control mice and $\beta_1$ for treated mice). To assess the influence of time, β was modelled as $\beta_0$ before a threshold time τ and $\beta_1$ afterwards, and the value of τ was jointly inferred. In these cases, a significant difference between $\beta_0$ and $\beta_1$ implies a difference in segregation behavior between treated and control mice, or before and after the threshold time. Models were selected using the likelihood ratio test to distinguish the most supported model. When reporting a change in heteroplasmy h with respect to a reference value $h_0$ (for example, a previous measurement or a reference tissue), we used the transformation $\Delta h = \beta t = \log((h(h_0-1))/(h_0(h-1)))$, motivated by the above model as in[49].

*Survival Study:* Survival analysis was performed using the Kaplan–Meier method and a comparison of the survival curves was performed with the log-rank test log-rank (Mantel-Cox) and Gehan-Breslow-Wilcoxon test, which gives more weight to earlier time points. We considered both spontaneous deaths discovered in daily colony inspections and prescribed euthanasia as death events. The criteria for euthanasia were based on an independent assessment made by a veterinarian, and euthanasia deaths represented in the survival curves are exclusively those due to sacrifice of an animal whose condition was considered incompatible with continued survival.

*Micro-CT imaging:* Micro-CT was performed with a nanoScan PET/CT small-animal imaging system (Mediso Medical Imaging Systems, Budapest, Hungary). Whole-body helical scans were captured with a 142-micrometer spatial resolution at a voltage of 45 kVp. Curve progression in kyphosis was measured using a modification of the Cobb method. High-resolution CT images were processed using Osirix software (Pixmeo, Switzerland). The rotation angle was calculated using properly aligned reference lines drawn perpendicular to lines drawn along the fourth thoracic (T4) endplate and the inferior fourth lumbar (L4) end vertebra. The angle between these two lines was measured as the kyphotic angle. Tibia and femur length were measured on volume-rendering 3D

CT scans with Osirix imaging software. The measurements were calculated after tracing the line between the anatomical landmarks of these two bones using standard length ROI tools included in the software. Investigators were blinded to animal genotype and strain.

*PET-CT imaging:* In vivo positron emission tomography–computed tomography (PET-CT) imaging was performed with the nanoScan PET/CT small-animal imaging system (Mediso Medical Imaging Systems, Budapest, Hungary). List-mode PET data acquisition started 30 min after intravenous injection of ~15 MBq [$^{18}$F]-FDG into 20-week-old mice continued for 20 minutes PET acquisition. At the end of PET, microCT was performed for attenuation correction and to identify anatomical references. The resulting dynamic PET images were reconstructed in a 105x105 matrix (frame rates, 3 x 10 min, 1 x 30 min, 1 x 60 min) using a Tera-Tomo 3D iterative algorithm. Acquisition and reconstruction were performed with proprietary Nucline software (Mediso, Budapest, Hungary) and saved in Dicom format. The images were imported into OsiriX software (Pixmeo, Switzerland v.8.0.1) and were analysed quantitatively by comparing [$^{18}$F]-FDG uptake in manually segmented anatomical heart regions. Investigators were blinded to animal genotype and strain.

*Ex vivo biodistribution*: Biodistribution studies were performed 6 hours after intravenous injection of ~15 MBq [$^{18}$F]-FDG into 20-week-old mice in a Wizard 1470 Gamma Counter (Perkin Elmer). Animals were sacrificed in a CO2 chamber. Blood was extracted and animals were perfused with 8 mL of PBS. Tissues were extracted and counted for 1 min in the gamma counter. Decay correction was performed, and data represented as relative %ID/ g.

*Cardiac magnetic resonance imaging (MRI):* Animals were anaesthetized with isoflurane and were monitored for core body temperature, cardiac rhythm, and respiration rate using a specific monitoring system (SA Instruments Inc. New York, NY). *In vivo* cardiac images were acquired using an Agilent VNMRS DD2 7T MRI system (Agilent Varian, Palo Alto, USA) with an active shielded gradient 115/60 insert coil with 433 maximum gradient strengths and a four-channel receiver. A k-space segmented ECG-triggered short-axis cine gradient-echo sequence was used with a Helmet hundred-element surface coil (Extend MR, Silicon Valley, California). After shimming optimization, cardiac four-chamber and left two-chamber views were acquired and used to plan the short axis sequence. Mice were imaged with the following parameter settings: number of slices, 11-13; slice thickness, 0.8 mm; gap, 1 mm; data matrix size, 128x128; field of view, 30x30 mm$^2$; gating, ECG and respiratory triggered; cardiac phases, 20; averages, 4; effective TE, ~1.25 ms, minimum TR, 7 ms; flip angle, 15º. MRI images were exported to Dicom format and analyzed with Segment software (Mediso AB v.1.9R3819). The short-axis dataset and cine modus short-axis view were analyzed quantitatively by manual detection of RV and LV endocardial borders at end diastole and end systole without exclusion of papillary muscles and trabeculae, thus yielding left and right end-diastolic volumes, left and right end-systolic volumes, left and right ventricular ejection fractions (LV-EF, RV-EF), and cardiac output (CO).

$^{31}$P/$^{1}$H magnetic resonance spectroscopy ($^{31}$P/$^{1}$H-MRS). Spectroscopy examinations were performed *in vivo* with a 7T preclinical system (Agilent Varian, Palo Alto, USA) equipped with a DD2 console and an active shielded 115/60 gradient insert coil with 433 mT/m maximum strength. Double-tuned circular transmit/receive coils were (20 mm) placed over soleus muscle, cardiac muscle, and liver (Rapid Biomedical GmBH, Rimpar Germany). Proton spectra were acquired with 128 transients with 2048 complex points at a spectral bandwidth of 10 kHz and a repetition time of 1.2 ms. Phosphorus spectra were acquired with 1000 transients with 8192 complex points at a spectral bandwidth of 7 kHz and a repetition time of 800 ms. Spectra were acquired with adiabatic radiofrequency pulses to improve sensitivity and minimize spectral distortions with an Ernst flip

angle. Chemical shifts were expressed relative to the water signal (4.7-4.8 ppm) in $^{1}$H-MRS and relative to phosphocreatine (0 ppm) in $^{31}$P-MRS. Nuclear magnetic resonance spectra signals were determined quantitatively by integration after automatic or manual baseline correction, with fitting of each spectrum peak (after phase and baseline correction) to a Lorentzian function using the Mestrenova program (Mestrelab Research, Santiago de Compostela, Spain; released 2015-02-04 version:10.0.1-14719) on a Macintosh computer. An exponential line broadening (3 Hz for proton) was applied before Fourier transformation. Investigators were blinded to animal genotype and strain.

*Ventricular pressure measurements:* Mice were anesthetized using inhaled sevoflurane through a custom-built mask. Once the animal was anesthetized, skin was dissected at the right anterior portion of the neck and the dissection by planes was performed until finding the superior cava vein. Superior cava vein was carefully dissected and then a small hole was made with a needle to introduce a mikrotip catheter (SPR- 671NR Millar Catheter, ADInstruments, Dunedin, New Zealand). The mikrotip catheter was advanced through the cava vein until reaching the right ventricle, which was attested by observation of the typical square waveform characteristic. Then the right ventricle pressure was continuously acquired during free breathing. The mikrotip catheter was attached to a bridge amplifier (Bridge Amp FE221, ADInstruments, Dunedin, New Zealand) and data were collected using a PowerLab system (PL3504/P, ADInstruments, Dunedin, New Zealand) at 1,000 Hz sample rate. A second order Savitzky- Golay filter with a 35 samples window was applied to smooth the data. The systolic pressure was defined as the maximum pressure during each cardiac cycle and measured beat by beat during a 15 s period at which a stable measurement was observed. The average of this period length was used as representative of each animal and studied situation. Data were processed using custom-written routines programmed in Labchart v7.3.8 (ADInstruments, Dunedin, New Zealand) and Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA).

*Transthoracic echocardiography:* Two-dimensional and M-mode echocardiography longitudinal studies were performed in 3 months intervals over the entire mouse lifespan. Males (n=15-20) were anesthetized (1.5-2% isoflurane in a mixture with oxygen) and the analysis was carried out using a Vevo770 system (Vevo 2100, Visualsonics Inc., Canada) equipped with a 30-MHz linear transducer probe. To avoid night-day circadian variations, echocardiography analyses were performed in the morning. Before echocardiography, the animal was depilated and warmed to maintain body temperature. The heart was imaged in the 2D parasternal long- and short-axis projections with guided M-mode recordings at the midventricular level in both views. Images were recorded and transferred to a computer for posterior blinded analysis using the Vevo 2100 Workstation software. LV end-diastolic diameter (LVEDD), LV-systolic diameter (LVESD), end-diastolic LV anterior wall thickness (LVAW), and LV posterior wall thickness (LVPW) were measured from images obtained by M-mode echocardiography. LV fractional shortening (FS) and aortic cardiac output (CO) was calculated.

*Biochemical analysis:* Blood was collected in EDTA tubes. For plasma isolation, samples were centrifuged at 2000 x g for 15 minutes at 4ºC. The biochemical profile was assessed with a Dimension RxL Max® automated analyzer. All reagents used for the tests were purchased from Siemens.

*Histology:* Tissue samples were fixed in 4% paraformaldehyde, processed, and embedded in paraffin. Sections (10 μm) were prepared and mounted on coverslips for staining with hematoxylin and eosin or Masson's trichrome. Eyes were dissected and fixed in Davidson's fixative (2% formaldehyde, 35% ethanol, 10% glacial acetic acid), embedded in paraffin, sectioned with a

microtome (10 µm), and mounted on glass slides. The retina was divided into superior and inferior sections. Tissue images were captured using a NanoZoomer 2.0RS digital slide scanner (Hamamatsu) and analyzed with ImageJ v.1.6.0 Software (National Institutes of Health, USA). Collagen fibre and F4/80$^{+}$ areas were calculated as the ratio of positively stained area to total area for heart sections or of positively stained area to total perimeter for tissue sections. The ratios were calculated using Fiji (http://fiji.sc/Fiji) [50] and ImageJ v.2.0.0 [51].

*Evans Blue assay:* Evans Blue dye (Sigma) was used as an indicator of cardiomyocyte membrane permeability and viability. 100mg/kg of Evans Blue compound was injected intravenously into 20-week-old mice. After 24 hours, mice were sacrificed, and tissue slides were prepared and processed for counterstaining. Cardiac sections were quantified using ImageJ.

*Transmission electron microscopy (TEM)*: Samples of cardiac muscle (LV), liver, lung, eye and skeletal muscle (gastrocnemius) were obtained from 12-week-old and 80-week-old mice. Tissues were dissected immediately after sacrifice of animals by cervical dislocation. Samples were fixed for 4-5h at 4ºC in 2.5% glutaraldehyde (Sigma-Aldrich) and 4% formaldehyde (Electron Microscopy Sciences) in 0.1M HEPES. After buffer washes, samples were post-fixed for 1h at room temperature in a 1:1 mix of 2% osmium tetroxide (Electron Microscopy Sciences) and 3% aqueous potassium ferrocyanide (Sigma-Aldrich). Samples were rinsed in distilled $H_20$. Tissues were dehydrated through a graded acetone series and embedded in Spurr's Low Viscosity embedding mixture (Electron Microscopy Sciences). Ultrathin sections (60 nm) were then mounted on copper grids and stained with lead citrate. Samples were examined with a JEOL 10-10 electron microscope and images were analyzed with ImageJ v.1.6.0.

*Determination of superoxide in cardiac tissue:* Cardiac tissue was dissected, mounted in OCT embedding compound, and cryostat-sectioned (12 µm). Slides were defrosted and incubated for 45 minutes at 37ºC with 5 mM dihydroethidium (DHE). Samples were washed twice with PBS and incubated for 5 minutes at RT with DAPI (0.1 mg/ml) to counterstain nuclei. Slides were washed twice with PBS and water and mounted in ProLong® Gold Antifade Reagent. Slides were examined with a Leica SPE-Upright confocal microscope fitted with a HCX PL APO CS 40x 1.25 oil objective, and images were acquired with LAS-AF V 2.3.6 acquisition software (Leica Microsystems). DHE and DAPI fluorescence was quantified using ImageJ v. 2.0.0. Mean DHE fluorescence was expressed relative to mean DAPI fluorescence.

*Isolation and functional assessment of mitochondria:* Mitochondria were isolated from mice as described [52]. ATP synthesis was assessed in isolated cardiac mitochondria (15–25 µg mitochondrial protein) using glutamate plus malate or succinate as substrates in the presence of ADP in a kinetic luminescence assay based on the luciferin/luciferase reaction [53].

*Generation of mouse adult fibroblast (MAF) cell lines:* Fibroblasts mice were isolated from mouse ear and immortalized by transfection with pLOX-Ttag-iresTK (Addgene). Cell lines were grown in Dulbecco's modified Eagle's medium (DMEM; Gibco™) supplemented with 5% fetal bovine serum, 1% penicillin-streptomycin (Lonza) and 1 mM sodium pyruvate (Sigma). Where indicated, the carbon source was 5 mM glucose (Sigma), 5 mM galactose (Sigma), or Albumax™ lipid-rich BSA (1 mg/ml).

*Blue-native gel electrophoresis (BNE).* Purified mitochondria from MAFs and heart tissue were solubilized with 10% digitonin (4g/g) and separated on a 3%–13% gradient blue native gel. The gel gradient was prepared using a gradient former connected to a peristaltic pump. Proteins were electroblotted onto PVDF transfer membrane for 1 h at 100 V in transfer buffer (48 mM Tris, 39 mM glycine, 20 % EtOH). To study supercomplex assembly, membranes

were sequentially probed with antibodies to complex I (anti-NDUFA9, Abcam), complex III (anti-UQCRC2, ProteinTech) complex IV (anti-CO, Invitrogen), and complex II (anti-SDHA, ThermoFisher).

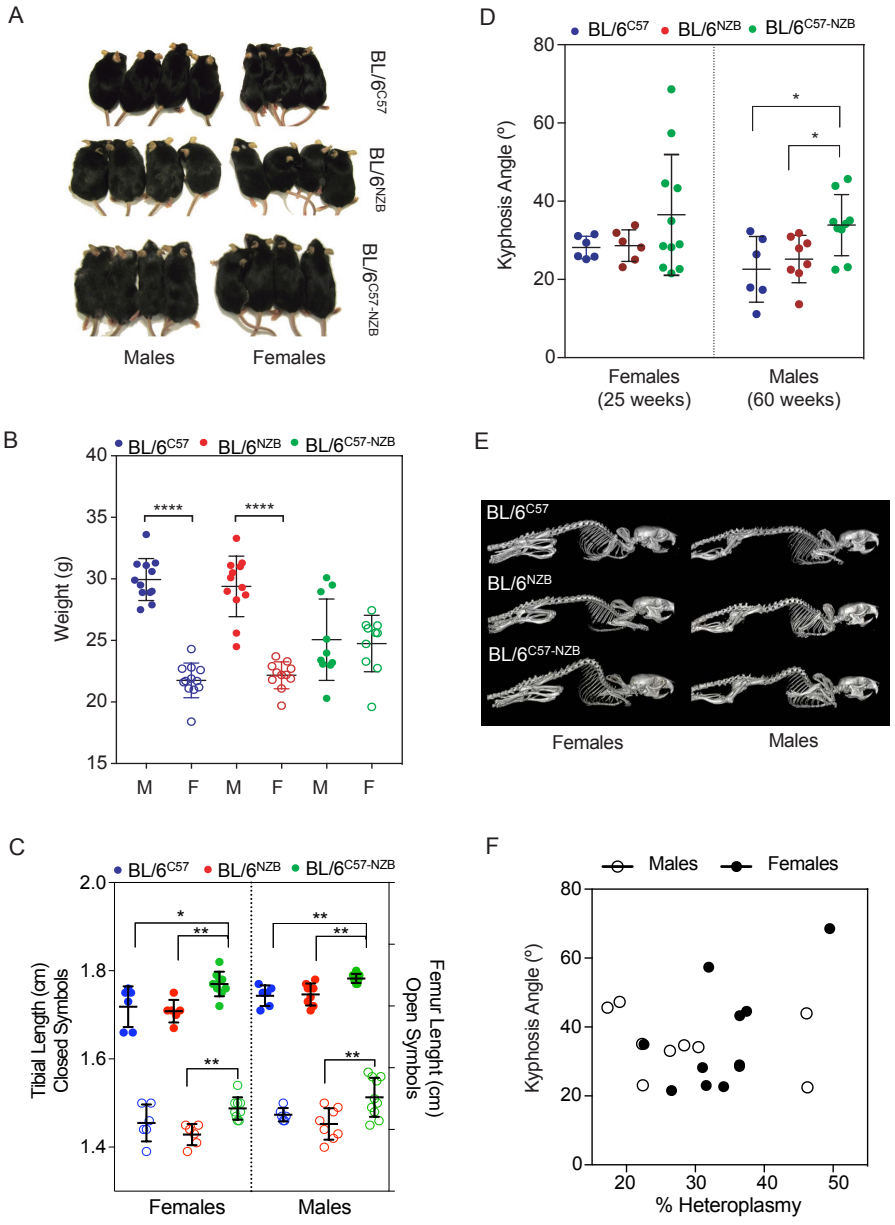
*SeaHorse$^{TM}$ analysis:* Oxygen consumption in mouse adult fibroblasts (MAFs) was measured using the XF96 MitoStress Test (Seahorse Bioscience). Oxygen consumption rates were normalized to cell number using CyQuant (Molecular Probes).

*Immunocytochemistry:* For in vitro immunofluorescence assays, cells were seeded on coverslips, fixed in 4% formaldehyde for 15 minutes at RT, and permeabilized in PBS containing 0.1% Triton X-100. Cells were then stained in blocking solution (5% BSA in PBS) with the F-actin marker phalloidin (LifeTechnologies) and TOM20 primary antibody (Santa Cruz Biotech) followed by Alexa Fluor secondary antibody (LifeTechnologies). Coverslips were mounted in ProLong® Gold Antifade Reagent, and images were acquired using a Leica TCS SP5 confocal system and analisedusing LAS AF software (Leica Microsystems, Germany) and ImageJ 1.48v.

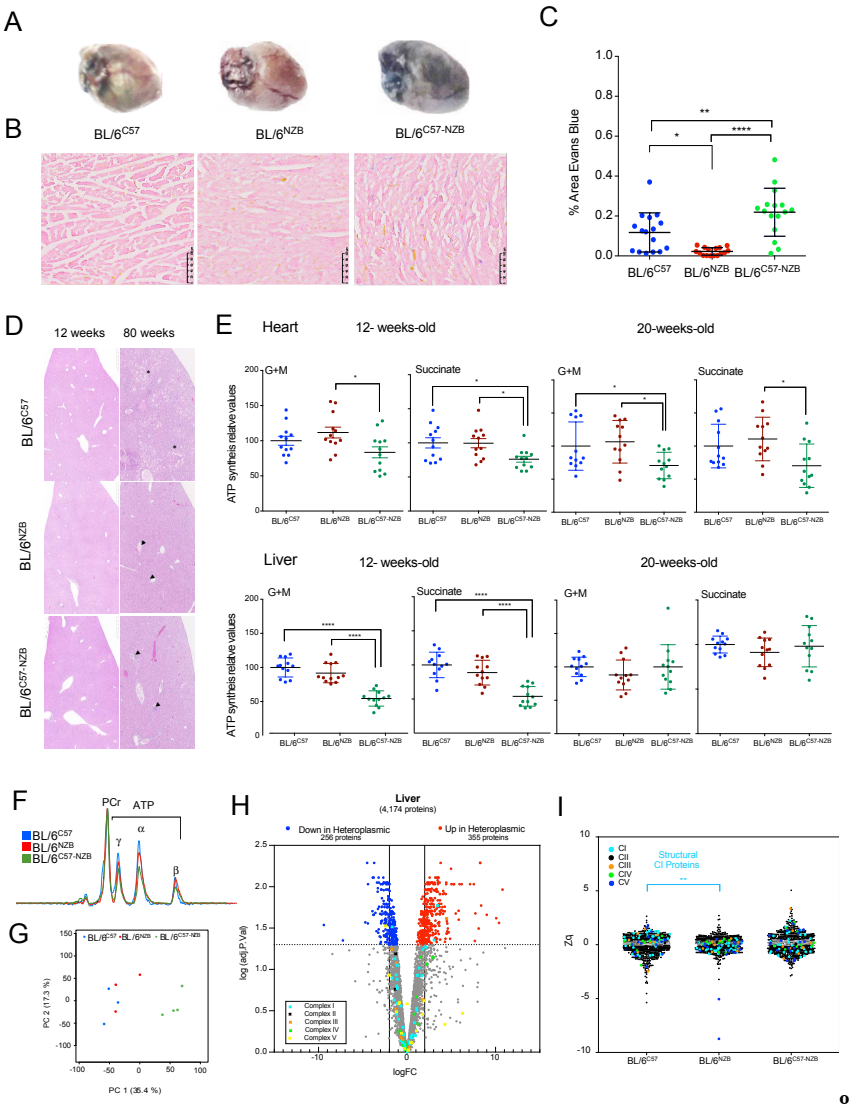
*Proteomics analysis:* Total heart, skeletal muscle (gastrocnemius), lung and liver protein extracts from 12-week-old BL/6$^{C57}$, BL/6$^{NZB}$, and BL/6$^{C57\text{-}NZB}$ male mice were digested with trypsin and labelled with TMT (10plex) according to the manufacturer's instructions (ThermoFisher), separated into 7 fractions using the high pH reversed-phase peptide fractionation kit (Thermo Fisher Scientific) and dried-down before MS analysis. The resulting labelled peptides were injected onto a C-18 reversed phase (RP) 2 cm trap column for desalting, and then onto an analytical nano-column (75 µm I.D. and 50 cm, Acclaim PepMap, Thermo Fisher, San José, CA, USA) to be analyzed on a continuous acetonitrile gradient consisting of 8-31% B for 240 min and 50-90% B for 1 min (B=0.5% formic acid in acetonitrile). Peptides were eluted from the RP nano-column at ~200 nL/min to an emitter nanospray needle for real-time ionization and peptide fragmentation in a Q-Exactive HF mass spectrometer (Thermo Fisher). Mass spectra were acquired in a data-dependent manner, with an automatic switch between MS and MS/MS using the top 15 method. Protein identification was performed by closed search (i.e. narrow-precursor window search) on SEQUEST HT using Proteome Discoverer (version 2.1, Thermo Fisher Scientific), against a Mouse Swissprot database (Uniprot release 57.3 May 2009; 26.885 entries). For database searching, parameters were selected as follows: trypsin digestion with 2 maximum missed cleavage sites; precursor mass tolerance, 800 ppm; fragment mass tolerance, 0.03 amu. We allowed variable methionine oxidation and fixed cysteine carbamidomethylation, and all lysines and N-terminals were modified in + 229.2634 amu. MS/MS spectra were also searched against decoy databases constructed from the same target databases. Peptides were identified from MS/MS data using the probability ratio method [54]. False discovery rates (FDR) of peptide identifications were calculated using the refined method, with an additional filter for precursor mass tolerance of 15 ppm [55], and with 1%FDR used as the criterion for peptide identification. Each peptide was assigned only to the top protein proposed by the Proteome Discoverer algorithm. Modified peptides were identified by open search on Comet-PTM in the same conditions as the closed search, except that precursor tolerance was set to 500 Da and no missed cleavages were allowed [23]. The FDR (<1%) of PTM identification was controlled at the global, local and peak levels by SHIFTS [23]. Quantitative information was extracted from MS/MS spectra of TMT-labeled peptides using Proteome Discoverer. Modified peptides and protein quantification was performed using the Generic Integration Algorithm [56,57] based on the WSPP model [58] with some modifications [23]. Briefly, proteins were quantified from the non-modified peptides, and the standardized log2-ratios of the modified peptides with respect to these protein values ($z_{pq}$) was calculated. Thus, modified peptides deviating more than expected from the non-modified peptides of the same protein were

detected (FDRpq<0.05%) regardless of the protein abundance change [23]. Conservation of amino acids showing PTM alterations in complexes I, III, IV, and V was calculated with Consurf [59].

*Transcriptomics analysis:* 12-week-old BL/6$^{C57\text{-}NZB}$, BL/6$^{C57}$, and BL/6$^{NZB}$ males were euthanized by cervical dislocation, and heart and liver tissues were dissected and immediately frozen in liquid nitrogen. Total RNA was extracted with TRIzol reagent and then purified on RNeasy spin columns (Qiagen). RNA integrity (RNA Integrity Score ≥ 7.9) and quantity were determined with an Agilent 2100 Bioanalyzer. Heart RNA libraries were produced using the TrueSeq RNASeq kit (Illumina), and liver RNA libraries were produced with the TrueSeq and NEB Next kits. RNAs were sequenced in a HiSeq2500 Illumina Sequencer. Reads were pre-processed with Cutadapt 1.7.1 to remove the Illumina adaptors. RSEM v1.2.20 and Bowtie 0.12.4 were used to map the reads against the GRCm38 reference and ensembl genebuild v84 (www.ensembl.org) modified by removing pseudogenes and adding the genes of the two mitochondrial references (C57 and NZB) as isoforms. Expected counts at the gene level were normalized (TMM method) and transformed using the VOOM function from the "limma" R package. Conplastic and wild type samples were retrieved from GSE56933and reanalyzed. In the new experiment (GSE93920), one sample from conplastic hearts and one sample from WT hearts were also included to allow for batch correction. Differential expression analysis was tested using a linear model with sequencing batch as a random variable to correct for batch effects and library kit biases. Significant differential expression was considered at an adjusted p-value of < 0.05. Liver samples did not produce any significant difference at this level and a threshold of < 0.2 was used to obtain a list of genes which produces a profile to cluster samples. For representation purposes in PCA, normalized data were corrected using ComBat as implemented in the "sva" R package. Functional analysis of sets of differentially expressed genes across conditions were analyzed using DAVID and Ingenuity Pathways Analysis Suite (IPA). Terms with an adjusted p-value or FDR < 0.2 were filtered and from those the top 50 from each analysis were selected for clustering. We implemented a novel function to compare IPA and GO terms based on the number of genes they share. This approach allows us to cluster functional terms and to identify the most relevant terms without the loss of information that results from choosing a given level in the hierarchy of ontologies. Plotting the average expression of the genes in each functional category and in each condition, gives us also information about the coordination of the expression profiles. All significant terms at an adjusted p-value or FDR < 0.05 are indicated by an asterisk. The rest of clustered non-significant terms are maintained to help to contextualize the significant ones. Benjamini-Hochberg adjusted p-values were used for both differential gene expression and functional analysis.

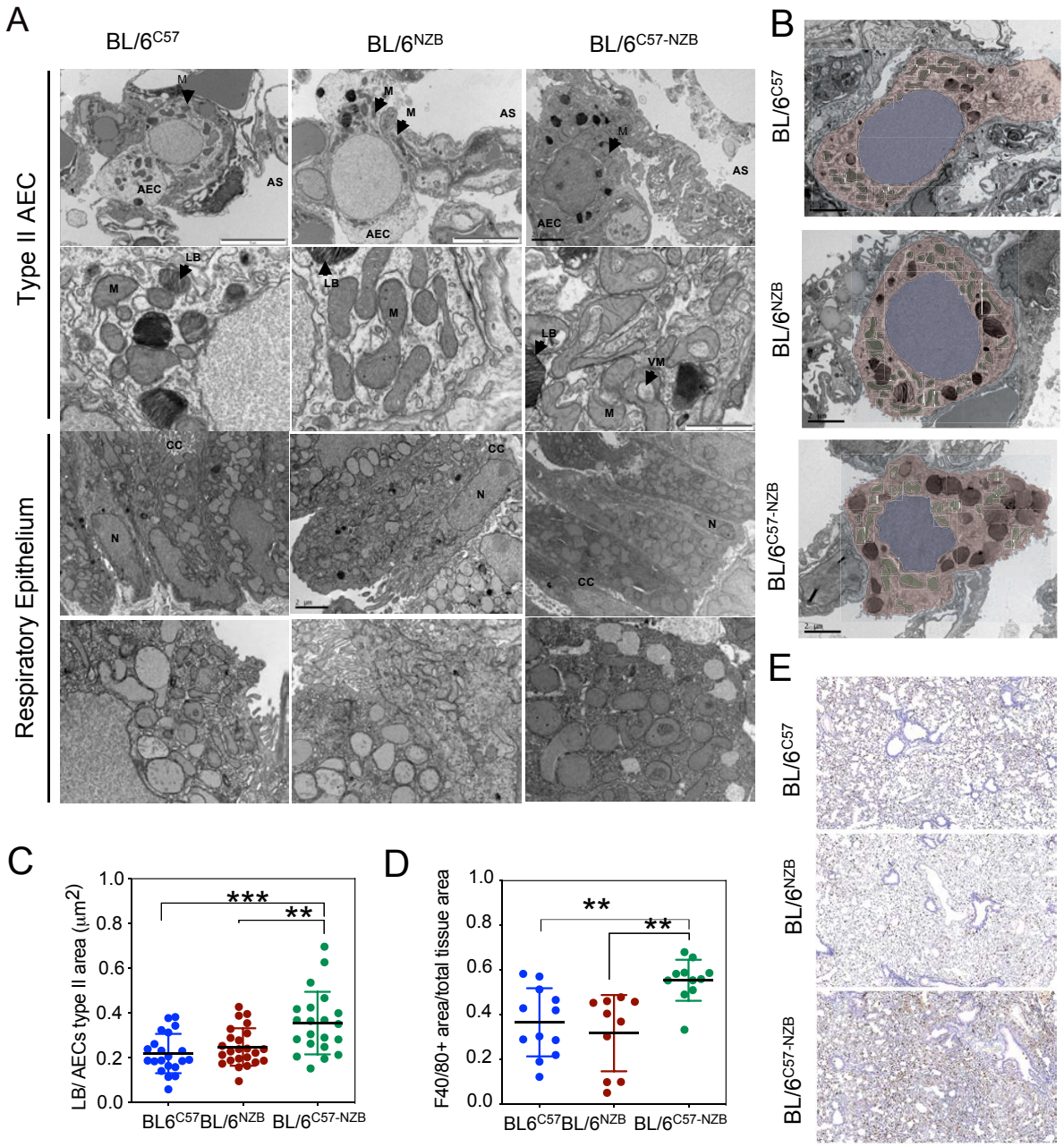



**Supplementary Fig. 1.- DNPH mice demonstrate altered sexual dimorphism and premature aging. A.** Picture illustrating the unusual increase in size of heteroplasmic females. **B.** Body weight differences between male and female in homoplasmic and heteroplasmic animals (≈100 days old; M: males, F: females). Each dot corresponds to a different individual. Asterisks marks indicate significant differences assessed by two-way ANOVA analysis, ****$p < 0.0001$. Data are given as mean ± s.d. **C,** MicroCT assessment of tibia and femur length differences between homoplasmic and heteroplasmic males and females. Each dot corresponds to a different individual. *$p < 0.05$, **$p < 0.01$, one-way ANOVA analysis. **D-E.** Measured kyphosis angle (D) and representative microCT scans (E) in 25 weeks-old-females and 60 weeks-old-males homoplasmic and heteroplasmic mice. *$p < 0.05$, one-way ANOVA, non-parametric Kruskal-Wallis test (n=6 homoplasmic mice per sex; n= 11 heteroplasmic females and n= 9 heteroplasmic males). **F.** Evaluation of the correlation of kyphosis angle and heteroplasmy level in heteroplasmic males and females measured in panel D. When adjusted by sex, no significant correlation was observed in males and females by Spearman's Rank Correlation Coefficient.

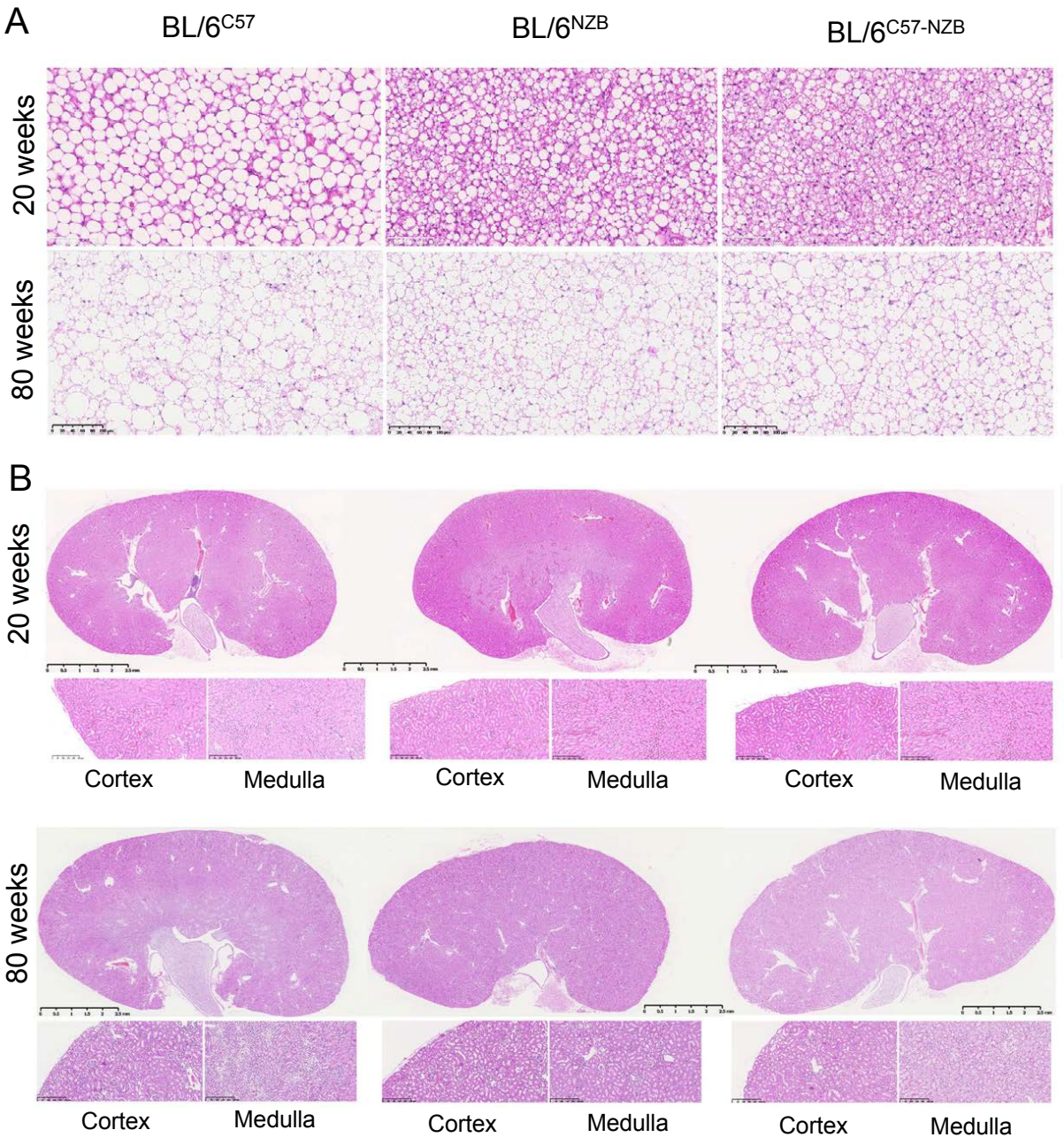



**Supplementary Figure 2.- Analysis of the heart and liver function. A.** Representative Evans blue stained hearts from 20-week-old homplasmic and heteroplasmic mice. Blue intensity indicates necrosis. **B.** Representative histological sections of the hearts in panel A, showing the highly dispersed distribution of damaged cardiomyocytes in heteroplasmic cardiac tissue. **C.** Quantification of Evans blue staining in histological sections: n= 3 homoplasmic hearts and n=5 heteroplasmic hearts; each dot represents a different tissue section. **D.** Hematoxylin and eosin staining in liver sections from 12-week-old and 80-week-old homoplasmic and heteroplasmic (DNPH) mice. In aged mice, asterisks indicate areas with microvesicular and macrovesicular steatosis and arrowheads mark perivascular infiltration of inflammatory cells. Representative images from n=4 mice per strain and age. **E.** ATP synthesis capacity of isolated mitochondria from heart and liver of 12- and 20-week-old mice of the indicated strain fed either with glutamate plus malate (G+M) or succinate (n=6 individuals per genotype and age). Each dot corresponds to an individual measurement (technical duplicates). **F.** Representative cardiac $^{31}$P-MRS stacked spectra in 20-week-old mice (n=5 mice per strain). **G-I.** Liver proteomic analysis in 12-week-old mice (n= 3 homoplasmicmice and n=4 heteroplasmic mice). (G) Principal component analysis. (H) Volcano plot illustrating proteins with significantly differential abundance between homoplasmic and heteroplasmic liver. The log (adjusted p-value) is plotted against the LogFC (fold difference between homoplasmic and heteroplasmic mice); 355 proteins are significantly over-represented and 256 are significantly under-represented. X-axis, vertical lines denote ±2-fold change. Y-axis, the horizontal line denotes the significance threshold ($p < 0.05$ before

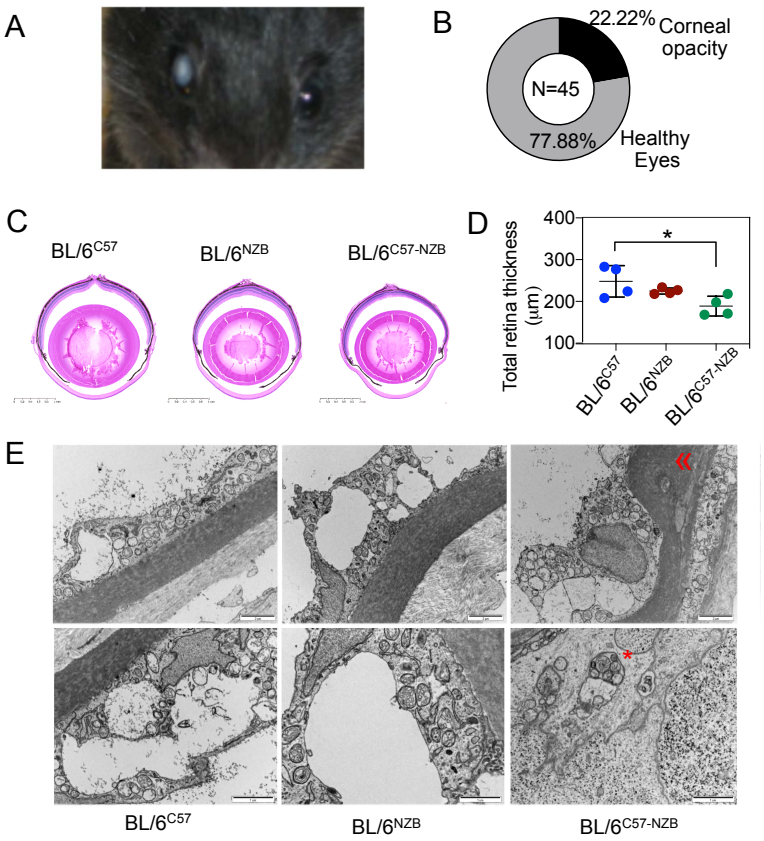
logarithmic transformation). (I) Mitoproteome analysis, highlighting structural OXPHOS components. Data in panels (C), (E) and (I) are presented as mean +/- s.d.; *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$, one-way ANOVA test.



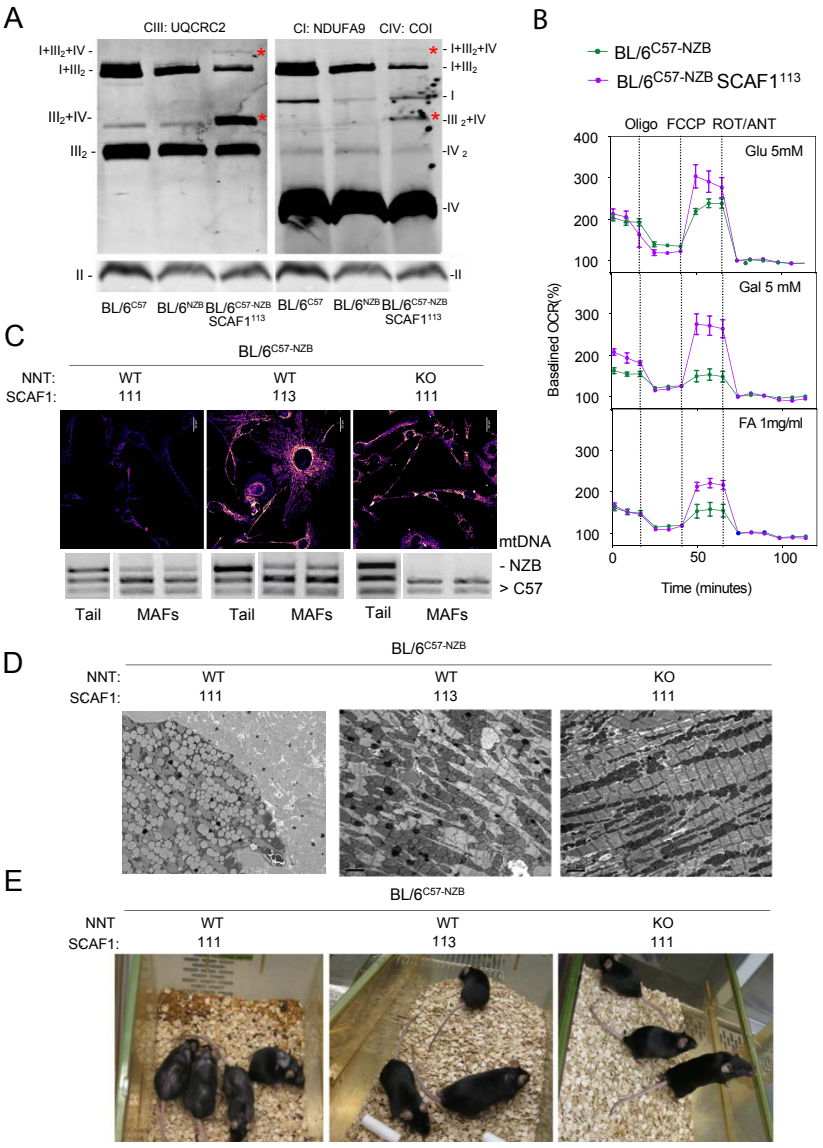

**Supplementary Figure 3.- Lung tissue morphology and ultrastructure. A.** Representative TEM images of type II lung alveolar epithelial cells (AEC) in the respiratory epithelium of 12-week-old mice of the indicated strains (n=3-4 mice per genotype). AS, alveolar space; M, mitochondria; VM, vacuolated mitochondria; LB, lamellar bodies; N, nucleus; CC, clara cell. **B.** Representative images of individual type II AECs showing between-strain differences in lamellar body composition and ultrastructure. **C.** Area density of lamellar bodies in type II AECs. $^{**}p < 0.01$, $^{***}p < 0.001$, one-way ANOVA; n=21-24 individual type II AECs per strain. **D-E.** F4/80 staining on lung sections of 12-week-old mice. Relative density of F4/80 expression in lung tissue sections (D) and representative images of the F4/80 staining in lung tissue from the indicated strains (E). $^{**}p < 0.01$, one-way ANOVA test (n=5 mice per strain).




**Supplementary Figure 4.- Histological analysis of brown adipose tissue and kidney. A-B.** Representative hematoxylin and eosin staining on sections of brown adipose tissue (A) and kidney (B) from 20- and 80-week-old homoplasmic and heteroplasmic mice (n=4 per genotype and age).




**Supplementary Figure 5.- Eye structure and function in DNPH mice. A.** Representative picture of corneal opacity in heteroplasmic mice (BL/6$^{C57\text{-}NZB}$). **B.** Percentage of heteroplasmic mice presenting corneal dysfunction during life (n = 45 heteroplasmic mice). **C-D.** Histological analysis of eyes from homoplasmic mice (BL/6$^{C57}$, BL/6$^{NZB}$) and heteroplasmic mice (BL/6$^{C57\text{-}NZB}$). Representative hematoxylin and eosin staining (C) and quantification of total retina thickness in the three strains (D). Each dot corresponds to a different animal. *$p < 0.05$, one-way ANOVA test. Scale bar, 1 mm. **E.** TEM analysis of corneal endothelial cells and Descemet's membrane (n = 3 homoplasmic and n= 5 heteroplasmic mice). Arrowhead: thickening of the Descemet's membrane, an indicator of corneal impairment. Asterisk: autolysosome with mitochondrial content in the heteroplasmic corneal endothelium.




**Supplementary Figure 6.- Impact of nuclear gene modifications on the DNPH phenotype. A.** Blue native electrophoresis showing the assembly and superassembly status of the respiratory complexes in BL/6$^{C57}$, BL/6$^{NZB}$ and BL/6$^{C57\text{-}NZB}$:SCAF1$^{113}$ heteroplasmic MAFs using 5 mM galactose as the carbon source for 48 hours. Left: Complex III (UQCRC2). Right: Complex I (NDUFA9) and Complex IV (COI). Bottom: Complex II (FP70). Red asterisks indicate the SCAF1(113)-dependent supercomplexes. **B.** Respiratory performance of control heteroplasmic MAFs and heteroplasmic MAFs expressing functional SCAF1(113) with different carbon sources: 5 mM glucose, 5 mM galactose, or Albumax$^{TM}$ lipid-rich BSA (1 mg/ml). Seahorse profile data are presented as mean ± s.e.m. **C.** Top: mitochondrial network distribution within control and genetically modified heteroplasmic MAFs using confocal imaging analysis of the outer mitochondrial membrane marker TOM20. Bottom: PCR-RFLP gel showing differences in heteroplasmy levels between i) tail tissue from the individuals used for MAFs generation and ii) immortalized MAFs at passage 8. Purple color indicates low mitochondria density; yellow, high mitochondria density. **D.** Representative TEM images at 5x magnification showing specific features observed in hearts of heteroplasmic mice with nuclear genetic modifications (n= 5 12-weeks-old mice per genotype). **E.** Representative pictures of aged females of the indicated strains. BL/6$^{C57\text{-}NZB}$, 349 days old; BL/6$^{C57\text{-}NZB}$:SCAF1(113), 492 days old; BL/6$^{C57\text{-}NZB}$:NNT$^{KO}$, 578 days old.

- Captions for supplementary Data Tables 1 to 2:

**Table S1.-** Sets of differentially expressed genes in heart across conditions analyzed using DAVID Bioinformatics Resources.
**Table S2.-** Sets of differentially expressed genes in heart across conditions analyzed using IPA suite.

- Captions for supplementary Movies 1 to 2:

**Supplementary Movies 1 and 2.-** Rear-leg paralysis in elderly heteroplasmic mice.