



OPEN

A genome-wide association and polygenic risk score study on abnormal electrocardiogram in a Chinese population

Mengqiao Wang^{1✉}, Jiaqi Gao¹, Yang Shi^{2✉} & Xing Zhao^{1✉}

Electrocardiography is a common and widely-performed medical examination based on the measurement and evaluation of electrocardiogram (ECG) to assess the up-to-date cardiac rhythms and thus suggest the health conditions of cardiovascular system and on a larger level the individual's wellness. Abnormal ECG assessment from the detection of abnormal heart rhythms may have clinical implications including blood clots in formation, ongoing heart attack, coronary artery blockage, etc. Past genetic-phenotypic research focused primarily on the physical parameters of ECG but not the medical evaluation. To unbiasedly uncover the underlying links of genetic variants with normal vs. abnormal ECG assessment, a genome-wide association study (GWAS) is carried out in a 1006-participant cohort of Chinese population effectively genotyped for 243487 single nucleotide polymorphisms (SNPs). Both age and sex are influential factors, and six novel SNPs are identified for potential association with abnormal ECG. With the selected SNPs, a polygenic risk score (PRS) differentiates the case-control subgroups, and correlates well with increased risk of abnormal ECG. The findings are reproduced in an independent validation cohort. The derived PRS may function as a potential biomarker for prospectively screening the high-risk subgroup of heart issues in the Chinese population.

Electrocardiogram (abbreviated as ECG or EKG) is a non-invasive test of the electrical activity of cardiac rhythms that may indicate normal or abnormal cardiovascular conditions including cardiac hypertrophy, arrhythmias, conduction defects etc. ECG tracings consist of three major components during a single heartbeat: the P wave (generally upwards and indicating atrial depolarization)¹, the QRS complex (of greater amplitude and indicating ventricular depolarization)², and the T wave (smaller upwards wave than the QRS and indicating ventricular repolarization)³. Additionally, the PR reflects the time the electrical impulse takes to travel from the sinus node through the atrioventricular node. Abnormal rhythms exist as abnormally slow or fast heart rate or an irregular cardiac rhythm. An abnormal P-wave may serve as predictors for atrial arrhythmias and atrial enlargement⁴. A premature ventricular contraction or a ventricular rhythm can be revealed as a prolonged QRS duration which is associated with mortality in both the general population and in patients with cardiac diseases⁵⁻⁹. T waves can be helpful in a variety of pathologies: as an example, tall T waves in anterior chest lead III, aVR, and V1 with a negative QRS complex may suggest acute myocardial ischemia¹⁰. QT interval measures myocardial depolarization and repolarization time, while shortening or prolongation of QT interval is associated with an increased risk for arrhythmias and sudden cardiac death (SCD)^{11,12}. Abnormal PR interval is a risk factor for atrial fibrillation and heart block¹³. Through ECG examination and the following evaluation by certified medical staffs (e.g. assessment of rhythm, range of heart rate, measurement of tracing intervals and segments), abnormality on cardiac rhythms could be detected early. Therefore, normal or abnormal ECG assessment is a key medical and clinical evaluation that provides the doctors with instrumental evidence-based revelation on individuals' health conditions.

Both genetic and environmental factors contribute to abnormal cardiac rhythms, and regarding the mode of inheritance for such a complex trait, it is highly likely that a collection of rather than a single genetic variant may be involved. Genome-wide association study, or GWAS, aims to genotype the genome of a population in

¹Department of Epidemiology and Biostatistics, West China School of Public Health and West China Fourth Hospital, Sichuan University, Renmin South Road 16, Chengdu 610041, Sichuan, People's Republic of China. ²Department of Population Health Science, Medical College of Georgia, Augusta University, 1120 15th Street, Augusta, GA 30912, USA. ✉email: mengqiaowang@gmail.com; yshi@augusta.edu; zhaoxing731@gmail.com

high-throughput manner and unbiasedly screen for loci significantly associated with the phenotype-of-interest. A few pioneering GWASs for various ECG phenotypes have found over 500 risk variants¹⁴, validating the underlying genetic basis^{15–19}. The QT interval was one of the first traits investigated at the outset of the GWAS era²⁰. Heritability estimates of QT interval are between 30 and 40%²¹. Previous GWAS identified 45 loci, of which 35 have been proved in QT-IGC, containing variants in genes that modestly influence QT interval, such as *GBF1*, *AZINI*, *CREBBP*, and *KCNQ1*^{22,23}. QRS interval duration is a quantitative trait influenced by multiple genetic factors and as well by both age and gender^{24,25}. In addition, the heritability of QRS duration is estimated to be 35–55% from twin and family studies^{26–29}. An exome-chip meta-analysis has identified 34 single nucleotide polymorphisms (SNPs) associated with QRS duration³⁰, confirming 20 out of the 29 previously identified loci, including *KLHL38*^{31,32}, *DLEC1*³³, *NACA*³⁴, and *ADAMTS6*³⁵. However, there were so far few GWASs with the qualitative assessment of ECG results as the outcome so it remains unclear if any genetic variants may be linked to abnormal ECG status.

Moreover, like many biometric parameters, ECG features are known to vary by ethnicity^{36,37}. Most of the published GWASs for ECG traits have been carried out in populations of European descent^{13,38–40}, but not in populations residing in populous regions of high incidence in cardiovascular diseases such as China. In-parallel GWASs in different ethnic groups are informative in both validating common biomarkers and discovering ethnicity-specific genetic links. To our knowledge, no large-sample GWAS on the assessment of ECG has been conducted for a cohort of Chinese population, and this study aims to fill this gap to explore the association of any known or novel SNPs with normal vs. abnormal ECG assessment in a Chinese cohort.

Methods

Study design and procedure. A group of adult residents living in Sichuan Province of P.R. China were recruited as a regional cohort and among them, 1006 individuals participated in the project studying the association between genomic polymorphisms and the evaluation of electrocardiogram status. For ECG measurement, individual participant was informed of the standard procedure in advance and was given a grace period of 20 min to calm down before measurement by certified medical staffs using a portable electrocardiogram machine that measures and records heart activity (setting at 1 cm/mV). A conventional 12-lead electrocardiogram (I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6) measures the overall magnitude and direction of the individuals' electrical depolarization throughout the cardiac cycle. The ECG results were assessed by a group of certified ECG medical practitioners. For each individual's ECG, the certified staffs were responsible for evaluating the examination to label any identifiable issues (multiple issues allowed for one ECG). The ECG issues include: sinus bradycardia (heart rate < 60 beats/min), sinus tachycardia (heart rate > 100 beats/min), sinus arrhythmia (irregular timing between successive P waves), T wave change (T wave deviates from normal and appears as flat, oddly-shaped, or inverted), high voltage of the left ventricle ($RV5 + SV1 > 4.0$ mV for males and > 3.5 mV for females) etc. After ECG evaluation was completed for all cohort members, each individual was categorized into a binary phenotype of normal vs. abnormal ECG based on the presence or absence of labelled ECG issues. The "normal" refers to the group with none of the labelled issues (issue count = 0), and the "abnormal" covers individuals who display at least one out of the various subtypes of ECG issues (issue count ≥ 1). It should be noted that as the phenotype-of-interest, abnormal ECG in this study refers to the feature assessment regarding general ECG issues but does not directly infer clinical abnormality (as certain ECG issues could be non-specific and have no clinical implications).

Related health and personal information of the participants were collected through physical examination and face-to-face survey by trained staffs. For biological samples, participants spitted 1–2 ml of saliva into sterilized sample tubes, which were transported back to a genetic laboratory within 24 h and subjected to DNA purification. Upon quality validation, DNA samples were genotyped on a high-throughput Affymetrix Axiom Precision Medicine Array chip on a GeneTitan Multi-Channel Instrument platform (Thermo Fisher Scientific). The methods were carried out in accordance with the relevant guidelines and regulations.

Data analysis. A dataset of cohort individuals with their categorized ECG status, genotypes, and related personal information (sex, age etc.) was assembled by the completion of on-site ECG measurements and lab experiments. Only SNPs on autosomes were selected. Genetic dataset is pre-processed for quality checks at both individual and SNP levels. Identity-by-descent (IBD) analysis iteratively disqualifies individuals exceeding a kinship threshold to remove cryptic relatedness. Hardy–Weinberg equilibrium (HWE) test excludes potential population substructure or genotyping errors. Details of the data pre-processing steps were annotated (Table S1). Principal component analysis (PCA) on SNPs in linkage equilibrium leads to principal components (PCs) essential for the adjustment of remaining population substructure (the top ten PCs were included in statistical models as covariates for adjustment). For GWAS, the model of multivariate logistic regression was fitted with the binary ECG status phenotype as the dependent variable, and the genotype and covariates (sex, age, PCs) as the independent variables. Genotypes were coded as the copies of minor allele, and accordingly the additive effect model was selected. A stringent Bonferroni-corrected threshold of 2×10^{-7} (significance level of 0.05 adjusted for about 250,000 SNPs; 243,487 SNPs analyzed for GWAS) and a less stringent candidate threshold of 2×10^{-5} were used exclusively for screening hit and candidate SNPs. Statistical analysis and data visualization were conducted in version 4.0.2 of the R statistical environment (R core team, 2020); parallel computing in GWAS was run on a Windows computer with an Intel CPU (32G RAM, 3.6 GHz, 8 cores).

Ethics approval. All participants in this study have signed the Informed Consent Form before participation to allow their data for scientific research purposes only. The health and genetic data from measurements were subjected to strict anonymization for the respect and protection of personal privacy. In observance of the Regulatory Articles of Human Genetic Resources of P.R. China, all work of genotyping and data analysis were

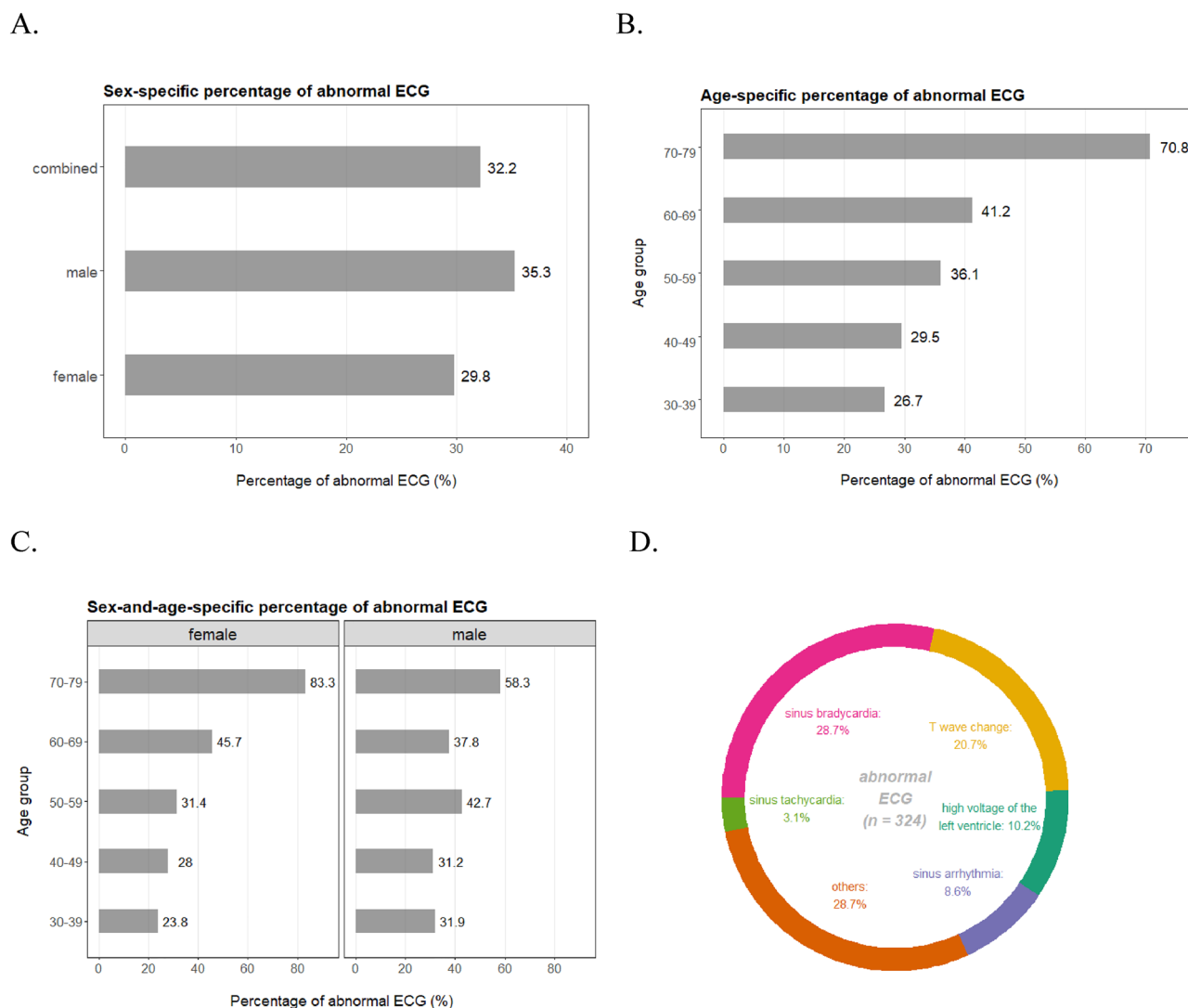


Figure 1. Percentage of abnormal ECG assessment by sex (A), age (B), sex and age (C), and subtypes of ECG issues (D) in the cohort.

conducted in China, and individuals' genetic data could not be made accessible without official authorization and approval from the supervisory agencies. The study design and procedure were approved by the Research Ethics Committee at West China School of Public Health, Sichuan University.

Results

With regard to the sex ratio, 439 (43.6%) of the participants are males while the other 567 (56.4%) are females so the cohort is slightly unbalanced towards the females (two-sample proportion test $P = 6.2 \times 10^{-5}$). The median age of the cohort is 46 years-old, and there is no apparent difference in age distribution regarding sex (Figure S1) or the ECG phenotype (Figure S2). For the outcome-of-interest of professional ECG assessment, 682 (67.8%) participants are evaluated as normal while the other 324 (32.2%) individuals are attributed to at least one ECG subtype issue so classified as abnormal. Percentages of abnormal ECG are 35.3% for the males and 29.8% for the females, with the difference being marginally non-significant (two-sample proportion test $P = 0.07$) (Fig. 1A). In comparison, there is apparently an upward trend in abnormal ECG with increasing ages (Cochran-Armitage trend test $P = 7.8 \times 10^{-6}$): notably, 70.8% of the most senior group (70–79 years-old) displays at least one subtype of issues in ECG assessment (Fig. 1B). Such observation is consistent with the general notion that as people age, their cardiac functions would gradually degenerate. The age-dependent sequential trends are well observed in the females but to a less extent in the males (Fig. 1C). It is also of interest that for abnormal ECG, females display lower percentage than males during middle-ages (30–59 years-old) but the contrast is reversed for senior ages (60–79 years-old). Nevertheless, sex and age are influential factors associated with the ECG results, and thus would be included as covariates for adjustment in the following analysis of GWAS. Among the detailed composition of ECG issues, the top three types combined account for around 60% of all cases and are respectively sinus bradycardia (28.7%), T wave change (20.7%), and high-voltage of the left ventricle (10.2%) (Fig. 1D).

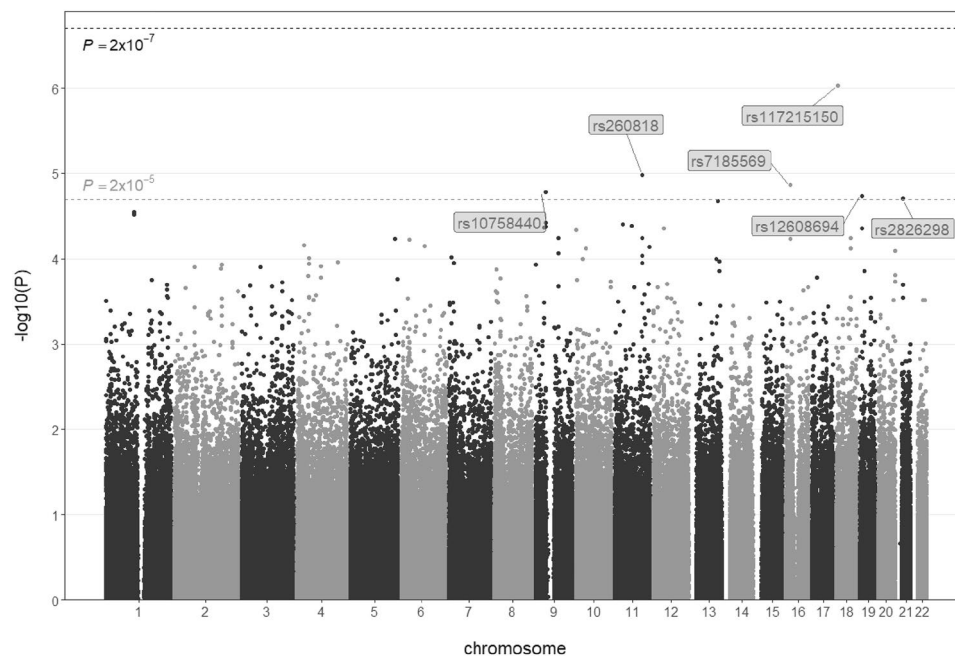


Figure 2. Manhattan plot of GWAS on the phenotype of normal vs. abnormal ECG. Bonferroni corrected threshold (2×10^{-7}) and less stringent candidate threshold (2×10^{-5}) respectively correspond to levels of 6.7 and 4.7 for $-\log_{10}(P)$.

SNP	Chr. (position)	Gene	Ref/minor allele (%)	HWE	Coefficient	P value	OR	95% CI _{OR}
rs117215150	18 (6717603)	Intergenic*	T/C (8.3%)	0.85	0.868	9.3×10^{-7}	2.38	[1.68, 3.37]
rs260818	11 (103920105)	PDGFD (intron)	A/C (40.7%)	0.21	-0.486	1.0×10^{-5}	0.61	[0.50, 0.76]
rs7185569	16 (19446926)	TMC5 (intron)	T/C (13.8%)	0.41	0.612	1.3×10^{-5}	1.84	[1.40, 2.43]
rs10758440	18 (37573330)	FBXO10 (intron)	A/C (12.9%)	0.91	0.645	1.6×10^{-5}	1.91	[1.42, 2.56]
rs12608694	19 (9361302)	OR7E24 (upstream)	T/C (38.0%)	0.13	-0.464	1.8×10^{-5}	0.63	[0.51, 0.78]
rs2826298	21 (21848839)	Intergenic*	C/T (7.1%)	0.55	0.813	2.0×10^{-5}	2.25	[1.55, 3.27]

Table 1. Summary of candidate SNPs from GWAS. *Chr.* Chromosome, *position* GRCh37 coordinate, *HWE* hypothesis test for Hardy–Weinberg equilibrium, *OR* odds ratio. *rs117215150 lies between LOC100130480 and ARHGAP28; rs2826298 lies between TMPRSS15 and NCAM2.

A number of previous GWASs aimed to investigate the genetic association underlying electrocardiography, and their phenotypes-of-interest are specific quantitative metrics of the ECG such as QT interval, PR interval, QRS-T angle etc. A collection of 27 highly significant associations have been annotated with electrocardiography in the GWAS Catalog database⁴¹ (Table S2). 11 out of the curated SNPs are available on our gene chip, and none of them display any apparent association with ECG abnormality (Table S3). A major reason for this disparity may lie in the different types of measurements for the phenotype: a qualitative normal vs. abnormal ECG assessment in this study compared to quantitative features of ECG in all previous studies.

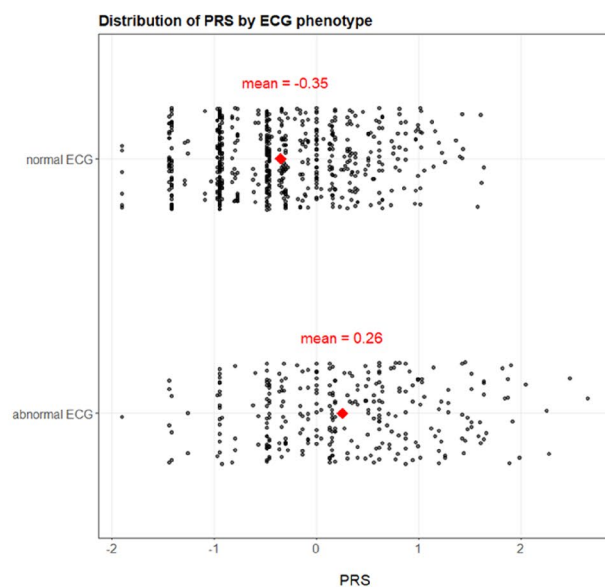
All participants in this study are Chinese of Han ethnicity, and principal component analysis validates the relative homogeneity of the cohort. A GWAS is conducted to map the genetic components linked to the observed abnormal ECG in this cohort. Due to the nature of ECG abnormality as a complex trait and as well the limited sample size in such pilot trial, both a Bonferroni-corrected level of 2×10^{-7} and a 100-times less stringent candidate level of 2×10^{-5} are defined as thresholds of significant and candidate associations. No variant scores high enough above the Bonferroni-corrected level, but a list of six SNPs surpass the candidate level (Fig. 2; Table 1). The candidate SNPs locate to different chromosomes without leading to any changes to the primary sequence of a gene product. In Hardy–Weinberg equilibrium, four of the six SNPs confer higher risk and the other two contribute to lower risk of abnormal ECG, and the allele frequency is comparable among different ethnicities across the 1000 Genomes⁴² (Figure S4). The candidate SNPs display significant but relatively weak association with abnormal ECG as revealed by the respective parameters of classification performance (Table 2). For the covariates in GWAS, age is significant with positive coefficients, and the females have higher risk compared to the males (Table S4), both in agreement with the above descriptive analysis of the cohort.

With ECG assessment being a complex trait and thus a non-Mendelian phenotype, single SNPs would not, in agreement with our GWAS findings, confer strong predictive power. Instead, combination of the candidate SNPs

SNP	Accuracy	Sensitivity	Specificity	PPV	NPV	F1	AUC	Youden
rs117215150	0.635	0.505	0.696	0.439	0.749	0.470	0.632	0.209
rs260818	0.594	0.599	0.591	0.408	0.758	0.486	0.626	0.210
rs7185569	0.625	0.556	0.658	0.432	0.760	0.487	0.632	0.232
rs10758440	0.611	0.541	0.645	0.421	0.747	0.473	0.626	0.206
rs12608694	0.595	0.590	0.598	0.408	0.757	0.483	0.633	0.210
rs2826298	0.604	0.497	0.654	0.402	0.736	0.444	0.617	0.187

Table 2. Classification performance by candidate SNPs. Parameters of classification performance were calculated using the percentage of abnormal ECG in the cohort as the threshold. *PPV/NPV* positive/negative predictive value, *AUC* area under the ROC curve, *Youden* Youden index.

A.



B.

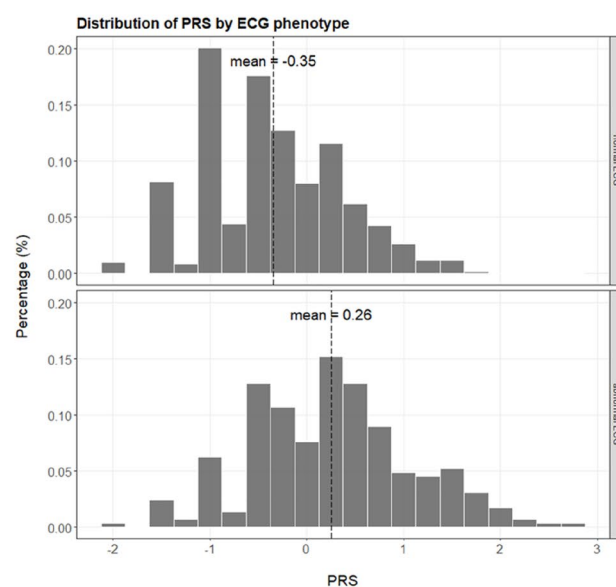
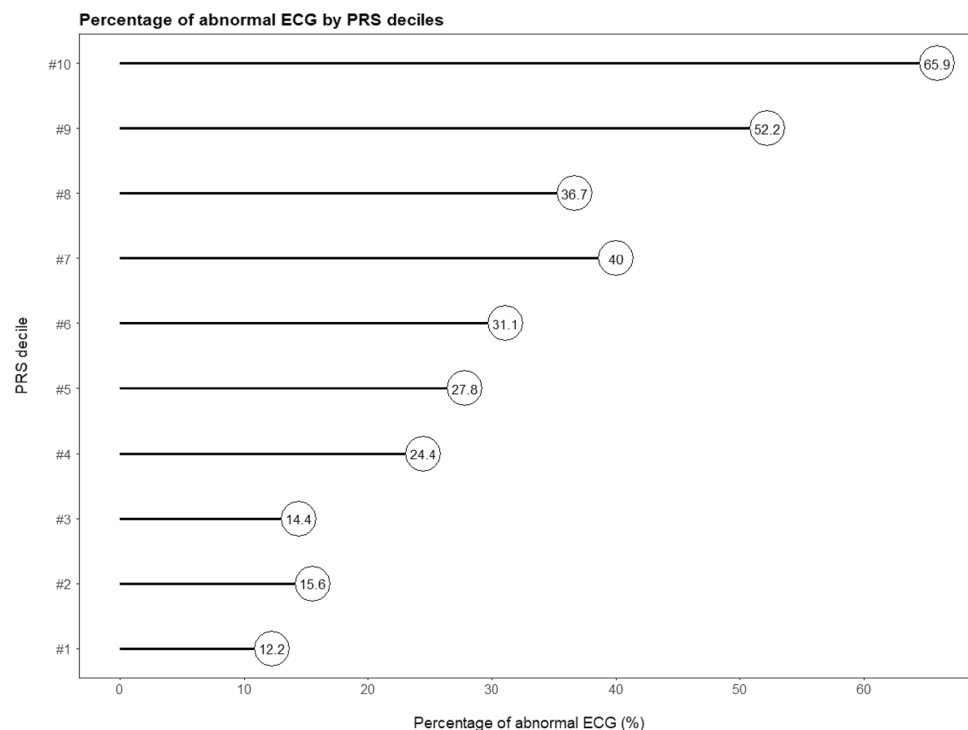


Figure 3. Distribution of PRS by ECG phenotype in the jittered dot plot (A) and the histogram (B).

would result in more informative and robust features, such as polygenic risk scores. We calculate a polygenic risk score (PRS) with the six candidate SNPs weighted by their respective coefficients (Figure S5). The PRS is significantly different in the case–control perspective between normal ECG and abnormal ECG (Wilcoxon rank sum test $P = 2 \times 10^{-16}$, Fig. 3). More convincingly, the PRS displays strong association with abnormal ECG as increasing PRS correlates well with the susceptibility to abnormal ECG (Fig. 4). The highest and lowest deciles of PRS display proportion of abnormal ECG at 65.9% and 12.2% respectively, well differentiating the cohort (abnormal ECG at 32.3%) into the high vs. low risk segments (Fig. 4A). The PRS values also match the observed pattern, and the risk is guaranteed for the few individuals with PRS over 2 (Fig. 4B).

The GWAS and PRS evaluations above are carried out in the same cohort, so there would be potential issues of overfitting. To address such concerns, we recruit an independent cohort of 100 individuals as the validation set to objectively assess the robustness and effectiveness of the candidate SNPs and the PRS. The limited sample-size of the validation set leads to the rarity of homogenous minor allele individuals, and consequently, associations of the candidate SNPs to abnormal ECG could not be effectively confirmed in the perspective of hypothesis test (P value under 0.05 for only one SNP) (Table S5). However, all six candidate SNPs display, in the validation set, classification performances that are better than random guess and similar to those in the initial discovery cohort (Table S5), thus confirming the presence of candidate associations in an independent population. In the validation set, the PRS displays the similar unimodal distribution (Figure S6), is different between the normal and abnormal groups (Wilcoxon rank sum test $P = 0.01$, Fig. 5), and reproducibly correlates well in a unidirectional trend with the susceptibility of abnormal ECG (Fig. 6). The AUC reaches 0.72 and the Youden index is at 0.40 in the ROC curve for the validation set (Fig. 7), further substantiating the PRS as a potential biomarker for high-risk subgroup of abnormal ECG in the Chinese population.

A.



B.

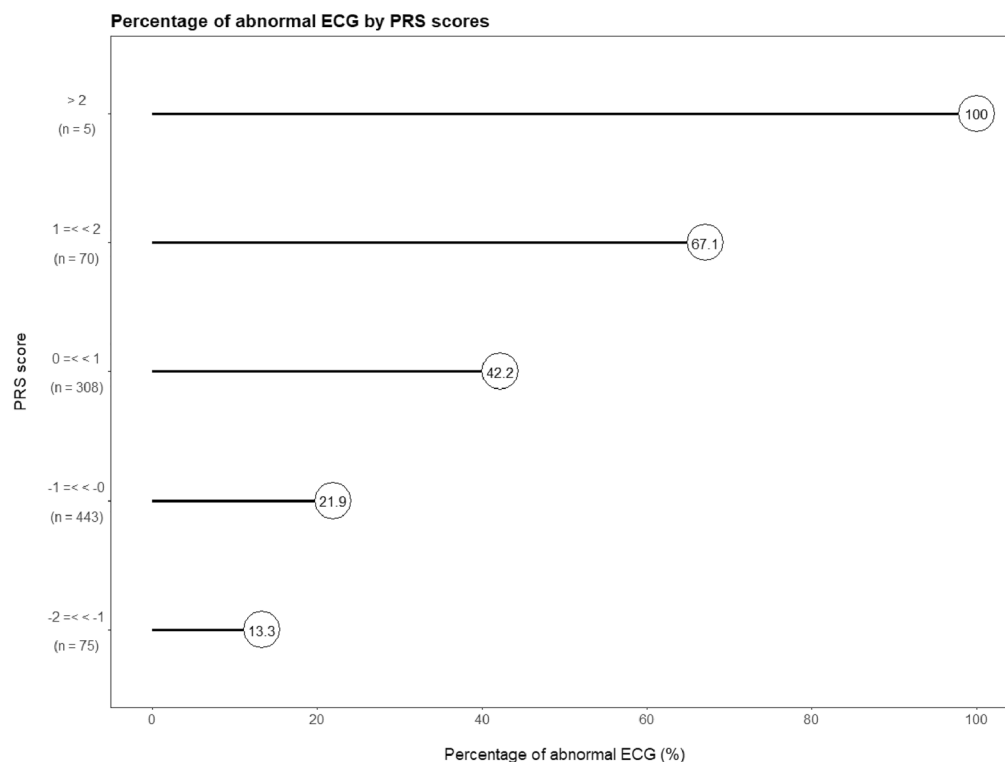


Figure 4. Percentage of abnormal ECG in subgroups of PRS deciles (A) and PRS scores (B). The ten deciles refer to the sequential 10% segments of the cohort ordered by the PRS values. For example, decile #1 annotates individuals with the lowest 10% PRS, and decile #10 annotates individuals with the highest 10% PRS.

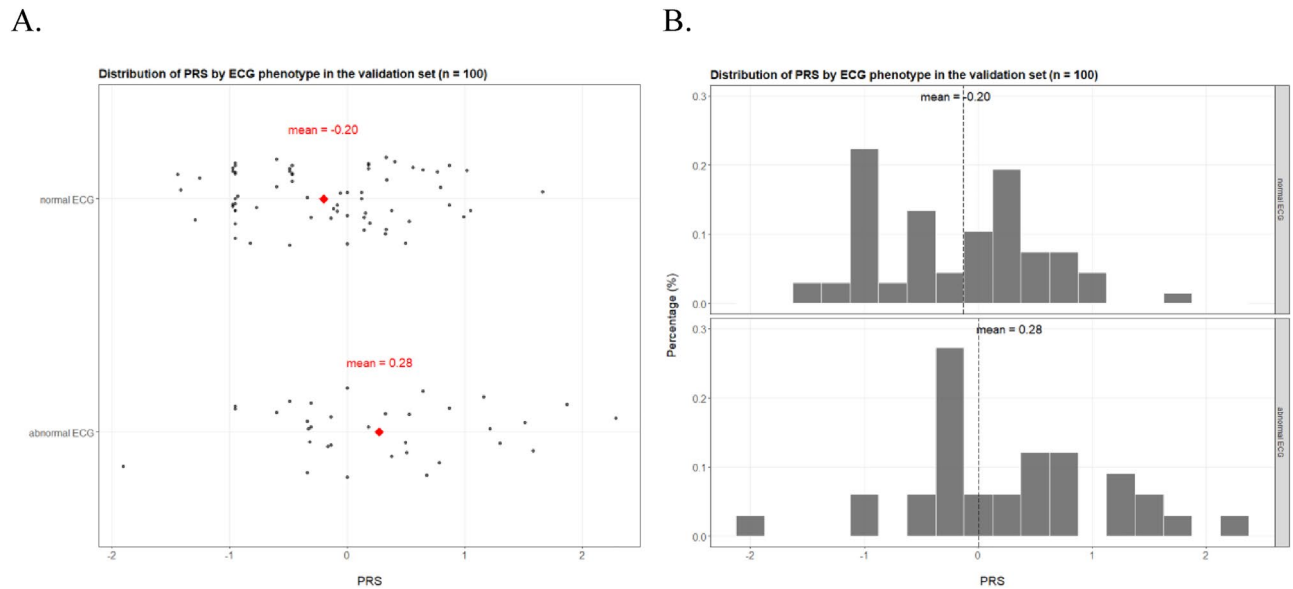


Figure 5. Distribution of PRS by ECG phenotype in the jittered dot plot (A) and the histogram (B) for the validation set.

Discussion

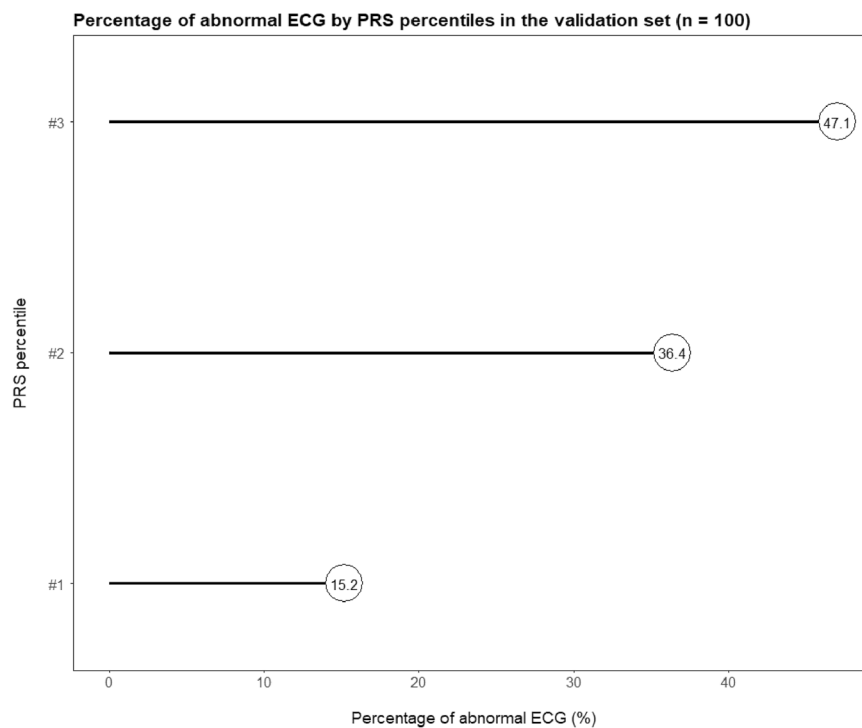
A clear uptrend of abnormal ECG with regard to aging is observed in the cohort, matching the progressive deterioration of cardiovascular functions throughout individuals' lifetime. On the other hand, sex also plays a role as the males and the females display similar but not identical pattern regarding the percentage of abnormal ECG. Therefore, incorporating age and sex into a statistical learning model is instrumental to facilitate the robust forecast of individuals' risk in being assessed for abnormal ECG.

While previous GWASs on electrocardiography were mainly focused on the specific quantitative metrics of ECG, one unique feature of this study is the categorization of ECG assessments into a general normal vs. abnormal binary phenotype based on the labelling of potential issues by certified medical staffs. This partially explains why previously identified genetic variants do not score significantly in this study, similar to the phenomenon observed in the treatment of blood pressure as a quantitative trait (e.g. systolic pressure) vs. as a qualitative trait (e.g. hypertension)^{43,44}. In addition, different backgrounds of ethnicity in these studies might also contribute to the disparate hits of genetic variants. Undoubtedly, there are various subtypes of ECG issues so generalization of all issues into a single abnormal group may be possibly confounding. However, participants may display multiple subtypes (some individuals in this study were labelled for up to 5 ECG issues) so the subtypes are themselves complex and non-independent outcomes to evaluate for individuals. Moreover, both the examination and the evaluation of ECG and as well the genotyping of a cohort are time and resource consuming, so subtype-based analysis suffers from significantly reduced sample size which is the major obstacle for discovering positive hits in heavily-adjusted genomic analysis. By evaluating ECG assessments into a binary outcome, this study benefits from the statistical power conferred by the full cohort. Subtype association analysis would be better suited if sample size would expand for our cohort in the future or when established large-scale cohorts such as the UK Biobank are analyzed⁴⁵.

Abnormal ECG proves to be a complex trait for the mapping to hit genetic variants. None of the six candidate SNPs identified in this study lead to alterations in the primary sequence of proteins, but four of them locate to the noncoding regions of certain genes: *PDGFD* (platelet derived growth factor D), *TMC5* (transmembrane channel like 5), *FBXO10* (F-box protein 10), and *OR7E24* (olfactory receptor family 7 subfamily E member 24). The mechanisms underlying these potential associations remain elusive but factors such as transcriptional regulation or alternative splicing might be involved. Higher statistical power from a larger-size cohort would better reveal the genetic-phenotypic link. Nevertheless, assembling information from all six candidate SNPs, a polygenic risk score displays robust and reproducible difference between the normal and abnormal groups, and sequentially correlates with the increased risk of abnormality in ECG. Validation of the PRS features in the independent set not only supports the credibility of our GWAS but also suggests the potential application of the PRS as a novel biomarker to early detecting Han Chinese individuals with high-risk of heart issues in a precision medicine perspective.

To our knowledge, this is the first GWAS using the professional assessments of ECG as the phenotype-of-interest, and it remains to be investigated to what extent our discoveries in a Han Chinese cohort could be reproduced in other ethnicities. It should also be noted that neither the SNPs nor the PRS lead to optimal classification performance, suggesting the genetic association only contributes partially to the phenotype of abnormal ECG. This should not be surprising because ECG status is itself a complex trait that should be influenced and determined by many factors (genetic, environmental, life-style etc.). Future research should aim to collect information on such variables within large cohorts in a longitudinal timeframe to enhance both the association analysis and the predictive modeling of ECG features and status.

A.



B.

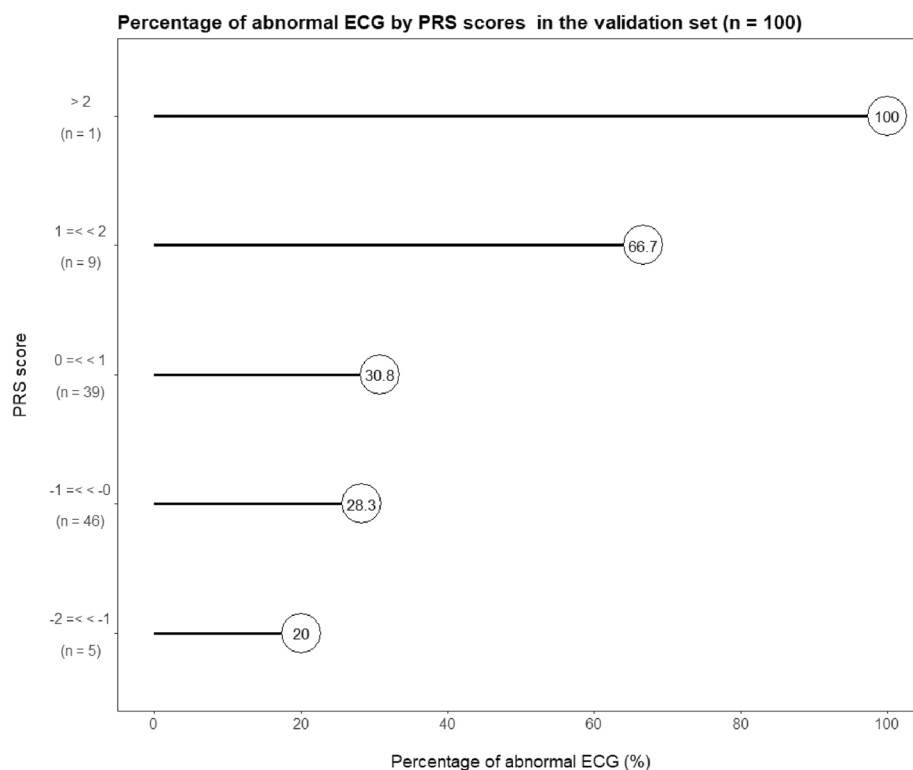


Figure 6. Percentage of abnormal ECG in subgroups of PRS percentiles (A) and PRS scores (B). Percentile #1, #2, and #3 respectively annotates the low, middle, and high one-third of PRS individuals. The validation set is divided into three percentiles rather than ten deciles due to its limited sample size.

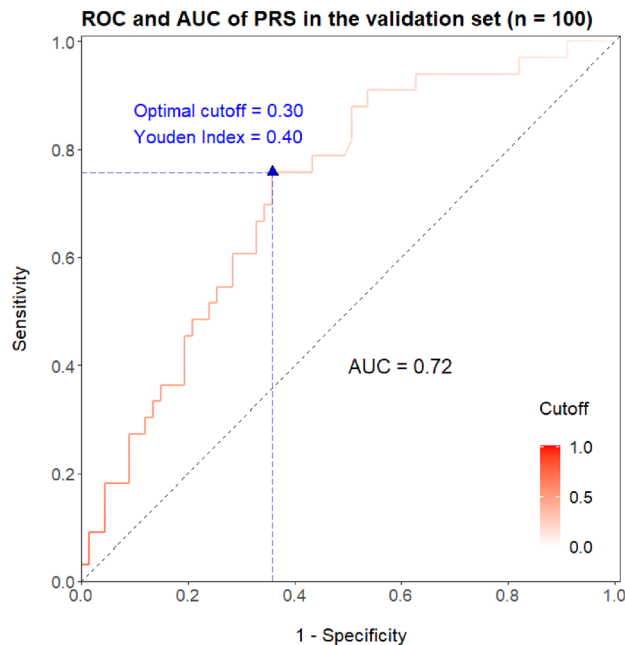


Figure 7. ROC plot of the PRS in the validation set. Cutoff (in the range of [0, 1]) refers to the threshold to dichotomize predicted probability into binary classes. AUC area under the ROC curve. Note: sex and age are included in the predictive model.

Data availability

The Informed Consent Form for the cohort study states that only population-based results would be published and any individual's raw genetic data are prohibited from unauthorized release. Consequently, cohort summary results are reported for GWAS in the perspective of population genetics. Summary statistics for the set of SNPs can be individually inquired from a publicly accessible user-interactive website (https://mengqiao-wang.shinyapps.io/GWAS_ECG).

Received: 27 August 2020; Accepted: 12 February 2021

Published online: 25 February 2021

References

- Baranchuk, A. & Bayes de Luna, A. The P-wave morphology: What does it tell us?. *Herzschrittmacherther. Elektrophysiol.* **26**, 192–199 (2015).
- Delewi, R. *et al.* Pathological Q waves in myocardial infarction in patients treated by primary PCI. *JACC Cardiovasc. Imaging* **6**, 324–331 (2013).
- Channer, K. & Morris, F. ABC of clinical electrocardiography: Myocardial ischaemia. *BMJ* **324**, 1023–1026 (2002).
- Christoffersen, I. E. *et al.* Fifteen genetic loci associated with the electrocardiographic P wave. *Circ. Cardiovasc. Genet.* **10**, 1023 (2017).
- Dhingra, R. *et al.* Electrocardiographic QRS duration and the risk of congestive heart failure: The Framingham Heart Study. *Hypertension* **47**, 861–867 (2006).
- Aro, A. L. *et al.* Intraventricular conduction delay in a standard 12-lead electrocardiogram as a predictor of mortality in the general population. *Circ. Arrhythm Electrophysiol.* **4**, 704–710 (2011).
- Karikari, Y. *et al.* The progress and significance of QRS duration by electrocardiography in hypoplastic left heart syndrome. *Pediatr. Cardiol.* **41**, 141–148 (2020).
- Badheka, A. O. *et al.* QRS duration on electrocardiography and cardiovascular mortality (from the National Health and Nutrition Examination Survey-III). *Am. J. Cardiol.* **112**, 671–677 (2013).
- Laukkanen, J. A. *et al.* T-wave inversion, QRS duration, and QRS/T angle as electrocardiographic predictors of the risk for sudden cardiac death. *Am. J. Cardiol.* **113**, 1178–1183 (2014).
- Nadour, W. *et al.* Does the presence of Q waves on the EKG accurately predict prior myocardial infarction when compared to cardiac magnetic resonance using late gadolinium enhancement? A cross-population study of noninfarct vs infarct patients. *Heart Rhythm* **11**, 2018–2026 (2014).
- O'Neal, W. T. *et al.* Association between QT-interval components and sudden cardiac death: The ARIC Study (Atherosclerosis Risk in Communities). *Circ. Arrhythm. Electrophysiol.* **10**, 20 (2017).
- Patel, S. I. *et al.* QT prolongation and sudden cardiac death risk in hypertrophic cardiomyopathy. *Acta Cardiol.* **74**, 53–58 (2019).
- van Setten, J. *et al.* PR interval genome-wide association meta-analysis identifies 50 loci associated with atrial and atrioventricular electrical activity. *Nat. Commun.* **9**, 2904 (2018).
- Benaglio, P. *et al.* Allele-specific NKX2-5 binding underlies multiple genetic associations with human electrocardiographic traits. *Nat. Genet.* **51**, 1506–1517 (2019).
- Newton-Cheh, C. *et al.* Genome-wide association study of electrocardiographic and heart rate variability traits: The Framingham Heart Study. *BMC Med. Genet.* **8**(Suppl 1), S7 (2007).
- Pfeufer, A. *et al.* Genome-wide association study of PR interval. *Nat. Genet.* **42**, 153–159 (2010).

17. Silva, C. T. *et al.* Heritabilities, proportions of heritabilities explained by GWAS findings, and implications of cross-phenotype effects on PR interval. *Hum. Genet.* **134**, 1211–1219 (2015).
18. Pilia, G. *et al.* Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet.* **2**, e132 (2006).
19. Butler, A. M. *et al.* Novel loci associated with PR interval in a genome-wide association study of 10 African American cohorts. *Circ. Cardiovasc. Genet.* **5**, 639–646 (2012).
20. Visscher, P. M. *et al.* 10 years of GWAS discovery: Biology, function, and translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
21. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, 5890 (2015).
22. Bihlmeyer, N. A. *et al.* ExomeChip-wide analysis of 95 626 individuals identifies 10 novel loci associated with QT and JT intervals. *Circ. Genom. Precis. Med.* **11**, e001758 (2018).
23. Consortium, G.T. Human genomics. The genotype-tissue expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
24. Mieszczańska, H. *et al.* Gender-related differences in electrocardiographic parameters and their association with cardiac events in patients after myocardial infarction. *Am. J. Cardiol.* **101**, 20–24 (2008).
25. Vicente, J., Johannesen, L., Galeotti, L. & Strauss, D. G. Mechanisms of sex and age differences in ventricular repolarization in humans. *Am. Heart J.* **168**, 749–756 (2014).
26. Nolte, I. M. *et al.* A comparison of heritability estimates by classical twin modeling and based on genome-wide genetic relatedness for cardiac conduction traits. *Twin Res. Hum. Genet.* **20**, 489–498 (2017).
27. Li, J. *et al.* Familial aggregation and heritability of electrocardiographic intervals and heart rate in a rural Chinese population. *Ann. Noninvasive Electrocardiol.* **14**, 147–152 (2009).
28. Mutikainen, S. *et al.* Genetic influences on resting electrocardiographic variables in older women: A twin study. *Ann. Noninvas. Electrocardiol.* **14**, 57–64 (2009).
29. Holm, H. *et al.* Several common variants modulate heart rate, PR interval and QRS duration. *Nat. Genet.* **42**, 117–122 (2010).
30. Prins, B. P. *et al.* Exome-chip meta-analysis identifies novel loci associated with cardiac conduction, including ADAMTS6. *Genome Biol.* **19**, 87 (2018).
31. Sano, M. *et al.* Genome-wide association study of electrocardiographic parameters identifies a new association for PR interval and confirms previously reported associations. *Hum. Mol. Genet.* **23**, 6668–6676 (2014).
32. den Hoed, M. *et al.* Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nat. Genet.* **45**, 621–631 (2013).
33. Wang, Z. *et al.* Epigenetic silencing of the 3p22 tumor suppressor DLEC1 by promoter CpG methylation in non-Hodgkin and Hodgkin lymphomas. *J. Transl. Med.* **10**, 209 (2012).
34. Park, C. Y. *et al.* skNAC, a Smyd1-interacting transcription factor, is involved in cardiac development and skeletal muscle growth and regeneration. *Proc. Natl. Acad. Sci. USA* **107**, 20750–20755 (2010).
35. Lee, S., Abecasis, G. R., Boehnke, M. & Lin, X. Rare-variant association analysis: Study designs and statistical tests. *Am. J. Hum. Genet.* **95**, 5–23 (2014).
36. Macfarlane, P. W., McLaughlin, S. C., Devine, B. & Yang, T. F. Effects of age, sex, and race on ECG interval measurements. *J. Electrocardiol.* **27**(Suppl), 14–19 (1994).
37. Macfarlane, P. W. *et al.* Racial differences in the ECG-selected aspects. *J. Electrocardiol.* **47**, 809–814 (2014).
38. Kerr, K. F. *et al.* Genome-wide association study of heart rate and its variability in Hispanic/Latino cohorts. *Heart Rhythm* **14**, 1675–1684 (2017).
39. Swenson, B. R. *et al.* GWAS of QRS duration identifies new loci specific to Hispanic/Latino populations. *PLoS One* **14**, e0217796 (2019).
40. Mendez-Giraldez, R. *et al.* GWAS of the electrocardiographic QT interval in Hispanics/Latinos generalizes previously identified loci and identifies population-specific signals. *Sci. Rep.* **7**, 17075 (2017).
41. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
42. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
43. Padmanabhan, S. & Joe, B. Towards precision medicine for hypertension: A review of genomic, epigenomic, and microbiomic effects on blood pressure in experimental rat models and humans. *Physiol. Rev.* **97**, 1469–1528 (2017).
44. Hastie, C. E., Padmanabhan, S. & Dominiczak, A. F. Genome-wide association studies of hypertension: Light at the end of the tunnel. *Int. J. Hypertens.* **20**, 509581 (2010).
45. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

Acknowledgements

The authors would like to thank the staffs and volunteers of the Southwestern Natural Population Cohort Project for the assistance with participant recruitment and medical electrocardiography. The authors would appreciate all individuals who participated in this research project. M.W. appreciates the authors of various R packages (*snStats*, *SNPRelate*, *doParallel*, *foreach*, *LDheatmap*, *dplyr*, *caret*, *ggplot2*, *grid*, *shiny* etc.) he has used in the data analysis pipeline, and RStudio, PBC for providing free hosting of his Shiny-based GWAS-ECG inquiry website.

Author contributions

M.W. and Y.S. designed the study. X.Z. supervised the cohort recruitment/management and the electrocardiogram examinations/assessments. M.W. and J.G. conducted the genotyping experiments. M.W. executed the data analysis, and produced all figures and tables. M.W. wrote and revised the manuscript. All authors have read, reviewed, and approved the manuscript.

Funding

This work was supported by the National Key Research and Development Program of China (2017YFC0907305 and 2017YFC0907300), the Youth Scholar Grant from the National Natural Science Foundation of China (81800071), and the Basic Research Grant from the Department of Science and Technology of Sichuan Province (2019YJ0035).

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/s41598-021-84135-7>.

Correspondence and requests for materials should be addressed to M.W., Y.S. or X.Z.

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, corrected publication 2021