Distinct Genetic Loci and Variations in Blood Pressure and Pulse Rate in Europeans and Africans from the UK Biobank

Musalula Sinkala[1Φ], Samar S. M. Elsheikh[2Φ], Mamana Mbiyavanga[1], Joshua Cullinan[1], Nicola J. Mulder[1]

1, University of Cape Town, Faculty of Health Sciences, Institute of Infectious Disease and Molecular Medicine, Computational Biology Division, Anzio Rd, Observatory, 7925, Cape Town, South Africa

2, Pharmacogenetics Research Clinic, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health,
Toronto, ON, Canada

[Φ,] Contributed equally.

Correspondence
musalula.sinkala@uct.ac.za, samar.salah119@gmail.com
[φ]Contributed equally.

Abstract

Trans-ancestral variations exist for various anthropometric traits. These traits, including cardiovascular traits, are likely associated with different loci in individuals of different ancestry groups. We assessed the difference in systolic blood pressure, diastolic blood pressure, pulse rate and maximum heart rate among African and European ancestry individuals. Furthermore, we conducted a genome-wide association study of cardiovascular traits in 383,471 Europeans and 5,978 Africans represented in the UK Biobank. Here, we report 2 and 1,202 variants associated with cardiovascular traits in Africans and Europeans, respectively. We identify 2 novel variants in Africans, including rs9388010 located in the *GJA1* gene previously associated with numerous cardiomyopathies. Remarkably, we find the associated variants are primarily unique to each ancestry group and map to largely different genes. Through integrative enrichment analyses we find that gene sets within each ancestral group are significantly enriched for pathways and phenotypes related to cardiovascular physiology and disease. Our discoveries provide a better understanding of variations in cardiovascular traits and the different associated variants in Africans and Europeans.

Introduction

Cardiovascular traits are complex phenotypes influenced by many different variables, including genetics and environmental factors [1,2]. Various cardiovascular measurements significantly vary among populations of diverse ancestry backgrounds [3,4]. Over the last years, genome-wide association studies (GWAS) have unearthed hundreds of genomic loci associated with cardiovascular traits [5–7]. Through these studies, we now know that loci associated with various cardiovascular traits, including systolic blood pressure (SBP), diastolic blood pressure (DBP) and heart rate overlap with those associated with cardiovascular pathology such as stroke, myocardial infarction and often predict early death and the overall well-being of individuals [8–11]. For example, a recent study of over 1 million individuals of European ancestry identified 535 novel variants associated with blood pressure traits [7]. In addition, another study of 375,256 individuals from the UK Biobank cohort identified 327 and 364 variants associated with systolic BP and diastolic BP, respectively [12]. Notably, these studies and others have identified variants that affect cardiovascular function in individuals predominantly of European descent.

Based on evidence from other GWA studies of other traits[13–19], we suggest that the variants that influence cardiovascular traits differ among individuals of European and African ancestry. The UK Biobank project has profiled over 500,000 individuals of various races and ancestry groups, offering new prospects for GWAS studies to identify and compare variants associated with cardiovascular traits among Europeans and Africans[20]. Identifying variants associated with cardiovascular traits for individuals of African ancestry compared to those of European ancestry is relevant to understanding population differences in cardiovascular pathophysiology. Nevertheless, to our knowledge, there have been no GWA studies to identify and compare genomic loci associated with cardiovascular traits in Africans and Europeans using many cardiovascular measurement parameters in the same dataset.

Here, we evaluated the differences in cardiovascular traits, including SBP, DBP, pulse rate, and Maximum Heart Rate During Fitness Test (MHR) in African and Europeans represented in the UK biobank. Then, we conducted GWA studies to identify variants that affect the four quantitative cardiovascular traits SBP, DBP, pulse rate, and MHR, separately in Africans and Europeans. Finally, we applied integrative enrichment analyses to identify genes, pathways and phenotypes associated with the variants. The aim was to determine whether different genetic variants influence cardiovascular traits in African and European populations.

Results

Cardiovascular parameters vary between individuals of European and African ancestry

We assessed the mean SBP, DBP, pulse rate, and MHR between 383,471 Europeans and 5,978 Africans represented in the UK Biobank datasets. We found that the DBP was significantly higher in Africans (mean DBP = 84.7 mmHg) compared to the Europeans (82.0 mmHg), (Welch test: $t = 13.9$, $p = 1.2 \times 10^{-42}$; Figure 1a). However, we found no significant difference in the mean SBP of Africans (mean = 139.8 mmHg) and Europeans (139.4 mmHg), p = 0.187; Figure 1b). Furthermore, we found that the pulse rate and the MHR were both significantly higher in Africans (mean pulse rate = 70.3 bps, mean MHR = 116.0 bps) than those measured in the Europeans (pulse rate = 68.9 bps, MHR = 110.3 bps), pulse rate; $t = 7.5$, $p = 8.2 \times 10^{-14}$ (Figure 1c) and MHR; $t = 7.5$, $p = 1.9 \times 10^{-13}$; Figure 1d.

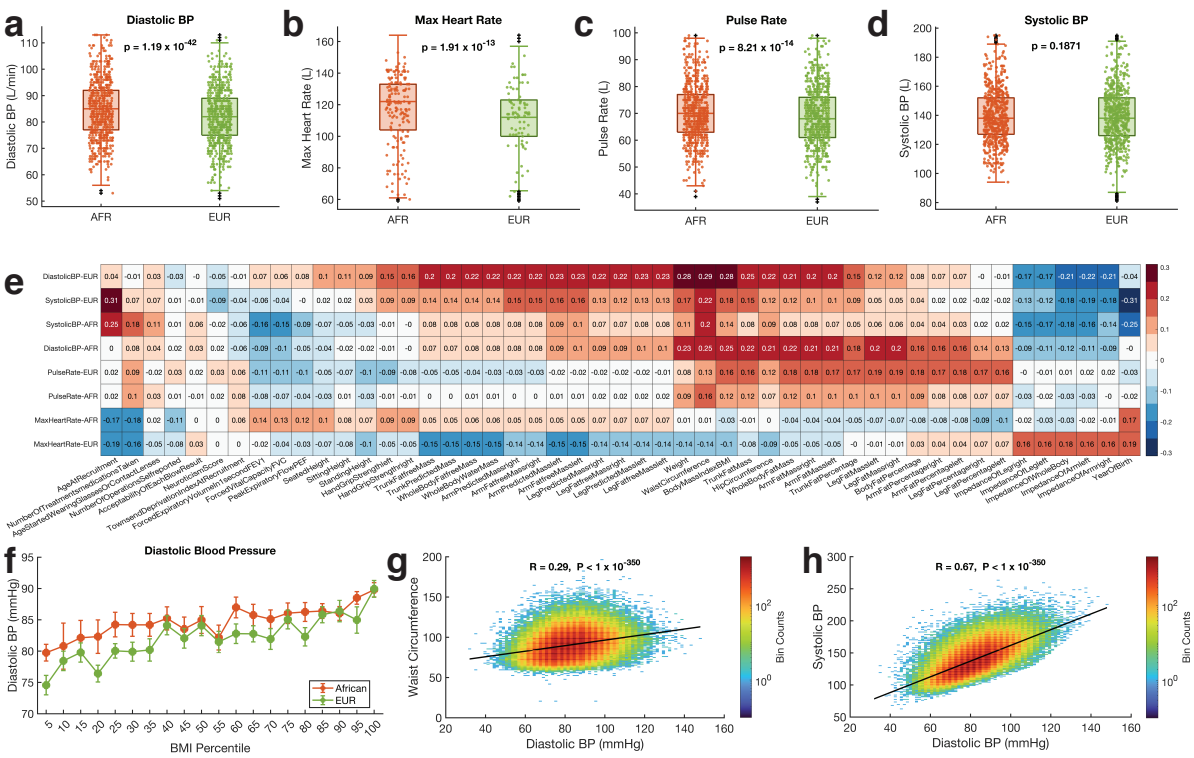

Figure 1: *Comparison of the cardiovascular function parameter in Africans and Europeans for* ***(a)*** *diastolic blood pressure,* ***(b)*** *systolic blood pressure,* ***(c)*** *pulse rate, and* ***(d)*** *maximum heart rate during fitness test. The boxplots indicate the distribution of each measurement among Africans and Europeans. The p-values shown for each comparison were calculated from Welch's t-test. On each*

box, the central mark indicates the median, and the left and right edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol. To make the visualisation clearer, the filled circle mark showing the distribution only include 1000 randomly sampled point from the entire samples size of each group. ***(e)*** *Correlation between SBP, DBP, pulse rate, and MHR with other anthropometric measurements in Africans and Europeans.* ***(f)*** *Error bars showing the variation in DBP across BMI percentiles among Africans and Europeans. The middle point indicated the mean DBP, and the error bars indicate the standard error of the mean at the particular BMI percentile. Binned scatter plot showing the Pearson's linear correlation between* ***(g)*** *DBP and waist circumference, and* ***(h)*** *DBP and SBP. The data points are spaced into rectangular bins, and each point is coloured based on logarithm bin size with redder colours indicating a higher number of plots. The colour bar shows the colour scale.*

Previous studies have shown that the SBP, DBP, and pulse rate vary by age and other anthropometric measurements of the individuals[21–24]. Therefore, we evaluated the Pearson's linear correlation between the SBP, DBP, pulse rate and MHR, with other anthropometric measurements (Figure 1d, Figure 1e and Figure 1f). We found a positive correlation between the SBP and DBP of the participants (Figure 1fd).

Genetic loci associated with cardiovascular traits among Europeans and Africans

We collected the GWAS summary statistics for each cardiovascular trait within each ethnic group (see methods sections). The total number of significant variants (including those in linkage disequilibrium) identified for each pulmonary function parameter is shown in Supplementary Figures 2a, 2b and 2c.

Separately, we identified 266 loci significantly associated (p-values $< 5 \times 10^{-8}$ and causal probability > 0.1; see methods section) with DBP in Europeans and none in Africans (Figure 2a). For SBP, we found 279 significant loci in Europeans and none in Africans (Figure 2b). Furthermore, we found 793 significant loci associated with pulse rate in Europeans and 2 in Africans (Figure 2c), whereas for the MHR, we found 19 significant loci in Europeans and one locus in Africans (Supplementary Figure 2). The difference in numbers of significant SNPs is almost certainly due to the sample size differences.

Next, we sought to identify the shared loci among the traits within each ancestral group. We found that the SBP and DBP shared 148 associated loci in Europeans, but none were shared among the African SNPs, and these loci were unique to each ancestral group (Figure 3e, Figure 3f and Supplementary File 1). When we compared the 1,202 loci in Europeans with the 2 loci in Africans significantly associated with the three cardiovascular function parameters, there were no shared loci between the two sets (Figure 3g).

We then compared the minor allele frequency of the 1,204 SNPs (1,202 loci in Europeans and 2 loci in Africans) between Europeans and Africans in the UK Biobank data. We found that 1,163 out of 1,204 SNPs differed significantly in their frequency among the African and Europeans (Supplementary File 2). The top variant whose frequency was significantly higher in European than Africans was rs7750412 (frequency in Europeans = 0.80, in Africans = 0.25, Fisher test p-values = $7.9 \times 10^{-}$

323), a *ZBTB9* gene intergenic variant that we found to be significantly associated with the pulse rate in Europeans (see Supplementary File 2 and Supplementary Figure 2). Conversely, several other variants, for example, rs9388010 (EUR = 0.1, AFR = 0.54, $p = 9.4 \times 10^{-323}$), exhibited significantly higher frequencies in Africans than Europeans (see Supplementary File 2 and Supplementary Figure 2). rs9388010, that was found to be associated with pulse rate in Africans, is an intron variant that falls within the *GJA1* gene.

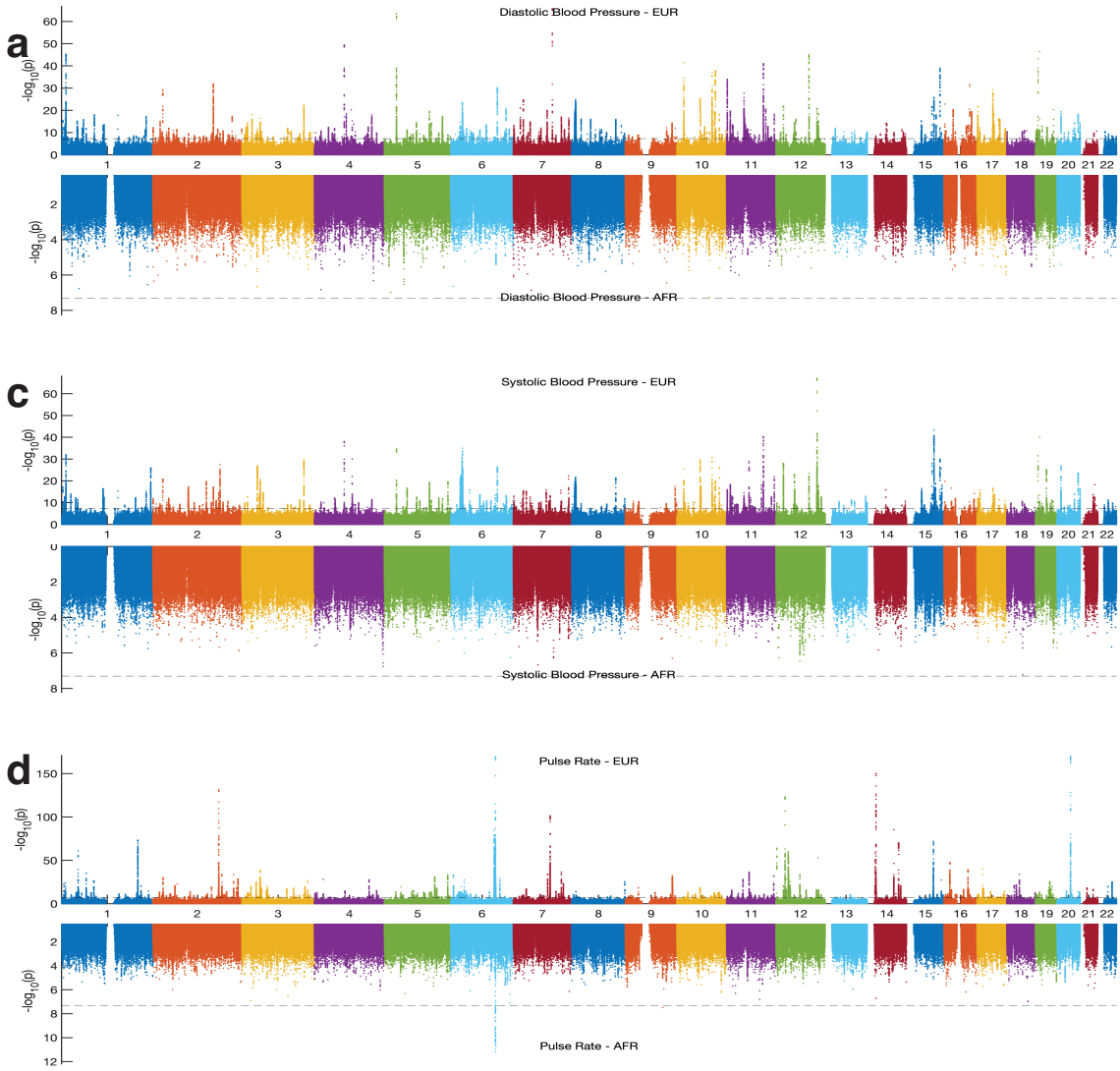

Figure 2: *Manhattan plots of the variant associated with* ***(a)*** *DBP,* ***(b)*** *SBP, and* ***(c)*** *pulse rate in Europeans and African for each chromosome.*

Next, we sought to identify the shared loci among the traits within each ancestral group. We found that the SBP and DBP shared 148 associated loci in Europeans, but none were shared among the African SNPs, and these loci were unique to each ancestral group (Figure 3e, Figure 3f and Supplementary File 1). When we compared the 1,202 loci in Europeans with the 2 loci in Africans significantly associated with the three cardiovascular function parameters, there were no shared loci between the two sets (Figure 3g).

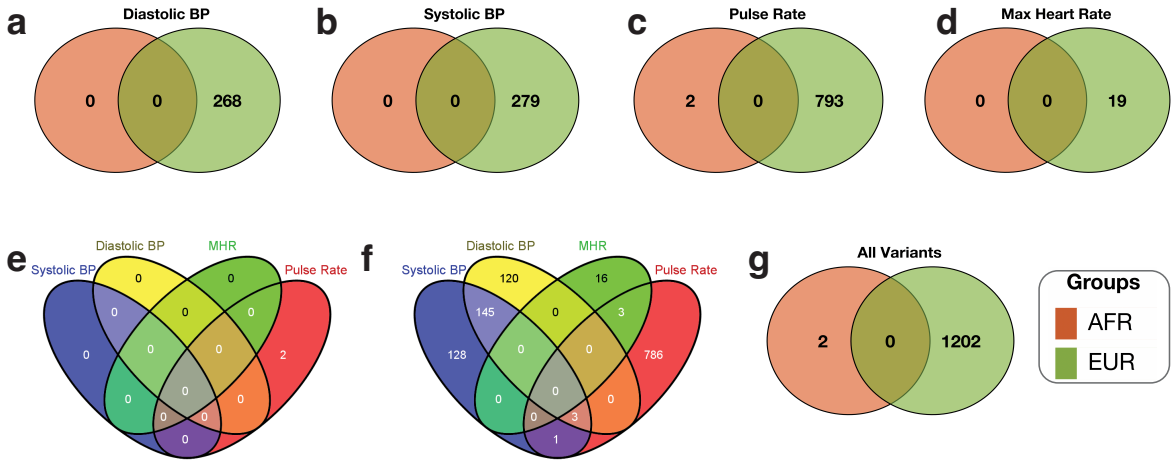

Figure 3: *Venn diagrams showing the overlap among the significant variants associated with* ***(a)*** *DBP,* ***(b)*** *SBP,* ***(c)*** *pulse rate, and* ***(d)*** *MHR in Africans and Europeans. The distribution of genetic variants associated with four cardiovascular traits in* ***(e)*** *Africans and* ***(f)*** *Europeans.* ***(g)*** *Distribution of all variation in Africans and Europeans. Refer to Supplementary File 1 for details concerning individual variants and their frequencies among Africans and Europeans.*

These analyses revealed the genetic loci associated with cardiovascular function among Africans and Europeans and the variations in the minor allele frequencies among these two ancestry groups.

Integrative pathway and phenotypes enrichment analyses

We mapped the SNPs onto genes and extracted those predicted to be associated (using a cutoff p-value of 1×10^{-6} due to the small number of hits in Africans) with cardiovascular function in Europeans and Africans. These gene lists were used to identify the enriched phenotypes, GWAS catalog[25] annotations terms and Elsevier Pathways[26] in Africans or Europeans based on prior knowledge from databases (see Methods section).

The GWAS catalogue term enrichment analyses for the loci in Europeans were associated with "Systolic blood pressure" ($p = 4.2 \times 10^{-124}$), "Diastolic blood pressure" ($p = 1.5 \times 10^{-91}$), "Pulse pressure" ($p = 3.7 \times 10^{-76}$), among other terms related to cardiovascular function (Figure 4a, Supplementary File 3). On the other hand, in Africans, we found that the genes were significantly enriched for other relevant GWAS terms, including "Resting heart rate" ($p = 0.002$) and "Heart rate" ($p = 0.002$), see Figure 4a, Supplementary File 3.

The Elsevier pathway analyses showed the strongest evidence that the genes within which significant loci are located are involved in cardiovascular function in Europeans. The genes were significantly enriched for "Proteins Involved in Dilated Cardiomyopathy" ($p = 1.7 \times 10^{-07}$), "Alcoholic Dilated Cardiomyopathy Model" ($p = 2.4 \times 10^{-06}$), "Cardiomyocyte Dysfunction" ($p = 1.0 \times 10^{-05}$), and other pathways related to cardiovascular pathophysiology (Figure 4b). In African ancestry individuals, we found that the associated genes (mapped based on the suggestive p-value cut-off point) were significantly enriched for pathways associated with "Gap Junction Assembly" ($p = 7.6 \times 10^{-4}$), "Proteins Involved in Pathogenesis of Arrhythmogenic Right Ventricular Cardiomyopathy" ($p = 8.4 \times 10^{-4}$), "Leukotriene

Effect on Vascular Endothelial Cell Response" ($p = 2.2 \times 10^{-4}$), among others. Gap junctions in the heart provide the pathways for intercellular current flow, enabling coordinated action potential propagation [27].

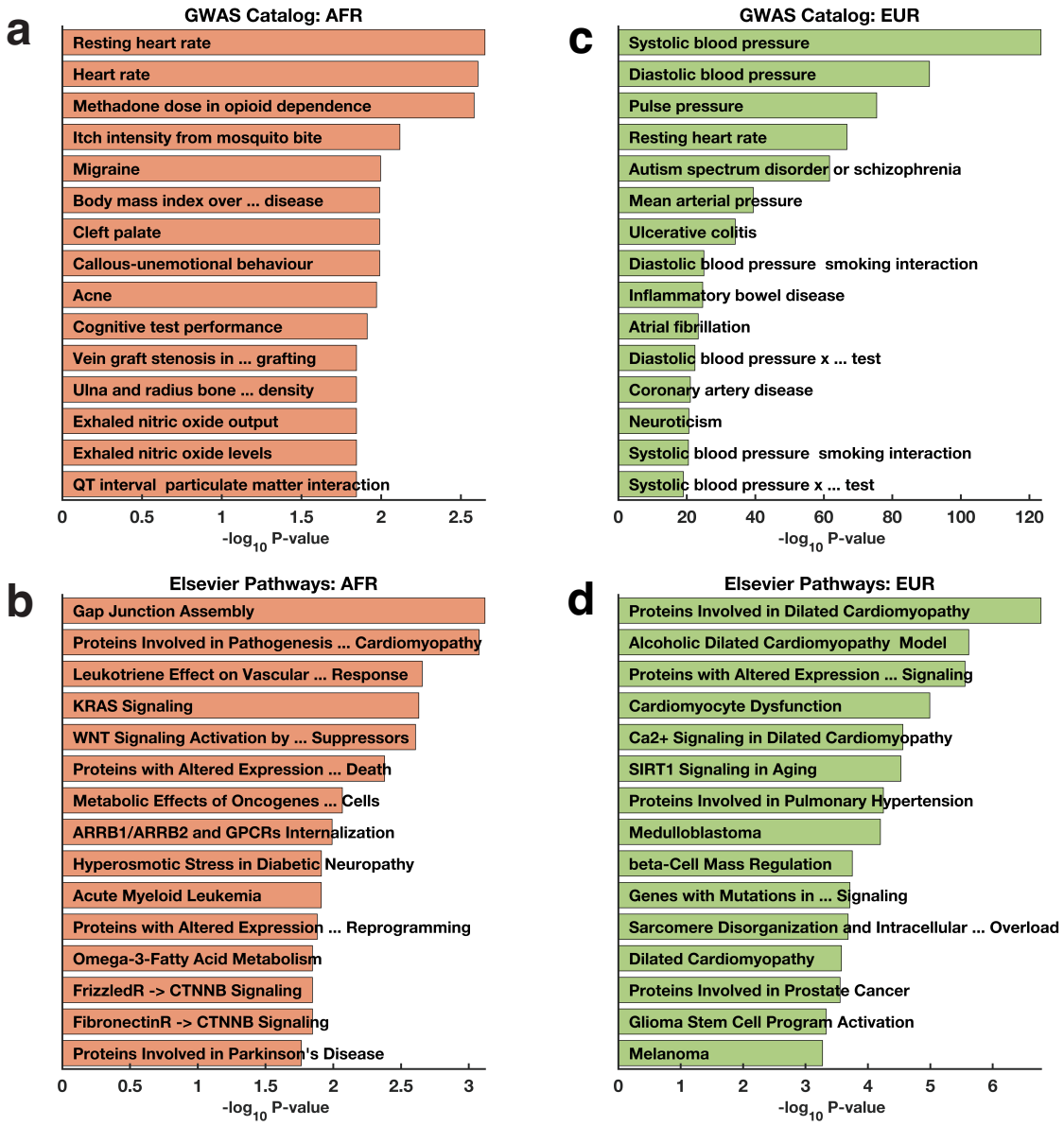

Figure 4: *Enrichment analysis plots of* ***(a)*** *GWAS catalog terms,* ***(b)*** *KEGG pathways,* ***(c)*** *WikiPathways, and* ***(d)*** *Mouse Genotype Ontology terms enriched for genes in which significant variants are located. Refer to Supplementary File 2 for the complete list of integrative enrichment results. Also, see Supplementary Figure 3.*

Overall, our findings demonstrate that while the variants we identified among Europeans and Africans are different, the different genes in which the SNPs are located are mostly involved in cardiovascular pathophysiology or related functions (see Supplementary Figure 4).

Causal variants spanning loci associated with cardiovascular traits

We identified several causal variants (see methods section) spanning various genomic loci. Here, among the several significant variants associated with DBP in Europeans at loci 6p22.3 (see Figure 1a), we found the that most likely causal variant was rs196023 (Probabilistic Identification of Causal SNPs [PICS] value = 1), an intron variant and non-coding variant located in the gene *CASC15* (Figure 5a and Supplementary File 4). This variant rs196023 has not previously been linked to DBP, however, the gene which it maps to, *CASC15*, is associated with various cardiovascular traits, including DBP[28] and SBP[28–30] and cardiovascular disease[28,31,32].

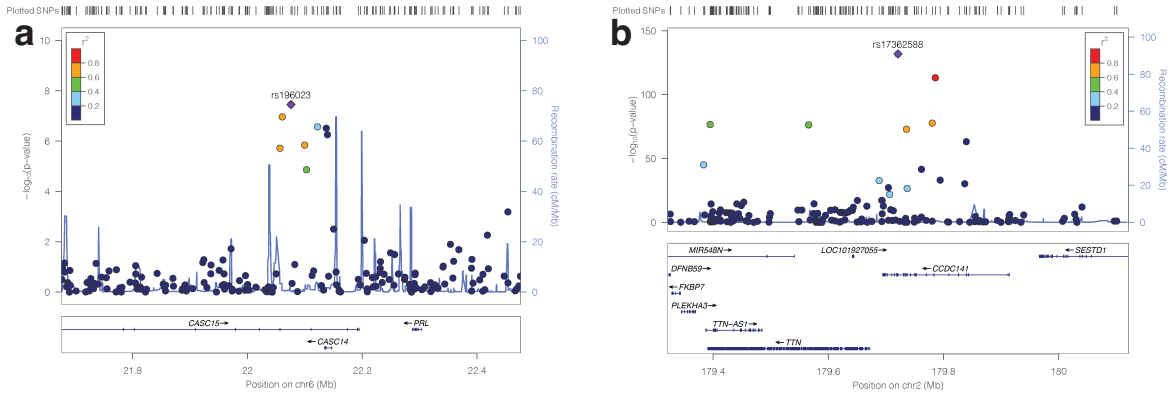

Figure 5: *Regional association plots for genome-wide significant genomic loci for the lead SNPs* ***(a)*** *rs196023 at loci 6p22.3* ***(b)*** *rs17362588 at 2q31.2. The genes within the chromosomal loci are shown in the lower panel. The blue line indicates the recombination rate. The filled circles show the position of the SNPs along the region on the x-axis and the negative logarithm of the association p-value on the y-axis. The lead SNP is shown in purple, and the SNPs within the locus are coloured based on the linkage disequilibrium correlation value (r^2) with the lead SNP based on the European HapMap haplotype from the 1000 genome project.*

In addition, at locus 2q31.2, the lead SNPs among several associated with pulse rate and maximum heart rate was rs17362588 (PICS pulse rate = 0.994) located in the *CCDC141* genes (Figure 5b). This variant has previously been linked to cardiovascular traits, including Pulse pressure[33] and DBP[33], 'Heart rate response to recovery post-exercise (10 sec)' [34], and Heart rate[35]. Refer to Supplementary File 4 for the complete list of variants and their predicted causal probabilities.

Novel variants associated with cardiovascular function

Next, we aimed to identify the novel and known variants among the predicted causal variants associated with cardiovascular function.

We found two novel causal variants, rs9388010 and rs750086835 associated with cardiovascular traits in Africans (Figure 6). rs9388010 is an intergenic variant that falls within the *GJA1* gene that encodes a gap junction protein known as connexin 43. Interestingly, connexin 43 has previously been implicated in several cardiomyopathies including, hypoplastic left heart syndrome, atrioventricular septal defect, cardiac arrhythmias, atrioventricular septal defect, coronary aneurysm, hypertension, and several other cardiomyopathies [27,36–41]. rs750086835 is an intergenic variant that falls within the *SEC61B* gene which has been linked to the

development of atherosclerosis which is strongly associated with cardiovascular function[42].

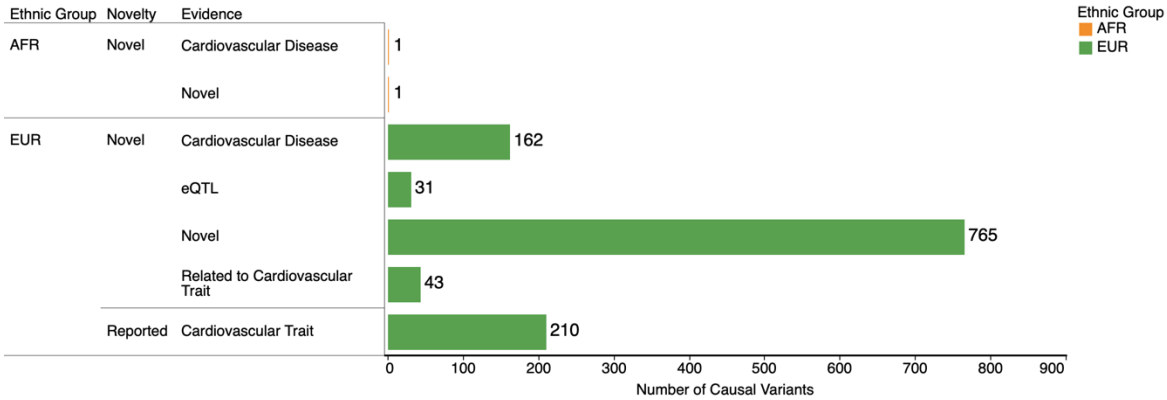

Figure 6: *Novelty of variant: breakdown of the known and novel variants in African and Europeans.*

In Europeans, we found that 210 of the 1,202 causal variants have previously been associated at least one cardiovascular trait (Figure 5). These include rs10260816[43], rs1050288[44], and rs10789207[45]. Refer to Supplementary File 4 for the complete list of variants, associated traits, and publications reporting the variants. Furthermore, we found 992 novel variants in Europeans, among which 43 are reported associated with traits correlated with cardiovascular function (e.g., BMI, Weight, and Age, See Figure 1), and 162 variants reported associated with cardiovascular disease. Interestingly, we found that 31 of the novel variants were significant eQTLs in cardiovascular tissues, including the heart and major blood vessels (refer to Supplementary File 4 for details on each of these variants).

Discussion

We used the UK Biobank dataset to assess variations in cardiovascular traits and the associated genomic loci in individuals of African and European ancestry. Here, we report differences between Africans and Europeans in the DBP, pulse rate, and MHR. These findings are similar to those reported elsewhere [33,46,47]. However, the analyses of these previous studies were limited to smaller cohorts and the GWA models accounted for only a few covariates.

We found different variants associated with cardiovascular traits among Africans and Europeans using GWAS analysis. Differences in genetic variants associated with various traits in different ancestry groups have been reported for nonalcoholic fatty liver disease[13], white blood cell count [15], and others[48–51]. Here, we identified only 2 significant novel variants in the African ancestry group, primarily because of the smaller sample size of 5,978 (compared with the European group of 383,471 individuals), which would likely be underpowered and result in the analysis missing many associations. Some of these associations may also differ among the two ancestry groups[52,53]. Therefore, we look forward to future studies comparing the

genetic association of cardiovascular traits among Europeans and Africans with equally as many African individuals.

We found that the frequency of 1,163 out of 1,204 (96.6%) variants associated with at least one cardiovascular trait significantly varies among Africans and Europeans; a finding that may explain why different variants are associated with cardiovascular function in these two groups[52]. The frequency difference of 96.6% observed for the variants associated with cardiovascular traits is higher than the overall background differences in frequency of 89% that we observed between the populations. Of note, we discovered that the variant rs9388010 (associated with pulse rate in Africans) is present in 53.9% of Africans and 9.8% of European. Conversely, the variant rs7750412, associated with pulse rate in Europeans, is present at a frequency of 80% in Europeans and 25% in Africans. Several phenotypes and diseases are driven by the frequency variants in a particular ancestry group[52,54–57]. For example, serum uric acid levels[14], fatty acid desaturases[16], and other phenotypes[17–19] are associated with different alleles among Africans and Europeans that are sometimes located on the same gene but whose frequencies vary between ethnic groups.

Our integrative enrichment analyses unearthed several pathways and phenotypes related to cardiovascular traits for both sets of variants separately in Africans and Europeans. Notably, the Elsevier pathway analyses show that the genes within which the variants are located are involved in gap junction assembly and other cardiovascular related phenotypes in African ancestry individuals (see Figure 4). Also, in Europeans, genes within which the variants are located are enriched for various cardiovascular related phenotypes. Therefore, our results have identified high confidence variants associated with cardiovascular traits in Africans and Europeans. Intriguingly, only 6 genes (82 SNPs; based on the association cut-off of 10^{-6}) were common between the Europeans and Africans, which suggests that different genes may contribute significantly to the cardiovascular traits among the two ancestral groups (see Figure S4).

Overall, we have shown that mean values of cardiovascular parameters, including DBP, pulse rate, and MHR traits, differ significantly among Africans and Europeans. Furthermore, we have revealed that these cardiovascular traits are principally associated with different genetic variants in these population groups. Encouragingly, the genes in both analyses map onto numerous pathways and GWAS catalog phenotypes linked to cardiovascular function in physiology and disease.

Methods

We analysed datasets from the UK Biobank[20] of genotyping array data, SBP, DBP, pulse rate, and MHR, and other anthropometric measurements. We utilised a subset of these datasets to include 383,471 Europeans (individuals labelled as White,

British, Irish, and "any other White background") and 5,978 individuals of African ancestry (labelled as Blacks). Then, we calculated the allele frequencies in Africans and European from the UK Biobank array data. Details of the study, UK Biobank participants, recruitment method to the study, the sample collection and analysis protocols are expansively described elsewhere[20,58].

Comparison of cardiovascular traits in Europeans and Africans

We compared the mean values of the SBP, DBP, pulse rate and MHR between the 383,471 Europeans and 5,978 Africans using the student t-test with the unequal variance assumption. Furthermore, we calculated the Pearson's linear correlation coefficient to evaluate the correlation of SBP, DBP, pulse rate, and MHR with other anthropometric measurements. Finally, to assess the extent to which the participant's BMI influence the DBP, we plotted the trend with error bars of the DBP over each 10th percentile bin.

Genome-wide identification of genetic loci associated with cardiovascular traits

The methods applied for genotyping of participants in the UK Biobank are reported elsewhere [20,58]. We obtained the GWAS summary statistics computed by the UK Biobank project of each cardiovascular function parameter on 7th November 2021 at https://pan.ukbb.broadinstitute.org/downloads. The methods used to perform the GWA analyses are also described elsewhere [59]. Briefly, the GWAS was performed for the cardiovascular phenotypes and ancestry groups using the Scalable and Accurate Implementation of GEneralized mixed model approach[60], using a linear or mixed logistic model including a kinship matrix as a random effect and covariates as fixed effects. The covariates included the participant's age, sex, age multiplied by sex, the square of the age, the square of the age multiplied by the sex, and the first 10 principal components calculated from the genotype datasets. Variants were considered statistically significantly associated with a trait when the GWAS p-value was less than 5×10^{-8}. In this study, the variants were not evaluated for replication. The Manhattan plots were produced in MATLAB using the software described here[61]. Furthermore, fine mapping of SNPs to identify the most credible causal SNPs within each linkage disequilibrium block conditioning on the lead SNP signal in each locus ±50 kb was done using the PICS2 software[62].

Integrative Enrichment Analyses

We used the chromosome positions of the suggestively associated variants (cut off p-value point of 1×10^{-6}) and the hg19 genome coordinates to map the variants onto genes. Then, separately for African and Europeans, we used the gene lists to perform enrichment analyses[26 63] using annotations from various databases. To identify the genes' pathways, we used the Elsevier pathways[26]. Furthermore, to identify the GWAS traits associated with the genes, we used the annotation from the GWAS catalog[25]. Finally, to identify the phenotypes enriched for in our genes lists,

we use information from the MGI phenotypes[64], Human Phenotype Ontology[65], and OMIM database[66].

Identifying novel of eQTLs variants

We obtained data on all previously published GWAS studies from the GWAS catalog[25]. We then filtered these data to return only GWAS studies that report variants associated with cardiovascular traits. Next, we used this filtered information to determine which among the variants identified in our study are previously reported as being associated with cardiovascular traits. We used the subset of the GWAS catalog data of studies unrelated to cardiovascular function to determine previously known variants associated with any other traits.

In addition, we retrieved information for variants that are eQTLs (expression quantitative trait loci) in various cardiovascular tissues, including the heart, ventricles, atrium and the major blood vessels, from the GTEx consortium database[67]. Then, we used this information to identify which variants in our GWAS significant set are eQTLs.

Statistics and Reproducibility

We performed the statistical analyses using various software, including R programming language, Python, MATLAB 2021a and Bash. We considered all statistical tests significant if the calculated two-sided p-value was < 0.05 for single comparisons. In addition, we corrected for multiple hypotheses testing using the Benjamini & Hochberg procedure[68].

Data Availability

The datasets that support results presented in this manuscript are available from: the UK Biobank; https://www.ukbiobank.ac.uk, dbSNP; https://www.ncbi.nlm.nih.gov/snp, and the GWAS catalogue; https://www.ebi.ac.uk/gwas.

Code Availability

The code to reproduce most of the results presented here is available from the corresponding authors upon request.

Acknowledgements

This research has been conducted using the UK Biobank Resource under Application Number 53163. The funding for this project was provided by H3ABioNet, supported by the National Institutes of Health Common Fund under grant number U24HG006941. The content of this publication is solely the authors' responsibility and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions

M.S., S.E., and N.M conceptualized the study. M.S., N.M., S.E., and J.C. designed the methodology, and M.M. M.S, M.M., J.C., and S.E. performed the formal analysis of the data. M.S., N.M., and S.E., drafted manuscript. Editing and reviewing the manuscript was carried out by M.S., N.M., S.E., J.C., and M.M. Data visualisations were produced by M.S. and S.E.

Competing interests

The authors declare that they have no competing interests

References

1. Mohlke, K. L., Boehnke, M. & Abecasis, G. R. Metabolic and cardiovascular traits: an abundance of recently identified common genetic variants. *Hum. Mol. Genet.* **17**, R102–R108 (2008).
2. Pilia, G. *et al.* Heritability of Cardiovascular and Personality Traits in 6,148 Sardinians. *PLOS Genet.* **2**, e132 (2006).
3. Ma, R. C. Genetics of cardiovascular and renal complications in diabetes. *J. Diabetes Investig.* **7**, 139–154 (2016).
4. Winham, S. J., de Andrade, M. & Miller, V. M. Genetics of cardiovascular disease: Importance of sex and ethnicity. *Atherosclerosis* **241**, 219–228 (2015).
5. Zhang, X. *et al.* Genetic associations with expression for genes implicated in GWAS studies for atherosclerotic cardiovascular disease and blood phenotypes. *Hum. Mol. Genet.* **23**, 782–795 (2014).
6. Yao, C. *et al.* Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease. *Nat. Commun. 2018 91* **9**, 1–11 (2018).
7. Evangelou, E. *et al.* Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat. Genet. 2018 5010* **50**, 1412–1425 (2018).
8. Georgakis, M. K. *et al.* Genetically Predicted Blood Pressure Across the Lifespan. *Hypertension* 953–961 (2020). doi:10.1161/HYPERTENSIONAHA.120.15136
9. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun. 2018 91* **9**, 1–12 (2018).
10. Liu, C. *et al.* Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. *Nat. Genet. 2016 4810* **48**, 1162–1170 (2016).
11. Pilling, L. C. *et al.* Human longevity is influenced by many genetic variants: evidence from 75,000 UK Biobank participants. *Aging (Albany NY)* **8**, 547 (2016).
12. Wan, E. Y. F. *et al.* Blood Pressure and Risk of Cardiovascular Disease in UK Biobank. *Hypertension* **77**, 367–375 (2021).
13. Palmer, N. D. *et al.* Characterization of european ancestry nonalcoholic fatty liver disease-associated variants in individuals of african and hispanic descent. *Hepatology* **58**, 966–975 (2013).

14. Rule, A. D. *et al.* Association between SLC2A9 transporter gene variants and uric acid phenotypes in African American and white families. *Rheumatology* **50**, 871–878 (2011).
15. Reiner, A. P. *et al.* Genome-Wide Association Study of White Blood Cell Count in 16,388 African Americans: the Continental Origins and Genetic Epidemiology Network (COGENT). *PLOS Genet.* **7**, e1002108 (2011).
16. Buckley, M. T. *et al.* Selection in Europeans on Fatty Acid Desaturases Associated with Dietary Changes. *Mol. Biol. Evol.* **34**, 1307–1318 (2017).
17. Batai, K. *et al.* Common vitamin D pathway gene variants reveal contrasting effects on serum vitamin D levels in African Americans and European Americans. *Hum. Genet. 2014 13311* **133**, 1395–1405 (2014).
18. Mathias, R. A. *et al.* A combined genome-wide linkage and association approach to find susceptibility loci for platelet function phenotypes in European American and African American families with coronary artery disease. *BMC Med. Genomics 2010 31* **3**, 1–11 (2010).
19. Larkin, E. K. *et al.* A Candidate Gene Study of Obstructive Sleep Apnea in European Americans and African Americans. *https://doi.org/10.1164/rccm.201002-0192OC* **182**, 947–953 (2012).
20. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nat. 2018 5627726* **562**, 203–209 (2018).
21. Buchman, A. S. *et al.* Pulmonary function, muscle strength and mortality in old age. *Mech. Ageing Dev.* **129**, 625–631 (2008).
22. Shapira, N. *et al.* Determinants of pulmonary function in patients undergoing coronary bypass operations. *Ann. Thorac. Surg.* **50**, 268–273 (1990).
23. Schoenberg, J. B., Beck, G. J. & Bouhuys, A. Growth and decay of pulmonary function in healthy blacks and whites. *Respir. Physiol.* **33**, 367–393 (1978).
24. Park, J. E., Chung, J. H., Lee, K. H. & Shin, K. C. The Effect of Body Composition on Pulmonary Function. *Tuberc. Respir. Dis. (Seoul).* **72**, 433–440 (2012).
25. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
26. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
27. Jongsma, H. J. & Wilders, R. Gap Junctions in Cardiovascular Disease. *Circ. Res.* **86**, 1193–1197 (2000).
28. TJ, H. *et al.* Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat. Genet.* **49**, 54–64 (2017).
29. HR, W. *et al.* Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat. Genet.* **49**, 403–415 (2017).
30. G, K. *et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am. J. Hum. Genet.* **104**, 65–75 (2019).
31. H, M. *et al.* Transethnic Meta-Analysis of Genome-Wide Association Studies Identifies Three New Loci and Characterizes Population-Specific Differences for Coronary Artery Disease. *Circ. Genomic Precis. Med.* **13**, (2020).
32. S, K. *et al.* Population-specific and trans-ancestry genome-wide analyses

identify distinct and shared genetic risk loci for coronary artery disease. *Nat. Genet.* **52**, 1169–1177 (2020).
33. Giri, A. *et al.* Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat. Genet. 2018 511* **51**, 51–62 (2018).
34. JL, H., P, H. & V, B. In vitro transcription of the Leptomonas seymouri SL RNA and U2 snRNA genes using homologous cell extracts. *Mol. Biochem. Parasitol.* **90**, 183–192 (1997).
35. M, den H. *et al.* Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nat. Genet.* **45**, 621–631 (2013).
36. Salameh, A. *et al.* On the Role of the Gap Junction Protein Cx43 (GJA1) in Human Cardiac Malformations with Fallot-Pathology. A Study on Paediatric Cardiac Specimen. *PLoS One* **9**, (2014).
37. Paznekas, W. A. *et al.* Connexin 43 (GJA1) Mutations Cause the Pleiotropic Phenotype of Oculodentodigital Dysplasia. *Am. J. Hum. Genet.* **72**, 408–418 (2003).
38. Huang, G. Y. *et al.* Evaluating the role of connexin43 in congenital heart disease: Screening for mutations in patients with outflow tract anomalies and the analysis of knock-in mouse models. *J. Cardiovasc. Dis. Res.* **2**, 206–212 (2011).
39. Dupont, E. *et al.* Altered Connexin Expression in Human Congestive Heart Failure. *J. Mol. Cell. Cardiol.* **33**, 359–371 (2001).
40. Michela, P., Velia, V., Aldo, P. & Ada, P. Role of connexin 43 in cardiovascular diseases. *Eur. J. Pharmacol.* **768**, 71–76 (2015).
41. Severs, N. J. *et al.* Remodelling of gap junctions and connexin expression in heart disease. *Biochim. Biophys. Acta - Biomembr.* **1662**, 138–148 (2004).
42. Sivapalaratnam, S. *et al.* Identification of candidate genes linking systemic inflammation to atherosclerosis; Results of a human in vivo LPS infusion study. *BMC Med. Genomics* **4**, 1–8 (2011).
43. Hoffmann, T. J. *et al.* Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat. Genet.* **49**, 54–64 (2017).
44. Warren, H. R. *et al.* Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat. Genet.* **49**, 403–415 (2017).
45. Eppinga, R. N. *et al.* Identification of genomic loci associated with resting heart rate and shared genetic predictors with all-cause mortality. *Nat. Genet.* **48**, 1557–1563 (2016).
46. Fumo, M. T. *et al.* Diurnal Blood Pressure Variation and Cardiac Mass in American Blacks and Whites and South African Blacks. *Am. J. Hypertens.* **5**, 111–116 (1992).
47. Howard, G. *et al.* Racial Differences in the Impact of Elevated Systolic Blood Pressure on Stroke Risk. *JAMA Intern. Med.* **173**, 46–51 (2013).
48. Barnes, K. C. Genomewide association studies in allergy and the influence of ethnicity. *Curr. Opin. Allergy Clin. Immunol.* **10**, 427 (2010).
49. Chan, S. L., Jin, S., Loh, M. & Brunham, L. R. Progress in understanding the genomic basis for adverse drug reactions: a comprehensive review and focus on the role of ethnicity. *http://dx.doi.org/10.2217/PGS.15.54* **16**, 1161–1178 (2015).

50. Ueta, M. *et al.* Genome-wide association study using the ethnicity-specific Japonica array: identification of new susceptibility loci for cold medicine-related Stevens–Johnson syndrome with severe ocular complications. *J. Hum. Genet. 2017 624* **62**, 485–489 (2017).
51. Jorgenson, E. *et al.* Genetic contributors to variation in alcohol consumption vary by race/ethnicity in a large multi-ethnic genome-wide association study. *Mol. Psychiatry 2017 229* **22**, 1359–1367 (2017).
52. Asif, H. *et al.* GWAS significance thresholds for deep phenotyping studies can depend upon minor allele frequencies and sample size. *Mol. Psychiatry 2020* 1–8 (2020). doi:10.1038/s41380-020-0670-3
53. Ball, R. D. Designing a GWAS: Power, Sample Size, and Data Structure. *Methods Mol. Biol.* **1019**, 37–98 (2013).
54. Emison, E. S. *et al.* Differential Contributions of Rare and Common, Coding and Noncoding Ret Mutations to Multifactorial Hirschsprung Disease Liability. *Am. J. Hum. Genet.* **87**, 60–74 (2010).
55. Witte, J. S., Visscher, P. M. & Wray, N. R. The contribution of genetic variants to disease depends on the ruler. *Nat. Rev. Genet. 2014 1511* **15**, 765–776 (2014).
56. Katsanis, N. The continuum of causality in human genetic disorders. *Genome Biol. 2016 171* **17**, 1–5 (2016).
57. Minikel, E. V. *et al.* Quantifying prion disease penetrance using large population control cohorts. *Sci. Transl. Med.* **8**, (2016).
58. Bycroft, C. *et al.* Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv* 166298 (2017). doi:10.1101/166298
59. Team, P.-U. Quality Control (QC) | Pan UKBB. Available at: https://pan-dev.ukbb.broadinstitute.org/docs/qc/index.html. (Accessed: 7th November 2021)
60. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet. 2018 509* **50**, 1335–1341 (2018).
61. Green, H. Manhattan Plots for visualisation of GWAS results - File Exchange - MATLAB Central. Available at: https://www.mathworks.com/matlabcentral/fileexchange/69549-manhattan-plots-for-visualisation-of-gwas-results?s_tid=srchtitle. (Accessed: 4th September 2021)
62. Taylor, K. E., Ansel, K. M., Marson, A., Criswell, L. A. & Farh, K. K.-H. PICS2: next-generation fine mapping via probabilistic identification of causal SNPs. *Bioinformatics* (2021). doi:10.1093/BIOINFORMATICS/BTAB122
63. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–50 (2005).
64. Smith, C. L. & Eppig, J. T. The mammalian phenotype ontology: enabling robust annotation and comparative analysis. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **1**, 390–9 (2009).
65. Köhler, S. *et al.* The Human Phenotype Ontology in 2017. *Nucleic Acids Res.* **45**, D865–D876 (2017).
66. Amberger, J. S., Bocchini, C. A., Scott, A. F. & Hamosh, A. OMIM.org: leveraging knowledge across phenotype–gene relationships. *Nucleic Acids*

Res. **47**, D1038–D1043 (2019).
67. Ardlie, K. G. *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science (80-.).* **348**, 648–660 (2015).
68. Benjamini, Y. Discovering the false discovery rate. *J. R. Stat. Soc. Ser. B (Statistical Methodol.* **72**, 405–416 (2010).