



Genomic approaches for the elucidation of genes and gene networks underlying cardiovascular traits

M. E. Adriaens¹ · C. R. Bezzina²

Received: 18 May 2018 / Accepted: 13 June 2018 / Published online: 22 June 2018
© The Author(s) 2018

Abstract

Genome-wide association studies have shed light on the association between natural genetic variation and cardiovascular traits. However, linking a cardiovascular trait associated locus to a candidate gene or set of candidate genes for prioritization for follow-up mechanistic studies is all but straightforward. Genomic technologies based on next-generation sequencing technology nowadays offer multiple opportunities to dissect gene regulatory networks underlying genetic cardiovascular trait associations, thereby aiding in the identification of candidate genes at unprecedented scale. RNA sequencing in particular becomes a powerful tool when combined with genotyping to identify loci that modulate transcript abundance, known as expression quantitative trait loci (eQTL), or loci modulating transcript splicing known as splicing quantitative trait loci (sQTL). Additionally, the allele-specific resolution of RNA-sequencing technology enables estimation of allelic imbalance, a state where the two alleles of a gene are expressed at a ratio differing from the expected 1:1 ratio. When multiple high-throughput approaches are combined with deep phenotyping in a single study, a comprehensive elucidation of the relationship between genotype and phenotype comes into view, an approach known as systems genetics. In this review, we cover key applications of systems genetics in the broad cardiovascular field.

Keywords GWAS · Cardiovascular · RNA-seq · Allele-specific expression · Systems genetics

Introduction

Over the last decade, genome-wide association (GWA) studies have shed light on the association between natural genetic variation and cardiovascular traits. GWA studies have been conducted, largely in individuals of European descent, for sudden cardiac death (Arking et al. 2011; Bezzina et al. 2010); atrial fibrillation (Benjamin et al. 1994; Christophersen et al. 2017b; Ellinor et al. 2010, 2012; Gudbjartsson et al. 2007, 2009; Lee et al. 2017; Low et al. 2017); resting heart rate (Cho et al. 2009; den Hoed et al.

2013; Deo et al. 2013; Eijgelsheim et al. 2010; Eppinga et al. 2016; Holm et al. 2010; Marroni et al. 2009; Mezzavilla et al. 2014; Nagy et al. 2017; Noordam et al. 2017); heart rate variability (Kerr et al. 2017; Nolte et al. 2017); PR interval (Butler et al. 2012; Chambers et al. 2010; Christophersen et al. 2017a; Hong et al. 2014; Pfeufer et al. 2010; Smith et al. 2011; Verweij et al. 2014); QRS duration (Evans et al. 2016; Floyd et al. 2018; Holm et al. 2010; Hong et al. 2014; Ritchie et al. 2013; Sotoodehnia et al. 2010; van der Harst et al. 2016); ST and JT duration (Floyd et al. 2018; Verweij et al. 2016); and QT interval (Arking et al. 2006, 2014; Chambers et al. 2010; Floyd et al. 2018; Holm et al. 2010; Kim et al. 2012; Marroni et al. 2009; Newton-Cheh et al. 2009; Nolte et al. 2009; Pfeufer et al. 2009; Sano et al. 2014; Smith et al. 2012). However, of the hundreds of loci that have been identified within these studies, most are intergenic and do not directly affect protein sequence or function, instead impacting on regulation of gene expression and splicing (Heinig et al. 2017; Koopmann et al. 2014), thereby requiring the generation of additional gene expression data in relevant tissues. Secondly, taking into account the non-linear nature of regulatory chromatin interactions, a locus can interact with any one or

This article is part of a Special Issue on “Heart Failure Due to Non-Myofibrillar Defects” edited by Elisabeth Ehler and Katja Gehmlich.

✉ C. R. Bezzina
c.r.bezzina@amc.uva.nl

¹ Maastricht Centre for Systems Biology, Maastricht University, Universiteitssingel 60, 6229 ER Maastricht, The Netherlands

² Department of Clinical and Experimental Cardiology, Heart Center, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands

multiple genes in *cis*, as shown previously for the *SCN5A* locus which houses several important sodium channel genes (van den Boogaard et al. 2014). Thirdly, the underlying statistical modeling foundations of GWA studies seldom take epistatic effects into account, due to the low power associated with formal testing of such gene-gene interactions (Ma et al. 2015). Hence, linking a cardiovascular trait associated locus to a candidate gene or set of candidate genes for prioritization for follow-up mechanistic studies is all but straightforward.

Genomic technologies based on next-generation sequencing technology nowadays offer multiple opportunities to dissect gene regulatory networks underlying genetic cardiovascular trait associations, thereby aiding in the identification of candidate genes at unprecedented scale. Genome-wide chromosome conformation capture techniques (3C, 4C, 5C, and Hi-C) have demonstrated that chromatin interactions in human cells occur predominantly within domains with an average size of 880 Kb, known as topologically associated domains (TADs). These TADs are shared across a variety of cell types and are suggested to align with replication sites (Dixon et al. 2016) and causal genes underlying a GWAS locus are likely to reside in the same TAD as the haplotype carrying the associated SNP. Meanwhile, chromatin immunoprecipitation (ChIP) sequencing has opened the path to studying a plethora of histone modifications, which when combined can pinpoint specific regulatory regions within TADs, such as enhancers, promoters, silencers, and insulators, at an unprecedented resolution (ENCODE-Project-Consortium 2012; Zhou et al. 2011). Additional high-resolution chromosome conformation capture techniques subsequently enables revealing of the interactions between regulatory regions, for example showing the gene promoters that a specific enhancer associates with as in the case of the *SCN10A* enhancer interacting with the *SCN5A* locus (van den Boogaard et al. 2014). Finally, RNA sequencing (RNA-seq) technology can reveal the consequence of these chromatin interactions on the expression and splicing of genes.

RNA-seq becomes a particularly powerful tool when combined with genotyping to identify loci that modulate transcript abundance, known as expression quantitative trait loci (eQTL), or loci modulating transcript splicing known as splicing quantitative trait loci (sQTL). Additionally, the allele-specific resolution of RNA-seq technology enables estimation of allelic imbalance, a state where the two alleles of a gene are expressed at a ratio differing from the expected 1:1 ratio. Key mechanisms underlying allelic imbalance include the presence of truncating variants leading to nonsense-mediated decay of one of the alleles (Castel et al. 2015), genetic regulation via *cis* eQTLs and sQTLs, as well as epigenetic modifications through imprinting and interactions with the environment (Castel et al. 2015; Gutierrez-Arcelus et al. 2013; Gutierrez-Arcelus et al. 2015). The genetic effects can

be calculated using standardized procedures (Shabalín 2012), after which the epigenetic effect can be reliably established through statistical inference (Knowles et al. 2017).

Genomic datasets like the ones described above are increasingly available in the public domain for data reuse and integration. A prime example, the ambitious database of GTEx (GTEx-Consortium et al. 2017) offers data on genetic regulation of expression (eQTL) and splicing (sQTL) in 44 human tissues, including cardiac tissues, at respectable sample sizes. Meanwhile, the current Ensembl database offers integrated information of regulatory regions, such as enhancers and promoters, in a plethora of different cell types and tissues (Zerbino et al. 2018), including cardiac tissues, based in part on the results from the ENCODE project (ENCODE-Project-Consortium 2012). One step further, the integrative web-based platform FUMA (Watanabe et al. 2017) uses information from multiple biological resources to facilitate functional annotation of GWA study results and gene prioritization by accommodating positional, eQTL, and chromatin interaction mappings, and provides gene-based, pathway, and tissue enrichment results. FUMA has recently been successfully applied in GWA studies on insomnia (Hammerschlag et al. 2017) and dementia (Sniekers et al. 2017). It should be noted that all these resources are updated on a regular basis, hence using the most recent online version should be common practice in their application.

The above would suggest that integrative interpretation is mainly restricted by the available computational expertise and resources. However, an important limitation of public resources is that the majority of data is derived from tissues of non-diseased origin. Additionally, tissues are composed of many different cell types, with each comprising a specific gene expression profile and set of gene regulatory networks (GTEx-Consortium et al. 2017). In assessing differences between experimental conditions, this may result in spurious findings due to variations in cell type composition, or conversely, the inability to identify differences due to underlying mechanisms being cell type specific. Finally, biological differences between different study populations, for example with respect to genetic and epigenetic background and phenotypic characteristics, as well as technical differences, such as experimental conditions and sample collection and handling, are often unknown, leading to residual latent confounding that comes at the cost of statistical power. Ideally, multiple high-throughput approaches are combined with deep phenotyping in a single study, integrating genetic variants with molecular phenotypes to more comprehensively elucidate the relationship between genotype and phenotype, an approach known as systems genetics (Civelek and Lusis 2014; Moreno-Moral et al. 2017). Here, we cover several key applications of systems genetics in the broad cardiovascular field.

The principles underlying systems genetics approaches

The starting point of systems genetics approaches remains quantitative trait locus (QTL) analysis, identifying genetic variants modulating a specific quantitative trait. By subsequently overlapping the QTL with eQTL and sQTL data, either generated within the study or retrieved from public repositories (GTEx-Consortium 2013; GTEx-Consortium et al. 2017), candidate genes that are under genetic control with regard to overall expression or splicing can be identified. In case such genes are found, their respective regulatory networks should be further explored by integrating additional information. This may consist of predictive information, such as protein-protein interactions (Szklarczyk et al. 2017), shared function via Gene Ontology (Alexa et al. 2006), shared transcription factor binding sites (Roeder et al. 2009) shared interacting regulatory molecules such as miRNAs (Backes et al. 2016), or a combination of these. Better yet, regulatory interactions can be implied by using molecular data generated within the study. The most common approach is using the genome-wide transcript abundance to construct co-expression networks (Langfelder and Horvath 2008). Alternatively, one could look at shared epigenetic markings, such as histone modifications (Rintisch et al. 2014) or DNA methylation (Gutierrez-Arcelus et al. 2013, 2015; Lappalainen et al. 2013), to infer regulatory relationships between genes.

Historically, the statistical framework underlying QTL analyses has been based on standard linear regression approaches, where the individual relation between each variant and a trait is tested independently. Besides feeding into a substantial multiple testing problem, which can be partly compensated by appropriate methods (Johnson et al. 2010), it ignores the linkage disequilibrium (LD) between SNPs (i.e., *cis* inter-dependence), as well as the presence of epistatic effects (i.e., *trans* inter-dependence). Bayesian approaches offer a means for true multi-variate statistical analysis that considers these aspects. Bayesian approaches for QTL analysis have been successfully applied to genetic association studies in animal models (Bottolo et al. 2011a, b). While statistically more robust and—by taking into account interdependence of genetic variants—more biologically sound than standard one-by-one linear regression approaches, a clear limitation is the required computational time. Hence, while it has been successfully applied in animal models (Bottolo et al. 2011b; Heinig et al. 2010; Moreno-Moral et al. 2013; Petretto et al. 2008, 2010; Rintisch et al. 2014), the genetic heterogeneity of human populations hampers the application of Bayesian methods.

Recent advances however have enabled the development of computationally accelerated and parallelized algorithms (Lewin et al. 2015), enabling the application of Bayesian methods to human genetic studies as well. Das et al. (2015) developed eQTeL, a Bayesian algorithm that infers *cis*

regulatory polymorphisms underlying gene expression variability by integrating (i) genotype and gene-expression variance across individuals; (ii) epigenetic data, (iii) DNase I hypersensitivity variance of SNPs and promoters, and (iv) expression variance of genes across multiple cell types; in addition to (v) LD blocks and (vi) imputed haplotypes inferred from the 1000 Genomes Project (Das et al. 2015). Importantly, eQTeL is scalable to large datasets.

Applications of systems genetics in the study of cardiovascular traits

Concerning systems genetics studies of complex traits in the cardiovascular domain, the HXB/BXH panel of recombinant inbred (RI) rat strains (Pravenec et al. 1989) remains a leading model (Heinig et al. 2010; Hubner et al. 2005; Moreno-Moral and Petretto 2016; Printz et al. 2003). The panel was generated by reciprocal crossing of the Brown Norway (BN) rat and the spontaneously hypertensive rat (SHR). Briefly, the construction of an RI panel involves mating two inbred genetically distant progenitor strains to produce F2 hybrids that carry a unique combination of maternal and paternal loci. Subsequently, individual homozygous RI strains are generated by inbreeding of randomly chosen pairs of F2 animals and brother-sister mating for more than 20 generations (Silver 1995). This results in a panel that offers a controllable, renewable resource that combines genetic identity within strains with genetic diversity across strains. A comprehensive overview of the HXB/BXH RI panel and its applications can be found in Moreno-Moral and Petretto (2016).

RI panels are particularly powerful for genetic studies, as the study of genetically identical biological replicates optimizes estimation of trait heritability by reducing environmental variance, while the constant genetic background within each RI strain allows for the accumulation of genetic, omics, and phenotypic data over time. For the HXB/BXH RI rat panel, genotyping of the RI strains has led to the identification of 1384 strain distribution patterns (SDPs) of single nucleotide polymorphisms (SNPs) for use in genetic mapping (Star-Consortium et al. 2008). Full genomic sequencing data is available for both parental strains (Atanur et al. 2010; Simonis et al. 2012), in addition to RNA-seq-based transcript abundance data for multiple tissues across the panel (Johnson et al. 2014) and a plethora of cardiovascular physiological traits such as blood pressure and cardiac mass (Pravenec et al. 1995). The HXB/BXH RI panel therefore lends itself exceptionally well to the study of complex cardiovascular traits.

One of the first systems genetics studies performed within the HXB/BXH panel was published in 2008 by Petretto et al. (2008). In that study on the genetics underlying left ventricular mass (LVM), a combination of QTL and transcriptomics analysis pinpointed a single gene, *Ogn*, as a major candidate

regulator. Two years later, Heinig et al. (2010) through combined analyses of gene networks and genetic variation identified an interferon regulatory factor 7 (*Irf7*)-driven inflammatory network enriched for viral response genes. The expression of genes in this network was shown to be regulated in multiple tissues by a single locus on chromosome 15. Subsequent analysis of the orthologous region and orthologous genes in human demonstrated association with type 1 diabetes for both, implicating the implicate *Irf7* network genes and their regulatory locus in the pathogenesis of this complex cardiometabolic disease. Petretto et al. (2010) extended this approach by hunting for eQTL hotspots, loci that modulate gene expression across a wide variety of tissues, using a multi-variate Bayesian approach (Bottolo et al. 2011b; Lewin et al. 2015). They demonstrated common genetic regulation of gene expression across four tissues for approximately 27% of all expressed transcripts, providing a more than fivefold increase in eQTL detection rate compared to single tissue analyses. Further extending this integrative approach, Morrissey et al. (2011) combined a quantitative trait transcript (QTT) analysis, i.e. genes that correlate with clinically relevant cardiometabolic traits such as systolic blood pressure and blood glucose, with *cis* eQTLs and cardiometabolic trait QTLs. They proposed that these co-localizing correlated *cis* eQTLs (c3-eQTLs) are highly attractive for prioritization and identified multiple candidate genes as strong positional candidates for the investigated cardiometabolic traits. Lastly, Moreno-Moral et al. (2013) combined left ventricular myocardium co-expression networks with histological and histomorphometric data of the heart and coronary vasculature, to provide a large catalog of gene co-expression networks in the heart that are significantly associated with quantitative variation in left ventricular hypertrophy, microvascular remodeling, and fibrosis-related traits. Additionally, they demonstrated the relevance of co-expression networks identified in rat for human heart disease, by showing that many of these networks were shown to be significantly conserved in left ventricular myocardium of human idiopathic and ischemic cardiomyopathy patients.

More recently, the HXB/BXH panel was used to study the interplay between genetic variation and the presence of specific histone modifications (Rintisch et al. 2014). These epigenetic marks known to be important in the modulation of biological processes, including regulation of gene expression through control of chromatin structure. The study by Rintisch et al. (2014) focused on the histone methylation marks H3K4me3 (associated with active gene promoters), H3K27me3 (associated with silenced gene promoters, and more specifically Polycomb-repressed regions), H3K4me1 (associated with gene promoter and enhancer regions), and H4K20me1 (associated with recently transcribed genomic regions). Additionally, genome-wide genetic variation and left ventricular transcript abundance were assessed for the entire

panel. The results of this study showed that genetic variants associated with differential histone modifications, referred to as histoneQTLs, are a predictor of gene expression, with the most prominent QTL identified influencing H3K4me3 levels at 899 gene promoters in the heart, proving the important role of histone modifications in genotype-phenotype relationships.

One of our own recent efforts (*under review*) pertaining to the HXB/BXH RI rat panel involved mapping genetic loci impacting on heart rate and ECG indices of cardiac conduction. QTL analysis on ECG parameters identified two QTL for PR-interval, respectively, on chromosomes 10 and 17. By further intersection of this data with cardiac eQTL data, we identified novel candidate genes associated with cardiac conduction. Subsequently expanding these findings by constructing co-expression networks enabled us to identify multiple gene networks associated with cardiac conduction, shown to be enriched for cardiac related processes through Gene Ontology enrichment analysis.

However powerful, translating genetic findings from animal models to human without functional validation is by nature limited through inference (Heinig et al. 2010; Petretto et al. 2008). Several large studies in human have focused on investigating gene expression regulation in the heart. Koopmann et al. (2014) reported on an eQTL analysis of human left ventricular tissue of 129 non-diseased donors, creating the first comprehensive eQTL map of human heart. The study identified 771 eQTLs regulating 429 unique genes, while overlaying these eQTLs with cardiovascular trait, GWAS loci identified novel candidates for follow-up studies to further elucidate the functional and transcriptional impact of these loci. In the same year, Lin et al. (2014) reported on the characterization of gene expression and genetic variation in 53 human left atrial and 52 right atrial tissue samples collected from the Myocardial Applied Genomics Network (MAGNet) repository. Next to identifying a distinct transcriptional profile between the right and left atrium and extensive *cis* associations between atrial transcripts and common genetic variants, their approach identified a causative gene for AF, *MYOZ1*.

More recently, the study from Heinig et al. (2017) reports on the first in-depth analysis of the dilated cardiomyopathy (DCM) transcriptome. Myocardial ischemia as well as toxic, metabolic immunologic factors can lead to the DCM phenotype. Moreover, genetic susceptibility plays an important role, with at least 23% of DCM cases being familial and more than 50 genes linked to inherited DCM. By performing RNA-sequencing analysis in left ventricular biopsies or 97 patients with DCM and 108 non-diseased controls, they revealed extensive differences of gene expression and splicing between patients and controls, in addition to a widespread effect of genetic variation on the regulation of transcription, isoform usage, and allele-specific expression. Strikingly, these differences did not just pertain to known DCM genes but extended

to a wide range of novel candidates. Overlaying their findings with genome-wide association, SNPs identified 60 functional candidate genes for cardiovascular traits. This represented 20% of all published heart genome-wide association loci at the time, a much larger set of candidates than previously identified (Koopmann et al. 2014). Finally, it was shown that eQTL variants are also enriched for dilated cardiomyopathy genome-wide association signals in two independent cohorts. Their conclusion that genetic regulation of RNA transcription, splicing, and allele-specific expression are important determinants for the DCM phenotype likely extends to many other cardiovascular traits.

Taken together, these studies suggest that combining genome-wide genotyping with transcript abundance data forms the powerful foundations of systems genetic approaches.

Conclusions and prospects

Throughout the genetic studies discussed within this review, the loci and gene networks associated with the complex cardiovascular traits provide a starting point for future studies with the potential of identifying novel underlying mechanisms. The take-home message is that the main conclusion of all these studies could not have been reached without the applied integrative approaches: combining multiple measures of transcriptional control in a presence of a complex genetic background with multiple multi-variate analyses techniques.

Lastly, early detection combining molecular markers and clinical parameters is pivotal for preventive medicine in cardiology, but due to the heterogeneity of many diseases, this has proven to be exceptionally difficult. In the recent paper by Heinig et al. (2017), by exploiting the allele specificity of RNA-sequencing through allele-specific expression (ASE) analysis, it was shown that in DCM hearts, dysregulation occurs of only one of the two alleles for known DCM progressing genes. Allelic imbalance was for example observed for *TTN*, a gene for which many truncating variants have been identified that can lead to nonsense-mediated decay (Schafer et al. 2017). Strikingly, there was no clear pattern emerging, in line with earlier observations that were made in a subset of DCM samples (Roberts et al. 2015). For the DCM phenotype, this could indicate that imbalance shifts towards disease-contributing alleles during disease progression. Looking beyond *TTN*, both the extent of the allelic imbalance and the number of known DCM genes affected varied greatly from patient to patient, yet all differential allelic imbalanced genes combined showed enrichment for DCM-related processes, as well as differential splicing and miRNA interference. This suggests a complex, heterogeneous genetic dosage effect across individuals, that in many cases will not be identifiable by standard GWAS, eQTL, and sQTL analyses.

Additionally, this could suggest that such allelic imbalances are related to disease stage and may accumulate over time by interplay between genetic and epigenetic mechanisms, enabling early detection. Contrasting regular gene expression analysis, allelic imbalance can be reliably measured within a single individual, and several recent developments enable a statistically robust analysis (Knowles et al. 2017; Leon-Novelo et al. 2014). Taken together, ASE analysis has the potential to enable a full personalized, quantitative genetic typing for clinical prognosis that goes beyond established qualitative assays detecting solely the presence of genetic variants, as well as enable studying the interaction between genetic and epigenetic regulation in the establishment of the imbalance for complex cardiovascular traits.

Funding information M.E.A. is supported by the Dutch Province of Limburg through the Maastricht Centre for Systems Biology at Maastricht University. C.R.B. is supported by the Dutch Heart Foundation (CVON PREDICT and CONCOR-genes projects) and by a VICI grant from the Netherlands Organization for Scientific Research (016.150.610).

Compliance with ethical standards

Conflicts of interest M.E. Adriaens declares that he has no conflict of interest. C.R. Bezzina declares that she has no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alexa A, Rahnenfuhrer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22:1600–1607. <https://doi.org/10.1093/bioinformatics/btl140>
- Arking DE et al (2006) A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization. *Nat Genet* 38:644–651. <https://doi.org/10.1038/ng.1790>
- Arking DE et al (2011) Identification of a sudden cardiac death susceptibility locus at 2q24.2 through genome-wide association in European ancestry individuals. *PLoS Genet* 7:e1002158. <https://doi.org/10.1371/journal.pgen.1002158>
- Arking DE et al (2014) Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. *Nat Genet* 46:826–836. <https://doi.org/10.1038/ng.3014>
- Atanur SS et al (2010) The genome sequence of the spontaneously hypertensive rat: analysis and functional significance. *Genome Res* 20:791–803. <https://doi.org/10.1101/gr.103499.109>

- Backes C, Khaleeq QT, Meese E, Keller A (2016) miEAA: microRNA enrichment analysis and annotation. *Nucleic Acids Res* 44:W110–W116. <https://doi.org/10.1093/nar/gkw345>
- Benjamin EJ, Levy D, Vaziri SM, D'Agostino RB, Belanger AJ, Wolf PA (1994) Independent risk factors for atrial fibrillation in a population-based cohort. The Framingham Heart Study. *Jama* 271:840–844
- Bezzina CR et al (2010) Genome-wide association study identifies a susceptibility locus at 21q21 for ventricular fibrillation in acute myocardial infarction. *Nat Genet* 42:688–691. <https://doi.org/10.1038/ng.623>
- Bottolo L et al (2011a) ESS++: a C++ objected-oriented algorithm for Bayesian stochastic search model exploration. *Bioinformatics* 27:587–588. <https://doi.org/10.1093/bioinformatics/btq684>
- Bottolo L, Petretto E, Blankenberg S, Cambien F, Cook SA, Tiret L, Richardson S (2011b) Bayesian detection of expression quantitative trait loci hot spots. *Genetics* 189:1449–1459. <https://doi.org/10.1534/genetics.111.131425>
- Butler AM et al (2012) Novel loci associated with PR interval in a genome-wide association study of 10 African American cohorts. *Circ Cardiovasc Genet* 5:639–646. <https://doi.org/10.1161/CIRCGENETICS.112.963991>
- Castel SE, Levy-Moonshine A, Mohammadi P, Banks E, Lappalainen T (2015) Tools and best practices for data processing in allelic expression analysis. *Genome Biol* 16:195. <https://doi.org/10.1186/s13059-015-0762-6>
- Chambers JC et al (2010) Genetic variation in SCN10A influences cardiac conduction. *Nat Genet* 42:149–152. <https://doi.org/10.1038/ng.516>
- Cho YS et al (2009) A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat Genet* 41:527–534. <https://doi.org/10.1038/ng.357>
- Christophersen IE et al. (2017a) Fifteen genetic loci associated with the electrocardiographic P wave. *Circ Cardiovasc Genet* 10 doi:<https://doi.org/10.1161/CIRCGENETICS.116.001667>
- Christophersen IE et al (2017b) Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. *Nat Genet* 49:946–952. <https://doi.org/10.1038/ng.3843>
- Civelek M, Lusis AJ (2014) Systems genetics approaches to understand complex traits. *Nat Rev Genet* 15:34–48. <https://doi.org/10.1038/nrg3575>
- Das A et al (2015) Bayesian integration of genetics and epigenetics detects causal regulatory SNPs underlying expression variability. *Nat Commun* 6:8555. <https://doi.org/10.1038/ncomms9555>
- van den Boogaard M et al. (2014) A common genetic variant within SCN10A modulates cardiac SCN5A expression *The J Clin Investig* 124:1844–1852 doi:<https://doi.org/10.1172/JCI73140>
- Deo R et al (2013) Common genetic variation near the connexin-43 gene is associated with resting heart rate in African Americans: a genome-wide association study of 13,372 participants. *Heart Rhythm* 10:401–408. <https://doi.org/10.1016/j.hrthm.2012.11.014>
- van der Harst P et al (2016) 52 genetic loci influencing myocardial mass. *J Am Coll Cardiol* 68:1435–1448. <https://doi.org/10.1016/j.jacc.2016.07.729>
- Dixon JR, Gorkin David U, Ren B (2016) Chromatin domains: the unit of chromosome organization. *Mol Cell* 62:668–680. <https://doi.org/10.1016/j.molcel.2016.05.018>
- Eijgelsheim M et al (2010) Genome-wide association analysis identifies multiple loci related to resting heart rate. *Hum Mol Genet* 19:3885–3894. <https://doi.org/10.1093/hmg/ddq303>
- Ellinor PT et al (2010) Common variants in KCNN3 are associated with lone atrial fibrillation. *Nat Genet* 42:240–244. <https://doi.org/10.1038/ng.537>
- Ellinor PT et al (2012) Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat Genet* 44:670–675. <https://doi.org/10.1038/ng.2261>
- ENCODE-Project-Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74. <https://doi.org/10.1038/nature11247>
- Eppinga RN et al (2016) Identification of genomic loci associated with resting heart rate and shared genetic predictors with all-cause mortality. *Nat Genet* 48:1557–1563. <https://doi.org/10.1038/ng.3708>
- Evans DS et al (2016) Fine-mapping, novel loci identification, and SNP association transferability in a genome-wide association study of QRS duration in African Americans. *Hum Mol Genet* 25:4350–4368. <https://doi.org/10.1093/hmg/ddw284>
- Floyd JS et al (2018) Large-scale pharmacogenomic study of sulfonylureas and the QT, JT and QRS intervals: CHARGE Pharmacogenomics Working Group. *Pharmacogenomics J* 18:127–135. <https://doi.org/10.1038/tpj.2016.90>
- GTEx-Consortium (2013) The genotype-tissue expression (GTEx) project. *Nat Genet* 45:580–585. <https://doi.org/10.1038/ng.2653>
- GTEx-Consortium et al. (2017) Genetic effects on gene expression across human tissues. *Nature* 550:204–213. <https://doi.org/10.1038/nature24277>
- Gudbjartsson DF et al (2007) Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature* 448:353–357. <https://doi.org/10.1038/nature06007>
- Gudbjartsson DF et al (2009) A sequence variant in ZFX3 on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat Genet* 41:876–878. <https://doi.org/10.1038/ng.417>
- Gutierrez-Arcelus M et al (2013) Passive and active DNA methylation and the interplay with genetic variation in gene regulation. *eLife* 2:e00523. <https://doi.org/10.7554/eLife.00523>
- Gutierrez-Arcelus M et al (2015) Tissue-specific effects of genetic and epigenetic variation on gene regulation and splicing. *PLoS Genet* 11:e1004958. <https://doi.org/10.1371/journal.pgen.1004958>
- Hammerschlag AR et al (2017) Genome-wide association analysis of insomnia complaints identifies risk genes and genetic overlap with psychiatric and metabolic traits. *Nat Genet* 49:1584–1592. <https://doi.org/10.1038/ng.3888>
- Heinig M et al (2010) A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. *Nature* 467:460–464. <https://doi.org/10.1038/nature09386>
- Heinig M et al (2017) Natural genetic variation of the cardiac transcriptome in non-diseased donors and patients with dilated cardiomyopathy. *Genome Biol* 18:170. <https://doi.org/10.1186/s13059-017-1286-z>
- Holm H et al (2010) Several common variants modulate heart rate, PR interval and QRS duration. *Nat Genet* 42:117–122. <https://doi.org/10.1038/ng.511>
- Hong KW et al (2014) Identification of three novel genetic variations associated with electrocardiographic traits (QRS duration and PR interval) in East Asians. *Hum Mol Genet* 23:6659–6667. <https://doi.org/10.1093/hmg/ddu374>
- Hubner N et al (2005) Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nat Genet* 37:243–253. <https://doi.org/10.1038/ng1522>
- Johnson RC, Nelson GW, Troyer JL, Lautenberger JA, Kessing BD, Winkler CA, O'Brien SJ (2010) Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics* 11:724. <https://doi.org/10.1186/1471-2164-11-724>
- Johnson MD et al (2014) Genetic analysis of the cardiac methylome at single nucleotide resolution in a model of human cardiovascular disease. *PLoS Genet* 10:e1004813. <https://doi.org/10.1371/journal.pgen.1004813>
- Kerr KF et al (2017) Genome-wide association study of heart rate and its variability in Hispanic/Latino cohorts. *Heart Rhythm* 14:1675–1684. <https://doi.org/10.1016/j.hrthm.2017.06.018>
- Kim JW et al (2012) A common variant in SLC8A1 is associated with the duration of the electrocardiographic QT interval. *Am J Hum Genet* 91:180–184. <https://doi.org/10.1016/j.ajhg.2012.05.019>

- Knowles DA et al (2017) Allele-specific expression reveals interactions between genetic variation and environment. *Nat Methods* 14:699–702. <https://doi.org/10.1038/nmeth.4298>
- Koopmann TT et al (2014) Genome-wide identification of expression quantitative trait loci (eQTLs) in human heart. *PLoS One* 9. <https://doi.org/10.1371/journal.pone.0097380>
- Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. <https://doi.org/10.1186/1471-2105-9-559>
- Lappalainen T et al (2013) Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* 501:506–511. <https://doi.org/10.1038/nature12531>
- Lee JY et al (2017) Korean atrial fibrillation network genome-wide association study for early-onset atrial fibrillation identifies novel susceptibility loci. *Eur Heart J* 38:2586–2594. <https://doi.org/10.1093/eurheartj/ehx213>
- Leon-Novelo LG, McIntyre LM, Fear JM, Graze RM (2014) A flexible Bayesian method for detecting allelic imbalance in RNA-seq data. *BMC Genomics* 15:920. <https://doi.org/10.1186/1471-2164-15-920>
- Lewin A et al (2015) MT-HESS: an efficient Bayesian approach for simultaneous association detection in OMICS datasets, with application to eQTL mapping in multiple tissues. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btv568>
- Lin H et al (2014) Gene expression and genetic variation in human atria. *Heart Rhythm* 11:266–271. <https://doi.org/10.1016/j.hrthm.2013.10.051>
- Low SK et al (2017) Identification of six new genetic loci associated with atrial fibrillation in the Japanese population. *Nat Genet* 49:953–958. <https://doi.org/10.1038/ng.3842>
- Ma L, Keinan A, Clark AG (2015) Biological knowledge-driven analysis of epistasis in human GWAS with application to lipid traits. *Methods Mol Biol* 1253:35–45. https://doi.org/10.1007/978-1-4939-2155-3_3
- Marroni F et al (2009) A genome-wide association scan of RR and QT interval duration in 3 European genetically isolated populations: the EUROSPAN project. *Circ Cardiovasc Genet* 2:322–328. <https://doi.org/10.1161/CIRCGENETICS.108.833806>
- Mezzavilla M et al (2014) Insight into genetic determinants of resting heart rate. *Gene* 545:170–174. <https://doi.org/10.1016/j.gene.2014.03.045>
- Moreno-Moral A, Petretto E (2016) From integrative genomics to systems genetics in the rat to link genotypes to phenotypes. *Dis Model Mech* 9:1097–1110. <https://doi.org/10.1242/dmm.026104>
- Moreno-Moral A, Mancini M, D'Amati G, Camici P, Petretto E (2013) Transcriptional network analysis for the regulation of left ventricular hypertrophy and microvascular remodeling. *J Cardiovasc Transl Res* 6:931–944. <https://doi.org/10.1007/s12265-013-9504-x>
- Moreno-Moral A, Pesce F, Behmoaras J, Petretto E (2017) Systems genetics as a tool to identify master genetic regulators in complex disease. *Methods Mol Biol* 1488:337–362. https://doi.org/10.1007/978-1-4939-6427-7_16
- Morrissey C et al (2011) Integrated genomic approaches to identification of candidate genes underlying metabolic and cardiovascular phenotypes in the spontaneously hypertensive rat. *Physiol Genomics* 43:1207–1218. <https://doi.org/10.1152/physiolgenomics.00210.2010>
- den Hoed M et al (2013) Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nat Genet* 45:621–631. <https://doi.org/10.1038/ng.2610>
- Nagy R et al (2017) Exploration of haplotype research consortium imputation for genome-wide association studies in 20,032 generation Scotland participants. *Genome Med* 9:23. <https://doi.org/10.1186/s13073-017-0414-4>
- Newton-Cheh C et al (2009) Common variants at ten loci influence QT interval duration in the QTGEN Study. *Nat Genet* 41:399–406. <https://doi.org/10.1038/ng.364>
- Nolte IM et al (2009) Common genetic variation near the phospholamban gene is associated with cardiac repolarisation: meta-analysis of three genome-wide association studies. *PLoS One* 4:e6138. <https://doi.org/10.1371/journal.pone.0006138>
- Nolte IM et al (2017) Genetic loci associated with heart rate variability and their effects on cardiac disease risk. *Nat Commun* 8:15805. <https://doi.org/10.1038/ncomms15805>
- Noordam R et al (2017) A genome-wide interaction analysis of tricyclic/tetracyclic antidepressants and RR and QT intervals: a pharmacogenomics study from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. *J Med Genet* 54:313–323. <https://doi.org/10.1136/jmedgenet-2016-104112>
- Petretto E et al (2008) Integrated genomic approaches implicate osteoglycin (Ogn) in the regulation of left ventricular mass. *Nat Genet* 40:546–552. <https://doi.org/10.1038/ng.134>
- Petretto E et al (2010) New insights into the genetic control of gene expression using a Bayesian multi-tissue approach. *PLoS Comput Biol* 6:e1000737. <https://doi.org/10.1371/journal.pcbi.1000737>
- Pfeuffer A et al (2009) Common variants at ten loci modulate the QT interval duration in the QTSCD Study. *Nat Genet* 41:407–414. <https://doi.org/10.1038/ng.362>
- Pfeuffer A et al (2010) Genome-wide association study of PR interval. *Nat Genet* 42:153–159. <https://doi.org/10.1038/ng.517>
- Pravenec M, Klir P, Kren V, Zicha J, Kunes J (1989) An analysis of spontaneous hypertension in spontaneously hypertensive rats by means of new recombinant inbred strains. *J Hypertens* 7:217–221
- Pravenec M et al (1995) Mapping of quantitative trait loci for blood pressure and cardiac mass in the rat by genome scanning of recombinant inbred strains. *J Clin Invest* 96:1973–1978. <https://doi.org/10.1172/JCI118244>
- Printz MP, Jirout M, Jaworski R, Alemayehu A, Kren V (2003) Genetic models in applied physiology. HXB/BXH rat recombinant inbred strain platform: a newly enhanced tool for cardiovascular, behavioral, and developmental genetics and genomics. *J Appl Physiol* 94:2510–2522. <https://doi.org/10.1152/japplphysiol.00064.2003>
- Rintisch C et al (2014) Natural variation of histone modification and its impact on gene expression in the rat genome. *Genome Res* 24:942–953. <https://doi.org/10.1101/gr.169029.113>
- Ritchie MD et al (2013) Genome- and phenome-wide analyses of cardiac conduction identifies markers of arrhythmia risk. *Circulation* 127:1377–1385. <https://doi.org/10.1161/CIRCULATIONAHA.112.000604>
- Roberts AM et al (2015) Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. *Sci Transl Med* 7:270ra276. <https://doi.org/10.1126/scitranslmed.3010134>
- Roider HG, Manke T, O'Keefe S, Vingron M, Haas SA (2009) PASTA A: identifying transcription factors associated with sets of co-regulated genes. *Bioinformatics* 25:435–442. <https://doi.org/10.1093/bioinformatics/btn627>
- Sano M et al (2014) 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. <https://doi.org/10.1093/hmg/ddu375>
- Schafer S et al (2017) Titin-truncating variants affect heart function in disease cohorts and the general population. *Nat Genet* 49:46–53. <https://doi.org/10.1038/ng.3719>
- Shabalín AA (2012) Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 28:1353–1358. <https://doi.org/10.1093/bioinformatics/bts163>
- Silver LM (1995) Mouse genetics: concepts and applications. Oxford University Press,
- Simonis M et al (2012) Genetic basis of transcriptome differences between the founder strains of the rat HXB/BXH recombinant inbred

- panel. *Genome Biol* 13:r31. <https://doi.org/10.1186/gb-2012-13-4-r31>
- Smith JG et al (2011) Genome-wide association studies of the PR interval in African Americans. *PLoS Genet* 7:e1001304. <https://doi.org/10.1371/journal.pgen.1001304>
- Smith JG et al (2012) Impact of ancestry and common genetic variants on QT interval in African Americans. *Circ Cardiovasc Genet* 5:647–655. <https://doi.org/10.1161/CIRCGENETICS.112.962787>
- Sniekers S et al (2017) Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat Genet* 49:1107–1112. <https://doi.org/10.1038/ng.3869>
- Sotoodehnia N et al (2010) Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. *Nat Genet* 42:1068–1076. <https://doi.org/10.1038/ng.716>
- Star-Consortium et al. (2008) SNP and haplotype mapping for genetic analysis in the rat. *Nat Genet* 40:560–566. <https://doi.org/10.1038/ng.124>
- Szklarczyk D et al (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 45:D362–D368. <https://doi.org/10.1093/nar/gkw937>
- Verweij N et al (2014) Genetic determinants of P wave duration and PR segment. *Circ Cardiovasc Genet* 7:475–481. <https://doi.org/10.1161/CIRCGENETICS.113.000373>
- Verweij N et al (2016) Twenty-eight genetic loci associated with ST-T-wave amplitudes of the electrocardiogram. *Hum Mol Genet* 25:2093–2103. <https://doi.org/10.1093/hmg/ddw058>
- Watanabe K, Taskesen E, van Bochoven A, Posthuma D (2017) Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 8:1826. <https://doi.org/10.1038/s41467-017-01261-5>
- Zerbino DR et al (2018) Ensembl 2018. *Nucleic Acids Res* 46:D754–D761. <https://doi.org/10.1093/nar/gkx1098>
- Zhou VW, Goren A, Bernstein BE (2011) Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet* 12:7–18. <https://doi.org/10.1038/nrg2905>