

Genetic Architecture of Quantitative Cardiovascular Traits: Blood Pressure, ECG and Imaging Phenotypes

Nay Aung, William J Young, Stefan van Duijvenboden, Julia Ramírez, Steffen E Petersen, Patricia B Munroe

William Harvey Research Institute, NIHR Cardiovascular Biomedical Research Centre at Barts, Queen Mary University of London, UK

Abstract

Background: We provide an overview of the genetic architecture of quantitative cardiovascular phenotypes such as blood pressure (BP), electrocardiogram (ECG) and cardiac imaging measurements which play a critical and prognostic role in the management of numerous diseases.

Methods: The genetics of BP, ECG and cardiac imaging traits have been studied in large-scale genome-wide association studies (GWASs).

Results: To-date more than 1,400 BP loci have been discovered. These genetic loci harbouring known and novel BP-regulating genes, several of which are linked to existing drugs, that can be repurposed for BP treatment. Regarding the ECG indices, 437 independent signals have been reported for resting heart rate ($n \sim 400,000$), mainly modulating the autonomic nervous system, as well as 202 loci for PR interval, 29 loci for QRS duration and 35 loci for QT interval. The LV GWASs ($n \sim 17,000$) identified 14 loci harbouring genes regulating the cardiac developmental pathways.

Conclusion: Large-scale genetic analyses of quantitative cardiovascular traits have yielded hundreds of susceptibility loci, candidate genes and key biological pathways, which significantly advance our understanding of their genetic architecture and shed lights on potential novel therapeutic targets.

1. Introduction

Quantitative measurements such as blood pressure (BP) and electrocardiogram (ECG) and cardiac imaging parameters play a dominant role in the assessment of cardiovascular (CV) health. They are important endophenotypes which represent a wide spectrum of variation reflecting normal physiology to an overt pathology. Hypertension, diagnosed by the presence of elevated systolic and/or diastolic blood pressure (SBP and DBP), is the most important modifiable risk factor attributable to 10.4 million deaths worldwide in 2017 (1). Abnormalities in cardiac electrical phenotypes derived

from ECG, such as PR interval and QT interval, are predictive of future risk of arrhythmia and sudden death (2,3). Lastly, cardiac imaging technologies such as echocardiography (ECHO) and cardiovascular magnetic resonance (CMR) provide a plethora of measurements related to cardiac structure and function, instrumental in the management of many CV diseases. All of these phenotypes are the so-called common and complex traits under the influence of genes, environment and their intricate interplay. Typically, the genetic basis of such phenotypes is investigated with genome-wide association studies (GWASs). Here, we provide a brief overview of key findings from the published GWASs which illuminates the genetic architecture of clinically relevant quantitative CV traits.

2. Genome-wide association studies

At the turn of 21st century, the human genome project completed sequencing of the entire human genome consisting of 3 billion nucleotide base pairs at a cost of ~2.7 billion US dollars. Since then, the genotyping cost has fallen exponentially, bringing the genomic technology to the masses. Over the last decade, several population-based studies had conducted whole-genome genotyping – most commonly with DNA micro-arrays – in thousands to millions of individuals. Although the human genome is 99.9% identical between individuals, each person's genome differs from the reference human genome at 4 to 5 million sites, the majority of which are single nucleotide polymorphisms (SNPs). GWASs systematically examines these genetic variations across the genome, in order to identify regions (loci) associated with a particular trait or disease. They provide an important avenue to agnostically evaluate the association between human genetic variations and complex traits and are well-suited to investigate the genetic determinants of both quantitative traits and binary disease status. Downstream utilities of GWAS results include: (i) estimation of heritability to understand the genetic architecture of a trait; (ii) examination of genetic correlations which detect and quantify pleiotropy; (iii) identification of candidate causal genes which explain the biology and may lead to novel therapies and drug

repurposing opportunities; (iv) construction of polygenic risk scores; and (v) creation of genetic instruments for Mendelian randomisation studies to explore causal relationships.

2.1. Genomic studies of blood pressure traits

Hypertension is an important root cause of stroke, coronary artery disease, heart failure and kidney disease. Given its clinical importance and the simplicity of data acquisition, BP is perhaps the most well-studied quantitative CV trait with a large genomic knowledge base. The discovery of genetic loci associated with BP traits is further aided by the emergence mega studies and consortia such as the UK Biobank (n ~ 500,000) (4), the US Million Veteran Program (n ~ 318,891) (5) and the International Consortium of Blood Pressure (ICBP) (n ~ 299,024) (6). BP is highly heritable with 20 – 50% of variability attributable to genetics (7). However, even a very large recent study in over a million individuals only recovered 27% of expected heritability (8), highlighting the magnitude of gap in our knowledge. To date, a total of 1477 BP risk loci have been reported in the literature, consistent with the polygenicity of BP regulation which creates a significant challenge in disentangling the biological function of each locus. Increasingly, in-silico analyses leveraged on publicly available genomic databases are used to prioritise candidate genes which can be followed up in experimental studies using techniques like CRISPR gene editing. This approach has highlighted a number of highly promising genes that led to an improved understanding of BP physiology and drug repurposing. An example of such success story is the *UMOD* (uromodulin) gene identified in an early BP GWAS (9). Hypertensive individuals with AA genotype at rs13333226 variant have higher *UMOD* excretion, greater salt sensitivity and greater BP drop in response to loop diuretics when compared with those with GG genotype (9). This observation alongside other supporting evidence in a transgenic mouse model (10) has resulted in a phase-4 clinical trial that investigates if the therapeutic response of a known antihypertensive agent (Torsemide) differs depending on the genotype (AA vs AG/GG) (<https://clinicaltrials.gov/ct2/show/NCT03354897>). The abundance of candidate BP genes allows the enrichment (also known as over-representation) analysis which highlights key biological pathways (11). Some notable pathways include G-protein coupled receptor signalling, extracellular matrix regulation and olfactory signalling which can be dissected in future studies for potential therapeutic targets.

2.2. Genomic studies of ECG traits

Before the advent of GWASs, the genetic underpinning of cardiac electrical activity was primarily investigated in the context of Mendelian channelopathies such as the long QT syndrome (LQTS). Inclusion of ECG in more recent large-scale population genomic studies enabled investigations into the genetic basis of prognostically important ECG-derived electrical traits. Observationally, elevated resting HR is associated with an increased risk of CV morbidity and mortality (12). A recent GWAS of resting HR in the UK Biobank (n = 428,250) discovered a total of 437 genetic loci (13) and evidence of genetic and causal association between resting HR and type 2 diabetes. The analysis of shared genes between the two traits highlighted the role of glucose homeostasis, energy metabolism and autonomic activity which may explain the co-existence high resting heart rate and type 2 diabetes. A multi-ancestry GWAS of PR interval, a conduction trait, reported 202 loci (141 of which were novel) enriched in cardiac muscle development and cytoskeletal genes (14). The polygenic risk score of PR interval is predictive of atrial fibrillation (AF) and distal conduction disease, recapitulating the findings from prior observational studies. As for QRS duration which represents ventricular depolarisation and conduction, a total of 29 loci have been reported to date by studies in a diverse set of ethnicities (15,16). Interestingly, PR-associated *MYOCD* locus containing *MYOCD* gene that plays a role in cardiogenesis was discovered only in Hispanic individuals underlining the importance of expanding the genome-wide discovery analysis into non-European ancestries (15). The QT interval which reflects myocardial repolarisation is a key measurement for the diagnosis of Mendelian long and short QT syndromes and can be modulated by electrolyte imbalance, myocardial ischaemia and medications. QT prolongation has been associated with malignant arrhythmias and sudden cardiac death (17). A large genome-wide association meta-analysis (n ~ 103,331) of QT interval yielded a total of 35 genome-wide significant loci which harbour not only known genes implicated in LQTS, but novel genes such as *ATP2A2*, *PLN*, *PRKCA*, *SRL* and *SLC8A1* that modulate the calcium signalling pathway (18). A more refined assessment of repolarisation, namely T-peak-to-T-end-interval (Tpe), has been investigated in a GWAS which identified 32 loci comprising genes involved in cellular processes of ventricular repolarisation (19). More recently, several studies have looked into exercise ECG measurements which are known to be prognostically important. A GWAS of HR response to exercise and recovery discovered 30 loci containing genes involved in neural development and regulation of autonomic nervous system (20). Another study on exercise QT dynamics revealed 20 loci comprising genes implicated in LQTS, mitochondrial calcium handling and autonomic regulation (21). A genome-wide investigation into exercise-related variation in the T-wave morphology, known as T-wave morphology

restitution (TMR), identified 12 loci and the polygenic risk score developed from the loci of exercise-recovery TMR was predictive of future CV events (22).

2.3. Genomic studies of cardiac imaging traits

Cardiac imaging phenotypes mediate considerably heterogeneous syndromes such as heart failure and inherited cardiomyopathies. They can be measured with a high precision using current imaging techniques. In particular, ECHO and CMR are well-suited for large-scale population studies due to their non-ionising properties. Imaging phenotypes can be objectively quantified and evaluated as indicators of both normal biological and pathogenic processes. They may be superior to binary disease traits such as heart failure in a GWAS due to their ability to detect pre-clinical changes with a diminished misclassification bias. The largest published GWAS of ECHO-derived measurements (16 left ventricular [LV] and aortic phenotypes) conducted in 46,533 European individuals discovered 12 loci for aortic diameter, 3 loci for LV end-diastolic diameter, one locus each for LV fractional shortening, LV mass and peak velocity of mitral valve inflow corresponding to atrial contraction (Mv-A) (23). Another ECHO-based study in Japanese individuals (n max ~ 19000) reported a total of 14 novel loci: 2 loci for LV end-diastolic diameter, 3 loci for LV ejection fraction, 4 loci each for LV end-systolic diameter and LV fractional shortening and one locus for LV relative wall thickness (24). Limited follow-up data are available for these GWAS signals. Of note, the *MTSS1* locus for LV end-diastolic diameter, identified by both European and Japanese studies contains *MTSS1* gene which is expressed in the heart and encodes a scaffold protein that regulates actin dynamics (25). A more recent study of LV phenotypes (n ~ 17000) using CMR, which leveraged on the breakthroughs in automatic segmentation by deep learning, discovered 14 locus-phenotype associations of which 6 loci were novel for LV imaging traits (26). Some noteworthy candidate genes in these loci include: *TTN* involved in sarcomere assembly and stretch sensing with its truncating variants found to be causal in 10% to 20% of dilated cardiomyopathy (DCM) cases (27); *BAG3* (associated with LV end-systolic volume and LV ejection fraction) essential for Z-disc integrity and also implicated in DCM (28); *ALPK3* that regulates cardiomyocyte differentiation (29), and *MMP11* involved in extracellular matrix homeostasis (30). Polygenic risk scores constructed from the LV GWASs were predictive of incident heart failure supporting the role of imaging phenotypes (specifically, their genetic risk) as intermediate biomarkers for hard outcomes.

3. Conclusion

Significant progress has been made in unravelling the genetic underpinning of quantitative cardiovascular phenotypes based on the large-scale GWASs of common variants. The time is upon us to capitalise on the gain in knowledge by conducting mechanistic studies of the promising candidate genes which may lead to novel therapies. Delving into the rare variants aided by the emerging data from whole genome sequencing may shine the light on the hitherto missing heritability and potentially identify new variants with larger effect sizes. Certain genetic loci demonstrated pleiotropic associations with multiple cardiovascular traits (for instance, *SH2B3* locus with BP and LV end-diastolic volume; *TTN* with multiple imaging traits, PR interval and AF). This observation underlines their intricate physiological and mechano-electrical relationships which could be exploited by joint multi-trait and genetic overlap analyses in future studies.

References

1. Stanaway JD, Afshin A, Gakidou E, Lim SS, Abate D, Abate KH, et al. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*. 2018 Nov 10;392(10159):1923–94.
2. Cheng S, Keyes MJ, Larson MG, McCabe EL, Newton-Cheh C, Levy D, et al. Long-term outcomes in individuals with a prolonged PR interval or first-degree atrioventricular block. *JAMA J Am Med Assoc*. 2009 Jun 24;301(24):2571–7.
3. Algra A, Tijssen JG, Roelandt JR, Pool J, Lubsen J. QTc prolongation measured by standard 12-lead electrocardiography is an independent risk factor for sudden death due to cardiac arrest. *Circulation*. 1991 Jun;83(6):1888–94.
4. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018 Oct;562(7726):203–9.
5. Gaziano JM, Concato J, Brophy M, Fiore L, Pyarajan S, Breeling J, et al. Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol*. 2016 Feb 1;70:214–23.
6. Wain LV, Vaez A, Jansen R, Joehanes R, van der Most PJ, Erzurumluoglu AM, et al. Novel blood pressure locus and gene discovery using genome-wide association study and expression data sets from blood and the kidney. *Hypertens Dallas Tex 1979 [Internet]*. 2017 Jul 24 [cited 2020 Aug 25]; Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5783787/>
7. Salfati E, Morrison AC, Boerwinkle E, Chakravarti A. Direct estimates of the genomic contributions to blood

- pressure heritability within a population-based cohort (ARIC). *PLOS ONE*. 2015 Jul 10;10(7):e0133031.
8. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet*. 2018 Oct;50(10):1412.
 9. Padmanabhan S, Melander O, Johnson T, Di Blasio AM, Lee WK, Gentilini D, et al. Genome-wide association study of blood pressure extremes identifies variant near UMOD associated with hypertension. *PLoS Genet*. 2010 Oct 28;6(10):e1001177.
 10. Graham LA, Padmanabhan S, Fraser NJ, Kumar S, Bates JM, Raffi HS, et al. Validation of uromodulin as a candidate gene for human essential hypertension. *Hypertens Dallas Tex* 1979. 2014 Mar;63(3):551–8.
 11. Cabrera CP, Ng FL, Nicholls HL, Gupta A, Barnes MR, Munroe PB, et al. Over 1000 genetic loci influencing blood pressure with multiple systems and tissues implicated. *Hum Mol Genet*. 2019 Nov 21;28(R2):R151–61.
 12. Aune D, Sen A, ó'Hartaigh B, Janszky I, Romundstad PR, Tonstad S, et al. Resting heart rate and the risk of cardiovascular disease, total cancer, and all-cause mortality - A systematic review and dose-response meta-analysis of prospective studies. *Nutr Metab Cardiovasc Dis NMCD*. 2017 Jun;27(6):504–17.
 13. Guo Y, Chung W, Zhu Z, Shan Z, Li J, Liu S, et al. Genome-wide assessment for resting heart rate and shared genetics with cardiometabolic traits and type 2 diabetes. *J Am Coll Cardiol*. 2019 Oct 29;74(17):2162–74.
 14. Ntalla I, Weng L-C, Cartwright JH, Hall AW, Sveinbjornsson G, Tucker NR, et al. Multi-ancestry GWAS of the electrocardiographic PR interval identifies 202 loci underlying cardiac conduction. *Nat Commun*. 2020 May 21;11(1):2542.
 15. Br S, T L, Hj L, R M-G, Je B, Cc L, et al. GWAS of QRS duration identifies new loci specific to Hispanic/Latino populations. *Plos One*. 2019 Jun 28;14(6):e0217796–e0217796.
 16. van Setten J, Verweij N, Mbarek H, Niemeijer MN, Trompet S, Arking DE, et al. Genome-wide association meta-analysis of 30,000 samples identifies seven novel loci for quantitative ECG traits. *Eur J Hum Genet*. 2019 Jun;27(6):952–62.
 17. Al-Khatib SM, LaPointe NMA, Kramer JM, Califf RM. What clinicians should know about the QT interval. *JAMA*. 2003 Apr 23;289(16):2120–7.
 18. Arking DE, Pulit SL, Crotti L, van der Harst P, Munroe PB, Koopmann TT, et al. Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. *Nat Genet*. 2014 Aug;46(8):826–36.
 19. Ramírez J, van Duijvenboden S, Young WJ, Orini M, Lambiase PD, Munroe PB, et al. Common genetic variants modulate the electrocardiographic Tpeak-to-Tend interval. *Am J Hum Genet*. 2020 Jun 4;106(6):764–78.
 20. Ramírez J, Duijvenboden S van, Ntalla I, Mifsud B, Warren HR, Tzani E, et al. Thirty loci identified for heart rate response to exercise and recovery implicate autonomic nervous system. *Nat Commun [Internet]*. 2018 May 16 [cited 2018 Jun 4];9. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5955978>
 21. van Duijvenboden Stefan, Ramírez Julia, Young William J., Mifsud Borbala, Orini Michele, Tinker Andrew, et al. Genetic basis and prognostic value of exercise QT dynamics. *Circ Genomic Precis Med*. 2020 Aug 1;13(4):e002774.
 22. Ramírez Julia, van Duijvenboden Stefan, Aung Nay, Laguna Pablo, Pueyo Esther, Tinker Andrew, et al. Cardiovascular predictive value and genetic basis of ventricular repolarization dynamics. *Circ Arrhythm Electrophysiol*. 2019 Oct 1;12(10):e007549.
 23. Wild PS, Felix JF, Schillert A, Teumer A, Chen M-H, Leening MJG, et al. Large-scale genome-wide analysis identifies genetic variants associated with cardiac structure and function. *J Clin Invest*. 2017 May 1;127(5):1798–812.
 24. Kanai M, Akiyama M, Takahashi A, Matoba N, Momozawa Y, Ikeda M, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet*. 2018;50(3):390–400.
 25. Saarikangas J, Mattila PK, Varjosalo M, Bovellan M, Hakanen J, Calzada-Wack J, et al. Missing-in-metastasis MIM/MTSS1 promotes actin assembly at intercellular junctions and is required for integrity of kidney epithelia. *J Cell Sci*. 2011 Apr 15;124(Pt 8):1245–55.
 26. Aung Nay, Vargas Jose D., Yang Chaojie, Cabrera Claudia P., Warren Helen R., Fung Kenneth, et al. Genome-wide analysis of left ventricular image-derived phenotypes identifies fourteen loci associated with cardiac morphogenesis and heart failure development. *Circulation*. 2019 Oct 15;140(16):1318–30.
 27. Herman DS, Lam L, Taylor MRG, Wang L, Teekakirikul P, Christodoulou D, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med*. 2012 Feb 16;366(7):619–28.
 28. Villard E, Perret C, Gary F, Proust C, Dilanian G, Hengstenberg C, et al. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J*. 2011 May;32(9):1065–76.
 29. Hosoda T, Monzen K, Hiroi Y, Oka T, Takimoto E, Yazaki Y, et al. A novel myocyte-specific gene midori promotes the differentiation of P19CL6 cells into cardiomyocytes. *J Biol Chem*. 2001 Sep 21;276(38):35978–89.
 30. Li YY, McTiernan CF, Feldman AM. Interplay of matrix metalloproteinases, tissue inhibitors of metalloproteinases and their regulators in cardiac matrix remodeling. *Cardiovasc Res*. 2000 May 1;46(2):214–24.

Address for correspondence:

Dr Nay Aung
 William Harvey Research Institute
 Queen Mary University of London
 Charterhouse Square
 London EC1M 6BQ
 n.aung@qmul.ac.uk