

SUPPLEMENTARY APPENDIX

Association of mitochondrial DNA copy number with cardiometabolic diseases in a large cross-sectional study of multiple ancestries

Xue Liu¹, Ryan J. Longchamps², Kerri Wiggins³, Laura M. Raffield⁴, Lawrence F. Bielak⁵, Wei Zhao⁵, Achilleas Pitsillides⁶, Thomas Blackwell⁷, Xiuqing Guo⁸, Nuzulul Kurniansyah⁹, Bharat Thyagarajan¹⁰, Nathan Pankratz¹¹, Stephen S. Rich¹², Kent D. Taylor⁸, Patricia A. Peyser⁵, Susan Heckbert³, Sudha Seshadri¹³, L Adrienne Cupples⁶, Eric Boerwinkle¹⁴, Megan L. Grove¹⁴, Nicholas Larson¹⁵, Jennifer A. Smith⁵, Ramachandran S Vasani^{16,17}, Tamar Sofer^{9,18}, Annette L Fitzpatrick¹⁹, Myriam Fornage²⁰, Jun Ding²¹, Adolfo Correa²², Gonçalo Abecasis⁷, Bruce M. Psaty^{3,23}, James G. Wilson²⁴, Daniel Levy^{16,25}, Jerome I. Rotter⁸, Joshua C. Bis³, †Claudia L Satizabal¹³, †Dan E. Arking², and †Chunyu Liu⁶

Representing the TOPMed mtDNA Working Group
NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium

¹Boston University School of Medicine, Department of Medicine, Boston, MA 20118, USA;

²McKusick-Nathans Institute, Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, MD 21205, USA;

³Cardiovascular Health Research Unit and Department of Epidemiology, University of Washington, Seattle, WA 98101, USA

⁴Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA;

⁵University of Michigan School of Public Health, Department of Epidemiology, Ann Arbor, MI 48109, USA;

⁶Department of Biostatistics, School of Public Health, Boston University, Boston, MA 02118, USA;

⁷TOPMed Informatics Research Center, University of Michigan, Ann Arbor, MI 48109, USA;

⁸The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA;

⁹Brigham and Women's Hospital, Boston, MA 02115, USA;

¹⁰Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN 55455, USA;

¹¹Department of Computational Pathology, University of Minnesota, Minneapolis, MN 55455

¹²Department of Public Health Services, Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA;

¹³University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA;

¹⁴Department of Epidemiology, Human Genetics and Environmental Sciences, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA;

¹⁵Division of Biomedical Statistics and Informatics, Mayo Clinic College of Medicine and Science, Rochester, MN 55905, USA

¹⁶Framingham Heart Study, NHLBI, Framingham, MA, USA.

¹⁷Sections of Preventive Medicine and Epidemiology, and Cardiovascular Medicine, Boston University School of Medicine, Boston, MA, USA.

¹⁸Harvard Medical School, Boston, MA 02115, USA;

¹⁹University of Washington School of Public Health, Department of Epidemiology, Seattle, WA 98101, USA;

²⁰Center for Human Genetics, University of Texas Health Science Center at Houston, Houston, TX 77030, USA;

²¹Laboratory of Genetics and Genomics, NIA/NIH, Baltimore, MD 21224

²²Department of Medicine, University of Mississippi Medical Center, Jackson, MS, 39216, USA;

²³Kaiser Permanente Washington Health Research Institute, Seattle, WA 98101, USA;

²⁴Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS

²⁵The Framingham Heart Study, Framingham, MA 01702, USA;

²⁶Population Sciences Branch National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA;

† These authors contribute equally to this work

Disclaimer: The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services.

Methods

Disclaimer: The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services.

Study participants:

This study included eight well-known population-based cohorts.

Atherosclerosis Risk in Communities study (ARIC) (n=2,971): The Atherosclerosis Risk in Communities Study (ARIC) is a prospective epidemiologic study conducted in four communities which are Forsyth County, NC; Jackson, MS; the northwest suburbs of Minneapolis, MN; and Washington County, MD.¹ Focusing on cardiovascular disease outcomes, event adjudication through 2017 consisted of expert committee review of death certificates, hospital records and telephone interviews. DNA for mtDNA-CN estimation used blood samples collected from different visits. Buffy coat was purified using the Genra Puregene Blood Kit (Qiagen). mtDNA-CN was available for 2,971 participants of European and African American ancestry with TOPMed WGS sequencing.

Cohort acknowledgement/support: The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I,

HHSN268201700003I, HHSN268201700004I and HHSN268201700005I), R01HL087641, R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Funding support for “Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium” was provided by the NIH through the American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419). Sequencing was carried out at the Baylor College of Medicine Human Genome Sequencing Center and supported by the National Human Genome Research Institute grants U54 HG003273 and UM1 HG008898. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. Whole genome sequencing (WGS) for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung and Blood Institute (NHLBI). WGS for “NHLBI TOPMed: Atherosclerosis Risk in Communities” (phs001211.v3.p2.c1) was performed at the Broad Institute for MIT and Harvard (3R01HL092577-06S1). WGS for “Trans-Omics for Precision Medicine (TOPMed) Whole Genome Sequencing (WGS) Project: ARIC” (phs001211.v3.p2.c1) was performed at the Baylor College of Medicine Human Genome Sequencing Center. Centralized read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1). Phenotype harmonization, data management, sample-identity QC, and general study coordination, were provided by the TOPMed Data Coordinating

Center (3R01HL-120393-02S1). We gratefully acknowledge the studies and participants who provided biological samples and data for TOPMed.

Coronary Artery Risk Development in Young Adults Study (CARDIA) (n=3,452): CARDIA is a prospective cohort study which was initiated in 1984 to investigate life-style and other factors that influence cardiovascular disease and their risk factors during young adulthood. The study recruited and examined 5,116 black and white women and men aged 18-30 years in four urban areas: Birmingham, Alabama; Chicago, Illinois; Minneapolis, Minnesota, and Oakland, California.² The initial examination included carefully standardized measurements of major risk factors as well as assessments of psychosocial, dietary, and exercise-related characteristics that might influence them, or that might be independent risk factors. mtDNA-CN was available for 3,452 participants with TOPMed WGS sequencing.

Cohort acknowledgement/support: Whole genome sequencing (WGS) for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung and Blood Institute (NHLBI). WGS for NHLBI TOPMed: CARDIA (phs001612) was performed at the Baylor Human Genome Sequencing Center (HHSN268201600033I). Centralized read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1). Phenotype harmonization, data management, sample-identity QC, and general study coordination, were provided by the TOPMed Data Coordinating Center (3R01HL-120393-02S1). We gratefully acknowledge the studies and participants who provided biological samples and data for TOPMed. The Coronary Artery

Risk Development in Young Adults Study (CARDIA) is supported by contracts HHSN268201300025C, HHSN268201300026C, HHSN268201300027C, HHSN268201300028C, HHSN268201300029C, and HHSN268200900041C from the National Heart, Lung, and Blood Institute (NHLBI), the Intramural Research Program of the National Institute on Aging (NIA), and an intra-agency agreement between NIA and NHLBI (AG0005)." as noted in dbGaP. The authors also wish to thank the staffs and participants of the CARDIA.

The Cardiovascular Health Study (CHS) (n=3,554): The Cardiovascular Health Study (CHS) is a population based, longitudinal, multicenter study of coronary heart disease and stroke in 5,888 elderly adults aged 65 years and older. The CHS originated in 1988 to recruit participants from four U.S. communities.³ The original cohort recruited 5201 participants and 687 predominately African-American participants were recruited at three of the four field centers in 1992. The first exam began in June 1989. A second comprehensive exam began 3 years after the first exam. A total of n=3,554 CHS participants (mean age 74 and 58% women) with WGS were included in this study.

Cohort acknowledgement/support: The Cardiovascular Health Study (CHS) (phs001368.v1.p1) was supported by contracts HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and grants U01HL080295 and U01HL130114 from the National Heart, Lung, and Blood Institute (NHLBI), with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by R01AG023629 from the National

Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. Sequencing was supported and conducted in collaboration with Baylor University (HHSN268201600033I, 3U54HG003273-12S2, HHSN268201500015C) contracts from NHLBI.

The Framingham Heart Study (FHS) (n=4,124): The FHS is a single-site, community-based, prospective study that was initiated in 1948 to investigate the risk factors for CVD⁴. The second generation⁵ was recruited in 1971 and the third generation⁶ was recruited between 2002 and 2005. The first generation has been examined every two years. The second generation has been examined every 4-8 years. The third generation has had three examinations. A small number of spouse individuals of the second generation was examined at the same time when the third generation had their first examination. A total of 4,196 FHS participants were whole genome sequenced (WGS) by TOPMed; of those, 376 were the first generation, 2218 were the second generation and 95 were spouses of the second generation participants; and 1507 were the third generation participants. This study included 4,123 FHS participants.

The WGS for FHS (phs000974) was performed at the Broad Institute of MIT and Harvard (3R01HL092577-06S1 and 3U54HG003067-12S2). The FHS acknowledges the support of contracts NO1-HC-25195, HHSN268201500001I and 75N92019D00031 from the National Heart, Lung and Blood Institute and grant supplement R01 HL092577-06S1 for this research. We also acknowledge the dedication of the FHS study participants without whom this research would not be possible. Dr. Vasan is supported in part by the Evans Medical Foundation and the Jay and Louis Coffman Endowment

from the Department of Medicine, Boston University School of Medicine. X.L., S.S., C.L.S, and C.L. are also supported by individual funding (R01AG059727).

The Genetic Epidemiology Network of Arteriopathy (GENOA) (n=1,253): The Genetic Epidemiology Network of Arteriopathy (GENOA) study enrolled sibships in which at least 2 siblings had essential hypertension diagnosed prior to age 60 years. From 1995 to 2000, the first exam enrolled 1583 non-Hispanic white Americans from Rochester, Minnesota, and 1854 African Americans from Jackson, Mississippi.⁷ All siblings within the sibship were invited to participate, including both normotensive and hypertensive siblings. The second exam re-recruited 80% of participants from 2000 to 2005. The GENOA data consists of biological samples (DNA, serum, urine) as well as demographic, anthropometric, environmental, clinical, biochemical, physiological, and genetic data for understanding the genetic predictors of diseases of the heart, brain, kidney, and peripheral arteries.

Cohort acknowledgement/support: Support for GENOA was provided by the National Heart, Lung and Blood Institute (HL054457, HL054464, HL054481, HL119443, and HL087660) of the National Institutes of Health. Sequencing for the Genetic Epidemiology Network of Arteriopathy (GENOA) (phs001345.v1.p1) was performed by the University of Washington Northwest Genomics Center (3R01HL055673-18S1 from the National Heart, Lung, and Blood Institute (NHLBI) and at the Broad Institute of MIT and Harvard (HHSN268201500014C)). The authors also wish to thank the staff and participants of GENOA.

The Jackson Heart Study (n=3,286): The JHS cohort is one of the largest prospective, epidemiologic investigation of CVD among African-Americans residing in the three counties (Hinds, Madison, and Rankin) that make up the Jackson, Mississippi metropolitan area^{8,9}. Data and biologic materials have been collected from 5306 participants, including a nested family cohort of 1498 members of 264 families. The age at enrollment for the unrelated cohort was 35-84 years; the family cohort included related individuals >21 years old. Participants provided extensive medical and social history and had an array of physical and biochemical measurements and diagnostic procedures during a baseline examination (2000-2004), two follow-up examinations (2005-2008 and 2009-2012), and ancillary studies. Samples for genomic DNA were collected during the first two examinations. Consent for genetic studies and broad sharing of genetic data was provided by 3,482 participants. After all quality control procedures, whole genome sequence data are available for 3,406 participants. Follow-up information on vital status, major illnesses or injuries, and hospitalizations to identify intervening clinical events is done annually by phone. Medical records of cardiovascular disease related hospitalizations and death certificates are abstracted and used for adjudication of cardiovascular events and related deaths. This study included 3,286 individuals with WGS and matched metabolic phenotypes.

Cohort acknowledgement/support: The Jackson Heart Study (JHS) (phs000964.v1.p1) is supported and conducted in collaboration with Jackson State University (HHSN268201800013I), Tougaloo College (HHSN268201800014I), the Mississippi State Department of Health (HHSN268201800015I) and the University of

Mississippi Medical Center (HHSN268201800010I, HHSN268201800011I and HHSN268201800012I) contracts from the National Heart, Lung, and Blood Institute (NHLBI) and the National Institute on Minority Health and Health Disparities (NIMHD). WGS was performed at University of Washington (HHSN268201100037C).

The authors also wish to thank the staffs and participants of the JHS.

Multi-Ethnic Study of Atherosclerosis Study (MESA) (n=4,596): The Multi-Ethnic Study of Atherosclerosis (MESA) is a study of the characteristics of subclinical cardiovascular disease and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease. MESA researchers study a diverse, population-based sample of more than 6,800 asymptomatic men and women free of prevalent clinical CVD aged 45-84 from six field centers across the United States over a two-year period.¹⁰ Starting in 2000, event adjudication through 2015 consisted of expert committee review of death certificates, hospital records and telephone interviews. DNA for mtDNA-CN analyses was isolated from exam 1 peripheral leukocytes using the Gentra Puregene Blood Kit. mtDNA-CN was available for 4,977 individuals (24.4% Black, 22.2% Hispanic, 13.1% Chinese) derived from TOPMed WGS sequencing.

Cohort acknowledgement/support: Whole genome sequencing (WGS) for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung and Blood Institute (NHLBI). WGS for “NHLBI TOPMed: Multi-Ethnic Study of Atherosclerosis (MESA)” (phs001416.v1.p1) was performed at the Broad Institute of MIT and Harvard (3U54HG003067-13S1). Centralized read mapping and genotype

calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1). Phenotype harmonization, data management, sample-identity QC, and general study coordination, were provided by the TOPMed Data Coordinating Center (3R01HL-120393-02S1). MESA and the MESA SHARe project (phs001416.v1.p1) are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, UL1-TR-001420. Also supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

Hispanic Community Health Study / Study of Latinos (HCHS_SOL) is a longitudinal cohort study established in 2008 following Hispanics/Latinos from four US cities: Bronx, NY; Chicago, IL; Miami, FL; San Diego, CA. All participants provided written informed consent during their field center visit and the study was approved by the IRB in all field centers. At the present analysis there are 3868 participants with available whole-genome sequencing data from blood drawn in their first field center visit. Detailed information on HCHS/SOL is provided in Sorlie et al. (2010) and Lavange et al. (2010). Statistical methods used to analyze the present data account for the complex study design, including stratified sampling, clustering, and sampling probabilities. We also

adjusted for 11 principal components, estimated from the TOPMed DCC. Sequencing center for HCHS/SOL was Baylor College of Medicine Human Genome Sequencing Center.

HCHS_SOL of Latinos is a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (HHSN268201300001I / N01-HC-65233), University of Miami (HHSN268201300004I / N01-HC-65234), Albert Einstein College of Medicine (HHSN268201300002I / N01-HC-65235), University of Illinois at Chicago – HHSN268201300003I / N01-HC-65236 Northwestern Univ), and San Diego State University (HHSN268201300005I / N01-HC-65237). The WGS of HCHS/SOL (phs001395) was performed at Baylor Human Genome Sequencing Center (HHSN268201600033I). The following Institutes/Centers/Offices have contributed to the HCHS/SOL through a transfer of funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, NIH Institution-Office of Dietary Supplements.

UK Biobank (n=39,106): UK Biobank is a prospective cohort study, with the aim of improving the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses – including cancer, heart diseases, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression and forms of dementia. UK Biobank recruited

500,000 individuals from across the United Kingdom aged between 40-69 years in 2006-2010, who have undergone measures, provided blood, urine and saliva samples for future analysis.¹¹ 39,106 participants with whole exome sequencing and matched metabolic phenotypes were included in this study.

This research has been conducted using the UK Biobank Resource under Application Number 17731.

mtDNA CN estimation from whole exome sequencing (WES) in the UK Biobank

We started with 49,997 Exome SPB CRAM files (version Jul 2018) downloaded from the UKB data repository, and used Samtools (ver1.9) to extract read summary statistics ('idxstats' command). A custom perl script was used to aggregate the summary statistics from each individual file into the following categories (see perl script and example stats file): 1) Total Reads (sum of columns 3 and 4, across all rows), 2) Mapped Reads (sum of column 3, across all rows), 3) Unmapped Reads (some of column 4 across all rows), 4) Autosomal Reads (sum of column 3, rows 1-22), 5) Chr X, 6) Chr Y, 7) Chr MT, 8) 'Random' Reads (sum of column 3, across rows 26-67), 9) 'Unknown' Reads (sum of column 3 across rows 68-194), 10) EBV Reads, 11) 'Decoy1' Reads (sum of column 3 across rows 196-582), 12) 'Decoy2' Reads (sum of column 3 across rows 583-2580). Linear regression models were used to adjust for total DNA and potential technical artifacts. Specifically, we used 10-fold cross validation for variable selection, using the 'leaps' R package (version 3.0), with an initial model with chrMT read count as the dependent variable, and 'Total', 'Mapped', 'unknown', 'random', 'decoy1' and 'decoy2' read counts as the independent variables. For each of the independent variables, we

included a natural spline with $df=4$ to allow for non-linear effects. The independent variables 'Total', 'unknown', 'decoy1' and 'decoy2' read counts were selected. We then increased the natural spline df to 15, and then used backward selection to reduce model complexity, requiring $P < 0.005$ to keep a term in the model. The final regression model residuals were generated with the following R (version 3.6.0) code:

```
resid.mtDNA=residuals(lm(chrMT ~ ns(Total,df=3) + ns(unknown,df=4) + ns(decoy1,df=7)
+ decoy2))
```

Residuals from this model were then adjusted for age, with a natural spline $df=2$, and sex. For analyses adjusted for cell type counts, we included terms in the regression model for Lymphocytes, Platelets, Monocytes, Nucleated RBCs, Neutrophils, Basophils, and Eosinophils. For these analyses, outliers of cell type counts were removed based on the following criteria:

$\log(WBC) \leq 1.25$ or ≥ 3 , $\log(RBC) \leq 1.4$ or ≥ 2 , $\text{platelet} \leq 10$ or ≥ 500 , $\log(\text{Lymphocytes}) \leq 0.10$ or ≥ 2 , $\log(\text{Monocytes}) \geq 0.9$, $\log(\text{Neutrophils}) \leq 0.75$ or ≥ 2.75 , $\log(\text{Eosinophils}) \geq 0.75$, $\log(\text{Basophils}) \geq 0.45$. For all analyses, mtDNA-CN was standardized by subtracting the mean and dividing by the standard deviation.

mtDNA CN estimation from Affymetrix Genome-Wide Human SNP Array 6.0 in ARIC cohort

mtDNA-CN was determined using the Genvisis15 software package for the Affymetrix Genome-Wide Human SNP Array 6.0. A list of high-quality mitochondrial SNPs were hand-curated by employing BLAST to remove SNPs without a perfect match to the

annotated mitochondrial location and SNPs with off-target matches longer than 20bp. The probe intensities of the remaining mitochondrial SNPs (25 Affymetrix, 58 Illumina Exome Chip) were determined using quantile sketch normalization (apt-probeset-summarize) as implemented in the Affymetrix Power Tools software. The median of the normalized intensity, log R ratio (LRR) for all homozygous calls was GC corrected and used as initial estimates of mtDNA-CN for each sample.

Technical covariates such as DNA quality, DNA quantity, and hybridization efficiency were captured via surrogate variable analysis described². Surrogate variables were applied to the BLAST filtered, GC corrected LRR of the remaining autosomal SNPs (43,316 Affymetrix, 47,512 Exome Chip). These autosomal SNPs were selected based on the following quality filters: call rate > 98%, HWE p value > 0.00001, PLINK mishap for non-random missingness p value > 0.0001, association with sex p value > 0.00001, linkage disequilibrium pruning ($r^2 < 0.30$), with maximal spacing between autosomal SNPs of 41.7 kb.

mtDNA CN estimation from low-pass whole genome sequencing in ARIC cohort

Low-pass whole genome sequencing data for ARIC was generated at the Baylor College of Medicine Human Genome Sequencing Center using Nano or PCR-free DNA libraries on the Illumina HiSeq 2000. Sequence reads were mapped to the hg19 reference genome using BWA¹². Quality control was performed as previously described¹³. A count for the total number of reads in a sample was scraped from the NCBI sequence read archive using the R package RCurl¹⁴ while reads aligned to the mitochondrial genome were downloaded directly through Samtools (version 1.3.1). A

raw measure of mtDNA-CN was calculated as the ratio of mitochondrial reads to the number of total aligned reads.

The final mtDNA-CN phenotype for all measurement techniques is represented as the standardized residuals from a linear model adjusting the raw measure of mtDNA-CN for age, sex, DNA collection center, any technical covariates. As mtDNA-CN was standardized, the effect size estimates are in units of standard deviations, with positive betas corresponding to an increase in mtDNA-CN.

Supplemental Table 1. Cohort characteristics

Variable Mean (sd) or N (%)	ARIC (EA) (N=2770)	CARDIA (EA) (N=1840)	CHS (EA) (N=2788)	FHS (EA) (N=4124)	MESA (EA) (N=1863)	UKB (EA) (N=39,106)
Ancestry	EA	EA	EA	EA	EA	EA
mtDNA CN (men)	271 (77)	197 (71)	249 (119)	145 (50)	238 (70)	-0.047 (0.29)
mtDNA CN (women)	313 (75)	221 (94)	281 (136)	151 (49)	255 (74)	0.038 (0.29)
Age	59 (6)	45 (7)	74 (6)	60 (16)	62 (10)	57 (8)
Women	(1425) 51%	989 (54%)	1589 (57%)	2227 (54%)	911 (49%)	21148 (54%)
Obesity	750 (27%)	500 (27%)	502 (18%)	1116 (28%)	519 (28%)	9203 (24%)
HTN	1013 (37%)	275 (15%)	1868 (67%)	1909 (48%)	781 (42%)	19754 (51%)
DIAB	362 (13%)	88 (5%)	390 (14%)	441 (11%)	96 (5%)	1857 (5%)
Hyperlipidemia	1911 (69%)	943 (51%)	1979 (71%)	2616 (66%)	1240 (67%)	31046 (80%)
Metabolic syndrome	947 (34%)	389 (21%)	1310 (47%)	1722 (43%)	688 (37%)	12929 (33%)

Variable Mean (sd) or (%)	ARIC (AA) (N=201)	JHS (AA) (N=3286)	GENOA (AA) (N=1253)	CHS(AA) (N=705)	CARDIA (AA) (N=1612)	MESA (AA) (N=1108)
Ancestry	AA	AA	AA	AA	AA	AA
mtDNA CN (men)	304.2 (86)	122 (48)	238 (77)	278 (118)	220 (83)	243 (71)
mtDNA CN (women)	292 (77)	134 (42)	255 (76)	283 (121)	247 (112)	268 (84)
Age	59 (6)	56 (13)	62 (10)	74 (6)	44 (7)	63 (10)
Women	118 (59%)	2037 (62%)	877 (70%)	444 (63%)	948 (59%)	499 (45%)
Obesity	80 (40%)	1774 (54%)	677 (54%)	247 (35%)	822 (51%)	510 (46%)
HTN	128 (64%)	1972 (60%)	977 (78%)	557 (79%)	557 (35%)	676 (61%)
DIAB	66 (33%)	756 (23%)	363 (29%)	169 (24%)	169 (10%)	199 (18%)
Hyperlipidemia	138 (69%)	1939 (59%)	852 (68%)	444 (63%)	707 (44%)	499 (45%)
Metabolic syndrome	73 (36%)	1249 (38%)	589 (47%)	338 (48%)	476 (30%)	454 (41%)

Variable Mean (sd) or (%)	MESA (EAS) (N=601)	MESA (HA) (N=1024)	SOL (HA) (N=3868)
Ancestry	East Asian	Hispanic	Hispanic
mtDNA CN (men)	227 (64)	232 (64)	202 (36)
mtDNA CN (women)	245 (70)	254 (69)	218 (40)
Age	62 (10)	62 (11)	47 (14)
Women	299 (50%)	526 (51%)	2294 (60%)
Obesity	33 (4%)	455 (39%)	1760 (46%)
HTN	301 (39%)	508 (43%)	839 (22%)
DIAB	401 (13%)	219 (19%)	827 (21%)
Hyperlipidemia	419 (54%)	762 (65%)	634 (16%)
Metabolic syndrome	199 (33%)	480 (47%)	808 (21%)

ARIC, Atherosclerosis Risk in Communities study; CARDIA, Coronary Artery Risk Development in Young Adults Study; CHS, Cardiovascular Health Study; FHS, Framingham Heart Study; GENOA, Genetic Epidemiology Network of Arteriopathy Study; JHS, Jackson Heart Study; MESA, Multi-Ethnic Study of Atherosclerosis; UKB, UK Biobank.

Supplemental Table 2. Association analysis of mtDNA CN with cell counts

Covariates	FHS (N=~2,600)		GENOA (N=~880)		JHS (N=~2,800)		UKB (N=~39,100)		SOL (N=~3,600)	
	Coef	P-value	Coef	P-value	Coef	P-value	Coef	P-value	Coef	P-value
Neutrophil count	-0.048	0.56	na	na	na	na	-0.072	<2e-16	-0.5528	5.73e-47
Platelet count	0.0032	2.0e-11	na	na	0.0010	0.00026	0.00074	<2e-16	20.4756	3.53e-28
White blood cell count	-0.10	6.8e-12	na	na	-0.064	5.0e-09	na	na	-0.6017	6.68e-44
Lymphocyte count	-0.027	0.74	na	na	0.025	<2e-16	0.026	<2e-16	0.0205	1.12e-01
Red blood cell count	0.17	0.0012	na	na	-0.011	0.76	0.22	0.0077	-0.0074	1.41e-13
Eosinophil count	-0.024	0.78	na	na	0.0015	0.86	-0.024	0.040	0.1163	3.82e-01
Basophil count	na	na	na	na	0.057	0.22	-0.11	0.0022	-0.0017	4.29e-02
Monocyte count	-0.025	0.77	-12.20	0.095	0.015	0.063	-0.055	4.4e-08	-0.0349	1.83e-04
CD8T cell count	na	na	-14.51	0.042	na	na	na	na		
CD4T cell count	na	na	-13.70	0.057	na	na	na	na		
Natural killer cell count	na	na	-13.49	0.040	na	na	na	na		
B lymphocytes cell count	na	na	-13.27	0.062	na	na	na	na		
Granulocytes cell count	na	na	-16.72	0.019	na	na	na	na		

Association analysis between mtDNA CN residuals with cell counts. mtDNA CN residuals was obtained by regressing mtDNA CN on batch effect ,age and sex.

Supplemental Table 3. Meta-analysis using inverse-variance weighting in the TOPMed participants of African American ancestry

Traits	African Americans Ancestry in TOPMed (N=8,100)		
	Beta	SE	P-value
Obese	0.009	0.027	0.74
HTN	0.055	0.03	0.068
Diabetes	0.13	0.036	2.5e-04
Hyperlipids	-0.0082	0.028	0.77
MetS	0.070	0.028	0.014
BMI	0.032	0.076	0.68
DBP	0.32	0.13	0.017
SBP	0.68	0.24	4.0e-03
FBG	0.15	0.13	0.25
HDL	-0.0003	0.0032	0.92
LDL	-0.0047	0.004	0.24
TRIG	-0.0047	0.004	0.24

Association analysis of mtDNA CN with metabolic traits was performed in each cohort of African American (AA) in TOPMed (ARIC, CARDIA, CHS, GENOA, JHS, and MESA). Meta-analysis using fixed effects inverse variance method was used to summarize the results in AA in TOPMed. The effect size estimates are in units of cardiometabolic traits corresponding to one s.d. decrease in mtDNA-CN. DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein; TRIG, triglyceride; Obese, obesity; HTN, hypertension; Diabetes, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome. na, not applicable.

Supplemental Table 4. Association analysis between mtDNA CN and metabolic phenotypes in participants of Hispanic and East Asians

Traits	HA (N=4,880)		EAS (N=601)	
	Beta/SE	P	Beta/SE	P
Obese	0.0022/0.025	0.93	-0.069/0.22	0.75
HTN	0.047/0.042	0.26	-0.037/0.099	0.71
DIAB	0.14/0.045	0.0012	0.20/0.14	0.16
Hyperlipids	-0.011/0.048	0.81	0.36/0.094	0.00014
Metabolic syndrome	0.017/0.0056	0.0027	0.069/0.096	0.47
BMI	0.23/0.12	0.058	-0.16/0.14	0.24
DBP	0.03/0.21	0.88	-0.11/0.49	0.82
SBP	0.12/0.32	0.72	-0.87/0.94	0.36
FBG	-0.084/0.19	0.66	-0.070/0.46	0.88
HDL	-0.47/0.49	0.34	0.0001/0.0093	0.99
LDL	-0.019/0.0055	6.6e-04	0.020/0.011	0.071
TRIG	0.028/0.0091	0.0020	0.027/0.020	0.18

Association analysis of mtDNA CN with cardiometabolic traits was performed in Hispanic participants (HA) in TOPMed MESA and SOL study, and East Asia participants (EAS) in TOPMed MESA study. Meta-analysis using fixed ($P_Q \geq 0.01$) or random ($P_Q < 0.01$) effects inverse variance method was used to summarize the HA results. The effect size estimates are in units of cardiometabolic traits corresponding to one s.d. decrease in mtDNA-CN. DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein, TRIG, triglyceride. Obese, obesity; HTN, hypertension; Diabetes, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome.

Asterisk (*) denotes inverse variance random effect meta-analysis with random effects method.

Supplemental Table 5. Meta-analysis using inverse-variance weighting in participants of all ancestries in TOPMed and UK Biobank

Traits	Beta	SE	P-value
Obese	-0.057*	0.036*	0.12*
HTN	-0.052	0.0096	5.4e-08
Diabetes	-0.11	0.017	3.5e-11
Hyperlipids	-0.025	0.011	0.023
MetS	-0.085*	0.035*	0.014*
BMI	-0.14*	0.074*	0.054*
DBP	-0.19	0.048	8.0e-05
SBP	-0.43	0.077	3.9e-08
FBG	-0.15	0.042	6.2e-04
lnHDL	0.0052	0.002	0.0093
lnLDL	0.0042*	0.0048*	0.38*
lnTRIG	-0.022*	0.0054*	4.0e-05*

Association analysis of mtDNA CN with cardiometabolic traits was performed in cohorts of European participants (EA), African Americans (AA), Hispanic participants (HA), and East Asian participants (EAS) in TOPMed and UK Biobank. Meta-analysis using fixed ($P_Q \geq 0.01$) or random ($P_Q < 0.01$) effects inverse variance method was used to summarize the results. The effect size estimates are in units of cardiometabolic traits corresponding to one s.d. decrease in mtDNA-CN. DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein, TRIG, triglyceride; Obese, obesity; HTN, hypertension; Diabetes, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome. na, not applicable.

Asterisk (*) denotes inverse variance random effect meta-analysis with random effects method.

Supplemental Table 6. Meta-analysis using inverse-variance weighting in EA participants with cell counts in TOPMed and UK Biobank

Traits	TOPMed WGS EA and UK Biobank EA (N=~44,100)			
	w/o adjusting for cell counts		Adjusting for cell counts	
	Beta/SE	P-value	Beta/SE	P-value
Obesity	0.14/0.012	7.7e-31	0.056/0.012	2.5e-06
HTN	0.060/0.11	1.2e-07	0.017/0.011	0.13
Diabetes	0.097/0.024	5.8e-05	0.0031/0.024	0.89
Hyperlipids	0.031/0.014	0.028	0.058/0.014	3.9e-05
MetS	0.14/0.011	1.1e-32	0.035/0.011	0.0016
BMI	0.32/0.023	7.7e-44	0.15/0.023	1.3e-11
DBP	0.19/0.052	2.4e-04	0.071/0.052	0.17
SBP	0.47/0.091	2.2e-07	0.14/0.091	0.11
FBG	0.20/0.050	1.1e-04	0.18/0.05	2.3e-04
lnHDL	-0.16/0.071	0.027	0.0026/0.0066	0.69
lnLDL	0.0021/0.0014	0.14	0.012/0.0014	7.8e-18
lnTrig	0.021/0.0024	9.8e-19	0.016/0.0024	6.7e-12

Association analysis of mtDNA CN with cardiometabolic traits was performed in the participants with imputed cell counts in European participants (EA) in TOPMed and in UK Biobank. The effect size estimates are in units of cardiometabolic traits corresponding to one s.d. decrease in mtDNA-CN. DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein; TRIG, triglyceride; Obese, obesity; HTN, hypertension; Diabetes, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome. FHS, Framingham Heart Study; JHS, Jackson Heart Study; GENOA, Genetic Epidemiology Network of Arteriopathy Study; UKB, UK Biobank. HDL, LDL and TRIG values were log-transformed.

Supplemental table 7. Comparison of results of mtDNA CN with cardiometabolic disease phenotypes in WGS, Affymetrix, and low-pass in non-overlap participants in ARIC.

	Traits	TOPMed WGS (N=~2,770)			Affymetrix (N=~5,200)			Low-pass sequencing (N=~1,440)		
		Beta	SE	P-value	Beta	SE	P-value	Beta	SE	P-value
Binary outcomes	Obese	0.090	0.044	0.041	0.059	0.033	0.074	0.12	0.062	0.053
	HTN	0.14	0.044	0.0015	0.084	0.032	0.0087	0.075	0.061	0.22
	Diabetes	0.39	0.063	6.0e-10	0.26	0.045	7.6e-09	0.33	0.10	0.00097
	Hyperlipids	0.10	0.043	0.020	-0.0059	0.030	0.84	0.013	0.059	0.83
Continuous outcomes	BMI	0.23	0.098	0.019	0.11	0.070	0.12	0.29	0.14	0.038
	DBP	0.30	0.22	0.17	0.28	0.15	0.062	0.12	0.29	0.68
	SBP	1.00	0.38	0.0085	0.95	0.26	0.00026	0.71	0.51	0.16
	FBG	0.11	0.20	0.58	0.16	0.13	0.22	0.073	0.27	0.79
	HDL	-1.40	0.28	5.7e-07	-1.37	0.21	6.9e-11	-0.63	0.39	0.11
	TRIG	0.055	0.0091	1.50e-09	0.027	0.0069	9.1e-05	0.013	0.013	0.32

The beta estimates are in units of cardiometabolic traits corresponding to one s.d. lower mtDNA-CN. Association analysis of mtDNA CN with cardiometabolic traits was performed in cohorts of European participants (EA) with WGS, Affymetrix Genome-Wide Human SNP Array 6.0 and low-pass whole genome sequencing. DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; Trig, triglyceride. Obese, obesity; HTN, hypertension; Diabetes, Diabetes; MetS, metabolic syndrome. Asterisk (*) denotes inverse variance random effect meta-analysis.

Supplemental Table 8. The investigation of effect modification by sex or age on associations of mtDNA CN with cardiometabolic traits.

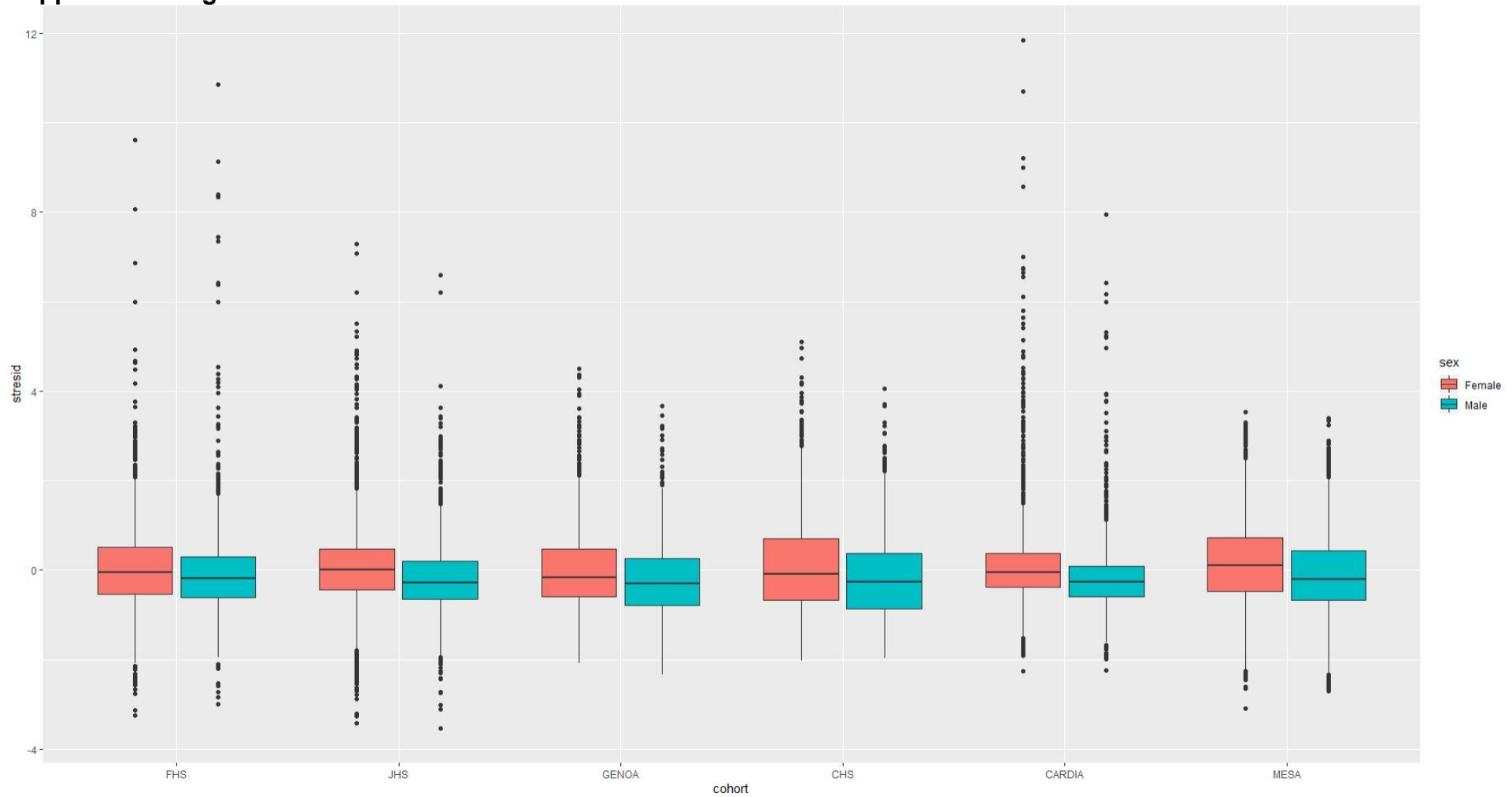
Traits	CARDIA EA		CHS EA		FHS		MESA EA		UKB		META	
	P(age*)	P(sex*)	P(age*)	P(sex*)	P(age*)	P(sex*)	P(age*)	P(sex*)	P(age*)	P(sex*)	P(age*)	P(sex*)
BMI	0.72	0.74	0.37	0.13	0.73	0.98	0.12	0.84	0.59	0.22	0.16	0.34
DBP	0.50	0.31	0.75	0.67	0.52	0.17	0.28	0.33	0.54	0.62	0.35	0.36
SBP	0.95	0.16	0.68	0.22	0.31	0.07	0.052	0.34	0.14	0.22	0.15	0.16
FBG	0.27	0.79	0.0060	0.14	0.65	0.71	0.47	0.16	0.58	0.33	0.94	0.43
HDL	0.056	0.93	0.45	0.80	0.35	0.29	0.52	0.58	0.58	0.15	0.49	0.72
LDL	0.72	0.45	0.66	0.18	0.32	0.29	0.82	0.11	0.74	0.31	0.78	0.56
TRIG	0.030	0.46	0.66	0.34	0.35	0.53	0.49	0.23	0.99	0.85	0.62	0.73
Obesity	0.75	0.80	0.24	0.21	0.45	0.78	0.33	0.41	0.69	0.99	0.92	0.92
HTN	0.32	0.73	0.85	0.88	0.20	0.07	0.095	0.81	0.49	0.31	0.87	0.69
DIAB	0.22	0.062	0.18	0.20	0.02	0.66	0.34	0.31	0.81	0.97	0.09	0.58
Hyperlipids	0.25	0.44	0.41	0.55	0.93	0.32	0.34	0.19	0.78	0.36	0.80	0.97
MetS	0.72	0.15	0.53	0.91	0.65	0.82	0.59	0.26	0.17	0.16	0.11	0.25

Traits	CARDIA AA		CHS AA		GENOA		JHS		MESA AA		META	
	P(age*)	P(sex*)	P(age*)	P(sex*)	P(age*)	P(sex*)	P(age*)	P(sex*)	P(age*)	P(sex*)	P(age*)	P(sex*)
BMI	0.48	0.59	0.13	0.77	0.51	0.34	0.74	0.08	0.21	0.96	0.28	0.38
DBP	0.74	0.25	0.68	0.13	0.28	0.53	0.56	0.58	0.88	0.18	0.95	0.42
SBP	0.48	0.46	0.69	0.37	0.51	0.73	0.92	0.13	0.33	0.83	0.76	0.29
FBG	0.48	0.67	0.32	0.16	0.94	0.89	0.014	0.87	0.31	0.30	0.02	0.91
HDL	0.019	0.60	0.61	0.27	0.39	0.66	0.92	0.82	0.094	0.84	0.31	0.82
LDL	0.053	0.47	0.012	0.60	0.92	0.37	0.09	0.12	0.81	0.093	0.04	0.09
TRIG	0.28	0.40	0.04	0.73	0.38	0.57	0.11	0.95	0.92	0.81	0.35	0.52
Obesity	0.15	0.69	0.84	0.089	0.63	0.69	0.43	0.071	0.11	0.32	0.63	0.31
HTN	0.61	0.61	0.69	0.82	0.24	0.24	0.24	0.086	0.86	0.89	0.26	0.12
DIAB	0.89	0.51	0.86	0.28	0.70	0.60	0.012	0.85	0.78	0.96	0.02	0.66
Hyperlipids	0.56	0.35	0.0045	0.98	0.32	0.86	0.054	0.30	0.44	0.056	0.03	0.06
MetS	0.83	0.61	0.87	0.039	0.70	0.64	0.051	0.98	0.59	0.83	0.05	0.38

Association analysis of mtDNA CN with cardiometabolic traits adjusting with an interaction term was performed in each cohort of European participants (EA) and African American (AA) in TOPMed and also in UK Biobank EA participants. Fixed-effect inverse variance meta-analysis was used to summarize the results in EA or AA in TOPMed. An interaction term of mitoage (residual mtDNA CN*age) or mitosex (residual mtDNA CN*sex) was constructed and included in the model to investigate whether age or sex was effect modifier of the association between mtDNA CN and cardiometabolic traits. Residual mtDNA CN was obtained by regressing mtDNA CN on batch effect, obtained the residuals then multiply age

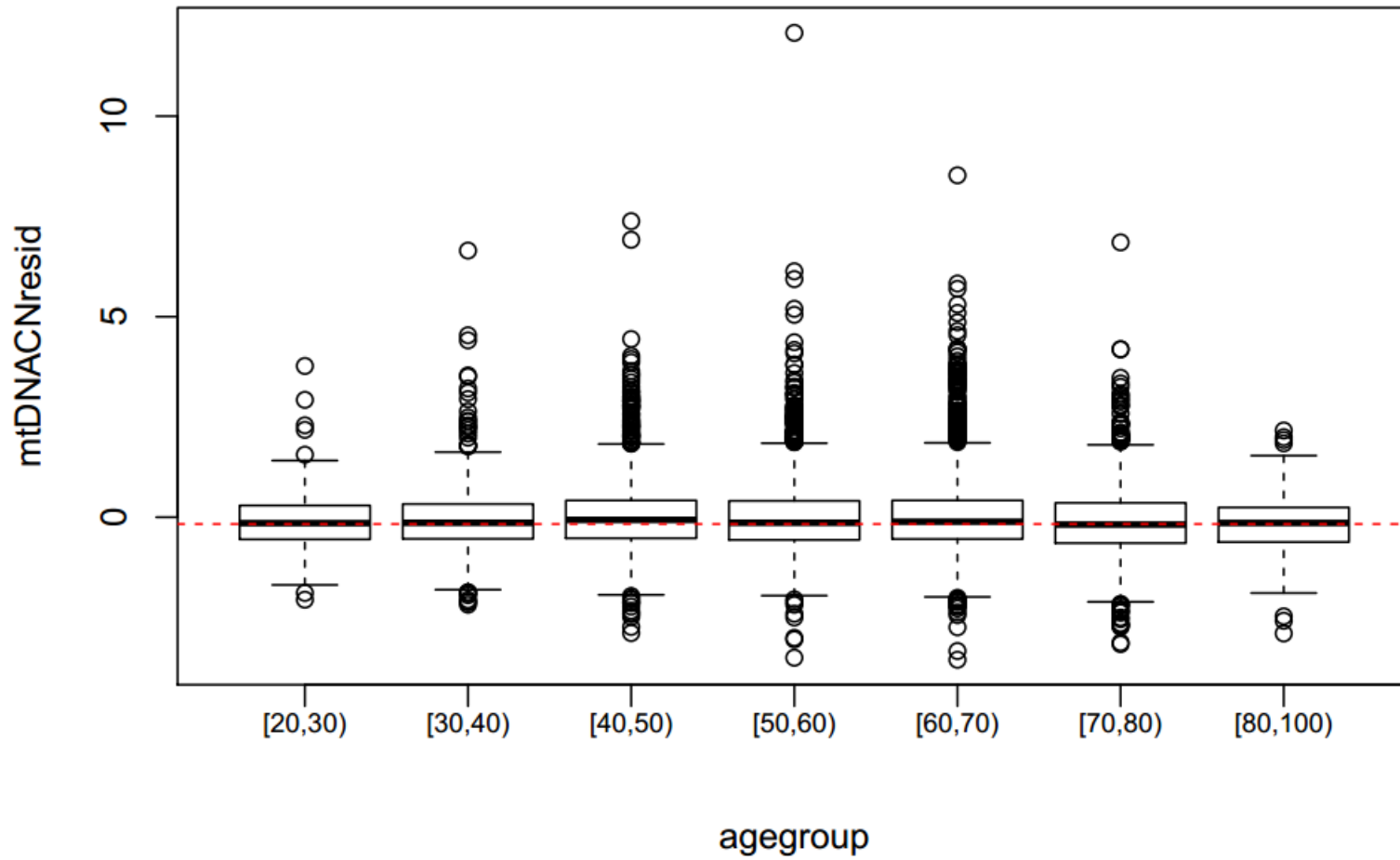
or sex. DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; TRIG, triglyceride. Obese, obesity; HTN, hypertension; Diabetes, Diabetes; MetS, metabolic syndrome.

Supplemental Figure 1. Distribution of mtDNA CN residuals versus sex in each cohort



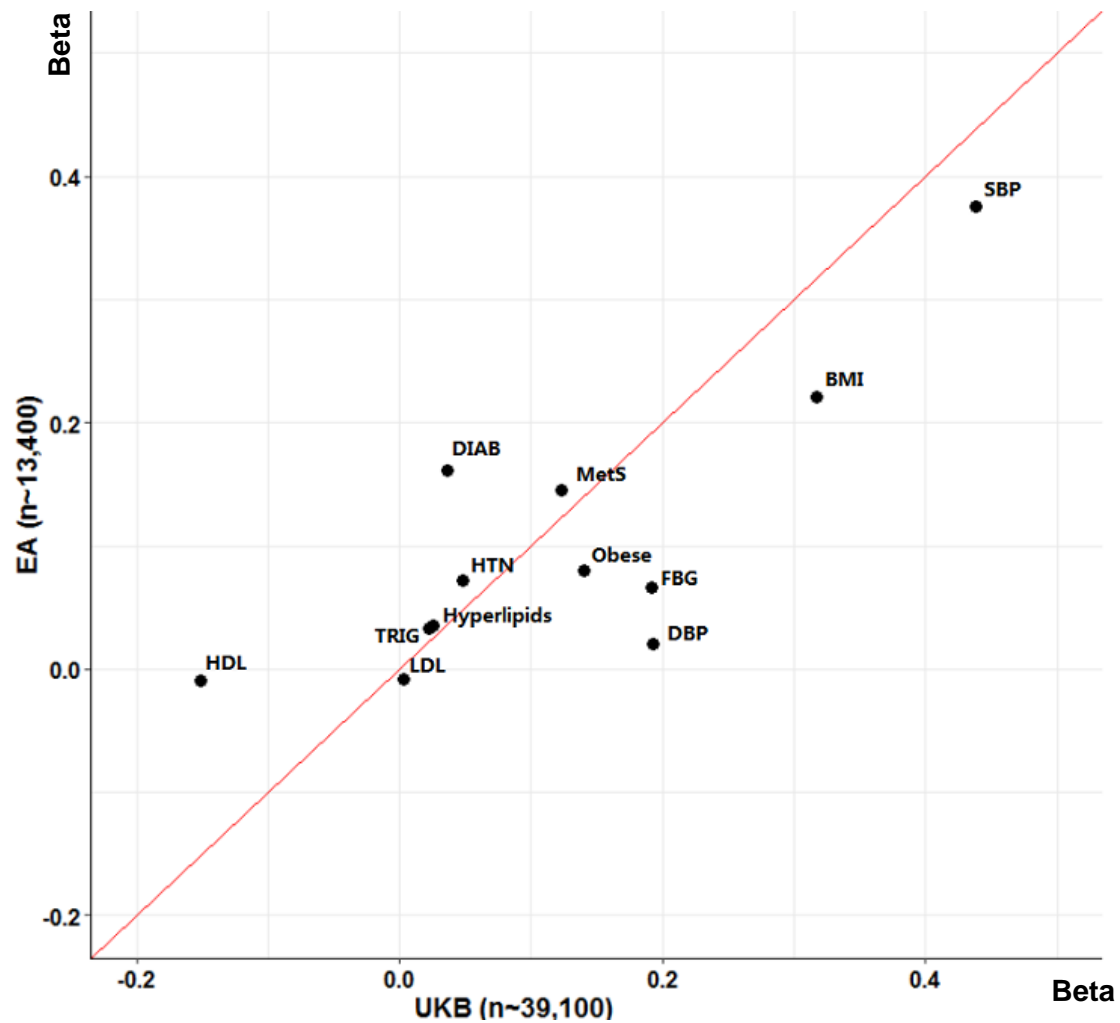
Y-axis is mtDNA CN residuals, x-axis is cohort. mtDNA CN residuals was obtained by regressing mtDNA CN on batch effect and age.

Supplemental Figure 2. Plot of mtDNA CN residuals adjusted for cell counts vs age



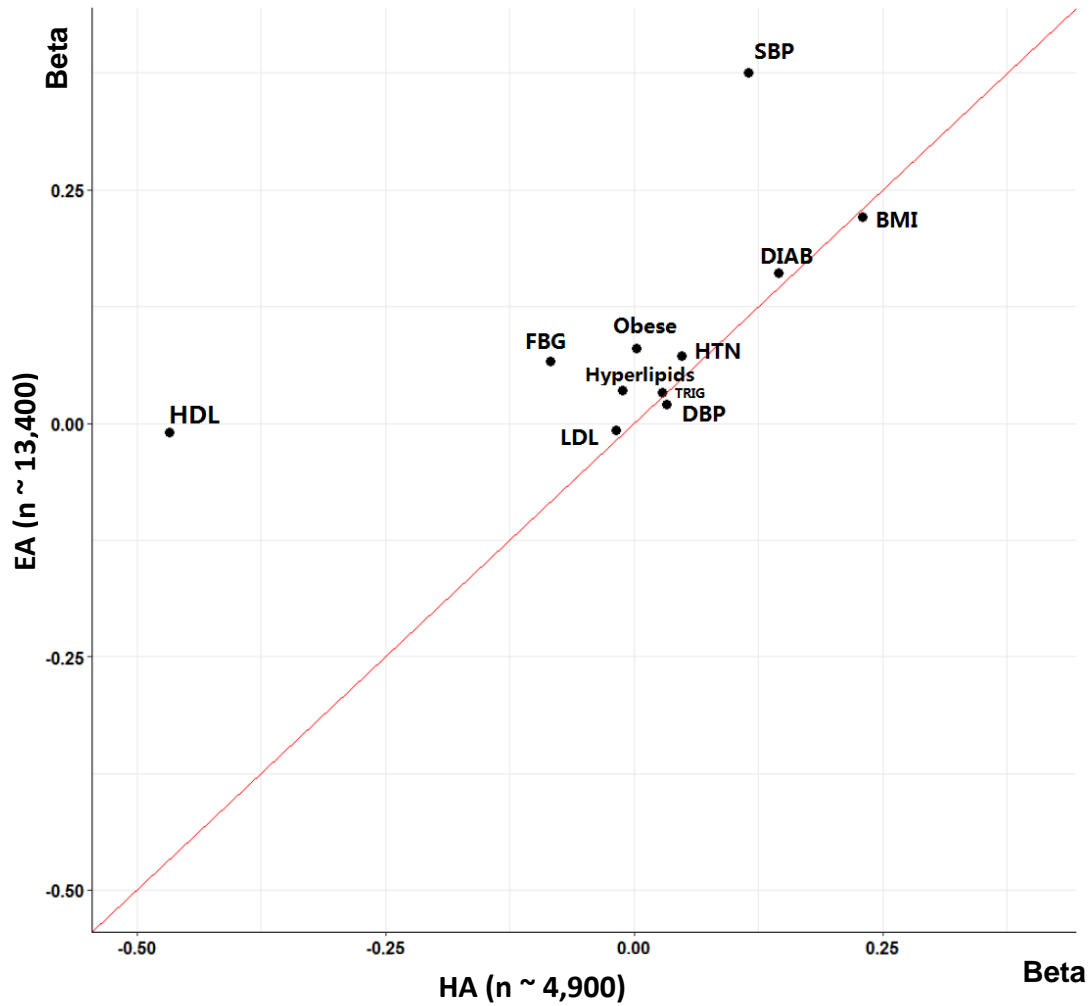
The relationship of mtDNA CN residuals with age in TOPMED participants. Y-axis is mtDNA CN residuals, x-axis is age. mtDNA CN residuals were obtained by regressing mtDNA CN on batch effect and cell count.

Supplemental Figure 3. Comparison of regression coefficients of mtDNA CN with cardiometabolic traits in participants of Europeans in TOPMed vs UK Biobank



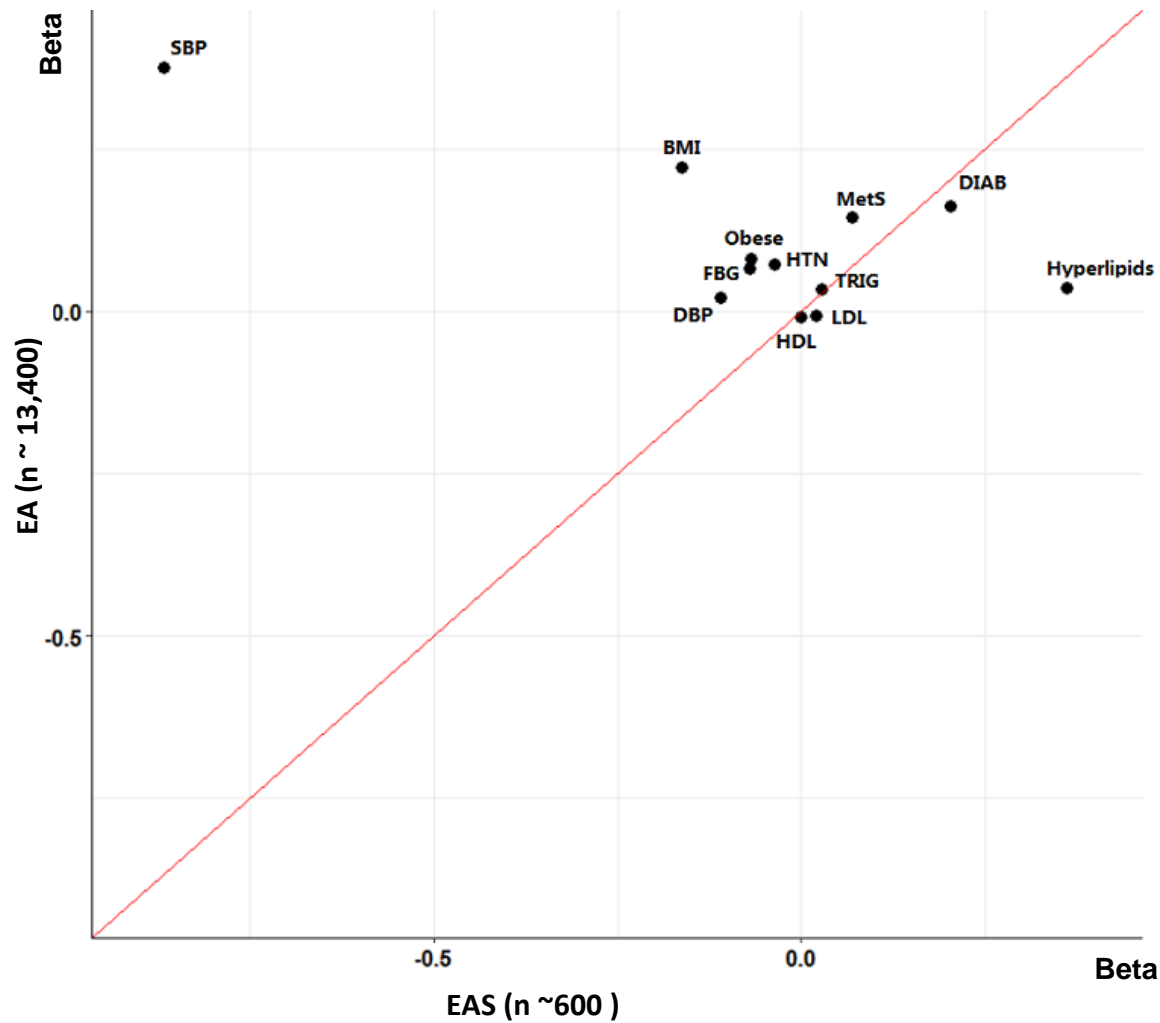
Comparison of beta of cardiometabolic traits in the participants of European (EA) in TOPMed and UK Biobank (UKB). DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein; TRIG, triglyceride; Obese, obesity; HTN, hypertension; DIAB, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome.

Supplemental Figure 4. Comparison of regression coefficients of mtDNA CN with cardiometabolic traits in participants of European ancestry vs Hispanic participants in TOPMed



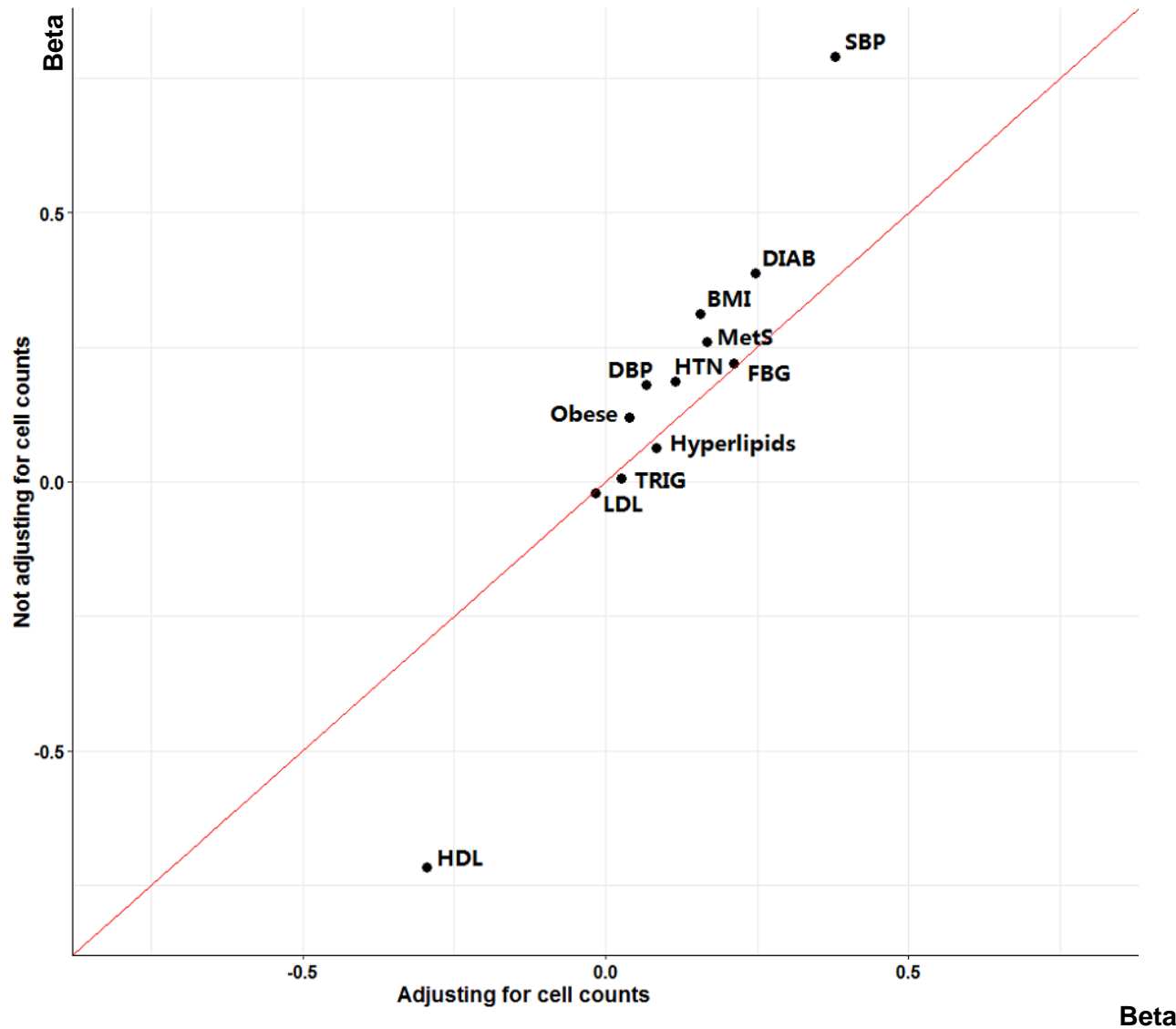
Comparison of beta of cardiometabolic traits in the participants of European (EA) and Hispanic (HA) in TOPMed. DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein; TRIG, triglyceride; Obese, obesity; HTN, hypertension; DIAB, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome.

Supplemental Figure 5. Comparison of regression coefficients of mtDNA CN with cardiometabolic traits in participants of European vs East Asian ancestry in TOPMed



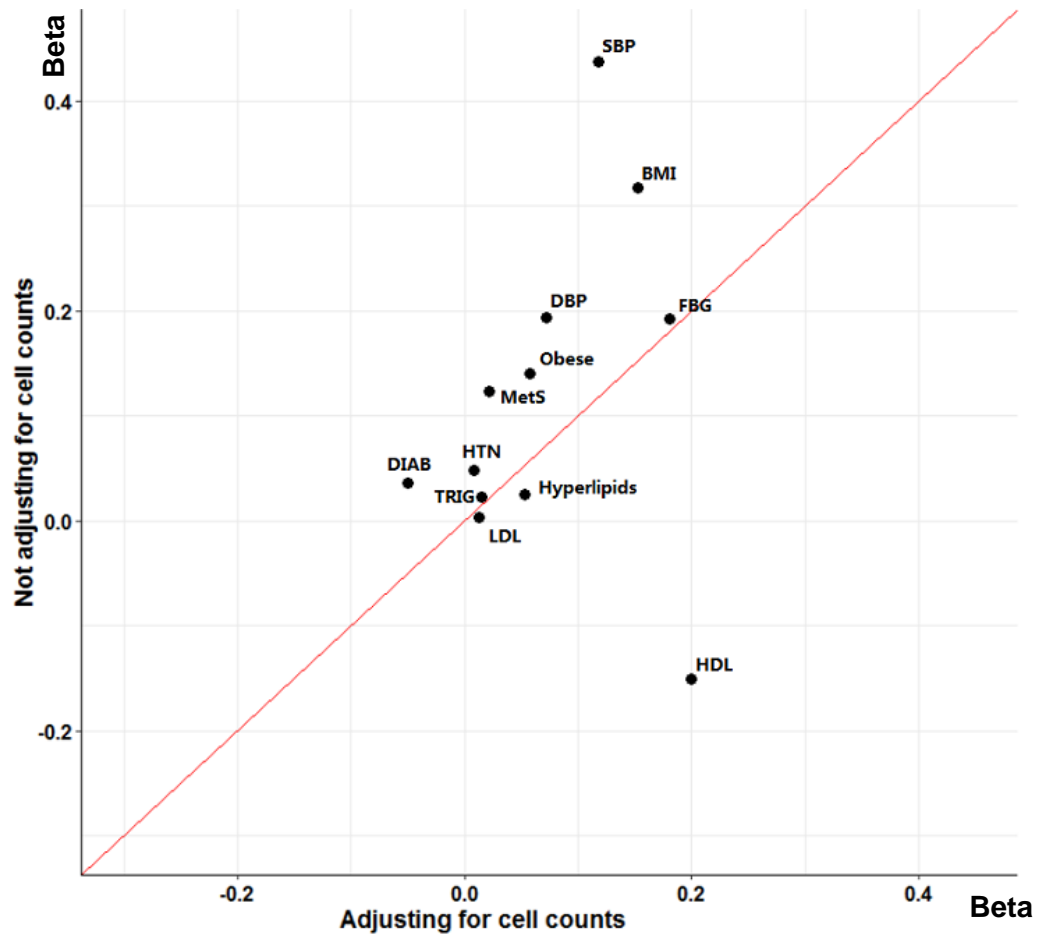
Comparison of beta of cardiometabolic traits in the participants of European (EA) and Chinese (EAS) in TOPMed. DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein; TRIG, triglyceride; Obese, obesity; HTN, hypertension; DIAB, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome.

Supplemental Figure 6. Comparison of regression coefficients of mtDNA CN with cardiometabolic traits in TOPMed participants of European ancestry not adjusting for cell counts vs adjusting for cell counts



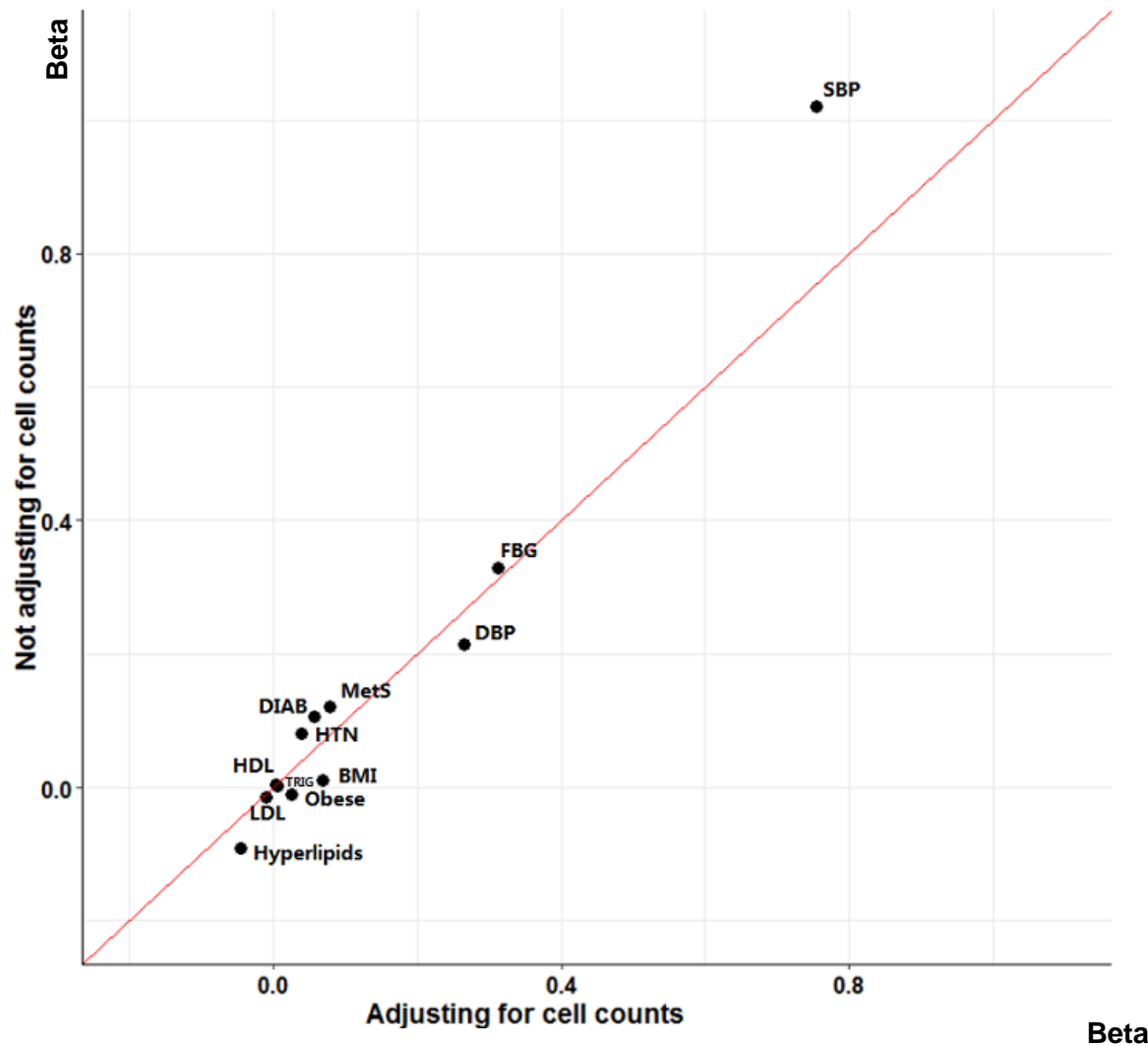
Comparison of beta of cardiometabolic traits of model not adjusting for cell counts vs adjusting for cell counts in the same participants of Europeans (EA) in TOPMed (n~5,000). DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein; TRIG, triglyceride; Obese, obesity; HTN, hypertension; DIAB, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome.

Supplemental Figure 7. Comparison of regression coefficients of mtDNA CN with cardiometabolic traits in participants of UK Biobank not adjusting for cell counts vs adjusting for cell counts



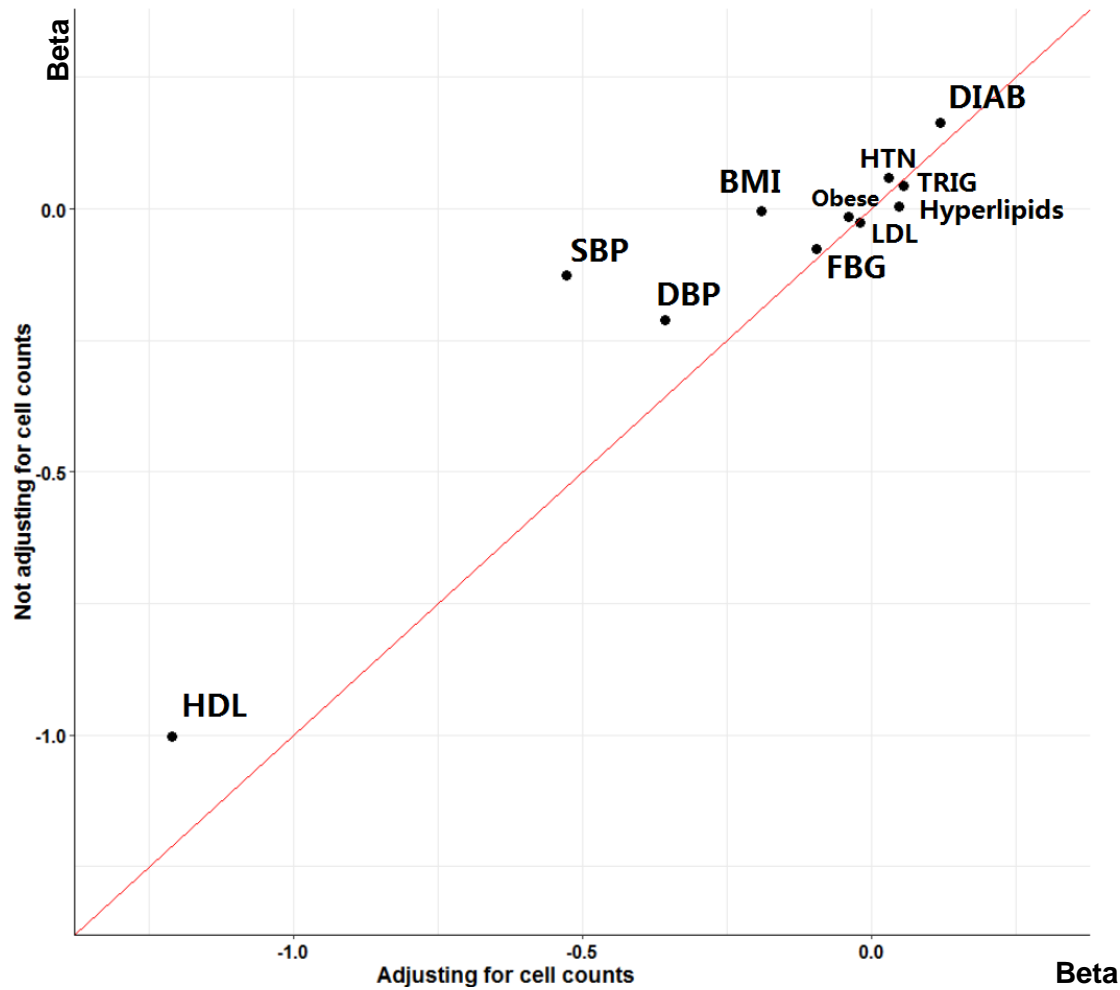
Comparison of beta of cardiometabolic traits not adjusting for cell counts vs adjusting for cell counts in the participants of UK Biobank (UKB) (n~39,000). DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein, Trig, triglyceride. TC, total cholesterol; Obese, obesity; HTN, hypertension; Diabetes, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome.

Supplemental Figure 8. Comparison of regression coefficients of mtDNA CN with cardiometabolic traits in TOPMed African American participants not adjusting for cell counts vs adjusting for cell counts



Comparison of beta of cardiometabolic traits of model not adjusting for cell counts vs adjusting for cell counts in the participants of African Americans (AA) in TOPMed (n~3,700). DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein; TRIG, triglyceride; Obese, obesity; HTN, hypertension; DIAB, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome.

Supplemental Figure 9. Comparison of regression coefficients of mtDNA CN with cardiometabolic traits in TOPMed Hispanic participants not adjusting for cell counts vs adjusting for cell counts



Comparison of beta of cardiometabolic traits of model not adjusting for cell counts vs adjusting for cell counts in the participants of Hispanic Ancestry (HA) in TOPMed (n~3,600). DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; FBG, fasting blood glucose; HDL, high density lipoprotein; LDL, low density lipoprotein; TRIG, triglyceride; Obese, obesity; HTN, hypertension; DIAB, Diabetes; Hyperlipids, hyperlipidemia; MetS, metabolic syndrome.

Reference:

1. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* 1989;129:687-702.
2. FRIEDMAN GD, TEKAWA I, GRIMM RH, MANOLIO T, SHANNON SG, SIDNEY S. The Leucocyte Count: Correlates and Relationship to Coronary Risk Factors: The CARDIA Study. *International Journal of Epidemiology* 1990;19:889-93.
3. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1991;1:263-76.
4. Dawber TR, Meadors GF, Moore FE, Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health* 1951;41:279-81.
5. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med* 1975;4:518-25.
6. Splansky GL, Corey D, Yang Q, et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* 2007;165:1328-35.
7. Daniels PR, Kardia SL, Hanis CL, et al. Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *Am J Med* 2004;116:676-81.
8. Wilson JG, Rotimi CN, Ekunwe L, et al. Study design for genetic analysis in the Jackson Heart Study. *Ethn Dis* 2005;15:S6-30-7.
9. Sempos CT, Bild DE, Manolio TA. Overview of the Jackson Heart Study: a study of cardiovascular diseases in African American men and women. *Am J Med Sci* 1999;317:142-6.
10. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: Objectives and Design. *American Journal of Epidemiology* 2002;156:871-81.
11. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12:e1001779.
12. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25:1754-60.
13. Morrison AC, Voorman A, Johnson AD, et al. Whole-genome sequence-based analysis of high-density lipoprotein cholesterol. *Nat Genet* 2013;45:899-901.
14. Duncan Temple Lang and the CRAN team (2018). RCurl: General Network (HTTP/FTP/...) Client Interface for R. R package version 1.95-4.11. <https://CRAN.R-project.org/package=RCurl>.