




OPEN

Examining the effect of mitochondrial DNA variants on blood pressure in two Finnish cohorts

Jaakko Laaksonen¹ , Pashupati P. Mishra¹, Ilkka Seppälä¹, Leo-Pekka Lyytikäinen¹, Emma Raitoharju¹, Nina Mononen¹, Maija Lepistö², Henrikki Almusa², Pekka Ellonen², Nina Hutri-Kähönen³, Markus Juonala^{4,5,6}, Olli Raitakari^{7,8,9}, Mika Kähönen¹⁰, Jukka T. Salonen^{11,12} & Terho Lehtimäki¹

High blood pressure (BP) is a major risk factor for many noncommunicable diseases. The effect of mitochondrial DNA single-nucleotide polymorphisms (mtSNPs) on BP is less known than that of nuclear SNPs. We investigated the mitochondrial genetic determinants of systolic, diastolic, and mean arterial BP. MtSNPs were determined from peripheral blood by sequencing or with genome-wide association study SNP arrays in two independent Finnish cohorts, the Young Finns Study and the Finnish Cardiovascular Study, respectively. In total, over 4200 individuals were included. The effects of individual common mtSNPs, with an additional focus on sex-specificity, and aggregates of rare mtSNPs grouped by mitochondrial genes were evaluated by meta-analysis of linear regression and a sequence kernel association test, respectively. We accounted for the predicted pathogenicity of the rare variants within protein-encoding and the tRNA regions. In the meta-analysis of 87 common mtSNPs, we did not observe significant associations with any of the BP traits. Sex-specific and rare-variant analyses did not pinpoint any significant associations either. Our results are in agreement with several previous studies suggesting that mtDNA variation does not have a significant role in the regulation of BP. Future studies might need to reconsider the mechanisms thought to link mtDNA with hypertension.

High blood pressure (BP) is a major public health problem. It is an established risk factor for many noncommunicable diseases, such as cardiovascular disease, renal dysfunction, and dementia, and causes over 9 million premature deaths globally per year¹. It is generally accepted that the regulation of BP is a multi-factorial trait, involving lifestyle, environmental, and genetic factors. To date, the majority of the genetic variants have been identified in studies of the nuclear genome^{2–4}, and a limited number of studies have explicitly investigated the associations with variation in the mitochondrial genome. This gap in genetic knowledge is of a particular interest because, in addition to cellular energy production, mitochondria modulate, for example, the intracellular dynamics of nitric oxide, reactive oxygen species, and Ca²⁺, which, in turn, control endothelial function in blood vessels⁵.

¹Department of Clinical Chemistry, Fimlab Laboratories and Finnish Cardiovascular Research Center Tampere, Faculty of Medicine and Health Technology, Tampere University, Arvo Ylpön katu 34, PO Box 100, 33014 Tampere, Finland. ²Institute for Molecular Medicine (FIMM), University of Helsinki, Helsinki, Finland. ³Department of Paediatrics, Tampere University Hospital and Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. ⁴Department of Medicine, University of Turku, Turku, Finland. ⁵Division of Medicine, Turku University Hospital, Turku, Finland. ⁶Murdoch Children's Research Institute, Parkville, VIC, Australia. ⁷Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland. ⁸Research Centre for Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland. ⁹Department of Clinical Physiology and Nuclear Medicine, University of Turku and Turku University Hospital, Turku, Finland. ¹⁰Department of Clinical Physiology, Tampere University Hospital and Finnish Cardiovascular Research Center Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. ¹¹Department of Public Health, Faculty of Medicine, University of Helsinki, Helsinki, Finland. ¹²MAS-Metabolic Analytical Services Oy, Helsinki, Finland. ✉email: jaakko.h.laaksonen@tuni.fi

Mitochondrial DNA (mtDNA) is a maternally inherited, double-stranded circular molecule containing 16,569 nucleotides that encode 22 transfer RNAs (tRNAs), two ribosomal RNAs (rRNAs), and 13 protein subunits of the four oxidative phosphorylation (OXPHOS) complexes⁶. The majority of the proteins functioning in the mitochondria are, however, encoded in nuclear DNA. There are multiple copies of mtDNA within each cell, and the mtDNA mutation rate is significantly higher than that of nuclear DNA⁷. Given the high mutation rate and the fact that both common and rare variants may influence the disease phenotypes⁸, the most appropriate method to study mtDNA is through sequencing. One possible mechanism through which mutations in the mtDNA contribute to BP variation is the oxidative stress due to an increased production of reactive oxygen species, which in turn causes cardiovascular and renal damage^{9–11}. In addition to mutations in the mtDNA, the mitochondrial dysfunction may also be the consequence of mutations in the nuclear-encoded mitochondrial genes, or result from environmental and lifestyle factors.

Several studies with varying ethnic groups have demonstrated that a mutational hot spot for mitochondrial single-nucleotide polymorphisms (mtSNPs) associated with hypertension is in the tRNA-encoding genes^{10,12,13}. However, many of these studies have been conducted on related individuals, and it is likely that most of these hypertension-associated variants are inherited and rare on a population level. While the role of naturally occurring mtDNA variation is still incompletely understood¹⁴, some evidence for the association of mtSNPs with BP exists from well-established cohorts. In the Framingham Heart Study, one mtSNP in OXPHOS complex IV correlated with systolic blood pressure (SBP)¹⁵. In a recent two-cohort sequencing study of older North American adults, both common and rare mtSNPs in the tRNA region were associated with variation in SBP in white participants¹⁶. Negative results have also been reported: 64 tagging mtSNPs that efficiently capture common mtDNA variation in the European population were not associated with BP in a study consisting of over 2000 individuals¹⁷, and a sequencing study conducted in South African population which utilized the MutPred pathogenicity prediction scores did not find a significant role for mtDNA variation in association with blood pressure levels¹⁸. A lexical tree analysis with phylogenetically related mtDNA variants in European population identified significant relationships with some common diseases, e. g. multiple sclerosis, but not with hypertension¹⁹.

The goal of the present study was to investigate the effect of mtDNA variants on SBP, diastolic blood pressure (DBP), and mean arterial pressure (MAP) among participants of two independent Finnish cohorts. We sought to both discover new and replicate the previously reported mitochondrial genetic determinants of BP. We studied both the effects of common single mtSNPs and the pooled effects of rare variants across seven mtDNA regions.

Materials and methods

Study participants. The Young Finns Study (YFS, <http://youngfinnsstudy.utu.fi>) is a Finnish longitudinal population study on the evolution of cardiovascular risk factors from childhood to adulthood²⁰. We used the phenotype data from the follow-up in 2011 which included 2060 participants in total. The blood samples for mtDNA sequencing were obtained during the 2007 follow-up, when 2204 participants were examined. The study was approved by the ethical committee of the Hospital District of Southwest Finland (ETMK:68/1801/2017), and the study protocol of each study phase corresponded with the proposal by the World Health Organization.

The Finnish Cardiovascular Study (FINCAVAS) participant pool consists of patients recruited during 2001–2007 who underwent exercise stress tests at Tampere University Hospital²¹. A total of 4,068 participants completed a technically successful exercise test. The main indications for the exercise test were a suspicion of coronary heart disease (frequency 46%), the evaluation of work capacity (26%), testing for vulnerability to arrhythmia during exercise (25%), and assessing the adequacy of coronary heart disease treatment (13%); some patients had more than one indication. The study protocol was approved by the Ethical Committee of the Pirkanmaa Hospital District, Finland (R00153).

All participants in both study cohorts gave their written informed consent, and the studies were conducted in accordance with the Declaration of Helsinki.

Blood pressure and other clinical measurements. In the YFS, BP was determined as the average of three measurements taken at two-minute intervals in a sitting position from the right arm brachial artery with a random zero sphygmomanometer. Korotkoff's first phase was used as the sign of SBP and the fifth phase as the sign of DBP. In the FINCAVAS, the patients lay in the supine position for 10 min, after which BP was measured once by an experienced nurse using a brachial cuff according to the Korotkoff's method.

In both cohorts, the observed BPs were adjusted for antihypertensive medication usage. The medications were self-reported by the study participants, the duration of treatment was not known, and adherence was not assessed. Adjusted SBP was calculated by increasing the recorded measure by 8, 14, and 20 mmHg for 1, 2, and ≥ 3 medication classes taken, respectively. DBP measurements were adjusted similarly by increasing the recorded measure by 4, 10, and 16 mmHg for 1, 2, and ≥ 3 medication classes taken, respectively²². This adjustment method has proven to work with both simulated and real-life data²³, and it maximizes the genetic and shared environmental variance components while minimizing individual-specific components²². FINCAVAS participants displayed clearly more antihypertensive medication usage (31%, 24%, and 16% on 1, 2, and ≥ 3 medication classes taken, respectively) than YFS participants (7%, 3%, and 0.3%, respectively). The mean adjustment was 10/6 and 13/9 mmHg for treated YFS and FINCAVAS participants, respectively. Adjusted MAP was calculated from the adjusted SBP and DBP values as $\text{MAP} = \text{DBP} + 0.333 \times (\text{SBP} - \text{DBP})$.

Height and weight were measured, and body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared.

MtDNA sequencing and quality control in the YFS. Genomic DNA sample ($n=1807$) concentrations were measured from whole-blood samples with the Qubit BR dsDNA kit (Life Technologies). MtDNA

was amplified from the genomic DNA using the REPLI-g mtDNA kit (Qiagen) in a 50 μ l reaction volume. The primer sites in the REPLI-g kit have been previously validated²⁴. After the amplification mtDNA samples were processed into Illumina sequencing-compatible libraries with the Nextera DNA sample preparation kit (Illumina). The mtDNA concentrations were measured with Qubit dsDNA for Nextera tagmentation reaction. The reaction volume in the Nextera tagmentation and amplification steps was 20 μ l, and after both steps the libraries were purified with an EdgeBio Performa V3 96-Well Short Plate (Edge BioSystems). After the amplification, the libraries were first incubated with 4 μ l of EdgeBio SOPE resin and then purified with EdgeBio Performa plates. After purification, 48 samples with different index tags were pooled together (2 μ l each) in each pool and concentrated with DNA Clean & Concentrator-5 (Zymo Research). The final volume of the concentrated pool was 15 μ l. The sequencing-ready libraries were quantitated with an Agilent 2100 Bioanalyzer High Sensitivity kit (Agilent). The libraries were sequenced with the Illumina HiSeq system. All samples ($n = 1658$) that achieved any mean bait coverage were included in the further processing steps.

Paired-end FASTQ files were aligned with the revised Cambridge Reference Sequence (rCRS)⁶ using the sequence aligner BWA-MEM v. 0.7.17²⁵. Reads mapped to the rCRS were sorted according to the start position and written into a BAM file using SAMtools v. 1.8²⁶. Homo- and heteroplasmic variants were then detected with Mutserve v. 1.2.1, a stand-alone version of the web tool mtDNA-server²⁷. We applied the default thresholds for mapping, base, and alignment quality scores in the Phred scale: 20, 20, and 30, respectively. Mutations in the mtDNA can affect either all or a varying proportion of the mtDNA molecules, and the terms used for these phenomena are homoplasmy and heteroplasmy, respectively⁷. The minimum heteroplasmy level was set to 0.05—we defined sites with a heteroplasmy level below this threshold as homoplasmic wild-type alleles and sites with a heteroplasmy level above 0.95 as homoplasmic variants. Mutserve applies a Bayesian model to the detection of homoplasmic variants, and in order to call a wild-type allele, we required a minimum sequencing coverage of five on both the forward and reverse strands. Mutserve identified mtSNPs in 1,365 different nucleotide positions from 1657 samples. Mean coverages per sample and per mtSNP were 497 and 525, respectively. With this coverage, the 0.05 heteroplasmy detection level is quite reliable²⁸. Genotypes for heteroplasmic variants overlapping with any mtDNA-like sequence in the nucleus (NUMTs) were set to missing. The list of NUMTs insertions was based on the work by Dayama et al.²⁹ We required each sequenced sample to have an overall mean coverage of ≥ 50 , and 1,434 samples had a mean coverage above this threshold. Then, samples that did not have both phenotype and mtDNA data available were excluded, after which 1,150 samples remained for further analysis. After this, mean coverages per sample and per mtSNP were 563 and 568, respectively. Of the 1,365 identified common and rare mtSNPs, 249 had an allele frequency of ≥ 0.01 . Heteroplasmy level between 0.05 and 0.95 in at least 1% of the samples was observed only for one mtSNP, m.16192C>T in the hypervariable region. This mtSNP was not genotyped in the FINCAVAS arrays and hence not included in the meta-analysis. In general, heteroplasmy rate was low in most sites, for 99% of the sites the number of heteroplasmic samples was three or less. This finding is in line with a previous study which found the incidence of heteroplasmy to be higher in tissues with high metabolic activity³⁰.

MtDNA genotyping and quality control in the FINCAVAS. Genomic DNA was extracted from peripheral blood leukocytes by using the QIAamp DNA Blood Minikit and automated biorobot M48 extraction (Qiagen). We applied the Illumina Cardio-MetaboChip and HumanCoreExome-12 v1.1 SNP arrays for genotyping mtSNPs. Genotyping was completed for 2,824 and 1,032 samples using the Cardio-MetaboChip and HumanCoreExome arrays, respectively. Genotypes were called using Illumina's GenomeStudio GenCall algorithm. Samples with call rate of < 0.95 , excess heterozygosity, cryptic relatedness (π -hat > 0.2) and sex mismatch, as well as genetic outliers based on multi-dimensional scaling (MDS) plots, were removed. mtSNPs with a call rate of < 0.95 and a Hardy–Weinberg equilibrium p -value of $\leq 10^{-6}$ were also removed. Heterozygous genotypes possibly due to mitochondrial heteroplasmy or technical error were coded as missing. Homozygosity at a genotyped mtSNP indicates genotype frequency close to zero or one. After quality control, 53 mtSNPs from 2273 samples genotyped with the Cardio-MetaboChip array and 146 mtSNPs from 926 samples genotyped with the HumanCoreExome array were available. For four individuals that were genotyped with both arrays, the genotypes for the 34 overlapping mtSNPs were set as missing in the Cardio-MetaboChip array. After including only samples with both phenotype and mtDNA data available, 2193 and 923 samples from the Cardio-MetaboChip and HumanCoreExome arrays, respectively, were available for association analyses. The total number of individuals was 3112 since four samples were genotyped with both arrays. Of the variants genotyped with these arrays, 36 Cardio-MetaboChip mtSNPs and 67 HumanCoreExome mtSNPs had an allele frequency of ≥ 0.01 .

Statistical analyses. *Association analyses of common variants.* In order to investigate the associations of SBP, DBP, and MAP with mtDNA variants, BP levels were modeled as a linear function of the presence (coded as 1) or absence (coded as 0) of the variant allele using the lm function in R. Heteroplasmic genotypes were set to missing. In order to achieve normality of the BP distributions, we applied the rescaled inverse normal transformation, i.e. multiplied the inverse normal transformed BP values by the standard deviation of the original trait values in each cohort. This strategy makes the distributions normal and controls the type I error, and restores the original scale of measurement and thus enhances the power of meta-analysis³¹. Inverse normal transformation also effectively deals with phenotypic outliers³². Age, sex, and BMI were added as covariates to the linear regression models and the p -values were calculated using a standard F test with one degree of freedom. Association analyses were performed separately for the three data sets, and a random-effect meta-analysis was then performed using GWAMA software³³. Variants with an allele frequency of ≥ 0.01 were included. The total number of mtSNPs included in the meta-analysis was 87, of which 22 were present in YFS and in both FINCAVAS data sets. The number of mtSNPs present in YFS and only in the FINCAVAS participants genotyped with

the HumanCoreExome array was 52, and 13 mtSNPs were present in YFS and only in participants genotyped with the Cardio-MetaboChip array. Using Matrix Spectral Decomposition (matSpDlite)³⁴ with the method of Li and Ji³⁵, we determined that 45 of the 87 mtSNPs represented an estimate of the number of independent genetic effects for mtDNA. This resulted in a Bonferroni-corrected significance level of 0.0011 (i.e. 0.05/45). In the correction for multiple testing, we did not account for the testing of three BP traits, since they were correlated (Pearson's correlation coefficients between adjusted SBP, DBP, and MAP values 0.78–0.96 in the YFS and 0.67–0.92 in FINCAVAS).

Study power. With a sample size of ~2000, meaning that an mtSNP with an allele frequency of ≥ 0.01 was present in the YFS and the smaller FINCAVAS data set, our single-variant analysis had ~80% power to detect an mtSNP that explains 2.5% of the variance in BP. When an mtSNP with an allele frequency of ≥ 0.01 was present in the YFS and both FINCAVAS data sets, our analysis had ~95% and ~80% power to detect mtSNPs explaining 2.5% and 1% of the variance, respectively³⁶.

Sex-specific analyses. The power of sex-combined analysis (i.e. males and females analyzed together) is reduced when heterogeneity is present in the allelic effects between the sexes. We examined the possible heterogeneity by applying the same linear model as described above to males and females separately, and by applying sex-differentiated meta-analysis in GWAMA^{33,37}. In total, 66 mtSNPs with an allele frequency of ≥ 0.01 in both sexes were included in the meta-analysis. A sex-differentiated *p*-value below the individual *p*-values for males and females is indicative of a significant association with both sexes. We also tested for sex-specific heterogeneity, which is equivalent to a formal test of interaction with sex. A significant heterogeneity *p*-value would suggest that there is a difference in the effect sizes between the sexes. By using the same Matrix Spectral Decomposition as above, we now determined that 33 of the 66 mtSNPs represented an estimate of the number of independent genetic effects for mtDNA. To account for the two sexes tested, the significance level was now defined as $0.05/33/2 = 7.6 \times 10^{-4}$.

Rare-variant analyses. Standard methods used to test for common variant associations are underpowered for detecting associations with rare variants³⁸. The power of haplogroup-based analysis may also be insufficient if potential causative variants at a same mtDNA site are scattered across divergent haplogroups. The most common approaches for the analysis of rare-variant associations are the burden test and the sequence kernel association test (SKAT)^{38,39}. The former collapses the rare variants within a specified region, assuming that all variants are either deleterious or protective. The latter aggregates and tests the collective effects of rare variants within a region without assuming similar directionality or effect size for each variant. Therefore, SKAT is superior to burden test for analyzing regions where both risk and protective variants as well as noncausal variants may be present^{39,40} and it has been successfully applied for both sequenced and genotyped mtDNA data^{16,41,42}.

We used Mitomap⁴³ to cluster the variants into seven regions, including each of the four OXPHOS complexes (I, III, IV, and V), all rRNAs combined, all tRNAs combined, and control region and non-coding regions combined. Homoplasmic alleles were coded as 0 or 2, corresponding to the reference and variant allele, respectively. Heteroplasmic genotypes were introduced in a dosage matrix, similarly to imputed genotypes. The dosage was calculated as twice the heteroplasmy rate. We employed a SKAT meta-analysis with minor allele frequencies (MAFs) of ≤ 0.01 (T1) and ≤ 0.05 (T5) for SBP, DBP, and MAP, using the seqMeta package in R, with the default beta weights, and with age, sex, and BMI as covariates. Bonferroni-corrected statistical significance was defined as 0.0071 (i.e. 0.05/7).

In order to assess the functional relevance of non-synonymous variants, we predicted the pathogenicity of the identified variants with MutPred^{44,45} and MitoTIP⁴⁶ pathogenicity scores. The MutPred algorithm assigns each variant in the protein-encoding mtDNA regions a pathogenicity score between 0 and 1. Variants with a MutPred score > 0.5 are potentially “harmful”, and variants with a score > 0.75 should be considered a high confidence “harmful”. MitoTIP predicts the pathogenicity of the variants in the tRNA regions, and the prediction scores have been interpreted within quartiles as “likely benign”, “possibly benign”, “possibly pathogenic” or “likely pathogenic”. We leveraged SKAT meta-analysis for the BP levels, similarly as described above, but including only variants with a MutPred score > 0.5 or a MitoTIP classification “possibly” or “likely pathogenic”.

Control for population stratification. The use of mtDNA principal components (PCs) as covariates has been demonstrated to be a robust method to adjust for population stratification in genetic association studies. In addition, the use of mitochondrial PCs effectively removes false-positive associations but does not cause a loss of power in detecting true associations^{47,48}. Logistic PC analysis was performed on all homoplasmic genotypes passing quality control and with a MAF of ≥ 0.01 using the logisticPCA package⁴⁹. For each data set, we selected the number of mitochondrial PCs for the single variant analyses so that the median chi-squared-based genomic inflation factor (λ_{GC}) was as close to one as possible. For SKAT, we selected the same number of PCs as were used in the sex-combined common variant analyses. Values of $\lambda_{GC} < 1.05$ are generally considered benign⁵⁰. For any result data sets with $\lambda_{GC} > 1.05$, genomic control was applied by multiplying the standard errors of regression coefficients by the square root of the inflation factor of the respective study.

Results

On average, participants in the FINCAVAS were older and had higher BMI and BP levels than participants in the YFS (Table 1). We evaluated the associations of 87 common mtDNA variants with SBP, DBP, and MAP using random-effect meta-analysis in these two Finnish cohorts, with sample sizes of up to 4,262. When both sexes were analyzed together, we did not observe any statistically significant associations. The top three associations

	YFS n = 1150	FINCAVAS n = 3112
Age, years	42.2 (5.0)	56.3 (13.0)
Women, n (%)	661 (57.5)	1,214 (39.0)
Body mass index, kg/m ²	26.5 (5.00)	27.5 (4.5)
Unadjusted SBP, mmHg	118.7 (13.9)	136.7 (18.9)
Unadjusted DBP, mmHg	74.8 (10.4)	80.0 (9.6)
Unadjusted MAP, mmHg	89.5 (10.8)	98.9 (11.3)
Medication-adjusted SBP, mmHg	119.7 (14.7)	145.6 (21.1)
Medication-adjusted DBP, mmHg	75.4 (10.8)	86.2 (11.2)
Medication-adjusted MAP, mmHg	90.2 (11.4)	106.0 (13.3)

Table 1. Baseline characteristics of the YFS and FINCAVAS cohorts. Values are means (SD). YFS Young Finns Study, FINCAVAS Finnish Cardiovascular Study, SBP systolic blood pressure, DBP diastolic blood pressure, MAP mean arterial pressure.

mtSNP	Trait	Beta (SE)	<i>p</i>	MAF	n
m.1243T>C	SBP	− 4.22 (1.79)	0.019	0.024	4,219
m.15257G>A	MAP	4.91 (2.20)	0.025	0.013	2,071
m.11674C>T	SBP	− 4.26 (1.91)	0.026	0.029	3,342

Table 2. Three most significant associations in the sex-combined meta-analysis. mtSNP mitochondrial single-nucleotide polymorphism, Trait the trait used for a specific test, Beta (SE) beta coefficient and the corresponding standard error, *p* unadjusted *p*-value from meta-analysis, MAF minor allele frequency, *n* sample size contributing in a particular mtSNP meta-analysis.

are shown in Table 2. All meta-analysis results and the quantile–quantile plots from the cohort-level analyses are provided in the Supplemental material, in Table S1 and Fig. S1.

There was no evidence for differences between the sexes in the mitochondrial genetic control of BP levels. All results from the sex-differentiated and heterogeneity meta-analysis as well as the quantile–quantile plots from the cohort-level analyses are provided in the Supplemental material, in Table S2 and Figs. S2–S3.

Finally, we conducted SKAT meta-analyses on all rare (T1 test) and low-frequency (T5 test) variants. We also employed SKAT taking into account the predicted pathogenicity of the variants. None of the analyses yielded significant associations with the BP traits over of the tested mtDNA regions. All results from SKAT meta-analysis are provided in the Supplemental material, Tables S3 and S4.

Discussion

We conducted one of the largest studies to date investigating the possible mitochondrial genetic determinants of BP. We did not find any associations that survived after correction for multiple testing.

We could not replicate the results where the common mtSNPs m.3197T>C and m.15924A>G were associated with higher SBP and MAP, respectively, in white North American individuals¹⁶. In the current study, the former mtSNP was sequenced or genotyped in over 2000 and the latter in over 4200 individuals, which casts doubt over the role of these two mtSNPs in the variation of BP. Possible confounding effects due to different genetic ancestry should not have had an effect on the associations, since both the previous and the current study included mitochondrial PCs as regression covariates^{47,48}. However, it should be noted that, since the North American individuals were significantly older than the Finns in our cohorts, it is possible that the regulative role of these two mtSNPs is activated only in later life.

The aforementioned North American study¹⁶ identified significant pooled effects on SBP across variants in the tRNA regions in white participants. Our analysis of rare and low-frequency variants did not yield any significant associations, not even when we accounted for the predicted pathogenicity of the variants within protein-encoding and the tRNA regions. We did not account for the functional relevance of the variants in the rRNA regions because to the best of our knowledge, no tools exist for annotating pathogenicity for variants in those regions.

Another mtSNP previously found to be associated with SBP is m.5913G>A in the Framingham Heart Study¹⁵. In the current study, this mtSNP was only sequenced in the YFS population with a MAF of < 0.01, while it was not genotyped in FINCAVAS, which left us unable to study the effect of this mtSNP in single-variant analysis. While the sample size of the current study is, to the authors' knowledge, one of the largest used in investigating mtDNA associations with BP, our inability to replicate earlier findings from well-characterized cohorts underlines the need for mtDNA association studies with significantly larger sample sizes from multiple cohorts. For example, a recent study including ~ 170,000 individuals from 45 cohorts reported associations with seven metabolic outcomes, but these did not include BP⁴². These large consortium studies can however be extremely costly, and another basis for future research could be achieved by introducing more homogeneous study groups with less confounding effects. We should also be ready to accept to the null hypothesis that mtDNA variants do not

contribute to BP variation on a significant level as it is implied by several previous studies^{17–19}. Another hypothesis to be tested is that instead of being causal of BP variation, the mtDNA variants would impact the hypertension complications or alter the course of the disease⁵¹.

The strength of the current study was that the mtSNPs in the YFS were obtained through sequencing, which allowed us to study the effects of both common and rare variants. Another strength was the large range of blood pressure variation in our two independent cohorts, and the method we used to adjust for antihypertensive treatment effects has proven to work across a wide variety of clinical scenarios^{22,23}. Some limitations should also be acknowledged—the mtSNPs in FINCAVAS were genotyped, resulting in a relatively small number of mtSNPs included in the meta-analyses. The majority of the rare mtSNPs were identified only in the YFS population which decreased the power of the SKAT meta-analysis. We also applied a stringent threshold to define homoplasmic variants and wild-type alleles, but relaxing the threshold to e. g. 10% and 90% would have increased the number of homoplasmic alleles only marginally. Another limitation is that we only investigated substitutions, whereas mtDNA deletions and insertions were not identified. In addition, BP measurements were performed only once for the FINCAVAS participants. Ambulatory BP monitoring could greatly increase the robustness of the future studies.

In conclusion, we found no support for the hypothesis that common or rare mtSNPs play a significant role in the regulation of BP. While studies with larger sample sizes might show different results, we should also be open to the idea that the non-significant outcomes reported by the current and previous studies are in fact correct, and that future studies concerning this topic need to reconsider the mechanisms thought to link mtDNA with hypertension.

Data availability

The data sets generated and analyzed during the current study comprise health-related participant data, and their use is therefore restricted under the regulations concerning professional secrecy (Act on the Openness of Government Activities, 612/1999) and sensitive personal data (Personal Data Act, 523/1999, implementing the EU data protection directive 95/46/EC). Due to these legal constraints, the Ethics Committee of the Hospital District of Southwest Finland has stated in 2016 that individual-level data cannot be stored in public repositories or otherwise made publicly available. Data are, however, available from the authors upon a reasonable request.

Received: 11 October 2020; Accepted: 10 December 2020

Published online: 12 January 2021

References

- Lim, S. S. *et al.* A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2224–2260 (2012).
- Newton-Cheh, C. *et al.* Genome-wide association study identifies eight loci associated with blood pressure. *Nat. Genet.* **41**, 666–676 (2009).
- Ehret, G. B. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103–109 (2011).
- Liu, C. *et al.* Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. *Nat. Genet.* **48**, 1162–1170 (2016).
- Davidson, S. M. & Duchon, M. R. Endothelial mitochondria: Contributing to vascular function and disease. *Circ. Res.* **100**, 1128–1141 (2007).
- Andrews, R. M. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**, 147 (1999).
- Chinnery, P. F. & Hudson, G. Mitochondrial genetics. *Br. Med. Bull.* **106**, 135–159 (2013).
- Schork, N. J., Murray, S. S., Frazer, K. A. & Topol, E. J. Common vs. rare allele hypotheses for complex diseases. *Curr. Opin. Genet. Dev.* **19**, 212–219 (2009).
- Dikalov, S. I. & Ungvari, Z. Role of mitochondrial oxidative stress in hypertension. *Am. J. Physiol. Heart Circ.* **305**, 1417–1427 (2013).
- Ding, Y., Xia, B., Yu, J., Leng, J. & Huang, J. Mitochondrial DNA mutations and essential hypertension (review). *Int. J. Mol. Med.* **32**, 768–774 (2013).
- Lahera, V., De Las Heras, N., López-Farré, A., Manucha, W. & Ferder, L. Role of mitochondrial dysfunction in hypertension and obesity. *Curr. Hypertens. Rep.* **19**, 11 (2017).
- Wilson, F. H. *et al.* A cluster of metabolic defects caused by mutation in a mitochondrial DNA. *Science* **306**, 1190–1194 (2004).
- Wang, S. *et al.* Maternally inherited essential hypertension is associated with the novel 4263A>G mutation in the mitochondrial tRNA Ile gene in a Large Han Chinese Family. *Circ. Res.* **108**, 862–870 (2011).
- Bray, A. W. & Ballinger, S. W. Mitochondrial DNA mutations and cardiovascular disease. *Curr. Opin. Cardiol.* **32**, 267–274 (2017).
- Liu, C. *et al.* Association of genetic variation in the mitochondrial genome with blood pressure and metabolic traits. *Hypertension* **60**, 949–956 (2012).
- Buford, T. W. *et al.* Mitochondrial DNA sequence variants associated with blood pressure among 2 cohorts of older adults. *J. Am. Heart Assoc.* **7**, e010009. <https://doi.org/10.1161/JAHA.118.010009> (2018).
- Saxena, R. *et al.* Comprehensive association testing of common mitochondrial DNA variation in metabolic disease. *Am. J. Hum. Genet.* **79**, 54–61 (2006).
- Venter, M., Malan, L., van Dyk, E., Elson, J. L. & van der Westhuizen, F. H. Using MutPred derived mtDNA load scores to evaluate mtDNA variation in hypertension and diabetes in a two-population cohort: The SABPA study. *J. Genet. Genomics* **44**, 139–149 (2017).
- Hudson, G., Gomez-Duran, A., Wilson, I. J. & Chinnery, P. F. Recent mitochondrial DNA mutations increase the risk of developing common late-onset human diseases. *PLoS Genet.* **10**, e1004369. <https://doi.org/10.1371/journal.pgen.1004369> (2014).
- Juonala, M., Viikari, J. S. & Raitakari, O. T. Main findings from the prospective cardiovascular risk in Young Finns Study. *Curr. Opin. Lipidol.* **24**, 57–64 (2013).
- Nieminen, T. *et al.* The Finnish Cardiovascular Study (FINCAVAS): Characterising patients with high risk of cardiovascular morbidity and mortality. *BMC Cardiovasc. Disord.* **6**, 9. <https://doi.org/10.1186/1471-2261-6-9> (2006).
- Cui, J. S., Hopper, J. L. & Harrap, S. B. Antihypertensive treatments obscure familial contributions to blood pressure variation. *Hypertension* **41**, 207–210 (2003).

23. Tobin, M. D., Sheehan, N. A., Scurrah, K. J. & Burton, P. R. Adjusting for treatment effects in studies of quantitative traits: Anti-hypertensive therapy and systolic blood pressure. *Stat. Med.* **24**, 2911–2935 (2005).
24. Laaksonen, J. *et al.* Discovery of mitochondrial DNA variants associated with genome-wide blood cell gene expression: A population-based mtDNA sequencing study. *Hum. Mol. Genet.* **28**, 1381–1391 (2019).
25. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
26. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
27. Weissensteiner, H. *et al.* mtDNA-Server: Next-generation sequencing data analysis of human mitochondrial DNA in the cloud. *Nucleic Acids Res.* **44**, W64–W69 (2016).
28. Weissensteiner, H. *et al.* Haplocheck: Phylogeny-based contamination detection in mitochondrial and whole-genome sequencing studies. *bioRxiv* [https://doi.org/10.1101/2020.05.06.080952v2\(2020\)](https://doi.org/10.1101/2020.05.06.080952v2(2020)) (2020).
29. Dayama, G., Emery, S. B., Kidd, J. M. & Mills, R. E. The genomic landscape of polymorphic human nuclear mitochondrial insertions. *Nucleic Acids Res.* **42**, 12640–12649 (2014).
30. Naue, J. *et al.* Evidence for frequent and tissue-specific sequence heteroplasmy in human mitochondrial DNA. *Mitochondrion* **20**, 82–94 (2015).
31. Tang, Z. Z. & Lin, D. Y. Meta-analysis for discovering rare-variant associations: Statistical methods and software programs. *Am. J. Hum. Genet.* **97**, 35–53 (2015).
32. Auer, P. L., Reiner, A. P. & Leal, S. M. The effect of phenotypic outliers and non-normality on rare-variant association testing. *Eur. J. Hum. Genet.* **24**, 1188–1194 (2016).
33. Mägi, R. & Morris, A. P. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinform.* **11**, 288. <https://doi.org/10.1186/1471-2105-11-288> (2010).
34. Nyholt, D. R. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.* **74**, 765–769 (2004).
35. Li, J. & Ji, L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* **95**, 221–227 (2005).
36. McRae, A. F., Byrne, E. M., Zhao, Z., Montgomery, G. W. & Visscher, P. M. Power and SNP tagging in whole mitochondrial genome association studies. *Genome Res.* **18**, 911–917 (2008).
37. Mägi, R., Lindgren, C. M. & Morris, A. P. Meta-analysis of sex-specific genome-wide association studies. *Genet. Epidemiol.* **34**, 846–853 (2010).
38. Li, B. & Leal, S. M. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am. J. Hum. Genet.* **83**, 311–321 (2008).
39. Wu, M. C. *et al.* Rare-variant association testing for sequencing data with the sequence Kernel association test. *Am. J. Hum. Genet.* **89**, 82–93 (2011).
40. Basu, S. & Pan, W. Comparison of statistical tests for disease association with rare variants. *Genet. Epidemiol.* **35**, 606–619 (2011).
41. Giorgi, E. E. *et al.* No association between the mitochondrial genome and prostate cancer risk: The multiethnic cohort. *Cancer Epidemiol. Biomarkers Prev.* **25**, 1001–1003 (2016).
42. Kraja, A. T. *et al.* Associations of mitochondrial and nuclear mitochondrial variants and genes with seven metabolic traits. *Am. J. Hum. Genet.* **104**, 112–138 (2019).
43. Lott, M. T. *et al.* mtDNA variation and analysis using mitomap and mitomaster. *Curr. Protoc. Bioinform.* **44**, 23–26 (2013).
44. Mort, M. *et al.* In silico functional profiling of human disease-associated and polymorphic amino acid substitutions. *Hum. Mutat.* **31**, 335–346 (2010).
45. Pereira, L., Soares, P., Radivojac, P., Li, B. & Samuels, D. C. Comparing phylogeny and the predicted pathogenicity of protein variations reveals equal purifying selection across the global human mtDNA diversity. *Am. J. Hum. Genet.* **88**, 433–439 (2011).
46. Sonney, S. *et al.* Predicting the pathogenicity of novel variants in mitochondrial tRNA with MitoTIP. *PLoS Comput. Biol.* **13**, e1005867. <https://doi.org/10.1371/journal.pcbi.1005867> (2017).
47. Biffi, A. *et al.* Principal-component analysis for assessment of population stratification in mitochondrial medical genetics. *Am. J. Hum. Genet.* **86**, 904–917 (2010).
48. Miller, B. *et al.* Comparing the utility of mitochondrial and nuclear DNA to adjust for genetic ancestry in association studies. *Cells* **8**, 306. <https://doi.org/10.3390/cells8040306> (2019).
49. Landgraf, A. J. & Lee, Y. Dimensionality reduction for binary data through the projection of natural parameters. *J. Multivar. Anal.* **180**, 104668. <https://doi.org/10.1016/j.jmva.2020.104668> (2020).
50. Price, A. L., Zaitlen, N. A., Reich, D. & Patterson, N. New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* **11**, 459–463 (2010).
51. Achilli, A. *et al.* Mitochondrial DNA backgrounds might modulate diabetes complications rather than T2DM as a whole. *PLoS ONE* **6**, e21029. <https://doi.org/10.1371/journal.pone.0021029> (2011).

Acknowledgements

This work was supported by personal Grants awarded to JL by the Juhani Aho Foundation, Laboratoriolääketieteen edistämissäätiö, the Aarne Koskelo Foundation, and the Finnish Medical Foundation. The Young Finns Study (YFS) has been financially supported by the Academy of Finland: Grants 322098, 286284, 134309 [Eye], 126925, 121584, 124282, 129378 [Salve], 117787 [Gendi], and 41071 [Skidi]; the Social Insurance Institution of Finland; Competitive State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals (Grant X51001); the Juho Vainio Foundation; the Paavo Nurmi Foundation; the Finnish Foundation for Cardiovascular Research; the Finnish Cultural Foundation; the Sigrid Juselius Foundation; the Tampere Tuberculosis Foundation; the Emil Aaltonen Foundation; the Yrjö Jahnsson Foundation; the Signe and Ane Gyllenberg Foundation; the Diabetes Research Foundation of Finnish Diabetes Association; the EU Horizon 2020 program (Grant 755320 for TAXINOMISIS; Grant 848146 for To_Aition); the European Research Council (Grant 742927 for MULTIEPIGEN project); the Tampere University Hospital Supporting Foundation; and the Finnish Society of Clinical Chemistry. The Finnish Cardiovascular Study (FINCAVAS) has been financially supported by the Competitive Research Funding of Tampere University Hospital (Grant 9M048 and 9N035), the Finnish Cultural Foundation, the Finnish Foundation for Cardiovascular Research, the Emil Aaltonen Foundation, the Tampere Tuberculosis Foundation, the EU Horizon 2020 program (Grant 755320 for TAXINOMISIS; Grant 848146 for To_Aition), and the Academy of Finland Grant 322098. In addition, we thank Jari Kaikkonen and Janne Asikainen for their technical assistance.

Author contributions

J.L. designed the study, conducted statistical analyses and wrote the main manuscript text. P.M., I.S. and L.P.L. contributed to the statistical analyses. E.R., N.M., M.L., H.A. and P.E. contributed to data collection. N.H.K., M.J.

and M.K. participated in cohort collection. O.R. handled funding and participated in cohort collection. J.S. contributed to data collection. T.L. supervised the study, handled funding, participated in cohort collection, in addition to reviewing and editing the manuscript. All authors have contributed to and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-020-79931-6>.

Correspondence and requests for materials should be addressed to J.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021