Fifty-two Genetic Loci Influencing Myocardial Mass

Pim van der Harst[1,2,3*†], MD, PhD, Jessica van Setten[4,5†], PhD, Niek Verweij[1†], PhD, Georg Vogler[6†], PhD, Lude Franke[2†], PhD, Matthew T. Maurano[7,8,9,10†], PhD, Xinchen Wang[11†], BSc, Irene Mateo Leach[1†], PhD, Mark Eijgelsheim[12,13], MD, PhD, Nona Sotoodehnia[14] MD, MPH, Caroline Hayward[15], PhD, Rossella Sorice[16], PhD, Osorio Meirelles[17], PhD, Leo-Pekka Lyytikäinen[18,19], MD, Ozren Polašek[20,21], MD, PhD, Toshiko Tanaka[22], PhD, Dan E. Arking[23], PhD, Sheila Ulivi[24], PhD, Stella Trompet[25,26], PhD, Martina Müller-Nurasyid[27,28,29,30], PhD, Albert V. Smith[31,32], PhD, Marcus Dörr[33,34], MD, Kathleen F. Kerr[35], PhD, Jared W. Magnani[36], MD, MSc, Fabiola Del Greco M.[37], PhD, Weihua Zhang[38,39], PhD, Ilja M. Nolte[40], PhD, Claudia T. Silva[41,42,43], MSc, Sandosh Padmanabhan[44], MD, PhD, Vinicius Tragante[4,5], PhD, Tõnu Esko[45,46], PhD, Gonçalo R. Abecasis[47], PhD, Michiel E. Adriaens[48,49], PhD, Karl Andersen[32,50], PhD, Phil Barnett[51], PhD, Joshua C. Bis[52], PhD, Rolf Bodmer[6], PhD, Brendan M. Buckley[53], MD, PhD, Harry Campbell[20], MD, Megan V. Cannon[1], PhD, Aravinda Chakravarti[23], PhD, Lin Y. Chen[54], MD, MS, Alessandro Delitala[55], PhD, Richard B. Devereux[56], MD, Pieter A. Doevendans[5], MD, PhD, Anna F. Dominiczak[44], MD, FRCP, FAHA, FESC, FRSE, FMediSci, Luigi Ferrucci[22], MD, PhD, Ian Ford[57], PhD, Christian Gieger[29,58,59], PhD, Tamara B. Harris[60], PhD, Eric Haugen[7], Matthias Heinig[61,62,63], PhD, Dena G. Hernandez[64], MSc, Hans L. Hillege[1], MD, PhD, Joel N. Hirschhorn[46,65,66],MD, PhD, Albert Hofman[12,13], MD, PhD, Norbert Hubner[61,67], MD, Shih-Jen Hwang[68], PhD, Annamaria Iorio[69], MD, Mika Kähönen[70,71], MD, PhD, Manolis Kellis[72,73], PhD, Ivana Kolcic[21], MD, PhD, Ishminder K. Kooner[39], MD, PhD, Jaspal S. Kooner[39,74], MBBS, FRCP, MD, Jan A. Kors[75], PhD, Edward G. Lakatta[76], MD, Kasper Lage[73,77,78], PhD, Lenore J. Launer[60], PhD, Daniel Levy[79], MD, Alicia Lundby[80,81], PhD, Peter W. Macfarlane[82], DSc, FRSE, Dalit May[83], MD, Thomas

Meitinger[30,84,85], MD, Andres Metspalu[45], PhD, Stefania Nappo[16], MSc, Silvia Naitza[55], PhD, Shane Neph[7], BS, Alex S. Nord[86,87], PhD, Teresa Nutile[16], PhD, Peter M. Okin[56], MD, Jesper V. Olsen[80], PhD, Ben A. Oostra[41], PhD, Josef M. Penninger[88], MD, Len A. Pennacchio[86,89], PhD, Tune H. Pers[46,65,90,91], PhD, Siegfried Perz[58,92], MSc, Annette Peters[30,58], PhD, Yigal M. Pinto[48], MD, PhD, Arne Pfeufer[37,93], MD, MSc, Maria Grazia Pilia[55], PhD, Peter P. Pramstaller[37,94,95], MD, Bram P. Prins[96], MSc, Olli T. Raitakari[97,98], MD, PhD, Soumya Raychaudhuri[99,100], MD, PhD, Ken M. Rice[35], PhD, Elizabeth J. Rossin[78,101], MD, PhD, Jerome I. Rotter[102], MD, Sebastian Schafer[61,67,103], PhD, David Schlessinger[17], PhD, Carsten O. Schmidt[104], PhD, Jobanpreet Sehmi[39,74], MRCP, PhD, MD, Herman H.W. Silljé[1], PhD, Gianfranco Sinagra[69], MD, FESC, Moritz F. Sinner[27] MD, MPH, Kamil Slowikowski[105], BS, Elsayed Z. Soliman[106], MD, MSc, Timothy D. Spector[107], MBBS, MRCP, MD, MSc, FRCP, Wilko Spiering[108], MD, PhD, John A. Stamatoyannopoulos[7], MD, Ronald P. Stolk[40], MD, PhD, Konstantin Strauch[28,29], PhD, Sian-Tsung Tan[39,74], MD, Kirill V. Tarasov[76], MD, PhD, Bosco Trinh[6], BSc, Andre G. Uitterlinden[12,13], PhD, Malou van den Boogaard[51], MD, Cornelia M. van Duijn[41], PhD, Wiek H. van Gilst[1], PhD, Jorma S. Viikari[109,110], MD, PhD, Peter M. Visscher[111,112], PhD, Veronique Vitart[15], PhD, Uwe Völker[34,113], PhD, Melanie Waldenberger[58,59], PhD, MPH, Christian X. Weichenberger[37], PhD, Harm-Jan Westra[65,114,115], PhD, Cisca Wijmenga[2], PhD, Bruce H. Wolffenbuttel[116], MD, PhD, Jian Yang[111], PhD, Connie R. Bezzina[48], PhD, Patricia B. Munroe[117,118], PhD, Harold Snieder[40], PhD, Alan F. Wright[15], MD, PhD, Igor Rudan[20], MD, PhD, Laurie A. Boyer[11], PhD, Folkert W. Asselbergs[3,5,119], MD, PhD, Dirk J. van Veldhuisen[1], MD, PhD, Bruno H. Stricker[12,13], MD, PhD, Bruce M. Psaty[120,121], MD, PhD, Marina Ciullo[16,122], PhD, Serena Sanna[55], PhD, Terho Lehtimäki[18,19], MD, PhD, James F. Wilson[15,20], PhD, Stefania Bandinelli[123], MSc, Alvaro Alonso[124], MD, PhD, Paolo Gasparini[24,125,126], MD, J. Wouter Jukema[25,127], MD, PhD, Stefan Kääb[27,30], MD, PhD, Vilmundur Gudnason[31,32], MD,

PhD, Stephan B. Felix[33,34], MD, Susan R. Heckbert[121,128], MD, PhD, Rudolf A. de Boer[1], MD, PhD, Christopher Newton-Cheh[65,129,130], MD, PHD, Andrew A. Hicks[37], PhD, John C. Chambers[38,39†], BA, MBBS, FRCP, PhD, Yalda Jamshidi[96†], PhD, Axel Visel[86,89,131†], PhD, Vincent M. Christoffels[51†], PhD, Aaron Isaacs[41,132†], PhD, Nilesh J. Samani[133,134†], MD, FRCP, Paul I.W. de Bakker[4,135†], PhD

Affiliations: are provided at the end of the manuscript

Word count: 4978

* To whom correspondence should be addressed:

Pim van der Harst

University of Groningen, University Medical Center Groningen,

Department of Cardiology & Department of Genetics

Hanzeplein 1, 9700RB Groningen, The Netherlands

Email: p.van.der.harst@umcg.nl

Abstract

BACKGROUND: Myocardial mass is a key determinant of cardiac muscle function and hypertrophy. Myocardial depolarization leading to cardiac muscle contraction is reflected by the amplitude and duration of the QRS complex on the electrocardiogram (ECG). Abnormal QRS amplitude or duration reflect changes in myocardial mass and conduction, and are associated with increased risk of heart failure and death.

OBJECTIVE: To gain insights into the genetic determinants of myocardial mass,

METHODS: We carried out a genome-wide association meta-analysis of 4 QRS traits in up to 73,518 individuals of European ancestry, followed by extensive biological and functional assessment.

RESULTS: We identified 52 genomic loci, of which 32 are novel, reliably associated with one or more QRS phenotypes at $P<1\times10^{-8}$. These loci are enriched in regions of open chromatin, histone modifications, and transcription factor binding suggesting that they represent regions of the genome that are actively transcribed in the human heart. Pathway analyses provide evidence that these loci play a role in cardiac hypertrophy. We further highlight 67 candidate genes at the identified loci that are preferentially expressed in cardiac tissue and associated with cardiac abnormalities in *Drosophila melanogaster* and *Mus musculus*. We validated the regulatory function of a novel variant in the *SCN5A/SCN10A* locus *in vitro* and *in vivo*.

CONCLUSIONS: Taken together, our findings provide new insights into genes and biological pathways controlling myocardial mass and may help identify novel therapeutic targets.

Keywords: Genetic association study, QRS, left ventricular hypertrophy, heart failure, electrocardiogram, genes

Abbreviations list

DHS	Deoxyribonuclease hypersensitivity sites
DNase	Deoxyribonuclease
ECG	Electrocardiogram
eQTL	Expression quantitative trait locus
FDR	False Discovery Rate
GWAS	Genome-wide association study
LD	Linkage Disequilibrium
RNAi	Ribonucleic acid interference
SNP	Single Nucleotide Polymorphisms
TF	Transcription Factor

Introduction

The role of the heart is to provide adequate circulation of blood to meet the body's requirements of oxygen and nutrients. The QRS complex on the ECG is the most widely used measurement of cardiac depolarization, which causes the ventricular muscle to contract, resulting in pulsatile blood flow. The amplitude and duration of the QRS complex reflects the conduction through the left ventricle and is well correlated with left ventricular mass as measured by echocardiography (1,2). ECG measurements of the QRS complex are important in clinical and preclinical cardiovascular diseases such as cardiac hypertrophy, heart failure, and various cardiomyopathies, and can also predict cardiovascular mortality(3-6).

Identification of specific genes influencing the QRS complex may thus enhance our understanding of the human heart and ultimately lead to prevention of cardiovascular disease and death. To further our understanding of the genetic factors influencing the QRS complex, we carried out a large scale GWAS and replication study of 4 related and clinically used QRS traits: the Sokolow-Lyon, Cornell and 12-lead-voltage duration products (12-leadsum), and QRS duration. We identified 52 loci that were subsequently interrogated using bioinformatics and experimental approaches to gain more insights into the biological mechanisms regulating cardiac mass and QRS parameters.

Methods

Additional details on the methods can be found in the **Supplementary Note**.

Genome wide analyses and replication testing

Our study design is summarized in **Fig. S1**. Briefly, we combined summary statistics from 24 studies for up to 2,766,983 autosomal SNPs using an inverse-variance fixed-effects meta-analysis for each QRS trait. We performed replication testing for loci showing suggestive association ($1\times10^{-8}<P<5\times10^{-7}$) (**table S1 and S2**). The threshold for genome-wide significance was set at $P<1\times10^{-8}$.

DNA Functional elements, coding variation and enrichment analyses

We performed an intersection between SNPs and regions of DHSs, covalently modified histones and genomic features (ChromHMM) of cardiac tissues mapped by the National Institutes of Health Roadmap Epigenomics Program, as well as various cardiac transcription factor binding sites (GATA4, MEF2, SRF, TBX5, TBX3, GATA4 and Nkx2-5) measured by Chip-seq.

Experimental cardiac enhancer studies

Single cell suspensions of human ventricular tissue was obtained by dissociation with IKA Ultra Turrax T5 FU, followed by dounce homogenization. 4C templates were mixed and sequenced simultaneously in one Illumina HiSeq 2000 lane. Enhancer candidate regions with major and minor allele for rs6781009 were obtained by PCR from human control DNA and cloned into the Hsp68-LacZ reporter vector. DNA was injected into the pronucleus of fertilized FVB strain egg and approximately 200 injections per construct were performed. Embryos were harvested, stained with X-gal to detect LacZ activity.

H10 cells, grown in 12-well plates in DMEM supplemented with 10% FCS (Gibco-BRL) and glutamine, were transfected using polyethylenimine 25 kDa (PEI, Brunschwick) at a 1:3 ratio (DNA:PEI). Transfections were carried out at least three times and measured in triplo. Luciferase measurements were performed using a Promega Turner Biosystems Modulus Multimode Reader luminometer.

Identification of candidate genes

We considered genes to be causal candidates based on 1) the nearest gene and any other gene located within 10kb of the sentinel SNP; 2) genes containing coding variants in LD with the ST-T wave SNPs at r^2>0.8; 3) GRAIL analyses using the 2006 dataset to avoid confounding by subsequent GWAS discovery, and 3) genes with an eQTL analyses in cis using 4 independent sets of cardiac left ventricle and blood tissues. Ingenuity Pathway Analysis (IPA) Knowledge Base March 2015 (Ingenuity Systems, CA, USA) was used to explore molecular pathways between proteins encoded by the 67 candidate genes from the 52 genome-wide significant loci.

Drosophila melanogaster and Mus Musculus methods

We queried a *D. melanogaster* dataset containing a genome-wide phenotypic screen of cardiac specific RNAi-silencing of evolutionarily conserved genes under conditions of stress. We also queried the international database resource for the laboratory mouse (MGI-Mouse Genome Informatics) and manually curated Mammalian Phenotypes (MP) identifiers related to cardiac phenotypes. To illustrate that prioritized genes may play a critical role in heart development we tested *CG4743/SLC25A26, Fhos/FHOD3, Cka/STRN, NACα/NACA*, *EcR/NR1H* and *Hand/HAND1* by performing heart-specific RNAi knockdown with the cardiac Hand4.2-Gal4 driver line.

Gene expression profiling and cardiomyocyte differentiation analysis

We collected 43,278 raw Affymetrix Human Genome U133 Plus 2.0 Arrays from the Gene Expression Omnibus (GEO) containing human gene expression data. RMA was used for normalization and subsequently conducted stringent quality control and processing of the data, which resulted in a tissue-expression matrix. After quality control 37,427 samples remained and assigned 54,675 different probesets to 19,997 different Ensembl genes used for human tissue expression profiling. To explore gene-expression of our candidate genes during cardiac differentiation we performed RNA-sequencing using E14 Tg(Nkx2-5-EmGFP) mouse embryonic stem cells cultured feeder-free conditions and subsequently differentiated.

Results

Large scale meta-analysis of genome wide association studies

Characteristics of studies, participants, genotyping arrays and imputation are summarized in **table S1 and S2**. Together our studies comprise 60,255 individuals of European ancestry ascertained in North America and Europe, with a maximum sample size of 54,993 for Sokolow-Lyon, 58,862 for Cornell, 48,632 for 12-leadsum, and 60,255 for QRS duration. Across the genome, 52 independent loci, 32 of which are novel, reached genome-wide significance for association with one or more QRS phenotypes (**Fig. 1**, **Fig. S2, table S3** and **Supplementary Note**). At each locus, we defined a single 'sentinel' SNP with the lowest *P*-value against any of the four phenotypes; regional association plots for the 52 loci are shown in **Fig. S3**. Among the 52 loci, 32 were associated with only one QRS phenotype, and 20 with at

least two phenotypes (**Fig. S4)**. The total number of locus-phenotype associations at $P<10^{-8}$ was 79 (72 SNPs), of which 59 are novel (**table S3**). Full lists of the sentinel SNPs and the SNPs associated with any phenotype at $P<10^{-6}$ are provided in **table S4 and S5**. All previously known QRS duration loci showed evidence for association ($P<10^{-6}$, **table S6**). Among the 32 novel loci, 8 demonstrated genome-wide significant association with Sokolow-Lyon, 9 with Cornell, 20 with 12-leadsum, and 9 with QRS duration (**table S5**). Collectively, the total variance explained by the 52 sentinel SNPs for the QRS traits was between 2.7% (Sokolow-Lyon) and 5.0% (QRS-duration) (**table S7**). At some loci we found evidence for multiple independent associations with QRS phenotypes at $P<10^{-8}$ in conditional analyses(7) (**table S8** and **Supplementary Note**). Among the 52 loci identified 8 have been associated previously with PR (reflecting atrial and atrioventricular node function), 5 with QT duration (ventricular repolarization) and 2 with heart rate (sinus node function) (**table S6**), indicating genetic overlap among the four cardiac measures studied. We further demonstrated that there was directional consistency of the association of common variants identified in this study with QRS phenotypes in other ethnic groups (**Fig. S1, table S9,** and **Supplementary note**).

Functional annotation of the QRS associations

To better capture common sequence variants at the 52 loci, we queried the 1000 Genomes Project dataset(8), and identified 41 non-synonymous SNPs in 17 genes that are in high LD ($r^2>0.8$) with 12 of the sentinel SNPs (**table S10**), which represent an initial set of candidate variants that may have a functional effect on the QRS phenotypes through changes in protein structure and function.

To assess the potential role of gene expression regulation, we tested the 52 loci for enrichment of deoxyribonuclease I (DNase I) hypersensitive sites (DHSs)(9). In an analysis across 349 diverse cell

lines, cultured primary cells and fetal tissues(10) mapped by the ENCODE project(11) and the National Institute of Health Roadmap Epigenomics Program(12), the majority (42 of 52) of sentinel SNPs were located in DHSs. In human fetal heart tissue we found that less than half (22 of 52) overlapped DHSs, which still represents a ~3.5-fold enrichment compared to the null expectation (P=7.7×10^{-12}, **Fig. 2A**). Further, the enrichment of genome-wide significant SNPs (P<10^{-8}) in DHSs was strongest within the first 100 bp around the sentinel variants (**Fig. 2B**). In addition, there was a strong enrichment for histone marks and chromatin states(13) associated with active enhancers, promoters and transcription in human heart; by contrast no enrichment was observed for transcriptionally repressive histone marks or states (**Fig. 2, C and D and Fig. S5**). Strikingly, we observed increasing enrichment of activating histone marks at the identified QRS loci during the process of differentiating mouse embryonic stem cells into cardiomyocytes (**Fig. S6**). Altogether, these findings are consistent with earlier observations of selective enrichment of trait-associated variants within DHSs of specific cell of tissue types(10), and point to a regulatory role of the QRS-associated loci during cardiac development.

We next surveyed our genome-wide significant SNPs in DHSs for perturbation of transcription factor (TF) recognition sequences, since these sites can point directly to binding events (**Supplementary Note**). Of the 22 sentinel SNPs in human fetal heart DHSs, 11 are predicted to alter TF recognition sequences (**table S11**). When considering all genome-wide significant SNPs (P<10^{-8}) as well as those in high LD (r^2>0.8), 402 SNPs in the co-localizing DHSs perturb transcription recognition sequences, including those of important cardiac and muscle developmental regulators like TBX, GATA-4, and MEF2. When we intersected the GWAS results with ChIP-seq data from mouse and human cardiac tissue(14-16), we found enrichment in enhancers marked by p300, sites bound by RNA Polymerase II (RNAP2), and the transcription factors NKX2-5, GATA-4, TBX3, TBX5, and SRF (**Fig. 2E**). Nine of

our 52 loci contained not only fetal heart DHSs but also ChIP-seq validated TF binding sites. SNPs overlapping TF binding sites were 5.65 fold enriched within DHSs (P=9.0×10^{-10}) but not outside DHSs (P=0.20). The associations of the 52 sentinel SNPs with all tested functional elements are summarized in **Fig. 1**. We validated several candidate regulatory regions identified above as heart enhancers *in vivo*. Activity of 4 exemplar novel human cardiac enhancers in embryonic transgenic mice stained for *LacZ* enhancer reporter activity are shown in **Fig. 3A**. Recently, rs6801957 (**Fig. 1**) in the *SCN5A/SCN10A* locus was reported to influence the activity of a regulatory element affecting SCN5A expression(16,17). Conditional analysis (**table S8**) revealed that rs6781009 (at 180-kb from the sentinel) is an additional novel independent signal at this locus. Our follow-up in silico and experimental results (**Fig. 3**) indicate the presence of *in vivo* heart enhancers in genome regions associated with QRS traits.

Identification of candidate genes

Across the 52 loci, 974 annotated genes are located within 1 Mb of all sentinel SNPs. Among these genes, we prioritized potential candidates using an established complementary strategy (18,19); we chose (i.) Genes nearest to the sentinel SNP, and any other genes within 10kb (56 genes; **Fig. 1**); (ii.) Genes containing a non-synonymous SNP in high LD (r^2>0.8) with the sentinel SNP (11 genes; **table S10**); (iii.) Protein-coding genes with *cis*-eQTL associated with sentinel SNP (14 genes; **table S12**), and (iv.) GRAIL literature analysis(20) (16 genes **table S13**) with 'cardiac', 'muscle' and 'heart' as the top 3 keywords describing the observed functional connections. In total, this strategy identified 67 candidate genes at the 52 loci (**Fig. 1**). Pathway analysis confirmed that the list of 67 candidate genes is strongly enriched for genes known to be involved in cardiovascular and muscular system development and function (P=1×10^{-56}; **table S14 and S15**). We have summarized the available functional annotations for all 67 candidates in **table S16**, including established links from the Online Mendelian Inheritance in

Man (OMIM) between candidate genes and familial cardiomyopathies (*TNNT2*, *TTN*, *PLN*, *MYBPC3*) and cardiac arrhythmias (*CASQ2*). We also identified genes that are associated with atrial septal defects (*TBX20*) and more complex syndromes involving cardiac abnormalities such as the Schinzel-Giedion midface retraction syndrome (*SETBP1*)(21) and the ulnar-mammary syndrome (*TBX3*)(22).

Insights from gene expression profiling and model organisms

We explored gene expression profiles of our candidate genes in data derived from 37,427 Affymetrix U133 Plus 2.0 arrays across 40 annotated tissues. We could reliably assign a probe for 63 of our 67 candidate genes. On average expression levels for these transcripts were higher in cardiac-derived samples compared to other transcripts in the same sample (P=9.8×10^{-6} for heart tissue; Wilcoxon test; **Fig. S7**) and also when compared to the same transcripts in other tissues (P=0.005 after Bonferroni correction; **Fig. S8**). To further investigate the potential role of these candidate genes in cardiac development, we assessed temporal gene expression patterns during *in vitro* differentiation of mouse embryonic stem cells (ESC) via mesoderm (MD) and cardiac precursor (CP) cells to cardiomyocytes (CM). Seven percent of genes are mainly expressed during the ESC stage, 22% during MD stage, 7% in the CP stage and 64% in the cardiomyocyte stage. Compared to other genes, the candidate genes were more highly expressed in cardiomyocytes (P=5.4×10^{-8}, Wilcoxon test; **Fig. S9**). These results suggest that our candidate gene set is enriched for genes differentially expressed in cardiac tissue and increasingly expressed during cardiac development.

Next, we analyzed data from model organisms to explore the function of the selected candidate genes. From cardiac tissue-specific RNAi knockdown data collected in *D. melanogaster*, we found that the 67 candidate genes were 2.3-fold enriched for stress-induced cardiac death (9 genes, P=1.84×10^{-2}; **Fig. S10**). To illustrate that prioritized genes may play a critical role in heart development we tested 4

(*CG4743/SLC25A26, Fhos/FHOD3, Cka/STRN, NACα/NACA*) of these 9 genes with unknown cardiac function by performing heart-specific RNAi knockdown with the cardiac Hand4.2-Gal4 driver line. We also re-tested *EcR/NR1H*, which has multiple homologous genes in mammals, as well as *Hand*/HAND1 as this gene was only tested in as a full-knockout in early development but not in adult *D. melanogaster* heart using cardiac specific knockdown. Adult hearts of Cka/STRN, *NACα*/NACA, and *EcR*/NR1H RNAi showed severe cardiac defects (**Fig. 4**). Knockdown of *Hand*/HAND1 and *Cka*/STRN both had a reduced cardiac heart rate. We also expanded on gene-by-gene analysis and identified 6 further genes causing cardiac abnormalities (**Supplementary Note** and **table S17**). From the Mouse Genome Informatics database, knockout models were annotated for 45 orthologues of the 67 candidate genes, of which 18 (40%) revealed a cardiac phenotype (**table S16**). This represents a 5.2-fold enrichment compared to randomly matched sets of 67 genes (P=3.4×10^{-14}; **Fig. S10**). Given the evolutionary conservation the observed heart phenotypes in these model organisms suggest potentially important roles for the significant GWAS loci in electrical and contractile properties of the human heart.

Interestingly, the 11p11.2 locus harbors multiple candidate genes (**Fig. 1**), including *MYBPC3*, *ACP2*, *MADD*, and *NR1H3*. *MYBPC3* deficiency is well established to cause hypertrophic and dilated cardiomyopathies in both human and mouse models and thus represents a plausible candidate gene (**table S16**). In addition to *MYBPC3*, eQTL and histone modification data also suggests a potential role for *NR1H3* (**Fig. S11**), as decreased expression of *NR1H3* was associated with higher QRS voltages. However, *NR1H3* deficient mice do not spontaneously develop a cardiac hypertrophic phenotype (MGI: 1352462). To study the potential cardiac effects of *NR1H3*, we created a transgenic mouse with cardiac-specific overexpression of *NR1H3* under the control of the *Myh6* promoter and found a diminished susceptibility to perturbations such as transverse aortic constriction and angiotensin II infusion that

provoke cardiac hypertrophy(23). This observation is in line with protective effects due to treatment with T0901317, a synthetic *NR1H3* agonist, in mice challenged with aortic constriction(24). These data highlight the importance of systematic approaches to identify causal genes beyond well-known candidates.

Insights from Data-Driven Expression-Prioritized Integration for Complex Traits (DEPICT)

As a complementary approach we employed the newly developed computational tool DEPICT(25) to analyze functional connections among associated loci (**Supplementary Note**). Enrichment of expression in 209 particular tissues and cell types identified heart and heart ventricles as the most relevant tissue for our association findings (**Fig. 5A; table S18**) and identified 404 significantly (FDR <5%) enriched gene sets (**table S19**). Comparing the names of these sets with those of the remaining 14,057 gene sets showed an over-representation of the common key words ‘Abnormal’, 'Muscle', 'Heart', 'Cardiac', 'Morphology' (**table S20**). We investigated similarities among gene sets by clustering them on the basis of the correlation between scores for all genes (**Supplementary Note**). Many of the resulting 43 meta-gene sets are correlated and relevant to cardiac biology (**Fig. 5B**). As an example we show the correlation structure within the second most significant meta-gene set “Dilated Heart Left Ventricle” (**Fig. S12**). When prioritizing genes based on functional similarities among genes from different associated regions DEPICT identified 35 genes (FDR<5%) at 27 of the 52 loci (**Fig. 1, table S21**).

Discussion

In this study, we performed a meta-analysis of GWAS in 73,518 individuals for 4 quantitative QRS phenotypes and identified 52 independent genetic loci influencing these traits with 79 locus-phenotype associations; the majority of these discoveries are novel. Our loci are co-localized with open chromatin, histone modification, and TF binding sites specifically in cardiac tissue, and contain *in vivo* functional enhancers. We also provide direct evidence that rs6781009, located in a cardiac enhancer, interacts with the promoter of SCN5A to modify expression levels. Based on multiple criteria, we defined a core set of 67 candidate genes which we believe are likely to influence cardiac mass and function. We have provided several exemplar experiments to further support this hypothesis.

We identified a number of loci containing genes that are directly or indirectly key the function of cardiomyocytes and cardiac function. *TTN*, *MYBPC3*, *TNNT2, SYNPO2L,* and *MYH7B* are essential components of the cardiac sarcomere; *PLN, CTNNA3, PRKCA, CASQ2,* and *STRN* are also examples of genes essential for cardiac myocyte function; while several key cardiac transcription factors are prominently involved in cardiac muscle and tissue development such as *MEF2D, HAND1, TBX20, TBX3 and NACA.* The abundance of candidate genes known to be involved in cardiac muscle function strengthen the hypothesis that the easily obtainable QRS-voltage phenotypes of the electrocardiogram are effective in capturing unknown loci that harbor genes that are likely to play an important role in left ventricular mass but are currently not well understood. The co-localization of our genetic loci with regulatory DNA elements (e.g. enhancers, promoters and transcription factor binding sites) that are active in cardiac tissues further support the relevance of the genes within these loci. The current work was not designed to provide an explanation for association of each loci and each individual gene. It is

clear that future translational efforts should be undertaken to resolve the causal genes and exact molecular machinery resulting the in the phenotype should consider mapping effect of genetic variants on the these functional elements at each of the identified loci. Nevertheless we have provided some exemplar preliminary elements to provide some early insights in strategies that can be undertaken to follow-up our findings. For example, we performed a series of experiments to demonstrate *in vivo* effects of rs6781009 on expression. Dedicated experiments might also elucidate loci containing effects on multiple plausible genes. In one of our loci we identified a very strong candidate gene (*MYBPC3*), well known to be involved in hypertrophic cardiomyopaties. However, using additional layers of information derived from gene expression and histone modifications we also considered *NR1H3* and were able to link overexpression of this gene to cardiac protection of hypertrophy. These examples fuel our expectation that the presented shortlist of SNP associations and the identified candidate genes provided in this work are a valuable resource that will help to prioritize and guide future translational studies to further our knowledge on the (patho)physiology of cardiac hypertrophy.

Our findings do have some limitations that warrant consideration. As for all current GWAS, we have only studied a finite number (~2.8 million) marker on the genome. Further fine mapping studies might be required to narrow the signal of association even further and to identify the potential causal variants with higher accuracy. Also additional exome focused arrays or whole genome sequencing might lead to a stronger signal within a locus or to multiple additional independent signals within a locus. To understand genetic mechanisms and to identify candidate genes, we have studied eQTLs. Although we studied the largest set of human cardiac eQTL available to date, the absolute number of studied samples is relatively small compared to eQTL data available in easily accessible peripheral blood. Finally, our electrocardiographic indices are generally considered markers of cardiac hypertrophy, they may also

reflect electrical remodeling of the action potential and not mass per se. Nevertheless, the variables studied here harbor important prognostic information, independently from cardiac mass parameters as assessed by echocardiography(26). This further underscores the relevance of the trait studied and the importance of understanding its genetic determinants.

In conclusion, we have identified 52 genomic loci, of which 32 are novel, associated with electrically active cardiac mass, prioritized 67 candidate genes and showed their relevance in cardiac biology using bioinformatics approaches and performed *in-vitro* and *in-vivo* experiments, going beyond the classical GWAS approach. To facilitate and accelerate future studies aimed at a better understanding of cardiac hypertrophy, heart failure and related diseases, we made our results of genome-wide associations publicly available.

Perspectives

COMPETENCY IN MEDICAL KNOWLEDGE: Abnormalities of cardiac mass are underlying many cardiovascular diseases such as heart failure. The lack of knowledge surrounding the basis of cardiomyocyte dysfunction and heart failure susceptibility is a major roadblock to understand risk for heart failure and to designing innovative strategies for therapy.

TRANSLATIONAL OUTLOOK: These findings will be a valuable resource for studying biological processes underlying cardiac mass, ultimately leading to prevention of cardiovascular disease and death.

References

1. Levy D, Labib SB, Anderson KM, Christiansen JC, Kannel WB, Castelli WP. Determinants of sensitivity and specificity of electrocardiographic criteria for left ventricular hypertrophy. Circulation. 1990;81:815-20.
2. Okin PM, Roman MJ, Devereux RB, Pickering TG, Borer JS, Kligfield P. Time-voltage QRS area of the 12-lead electrocardiogram: detection of left ventricular hypertrophy. Hypertension. 1998;31:937-42.
3. Kannel WB, Gordon T, Offutt D. Left ventricular hypertrophy by electrocardiogram. Prevalence, incidence, and mortality in the Framingham study. Ann Intern Med. 1969;71:89-105.
4. Verdecchia P, Schillaci G, Borgioni C, et al. Prognostic value of a new electrocardiographic method for diagnosis of left ventricular hypertrophy in essential hypertension. J Am Coll Cardiol. 1998;31:383-90.
5. Usoro AO, Bradford N, Shah AJ, Soliman EZ. Risk of mortality in individuals with low QRS voltage and free of cardiovascular disease. Am J Cardiol. 2014;113:1514-7.
6. Kamath SA, Meo Neto Jde P, Canham RM, et al. Low voltage on the electrocardiogram is a marker of disease severity and a risk factor for adverse outcomes in patients with heart failure due to systolic dysfunction. Am Heart J. 2006;152:355-61.
7. Yang J, Ferreira T, Morris AP, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet. 2012;44:369-75, S1-3.
8. Abecasis GR, Altshuler D, Auton A, et al. A map of human genome variation from population-scale sequencing. Nature. 2010;467:1061-73.
9. Stergachis AB, Neph SJ, Reynolds AP, et al. Epigenetic memory of developmental fate and time encoded in human regulatory DNA landscapes. Cell. 2013;154:888-903.
10. Maurano MT, Humbert R, Rynes E, et al. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012;337:1190-5.

11. Thurman RE, Rynes E, Humbert R, et al. The accessible chromatin landscape of the human genome. Nature. 2012;489:75-82.

12. Bernstein BE, Stamatoyannopoulos JA, Costello JF, et al. The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotechnol. 2010;28:1045-8.

13. Kundaje A, Meuleman W, Ernst J, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518:317-30.

14. He A, Kong SW, Ma Q, Pu WT. Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. Proc Natl Acad Sci U S A. 2011;108:5632-7.

15. May D, Blow MJ, Kaplan T, et al. Large-scale discovery of enhancers from human heart tissue. Nat Genet. 2012;44:89-93.

16. van den Boogaard M, Smemo S, Burnicka-Turek O, et al. A common genetic variant within SCN10A modulates cardiac SCN5A expression. J Clin Invest. 2014;124:1844-52.

17. van den Boogaard M, Wong LY, Tessadori F, et al. Genetic variation in T-box binding element functionally affects SCN5A/SCN10A enhancer. J Clin Invest. 2012;122:2519-30.

18. Gieger C, Radhakrishnan A, Cvejic A, et al. New gene functions in megakaryopoiesis and platelet formation. Nature. 2011;480:201-8.

19. van der Harst P, Zhang W, Mateo Leach I, et al. Seventy-five genetic loci influencing the human red blood cell. Nature. 2012;492:369-75.

20. Raychaudhuri S, Plenge RM, Rossin EJ, et al. Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. PLoS Genet. 2009;5:e1000534.

21. Hoischen A, van Bon BW, Gilissen C, et al. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. Nat Genet. 2010;42:483-5.

22. Linden H, Williams R, King J, Blair E, Kini U. Ulnar Mammary syndrome and TBX3: expanding the phenotype. Am J Med Genet A. 2009;149A:2809-12.

23. Cannon MV, Sillje HHW, J.W.A. S, et al. Cardiac LXRα overexpression protects against pathological hypertrophy and dysfunction by enhancing glucose uptake and utilization. EMBO Mol Med. 2015;7:1229-43.

24. Kuipers I, Li J, Vreeswijk-Baudoin I, et al. Activation of liver X receptors with T0901317 attenuates cardiac hypertrophy in vivo. Eur J Heart Fail. 2010;12:1042-50.

25. Pers TH, Karjalainen JM, Chan Y, et al. Biological interpretation of genome-wide association studies using predicted gene functions. Nat Commun. 2015;6:5890.

26. Sundstrom J, Lind L, Arnlov J, Zethelius B, Andren B, Lithell HO. Echocardiographic and electrocardiographic diagnoses of left ventricular hypertrophy predict mortality independently of each other in a population of elderly men. Circulation. 2001;103:2346-51.

Acknowledgements

A detailed list of acknowledgements is provided in the **Supplementary Information**.

Author contributions

See **Supplementary Information**

Figure legends

Fig. 1 Genome-wide associations and candidate genes

Overlay Manhattan plot showing the results for the genome-wide associations with QRS traits amongst Europeans. SNPs reaching genome-wide significance ($P<1x10^{-8}$) are colored red (novel loci) or blue (previously reported loci). Candidate genes have been identified by one or multiple strategies; [n] nearest; [c] coding non-synonymous variant; [g] GRAIL tool; [e] eQTL; [d] DEPICT tool. The presence of associated eQTL, coding SNPs, DNAse hypersensitivity sites, chromatin states, TF binding sites are indicated for lead SNPs (light blue) or those in high ($r^2>0.8$) LD (dark blue).

Fig. 2 Functional annotations

(**a**) The 52 sentinel SNPs are significantly enriched in DHSs of the human fetal heart compared to the matched random distribution of HapMap SNPs. (**b**) The impact of physical distance between SNPs that meet genome wide significance ($P<1\times10^{-8}$) on enrichment of fetal heart relative to all other tissues at DHs. The enrichment is strongest at the SNP's location and decreases after 100bp from the SNP sites. (**c**) SNPs associated with QRS traits are enriched for the activating histone modifications H3K27ac, H3K4me3, H3K4me1 and H3K36me3 in human left ventricle, which increased at more stringent GWAS *P*-value thresholds. The repressive mark H3K27me3 is not enriched while H3K9me3 is significantly reduced, suggesting that QRS-trait loci are predominantly expressed in the left ventricle. (**d**) To capture the greater complexity we performed an integrative analysis in an 18-state 'expanded' ChromHMM model representative of different functional regions of the genome. The left panel shows the enrichment of the 52 loci for the 18-state model using the six core histone marks. The right panel shows the total

number of the 52 loci overlapped by each feature. (**e**) SNPs ($P<1\times10^{-8}$) were also significantly enriched for various factors in the human heart, mouse heart and the HL-1 cell-line.

Fig. 3 Functional follow-up of rs6781009 in the SCN5A locus

(**a**) *In vivo* activity of 4 exemplar human cardiac enhancers in embryonic transgenic mice stained for *LacZ* enhancer reporter activity (dark blue). Additional examples of previously described enhancers near lead SNPs are provided in **Fig. S13** (**b**) Position of the regulatory element containing rs6781009 on the SCN5A-SCN10A locus. GWAS signals are plotted on -log(*P*) scale in dark blue. The regulatory element is bound by *TBX3*, *TBX5*, and P300 (lower black traces) in mouse, and the contact profile of the *SCN5A* promoter obtained by 4C-seq human cardiac ventricular tissue revealed an interaction between this regulatory element and the *SCN5A* promoter (upper black trace and contact profile). Normalized contact intensities (gray dots) and their running median trends (black line) are depicted for the *SCN5A* promoter viewpoint. Medians are computed for 4 kb windows and the gray band displays the 20-80% percentiles for these windows. Below the profile statistical enrichment across differently scaled window sizes (from 2 kb (top row) to 50 kb (bottom)) is depicted of the observed number of sequenced ligation products over the expected total coverage of captured products, with the latter being estimated based on a probabilistic background model. Local changes in color codes indicate regions statistically enriched for captured sequences. The lowest box shows the linkage disequilibrium pattern for the HapMap CEU population. (**c**) Luciferase assay performed in H10 cells showing a high constitutive activity for the enhancer core element (0.6kb) containing the major allele for rs6781009, which is reduced for the minor allele in both a large enhancer construct (1.5kb), as well as in the core enhancer element (0.6kb) **P*<0.01 (**d**) Dorsal views of hearts containing the human regulatory element with the major vs minor

allele for rs6781009 in a *LacZ* reporter vector, showing specific expression of the enhancer in the interventricular septum (ivs) for the major allele, which is absent for the minor allele **P*<0,05. ra, right atrium; la, left atrium; rv, right ventricle; lv, left ventricle.

Fig. 4 Heart-specific RNAi knockdown in *Drosophila*

Cardiac defects upon heart-specific RNAi knockdown in *Drosophila*. (**a**) Wild-type dorsal heart tube stained with the F-actin stain phalloidin. Magnified region (right) is highlighted. Arrowheads point to ostia (inflow tracks), arrow shows the circumferential orientation of myofibrils. (**b**) Cka/Striatin RNAi induces myofibrillar disarrangement. Myofibrils are oriented in a disorganized, mainly anterior-posterior orientation with gaps in between (arrow). (**c**) Knockdown of NACα/NACA causes severe cardiac tissue disintegration. Adult cardiomyocyte tissue may be completely absent (asterisk), while some heart-associated longitudinal muscles are still present (arrowheads). At larval stages the heart is much less affected, suggesting maturation or remodeling defect. (**d**) Knockdown of EcR/NR1H blocks cardiac remodeling and causes myofibrillar disarray (arrow). Ventral longitudinal muscles are also abnormal (arrowhead).

Fig. 5 Functional connections of gene expression networks

DEPICT analysis. (**a**) Plots showing the enrichment of loci associated with QRS traits in specific physiological systems. (**b**) Graphical display of DEPICT gene set enrichment analysis. Gene meta-sets are represented by nodes colored according to statistical significance, and similarities between them are indicated by edges scaled according to their correlation (only correlations with r >0.3 are shown).

Affiliations

1 University of Groningen, University Medical Center Groningen, Department of Cardiology, 9700 RB Groningen, The Netherlands

2 University of Groningen, University Medical Center Groningen, Department of Genetics, 9700 RB Groningen, The Netherlands

3 Durrer Center for Cardiovascular Research, Netherlands Heart Institute, 3501 DG Utrecht, The Netherlands

4 Department of Medical Genetics, University Medical Center Utrecht, 3584 CG Utrecht, The Netherlands

5 Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, 3508 GA Utrecht, The Netherlands

6 Development, Aging and Regeneration, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA 92037, United States of America

7 Department of Genome Sciences, University of Washington, Seattle, WA 98195, United States of America

8 Department of Medicine, Division of Oncology, University of Washington, Seattle, WA 98195, United States of America

9 Department of Pathology, New York University Langone Medical Center, New York, NY 10016, United States of America

10 Institute for Systems Genetics, New York University Langone Medical Center, New York, NY 10016, United States of America

11 Department of Biology, Massachusetts Institute of Technology, MA 02139 United States of America

12 Department of Epidemiology, Erasmus Medical Center, 3000 CA Rotterdam, The Netherlands

13 Department of Internal Medicine, Erasmus Medical Center, 3000 CA Rotterdam, The Netherlands

14 Division of Cardiology, Cardiovascular Health Research Unit, University of Washington, Seattle, WA 98195, United States of America

15 MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Crewe Road, Edinburgh EH4 2XU, Scotland

16 Institute of Genetics and Biophysics A. Buzzati-Traverso, CNR, 80131 Naples, Italy

17 Laboratory of Genetics, National Institute on Aging, Baltimore, MD 21224, United States of America

18 Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland

19 Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere 33014, Finland

20 Centre for Global Health Research, The Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland

21 Department of Public Health, Faculty of Medicine, University of Split, Split, Croatia

22 Translational Gerontology Branch, National Institute on Aging, Baltimore, MD 21250, United States of America

23 Center for Complex Disease Genomics. McKusick - Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, United States of America

24 Institute for Maternal and Child Health, IRCCS "Burlo Garofolo", 34137 Trieste, Italy

25 Department of Cardiology, Leiden University Medical Center, 2300 RC Leiden, The Netherlands

26 Department of Gerontology and Geriatrics, Leiden University Medical Center, 2300 RC Leiden, The Netherlands

27 Department of Medicine I, University Hospital Munich, Campus Grosshadern, Ludwig-Maximilians-University, 81377 Munich, Germany

28 Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, 81377 Munich, Germany.,

29 Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany

30 DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany,

31 Icelandic Heart Association, IS-201 Kópavogur, Iceland

32 University of Iceland, 101 Reykjavik, Iceland

33 Department of Internal Medicine B, University Medicine Greifswald, 17475 Greifswald, Germany

34 DZHK (German Centre for Cardiovascular Research), partner site Greifswald, Germany

35 Department of Biostatistics, University of Washington, Seattle, WA 98195, United States of America

36 Section of Cardiovascular Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA 02118, United States of America

37 Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), 39100 Bolzano, Italy - Affiliated Institute of the University of Lübeck, Lübeck, Germany,

38 Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, United Kingdom

39 Ealing Hospital NHS Trust, Middlesex UB1 3HW, United Kingdom

40 University of Groningen, University Medical Center Groningen, Department of Epidemiology, 9700 RB Groningen, The Netherlands

41 Genetic Epidemiology Unit, Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, 3000 CA Rotterdam, The Netherlands

42 Doctoral Program in Biomedical Sciences, Universidad del Rosario, Bogotá, Colombia

43 Department of Genetics (GENIUROS), Escuela de Medicina y Ciencias de la salud, Universidad del Rosario, Bogotá, Colombia

44 Institute of Cardiovascular and Medical Sciences, University of Glasgow, G12 8TA Glasgow, United Kingdom

45 Estonian Genome Center, University of Tartu, 51010 Tartu, Estonia

46 Division of Endocrinology and Center for Basic and Translational Obesity Research, Children's Hospital Boston, Boston, MA 02115, United States of America

47 Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI, United States of America

48 Department of Experimental Cardiology, University of Amsterdam, Amsterdam Medical Centre, 1105 AZ Amsterdam, The Netherlands

49 Maastricht Centre for Systems Biology, Maastricht University, 6229 ER Maastricht, The Netherlands

50 Landspitali University Hospital, 101 Reykjavik, Iceland

51 Department of Anatomy, Embryology and Physiology, University of Amsterdam, Academic Medical Center, 1105 AZ Amsterdam, The Netherlands

52 Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA 98101, United States of America

53 Department of Pharmacology and Therapeutics, University College Cork, Co. Cork, Ireland

54 Department of Medicine, Cardiovascular Division, University of Minnesota, Minneapolis, MN 55455, United States of America

55 Istituto di Ricerca Genetica e Biomedica, CNR, Monserrato, 09042, Cagliari, Italy

56 Department of Medicine, Division of Cardiology, Weill Cornell Medicine, New York, NY 10065, United States of America

57 Robertson Center for Biostatistics, University of Glasgow, G12 8QQ Glasgow, United Kingdom

58 Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany

59 Research Unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany

60 Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, United States of America

61 Cardiovascular and Metabolic Diseases, Max-Delbrück-Center for Molecular Medicine (MDC), 13125 Berlin, Germany

62 Department of computational biology, Max Planck Institute for molecular genetics, 14195 Berlin, Germany

63 Institute of Computational Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany

64 Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD 20892, United States of America

65 Medical and Population Genetics Program, Broad Institute, Cambridge, MA 02142, United States of America

66 Department of Genetics, Harvard Medical School, Boston, MA 02115, United States of America

67 DZHK (German Center for Cardiovascular Research), partner site Berlin, Berlin, Germany

68 Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institute of Health, Bethesda, MD 20824, United States of America

69 Cardiovascular Department and Postgraduate School of Cardiovascular Disease, University of Trieste, 34128 Trieste Italy

70 Department of Clinical Physiology, Tampere University Hospital, Tampere 33521, Finland

71 Department of Clinical Physiology, University of Tampere School of Medicine, Tampere 33014, Finland

72 Computer Science and Artificial Intelligence Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, United States of America

73 Broad Institute, Cambridge, MA 02142, United States of America

74 National Heart and Lung Institute, Imperial College London, London W12 0NN, United Kingdom

75 Department of Medical Informatics, Erasmus Medical Center, 3000 CA Rotterdam, The Netherlands

76 Laboratory of Cardiovascular Science, National Institute on Aging, Baltimore, MD 21224, United States of America

77 Department of Surgery, Massachusetts General Hospital, Boston, MA 2114, United States of America

78 Harvard Medical School, Harvard University, Boston, MA 2115, United States of America

79 Center for Population Studies, National Heart, Lung, and Blood Institute, National Institute of Health, Bethesda, MD 20824, United States of America

80 Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, DK-2200 Copenhagen N, Denmark

81 Department of Biomedical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, DK-2200 Copenhagen N, Denmark

82 Electrocardiology Section, Institute of Cardiovascular and Medical sciences, University of Glasgow, G12 8QQ Glasgow, United Kingdom

83 Department of Family Medicine, Clalit Health Services, and The Hebrew University-Hadassah Medical School, 91120 Jerusalem, Israel

84 Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany

85 Institute of Human Genetics, Technische Universität München, 81675 Munich, Germany

86 Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States of America

87 Center for Neuroscience, Departments of Neurobiology, Physiology, and Behavior and Psychiatry and Behavioral Sciences, University of California, Davis, CA 95618, United States of America

88 Institute of Molecular Biotechnology of the Austrian Academy of Sciences, 1030 Vienna, Austria

89 DOE Joint Genome Institute, Walnut Creek, United States

90 Novo Nordisk Foundation Centre for Basic Metabolic Research, University of Copenhagen, DK-2200 Copenhagen N, Denmark

91 Department of Epidemiology Research, Statens Serum Institut, DK-2300 Copenhagen, Denmark

92 Institute for Biological and Medical Imaging, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany

93 Department of Bioinformatics and Systems Biology IBIS, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany

94 Department of Neurology, General Central Hospital, Bolzano, Italy

95 Department of Neurology, University of Lübeck, 23562 Lübeck, Germany

96 Cardiogenetics Lab, Human Genetics Research Centre,, St George's University of London, SW17 0RE London, United Kingdom

97 Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku 20521, Finland

98 Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20520, Finland

99 Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, United States of America,

100 Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02142, United States of America,

101 Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA 02114, United States of America

102 Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute and Department of Pediatrics and Medicine, Harbor-UCLA Medical Center, Torrance, CA 90502, United States of America

103 National Heart Research Institute Singapore, National Heart Centre Singapore, 168752, Singapore, Singapore

104 Institute for Community Medicine, University Medicine Greifswald, 17475 Greifswald, Germany

105 Bioinformatics and Integrative Genomics, Harvard University, Cambridge, MA 02138, United States of America,

106 Epidemiological Cardiology Research Center, Wake Forest School of Medicine, Winston Salem, NC, United States of America

107 Department of Twin Research and Genetic Epidemiology, King's College London, SE1 7EH London, United Kingdom

108 Department of Vascular Medicine, University Medical Center Utrecht, 3508 GA Utrecht, The Netherlands

109 Division of Medicine, Turku University Hospital, Turku 20521, Finland

110 Department of Medicine, University of Turku, Turku 20014, Finland

111 Queensland Brain Institute, University of Queensland, St Lucia, QLD 4072, Australia

112 University of Queensland Diamantina Institute, Translational Research Institute, Woolloongabba, QLD 4102, Australia

113 Department of Functional Genomics, Interfaculty Institute of Genetics and Functional Genomics, University Medicine Greifswald, 17475 Greifswald, Germany

114 Divisions of Genetics and Rheumatology, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02446, United States of America

115 Partners Center for Personalized Genetic Medicine, Boston, MA 02446, United States of America

116 University of Groningen, University Medical Center Groningen, Department of Endocrinology, 9700 RB Groningen, The Netherlands

117 Clinical Pharmacology and Barts and The London Genome Centre, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, United Kingdom.

118 National Institute for Health Research Biomedical Research Unit, Barts and the London School of Medicine, Queen Mary University of London, London EC1M 6BQ, United Kindom

119 Institute of Cardiovascular Science, faculty of Population Health Sciences, University College London, WC1E 6BT, London, United Kingdom

120 Departments of Medicine, Epidemiology, and Health Services, Cardiovascular Health Research Unit, University of Washington, Seattle, WA 98195, United States of America

121 Group Health Research Institute, Group Health Cooperative, Seattle, WA 98101, United States of America

122 IRCCS Neuromed, 86077 Pozzilli, Isernia, Italy

123 Geriatric Unit, Azienda Sanitaria Firenze, Florence, Italy

124 Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA 30322, United States of America

125 University of Trieste, 34128 Trieste, Italy

126 Sidra Medical and Research Center, Doha, Qatar

127 Netherlands Heart Institute, 3511 EP Utrecht, The Netherlands

128 Department of Epidemiology, University of Washington, Seattle, WA 98195, United States of America

129 Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA 02129, United States of America

130 Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114, United States of America

131 School of Natural Sciences, University of California, Merced, CA 95343, United States of America

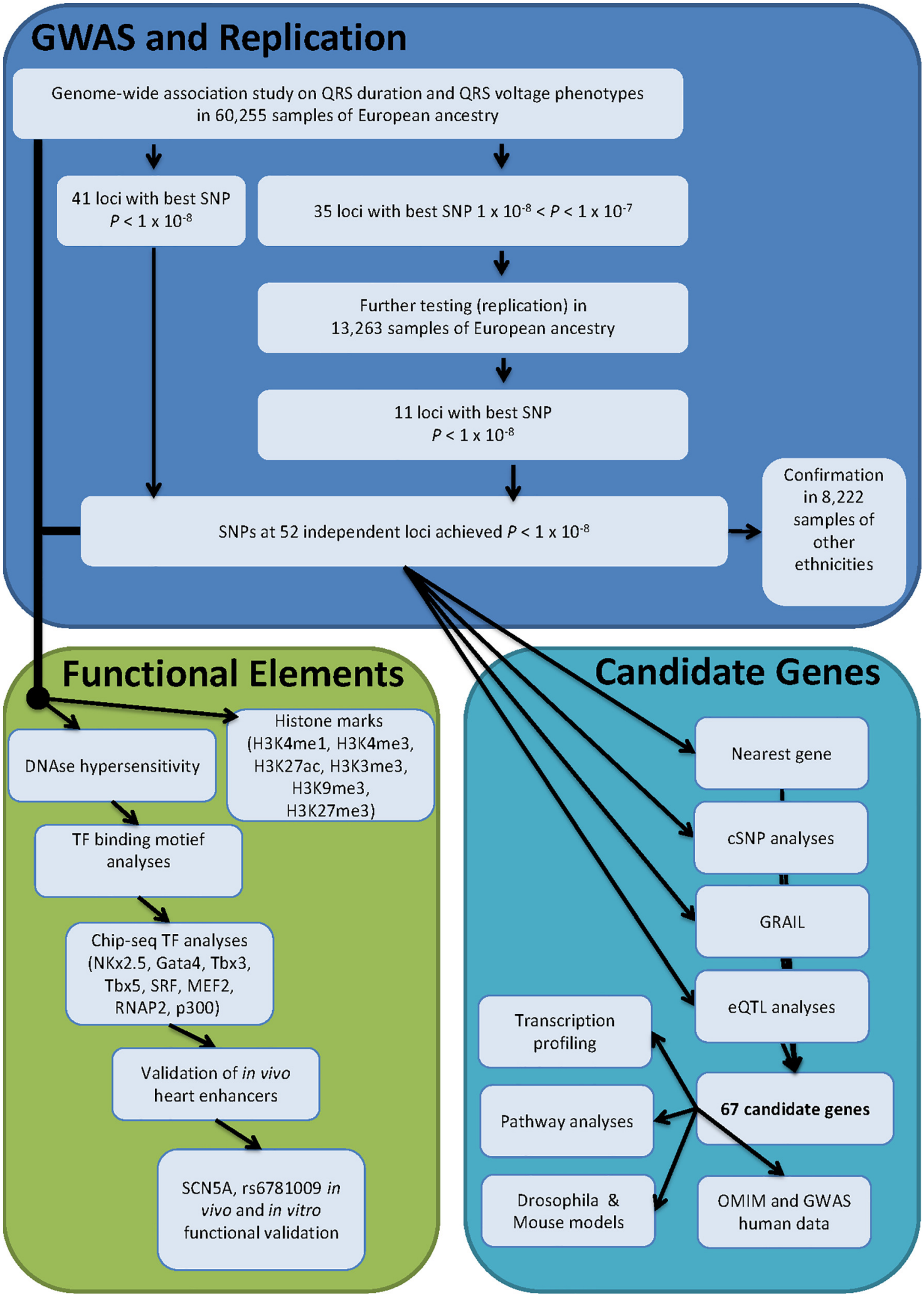
132 CARIM School for Cardiovascular Diseases, Maastricht Centre for Systems Biology (MaCSBio), Dept. of Biochemistry, Maastricht University, 6229 ER Maastricht, the Netherlands

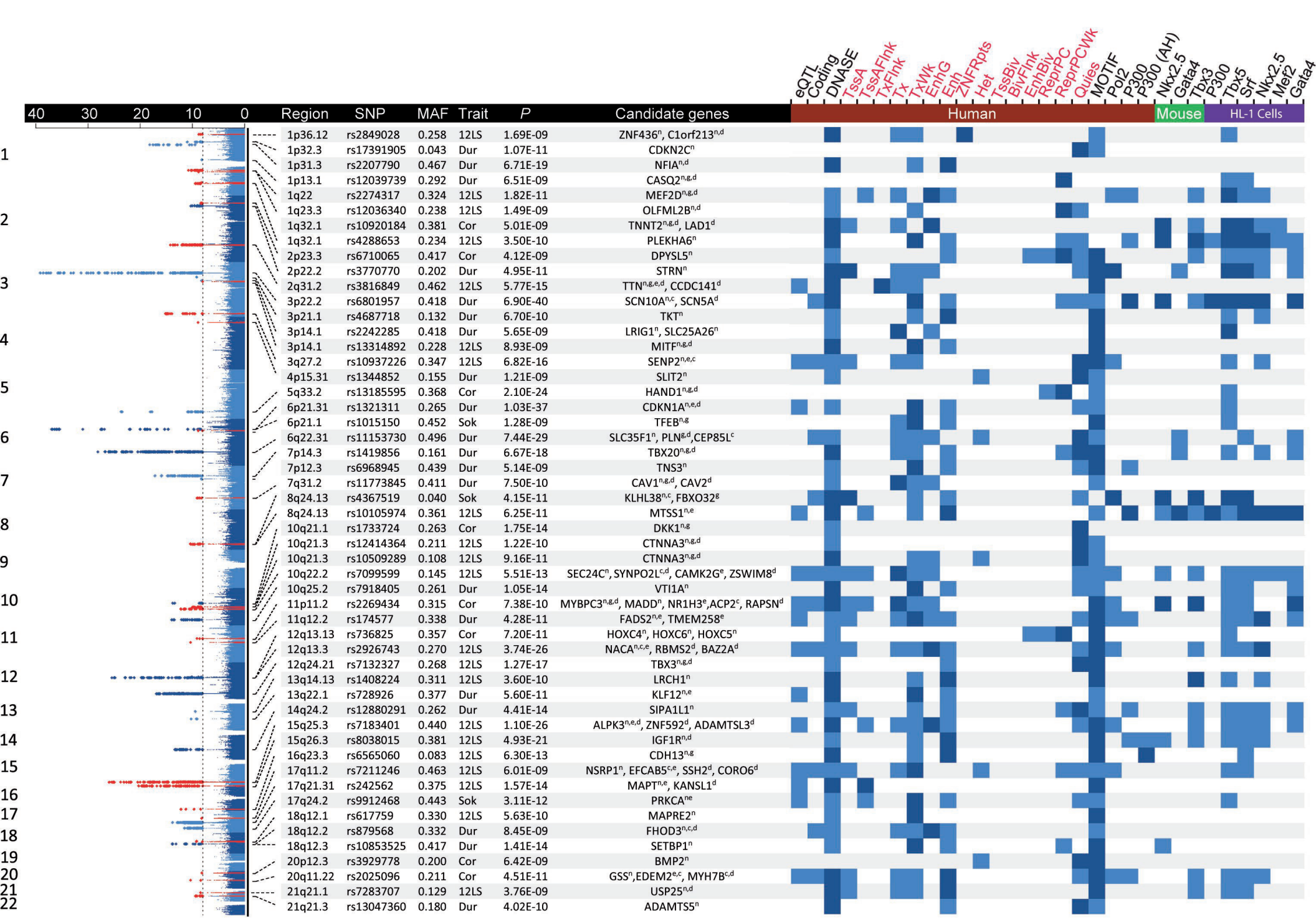
133 Department of Cardiovascular Sciences, University of Leicester, BHF Cardiovascular Research Centre, Glenfield Hospital, Leicester LE3 9QP , United Kingdom

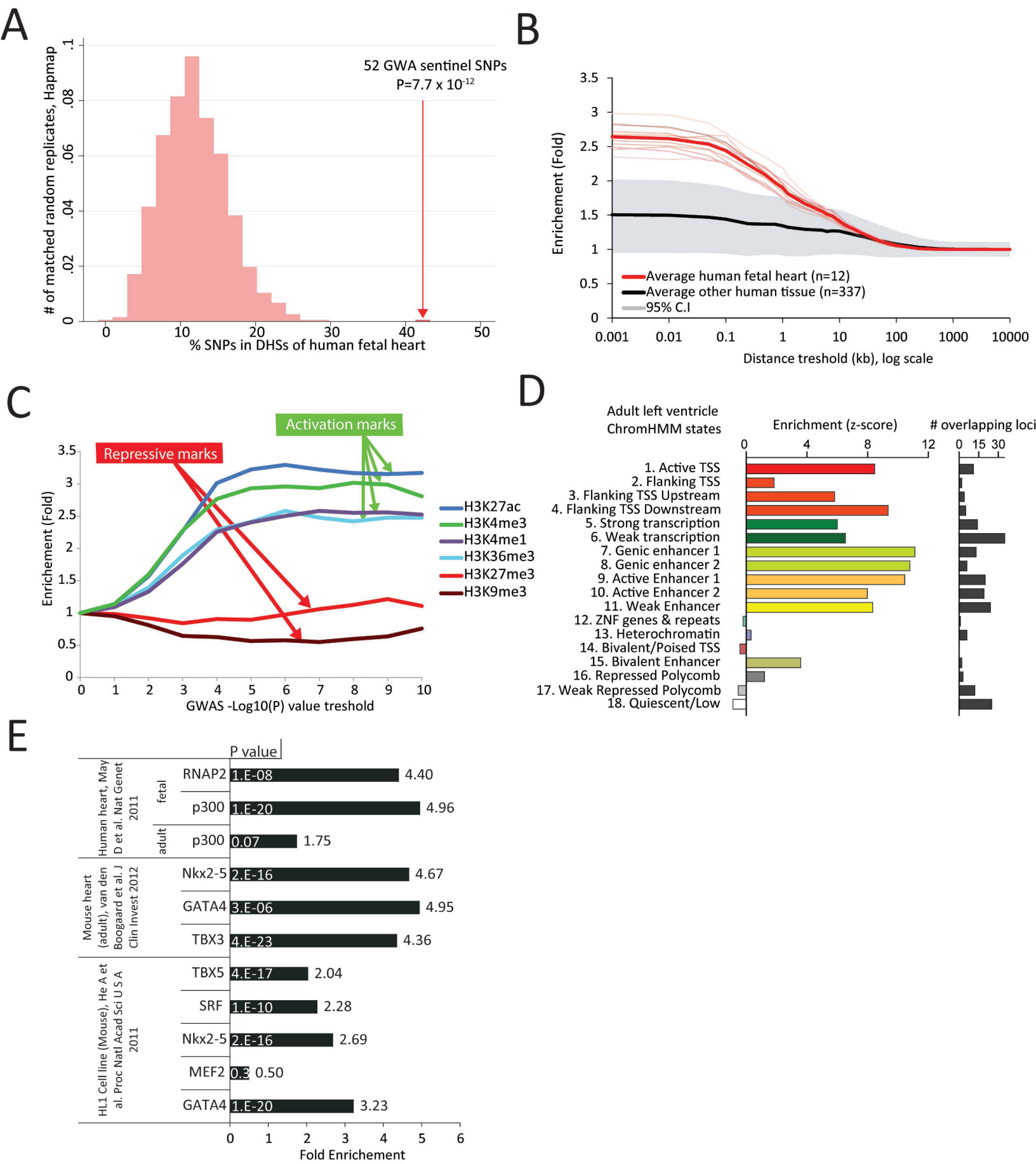
134 National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester LE3 9QP, United Kingdom

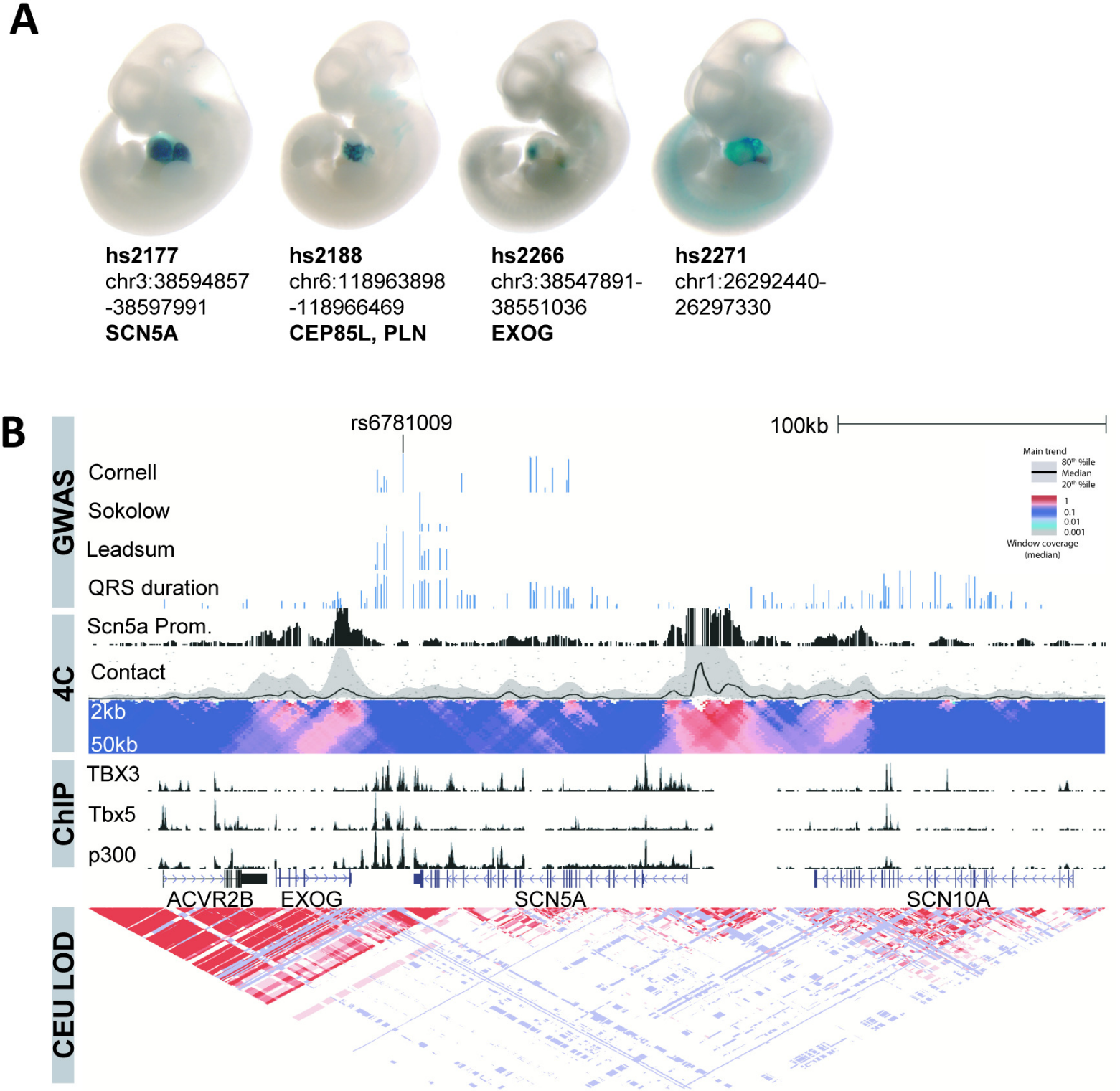
135 Department of Epidemiology, University Medical Center Utrecht, 3584 CG Utrecht, The Netherlands

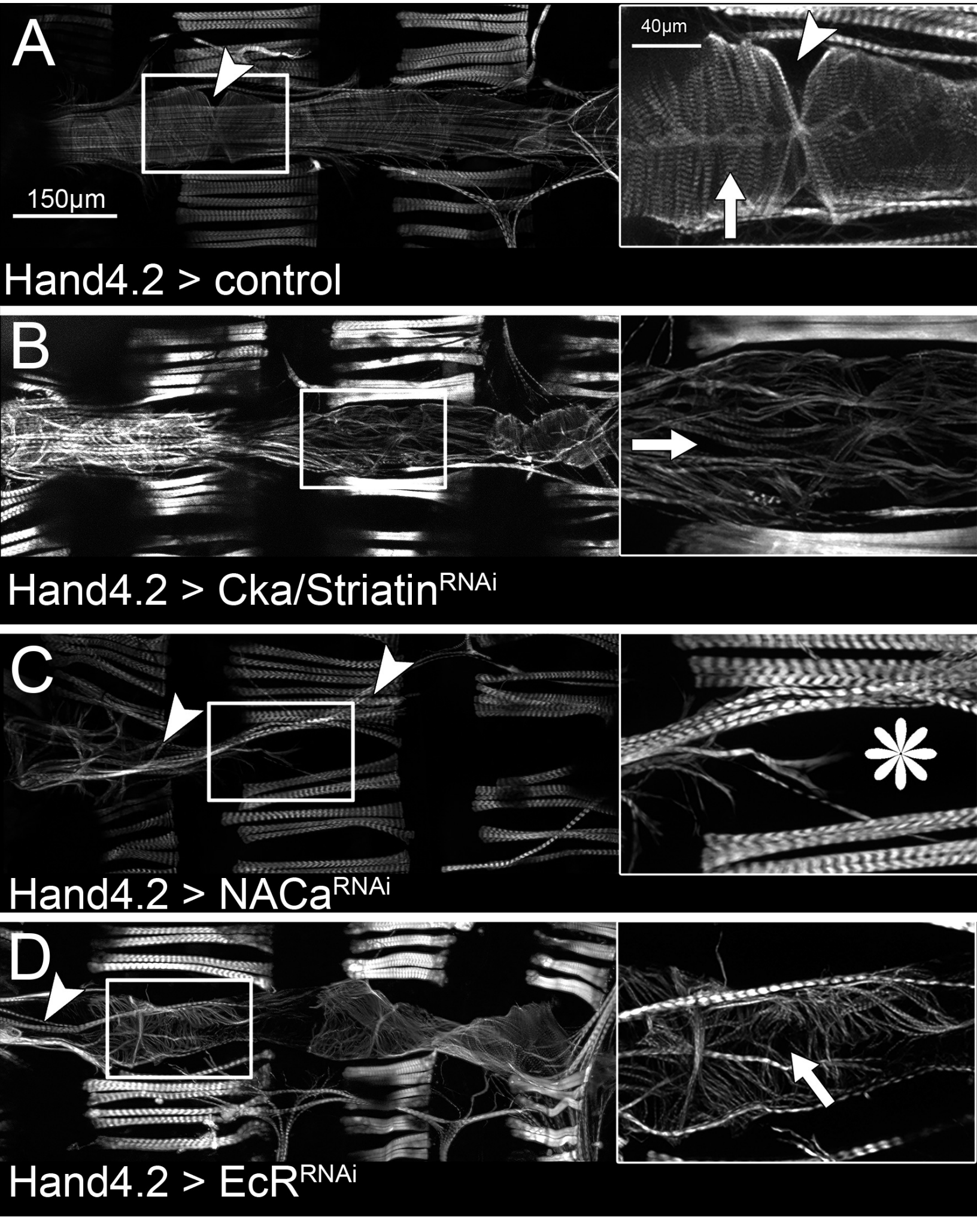
† These authors contributed equally

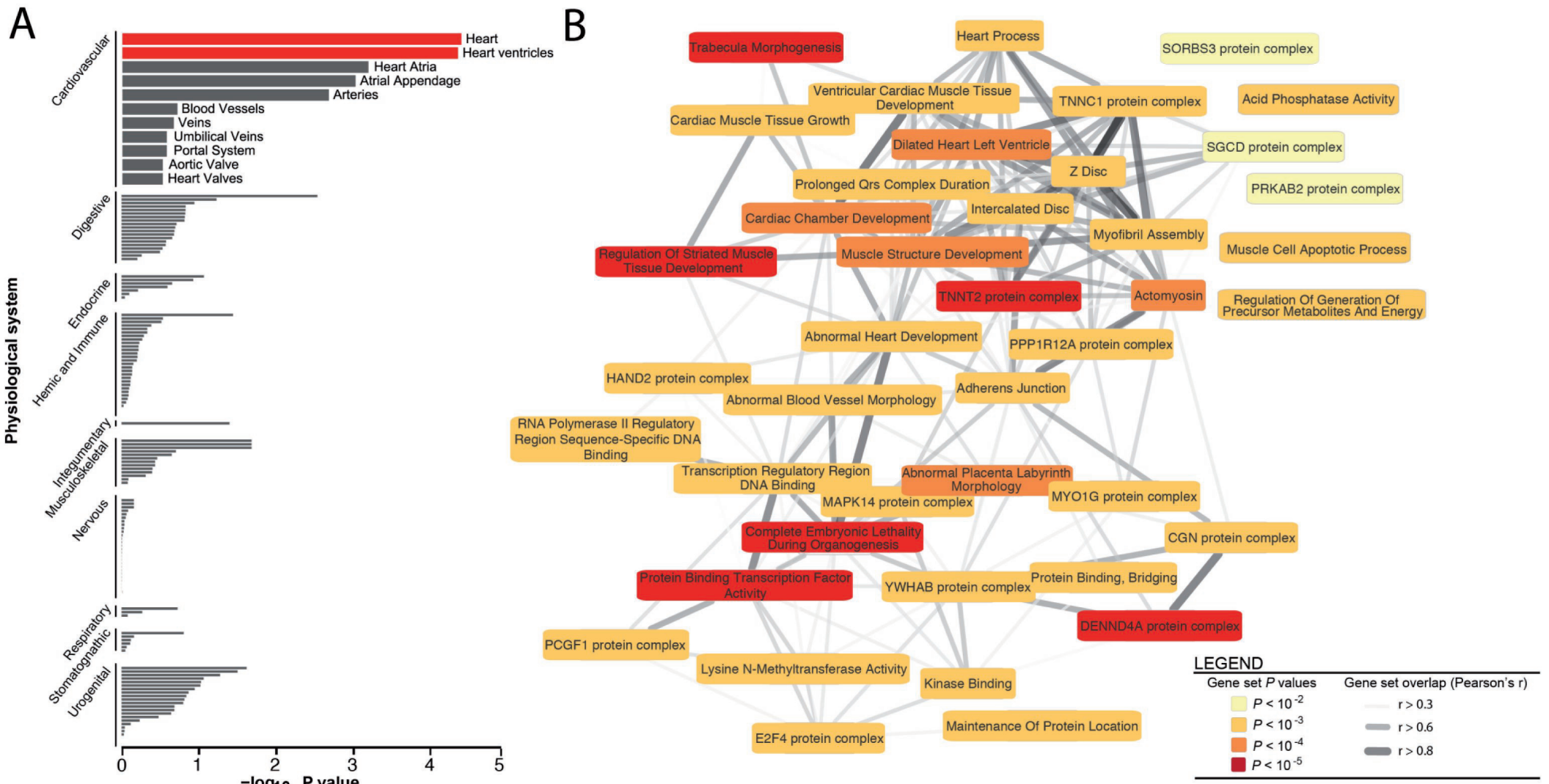

GWAS and Replication
Genome-wide association study on QRS duration and QRS voltage phenotypes in 60,255 samples of European ancestry
41 loci with best SNP $P < 1 \times 10^{-8}$
35 loci with best SNP $1 \times 10^{-8} < P < 1 \times 10^{-7}$
Further testing (replication) in 13,263 samples of European ancestry
11 loci with best SNP $P < 1 \times 10^{-8}$
SNPs at 52 independent loci achieved $P < 1 \times 10^{-8}$
Confirmation in 8,222 samples of other ethnicities
Functional Elements
DNAse hypersensitivity
Histone marks (H3K4me1, H3K4me3, H3K27ac, H3K3me3, H3K9me3, H3K27me3)
TF binding motief analyses
Chip-seq TF analyses (NKx2.5, Gata4, Tbx3, Tbx5, SRF, MEF2, RNAP2, p300)
Validation of *in vivo* heart enhancers
SCN5A, rs6781009 *in vivo* and *in vitro* functional validation
Candidate Genes
Nearest gene
cSNP analyses
GRAIL
eQTL analyses
Transcription profiling
Pathway analyses
Drosophila & Mouse models
67 candidate genes
OMIM and GWAS human data


eQTL
Coding
DNASE
TssA
TssAFlnk
TxFlnk
Tx
TxWk
EnhG
Enh
ZNFRpts
Het
TssBiv
BivFlnk
EnhBiv
ReprPC
ReprPCWk
Quies
MOTIF
Pol2
P300
P300 (AH)
Nkx2.5
Gata4
Tbx3
P300
Tbx5
Srf
Nkx2.5
Mef2
Gata4
Human
Mouse
HL-1 Cells
40
30
20
10
0
Region
SNP
MAF
Trait
P
Candidate genes
1p36.12 rs2849028 0.258 12LS 1.69E-09 ZNF436[n], C1orf213[n,d]
1p32.3 rs17391905 0.043 Dur 1.07E-11 CDKN2C[n]
1p31.3 rs2207790 0.467 Dur 6.71E-19 NFIA[n,d]
1p13.1 rs12039739 0.292 Dur 6.51E-09 CASQ2[n,g,d]
1q22 rs2274317 0.324 12LS 1.82E-11 MEF2D[n,g,d]
1q23.3 rs12036340 0.238 12LS 1.49E-09 OLFML2B[n,d]
1q32.1 rs10920184 0.381 Cor 5.01E-09 TNNT2[n,g,d], LAD1[d]
1q32.1 rs4288653 0.234 12LS 3.50E-10 PLEKHA6[n]
2p23.3 rs6710065 0.417 Cor 4.12E-09 DPYSL5[n]
2p22.2 rs3770770 0.202 Dur 4.95E-11 STRN[n]
2q31.2 rs3816849 0.462 12LS 5.77E-15 TTN[n,g,e,d], CCDC141[d]
3p22.2 rs6801957 0.418 Dur 6.90E-40 SCN10A[n,c], SCN5A[d]
3p21.1 rs4687718 0.132 Dur 6.70E-10 TKT[n]
3p14.1 rs2242285 0.418 Dur 5.65E-09 LRIG1[n], SLC25A26[n]
3p14.1 rs13314892 0.228 12LS 8.93E-09 MITF[n,g,d]
3q27.2 rs10937226 0.347 12LS 6.82E-16 SENP2[n,e,c]
4p15.31 rs1344852 0.155 Dur 1.21E-09 SLIT2[n]
5q33.2 rs13185595 0.368 Cor 2.10E-24 HAND1[n,g,d]
6p21.31 rs1321311 0.265 Dur 1.03E-37 CDKN1A[n,e,d]
6p21.1 rs1015150 0.452 Sok 1.28E-09 TFEB[n,g]
6q22.31 rs11153730 0.496 Dur 7.44E-29 SLC35F1[n], PLN[g,d],CEP85L[c]
7p14.3 rs1419856 0.161 Dur 6.67E-18 TBX20[n,g,d]
7p12.3 rs6968945 0.439 Dur 5.14E-09 TNS3[n]
7q31.2 rs11773845 0.411 Dur 7.50E-10 CAV1[n,g,d], CAV2[d]
8q24.13 rs4367519 0.040 Sok 4.15E-11 KLHL38[n,c], FBXO32[g]
8q24.13 rs10105974 0.361 12LS 6.25E-11 MTSS1[n,e]
10q21.1 rs1733724 0.263 Cor 1.75E-14 DKK1[n,g]
10q21.3 rs12414364 0.211 12LS 1.22E-10 CTNNA3[n,g,d]
10q21.3 rs10509289 0.108 12LS 9.16E-11 CTNNA3[n,g,d]
10q22.2 rs7099599 0.145 12LS 5.51E-13 SEC24C[n], SYNPO2L[c,d], CAMK2G[e], ZSWIM8[d]
10q25.2 rs7918405 0.261 Dur 1.05E-14 VTI1A[n]
11p11.2 rs2269434 0.315 Cor 7.38E-10 MYBPC3[n,g,d], MADD[n], NR1H3[e],ACP2[c], RAPSN[d]
11q12.2 rs174577 0.338 Dur 4.28E-11 FADS2[n,e], TMEM258[e]
12q13.13 rs736825 0.357 Cor 7.20E-11 HOXC4[n], HOXC6[n], HOXC5[n]
12q13.3 rs2926743 0.270 12LS 3.74E-26 NACA[n,c,e], RBMS2[d], BAZ2A[d]
12q24.21 rs7132327 0.268 12LS 1.27E-17 TBX3[n,g,d]
13q14.13 rs1408224 0.311 12LS 3.60E-10 LRCH1[n]
13q22.1 rs728926 0.377 Dur 5.60E-11 KLF12[n,e]
14q24.2 rs12880291 0.262 Dur 4.41E-14 SIPA1L1[n]
15q25.3 rs7183401 0.440 12LS 1.10E-26 ALPK3[n,e,d], ZNF592[d], ADAMTSL3[d]
15q26.3 rs8038015 0.381 12LS 4.93E-21 IGF1R[n,d]
16q23.3 rs6565060 0.083 12LS 6.30E-13 CDH13[n,g]
17q11.2 rs7211246 0.463 12LS 6.01E-09 NSRP1[n], EFCAB5[c,e], SSH2[d], CORO6[d]
17q21.31 rs242562 0.375 12LS 1.57E-14 MAPT[n,e], KANSL1[d]
17q24.2 rs9912468 0.443 Sok 3.11E-12 PRKCA[ne]
18q12.1 rs617759 0.330 12LS 5.63E-10 MAPRE2[n]
18q12.2 rs879568 0.332 Dur 8.45E-09 FHOD3[n,c,d]
18q12.3 rs10853525 0.417 Dur 1.41E-14 SETBP1[n]
20p12.3 rs3929778 0.200 Cor 6.42E-09 BMP2[n]
20q11.22 rs2025096 0.211 Cor 4.51E-11 GSS[n],EDEM2[e,c], MYH7B[c,d]
21q21.1 rs7283707 0.129 12LS 3.76E-09 USP25[n,d]
21q21.3 rs13047360 0.180 Dur 4.02E-10 ADAMTS5[n]
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22


A
52 GWA sentinel SNPs
P=7.7 x 10-12
# of matched random replicates, Hapmap
% SNPs in DHSs of human fetal heart
B
Enrichement (Fold)
Average human fetal heart (n=12)
Average other human tissue (n=337)
95% C.I
Distance treshold (kb), log scale
C
Activation marks
Repressive marks
Enrichement (Fold)
H3K27ac
H3K4me3
H3K4me1
H3K36me3
H3K27me3
H3K9me3
GWAS -Log10(P) value treshold
D
Adult left ventricle
ChromHMM states
Enrichment (z-score)
overlapping loci
1. Active TSS
2. Flanking TSS
3. Flanking TSS Upstream
4. Flanking TSS Downstream
5. Strong transcription
6. Weak transcription
7. Genic enhancer 1
8. Genic enhancer 2
9. Active Enhancer 1
10. Active Enhancer 2
11. Weak Enhancer
12. ZNF genes & repeats
13. Heterochromatin
14. Bivalent/Poised TSS
15. Bivalent Enhancer
16. Repressed Polycomb
17. Weak Repressed Polycomb
18. Quiescent/Low
E
P value
Human heart, May D et al. Nat Genet 2011
fetal
RNAP2 1.E-08 4.40
p300 1.E-20 4.96
adult
p300 0.07 1.75
Mouse heart (adult), van den Boogaard et al. J Clin Invest 2012
Nkx2-5 2.E-16 4.67
GATA4 3.E-06 4.95
TBX3 4.E-23 4.36
HL1 Cell line (Mouse), He A et al. Proc Natl Acad Sci U S A 2011
TBX5 4.E-17 2.04
SRF 1.E-10 2.28
Nkx2-5 2.E-16 2.69
MEF2 0.3 0.50
GATA4 1.E-20 3.23
Fold Enrichement


A
hs2177
chr3:38594857
-38597991
SCN5A
hs2188
chr6:118963898
-118966469
CEP85L, PLN
hs2266
chr3:38547891-
38551036
EXOG
hs2271
chr1:26292440-
26297330
B
rs6781009
100kb
GWAS
Cornell
Sokolow
Leadsum
QRS duration
Main trend
80th %ile
Median
20th %ile
1
0.1
0.01
0.001
Window coverage
(median)
4C
Scn5a Prom.
Contact
2kb
50kb
ChIP
TBX3
Tbx5
p300
ACVR2B
EXOG
SCN5A
SCN10A
CEU LOD
C
rs6781009
T>C
6
5
4
3
2
1
0
*
*
Major
Minor
Major
Minor
Enhancer
(1.5kb)
Enhancer core
(0.6kb)
D
Major
Minor
la
ra
lv
rv
ivs
5/6
la
ra
lv
rv
ivs
0/3*


A
150µm
40µm
Hand4.2 > control
B
Hand4.2 > Cka/Striatin^RNAi
C
Hand4.2 > NACa^RNAi
D
Hand4.2 > EcR^RNAi


A
Physiological system
Cardiovascular
Heart
Heart ventricles
Heart Atria
Atrial Appendage
Arteries
Blood Vessels
Veins
Umbilical Veins
Portal System
Aortic Valve
Heart Valves
Digestive
Endocrine
Hemic and Immune
Integumentary
Musculoskeletal
Nervous
Respiratory
Stomatognathic
Urogenital
0
1
2
3
4
5
−log10 P value
B
Trabecula Morphogenesis
Heart Process
SORBS3 protein complex
Ventricular Cardiac Muscle Tissue Development
TNNC1 protein complex
Acid Phosphatase Activity
Cardiac Muscle Tissue Growth
Dilated Heart Left Ventricle
SGCD protein complex
Z Disc
Prolonged Qrs Complex Duration
PRKAB2 protein complex
Cardiac Chamber Development
Intercalated Disc
Myofibril Assembly
Regulation Of Striated Muscle Tissue Development
Muscle Structure Development
Muscle Cell Apoptotic Process
TNNT2 protein complex
Actomyosin
Regulation Of Generation Of Precursor Metabolites And Energy
Abnormal Heart Development
PPP1R12A protein complex
HAND2 protein complex
Adherens Junction
Abnormal Blood Vessel Morphology
RNA Polymerase II Regulatory Region Sequence-Specific DNA Binding
Transcription Regulatory Region DNA Binding
Abnormal Placenta Labyrinth Morphology
MYO1G protein complex
MAPK14 protein complex
Complete Embryonic Lethality During Organogenesis
CGN protein complex
Protein Binding Transcription Factor Activity
YWHAB protein complex
Protein Binding, Bridging
DENND4A protein complex
PCGF1 protein complex
Lysine N-Methyltransferase Activity
Kinase Binding
Maintenance Of Protein Location
E2F4 protein complex
LEGEND
Gene set P values
P < 10-2
P < 10-3
P < 10-4
P < 10-5
Gene set overlap (Pearson's r)
r > 0.3
r > 0.6
r > 0.8

Supplementary Information

Fifty-two Genetic Loci Influencing Myocardial Mass

Pim van der Harst, Jessica van Setten, Niek Verweij, Georg Vogler, Lude Franke, Matthew T. Maurano, Xinchen Wang, Irene Mateo Leach, Mark Eijgelsheim, Nona Sotoodehnia, Caroline Hayward, Rossella Sorice, Osorio Meirelles, Leo-Pekka Lyytikäinen, Ozren Polašek, Toshiko Tanaka, Dan E. Arking, Sheila Ulivi, Stella Trompet, Martina Müller-Nurasyid, Albert V. Smith, Marcus Dörr, Kathleen F. Kerr, Jared W. Magnani, Fabiola Del Greco M., Weihua Zhang, Ilja M. Nolte, Claudia T. Silva, Sandosh Padmanabhan, Vinicius Tragante, Tõnu Esko, Gonçalo R. Abecasis, Michiel E. Adriaens, Karl Andersen, Phil Barnett, Joshua C. Bis, Rolf Bodmer, Brendan M. Buckley, Harry Campbell, Megan V. Cannon, Aravinda Chakravarti, Lin Y. Chen, Alessandro Delitala, Richard B. Devereux, Pieter A. Doevendans, Anna F. Dominiczak, Luigi Ferrucci, Ian Ford, Christian Gieger, Tamara B. Harris, Eric Haugen, Matthias Heinig, Dena G. Hernandez, Hans L. Hillege, Joel N. Hirschhorn, Albert Hofman, Norbert Hubner, Shih-Jen Hwang, Annamaria Iorio, Mika Kähönen, Manolis Kellis, Ivana Kolcic, Ishminder K. Kooner, Jaspal S. Kooner, Jan A. Kors, Edward G. Lakatta, Kasper Lage, Lenore J. Launer, Daniel Levy, Alicia Lundby, Peter W. Macfarlane, Dalit May, Thomas Meitinger, Andres Metspalu, Stefania Nappo, Silvia Naitza, Shane Neph, Alex S. Nord, Teresa Nutile, Peter M. Okin, Jesper V. Olsen, Ben A. Oostra, Josef M. Penninger, Len A. Pennacchio, Tune H. Pers, Siegfried Perz, Annette Peters, Yigal M. Pinto, Arne Pfeufer, Maria Grazia Pilia, Peter P. Pramstaller, Bram P. Prins, Olli T. Raitakari, Soumya Raychaudhuri, Ken M. Rice, Elizabeth J. Rossin, Jerome I. Rotter, Sebastian Schafer, David Schlessinger, Carsten O. Schmidt, Jobanpreet Sehmi, Herman H.W. Silljé, Gianfranco Sinagra, Moritz F. Sinner, Kamil Slowikowski, Elsayed Z. Soliman, Timothy D. Spector, Wilko Spiering, John A. Stamatoyannopoulos, Ronald P. Stolk, Konstantin Strauch, Sian-Tsung Tan, Kirill V. Tarasov, Bosco Trinh, Andre G. Uitterlinden, Malou van den Boogaard, Cornelia M. van Duijn, Wiek H. van Gilst, Jorma S. Viikari, Peter M. Visscher, Veronique Vitart, Uwe Völker, Melanie Waldenberger, Christian X. Weichenberger, Harm-Jan Westra, Cisca Wijmenga, Bruce H. Wolffenbuttel, Jian Yang, Connie R. Bezzina, Patricia B. Munroe, Harold Snieder, Alan F. Wright, Igor Rudan, Laurie A. Boyer, Folkert W. Asselbergs, Dirk J. van Veldhuisen, Bruno H. Stricker, Bruce M. Psaty, Marina Ciullo, Serena Sanna, Terho Lehtimäki, James F. Wilson, Stefania Bandinelli, Alvaro Alonso, Paolo Gasparini, J. Wouter Jukema, Stefan Kääb, Vilmundur Gudnason, Stephan B. Felix, Susan R. Heckbert, Rudolf A. de Boer, Christopher Newton-Cheh, Andrew A. Hicks, John C. Chambers, Yalda Jamshidi, Axel Visel, Vincent M. Christoffels, Aaron Isaacs, Nilesh J. Samani, Paul I.W. de Bakker

Supplementary Note

QRS traits

The amplitude and duration of the QRS complex reflects the conduction through the left ventricle and is well correlated with left ventricular mass as measured by echocardiography [1-3]. ECG measurements of the QRS complex are important in clinical and preclinical cardiovascular diseases such as cardiac hypertrophy, heart failure, and various cardiomyopathies, and can also predict cardiovascular mortality[4-7]. We studied 4 related and clinically applied QRS traits associated with left ventricular hypertrophy:

1. *Sokolow-Lyon*; trait (sum of S wave in V1 plus R wave in V5 or R wave in V6) originally reported by Maurice Sokolow and Thomas Lyon in 1949[8] and developed based on comparison of cases with abnormal electrocardiograms and in whom a cardiac disorder capable of producing increased strain on the left ventricle (such as hypertension, aortic valvular lesions, coarctation of the aorta, patent ductus arterious) was present to healthy controls. Ninety percent of cases had hypertension exceeding 155/95mmHg, with a mean blood pressure of 197/117 and a mean increase in roentgenologic transverse diameter of the heart of 15.8 percent.
2. *Cornell*; trait (sum of R in aVL and the S in V3) originally reported by Paul N Casale et al in 1985 based on comparison consecutive patients (including hypertension, valvular heart disease, coronary artery disease, cardiomyopathy, pericardial disease, mitral valve prolapse) undergoing M-mode echocardiography to determine left ventricular mass.[9]
3. *12-lead sum*; trait (sum of R to S in all 12 leads) originally reported by Siegel and Roberts as a marker for the severity of aortic stenosis in a post-mortem study[10], Molloy et al reported on the correlation with echocardiographic determined left ventricular mass.[11]
4. *QRS duration*; is frequently increased in left ventricular hypertrophy[12]. This is manifest by a diffuse increase in QRS duration or an increase in time from onset of QRS to the R-wave peak in V5 or V6. The increased QRS duration may be attributed to the increased thickness of the left ventricular wall and to intramural fibrosis[13], which distorts and prolongs the transmural impulse propagation.[14]

Multiple ECG markers of increased left ventricular mass were examined because of the relatively limited sensitivity of any one of these markers alone and because performance of

these markers can vary with gender, ethnicity and body characteristics. Therefore the AHA/ACCF/HRS also recommends not using a single criterion compared to others[14]. The three voltage criteria were multiplied with QRS duration to obtain voltage-duration products as an approximation of the area under the QRS complex which show stronger correlation with left ventricular mass as determined post-mortem[11] or by cardiac magnetic resonance or echocardiographic imaging[15,16].

Methods of genome-wide association and replication testing

In each study, we genotyped single nucleotide polymorphisms (SNPs) and imputed autosomal SNPs catalogued in HapMap Phase 2 CEU panel. Participants with atrial fibrillation, history of myocardial infarction or heart failure, QRS duration of >120ms, QRS-axis smaller -30 or larger than +90 degrees, and extreme measurements (more than ±4SD from mean on a per phenotype basis) were excluded. Optional exclusions, if available, were pacemaker or AICD implant, pacemaker activity on ECG, WPW, class I and class III blocking medication (ATC code prefix C01B). Characteristics of participants, genotyping arrays and imputation are summarized (**table S1 and S2**).

SNP associations with each phenotype were tested by linear regression using an additive genetic model. Associations were tested with age, gender, height and body mass index as covariates with principal components and other study specific factors to account of population substructure as described in **table S2**. Test statistics from each cohort were then corrected for their respective genomic control inflation factor to adjust for residual population sub-structure (**table S23**).

We then carried out meta-analysis of results from the 24 individual cohorts using inverse variance meta-analyses by two independent analysis groups using METAL[17] and MANTEL. Consistency was confirmed against z-scores weighted by square root of sample size meta-analyses method. In total, 60,255 individuals were included (maximum sample size 60,255 for QRS-duration, 54,993 for Sokolow-Lyon, 58,862 for Cornell, and 48,632 for 12-leadsum) and 2,766,983 autosomal SNPs. Genomic control was also applied to the final meta-analysis results.

We used PLINK[18] to cluster SNPs into genomic loci using a 2Mb window; clustering was done separately for each phenotype. There were 1,913 SNPs associated with one or more QRS traits at $P<10^{-8}$ distributed among 41 genomic loci (**Fig. 1 and Fig. S1**).

We found a further 35 loci with SNPs showing suggestive evidence of association to QRS phenotypes ($1\times10^{-8}<P<5\times10^{-7}$); at these loci we identified the SNP with the lowest *P*-value

against any trait and carried out additional replication testing using a combination of *in silico* data and direct genotyping amongst 13,263 individuals of European ancestry (**table S1 and S2**). At 11 of the 35 the lead SNP showed replication ($P<0.05$ for replication, and became significant ($P<1x10^{-8}$) in combined analysis with their discovery GWA data (**table S24**). Another 11 loci remained suggestive for association ($1x10^{-8}<P<5x10^{-8}$). Taken together the genome-wide and replication data identified 52 loci robustly associated with QRS traits at $P<1\times10^{-8}$ (**Fig. 1, table S4**).

Genome-wide significance and correction for multiple phenotypes

Our choice of the statistical threshold ($P<1\times10^{-8}$) was grounded on the guidelines derived from studies of the ENCODE regions which suggests that $P<5\times10^{-8}$ is the appropriate threshold for genome-wide significance in Europeans[19,20], but also was designed to provide us additional adjustment for the multiple phenotypes tested. This threshold is conservative, also considering our 4 QRS phenotypes are also inter-related: correlation coefficients between the phenotype pairs range from r=0.22 to 0.80 (**table S25**).

Conditional and joint analysis for the meta-analyses data

We performed the approximate conditional and joint analysis for the meta-analysis data after GC.[21] 6,654 unrelated individuals with individual-level genotype data selected from the ARIC cohort to approximate the LD structure between SNPs. The genotype data of the ARIC samples were imputed to HapMap2 by MACH. We used the best guess genotypes from imputation. In ARIC we excluded SNPs with imputation $r^2<0.3$ and HWE $P<1\times10^{-6}$. There were ~2.7M SNPs in the meta-analysis data. We removed SNPs that are not available in the ARIC data after QC, only considered SNPs with an estimated sample size of at least 10,000 and retained ~2.5M SNP for the conditional analysis. Assuming that the LD correlations between SNPs more than 10Mb away or on different chromosomes are zero, we performed a genome-wide stepwise selection procedure to select associated SNPs one-by-one at $P<1\times10^{-8}$.

Directional consistency of associations in non-Caucasians

To study the potential relevance of our findings in non-Europeans lead SNPs reported in **Fig. 1** were analysed in 3,603 African Americans (MESA and ARIC) and 4,619 Asian Indians (LOLIPOP). To assess whether the observed concordance of the directions of

effect in each ethnic group compared to the primary analyses was due to chance we performed a binomial draw with null expectation of probability of success 0.5.

Statistics of enrichment of coding SNPS, DHS and TF sites

We carried out permutation testing by randomly selecting 52 Hapmap2 SNPs matched to our sentinel SNPs and counted the number of non-synonymous coding variants in high LD (r_2>0.8). Matching was based on distance to gene, distance to transcription start site, minor allele frequency, and genomic annotation (intronic, intergenic or exonic). This was repeated 1,000 times to build up an expectation under the null hypothesis. Next we determined the number of non-synonymous SNPs in the 52 identified loci and compared it to the simulations of the null hypothesis.

DNA functional elements analyses

We investigated the enrichment of identified variants in regions of covalently modified histones as well as chromatin states representative of functional genomic regions predicted by the combination of histone modifications in human cardiac tissue mapped by the NIH Roadmap Epigenomics Program.[22,23] We collected data (bed files) from the Roadmap epigenomics project (release 8) for 6 chromatin mark assays (H3K4me1, H3k4me3, H3K27me3, H3K27ac, H3K36me3 and H3K9me3) that included cardiac coverage and a large number of other cell types[22]. Only samples with matching input DNA samples were included. If replicate experiments were available we aggregated the sequence reads. MACS (v1.4) software was used to identify significant peaks ($P \leq 1\times10^{-3}$) using a fixed DNA fragment size of 146[24]. Two samples could not be called with MACS due to inconsistencies in the original data. As a result, this data set included aligned sequence reads of 298 samples that were called with MACS (**table S26**). For genomic features we used annotations derived from combinations of histone modifications from the Roadmap Epigenomics Project using ChromHMM[25]. For adult left ventricle we had available all 6 chromatin marks and used the 18-state model. For fetal heart tissue, H3K27ac data was unavailable and therefore the 15-state model was used.

To study cardiac transcription factors we collected data on GATA4, MEF2, NKx2-5, SRF and TBX5 from[26] mouse hearts; p300 marks in human adult and foetal heart and RNAP2 from human foetal heart from[27]; and TBX3, GATA4 and nkx2-5 from the HL-1 cell line[28]. We

used the called peak data as has been described previously. Peaks from mice were lift-over to human using the UCSC Genome Browser liftOver tool after extending the regions by 1kb. Heart enhancers aren't well conserved between human & mouse, so by expanding the enhancer before liftOver, we stand a better chance of hitting the true human heart enhancer[29]. To study overlap of SNPs with peak data (DHS, TF or Histones marks) we converted genomic coordinates from hg18 to hg19 when appropriate. The DHS's encompassing 349 tissues from the ENCODE and Roadmap epigenomics project were processed using the hotspot algorithm[30].

To provide insight into tissue specific regulatory DNA mechanisms influencing QRS indices, (**Fig. 2**) we explored DNAse I hypersensitivity sites histone marks. We assigned the lowest P value from the QRS traits to each of the 2.3 M SNPs.. To gain insights into DHS, TFs and genomic features (ChromHMM) underlying the 52 QRS loci, we performed carried out permutation testing as described above.

Prediction of SNPs perturbing TF recognition sequences

Potential sites of TF binding were identified by scanning the human genome using position weight matrices from four major TF binding motif collections: TRANSFAC[31], JASPAR[32], UniPROBE[33], and a published SELEX dataset[34]. To avoid ascertainment bias for motifs better matching the reference allele of common polymorphisms, we created an alternate genome to complement the reference GRCh37/hg19 human genome. This alternate genome incorporates the non-reference allele at the location of each SNP identified in the CEU population of the 1000 Genomes Project[35]. Both the reference and alternate genomes were then scanned for motifs with a threshold $P \leq 10^{-4}$ using the program FIMO[36]. Then, for each SNP overlapping a motif, we computed the significance of the perturbation as the log-odds difference between the two alleles according to the position weight matrix. We considered SNPs in a DHS of any cell type affecting a motif with a log-odds difference greater than 4 as likely to significantly perturb accessibility. Motif models were mapped to gene names as previously described[37].

Cardiomyocyte differentiation and analysis

For gene-expression profiling and localization of histone modifications during development we used E14 Tg(Nkx2-5-EmGFP) mouse ES cells[38] cultured in feeder-free conditions using

standard techniques as reported before[39]. Directed differentiations and analyses were performed as described previously[40]. RNA-seq was performed on total RNA isolated from 5×10^6 cells using *TRIzol* Reagent and sequencing libraries prepared according to the Illumina RNA-Seq library kit. DESeq was used to normalize raw read counts and to analyse differential gene expression[41] while Useq 7.0 was used thereafter to generate gene-level read counts and estimate RPKM (reads per kilobase of exon per million reads mapped)[42]. Only genes with expression values >1 RPKM in at least one cell type were considered for analysis. Genome-wide localization of histone modifications (H3K4me3, H3K4me1, H3K27me3, and H3K27ac) for each stage was determined via chromatin immunoprecipitation, prepaired according to the Young lab protocol[43], followed by high throughput sequencing. Further details and the analysis pipeline can be found elsewhere[39].

Active and poised enhancers from four mouse cardiomyocyte differentiation time points were obtained from a previous study[39]. We extended mouse enhancer sizes by ±5kb to facilitate mapping from mouse to human. Extended mouse enhancers were mapped to human using the UCSC liftOver tool with parameters –minMatch=0.1 and –multiple. When mouse enhancers mapped to multiple sites on the human genome, the largest mapped region was chosen for subsequent analysis.

We used 1000 Genomes genotype data to identify all SNPs in strong LD ($r^2 > 0.8$) with the sentinel SNP. We then quantified the number of GWAS loci that contained a SNP that overlapped active or poised human enhancers at the four time points. To assess the significance and enrichment of these values, we generated background distribution for the number of loci that overlap each set of enhancers by sampling control SNPs spotted on an Affymetrix 660W genotyping array that have similar genetic properties (LD block size, minor allele frequency, distance to gene) as GWAS SNPs.

Experimental cardiac enhancer studies

In order to demonstrate the overlap between chromatin interactions and the distribution of all variants in this study, we used the negative natural log of the *P*-value to plot GWAS signals to the UCSC Genome Browser. Datapoints (WIG) for TBX3, TBX5 and P300 ChIP-seq on the Scn5a-Scn10a locus (mm9 chr9:119,307,167-119,613,764) were uploaded to the Galaxy Software Interface, transformed into interval files and converted into Hg18 (chr3:38,465,466-38,820,542) using the Lift-Over Tool.

4C templates and primers have been prepared as previously described. For this protocol, human ventricular tissue was crushed with a metal douncer in liquid nitrogen. Single cell suspensions were obtained by dissociation of tissue with IKA Ultra Turrax T5 FU, followed by dounce homogenization. Chromatin was crosslinked with 2% formaldehyde in PBS with 10% FCS for 10 min at room temperature, nuclei were isolated and cross-linked DNA was digested with a primary restriction enzyme recognizing a 4-bp restriction site (DpnII), followed by proximity ligation. Cross-links were removed and a secondary restriction enzyme digestion (Csp6I), followed again by proximity ligation. For all experiments, 200 ng of the resulting 4C template was used for the subsequent PCR reaction, of which 16 (total: 3.2 µg of 4C template) were pooled and purified for next-generation sequencing. The PCR products were purified using two columns per sample of the High Pure PCR Product Purification Kit (Roche cat. no. 11732676001).

4C templates were mixed and sequenced simultaneously in one Illumina HiSeq 2000 lane. The sequence tags generated by the procedure are prefixed by the 4C reading primer that includes the DpnII restriction site sequence (described in 4C primer design section). The 4C reading primer sequences are separated from multiplexed 4C-seq libraries and the suffixes are extracted for further processing. Mapping and filtering of the sequence reads was done as previously described[44].

Enhancer candidate regions with major and minor allele for rs6781009 were obtained by PCR from human control DNA and cloned into the Hsp68-LacZ reporter vector as previously described[45]. DNA was injected into the pronucleus of fertilized FVB strain eggs, which were subsequently transferred into the oviducts of CD-1 pseudo-pregnant foster females (Cyagen Biosciences). Approximately 200 injections per construct were performed. Embryos were harvested, stained with X-gal to detect LacZ activity. Chi-Square test was used to test statistical differences.

H10 cells, grown in 12-well plates in DMEM supplemented with 10% FCS (Gibco-BRL) and glutamine, were transfected using polyethylenimine 25 kDa (PEI, Brunschwick) at a 1:3 ratio (DNA:PEI). Reporter constructs were generated by ligating the enhancer regions with major and minor allele (hg19 chr3:38,584,695-38,586,171 (1.5kb) and chr3:38,585,064-38,585,625 (0.6kb)) to pGL2basic+SV40 promoter (control reporter). Standard transfections used 1 µg of reporter (or control reporter) vector co-transfected with 2 ng phRG-TK Renilla vector (Promega) as normalization control. Transfections were carried out at least three times and measured in triplo. Luciferase measurements were

performed using a Promega Turner Biosystems Modulus Multimode Reader luminometer. All data was statistically validated using a Student's T-Test.

Identification of candidate genes

We considered the nearest gene and any other gene located within 10kb of the sentinel SNP, to be a candidate for mediating the association with the electrocardiographic phenotype. We also used coding variant, eQTL and literature analyses to identify candidate genes.

Coding variation: We identified all non-synonymous SNPs (nsSNPs) that were in LD with one or more of the sentinel SNPs at $r^2 \geq 0.8$ in HapMap phase 2 CEU or 1000 Genomes Project dataset (May 2011 dataset). We considered the gene to be a candidate when the non-synonymous and the sentinel SNPs were in LD at $r^2 \geq 0.8$ and with no evidence for heterogeneity of effect on phenotype.

Expression analyses: To identify the possible genes influencing electrocardiographic traits at the 52 loci, we examined the association of the sentinel SNPs with eQTL data from different sources.

(1.) Human left ventricular tissue 1: 110 left ventricular samples were collected from non-diseased human donor hearts.[46] RNA-seq libraries were prepared from 1 µg of total RNA with the TruSeq RNA Sample Preparation Kit (Illumina). Polyadenylated transcripts were enriched, fragmented and cDNA fragments were subsequently sequenced on a HiSeq 2000 (Illumina) instrument using 2 x 100 bp PE chemistry. Gene expression levels were estimated by counting RNA-seq reads over protein coding genes using HT-seq[47]. Expression levels were normalized across samples by a quantile based scaling method[48]. Normalized expression levels were log transformed, and adjusted for covariates (age, sex, RNA quality, library preparation date, center) using linear regression in the eQTL analysis.

(2.) Human left ventricular tissue 2: eQTLs analysis were performed as part of The Myocardial Applied Genomics Network(MAGNet: www.med.upenn.edu/magnet). 313 Left ventricular free-wall tissue samples were harvested at time of cardiac surgery from subjects with heart failure undergoing transplantation and from unused donor hearts. DNA samples were genotyped using Affymetrix Genome Wide SNP Array 6.0 and imputed with 1000 Genomes phase 3. RNA was hybridized with Affymetrix

Genechip ST1.1 arrays. The genetic variants were tested for associations with 15,395 RMA expression levels using a joint effects model taking into account the patient group of each sample.

(3.) Human peripheral blood 1: derived from 1,469 unrelated individuals from the UK and the Netherlands.

(4.) Human peripheral blood lymphocytes RNA-seq from 2,116 individuals.

SNPs were tested for association with expression of nearby (1MB) genes (at $P<0.05$ after Bonferroni correction for number of SNP-expression associations tested). Where probable eQTLs were identified, we used the whole genome SNP data available in these datasets (imputed with HapMap phase 2 genotypes), to identify the SNP at the locus most closely associated with Transcript level (the Transcript SNP). We then tested whether the sentinel SNP and the Transcript SNP were coincident, defined as $r^2>0.8$ with no evidence of heterogeneity of effect on phenotype or transcript level ($P>0.05$).

GRAIL Analyses: literature analyses were performed using the Gene Relationships Among Implicated Loci (GRAIL) algorithm, a statistical tool based on text mining of PubMed abstracts to annotate candidate genes within a loci.[49] To avoid confounding by subsequent GWAS discoveries we used the 2006 data set. Results using the 2012 PubMed data are provided in **table S13** but were not used for the identification of candidate genes.

Pathway Analysis

Ingenuity Pathway Analysis (IPA) Knowledge Base March 2015 (Ingenuity Systems, CA, USA) was used to explore molecular pathways between proteins encoded by the 67 candidate genes from the 52 genome-wide significant loci. The genes were entered into the Ingenuity database and mapped to its corresponding object in the Ingenuity Knowledge Base. Networks were algorithmically generated based on their direct interactions, with a maximum size of 35 genes/proteins per network. IPA computes a P-value based on a Fisher's exact test for each network and biological function and/or disease with $\alpha=0.05$. This *P*-value represents the likelihood of the core genes in a network and biological function being found together due to random chance.

Gene-expression profiling

We collected human gene expression data from the Gene Expression Omnibus (GEO). We confined analyses to Affymetrix Human Genome U133 Plus 2.0 Arrays. We downloaded 43,278 raw CEL files, and used RMA for normalization (we ran RMA in eight batches due to its size, by randomly assigning the samples to one of these batches). We subsequently conducted quality control on the data. We first removed duplicate samples, and subsequently conducted a principal component analysis (PCA) on the sample correlation matrix. The first principal component (PCqc) on such a matrix describes nearly always a constant pattern (dominating the data) which explains around 80-90% of the total variance. This pattern can be regarded as probe-specific variance, independent of the biological sample hybridized to the array. The correlation of each individual microarray with this PCqc can be used to detect outliers, as arrays of lesser quality will have a lower correlation with the PCqc. We removed samples that had a correlation $R<0.75$. After quality control, in total, 37,427 samples remained. We used Ensembl to assign the 54,675 different probesets to 19,997 different Ensembl genes. If multiple probesets mapped within the same Ensembl gene, we averaged the probeset levels. Subsequently, the expression levels of each gene were standardized to a mean of zero and standard deviation of one and we collapsed the individual gene expression into groups of tissues, by using text-mining of the GEO sample descriptions in conjunction with Medical Subject Headings (MeSH) terms for the 'Anatomy' group. We subsequently used this tissue expression matrix to ascertain in what tissue the genes inside the associated loci were most abundantly expressed, and whether the expression of the candidate genes within those tissues were also higher, compared to all other genes that were measured in that tissue.

Drosophila melanogaster methods and results

We carried out permutation testing in a genome-wide phenotypic screen of cardiac specific RNAi-silencing of evolutionarily conserved genes under conditions of stress. Details of the genome-screen have been published previously.[50] We randomly selected sets of 67 genes, identified their *D. melanogaster* orthologs, and counted the number of orthologs with a cardiac phenotype in the RNAi screen. This was repeated 1,000,000 times to build up an expectation under the null hypothesis (**Fig. S10**). Next we determined which of the 67 candidate genes identified in the QRS GWAs had a phenotype in the genome-wide RNAi

screen and calculated the enrichment compared to the mean observed in simulations of the null hypothesis. We found that the 67 candidate genes were 2.3-fold enriched for stress-induced cardiac death (9 genes, P=1.84×10^{-2}; **Fig. S10**). Cardiac abnormalities in *D. melanogaster* had been reported for 3 of these 9 genes (*Mhc*/*MYH7B*[51], *Slit*/*SLIT2*[52], and *EcR/NR1H*[53]) and a role in cardiac genesis for one (*Hand*/*HAND1*[54]). Another gene (*TTN*) is a well-known cardiomyopathy gene in humans[55]

To illustrate that prioritized genes may play a critical role in heart development we tested 4 (*CG4743/SLC25A26, Fhos/FHOD3, Cka/STRN, NACα*/*NACA*) genes with unknown cardiac function by performing heart-specific RNAi knockdown with the cardiac Hand4.2-Gal4 driver line. We also re-tested *EcR/NR1H*, which has multiple homologous genes in mammals, as well as *Hand/HAND1* as this gene was only tested in as a full-knockout in early development but not in adult *D. melanogaster* heart using cardiac specific knockdown. Adult hearts of *EcR*/NR1H, *NACα*/NACA, *Hand*/HAND1 and Cka/STRN RNAi showed severe cardiac defects (**Fig. 4**). Knockdown of *Hand*/HAND1 and *Cka*/STRN both had a reduced cardiac heart rate. While *Hand*/HAND1 knockdown hearts appeared structurally normal, we observed severely disorganized and misoriented myofibrillar arrangements within the cardiomyocytes in *Cka*/STRN RNAi hearts (**Fig. 4**), which caused a reduction in diastolic diameters and contractility. NACα/NACA mutants had the most severe phenotype with a complete loss of cardiac tissue beginning at eclosion, while the hearts of NACα mutant larvae were still intact, indicating a critical role for NACA during cardiomyocyte remodelling. RNAi-mediated knockdown of *CG4743*/SLC25A26 and *Fhos*/FHOD3 did not reveal cardiac phenotypes. We also expanded on gene-by-gene analysis and identified further novel genes causing cardiac abnormalities. Of the 58 (67 minus 9 mentioned above) remaining candidate genes, 12 had no reported or otherwise discernible *D. melanogaster* homologs, while the other 46 genes had a total of 112 homologues (DIOPT score[56] ≥ 1). 27 human genes had 32 *Drosophila* orthologues with a DIOPT score 5-10, i.e. their homology was confirmed by 5-10 databases. Of these *D. melanogaster* homologues, we tested 21 genes for a potential cardiac function. Interestingly, cardiac knockdown of 6 of these genes caused cardiac arrhythmia and/or structural/contractility defects increasing the total number to 10 genes (out of 30 genes tested) with impaired cardiac structure or function (**table S17**).

***Experimental procedures**:* Flies were reared on standard cornmeal food at 25°C. For cardiac-specific knockdown we used the Hand4.2-Gal4 driver line[50,57], which expresses in cardiac and pericardial cells from the late embryo throughout all stages of the life cycle[58]. Hand4.2-Gal4 virgin females were crossed to male flies carrying RNAi constructs from the VDRC collection[59] targeting the genes to be tested. Female progeny from this cross was collected and aged for 3 weeks, after which they were dissected[60]. In brief, flies were immersed in artificial hemolymph (AHL) and abdomens were cut open ventrally and the intestine removed, which exposes the dorsally located beating heart tube. Specimen were allowed to rest for 15min in oxygenized AHL and then filmed at 120fps, using a Leica DM microscope and a Hamamatsu C9300 camera. Heart beat parameters were analysed using custom made software (SOHA[57,61]). After filming, the hearts were fixed and stained according to[62], and images were taken using a Zeiss Apotome. Fig.s were assembled using ImageJ[63] and Photoshop (Adobe).

Mus musculus models

We systematically searched the international database resource for the laboratory mouse (MGI-Mouse Genome Informatics) and manually curated Mammalian Phenotypes (MP) identifiers related to abnormal size and/or morphology of the heart, left ventricle, cardiac septum, cardiomyocyte, heart development and/or abnormal cardiac depolarisation and/or cardiac output. We identified 154 MP's (**table S27**) of a total of 9,338 available MPs. We selected random sets of 67 human genes, identified their mouse orthologs, and counted the number of orthologs with a kock-out mice model available which had been annotated with one or more of the identified 154 MPs. This was repeated 1,000,000 times to build up an expectation under the null hypothesis (**Fig. S10**). Next we determined which of the 67 candidate genes identified in the QRS GWAs had one or more of the identified 154 MPs and calculated the enrichment compared to the mean observed in simulations of the null hypothesis. Of the 67 candidate genes, of which 18 (40%) revealed a cardiac phenotype (**table S16**). This represents a 5.2-fold enrichment compared to randomly matched sets of 67 genes (P=3.4×10^{-14}; **Fig. S10**).

Data-driven Expression-Prioritized Integration for Complex Traits (DEPICT).

DEPICT systematically identifies the most likely causal gene at a given associated locus, tests gene sets for enrichment in associated SNPs, and identifies tissues and cell types in which genes from associated loci are highly expressed (see Pers et al.[64] for a detailed description of the method). First, DEPICT assigns genes to associated SNPs using LD r^2 > 0.5 distance to define locus boundaries, merges overlapping loci and discards loci mapping within the extended major histocompatibility complex region (chromosome 6, base pairs 25,000 – 35,000). Next, DEPICT prioritizes genes within the associated loci based on genes functional similarity to genes from other associated loci within the same GWAS (genes that are similar to genes from other loci obtain low prioritization P values), and adjusts for gene length bias as well as other potential confounders by use of simulated GWAS results. There can be several prioritized genes in a given locus. Next, DEPICT conducts gene set enrichment analysis by testing whether genes in associated loci are enriching for reconstituted versions of known biological pathways, gene sets as well as protein complexes. Leveraging the guilt by association hypothesis that genes co-expressing with genes from a given gene set are likely to be part of that gene set (See Cvejic et al.[65] for details), the gene set reconstitution is accomplished by identifying genes that were co-expressed with genes in a given gene set based on a panel of 77,840 gene expression microarrays. Gene sets from the following repositories were reconstituted: 5,984 protein complexes originating from 169,810 high-confidence experimentally-derived protein-protein interactions[66]; 2,473 phenotypic gene sets derived from 211,882 gene-phenotype pairs from the Mouse Genetics Initiative[67]; 737 Reactome database pathways[68]; 184 KEGG database pathways[69]; and 5,083 Gene Ontology database terms[70]. Finally, DEPICT conducts tissue and cell type enrichment analysis, by testing whether genes in associated loci are highly expressed in microarray-based gene expression data covering 209 Medical Subject Heading annotations (37,427 Affymetrix U133 Plus 2.0 Array samples are used for this analysis). See Wood *et al*[71]. , Geller *et al*[72]. and van der Valk
et al.[73] for previous applications of DEPICT. For this work we first used the PLINK software tool[18], to clump all SNPs with association *P* value < 1 x 10^{-5} as input (parameters: '--clump-p1 1e-5 --clump-kb 500 --clump-r2 0.05') resulting in 202 SNPs. DEPICT was run using 202 SNPs as input (resulting in the 149 independent loci using the DEPICT locus definition and

331 genes in total). The gene prioritization, gene set enrichment and tissue/cell type enrichment analyses were run using the default settings in DEPICT.

URLs

1000 Genomes; www.1000genomes.org

BEAGLE; http://faculty.washington.edu/browning/beagle/beagle.html

Birdseed; http://www.broadinstitute.org/mpg/birdsuite/birdseed.html

Chiamo; http://mathgen.stats.ox.ac.uk/genetics_software/chiamo/chiamo.html

DEPICT; http:/www.broadinstitute.org/depict

ENCyclopedia Of DNA Elements (ENCODE); http://www.encodeproject.org/ENCODE/

Ensembl project; http://www.ensembl.org/index.html

FIMO; http://meme.nbcr.net/meme/cgi-bin/fimo.cgi

Galaxy; http://galaxyproject.org/

GenABEL; http://www.genabel.org/packages

Gene Expression Omnibus (GEO); http://www.ncbi.nlm.nih.gov/geo/

IMPUTE; https://mathgen.stats.ox.ac.uk/impute/impute.html

Ingenuity; http://www.ingenuity.com

International HapMap Project; http://hapmap.ncbi.nlm.nih.gov/

JASPAR; http://jaspar.genereg.net/

MACH; http://www.sph.umich.edu/csg/abecasis/mach/

MANTEL; http://www.broadinstitute.org/~debakker/mantel.html

METAL; http://www.sph.umich.edu/csg/abecasis/metal/

MGI-Mouse Genome Informatics; www.informatics.jax.org

PLINK; http://pngu.mgh.harvard.edu/~purcell/plink/

ProABEL; http://www.genabel.org/packages/ProbABEL

R; http://www.r-project.org/

Roadmap epigenomics project; http://www.roadmapepigenomics.org/

SNAP; http://www.broadinstitute.org/mpg/snap/

SNPTEST; https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html

TRANSCompel; http://www.gene-regulation.com/pub/databases.html#transcompel

UCSC Genome Browser; http://genome.ucsc.edu/

UniProbe; http://the_brain.bwh.harvard.edu/uniprobe/

Cohorts methods

AGES: The Age, Gene/Environment Susceptibility (AGES) Reykjavik Study was initiated to examine genetic susceptibility and gene/environment interaction as these contribute to phenotypes common in old age, and represents a continuation of the Reykjavik Study cohort begun in 1967 and is comprised of 5776 randomly recruited survivors from the original cohort. QRS interval duration was automatically measured from 12-lead electrocardiograms using the Marquette 12 SL analysis program (General Electric Marquette Medical Division, Milwaukee, Wisconsin, USA).

ARIC: The Atherosclerosis Risk in Communities study (http://www.cscc.unc.edu/aric/) includes 15,792 men and women from four communities in the United States (Jackson, Mississippi; Forsyth County, North Carolina; Washington County, Maryland; suburbs of Minneapolis, Minnesota) enrolled in 1987–1989 and prospectively followed. ECGs were recorded at baseline using MAC PC ECG machines (Marquette Electronics) and processed initially by the Dalhousie ECG program in a central laboratory at the EPICORE Center (University of Alberta). Processing was later repeated for the present study using the GE Marquette 12-SL program (2001 version) at the EPICARE Center (Wake Forest University). All ECGs were visually inspected for technical errors and inadequate quality. QRS voltage was calculated from parameters automatically measured from baseline ECGs.

BRIGHT: The MRC BRIGHT study (http://www.brightstudy.ac.uk/) comprises 2000 severely hypertensive probands ascertained from families with multiplex affected sibships or as parent-offspring trios. Case ascertainment and phenotyping has been described previously. Briefly, cases have BP readings ≥150/100 mmHg based on one reading or ≥145/95 mmHg based on the mean of three readings. Twelve-lead ECG recordings (Siemens-Sicard 440; http://www.brightstudy.ac.uk/info/sop04.html), which produces an automated measurement of the QRS voltage and duration, were available for all subjects. All data were transferred from each recruitment centre by electronic modem to electrophysiologists from the West of Scotland Primary Prevention Study (Professor Peter MacFarlane) for central reporting. All individuals included in the analysis were of white British ancestry (up to level of grandparents).

CHS: The CHS is a population-based cohort study of risk factors for CHD and stroke in adults ≥65 years conducted across four field centers.[74] The original predominantly Caucasian cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists; subsequently, an additional predominantly African-American cohort of 687 persons were enrolled for a total sample of 5,888. DNA was extracted from blood samples drawn on all participants at their baseline examination in 1989-90. In 2007-2008, genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai using the Illumina 370CNV BeadChip system on 3980 CHS participants who were free of CVD at baseline, consented to genetic testing, and had DNA available for genotyping. *Participant-level exclusions:* A total of 1908 persons were excluded from the GWAS study sample due to the presence at study baseline of coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke or transient ischemic attack or lack of available DNA. Because the other cohorts were predominantly white, the African American participants were excluded from this analysis to reduce the possibility of confounding by population structure. Samples were excluded from analysis for sex mismatch, discordance with prior genotyping, or call rate < 95%. To date, genotyping has been successful among 3,271 of 3,373 European Ancestry participants; the 2642 of these participants with available QRS voltage phenotypes who satisfied the phenotypic exclusion criteria constitute the CHS sample for this study. *Genotyping Detail:* In CHS, genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai using the Illumina 370CNV BeadChip system. Genotypes were called using the Illumina BeadStudio software as above. *QRS Voltage Duration:* Study electrocardiograms were recorded using MAC PC ECG machines (Marquette Electronics, Milwaukee, Wisconsin) in all clinical centers. ECGs were initially processed in a central laboratory at the EPICORE Center (University of Alberta, Edmonton, Alberta, Canada) and during later phases of the study, at the EPICARE Center (Wake Forest University, Winston-Salem, North Carolina). All ECGs were visually inspected for technical errors and inadequate quality. QRS interval was measured using the baseline ECG for eligible subjects. Initial ECG processing was done by the Dalhousie ECG program, and processing was later repeated with the 2001 version of the GE Marquette 12-SL program (GE 342 Marquette, Milwaukee, Wisconsin).

Cilento: Cilento study is a population-based study that includes 2,137 subjects from isolated populations located in the area of the National Park of Cilento e Vallo di Diano, South Italy. The ECG and genotype data were available for a subset of 629 individuals. Standard echocardiography was used to assess the ventricular function and to obtain the ECG measurements included the QRS voltage duration. The study design was approved by the ethics committee of Azienda Sanitaria Locale Napoli 1. The study was conducted according to the criteria set by the declaration of Helsinki and each subject signed an informed consent before participating to the study.

ERF: The Erasmus Rucphen Family study is comprised of a family-based cohort embedded in the Genetic Research in Isolated Populations (GRIP) program in the southwest of the Netherlands[75]. The aim of this program is to identify genetic risk factors for the development of complex disorders. In ERF, twenty-two families that had a large number of children baptized in the community church between 1850 and 1900 were identified with the help of detailed genealogical records. All living descendants of these couples, and their spouses, were invited to take part in the study. Comprehensive interviews, questionnaires, and examinations were completed at a research center in the area; approximately 3,200 individuals participated. Examinations included 12 lead ECG measurements. Electrocardiograms were recorded on ACTA electrocardiographs (ESAOTE, Florence, Italy) and digital measurements of the QRS intervals were made using the Modular ECG Analysis System (MEANS). The QRS detector of MEANS operates on multiple simultaneously recorded leads, which are transformed to a detection function that brings out the QRS complexes among the other parts of the signal. MEANS was used to measure QRS complex duration and the three LVH proxies. Data collection started in June 2002 and was completed in February 2005. In the current analyses, 1,722 participants for whom complete phenotypic, genotypic and genealogical information was available were studied.

FHS: The Framingham Heart Study (http://www.framinghamheartstudy.org/) is a community-based, longitudinal cohort study comprising three generations of individuals in multigenerational pedigrees and additional unrelated individuals. The current study included individuals from Generation 1 (11th examination), Generation 2 (1st examination) and Generation 3 (1st examination). In FHS, paper electrocardiograms recorded on Marquette machines were scanned and digital caliper measurements were made using proprietary

software (eResearchTechnology, generations 1 and 2) or using Rigel 1.7.2. (AMPS, LLC, New York, NY, USA, generation 3). The QRS duration was measured from the Q-onset to S-offset in two cardiac cycles from lead II and averaged.

FVG: The Friuli Venezia Giulia Project was initiated in 2008. All population was invited to partecipate. The cohort included 1,739 people. At the moment we have genotypied 1,378 samples. Our cardiologists enrolled 1,548 unselected inhabitants, the people with EKG and genotype data are : 1,269. A free screening (anamnesis, electrocardiogram, echocardiogram) of the whole population was offered by a team of three cardiologists. A consent form either for clinical and genetic studies has been signed by each participant in the study. The method of measurement of QRS voltage duration: the QRS duration have been measured in milliseconds using the automatic refertation powered by the electrocardiographer used for the recordings (Mortara Instrument ELI 250). QRS voltages have been measured in millimeters by two cardiologists. The three phenotypes have been evaluated and calculated by the same cardiologists.

InChianti: The InCHIANTI study is a population-based epidemiological study aimed at evaluating the factors that influence mobility in the older population living in the Chianti region in Tuscany, Italy (http://inchiantistudy.net). The details of the study have been previously reported.[76] Briefly, 1,616 residents were selected from the population registry of Greve in Chianti (a rural area: 11,709 residents with 19.3% of the population greater than 65 years of age), and Bagno a Ripoli (Antella village near Florence; 4,704 inhabitants, with 20.3% greater than 65 years of age). The participation rate was 90% (n=1453), and the subjects ranged between 21-102 years of age. Overnight fasted blood samples were for genomic DNA extraction.[77] The study protocol was approved by the Italian National Institute of Research and Care of Aging Institutional Review and Medstar Research Institute (Baltimore, MD).

KORA F3 and S4: The KORA study is a series of independent population-based epidemiological surveys of participants living in the city of Augsburg, Southern Germany, or the two adjacent counties. All survey participants are residents of German nationality identified through the registration office and aged between 25 and 74 years at recruitment. The baseline survey KORA S3 was conducted in the years 1994/95 and KORA S4 in 1999-

2001. 3,006 participants from KORA S3 were reexamined in a 10-year follow-up (KORA F3) in the years 2004/05. Genomewide data for the analysis of the length of the QRS interval is available for random subsets of 1,644 persons from KORA F3 and 1,814 study participants from KORA S4. In both studies, 12-lead resting electrocardiograms were recorded with digital recording systems (F3: Mortara Portrait, Mortara Inc., Milwaukee, USA, S4: Hörmann Bioset 9000, Hörmann Medizinelektronik, Germany).[78,79]

KORCULA: The KORCULA study sampled Croatians from the Adriatic island of Korcula, between the ages of 18 and 88. The fieldwork was performed in 2007 in the eastern part of the island, targeting healthy volunteers from the town of Korčula and the villages of Lumbarda, Žrnovo and Račišće. Mortara ELI 350 was used in ECG recording.

LifeLines: LifeLines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviours of 165,000 persons living in the North East region of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multimorbidity and complex genetics. Details of the protocol have been described elsewhere (https://www.lifelines.nl/lifelines-research/news). Standard 12-lead electrocardiograms were recorded with CardioPerfect equipment (Cardio Control; currently Welch Allyn, Delft, The Netherlands) and digital measurements of the QRS intervals were extracted.

LOLIPOP: The London Life Sciences Population study (www.lolipopstudy.org) is an ongoing population-based cohort study of ~30,000 individuals (18,000 Indian Asians and 12,000 European white men and women), aged 35-75 years and recruited from the lists of 58 general practitioners in West London, United Kingdom. A nurse-led interviewer-administered questionnaire was used to collect data on medical history, family history, current prescribed medications and cardiovascular risk factors. Physical assessment included measurements of height, weight, waist and hip circumference as well as blood pressure. In addition a 12 lead ECG was recorded using GE Marquette CARDIOSOFT software. Following an 8 hour fast, blood was collected for biochemical analysis, and whole blood taken for extraction of DNA. Subgroups of European White (EW) participants were

genotyped on 3 different GWAS platforms – Affymetrix (EWA), Perlegen (EWP) and Illumina610 (EW610) arrays.[80,81]

MESA: http://www.mesa-nhlbi.org/Mesa-Internal/MESASHARe.asp. Standard 12-lead ECGs were digitally acquired using a Marquette MAC 1200 ECG machines (Marquette Electronics, Milwaukee, WI) at 10 mm/mV calibration and speed of 25 mm/s. ECGs were processed in a central laboratory at the EPICARE Center (Wake Forest University, Winston-Salem, NC). All ECGs were visually inspected for technical errors and inadequate quality. ECG processing was done by the 2001 version of the GE Marquette 12-SL program (GE Marquette, Milwaukee, WI). QRS interval and voltage measures were calculated automatically.

MICROS: The MICROS study is part of the genomic health care program 'GenNova' and was carried out in three villages of the Val Venosta on the populations of Stelvio, Vallelunga and Martello. This study was an extensive survey carried out in South Tyrol (Italy) in the period 2001-2003. Study participants were volunteers from three isolated villages located in the Italian Alps, in a German-speaking region bordering with Austria and Switzerland. Due to geographical, historical and political reasons, the entire region experienced a prolonged period of isolation from surrounding populations. Genotyping was performed on just under 1,400 participants with 1,334 available for analysis after data cleaning. Information on participants' health status was collected through a standardized questionnaire and clinical examinations, including digitized ECG measurements (Mortara Portrait, Mortara Inc., Milwaukee, USA). Individuals with identified U-waves were excluded from analysis. The Mortara portrait determines QRS complex by a proprietary algorithm[82]. Laboratory data were obtained from standard blood analyses.[83]

ORCADES: The Orkney Complex Disease Study (ORCADES) is an ongoing family-based, cross-sectional study in the isolated Scottish archipelago of Orkney. Genetic diversity in this population is decreased compared to Mainland Scotland, consistent with high levels of endogamy historically. Participants included here were aged 18-92 years and came from a subgroup of ten islands. The Cardioview ECG device was used in the phenotyping.

PREVEND: The Prevention of REnal and Vascular ENd stage Disease (PREVEND) study is an ongoing prospective study investigating the natural course of increased levels of urinary albumin excretion and its relation to renal and cardiovascular disease. Inhabitants 28 to 75 years of age (n=85,421) in the city of Groningen, The Netherlands, were asked to complete a short questionnaire, 47% responded, and individuals were then selected with a urinary albumin concentration of at least 10 mg/L (n = 7,768) and a randomly selected control group with a urinary albumin concentration less than 10 mg/L (n = 3,395). Details of the protocol have been described elsewhere (www.prevend.org). Standard 12-lead electrocardiograms were recorded with CardioPerfect equipment (Cardio Control; currently Welch Allyn, Delft, The Netherlands) and digital measurements of the QRS intervals were extracted.

PROSPER: The protocol of PROSPER has been described elsewhere.[84] PROSPER is a prospective multicenter randomized placebo-controlled trial to assess whether treatment with pravastatin diminishes the risk of major vascular events in elderly individuals. Between December 1997 and May 1999, subjects were screened and enrolled in Scotland (Glasgow), Ireland (Cork), and the Netherlands (Leiden). Men and women aged 70-82 years were recruited if they had pre-existing vascular disease or increased risk of such disease because of smoking, hypertension, or diabetes. A total number of 5804 subjects were randomly assigned to pravastatin or placebo. In this study, the predefined endpoints all-cause mortality and mortality were evaluated due to vascular events and coronary heart disease death. Mean follow-up was 3.2 years (range 2.8-4.0) and 604 (10.4%) patients died during the study.[85] The GWAS was performed in a random sub-sample of 5244 subjects.

RS: The Rotterdam Study is a prospective population-based cohort study comprising 7,983 subjects aged 55 years or older (RS-I), which started in 1990. In 2000-2001, an additional 3,011 individuals aged 55 years or older were recruited (RS-II). In the RS-I and RS-II, electrocardiograms were recorded on ACTA electrocardiographs (ESAOTE, Florence, Italy) and digital measurements of the QRS intervals were made using the Modular ECG Analysis System (MEANS). The QRS detector of MEANS operates on multiple simultaneously recorded leads, which are transformed to a detection function that brings out the QRS complexes among the other parts of the signal.

Sardinia: The SardiNIA GWAS examined a total of 5,014 related individuals participating in a longitudinal study of aging-related quantitative traits in the Ogliastra province of the Sardinia region, Italy. The study has been described in detail previously[86]. Included individuals had four Sardinian grandparents and were selected without regard to their phenotypes. The ECG was recorded on paper (ECG machine Cardiette 600) with the participant at rest. Images of paper records were digitalised and voltages were manually measured using ImageJ software.

SHIP: The Study of Health in Pomerania (SHIP, www.medizin.uni-greifswald.de/cm/fv/english/ship_en.html) is a longitudinal population-based cohort study in West Pomerania, a region in the northeast of Germany. From the total population comprising 212,157 inhabitants in 1995, a two-stage stratified cluster sample of adults aged 20 to 79 years was drawn. From the net sample of 6265 eligible subjects, 4308 subjects (2192 women) of Caucasian origin participated in the baseline examination, SHIP-0 (response 68.8%). All participants gave written informed consent. The study conformed to the principles of the Declaration of Helsinki and was approved by the Ethics Committee of the University of Greifswald. For the present analyses both electrocardiographic and genotyping data were available from 2978 participants of the SHIP baseline cohort without exclusion criteria.
QRS intervals and voltages in SHIP were measured from digitally stored electrocardiograms (Personal 120LD, Esaote, Genova, Italy) using MEANS according to the method described above for the Rotterdam Study (RS).

SPLIT: The SPLIT study samples Croatians from the town of Split, between the ages 18 and 85. The sampling started in 2008, and continues throughout 2010. Mortara ELI 350 was used in ECG recording.

Twins_UK: The Twins UK Registry comprises unselected, mostly female volunteers ascertained from the general population through national media campaigns in the UK (1; www.twinsuk.co.uk). Means and ranges of quantitative phenotypes in Twins UK were similar to an age-matched singleton sample from the general population. Zygosity was determined by standardized questionnaire and confirmed by DNA fingerprinting. Written informed consent was obtained from all participants before they entered the studies, which

were approved by the local research ethics committee. Using linear regression analysis, we adjusted QRS Voltage product for age, sex, height, and body mass index. The residuals were used for further analyses. Association between QRS Voltage product residuals and autosomal SNPs was tested with an F-test in SNPTEST version 1.1.4 using an additive model and the proper option to account for the uncertainty of the genotypes that were imputed. As the TwinsUK cohort data consisted partly of dizygotic twins, the variances of the regression coefficients were corrected for the sibship relations using the Huber-White method for robust variance estimation in R.

YFS: The Cardiovascular Risk in Young Finns (YFS) is a population-based 27 year follow up-study (http://med.utu.fi/cardio/youngfinnsstudy/). The first cross-sectional survey was conducted in 1980, when 3,596 Caucasian subjects aged 3-18 years participated. In adulthood, the latest 27-year follow-up study was conducted in 2007 (ages 30-45 years) with 2,204 participants. The study cohort for the present analysis comprised subjects who had participated in the ultrasound study in 2007 and had other risk factor data.[87] Method of measurement of QRS voltage duration: GE CardioSoft Ver 4.2 (Tampere cohort, approximately ½ of the samples) and GE MUSE Ver 7.1.0 (Turku cohort, approximately ½ of the samples)

Author contributions

Study Organisation: P.v.d.H., A.Isaacs, N.J.S., and P.I.W.d.B. Manuscript preparation: P.v.d.H., J.v.S., N.V., G.V., L.F., M.M., X.W., I.M.L., J.C.C., Y.J., A.V., V.M.C. A.Isaacs, N.J.S., and P.I.W.d.B. All authors reviewed and had the opportunity to comment on the manuscript. Data collection and analysis in the participating genome-wide association, replication and phenotype cohorts: **AGES**: K.A., V.G., T.B.H., L.J.L., A.V.S.; **ARIC**: A.A., D.E.A., A.C., E.Z.S.; **BRIGHT**: A.F.D., P.B.M., P.W.M., S.Padmanabhan, N.J.S.; **CHS**: N.S., K.M.R., J.C.B., B.M.P. **Cilento**: M.C., S.Nappo, T.N., R.S.; **ERF**: A.Isaacs, B.A.O., C.T.S., C.M.v.D.; **FHS**: D.L., J.W.M., C.N.-C., S-J.H.; **FVG**: P.G., A.Iorio, G.S., S.U.; **InChianti**: S.B., L.Ferruci, D.G.H., T.T.; **KORA F3 and S4**: C.G., S.K., T.M., M.M.-N., S.Perz, A.Peters, A.Pfeufer, M.F.S., K.Strauch, M.W.; **KORCULA**: C.H., I.K., I.R., V.V., A.F.W.; **LifeLines**: H.L.H., R.P.S., P.v.d.H., D.J.v.V., N.V., C.W., B.H.W.; **LOLIPOP**: J.C.C., I.K.K., J.S.C., J.S., S.-T.T., W.Z.; **MESA**: L.Y.C., S.R.H., K.F.K., J.I.R.; **MICROS**: F.D.G.M., F.S.D., A.A.H., P.P.P., C.X.W., A.Pfeufer; **ORCADES**: H.C., O.P., J.F.W.; **PREVEND**: M.V.C., R.A.d.B., I.M.L., P.v.d.H., W.H.v.G., D.J.v.V.; **PROSPER**: B.M.B., I.F., J.W.J., P.W.M., S.T.; **RS**: M.E., A.H., J.A.K., B.H. S., A.G.U.; **SARDINIA**: M.G.P., G.R.A., A.D., P.v.d.H., E.G.L., O.M., S.Naitza, S.Sanna, D.S., H.H.W.S., K.V.T.; **SHIP**: M.D., S.B.F., C.O.S., U.V.; **SMART**: F.W.A., P.A.D., W.S., V.T.; **TWINSUK**: Y.J., I.M.N., B.P.P., H.S., T.D.S.; **YFS**: M.Kähönen, T.L., L.-P.L., O.T.R., J.S.V. Functional studies: *Drosophila*, R.B., B.T., G.V., *Mus Musculus*, P.B., V.M.C., M.v.d.B., D.M., A.S.N., L.A.P., A.V., expression profiling: L.Franke, H.J.W., T.E., A.M., functional elements; N.V., L.A.B., X.W., E.H., M.T.M., S.Neph, J.A.S., P.B., V.M.C., M.v.d.B., M.Kellis, 4C: P.B., V.M.C., M.v.d.B. *RNA-seq human heart*: M.E.A., C.R.B., Y.M.P., M.H., S.Schafer, N.H., Data analysis and bioinformatics: P.v.d.H., N.V., J.v.S., A.Isaacs, I.M.L., P.I.W.d.B., P.M.V., J.Y., L.Franke, T.H.P., J.N.H., H.J.W., A.V., K.L., A.L., J.V.O., S.R., K.Slowikowski, L.A.B., X.W., P.B., V.M.C., M.v.d.B., D.M., E.J.R., A.S.N., L.A.P., A.V., E.H., M.T.M., S.Neph, J.A.S..

Acknowledgements

AGES: The Age, Gene/Environment Susceptibility Reykjavik Study has been funded by NIH contract N01-AG-12100, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, (VSN: 00-063) and the Data Protection Authority. The researchers are indebted to the participants for their willingness to participate in the study. **ARIC:** The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. **BRIGHT:** The BRIGHT study was funded by the Medical Research Council of Great Britain (G9521010D) and the British Heart Foundation (PG/02/128). Genotyping was funded by the Wellcome Trust (grant number; 076113/B/04/Z) as part of The Wellcome Trust Case Control Consortium. The BRIGHT study is extremely grateful to all the patients who participated in the study and the BRIGHT nursing team. N.J.S. holds a Chair funded by the British Heart Foundation and is a National Institute for Health Research (NIHR) Senior Investigator. This work forms part of the research themes contributing to the translational research portfolio for the Leicester (N.J.S.) and Barts Cardiovascular Biomedical Research (P.B.M.) Units which are supported and funded by the National Institute for Health Research. **CHS:** This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, and R01HL130114 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through AG023629 from the National Institute on Aging (NIA). A full list of CHS investigators and institutions can be found at http://www.chs-nhlbi.org/pi.htm. The provision

of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. NS was supported by HL111089 and HL116747. **Cilento:** We thank the populations of Cilento for their participation in the study. This work was supported by grants from the Italian Ministry of Universities (FIRB -RBIN064YAT, INTEROMICS Flaghip Project), the Assessorato Ricerca Regione Campania, the Fondazione con il SUD (2011-PDR-13), and the Fondazione Banco di Napoli to M.C. **ERF:** The ERF study was supported by grants from the Netherlands Organization for Scientific Research (NWO; Pioneergrant), Erasmus Medical Center, the Centre for Medical Systems Biology (CMSB), and the Netherlands Kidney Foundation. We are grateful to all patients and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection. **FHS:** The Framingham Heart Study is funded by National Institutes of Health contract N01-HC-25195. The laboratory work for this investigation was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health. The analytical component of this project was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, and the Center for Information Technology, National Institutes of Health, Bethesda, MD. This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. The authors acknowledge the essential role of the Cohorts for Heart and Aging Research in Genome Epidemiology (CHARGE) Consortium in development and support of this manuscript, especially acknowledging members from the CHARGE ECG working group. Measurement of the electrocardiographic voltage in the 3rd generation was supported by the Burroughs Wellcome Fund (Newton-Cheh, Doris Duke Charitable Foundation (C.N.-C.) and the NIH (HL080025, C.N.-C.). **INGI-FVG:** We thank Anna Morgan and Angela D'Eustacchio for technical support. We are very grateful to the municipal administrators for their collaboration

on the project and for logistic support. We would like to thank all participants to this study. We also thank Associazione Amici del Cuore and Fondazione Cassa di Risparmio Trieste. The study was supported by Regione FVG (L.26.2008) **InChianti:** The InCHIANTI study baseline (1998-2000) was supported as a "targeted project" (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336). **KORA:** The KORA research platform (KORA, Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München - German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. **KORCULA:** KORCULA was supported by the MRC Human Genetics Unit, The Croatian Ministry of Science, Education and Sports (grant 216-1080315-0302), the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947), EU FP7 BBMRI-LPC (Biobanking and biomolecular resources research infrastructure - Large prospective cohort, contract 313010) and the Croatian Science Foundation (grant 8875). We would like to acknowledge administrative team in Split and the people of Korcula. **LIFELINES:** The LifeLines Cohort Study, and generation and management of GWAS genotype data for the LifeLines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes Research Foundation. We thank Behrooz Alizadeh, Annemieke Boesjes, Marcel Bruinenberg, Noortje Festen, Ilja Nolte, Lude Franke, Mitra Valimohammadi for their help in creating the GWAS database, and Rob Bieringa, Joost Keers, René Oostergo, Rosalie Visser, Judith Vonk for their work related to data-collection and validation. The authors are grateful to the study participants, the staff from the LifeLines Cohort Study and Medical Biobank Northern Netherlands, and the participating general practitioners and pharmacists. LifeLines Scientific Protocol Preparation: Rudolf de Boer, Hans Hillege, Melanie van der Klauw, Gerjan Navis, Hans Ormel, Dirkje Postma, Judith Rosmalen, Joris Slaets, Ronald Stolk, Bruce Wolffenbuttel; LifeLines GWAS Working Group: Behrooz Alizadeh, Marike

Boezen, Marcel Bruinenberg, Noortje Festen, Lude Franke, Pim van der Harst, Gerjan Navis, Dirkje Postma, Harold Snieder, Cisca Wijmenga, Bruce Wolffenbuttel. The authors wish to acknowledge the services of the LifeLines Cohort Study, the contributing research centres delivering data to LifeLines, and all the study participants. **LOLIPOP:** The LOLIPOP study is supported by the National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust, the British Heart Foundation (SP/04/002), the Medical Research Council (G0601966,G0700931), the Wellcome Trust (084723/Z/08/Z), the NIHR (RP-PG-0407-10371), European Union FP7 (EpiMigrant, 279143), and Action on Hearing Loss (G51). The work was carried out in part at the NIHR/Wellcome Trust Imperial Clinical Research Facility. We thank the participants and research staff who made the study possible. **MESA:** MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-001079, UL1-TR-000040, and DK063491. **MESA SNP Health Association Resource (SHARe)**: Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-64278. Genotyping was performed at Affymetrix (Santa Clara, California, USA) and the Broad Institute of Harvard and MIT (Boston, Massachusetts, USA) using the Affymetrix Genome-Wide Human SNP Array 6.0. **MESA Air Acknowledgment;** This publication was developed under a STAR research assistance agreement, No. RD831697 (MESA Air), awarded by the U.S Environmental protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this publication." **MICROS:** For the MICROS study, we thank the primary care practitioners and the personnel of the Hospital of Silandro (Department of Laboratory Medicine) for their participation and collaboration in the research project. In South Tyrol, the study was supported by the Ministry of Health and Department for Innovation, Universities and Research of the Autonomous Province of Bolzano, South Tyrol, the South Tyrolean Sparkasse Foundation, and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). **ORCADES:** ORCADES was supported by the Chief Scientist Office of the Scottish Government, the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK and the European

Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney. **PREVEND:** PREVEND genetics is supported by the Dutch Kidney Foundation (Grant E033), the EU project grant GENECURE (FP-6 LSHM CT 2006 037697), the National Institutes of Health (grant 2R01LM010098), The Netherlands organisation for health research and development (NWO-Groot grant 175.010.2007.006, NWO VENI grant 916.761.70, ZonMw grant 90.700.441), and the Dutch Inter University Cardiology Institute Netherlands (ICIN). **PROSPER:** The PROSPER study was supported by an investigator initiated grant obtained from Bristol-Myers Squibb. J.W.J. is an Established Clinical Investigator of the Netherlands Heart Foundation (grant 2001 D 032). Support for genotyping was provided by the seventh framework program of the European commission (grant 223004) and by the Netherlands Genomics Initiative (Netherlands Consortium for Healthy Aging grant 050-060-810). **RS:** The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. M.E. is supported by grant 2007B221 if The Netherlands Heart Foundation. **SardiNIA:** The SardiNIA team was supported by Contract NO1-AG-1-2109 from the National Institute on Aging and in part by the Intramural Research Program of the NIH, National Institute on Aging. The efforts of G.R.A. were supported in part by contract 263-MA-410953 from the National Institute on Aging to the University of Michigan and by

research grants from the National Human Genome Research Institute and the National Heart, Lung, and Blood Institute (to G.R.A.). We thank the cardiologists Angelo Scuteri and Marco Orrù for their supervision of the work, Giuseppe Basciu for scanning ECG records, the team of physicians and nurses that carried out the physical examination and made the observation and the genotyping team of biologists. **SHIP:** SHIP is part of the Community Medicine Research net of the University of Medicine Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-West Pomerania. The University of Greifswald was a member of the 'Center of Knowledge Interchange' program of the Siemens AG. This study was furthermore carried out in collaboration with the German Centre for Cardiovascular Research (GCCR), which is funded by the Federal Ministry of Education and Research and the Ministry of Cultural Affairs of the Federal State of Mecklenburg, West Pomer- ania, Germany. **SPLIT:** SPLIT was supported by the MRC Human Genetics Unit, The Croatian Ministry of Science, Education and Sports (grant 216-1080315-0302) and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). We would like to acknowledge administrative teams in Split and Zagreb and the people of Split. **Twins-UK:** Twins UK (TUK): The study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013). The study also receives support from the National Institute for Health Research (NIHR) BioResource Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London and the British Heart Foundation. Tim Spector is holder of an ERC Advanced Principal Investigator award. SNP Genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR. Analyses were performed on the Genetic Cluster Computer, which is financed by an NWO Medium Investment grant 480-05-003 and by the Faculty of Psychology and Education of the VU University, Amsterdam, The Netherlands. Y.J. received funding for the project from the British Heart Foundation (PG/12/38/29615 and PG/06/094/21278). **YFS:** The Young Finns Study has been financially supported by the Academy of Finland: grants 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the

Social Insurance Institution of Finland; Kuopio, Tampere and Turku University Hospital Medical Funds (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research ; Finnish Cultural Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Jahnsson Foundation; and Signe and Ane Gyllenberg Foundation. **eQTL:** L. Franke is supported by the Netherlands Organization for Scientific Research [NWO-VENI grant 916.10.135] and a Horizon Breakthrough grant from the Netherlands Genomics Initiative [grant 92519031]. The research leading to these results has received funding from the European Community's Health Seventh Framework Programme (FP7/2007–2013) under grant agreement n° 259867. This study was financed in part by the SIA-raakPRO subsidy for project BioCOMP. **EGCUT**: was supported by grant from EstRC IUT20-60 and PerMed I, funding from the European Regional Development Fund for the development of Centre of Exellence "GENTRANSMED", and was also supported by EU H2020 grants 692145, 676550, 654248, Estonian Research Council Grant IUT20-60, NIASC, EIT – Health and NIH-BMI Grant No: 2R01DK075787-06A1 and EU through the European Regional Development Fund (Project No. 2014-2020.4.01.15-0012 GENTRANSMED. We acknowledge EGCUT technical personnel, especially Mr V. Soo and S. Smit. Data analyses were carried out in part in the High Performance Computing Center of University of Tartu. **Functional Genomics (LBNL):** A.V., L.A.P., A.S.N., and D.M. were supported by NHGRI grants R01HG003988, R24HL123879, U01DE024427, UM1HL098166, and U54HG006997. A.S.N. was supported by NIGMS fellowship F32GM105202. Research conducted at Lawrence Berkeley National Laboratory was performed under Department of Energy Contract DE-AC02-05CH11231, University of California. **SMART**: SMART genotyping was financially supported by BBMRI-NL, a Research Infrastructure financed by the Dutch government (NWO 184.021.007). F.W.A. is supported by UCL Hospitals NIHR Biomedical Research Centre and Dekker scholarship-Junior Staff Member 2014T001 – Netherlands Heart Foundation. **UQ data analysis**; this research was supported by the Australian National Health and Medical Research Council (APP1050218, APP1048853, APP1052684; the content is solely the responsibility of the authors and does not necessarily represent the official view of the funding body). **RNA-seq human heart**: supported by the Helmholtz Alliance ICEMED, the Deutsche Forschungsgemeinschaft (Forschergruppe 1054, HU 1522/1-1) Bonn, Germany, and the Fondation Leducq. We thank Nanette H. Bishopric, Alfred L. George, Andras Varro, and Cristobal dos Remedios for providing samples for genotyping and RNA-seq gene

expression analysis in human donor hearts. We thank Tamara Koopmann and Margriet Westerveld for their help with human donor heart sample processing. The eQTL analysis in human donor hearts was supported by the Cardiovascular Onderzoek Nederland PREDICT project to C.R. Bezzina). We acknowledge the support from the Netherlands CardioVascular Research Initiative: the Dutch Heart Foundation, Dutch Federation of University Medical Centres, the Netherlands Organisation for Health Research and Development and the Royal Netherlands Academy of Sciences. T.H.P. is supported by The Lundbeck Foundation and The Alfred Benzon Foundation. M.E.A.'s work at the Maastricht Centre for Systems Biology has been made possible with the support of the Dutch Province of Limburg. **Other**: American Heart Association 15GRNT25670044 to L.A.B The project described was supported by award Number T32GM007753 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

Supplementary Tables

1. Table S1. Characteristics of participants in genome-wide and replication cohorts

Cohort	N, participants with EKG and genotype data	N, participants after exclusion	Sex, (% women)	Age, years (mean ± SD)	Age (range)	Height, cm (mean ± SD)	BMI, kg/m2 (mean ± SD)	Hypertension (%)	Diabetes mellitus (%)	Heart rate, bpm (mean ± SD)
Discovery										
AGES	3,043	2,249	64	76.0±5.4	66-95	166±9	27.0±4.5	78	10	66.6±11.4
ARIC	8,961	7,414	55	54.1±5.7	45-64	168±9	27.1±4.8	26	8	66.5±9.9
Bright	1,021	882	62	55.8±11.3	21-89	166±9	27.6±3.8	100	0	63±11
Cilento	629	422	60	52.8±18.6	14-93	161±10	26.6±4.5	42	8	73.4±13.4
CHS	3,223	2,642	64	72.0±5.1	65-94	164±9	26.2±4.5	52	11	64.3±10.2
ERF	2,042	1,722	60	47.1±14.1	18-85	167±9.2	26.7±4.6	48	4	63.0±10.5
FHS	4,095	3,869	54	40±9	19-72	170±9	26.9±5.5	16	3	62±10
FVG	1,269	1,001	59	48.3±17.6	6-90	168 ± 10	24.9±4.5	48	6	66.1±11.7
Inchianti	1,073	861	57	66.1±15.4	21-95	160+10	27.1+4.1	42	9	69.0+11.6
KORA S4	1,786	1,491	56	53.5±8.7	25-74	167±9	27.6±4.6	32	3	65.4±10.2
KORAF3	1,644	1,393	52	61.4	35-79	167	27.9	42	9	64.1
Korcula	428	376	63	54.0±13.2	18-88	168 ± 9	27.9±4.2	29	6	65.6±9.6
LifeLines	8017	7,818	58	47.5±11	18-89	175±10	26.2±4.2	23	2	68.2±11.4

LOLIPOP_EW610	785	762	30	55.9±9.9	35-75	172±9	27.4±4.6	8	8	65.4±11.2
LOLIPOP_EW_P	263	181	0	56.1±8.9	35-67	176±7	29.1±5.5	18	14	67.9±12.4
LOLIPOP_EW_A	432	158	42	52.2±12.1	35-75	170±9	29.0±5.8	9	13	66.6±11.4
MESA	2,495	1,735	62	62.2±10.1	44-84	167±9	27.5± 5.1	37	5	63.2±9.3
MICROS	1,090	604	49	40.5±14.6	18-83	168±9	25.1±4.4	33	4	68.2±11.6
Orcades	719	690	55	53.3±15.7	18-92	167±9	27.6±4.9	25	3	60.7±10.0
PREVEND	3,880	3,513	52	48.8±12.2	28-75	173±9	26.0±4.3	31	4	68.9±12.3
PROSPER	5,135	3,639	57	75.2±3.3	70-83	165±9	26.8± 4.2	64	10	66.6±11.6
RS1	4,396	3,043	63	67.4±8.4	55-99	167±9	26.2±3.5	53	9	70.5±11.8
RS2	1,806	1,406	59	64.1±7.4	55-95	168±9	27.2±4.0	59	10	69.6±11.0
RS3	2,077	1,800	59	55.8±5.5	45-97	171±9	27.6±4.6	46	NA	69.4±10.3
Sardinia	5,014	4,372	43	42.3±17.3	81.80	160±9	25.1±4.6	20	15	67.2±11.3
SHIP	3,548	2,978	53	48.1±15.8	20-81	169±9	27.0±4.8	50	6	72.0±12.7
Split	433	390	63	50.0±15.0	18-85	171±9	26.6±4.2	25	4	65.7±10.6
Twins UK	2,637	2,396	95	51.8±12.5	18-84	163±7	25.6±4.5	15	2	66.9±10.4
YFS	479	448	55	38.6±5.1	30-47	172±9	26.4±4.5	4	1	68.9±10.5
Replication										
SMART	5,676	4,721	39	54.0±12.9	17-82	172±9	26.7±4.5	30	20	66.0±13.0
PREVEND	3,677	3,295	54	48.5±12.5	28-75	173±9	26.0±4.1	14	1	69.3±10.4
LifeLines	5,264	5,247	61	50.1± 11.8	21-90	174±9	26.6± 4.2	27	4	68.2±11.4

Table S1 *Continued*

Cohort	QRS interval, ms (mean ± SD)	QRS interval, ms (range)	12 lead sum product, mm.ms (mean ± SD)	12 lead sum product, mm.ms (range)	Cornell voltage product, mm.ms (mean ± SD)	Cornell voltage product, mm.ms (range)	Sokolow-Lyon product, mm.ms (mean ± SD)	Sokolow-Lyon product, mm.ms (range)
Discovery								
AGES	90±11	60-120	12819±3569	4349-33784	1424±582	167-4421	1962±694	385-6278
ARIC	96±9	61-120	11700±3139	3357-32194	1073±488	27-4450	2019±648	44-6113
Bright	93±10	66-120	12585±3446	4731-27459	1736±560	199-4205	2159±715	570-5021
Cilento	106±9	80-120	6444±2485	1201-16281	407 ± 284	17-1344	1067±477	95-2897
CHS	88±10	62-120	11580±3167	4093-24295	1149±494	0-3062	1955±662	413-4927
ERF	97±10	68-120	13733±3692	4993-39250	1167±495	89-3853	2320±682	884-5289
FHS	87±9	60-120	NA	NA	937±465	60-3900	NA	NA
FVG	96±11	72-149	12615±3644	4851-32540	960±584	46-4896	2086±670	630-4900
Inchianti	88+13	52-120	NA	NA	1206±613	39-4531	1907±691	374-4718
KORA S4	90±8	64-120	12798±3275	5101-27061	1165±505	168-3371	2097±664	493-5292
KORA F3	92	60-120	NA	NA	NA	NA	NA	NA
Korcula	96±9	76-119	13807±3580	6478-25645	1509± 651	49-5848	2236±676	93-4557
LifeLines_R2	95±12	57-145	14177±3846	4600-29730	1067±546	68-3293	2027±670	298-4708
LifeLines_R3	94±12	57-136	13923 ±3636	4042-25849	884±470	28-2543	2023±625	146-4093

LOLIPOP_EW610	93±10	66-119	NA	NA	1236±552	102-4144	1983±700	420-4802
LOLIPOP_EW_P	95±10	70-118	NA	NA	1347±658	100-4264	2055±756	624-5280
LOLIPOP_EW_A	92±11	64-118	NA	NA	1312±604	46-3168	1902±783	440-4644
MESA	89±7	68-120	10255±2768	3450-25408	973±434	28-2910	1765±586	356-5308
MICROS	96±13	71-229	14030±4041	4509-27550	1211±629	17-3357	2210±732	681-5194
Orcades	90±11	60-120	12435±3356	5152-29700	1335±659	80-5530	2090±626	80-4655
PREVEND	96±11	56-120	14179±3735	5173-34860	1095±552	0-4020	2250±682	560-5973
PROSPER	93±11	62-120	1441±1475	192-54616	1819±689	0-6301	2204±810	0-6821
RS1	96±11	66-120	14072±3498	6043-29491	1322±528	43-4119	2336±703	642-5669
RS2	97±11	70-120	14218±3102	6860-28616	1313±483	54-3837	2315±679	986-5690
RS3	97±11	62-120	13852±3284	5368-26053	1187±510	48-3543	2288±682	481-6141
Sardinia	90±12	50-119	NA	NA	1132±593	55-5175	1981±798	483-6237
SMART								
SHIP	97±11	60-120	13872±3497	6058-30311	1167±572	50-4025	2134±656	546-6228
Split	97±17	70-120	12326±3754	536-27280	1442±706	58-5822	2124±676	92-4565
Twins UK	87±8	58-119	NA	NA	887±440	68-4508	1953±610	624-5900
YFS	93±10	74-120	14031±3850	7060-25857	1163±537	88-3534	2282±703	1002-4820
Replication								
SMART	94±10	60-120	18470±5612	6006-36678	1438±590	0-3549	2503±862	570-5599
PREVEND	98±11	41-132	14652±3427	5147-28241	1100±513	0-3439	2325±667	614-5570
LifeLines	94±13	60-135	13753±3594	4042-25586	876±463	0-2497	1983 ± 609	146-3962

2. Table S2. Summary of study genotyping methods in the genome-wide association cohorts

Cohort	Array	Genotype calling software	SNP call rate excl.	SNP MAF excl.	pHWE excl.	Imputation software	NCBI Build for imputation	GWAS statistical analysis	Related individuals (yes/no)	Familial adjustment method (if applicable)
AGES	Illumina CNV370	Bead Studio	<97%	<1%	<10e-6	MACH 1 v1.0.16	36	ProbABEL, R	No	NA
ARIC	Affymetrix 6.0	Birdseed	<95%	<1%	<10e-5	MACH 1 v1.0.16	36	ProbABEL	No	NA
Bright	Affymetrix 500K	CHIAMO	<95%	None	None	IMPUTE v.1	36	SNPTEST	No	NA
Cilento	370k Illumina	Genome Studio	<95%	5%	None	MACH v1.0.16.a	36.22	GenABEL, ProbABEL, R	Yes	Mmscore in ProbABEL
CHS	Illumina CNV370	Bead Studio	<97%	None	<10e-5	BIMBAM	36	R	No	NA
ERF	Illumina 318K, 370K, Affy 250K	Bead Studio	<98%	None	<10e-6	MACH v1.0.15	36	GenABEL, ProbABEL, R	Yes	Kinship package in R
FHS	Affymetrix 500K, 50K MIP	BRLMM	<97%	<1%	<1e-6	MACH v1.0.15	36.2	Linear mixed effect models	Yes	Linear mixed effect models
FVG	Illumina 370K	Bead Studio	90%	<5%	<0.05	MACH	36	GenABEL	Yes	Kinship
Inchianti	Illumina 550K	Beadstudio	< 97%	<1%	<10e-4	MACH	36	MERLIN	Yes	Pedigree,GC
KORA S4	Affymetrix 6.0	Birdseed	<93%	None	None	MACH v1.0.16	36	MACH2qtl	No	NA

KORA F3	Affymetrix 500K	BRLMM	<95%	<1%	<10e-5	MACH v1.0.10	35	ProbABEL, R	No	NA
Korcula	Illumina CNV370	Bead Studio	<98%	<1%	<10e-6	MACH v1.0.15	36	GeneABEL ProbABEL, R	Yes	Mmscore in ProbABEL
Lifelines_R2	Illumina HumanCytoSNP-12	Genome Studio	<95%	<1%	<10e-5	BEAGLE v3.1.0	36.23a	PLINK v1.07	No	NA
Lifelines_R3	Illumina HumanCytoSNP-12	Genome Studio	<95%	<1%	<10e-5	BEAGLE v3.1.0	36.23a	PLINK v1.07	No	NA
LOLIPOP_EW 610	Illumina 610	Bead Studio	<90%	<1%	<10e-6	MACH	36	Additive, using MACH2qtl	No	Top 10 PCs
LOLIPOP_EW _P	Perlegen custom	NA	<90%	<1%	<10e-6	MACH	35	Additive, using MACH2qtl	No	Top 10 PCs
LOLIPOP_EW _A	Affymetrix 500k	BRLMM	<90%	<1%	<10e-6	MACH	35	Additive, using MACH2qtl	No	Top 10 PCs
MESA	Affymetrix Genome-Wide Human SNP Array 6.0	Birdseed v2	<95%	0.5%	None	IMPUTE 2.1.0	36	R	No	NA
MICROS	Illumina CNV370,	Bead Studio	<98%	<1%	<10e-6	MACH v1.0.16	36	ProbABEL	Yes	Mmscore in

	Illumina HumHap300v2									ProbABEL
Orcades	Illumina CNV370 & Illumina HumHap300v2	Bead Studio	<98%	<1%	<10e-6	MACH 1.0 ML	36	GeneABEL, ProbABEL, R	Yes	Mmscore in ProbABEL
PREVEND	Illumina HumanCytoSNP-12	Genome Studio	<98%	1%	<10e-5	BEAGLE 3.2	36.23	PLINK v1.07	No	NA
PROSPER	Illumina 660k Quad	Genome Studio	<97.5 %	None	None	MACH	36.22	Linear regression with ProbABEL	No	NA
RS1	Illumina 550k	Bead Studio	<98%	<1%	<10e-6	MACH v1.0.15	36	MACH2qtl implemented in GRIMP[88]	No	NA
RS2	Illumina 550k / 610k Quad	Genome Studio	<98%	<1%	<10e-6	MACH v1.0.16	36	MACH2qtl implemented in GRIMP[88]	No	NA
RS3	Illumina 610k Quad	Genome Studio	<98%	<1%	<10e-6	Mach v1.0.16	36	MACH2qtl implemented in GRIMP[88]	No	NA
Sardinia	Affymetrix 500K, 10K and 6.0	BRLMM (500K and 10K) ,	<98%	<5% 500K <5%	<10e-6	MACH	36.22	Yes	Yes	Variance component (Merlin --

		Birdseed v2 (6.0)		10K <1% 6.0						fastassoc)
SMART SHIP	Affymetrix 6.0	Birdseed v2	None	None	None	IMPUTE v0.5.0	36	QUICKTEST v0.95	No	NA
Split	Illumina CNV370	Bead Studio	<98%	<1%	<10e-6	MACH v1.0.15	36	GeneABEL, ProbABEL, R	Yes	Mmscore in ProbABEL
Twins UK	Illumina Hap300 Duo, Hap 300, Hap 550, Hap610	Illuminus	<95%	<1%	<10e-6	IMPUTE v0.3.2	36	SNPTEST v1.1.4	Yes	Huber-White method for robust variance estimation in R
YFS	Illumina 670k custom	PLINK 1.07	< 95%	< 1%	≤1e-6	MACH v1.0	36	PLINK 1.07 and ProbABEL 0.1-3	No	NA

3. Table S3. Comprehensive list of 79 locus-phenotype associations identified.

Locus sentinel: 1 = the discovery association for the locus (SNP with lowest *P*-value against any QRS phenotype); 0 = secondary phenotype(s) associated with the locus at $P<1\times10^{-8}$

Region	SNP	Coded Allele	Non-coded allele	Beta(SE)	N	P	Trait	Sentinel SNP	After additional genotyping	Novel
1p36.12	rs2849028	A	G	154.42(25.63)	44813	1.69E-09	leadsum	1	0	Y
1p32.3	rs17391905	G	T	-1.27(0.19)	55934	1.07E-11	duration	1	0	N
1p32.3	rs17106459	C	T	-416.16(72.03)	44625	7.60E-09	leadsum	0	0	Y
1p31.3	rs2103883	G	A	-27.82(3.40)	47624	2.87E-16	cornell	0	1	Y
1p31.3	rs2207790	A	G	-0.55(0.06)	50473	6.71E-19	duration	1	0	N
1p13.1	rs12039739	T	C	-0.41(0.07)	54447	6.22E-10	duration	1	1	N
1q22	rs2274317	T	C	170.79(25.42)	40366	1.82E-11	leadsum	1	0	Y
1q23.3	rs12036340	A	G	148.75(22.78)	62073	6.61E-11	leadsum	1	0	Y
1q32.1	rs10920184	T	C	-18.75(3.21)	56649	5.01E-09	cornell	1	0	Y
1q32.1	rs4288653	A	T	195.52(31.16)	31716	3.50E-10	leadsum	1	0	Y
2p23.3	rs6710065	T	C	-18.66(3.17)	56524	4.12E-09	cornell	1	1	Y
2p22.2	rs2216101	T	C	21.96(3.58)	44793	8.80E-10	cornell	0	0	Y
2p22.2	rs3770770	T	C	0.49(0.07)	54594	4.95E-11	duration	1	0	N
2q31.2	rs3731754	G	C	-44.14(5.83)	51650	3.81E-14	sokolow	0	1	Y

2q31.2	rs3816849	C	T	174.07(22.29)	45961	5.77E-15	leadsum	1	0	Y
3p22.2	rs6781009	C	T	171.08(26.07)	43637	5.30E-11	leadsum	0	0	Y
3p22.2	rs6801957	T	C	0.77(0.06)	58237	6.90E-40	duration	1	0	N
3p21.1	rs4687718	A	G	-0.57(0.09)	54188	6.70E-10	duration	1	0	N
3p14.1	rs2242285	A	G	0.34(0.06)	58414	5.65E-09	duration	1	0	N
3p14.1	rs13314892	A	G	-154.68(26.94)	45072	9.44E-09	leadsum	1	0	Y
3q27.2	rs10937226	A	G	-26.17(4.54)	52737	8.21E-09	sokolow	0	0	Y
3q27.2	rs10937226	A	G	-192.61(23.86)	43137	6.82E-16	leadsum	1	0	Y
4p15.31	rs1344852	C	G	0.46(0.08)	65737	1.45E-09	duration	1	1	Y
5q33.2	rs13165478	A	G	-0.59(0.07)	44824	8.06E-19	duration	0	0	N
5q33.2	rs13185595	A	G	-38.40(3.77)	39311	2.10E-24	cornell	1	0	Y
6p21.31	rs1321311	A	C	0.84(0.07)	57398	1.03E-37	duration	1	0	N
6p21.31	rs9462210	A	G	21.73(3.55)	59163	9.64E-10	cornell	0	1	Y
6p21.1	rs1015150	T	C	-128.92(22.41)	45364	8.72E-09	leadsum	0	0	Y
6p21.1	rs1015150	T	C	-26.28(4.33)	54075	1.28E-09	sokolow	1	0	Y
6q22.31	rs11153730	T	C	-0.63(0.06)	58553	7.44E-29	duration	1	0	N
6q22.31	rs11153730	T	C	-25.96(4.29)	52661	1.50E-09	sokolow	0	0	Y
7p14.3	rs1419856	G	A	0.67(0.08)	59752	6.67E-18	duration	1	1	N
7p12.3	rs6968945	C	T	0.34(0.06)	58002	5.14E-09	duration	1	0	N
7q31.2	rs11773845	C	A	0.36(0.06)	59758	7.50E-10	duration	1	1	Y
8q24.13	rs4367519	T	C	-73.76(11.18)	52255	4.15E-11	sokolow	1	0	Y
8q24.13	rs10105974	G	T	-151.37(23.15)	45767	6.25E-11	leadsum	1	1	Y
10q21.1	rs1194743	T	C	0.44(0.07)	50270	5.87E-09	duration	0	0	N

10q21.1	rs1733724	A	G	32.94(4.30)	41414	1.75E-14	cornell	1	0	Y
10q21.3	rs12414364	C	G	178.28(27.70)	43842	1.22E-10	leadsum	1	1	Y
10q21.3	rs10509289	G	C	-229.76(35.46)	46025	9.16E-11	leadsum	1	0	Y
10q22.2	rs7099599	T	C	226.10(31.35)	46281	5.51E-13	leadsum	1	0	Y
10q22.2	rs4114992	G	A	39.97(5.98)	55146	2.39E-11	sokolow	0	1	Y
10q25.2	rs7918405	A	G	0.50(0.07)	58206	1.05E-14	duration	1	0	N
10q25.2	rs7918405	A	G	35.27(5.00)	52273	1.68E-12	sokolow	0	0	Y
11p11.2	rs2269434	C	T	22.80(3.70)	46165	7.38E-10	cornell	1	0	Y
11q12.2	rs174577	A	C	-0.38(0.05)	70694	4.79E-12	duration	1	0	Y
12q13.13	rs736825	G	C	-21.63(3.32)	54078	7.20E-11	cornell	1	0	Y
12q13.3	rs2958153	A	G	-22.09(3.56)	54115	5.19E-10	cornell	0	0	Y
12q13.3	rs2926743	A	G	-30.54(5.00)	50297	9.72E-10	sokolow	0	0	Y
12q13.3	rs2926743	A	G	-275.14(26.01)	41454	3.74E-26	leadsum	1	0	Y
12q24.21	rs883079	C	T	0.52(0.06)	59255	4.58E-16	duration	0	1	N
12q24.21	rs1896312	C	T	-41.31(4.89)	50175	3.11E-17	sokolow	0	0	Y
12q24.21	rs2891537	T	G	24.27(3.87)	46812	3.53E-10	cornell	0	0	Y
12q24.21	rs7132327	C	T	-217.11(25.40)	45655	1.27E-17	leadsum	1	0	Y
13q14.13	rs1408224	A	G	-139.00(20.14)	65242	5.12E-12	leadsum	1	1	Y
13q22.1	rs728926	T	C	-0.40(0.06)	57432	5.60E-11	duration	1	1	N
14q24.2	rs12880291	T	G	-0.49(0.06)	59888	4.41E-14	duration	1	0	N
15q25.3	rs7183401	T	G	241.73(22.61)	44287	1.10E-26	leadsum	1	0	Y
15q25.3	rs6496452	T	A	22.17(3.40)	47633	7.36E-11	cornell	0	0	Y
15q25.3	rs3803405	A	G	-34.61(5.46)	40688	2.27E-10	sokolow	0	0	Y

15q26.3	rs8038015	C	T	-19.89(3.15)	58630	2.89E-10	cornell	0	0	Y
15q26.3	rs8038015	C	T	-212.76(22.61)	46463	4.93E-21	leadsum	1	0	Y
15q26.3	rs6598541	A	G	-34.52(4.59)	51407	5.39E-14	sokolow	0	0	Y
16q23.3	rs6565060	G	A	345.81(48.07)	31789	6.30E-13	leadsum	1	0	Y
17q11.2	rs7211246	A	G	-107.30(18.62)	65092	8.28E-09	leadsum	1	1	Y
17q21.31	rs1635291	G	A	-32.22(4.97)	54480	8.78E-11	sokolow	0	0	Y
17q21.31	rs242562	A	G	191.41(24.92)	40609	1.57E-14	leadsum	1	0	Y
17q21.32	rs17608766	C	T	0.52(0.09)	49724	8.98E-09	duration	0	1	N
17q24.2	rs12940610	A	G	161.75(24.09)	38879	1.88E-11	leadsum	0	0	Y
17q24.2	rs9910355	A	C	0.41(0.06)	53713	1.14E-11	duration	0	0	N
17q24.2	rs9912468	G	C	32.05(4.60)	48740	3.11E-12	sokolow	1	0	Y
18q12.1	rs617759	T	G	150.45(24.26)	42995	5.63E-10	leadsum	1	1	Y
18q12.2	rs879568	C	G	-0.34(0.06)	68500	1.88E-09	duration	1	0	Y
18q12.3	rs10853525	T	C	0.46(0.06)	56180	1.41E-14	duration	1	0	N
20p12.3	rs3929778	T	C	-21.42(3.33)	70143	1.22E-10	cornell	1	0	Y
20q11.22	rs2025096	A	G	-21.82(3.23)	73680	1.33E-11	cornell	1	0	Y
20q11.22	rs2025096	A	G	-0.37(0.06)	75678	4.08E-09	duration	0	1	Y
21q21.1	rs7283707	A	G	188.13(27.91)	65800	1.58E-11	leadsum	1	1	Y
21q21.3	rs13047360	G	A	0.47(0.07)	59742	4.02E-10	duration	1	0	Y

4. Table S4. Genome-wide association and replication test results for the 52 sentinel SNPs

ONLINE XLS File

5. Table S5. Full lists of the SNPs associated with phenotype at $P<10^{-6}$

ONLINE XLS File

6. Table S6. SNPs previously reported to be associated with the electrocardiographic traits

Highlighted in green are genome wide significant associations ($P<1\times10^{-8}$).

CHR	POS	SNP	AF	Previous Trait	Previous P-value	cornell P-value	leadsum P-value	sokolow P-value	Duration P-value	Reference
1p13.1	116112490	rs4074536	0.29	QRS	2.36E-08	1.33E-01	3.83E-01	1.13E-01	7.15E-07	Sotoodehnia[89]
1p13.2	112238867	rs2798334	0.29	P	7.98E-11	3.72E-01	9.63E-01	2.63E-01	4.44E-02	Verweij[90]
1p31.3	61646265	rs9436640	0.46	QRS	4.57E-18	3.75E-15	7.61E-01	9.00E-03	2.90E-18	Sotoodehnia[89]
1p32.3	51318728	rs17391905	0.05	QRS	3.26E-10	7.70E-05	1.64E-08	1.98E-03	1.07E-11	Sotoodehnia[89]
1p36	23583062	rs2298632	0.50	QT	1.00E-14	3.70E-04	7.53E-09	3.51E-06	1.66E-05	Arkin[91]
1p36.31	6201957	rs846111	0.28	QT	1.00E-16	5.36E-03	4.09E-01	1.14E-01	9.41E-01	Newton-Cheh[92]
1q23.3	160300514	rs12143842	0.26	QT	2.00E-78	2.73E-01	3.02E-08	1.74E-05	1.20E-03	Newton-Cheh[92]
1q23.3	160379534	rs16857031	0.14	QT	1.00E-34	2.66E-01	5.16E-04	6.84E-03	3.49E-03	Newton-Cheh[92]
1q23.3	160399741	rs12029454	0.15	QT	3.00E-45	1.22E-01	1.63E-04	3.01E-04	2.26E-03	Newton-Cheh[92]
1q32.2	206007476	rs11118555	0.12	HR	3.88E-26	7.96E-01	8.26E-01	8.85E-01	7.14E-01	Den Hoed[93]
1q32.2	206195345	rs2745967	0.37	RR	3.20E-08	4.58E-01	4.26E-01	3.25E-01	7.03E-01	Eijgelsheim[94]
2p14	66625504	rs11897119	0.39	PR	4.62E-11	7.42E-01	3.53E-03	8.52E-04	1.43E-01	Pfeufer [95]
2p22.2	36527059	rs7562790	0.40	QRS	8.22E-08	1.31E-06	1.35E-04	8.78E-02	2.65E-10	Sotoodehnia[89]
2p22.2	37101519	rs17020136	0.21	QRS	1.90E-08	2.21E-02	2.48E-03	1.98E-02	1.14E-09	Sotoodehnia[89]
2q31.1	174450854	rs938291	0.39	QT	6.00E-10	7.18E-01	8.52E-03	5.95E-02	1.82E-01	Arkin[91]
2q31.2	179398101	rs7561149	0.42	QT	7.00E-09	3.07E-01	1.02E-05	3.39E-02	5.42E-02	Arkin[91]
2q31.2	179429291	rs17362588	0.11	HR	3.57E-26	1.13E-02	2.84E-04	1.44E-02	1.10E-05	Den Hoed[93]

2q32.1	188041309	rs4140885	0.32	HR	4.72E-08	1.65E-01	2.13E-01	2.08E-01	2.27E-01	Den Hoed[93]
2q33	200868944	rs295140	0.42	QT	2.00E-11	7.02E-01	4.85E-01	2.67E-01	7.15E-02	Arkin[91]
2q37.1	231979528	rs13030174	0.27	HR	1.04E-10	1.15E-01	3.36E-01	5.71E-01	6.42E-03	Den Hoed[93]
3p14.1	66514292	rs2242285	0.42	QRS	1.09E-08	2.70E-01	5.82E-01	1.70E-02	5.68E-09	Sotoodehnia[89]
3p21	47519007	rs17784882	0.40	QT	3.00E-08	7.38E-01	8.44E-01	2.10E-01	7.50E-02	Arkin[91]
3p21.1	53257343	rs4687718	0.14	QRS	6.25E-08	6.46E-02	9.46E-03	5.66E-01	6.74E-10	Sotoodehnia[89]
3p22.2	38534753	rs2051211	0.26	QRS	1.57E-08	8.57E-04	4.14E-01	5.91E-01	1.36E-12	Sotoodehnia[89]
3p22.2	38552366	rs10865879	0.25	QRS	1.67E-24	7.95E-07	4.67E-06	1.84E-04	3.83E-27	Sotoodehnia[89]
3p22.2	38568397	rs12053903	0.34	QT	1.00E-14	7.99E-06	1.02E-08	4.21E-06	5.41E-31	Newton-Cheh[92]
3p22.2	38608927	rs11708996	0.15	PR	6.00E-26	1.97E-04	3.32E-03	6.13E-02	5.77E-23	Pfeufer [95]
3p22.2	38608927	rs11708996	0.16	QRS	1.26E-18	1.97E-04	3.32E-03	6.13E-02	5.77E-23	Sotoodehnia[89]
3p22.2	38632903	rs11710077	0.21	QRS	5.74E-22	8.38E-05	2.25E-03	1.34E-02	4.11E-26	Sotoodehnia[89]
3p22.2	38694939	rs9851724	0.33	QRS	1.91E-20	5.74E-04	8.13E-03	7.23E-01	6.97E-25	Sotoodehnia[89]
3p22.2	38741679	rs6795970	0.36	PR	9.50E-59	1.85E-03	2.18E-03	6.04E-02	8.87E-39	Holm[96]
3p22.2	38741679	rs6795970	0.36	QRS	3.50E-09	1.85E-03	2.18E-03	6.04E-02	8.87E-39	Holm[96]
3p22.2	38742319	rs6801957	0.41	QRS	1.10E-28	1.58E-03	4.04E-03	8.45E-02	7.09E-40	Sotoodehnia[89]
3p22.2	38749836	rs6800541	0.40	PR	2.10E-74	1.65E-03	1.81E-03	7.56E-02	1.43E-38	Pfeufer [95]
3q26.31	173267862	rs9647379	0.4	HR	1.17E-09	8.88E-02	1.07E-02	5.97E-06	6.08E-01	Den Hoed[93]
3q26.33	180655673	rs7612445	0.18	HR	1.86E-14	8.12E-01	8.69E-01	4.45E-01	2.93E-01	Den Hoed[93]
4q13	72357080	rs2363719	0.11	QT	8.00E-10	4.43E-01	3.76E-01	9.28E-01	6.52E-02	Arkin[91]
4q21.23	86860173	rs7692808	0.31	PR	5.99E-20	6.75E-01	1.42E-01	7.52E-02	4.06E-01	Pfeufer [95]
4q21.23	86870488	rs7660702	0.26	PR	2.50E-17	6.76E-01	1.41E-01	7.40E-02	4.66E-01	Holm[96]
4q22	95245457	rs3857067	0.46	QT	1.00E-09	6.18E-02	1.13E-01	8.81E-02	3.21E-01	Arkin[91]

5q31	137601624	rs10040989	0.13	QT	5.00E-11	8.43E-01	8.28E-01	1.22E-01	7.30E-01	Arkin[91]
5q33.2	153849233	rs13165478	0.36	QRS	7.36E-14	3.73E-24	9.52E-03	4.47E-02	8.16E-19	Sotoodehnia[89]
5q35.1	172412942	rs251253	0.40	PR	9.45E-13	4.62E-01	3.74E-01	1.91E-02	1.35E-02	Pfeufer [95]
5q35.1	172596769	rs6882776	0.32	HR	2.29E-12	4.62E-04	6.38E-01	1.04E-02	3.11E-05	Den Hoed[93]
6p21.2	36730878	rs1321311	0.21	QRS	2.70E-10	1.39E-09	2.45E-01	7.51E-01	1.05E-37	Holm[96]
6p21.2	36731357	rs9470361	0.25	QRS	3.00E-27	2.40E-09	3.66E-01	9.87E-01	1.52E-37	Sotoodehnia[89]
6p22	16402701	rs7765828	0.40	QT	3.00E-10	6.74E-02	8.42E-01	4.94E-01	3.91E-01	Arkin[91]
6q22.31	118680754	rs281868	0.50	RR	1.50E-10	2.66E-01	2.66E-02	7.45E-06	2.15E-27	Eijgelsheim[94]
6q22.31	118774215	rs11153730	0.51	HR	7.55E-21	3.87E-01	5.52E-03	1.46E-09	7.58E-29	Den Hoed[93]
6q22.31	118774215	rs11153730	0.51	QRS	1.26E-18	3.87E-01	5.52E-03	1.46E-09	7.58E-29	Sotoodehnia[89]
6q22.31	119100325	rs11756438	0.47	QT	5.00E-22	3.55E-01	1.63E-02	1.41E-08	2.01E-23	Newton-Cheh[92]
6q22.31	121790241	rs11154022	0.33	RR	3.50E-08	8.89E-01	4.08E-01	3.05E-01	5.41E-01	Eijgelsheim[94]
6q22.31	122173184	rs1015451	0.1	HR	1.14E-33	7.93E-01	8.50E-01	6.48E-01	1.58E-04	Den Hoed[93]
6q22.31	122187733	rs9398652	0.10	RR	7.70E-16	6.43E-01	9.03E-01	7.46E-01	2.03E-05	Eijgelsheim[94]
7p12.3	46586670	rs7784776	0.43	QRS	1.28E-08	2.30E-04	2.02E-01	8.48E-01	4.63E-07	Sotoodehnia[89]
7p14.2	35271831	rs1362212	0.18	QRS	1.12E-13	1.35E-01	7.72E-01	1.79E-01	2.63E-14	Sotoodehnia[89]
7q21.3	93387532	rs180242	0.33	HR	6.78E-12	4.07E-01	6.26E-01	8.79E-01	4.72E-01	Den Hoed[93]
7q22.1	100291144	rs314370	0.19	RR	2.30E-10	6.47E-01	6.72E-01	4.96E-03	8.67E-01	Eijgelsheim[94]
7q22.1	100324690	rs12666989	0.18	RR	9.40E-09	6.59E-01	5.10E-01	2.73E-02	6.29E-01	Eijgelsheim[94]
7q22.1	100335067	rs13245899	0.2	HR	7.67E-27	5.74E-01	7.40E-01	1.06E-02	7.27E-01	Den Hoed[93]
7q31	115987328	rs9920	0.09	QT	3.00E-08	4.67E-03	1.04E-01	1.47E-01	3.04E-07	Arkin[91]
7q31.2	115973477	rs3807989	0.40	PR	3.66E-28	1.36E-03	6.41E-02	1.30E-02	2.51E-09	Pfeufer [95]
7q31.2	115973477	rs3807989	0.40	PR	7.40E-13	1.36E-03	6.41E-02	1.30E-02	2.51E-09	Holm[96]

7q33	136293174	rs2350782	0.12	HR	1.26E-12	7.39E-01	8.83E-01	8.89E-01	7.03E-01	Den Hoed[93]
7q36.1	150253095	rs2968864	0.25	QT	8.00E-16	3.78E-01	2.75E-01	5.71E-01	3.59E-02	Newton-Cheh[92]
7q36.1	150268796	rs4725982	0.22	QT	5.00E-16	4.26E-02	6.41E-02	4.50E-01	4.62E-01	Newton-Cheh[92]
8q13	71351896	rs16936870	0.10	QT	1.00E-09	2.24E-01	2.26E-01	4.52E-01	1.13E-02	Arkin[91]
8q22.1	98919506	rs11779860	0.47	QT	2.00E-10	8.81E-01	7.74E-01	7.49E-01	1.04E-02	Arkin[91]
8q22.3	104002021	rs1961102	0.33	QT	3.00E-09	3.08E-01	3.11E-01	1.27E-01	2.19E-01	Arkin[91]
10q21.1	53893983	rs1733724	0.25	QRS	3.05E-08	1.86E-14	3.44E-06	4.45E-01	3.62E-08	Sotoodehnia[89]
10q24	104039996	rs2485376	0.39	QT	3.00E-08	5.11E-03	3.20E-01	9.19E-01	7.46E-01	Arkin[91]
10q25.2	114469252	rs7342028	0.27	QRS	4.95E-10	3.77E-03	2.76E-01	1.41E-07	1.12E-11	Sotoodehnia[89]
11p15.5	2441379	rs2074238	0.06	QT	3.00E-17	2.01E-01	2.29E-01	7.03E-01	4.15E-01	Newton-Cheh[92]
11p15.5	2458895	rs12576239	0.13	QT	1.00E-15	2.85E-03	1.30E-01	3.39E-01	4.15E-02	Newton-Cheh[92]
11q12	61366326	rs174583	0.34	QT	8.00E-11	1.44E-03	1.41E-04	5.79E-03	2.30E-07	Arkin[91]
11q12.2	61327359	rs174547	0.33	RR	8.20E-10	3.11E-03	2.74E-04	3.84E-03	2.13E-07	Eijgelsheim[94]
11q12.2	61327958	rs174549	0.31	HR	1.38E-22	5.64E-04	4.42E-05	2.76E-03	1.21E-06	Den Hoed[93]
11q12.2	61361390	rs174577	0.33	PRseg	7.62E-13	1.85E-03	1.90E-04	4.15E-03	4.28E-11	Verweij[90]
11q13.5	75587267	rs4944092	0.32	PR	3.22E-08	2.39E-01	8.32E-01	2.42E-01	3.75E-01	Pfeufer [95]
12p11.1	33468257	rs7980799	0.4	HR	6.22E-24	6.86E-02	7.28E-02	2.39E-02	8.55E-01	Den Hoed[93]
12p12.1	24662145	rs17287293	0.15	HR	3.07E-20	4.39E-02	9.92E-01	9.37E-01	1.27E-02	Den Hoed[93]
12p12.1	24662145	rs17287293	0.15	RR	5.70E-11	4.39E-02	9.92E-01	9.37E-01	1.27E-02	Eijgelsheim[94]
12p12.1	24679606	rs11047543	0.15	PR	3.34E-13	4.99E-02	9.21E-01	8.67E-01	1.56E-02	Pfeufer [95]
12q12	37392998	rs826838	0.44	HR	3.73E-09	3.24E-01	9.43E-02	1.29E-03	9.08E-01	Den Hoed[93]
12q23.3	105673552	rs2067615	0.49	HR	1.58E-09	9.25E-01	1.47E-01	1.86E-01	7.75E-01	Den Hoed[93]
12q24	109207586	rs3026445	0.36	QT	3.00E-12	4.72E-01	5.26E-03	1.12E-01	3.88E-04	Arkin[91]

12q24.21	113277623	rs883079	0.29	QRS	1.33E-10	8.88E-01	8.29E-06	1.09E-02	4.63E-16	Sotoodehnia[89]
12q24.21	113279826	rs3825214	0.22	PR	3.30E-12	6.95E-01	1.22E-02	4.87E-02	2.66E-12	Holm[96]
12q24.21	113279826	rs3825214	0.22	QRS	3.00E-13	6.95E-01	1.22E-02	4.87E-02	2.66E-12	Holm[96]
12q24.21	113830807	rs1896312	0.28	PR	3.13E-17	8.26E-07	3.04E-17	2.95E-17	4.99E-11	Pfeufer [95]
12q24.21	113866123	rs10850409	0.27	QRS	3.06E-10	1.09E-07	1.30E-17	6.38E-17	4.96E-13	Sotoodehnia[89]
13q12.11	20198909	rs2253017	0.15	PRseg	2.20E-08	7.59E-02	7.84E-01	7.46E-01	9.50E-02	Verweij[90]
13q12.11	21029897	rs2798269	0.40	PRseg	3.22E-10	8.64E-01	9.44E-02	1.85E-01	2.93E-01	Verweij[90]
13q14.13	46136718	rs9590974	0.35	PRseg	2.00E-08	2.07E-03	4.75E-06	2.04E-03	1.79E-01	Verweij[90]
13q22	73411123	rs728926	0.36	QT	2.00E-08	4.32E-01	4.28E-01	6.92E-01	5.60E-11	Arkin[91]
13q22.1	73418187	rs1886512	0.37	QRS	1.27E-08	3.10E-01	9.86E-01	2.34E-01	5.84E-10	Sotoodehnia[89]
14q11.2	22931651	rs365990	0.34	HR	9.40E-11	3.35E-01	4.13E-01	1.58E-01	9.56E-01	Holm[96]
14q11.2	22931651	rs365990	0.35	HR	5.39E-45	3.35E-01	4.13E-01	1.58E-01	9.56E-01	Den Hoed[93]
14q11.2	22931651	rs365990	0.37	RR	5.40E-14	3.35E-01	4.13E-01	1.58E-01	9.56E-01	Eijgelsheim[94]
14q11.2	22935725	rs452036	0.36	RR	8.10E-15	6.10E-01	4.74E-01	2.57E-01	7.44E-01	Eijgelsheim[94]
14q11.2	23046850	rs223116	0.24	RR	1.10E-08	2.06E-01	3.53E-01	2.74E-01	3.44E-01	Eijgelsheim[94]
14q24.2	71127108	rs11848785	0.27	QRS	1.04E-10	4.93E-04	4.26E-03	1.99E-03	5.57E-14	Sotoodehnia[89]
14q31.3	84879664	rs17796783	0.28	HR	2.69E-13	1.88E-01	9.06E-01	6.33E-01	7.78E-01	Den Hoed[93]
14q32	102044752	rs2273905	0.35	QT	4.00E-11	7.38E-01	7.51E-02	6.95E-01	3.87E-01	Arkin[91]
15q21	48632310	rs3105593	0.45	QT	3.00E-12	7.38E-01	8.09E-03	8.89E-02	6.71E-02	Arkin[91]
15q24.1	71452559	rs4489968	0.16	HR	3.82E-20	2.82E-01	7.62E-01	2.97E-02	2.29E-01	Den Hoed[93]
16p13.12	14302933	rs246185	0.34	QT	3.00E-13	1.52E-01	8.83E-01	2.43E-01	6.56E-03	Arkin[91]
16p13.13	11599254	rs8049607	0.51	QT	5.00E-15	7.58E-02	6.13E-05	4.74E-06	3.86E-01	Newton-Cheh[92]
16p13.3	3813643	rs1296720	0.20	QT	4.00E-10	4.34E-01	5.99E-02	1.51E-01	1.04E-01	Arkin[91]

16q21	57124739	rs37062	0.24	QT	3.00E-25	1.03E-02	5.56E-02	2.40E-01	9.23E-01	Newton-Cheh[92]
17q12	30348495	rs2074518	0.54	QT	6.00E-12	1.56E-01	4.81E-01	7.54E-01	1.15E-01	Newton-Cheh[92]
17q21.32	42368270	rs17608766	0.16	QRS	4.75E-10	2.30E-08	4.64E-05	4.49E-01	9.03E-09	Sotoodehnia[89]
17q24	61734255	rs9892651	0.43	QT	3.00E-14	3.71E-03	2.14E-11	1.72E-11	2.07E-11	Arkin[91]
17q24.2	61748819	rs9912468	0.43	QRS	1.06E-08	2.66E-03	4.87E-11	3.01E-12	1.66E-11	Sotoodehnia[89]
18q12.3	40693884	rs991014	0.42	QRS	6.20E-10	2.82E-01	1.49E-02	3.89E-02	1.69E-14	Sotoodehnia[89]
20q11.23	36277452	rs6127471	0.46	HR	5.22E-29	7.12E-03	7.50E-01	3.30E-01	1.34E-01	Den Hoed[93]

7. Table S7. Phenotypic variance explained by sentinel SNPs

	GWA cohorts	Non-GWA cohorts	Combined
Sample size (n)	11,156	5,032	16,188
Model 1 - All Sentinel SNPs			
QRS-duration	0.050	0.049	0.050
12-lead sum	0.044	0.036	0.041
Sokolow-Lyon	0.029	0.024	0.027
Cornell	0.028	0.040	0.032
Model 2 - Phenotype specific SNPs			
QRS-duration	0.044	0.047	0.045
12-lead sum	0.033	0.022	0.030
Sokolow-Lyon	0.017	0.012	0.016
Cornell	0.012	0.018	0.014

8. Table S8. Potential secondary SNPs with independent effects on phenotype

Region	Phenotype	GWAS Lead SNP	GWAS P	Position	SNPs at P<10-8 in conditional analysis
2p22.2	QRS	rs3770770	4.946E-11	37046370	rs3770770 (2.11E-10, STRN), rs3770900 (3.95E-10, CRIM1)
3p22.2	QRS	rs6801957	6.904E-40	38742319	rs6801957 (4.41E-42, SCN10A), rs12631864 (2.59E-11, EXOG), rs6781009 (6.79E-25, SCN5A), rs10154914 (7.04E-19, SCN5A), rs9851724 (3.75E-23, SCN10A), rs6776034(4.80E-09, SCN10A)
5q33.2	Cornell	rs13185595	2.099E-24	153852363	rs13185595 (1.08E-28, HAND1), rs17116169 (8.81E-09, SAP30L)
7p14.3	QRS	rs1419856	6.669E-18	35273508	rs1419856 (2.49E-17, TBX20), rs340383 (1.31E-09, TBX20)
12q24.21	QRS	rs7132327	1.27E-17	113865454	rs7132327 (2.14E-13, TBX3), rs883079 (2.33E-16, TBX5)
12q24.21	Sokolow	rs1896312	3.106E-17	113830807	rs1896312 (1.11E-16, TBX3), rs11067246 (1.24E-16, TBX3)

9. Table S9. Directional consistency in African Americans and Asian Indians

In the African American sample, 35 of 51 available locus-phenotype associations had the same direction of effect as seen in the European sample (P=5.49 ×10^{-3}, one-way binomial test) and in the Indian Asian sample, 22 of 29 available locus-phenotype associations showed the same direction of effect (P=4.07×10^{-3}). Freq EW; allele frequency of European White (Discovery), Freq AA; Allele frequency African Americans, Freq AI; Allele frequency Asian Indians.

								Discovery meta-analysis (n=65,275)		**African Americans (n=3,603)**		**Asian Indians (n=4,619)**	
Region	**SNP**	**Sentinel SNP**	**Coded Allele**	**Non-coded allele**	**Freq EW**	**Freq AA**	**Freq AI**	**Direction of effect**	**P-value**	**Direction of effect**	**P-value**	**Direction of effect**	**P-value**
1p13.1	rs12039739	duration	T	C	0.29	0.27	0.26	-	6.51E-09	-	5.33E-01	+	2.84E-02
1p31.3	rs2207790	duration	A	G	0.47	0.25	0.55	-	6.71E-19	-	3.61E-01	-	1.47E-01
1p32.3	rs17391905	duration	G	T	0.04	0.07		-	1.07E-11	-	2.33E-01		
1p36.12	rs2849028	leadsum	A	G	0.26	0.73		+	1.69E-09	+	4.19E-01		
1q22	rs2274317	leadsum	T	C	0.32	0.70		+	1.82E-11	-	3.03E-01		
1q23.3	rs12036340	leadsum	A	G	0.76	0.82		+	1.49E-09	+	1.59E-01		
1q32.1	rs10920184	cornell	T	C	0.38	0.67	0.38	-	5.01E-09	-	3.66E-02	-	5.31E-02
1q32.1	rs4288653	leadsum	A	T	0.23	0.10		+	3.50E-10	+	3.03E-01		

2p22.2	rs3770770	duration	T	C	0.20	0.09	0.29	+	4.95E-11	-	1.38E-02	+	4.64E-01
2p23.3	rs6710065	cornell	T	C	0.42	0.38	0.32	-	4.12E-09	+	7.53E-02	-	2.58E-02
2q31.2	rs3816849	leadsum	C	T	0.46	0.29		+	5.77E-15	+	7.23E-02		
3p14.1	rs13314892	leadsum	A	G	0.77	0.81		-	8.93E-09	+	2.22E-01		
3p14.1	rs2242285	duration	A	G	0.42	0.28	0.37	+	5.65E-09	+	2.37E-01	-	1.75E-01
3p21.1	rs4687718	duration	A	G	0.13	0.53	0.15	-	6.70E-10	-	4.66E-01	+	4.53E-01
3p22.2	rs6801957	duration	T	C	0.42	0.16	0.38	+	6.90E-40	+	6.03E-02	+	3.72E-01
3q27.2	rs10937226	leadsum	A	G	0.35	0.44		-	6.82E-16	-	1.12E-02		
4p15.31	rs1344852	duration	C	G	0.85	0.86	0.84	+	1.21E-09	-	2.38E-01	-	6.25E-01
5q33.2	rs13185595	cornell	A	G	0.37	0.54	0.26	-	2.10E-24	-	3.81E-03	-	3.82E-03
6p21.1	rs1015150	sokolow	T	C	0.45	0.53	0.34	-	1.28E-09	-	2.17E-02	+	4.30E-01
6p21.31	rs1321311	duration	A	C	0.27	0.39	0.32	+	1.03E-37	-	7.77E-01	+	6.19E-01
6q22.31	rs11153730	duration	T	C	0.50	0.71	0.57	-	7.44E-29	-	3.51E-01	-	1.49E-01
7p12.3	rs6968945	duration	C	T	0.44	0.26	0.52	+	5.14E-09	-	4.82E-01	+	1.19E-01
7p14.3	rs1419856	duration	G	A	0.16	0.04		+	6.67E-18	-	5.24E-01		
7q31.2	rs11773845	duration	C	A	0.41	0.64	0.40	+	7.50E-10	+	3.50E-02	+	8.93E-02
8q24.13	rs10105974	leadsum	G	T	0.36	0.51		-	6.25E-11	+	1.42E-01		
8q24.13	rs4367519	sokolow	T	C	0.04	0.04	0.19	-	4.15E-11			-	2.66E-01
10q21.1	rs1733724	cornell	A	G	0.26	0.05	0.29	+	1.75E-14	+	5.34E-01	+	3.59E-01
10q21.3	rs10509289	leadsum	G	C	0.11	0.23		-	9.16E-11	-	6.42E-04		
10q21.3	rs12414364	leadsum	C	G	0.21	0.24		+	1.22E-10	+	1.28E-01		
10q22.2	rs7099599	leadsum	T	C	0.15	0.08		+	5.51E-13	-	7.50E-01		

10q25.2	rs7918405	duration	A	G	0.26	0.74		+	1.05E-14	-	6.42E-01	+	3.12E-02
11p11.2	rs2269434	cornell	C	T	0.32	0.55	0.45	+	7.38E-10	-	9.68E-01	+	4.46E-01
11q12.2	rs174577	duration	A	C	0.34	0.36	0.19	-	4.28E-11	-	3.33E-01	-	7.87E-01
12q13.13	rs736825	cornell	G	C	0.36	0.26	0.33	-	7.20E-11	-	7.95E-02	-	3.73E-02
12q13.3	rs2926743	leadsum	A	G	0.27	0.10		-	3.74E-26	-	2.03E-01		
12q24.21	rs7132327	leadsum	C	T	0.27	0.23		-	1.27E-17	-	1.06E-03		
13q14.13	rs1408224	leadsum	A	G	0.69	0.59		-	3.60E-10	-	9.86E-03		
13q22.1	rs728926	duration	T	C	0.38	0.32	0.40	-	5.60E-11	+	6.91E-01	+	8.04E-01
14q24.2	rs12880291	duration	T	G	0.26	0.07	0.15	-	4.41E-14	-	4.39E-01	+	3.25E-01
15q25.3	rs7183401	leadsum	T	G	0.44	0.64		+	1.10E-26	+	2.72E-02		
15q26.3	rs8038015	leadsum	C	T	0.38	0.53		-	4.93E-21	-	4.75E-01		
16q23.3	rs6565060	leadsum	G	A	0.08	0.14		+	6.30E-13	+	6.35E-04		
17q11.2	rs7211246	leadsum	A	G	0.54	0.49		-	6.01E-09	+	1.03E-01		
17q21.31	rs242562	leadsum	A	G	0.38	0.30		+	1.57E-14	+	1.84E-01		
17q24.2	rs9912468	sokolow	G	C	0.44	0.37	0.39	+	3.11E-12	+	9.80E-01	+	4.58E-03
18q12.1	rs617759	leadsum	T	G	0.33	0.06		+	5.63E-10	+	4.17E-02		
18q12.2	rs879568	duration	C	G	0.33	0.53	0.44	-	8.45E-09	+	1.58E-01	-	3.43E-01
18q12.3	rs10853525	duration	T	C	0.42	0.24	0.25	+	1.41E-14	+	9.03E-01	+	1.37E-01
20p12.3	rs3929778	cornell	T	C	0.80	0.88	0.80	-	6.42E-09	+	8.10E-01	-	6.62E-01
20q11.22	rs2025096	cornell	A	G	0.21	0.17	0.25	-	4.51E-11	-	3.13E-01	-	1.04E-01
21q21.1	rs7283707	leadsum	A	G	0.13	0.46		+	3.76E-09	+	1.99E-01		
21q21.3	rs13047360	duration	G	A	0.18	0.07	0.21	+	4.02E-10	+	4.50E-01	+	6.23E-01

10. Table S10. Coding SNPs in LD with lead locus-phenotype SNPs.

Coding SNPs in transcribed genes in LD at r^2>0.8 (1000G; European ancestry) for all locus-phenotype associated lead SNPs (n=79). AF is frequency of allele in the default global population of 1000 Genome phase 1 genotype data from 1,094 worldwide individuals (May 2011 dataset). R^2 is LD between sentinel and non-synonymous SNP.

Sentinel SNP	CHR	BP (hg19)	AF	Non-syn SNP	CEU r^2	FIN r^2	GBR r^2	TSI r^2	Gene	Amino Acid change	Protein Position	leadSNP 52 Loci (1) or trait specific snp (0)
rs6801957	3	38767315	0.42	rs6795970	0.98	1.00	0.98	0.98	SCN10A	Val→Ala	1073	1
rs10937226	3	185302885	0.35	rs6762208	0.97	1.00	0.98	1.00	SENP2	Thr→Lys	301	1
rs11153730	6	118667522	0.50	rs3734381		0.80			CEP85L	Ser→Gly	137/140	1
rs4367519	8	124666429	0.04	rs72711231	0.85				KLHL38	Lys→Glu	508	1
rs4367519	8	124666429	0.04	rs16898691	1.00		1.00	0.87	KLHL38	Gly→Arg	394	1
rs7099599	10	75487081	0.15	rs34163229	0.88	0.82	0.87	0.88	SYNPO2L	Ser→Tyr	609/833	1
rs7099599	10	75487081	0.15	rs3812629	0.88	0.82	0.87	0.88	SYNPO2L	Pro→Leu	483/707	1
rs7099599	10	75487081	0.15	rs60632610	0.88	0.88	0.87	0.89	SYNPO2L	Gly→Ser	2	1
rs4114992	10	75566829	0.14	rs34163229	0.88	0.82	0.83	0.88	SYNPO2L	Ser→Tyr	609/833	0
rs4114992	10	75566829	0.14	rs3812629	0.88	0.82	0.83	0.88	SYNPO2L	Pro→Leu	483/707	0

rs4114992	10	75566829	0.14	rs60632610	0.88	0.88	0.83	0.89	SYNPO2L	Gly→Ser	2	0
rs2269434	11	47360412	0.32	rs2167079	0.82	0.82			ACP2	Arg→Gln	29	1
rs2958153	12	57081517	0.28	rs2958149	1.00	0.87	0.93	0.90	NACA	Leu→Pro	688	0
rs2958153	12	57081517	0.28	rs2926743	1.00	0.87	0.93	0.90	NACA	Phe→Ser	405	0
rs2926743	12	57114100	0.27	rs2958149	1.00	1.00	1.00	1.00	NACA	Leu→Pro	688	1
rs2926743	12	57114100	0.27	rs2926743	1.00	1.00	1.00	1.00	NACA	Phe→Ser	405	1
rs3803405	15	85383640	0.30	rs1051168	0.88	0.88	0.94	0.84	NMB	Pro→Thr	73	0
rs3803405	15	85383640	0.30	rs3803403	1.00	1.00	1.00	1.00	ALPK3	Thr→Ser	414	0
rs3803405	15	85383640	0.30	rs3803405	1.00	1.00	1.00	1.00	ALPK3	Gly→Glu	579	0
rs7211246	17	28485762	0.54	rs9897794	0.91	0.87	0.86		EFCAB5	Leu→Val	181/237	1
rs114860868 (rs1635291)	17	43751913	0.25	rs116966623	0.84	0.86	0.84	0.85	SPPL2C	Ser→Pro	224	0
rs114860868 (rs1635291)	17	43751913	0.25	rs117598307	0.84	0.86	0.84	0.85	SPPL2C	Ala→Thr	332	0
rs114860868 (rs1635291)	17	43751913	0.25	rs112235641	0.84	0.86	0.84	0.85	SPPL2C	Arg→Pro	461	0
rs114860868 (rs1635291)	17	43751913	0.25	rs112636016	0.84	0.86	0.84	0.85	SPPL2C	Ile→Val	471	0
rs114860868 (rs1635291)	17	43751913	0.25	rs113834859	0.84	0.86	0.84	0.85	SPPL2C	Ser→Pro	601	0
rs114860868 (rs1635291)	17	43751913	0.25	rs111430241	0.84	0.86	0.84	0.85	SPPL2C	Gly→arg	620	0
rs114860868 (rs1635291)	17	43751913	0.25	rs112560719	0.84	0.86	0.81	0.85	SPPL2C	Pro→Arg	643	0
rs114860868 (rs1635291)	17	43751913	0.25	rs118160437	0.84	0.86	0.84	0.85	MAPT	Pro→Leu	202	0
rs114860868 (rs1635291)	17	43751913	0.25	rs118082626	0.84	0.86	0.84	0.85	MAPT	Asp→Asn	285	0
rs114860868 (rs1635291)	17	43751913	0.25	rs117070738	0.84	0.86	0.84	0.85	MAPT	Val→Ala	289	0
rs114860868 (rs1635291)	17	43751913	0.25	rs117495416	0.84	0.86	0.84	0.85	MAPT	Arg→Trp	370	0
rs114860868 (rs1635291)	17	43751913	0.25	rs117701706	0.84	0.86	0.84	0.85	MAPT	Ser→Pro	447	0

rs114860868 (rs1635291)	17	43751913	0.25	rs117086266	0.84	0.86	0.84	0.85	STH	Gln→Arg	7	0
rs114860868 (rs1635291)	17	43751913	0.25	rs116937503	0.84	0.86	0.84	0.85	KANSL1	Ile→Thr	1085	0
rs114860868 (rs1635291)	17	43751913	0.25	rs117312607	0.84	0.86	0.84	0.85	KANSL1	Ser→Pro	718	0
rs114860868 (rs1635291)	17	43751913	0.25	rs117648158	0.84	0.86	0.84		KANSL1	Arg→Ser	247	0
rs114860868 (rs1635291)	17	43751913	0.25	rs138137490		0.80			KANSL1	Asn→His	225	0
rs114860868 (rs1635291)	17	43751913	0.25	rs117830374		0.80			KANSL1	Lys→Thr	104	0
rs879568	18	34311659	0.33	rs2303510		0.90		0.88	FHOD3	Val→Ile	1151	1
rs2025096	20	33540000	0.21	rs3746435	0.86	0.91		0.97	MYH7B	Lys→Asn	1552	1
rs2025096	20	33540000	0.21	rs3746429		0.85			EDEM2	Ala→Thr	419/456	1

11. Table S11. Motif scan for transcription factor recognition sites within DHSs.

Motifs in bold are within a DHS in foetal heart.

Region	Sentinel SNP	#cell types DHS active in	DHS in fHeart	Genetic location	Motifs perturbed by SNP
1p36.12	rs2849028	32	Y	coding	no
1p32.3	rs17391905	0	N	intergenic	no
1p31.3	rs2207790	2	Y	intron	no
1p13.1	rs12039739	5	N	intergenic	no
1q22	rs2274317	82	N	coding	no
1q23.3	rs12036340	0	N	intergenic	no
1q32.1	rs10920184	5	Y	intron	no
1q32.1	rs4288653	80	Y	intron	No
2p23.3	rs6710065	3	N	intron	EHF,ELF3,ELK1,ELK3,ELK4,ERG,ETS1,ETV1,ETV2,ETV3,ETV4,ETV5,ETV6,FLI1,GABPA,PRDM4
2p22.2	rs3770770	74	Y	intron	CEBPA,CEBPB,CEBPD,CEBPE,CEBPG
2q31.2	rs3816849	4	Y	intron	ZEB1
3p22.2	rs6801957	30	Y	intron	No
3p21.1	rs4687718	5	N	intron	MSC
3p14.1	rs2242285	1	N	intron	MEF2A,ZNF713
3p14.1	rs13314892	3	Y	intron	MGA,TBX21
3q27.2	rs10937226	10	N	promoter	AHR,ARNT,ATF1,CREB1
4p15.31	rs1344852	3	N	intergenic	no
5q33.2	rs13185595	0	N	intergenic	no
6p21.31	rs1321311	23	N	intergenic	no
6p21.1	rs1015150	11	N	intron	ESRRG,PAX4
6q22.31	rs11153730	0	N	intergenic	no

7p14.3	rs1419856	1	Y	intergenic	FOXP1,POU5F1
7p12.3	rs6968945	1	N	intergenic	MECOM,SRF,TBX1
7q31.2	rs11773845	0	N	intron	no
8q24.13	rs4367519	28	Y	promoter	no
8q24.13	rs10105974	15	Y	intergenic	no
10q21.1	rs1733724	4	N	intergenic	no
10q21.3	rs12414364	0	N	intron	no
10q21.3	rs10509289	8	Y	intron	no
10q22.2	rs7099599	0	N	intergenic	no
10q25.2	rs7918405	3	Y	intron	no
11p11.2	rs2269434	2	Y	intron	ZBTB12,ZBTB6,ZNF524
11q12.2	rs174577	1	N	intron	GLI1,GLI2,GLI3,TP53
12q13.13	rs736825	4	N	promoter	EGR1,EGR2,EGR3,EWSR1,FLI1,KLF11,SP1,SP2,SP3,ZBTB7B,ZNF148,ZNF281,ZNF350,ZNF740
12q13.3	rs2926743	3	N	coding	ATF5,HOXA5,POU1F1,SPI1
12q24.21	rs7132327	1	N	intergenic	VDR
13q14.13	rs1408224	51	Y	intron	GLIS3,KLF11
13q22.1	rs728926	12	Y	intron	no
14q24.2	rs12880291	2	N	intergenic	no
15q25.3	rs7183401	33	Y	intron	CEBPA
15q26.3	rs8038015	2	N	intron	NANOG
16q23.3	rs6565060	3	Y	intron	GATA1,GATA2,GATA3,GATA4,GATA5,GATA6,RORA,TAL1
17q11.2	rs7211246	0	N	intron	no
17q21.31	rs242562	78	Y	intron	RELA
17q24.2	rs9912468	0	N	intron	no
18q12.1	rs617759	10	N	intergenic	ZFP161,ZNF423
18q12.2	rs879568	0	N	intron	no
18q12.3	rs10853525	4	N	intron	HOXA13,HOXC13
20p12.3	rs3929778	1	N	intergenic	NFAT5,STAT5A
20q11.22	rs2025096	11	Y	promoter	NR1I2,RXRA
21q21.1	rs7283707	45	Y	intron	REST
21q21.3	rs13047360	1	Y	intergenic	no

12. Table S12. Relationship between sentinel SNPs and cis-eQTLs

Relationships between sentinel SNPs from the GWAS with expression of cis genes (+/- 1 MB) in 4 unrelated studies: (1) Peripheral blood lymphocytes from 1,469 unrelated individuals from the UK and Netherlands (2), Left ventricle tissue from 313 individuals (3) Left ventricular tissue from 110 non-diseased human hearts (RNA-seq), and (4) peripheral blood lymphocytes from 2,116 individuals (RNA-seq). Genes identified as eQTLs based on: $P<1\times10^{-5}$ for association of sentinel SNP with transcript expression (Tx P1) and $r^2\geq0.8$ between Sentinel SNP and Transcript SNP (the SNP most closely associated with transcript). Tx P2: association of transcript SNP with expression; LD between sentinel and peak SNPs (r^2) calculated from the 1,469 individuals.

Band	Sentinel SNP	Position1	Primary Pheno	Gene	Tx P1	Transcript SNP	Position2	Distance	Tx P2	r2	leadSNP 52 Loci (1) or trait specific snp (0)	Source
2q31.2	rs3816849	179375335	leadsum	*TTN*	6.58E-17	rs3816849	1,79E+08	0	6.58E-17	1.00	1	4
3q27.2	rs10937226	186785579	leadsum	*SENP2*	3.90E-80	rs3087964	1,87E+08	45517	1.70E-80	0.98	1	1
6p21.31	rs1321311	36730878	duration	*CDKN1A*	1.10E-13	rs9470361	36731357	479	3.00E-16	0.92	1	1
8q24.13	rs10105974	125923696	leadsum	*MTSS1*	4.15E-18	rs7461129	1,26E+08	6859	9.17E-23	0.83	1	2
10q22.2	rs7099599	75157087	leadsum	*CAMK2G*	5.00E-08	rs4746145	75187993	30906	1.98E-08	0.97	1	4
10q22.2	rs4114992	75236835	sokolow	*CAMK2G*	4.98E-08	rs4746145	75187993	-48842	1.98E-08	0.97	0	4
11p11.2	rs2269434	47316988	cornell	*NR1H3*	3.10E-31	rs7395581	47202973	-114015	1.80E-38	0.81	1	1
11p11.2	rs2269434	47316988	cornell	*NR1H3*	6.57E-56	rs326222	47216244	-100744	8.61E-69	0.82	1	4
11q12.2	rs174577	61361390	duration	*FADS2*	7.26E-11	rs174548	61327924	-33466	1.48E-11	0.81	1	2
11q12.2	rs174577	61361390	duration	*TMEM258*	9.10E-17	rs174538	61316657	-44733	1.11E-17	0.86	1	4
12q13.3	rs2958153	55367784	cornell	*BAZ2A*	8.20E-16	rs941207	55309551	-58233	3.60E-17	0.91	0	1
12q13.3	rs2958153	55367784	cornell	*NACA*	7.35E-26	rs941207	55309551	-58233	2.38E-27	0.84	0	4

12q13.3	rs2926743	55400367	leadsum	*NACA*	3.78E-27	rs941207	55309551	-90816	2.38E-27	0.85	1	4
15q25.3	rs7183401	83172948	leadsum	*SCAND2*	1.00E-09	rs7169629	82992278	-180670	4.80E-13	0.80	1	1
15q25.3	rs7183401	83172948	leadsum	*ALPK3*	9.94E-18	rs1975277	83130562	-42386	2.90E-19	0.93	1	4
15q25.3	rs6496452	83173649	cornell	*ALPK3*	1.12E-17	rs1975277	83130562	-43087	2.90E-19	0.93	0	4
15q25.3	rs3803405	83184644	sokolow	*NMB*	2.10E-51	rs62021193	82971587	-213057	3.52E-63	0.83	0	4
17q11.2	rs7211246	25509888	leadsum	*EFCAB5*	2.71E-35	rs4294865	25229862	-280026	8.85E-43	0.83	1	4
17q21.31	rs1635291	41107696	sokolow	*LRRC37A2*	9.46E-38	rs2668624	41708649	600953	8.85E-49	0.84	0	2
17q21.31	rs1635291	41107696	sokolow	*LRRC37A4*	6.57E-13	rs2957297	41723989	616293	1.91E-15	0.89	0	3
17q21.31	rs1635291	41107696	sokolow	*ARL17A*	7.69E-06	rs7225002	41544850	437154	1.47E-08	0.87	0	3
17q21.31	rs1635291	41107696	sokolow	*LRRC37A*	5.32E-12	rs34097347	41305238	197542	5.16E-15	0.87	0	3
17q21.31	rs1635291	41107696	sokolow	*LOC644246*	2.45E-15	rs2696455	41639348	531652	9.91E-19	0.85	0	3
17q21.31	rs1635291	41107696	sokolow	*LRRC37A4P*	3.27E-310	rs111370985	41208507	100811	3.27E-310	0.85	0	4
17q21.31	rs1635291	41107696	sokolow	*CRHR1-IT1*	3.27E-310	rs60814418	41206410	98714	3.27E-310	0.85	0	4
17q21.31	rs1635291	41107696	sokolow	*DND1P1*	3.27E-310	rs55974014	41113233	5537	3.27E-310	0.85	0	4
17q21.31	rs1635291	41107696	sokolow	*RP11-707O23.5*	3.27E-310	rs55974014	41113233	5537	3.27E-310	0.85	0	4
17q21.31	rs242562	41382599	leadsum	*MAPT*	5.49E-24	rs242557	41375573	-7026	4.08E-24	0.90	1	2
17q24.2	rs9912468	61748819	sokolow	*PRKCA*	1.14E-41	rs11658550	61742145	-6674	8.80E-42	0.99	1	2
20q11.22	rs2025096	33003661	cornell	*EDEM2*	5.20E-20	rs3746429	33167268	163607	5.00E-22	0.82	1	1
20q11.22	rs2025096	33003661	cornell	*EDEM2*	2.97E-99	rs7353271	33191092	187431	3.31E-106	0.80	1	4

13. Table S13. Candidate genes identified by GRAIL using Pubmed 2006 or 2012 datasets.

P-values are corrected for multiple testing.

			GRAIL 2006		GRAIL 2012	
Region	**SNP**	**Position**	**Gene**	**P**	**Gene**	**P**
1p36.12	rs2849028	23561520	HTR1D	5.26E-01	HTR1D	8.32E-01
1p32.3	rs17391905	51318728	CDKN2C	2.30E-01	CDKN2C	2.07E-01
1p31.3	rs2207790	61670555	NFIA	6.64E-01	NFIA	5.66E-01
1p13.1	rs12039739	116134634	CASQ2	1.84E-03	CASQ2	8.03E-03
1q22	rs2274317	154713527	MEF2D	1.16E-02	C1orf61	1.56E-02
1q23.3	rs12036340	160282364	NOS1AP	2.02E-01	NOS1AP	7.40E-03
1q32.1	rs10920184	199605519	TNNT2	4.75E-06	TNNT2	6.92E-06
1q32.1	rs4288653	202532651	PLEKHA6	9.78E-01	PLEKHA6	7.82E-01
2p23.3	rs6710065	26930061	MAPRE3	6.54E-01	MAPRE3	6.95E-01
2p22.2	rs3770770	37046370	STRN	1.64E-01	STRN	4.75E-01
2q31.2	rs3816849	179375335	TTN	1.23E-02	TTN	1.92E-02
3p22.2	rs6801957	38742319	SCN10A	7.68E-01	SCN10A	6.46E-01
3p21.1	rs4687718	53257343	DCP1A	1.05E-01	DCP1A	9.63E-01
3p14.1	rs2242285	66514292	MAGI1	4.53E-01	MAGI1	2.55E-01
3p14.1	rs13314892	69877742	MITF	2.25E-03	MITF	1.81E-02
3q27.2	rs10937226	186785579	SENP2	1.43E-01	SENP2	4.88E-01
4p15.31	rs1344852	19793035	SLIT2	1.02E-01	SLIT2	1.73E-01
5q33.2	rs13185595	153852363	HAND1	3.41E-02	HAND1	1.90E-01
6p21.31	rs1321311	36730878	CDKN1A	6.15E-02	CDKN1A	1.00E-01
6p21.1	rs1015150	41767282	TFEB	2.91E-03	TFEB	2.38E-02
6q22.31	rs11153730	118774215	PLN	3.11E-04	PLN	8.27E-05
7p14.3	rs1419856	35273508	TBX20	1.32E-03	TBX20	2.83E-02
7p12.3	rs6968945	46607425	N/A	N/A	N/A	N/A
7q31.2	rs11773845	115978537	CAV1	4.03E-04	CAV1	7.34E-04
8q24.13	rs4367519	124735610	FBXO32	1.58E-02	C8ORFK36	7.08E-01
8q24.13	rs10105974	125923696	MTSS1	7.94E-01	MTSS1	4.72E-01
10q21.1	rs1733724	53893983	DKK1	1.23E-02	DKK1	2.66E-02

10q21.3	rs12414364	67674620	CTNNA3	1.74E-02	CTNNA3	1.80E-01
10q21.3	rs10509289	68951501	CTNNA3	1.74E-02	CTNNA3	1.80E-01
10q22.2	rs7099599	75157087	MYOZ1	2.56E-01	CAMK2G	7.02E-02
10q25.2	rs7918405	114495455	VTI1A	9.85E-01	VTI1A	9.81E-01
11p11.2	rs2269434	47316988	MYBPC3	1.05E-04	MYBPC3	5.97E-04
11q12.2	rs174577	61361390	FADS2	9.75E-01	FADS1	8.42E-01
12q13.13	rs736825	52703843	HOXC10	4.47E-01	HOXC9	8.28E-01
12q13.3	rs2926743	55400367	RBMS2	9.95E-01	RBMS2	9.99E-01
12q24.21	rs7132327	113865454	TBX3	1.00E-02	TBX3	6.01E-03
13q14.13	rs1408224	46113219	LRCH1	1.72E-01	LRCH1	1.04E-01
13q22.1	rs728926	73411123	KLF12	8.17E-01	KLF12	4.86E-01
14q24.2	rs12880291	70954320	SIPA1L1	1.67E-01	SIPA1L1	1.44E-01
15q25.3	rs7183401	83172948	ALPK3	6.07E-02	ALPK3	3.58E-05
15q26.3	rs8038015	97080797	IGF1R	8.36E-02	IGF1R	6.69E-02
16q23.3	rs6565060	81307552	CDH13	1.96E-02	CDH13	4.45E-02
17q11.2	rs7211246	25509888	SLC6A4	5.05E-01	SLC6A4	3.63E-01
17q21.31	rs242562	41382599	MAPT	2.70E-01	MAPT	1.25E-01
17q24.2	rs9912468	61748819	PRKCA	1.49E-01	PRKCA	1.21E-01
18q12.1	rs617759	30976867	MAPRE2	2.74E-01	MAPRE2	6.98E-01
18q12.2	rs879568	32565657	BRUNOL4	5.97E-01	FHOD3	7.35E-01
18q12.3	rs10853525	40690650	SETBP1	4.45E-01	SETBP1	4.87E-01
20p12.3	rs3929778	6408290	BMP2	1.09E-01	BMP2	4.98E-01
20q11.22	rs2025096	33003661	ITCH	6.14E-01	MYH7B	1.90E-01
21q21.1	rs7283707	16048865	USP25	9.24E-01	USP25	8.80E-01
21q21.3	rs13047360	27773451	N/A	N/A	N/A	N/A

14. Table S14. Canonical pathway analysis.

Canonical pathways analysis using the IPA software tool (IPA, Ingenuity Systems, CA, USA). The IPA Knowledge Base was used to explore the functional relationship between proteins encoded by the 67 candidate genes identified at the 52 loci associated with QRS traits. Genes were analysed for direct interactions only and networks were generated with a maximum size of 35 molecules.

Core Genes	Additional Genes	P-value	Top Functions
24genes: CAMK2G, CAV1, CDKN1A, CDKN2C, DKK1, FBXO32, HAND1, IGF1R, KLHL38, MAPT, MEF2D, MITF, MTSS1, MYBPC3, MYH7B, NACA, PLN, PRKCA, STRN, TBX3, TFEB, TNNT2, TNS3, TTN	Akt, Calmodulin, G protein alphai, Histone H3, Hsp90, Myosin, PI3K (complex), RNA polymerase II, SRC (family)	10^{-56}	Cardiovascular System Development and Function, Organ Morphology, Skeletal and Muscular System Development and Function
16 genes: ADAMTS5, DPYSL5, FADS2, HOXC4, HOXC5, KLF12, LRCH1, MAPRE2, NFIA, NR1H3, SEC24C, SENP2, SLIT2, TKT, USP25, VTI1A	APP, C18orf21, C1orf110, C1orf131, C9orf41, C9orf142, CHAC2, DDHD2, METTL8, MRPL35, PASD1, RABL3, RNASE11, SBNO1, SUMO2, SUN5, TBCC, UBC, ZNF720	10^{-33}	Nervous System Development and Function, Organ Morphology, Organismal Development
13 genes: ACP2, CASQ2, CEP85L, CTNNA3, EDEM2, GSS, LRIG1, MADD, NSRP1, SETBP1, SLC25A26, TBX20, TMEM258	ANKLE2, ATP8A1, CALR, CTAG1A/CTAG1B, FAIM, GJC1, ISYNA1, KIAA1147, LCMT1, NUBP1, PP1R11, PPP1R14B, PPP2CA, PRR14, PRR14L, PXYLP1, RTFDC1, SPRY2, THUMPD1, TRIM51, UBC, VPS13D	10^{-26}	Cell Signaling, Small Molecule, Biochemistry, Vitamin and Mineral Metabolism
10 genes: BMP2, CDH13, FHOD3, HOXC6, PLEKHA6, SCN10A, SIPA1L1, SYNPO2L, ZNF436	ACTB AJAP1, BCL9L, BMP8A, CDH4, CDH6-10, CHD16, CDH18, COL27A1, CREB1, Ctnna, CTNNB1, DENND2A, DLG4, DSC3, ERMN, EZH2, GPR63, GPR123, PCDH20, PDE1C, SCN3B, UBD	10^{-16}	Cancer, Dermatological diseases and Conditions, Organismal Injury and Abnormalities

15. Table S15. Top biological functions of candidate genes using the IPA software tool.

Biological functions	P-value range	Candidate genes (N)
Diseases and Disorders		
Cardiovascular Disease	7.29E-09 - 3.20E-03	22
Organismal Injury and Abnormalities	7.29E-09 - 3.20E-03	62
Cancer	2.63E-09 - 3.20E-03	61
Gastrointestinal Disease	2.63E-09 - 3.20E-03	50
Developmental Disorder	3.83E-09 - 3.20E-03	25
Molecular and Cellular Functions		
Cellular Function and Maintainance	1.53E-08 - 3.20E-03	22
Cell Morphology	2.02E-08 - 3.20E-03	27
Cellular Assembly and Organization	2.22E-07 - 3.20E-03	23
Cellular Development	2.22E-07 - 3.20E-03	31
Cell Growth and Proliferation	2.22E-07 - 3.20E-03	35
Physiological System Development and Function		
Skeletal and Muscular System Development and Function	1.54E-09 - 3.20E-03	29
Cardiovascular System Development and Function	1.72E-09 - 3.20E-03	29
Organ Morphology	1.72E-09 - 3.20E-03	30
Embryonic Development	5.43E-09 - 3.20E-03	30
Organ Development	5.43E-09 - 3.20E-03	29

16. Table S16. Summary of known biology for the 67 candidate genes

Cardiac phenotypes in humans are highlighted (red).

Region	SNP	GENE	Mouse homolog	KO available	Mouse Pheno	OMIM	Gene summary
1p36.12	rs2849028	ZNF436	Zfp46	0			May be a negative regulator in gene transcription mediated by the MAPK signaling pathways.[97] Highly expressed in human foetal brain and heart.[97]
1p36.12	rs2849028	C1orf213	-	0			Putative uncharacterized protein.
1p32.3	rs17391905	CDKN2C	Cdkn2c	1	1		A member of the INK4 family of cyclin-dependent kinase inhibitors. This protein has been shown to interact with CDK4 or CDK6, and prevent the activation of the CDK kinases, thus function as a cell growth regulator that controls cell cycle G1 progression.[98] May play a role in hypoplastic left heart syndrome.[99]
1p31.3	rs2207790	NFIA	Nfia	1			Dimeric DNA-binding protein, function as cellular transcription factors and as replication factors for adenovirus DNA replication.[100]
1p13.1	rs12039739	CASQ2	Casq2	1	1	Ventricular tachycardia, catecholaminergic polymorphic, 2	Specifies the cardiac muscle family member of the calsequestrin family, which are calcium-binding proteins of the sarcoplasmic reticulum. The release of calsequestrin-bound calcium triggers muscle contraction.[101] Mutations can cause abnormal intracellular calcium

						(MIM:611938)	regulation and can facilitate the development of tachyarrhythmias.[102]
1q22	rs2274317	MEF2D	Mef2d	1	1		Transcriptional activator which binds specifically to the MEF2 element, 5'-YTA[AT](4)TAR-3', found in numerous muscle-specific, growth factor- and stress-induced genes. Plays diverse roles in the control of cell growth, survival and apoptosis via p38 MAPK signaling in muscle-specific and/or growth factor-related transcription.[103,104]
1q23.3	rs12036340	OLFML2B	Olfml2b	0			Olfactomedin-like protein 2B. Function unknown.
1q32.1	rs10920184	TNNT2	Tnnt2	1	1	Cardiomyopathy, dilated, 1D (MIM:601494); Cardiomyopathy, familial hypertrophic, 2(MIM:115195); Cardiomyopathy, familial restrictive, 3(MIM:612422); Left ventricular noncompaction 6(MIM:601494)	tropomyosin-binding subunit of the troponin complex, the thin filament regulatory coplex which regulates muscle contraction in response to alterations in intracellular calcium ion concentration. Mutations in this gene cause familial forms of hypertrophic[105], dilated[106] and restrictive cardiomyopathies[107].
1q32.1	rs4288653	PLEKHA6	Plekha6	0			Pleckstrin homology domain containing, family A member 6. Function

							unknown.
2p23.3	rs6710065	DPYSL5	Dpysl5	1			Encodes a member of the CRMP (collapsing response mediator protein) family thought to be involved in neural development (RefSeq).
2p22.2	rs3770770	STRN	Strn	0			Encodes a calmodulin-binding protein which may function as scaffolding or signaling protein and may play a role in (dendritic) Ca^{2+} signaling.[108] Striatin can directly bind to CAV1.[109]
2q31.2	rs3816849	TTN	Ttn	1	1	Cardiomyopathy, dilated, 1G (MIM:604145); Cardiomyopathy, familial hypertrophic, 9 (MIM:613765); Muscular dystrophy, limb-girdle, type 2J (MIM:608807); Myopathy, early-onset, with fatal cardiomyopathy (MIM:611705);	A large abundant protein of striated muscle. A N-terminal Z-disc region and a C-terminal M-line region bind to the Z-line and M-line of the sarcomere respectively so that a single titin molecule spans half the length of a sarcomere. Titin also contains binding sites for muscle-associated proteins so it serves as an adhesion template for the assembly of contractile machinery in muscle cells. It has also been identified as a structural protein for chromosomes. Considerable variability exists in the I-band, the M-line and the Z-disc regions of titin. Variability in the I-band region contributes to the differences in elasticity of different titin isoforms and, therefore, to the differences in elasticity of different muscle types. Mutations are the cause of several hereditary myopathies[110,111], familial hypertrophic[112] and dilated cardiomyopathies[55,113].

						Myopathy, proximal, with early respiratory muscle involvement (MIM;603689); Tibial muscular dystrophy, tardive (MIM:600334)	
3p22.2	rs6801957	SCN10A	Scn10a	1			A tetrodotoxin-resistant voltage-gated sodium channel subunit initially known from and primarily found in the peripheral sensory nervous system.[114] Recently the gene has also been identified in intracardiac neurons contributing to regulation of cardiac electric activity[115,116]
3p21.1	rs4687718	TKT	Tkt	1	1		A thiamine-dependent enzyme which plays a role in the channeling of excess sugar phosphates to glycolysis in the pentose phosphate pathway[117]
3p14.1	rs2242285	LRIG1	Lrig1	1			Act as a feedback negative regulator of signaling by receptor tyrosine kinases, through a mechanism that involves enhancement of receptor ubiquitination and accelerated intracellular degradation.[118]
3p14.1	rs2242285	SLC25A26	Slc25a26	1			A member of the mitochondrial solute carriers shutteling metabolites,

							nucleotides, and cofactors through the mitochondrial inner membrane.[119]
3p14.1	rs13314892	MITF	Mitf	1		Tietz albinism-deafness syndrome (MIM:103500); Waardenburg syndrome, type 2A (MIM:193510); Waardenburg syndrome/ocular albinism, digenic (MIM:103470); Melanoma, cutaneous malignant, susceptibility to, 8 (MIM:614456)	A basic helix-loop-helix leucine zipper transcription factor involved in melanocyte[120] and osteoclast development.[121] Mutations in this gene cause auditory-pigmentary syndromes.
3q27.2	rs10937226	SENP2	Senp2	1	1		Small ubiquitin-like protein that process newly synthesized SUMO1 into the conjugatable form and catalyze the deconjugation of SUMO1-containing species.(RefSeq) Overexpression of SENP2 resulted in

						premature death of mice with CHDs-atrial septal defects (ASDs) and/or ventricular septal defects (VSDs).[122]
4p15.31	rs1344852	SLIT2	Slit2	1		May have a role in axon guidance as repulsive ligands for Roundabout receptors.[123]
5q33.2	rs13185595	HAND1	Hand1	1	1	Belongs to the basic helix-loop-helix family of transcription factors. This gene product is one of two closely related family members, the HAND proteins, which are asymmetrically expressed in the developing ventricular chambers and play an essential role in cardiac morphogenesis. Working in a complementary fashion, they function in the formation of the right ventricle and aortic arch arteries, implicating them as mediators of congenital heart disease. HAND1 mutations have been reported in Chinese patients with VSD.[124] HAND1 deficient mice display defects in the left ventricle and endocardial cushions, and exhibited dysregulated ventricular gene expression.[125]
6p21.31	rs1321311	CDKN1A	Cdkn1a	1		A potent cyclin-dependent kinase inhibitor. The encoded protein binds to and inhibits the activity of cyclin-CDK2 or -CDK4 complexes, and thus functions as a regulator of cell cycle progression at G1. The expression of this gene is tightly controlled by the tumor suppressor protein p53.[126]
6p21.1	rs1015150	TFEB	Tfeb	1		Transcription factor that specifically recognizes and binds E-box sequences (5'-CANNTG-3') and CLEAR-box sequence (5'-GTCACGTGAC-

							3') present in the regulatory region of many lysosomal genes, leading to activate their expression.[127] TFEB overexpression in cultured cells induces lysosomal biogenesis and increases degradation of complex molecules, including glycosaminoglycans and other pathogenic proteins. Some lysosomal storage disorders are known to affect the heart, including Anderson-Fabry and Pompe disease for the latter TFEB is considered a therapeutic target.[128] Homozygotes mice for a targeted null mutation exhibit severe defects in placental vascularization with few vessels entering the placenta and little branching. Mutants die between embryonic days 9.5 and 10.5.
6q22.31	rs11153730	PLN	Pln	1	1	Cardiomyopathy, dilated, 1P(MIM:609909); Cardiomyopathy, familial hypertrophic, 18(MIM:613874)	A major substrate for the cAMP-dependent protein kinase in cardiac muscle. The encoded protein is an inhibitor of cardiac muscle sarcoplasmic reticulum Ca(2+)-ATPase in the unphosphorylated state, but inhibition is relieved upon phosphorylation of the protein. The subsequent activation of the Ca(2+) pump leads to enhanced muscle relaxation rates, thereby contributing to the inotropic response elicited in heart by beta-agonists.[129] Mutations in this gene are a cause of inherited human dilated cardiomyopathy with refractory congestive heart failure.[130]
6q22.31	rs11153730	SLC35F1	Slc35f1	0			Solute carrier family 35 member F1. Function unknown.
6q22.31	rs11153730	CEP85L	Cep85l	0			Centrosomal protein 85kDa-like. Function unknown.
7p14.3	rs1419856	TBX20	Tbx20	1	1	Atrial septal	Transcription factor essential for heart development. Tbx20 physically

						defect 4 (MIM:611363)	interacted with cardiac transcription factors Nkx2-5, GATA4, and GATA5, collaborating to synergistically activate cardiac gene expression.[131] Mutations in this gene are associated with diverse cardiac pathologies, including defects in septation, valvulogenesis and cardiomyopathy.[132]
7p12.3	rs6968945	TNS3	Tns3	1			Tensins are intracellular proteins thought to act as links between the extracellular matrix and the cytoskeletion. TNS3 also interacts with the EGF receptor.[133]
7q31.2	rs11773845	CAV1	Cav1	1	1	Lipodystrophy, congenital generalized, type 3 (MIM:612526)	Main component of the caveolae plasma membranes found in most cell types and links integrin subunits to the tyrosine kinase FYN, an initiating step in coupling integrins to the Ras-ERK pathway and promoting cell cycle progression.The gene is a tumor suppressor gene candidate and a negative regulator of the Ras-p42/44 mitogen-activated kinase cascade. Has been implicated in the compartmentalization and regulation of certain signalling events, including TGF-beta[134] and eNOS.[135] Cav-1/3 dKO mice develop a severe cardiomyopathy.[136]
8q24.13	rs4367519	FBXO32	Fbxo32	1			Subunits of the ubiquitin protein ligase complex with function in phosphorylation-dependent ubiquitination. Probably recognizes and binds to phosphorylated target proteins during skeletal muscle atrophy.[137] Is highly expressed during muscle atrophy, whereas mice deficient in this gene were found to be resistant to atrophy.

8q24.13	rs4367519	KLHL38	Klhl38	0		Kelch-like protein 38. Function unknown.
8q24.13	rs10105974	MTSS1	Mtss1	1		Putative metastasis suppressor protein, which is implicated in actin cytoskeletal control and interaction with protein tyrosine phosphatase.[138]
10q21.1	rs1733724	DKK1	Dkk1	1		Is involved in embryonic development through its inhibition of the WNT signaling pathway by inhibiting LRP5/6 interaction.[139]
10q21.3	rs12414364, rs10509289	CTNNA3	Ctnna3	1*	1*	Localizes to intercalated disks of cardiomyocytes and peritubular myoid cells of testis, and Colocalizes with CTNNA1 and CTNNA2. May be involved in formation of stretch-resistant cell-cell adhesion complexes.[140] Mouse Ctnna3 deficient mice exhibit progressive cardiomyopathy (model yet not included in MGI).[141]
10q22.2	rs7099599	SEC24C	Sec24c	0		Involved in vesicle trafficking. Component of the COPII coat, that covers ER-derived vesicles involved in transport from the ER to Golgi apparatus and may be implicated in cargo selection and concentration.[142]
10q22.2	rs7099599	SYNPO2L	Synpo2l	0		Cytoskeletal protein highly expressed in the Z-disc of the heart and skeletal muscle, associates with actin and interacts with alpha-actinin. Kockdown in zebrafish resulted in aberrant heart and skeletal muscle development, disorganized sarcomeres and diminished cardiac contractility.[143]
10q22.2	rs7099599	CAMK2G	Camk2g	1		Calcium/calmodulin-dependent protein kinase II (CaM kinase II) is a ubiquitous serine/threonine protein kinase that has been implicated

							in diverse effects of hormones and neurotransmitters that utilize Ca(2+) as a second messenger. A mouse model of cardiac Camk2 inhibition demonstrated substantial prevention of maladaptive remodeling from excessive beta-adrenergic receptor stimulation and myocardial infarction, and induction of balanced changes in excitation-contraction coupling that preserved baseline and beta-adrenergic receptor-stimulated physiologic increases in cardiac function.[144]
10q25.2	rs7918405	VTI1A	Vti1a	1			V-SNARE that mediates vesicle transport pathways through interactions with t-SNAREs on the target membrane. Along with VAMP7, involved in a non-conventional RAB1-dependent traffic route to the cell surface used by voltage-gated potassium (Kv) channel-interacting protein 1 (KCNIP1) and potassium voltage-gated channel, Shal-related subfamily, member 2 (KCND2).[145]
11p11.2	rs2269434	MYBPC3	Mybpc3	1	1	Cardiomyopathy, dilated(MIM:115200); Cardiomyopathy, familial hypertrophic, 4(MIM:115197)	Cardiac isoform of myosin-binding protein C, a myosin-associated protein found in the cross-bridge-bearing zone (C region) of A bands in striated muscle. Regulatory phosphorylation by cAMP-dependent protein kinases upon adrenergic stimulation is linked to modulation of cardiac contraction.[146]
11p11.2	rs2269434	NR1H3	Nr1h3	1			NR1 subfamily of the nuclear receptor superfamily. Plays an important

					role in cholesterol homeostasis, regulation of cholesterol uptake. Regulate renin expression in vivo by interacting with the renin promoter and is required for the adrenergic control of the renin-angiotensin system[147] and might be involved in cardiac hypertrophy.[148]
11p11.2	rs2269434	ACP2	Acp2	1	beta subunit of lysosomal acid phosphatase (LAP), which is found in all tissues and is widely used as a biochemical marker for lysosomes.
11p11.2	rs2269434	MADD	Madd	1	Domain-containing adaptor protein that interacts with the death domain of TNF-alpha receptor 1 to activate mitogen-activated protein kinase (MAPK) and propagate the apoptotic signal. Plays a significant role in regulating cell proliferation, survival and death through alternative mRNA splicing.[149]
11q12.2	rs174577	FADS2	Fads2	1	A member of the fatty acid desaturase (FADS) gene family. Component of a lipid metabolic pathway that catalyzes the biosynthesis of highly unsaturated fatty acids from precursor essential polyunsaturated fatty acids, linoleic acid, and alpha-linolenic acid.
11q12.2	rs174577	TMEM258	Tmem258	0	Uncharacterized protein.
12q13.13	rs736825	HOXC4	Hoxc4	1	Homeoprotein of the HOX family, expressed in activated and/or proliferating lymphocytes of the T-, B-, or NK-cell lineage.[150] This protein expands human hematopoietic immature cells ex vivo and improves the level of in vivo engraftmen, possibly by regulating factors involved in stem cell fate or expansion.[151]

12q13.13	rs736825	HOXC5	Hoxc5	1			Homeoprotein of the HOX family, regulated during embryogenesis and activated by retinoic acid in cultured embryonal carcinoma cells.[152]
12q13.13	rs736825	HOXC6	Hoxc6	1			Homeoprotein of the HOX family, plays a key role in a variety of developmental processes including heart development[153].
12q13.3	rs2926743	NACA	Naca	1	1		An isoform of this gene is specifically expressed in myotubes. NACA is converted into a tissue-specific DNA-binding activator, suggesting that this regulation may be an important event in the proper control of gene expression during myogenic differentiation[154]. Knockdown of Naca by antisense oligos in zebrafish embryos results in skeletal muscle defects[155]. NACA degradation also triggers ER stress responses and initiates apoptotic processes in hypoxic cells[156].
12q24.21	rs7132327	TBX3	Tbx3	1	1	Ulnar-mammary syndrome (MIM:181450)	Transcription factors involved in the regulation of developmental processes, it is thought to play a role in the anterior/posterior axis of the tetrapod forelimb[157]. TBX3 is important in heart development; it is involved in atrioventricular myocardial development and endocardial cushion formation [158] and induces important pacemaker properties in cardiomyocytes[159]. Mutations in TBX3 cause Ulnar-mammary syndrome[160].
13q14.13	rs1408224	LRCH1	Lrch1	0			This gene contains leucine-rich repeats and a calponin homology domain, its function is unknown, but this gene has been associated with knee osteoarthritis[161].

13q22.1	rs728926	KLF12	Klf12	0			Member of the Kruppel-like zinc finger protein family, can repress expression of the AP-2 alpha gene by binding to a specific site in the AP-2 alpha gene promoter.[162] AP-2alpha is important in neural crest differentiation and development[163] and gene expression levels are also increased in in human failing myocardium where it may trigger apoptosis[164].
14q24.2	rs12880291	SIPA1L1	Sipa1l1	0			Signal-induced proliferation-associated 1 like 1. Function unknown.
15q25.3	rs7183401	ALPK3	Alpk3	1*	1*		Plays a role in myocyte differentiation[165]. ALPK3 deficient mice develop a predominant hypertrophic cardiomyopathy with reduced cardiac function and impaired contractility[166].
15q26.3	rs8038015	IGF1R	Igf1r	1	1	Insulin-like growth factor I, resistance to (MIM:270450)	This receptor binds insulin-like growth factor with a high affinity. It is regulated by p53 and impairment of its function causes apoptosis of tumor cells and inhibition of tumor growth in animal models[167]. Endogenous IGF-IR signaling is required for conservation of cardiac function of the aging heart, but not for the integrity of cardiac structure and function of young hearts[168]. Igf signalling is important for heart development and myocardial regeneration in zebrafish.[169] Patients with mutation in this gene have intrauterine growth retardation and short stature[170].
16q23.3	rs6565060	CDH13	Cdh13	0			Atypical member of the cadherin family because it lacks the transmembrane and intracellular domains and is attached to the plasma membrane via a glycosylphosphatidylinositol anchor. This

						gene is expressed in endothelial and smooth muscle cells, and is an adiponectin receptor[171]. In vascular tissue, this gene is up-regulated in vivo under disease conditions associated with oxidative stress and concomitant cell migration, proliferation and apoptosis/survival[172].
17q11.2	rs7211246	NSRP1	Ccdc55	1		A nuclear speckle-related protein that is a splicing regulator and essentially required in early stages of embryonic development[173].
17q11.2	rs7211246	EFCAB5	Efcab5	0		EF-hand calcium binding domain 5. The EF hand is a helix-loop-helix structural domain or motif found in a large family of calcium-binding proteins.
17q21.31	rs242562	MAPT	Mapt	1	Dementia, frontotemporal, with or without parkinsonism (MIM:600274); Pick disease (MIM:172700); Supranuclear palsy, progressive (MIM:601104); Supranuclear palsy, progressive atypical	The neuron-specific transcript undergoes complex alternative splicing (PMID: 1420178), depending on stage of neuronal maturation and neuron type. This gene is a major regulator of microtubule formation in cells[174]. Patients with a microdeletion spanning this gene, suffer from typical facial appearance, cardiac and renal defects, and speech delay in addition to intellectual disability, hypotonia and seizures[175].

						(MIM:260540); Tauopathy and respiratory failure, Parkinson disease, susceptibility to (MIM:168600)	
17q24.2	rs9912468	PRKCA	Prkca	1	1		Serine/threonine kinase which has been implicated in a variety of cellular functions including proliferation, apoptosis, differentiation, motility, and inflammation[176]. Mice models have shown that it is a fundamental regulator of cardiac contractility and Ca(2+) handling in myocytes[177].
18q12.1	rs617759	MAPRE2	Mapre2	0			This gene is likely a component of the microtubule cytoskeleton in mammalian cells. Associating with the mitotic apparatus, EB1 may play a physiologic role connecting APC to cellular division, coordinating the control of normal growth and differentiation processes in the colonic epithelium[178].
18q12.2	rs879568	FHOD3	Fhod3	1	1		Member of the formin family of proteins that play pivotal roles in actin filament assembly, this protein is essential for myofibrillogenesis at an early stage of heart development[179]. Two splice variants of exist in a tissue-specific manner; the longer variant is the major form in the heart, whereas the kidney and brain predominantly express a shorter

							protein[180].
18q12.3	rs10853525	SETBP1	Setbp1	0		Schinzel-Giedion midface retraction syndrome (MIM:269150)	This gene binds to SET, a nuclear oncogene, which is involved in DNA replication and associated with leukemogenesis and tumorigenesis[181]. Mutations in this gene cause Schinzel-Giedion midface retraction syndrome, which includes congenital heart defects[182].
20p12.3	rs3929778	BMP2	Bmp2	1	1		Member of the TGF-beta supergene family, involved in bone and cartilage formation[183]. BMP2 is a critical factor for both extraembryonic and embryonic development[184]. The protein prevents apoptosis of myocytes by induction of Bcl-x(L) via a Smad1 pathway and without any hypertrophic effect on myocytes[185].
20q11.22	rs2025096	MYH7B	Myh7b	0			This gene encodes a heavy chain of sarcomeric myosin II molecule, the major contractile protein of cardiac/striated muscle[186].MYH7B is expressed in the myocardium[187] The deducted 1,692-amino acid protein shares 71% identity with MYH7.[188] A mutation in MYH7B has recently been linked to left ventricular non-compaction cardiomyopathy.[189]
20q11.22	rs2025096	GSS	Gss	1		Glutathione synthetase deficiency (MIM:266130); Hemolytic anemia	Functions as a homodimer to catalyze the second step of glutathione biosynthesis, which is the ATP-dependent conversion of gamma-L-glutamyl-L-cysteine to glutathione[190]. Mutations in this gene cause Glutathione synthetase deficiency, causing metabolic acidosis, hemolytic anemia, and mental retardation[191].

					due to glutathione synthetase deficiency (MIM:231900)	
20q11.22	rs2025096	EDEM2	Edem2	0		Regulates endoplasmic reticulum-associated glycoprotein degradation[192,193]
21q21.1	rs7283707	USP25	Usp25	1		USP25 belongs to a complex family of deubiquitinating enzymes that specifically cleave ubiquitin conjugates on a great variety of substrates[194]. This protein has a heart and skeletal muscle specific isoform[195].
21q21.3	rs13047360	ADAMTS5	Adamts5	1		This enzyme functions as aggrecanase to cleave aggrecan, a major proteoglycan of cartilage[196]. Furthermore, it has been implicated in the regulation of proteoglycan turnover and lipoprotein retention in atherosclerosis[197]. ADAMTS5-deficient mice show altered cleavage of versican, a critical cardiac proteoglycan[198].

* models not included in the MGI database, not included in enrichment analysis

17. Table S17. Drosophila Adult Heart Phenotypes

Human gene	*Drosophila* Orthologue	Similarity*	phenotype	P-value	n
ACP2	Acph-1	8	arrhythmicity	3.50E-03	15
			irregular SI rhythm	1.70E-02	
HAND1	Hand	5	reduced heart rate	4.00E-06	32
NACA	NACalpha	9	no adult heart formed		
IGF1R	InR	8	irregular SI rhythm	5.00E-02	15
MADD	rab3-GEF	10	irregular SI rhythm	2.40E-04	14
			arrhythmicity	1.39E-03	
MEF2D	Mef2	6	arrhythmicity	5.00E-09	49
			constricted heart	4.00E-02	
			irregular SI rhythm	4.00E-08	
			reduced contractility	4.00E-08	
			reduced heart rate	2.00E-09	
STRN	Cka	8	arrhythmicity	4.40E-04	77
			constricted heart	3.00E-09	
			irregular SI rhythm	3.00E-02	
			reduced contractility	2.00E-11	
			reduced heart rate	5.00E-03	
TNS3	by	6	Arrhythmicity	6.00E-03	16
			irregular SI rhythm	3.00E-02	
NR1H3	EcR	7	deformed adult heart		
TBX20	Nmr1	9	Arrhythmicity	6.00E-08	16
			irregular SI rhythm	1.00E-06	
			reduced contractility	7.00E-08	
			reduced heart rate	2.00E-05	

* DIOPT score (1-10; refers to the number of databases that report homology according to the method of Hu *et al* [56]). SI – Systolic Interval; HP – Heart Period; Arrhythmicity - standard deviation of HP normalized to the HP; SI rhythm – standard deviation of SI, normalized to the SI

18. Table S18. Tissue and cell type enrichment analysis by DEPICT

ONLINE XLS File

19. Table S19. Significant reconstituted gene sets by DEPICT

ONLINE XLS File

20. Table S20: Key words in enriched reconstituted gene sets by DEPICT.

Comparison of the count of common key words in 404 gene set names with FDR < 5% with the respective count in 14,461 gene set names with FDR > 5%.

key word	N total	N FDR < 5%	N FDR > 5%	% FDR < 5%	% FDR > 5%
all gene sets	14461	404	14057		
protein complex	6011	183	5828	45.30%	41.46%
Abnormal	1080	38	1042	9.41%	7.41%
Muscle	141	36	105	8.91%	0.75%
Heart	62	29	33	7.18%	0.23%
Cardiac	45	27	18	6.68%	0.13%
Morphology	571	26	545	6.44%	3.88%
Development	340	25	315	6.19%	2.24%
Cell	977	23	954	5.69%	6.79%
Regulation	1274	18	1256	4.46%	8.94%
Binding	336	18	318	4.46%	2.26%

N = number of gene sets with FDR > or < 5% for each key word, % = percentage of gene sets with FDR < 5% (or < 5%) for each key word relative to all gene sets with FDR < 5% (or < 5%)

21. Table S21. Gene prioritization by DEPICT

ONLINE XLS File

23. Table S23. Genomic control inflation factors.

	Cornell	Sokolow-Lyon	Lead sum	QRS duration
Meta-analysis				
Europeans	1.083	1.108	1.089	1.039
Individual cohorts				
AGES	1.068	1.049	1.069	1.037
ARIC	1.033	1.039	1.033	1.010
Bright	1.041	1.034	1.036	1.002
Cilento	0.994	1.004	0.991	1.016
CHS	1.014	1.025	1.012	1.021
ERF	0.994	1.026	1.017	1.013
FHS	1.020			1.034
FVG	1.023	0.998	1.014	0.989
Inchianti	0.989	1.021		0.988
KORA S4	1.004	1.001	1.002	1.011
KORA F3				1.015
Korcula	1.247	0.997	1.013	1.031
LifeLines	1.044	1.032	1.042	1.024
LOLIPOP_EW610	1.003	1.000		1.013
LOLIPOP_EW_P	0.999	1.005		1.017
LOLIPOP_EW_A	0.970	0.979		0.997
MESA	1.018	1.022	1.027	1.036
MICROS	1.005	0.997	0.995	1.001
Orcades	0.996	1.011	1.006	
Orkney				0.998
PREVEND	1.028	1.014	1.014	1.036
PROSPER	1.045	1.081	1.033	1.026
RS1	1.026	1.017	1.020	1.013
RS2	1.021	1.010	1.008	1.016
RS3	1.016	1.011	1.014	
Sardinia	1.055	1.096		1.085

SHIP	1.005	1.029	1.023	1.036
Split	0.986	0.973	0.963	1.056
Twins UK	0.996	1.027		1.021
YFS	1.003	1.027	1.011	1.000

24. Table S24. Results of replication testing for the 35 loci associated with QRS phenotypes at $1x10^{-8}<P<5x10^{-7}$

Highlighted in green are the 11 loci that replicated; both *P*-value replication <0.05 and combined P <1×10^{-8}. Highlighted in yellow 11 loci with combined P-value of $1\times10^{-8}<P<5\times10^{-8}$ or combined P <1×10^{-8} but replication P>0.05.

				Discovery		Replication		Combined	
Region	**Position**	**SNP**	**Trait**	**P**	**N**	**P**	**N**	**P**	**N**
1p36.32	3259183	rs6683273	Duration	1.60E-07	54926	6.69E-02	12785	3.40E-08	67711
1p36.22	11887303	rs7537765	Leadsum	2.38E-07	46246	2.46E-02	12806	1.79E-08	59052
1p36.11	26387423	rs2997447	Leadsum	1.85E-07	39724	4.10E-02	12820	2.36E-08	52544
1p13.1	116333111	rs12039739	Duration	4.10E-08	51205	7.33E-03	5183	6.51E-09	56388
1q23.3	162015740	rs12036340	Leadsum	2.51E-08	44291	1.86E-02	12801	1.49E-09	57092
2q31.1	175467769	rs1991601	Leadsum	1.85E-08	42755	9.62E-02	12791	6.34E-09	55546
3p25.2	12842223	rs4642101	Leadsum	3.74E-07	42533	4.27E-01	8276	6.99E-07	50809
3p14.1	69795052	rs13314892	Leadsum	4.17E-07	34683	5.18E-03	12785	8.93E-09	47468
4p15.31	20183937	rs1344852	Duration	9.33E-08	55781	2.10E-03	12760	1.21E-09	68541
4q26	120518064	rs17358860	cornell	1.08E-06	57212	1.99E-01	12746	1.28E-04	69958
5q35.2	173315866	rs359466	Sokolow	3.41E-07	54866	2.39E-02	12806	2.70E-08	67672
6p24.3	7502749	rs7771320	Leadsum	6.87E-08	31952	7.55E-02	12794	1.67E-08	44746
6p24.1	12159699	rs3777755	Leadsum	5.84E-08	33683	1.55E-01	12744	3.34E-08	46427
6p22.1	27413924	rs13195040	cornell	1.41E-07	58854	4.89E-02	12774	3.03E-08	71628
6p21.33	30787762	rs1264353	cornell	3.37E-07	48816	8.51E-01	12811	6.51E-06	61627
6p12.3	46629505	rs9296504	Sokolow	3.23E-07	55481	6.15E-01	12787	2.11E-06	68268
6q25.3	159893937	rs4708832	Duration	1.32E-07	60186	5.52E-02	12812	2.11E-08	72998
10p12.32	18695892	rs7909027	Sokolow	2.09E-07	54589	4.99E-02	12745	3.86E-08	67334
10q22.3	77891246	rs12764182	cornell	1.92E-07	53032	7.26E-01	12729	1.23E-06	65761
10q25.2	112491620	rs2419577	Sokolow	1.60E-06	43903	5.27E-01	12790	4.95E-06	56693
11p15.4	10342711	rs1562782	cornell	2.26E-07	55820	2.48E-01	12775	2.44E-07	68595
11p14.1	30502175	rs10488821	Leadsum	6.19E-07	40853	4.18E-01	12787	7.78E-07	53640
11q12.2	61604814	rs174577	Duration	1.54E-07	52290	1.63E-05	12779	4.28E-11	65069
12q21.31	82576220	rs10778876	cornell	2.79E-07	55138	8.22E-01	12784	3.11E-06	67922
13q14.13	47215218	rs1408224	Leadsum	2.69E-07	46149	1.76E-04	12818	3.60E-10	58967

16p13.13	11688891	rs7198919	Sokolow	1.45E-06	41765	1.25E-02	12799	5.76E-08	54564
17p12	12593743	rs6502201	Sokolow	1.82E-06	54042	1.45E-01	12788	4.95E-04	66830
17q11.2	28485762	rs7211246	Leadsum	1.27E-06	46166	8.34E-04	12790	6.01E-09	58956
17q22	53373550	rs11079159	Duration	3.16E-07	60317	4.55E-01	12772	6.20E-07	73089
18p11.31	6615920	rs4638681	Sokolow	8.76E-07	20593	7.87E-01	12810	4.15E-04	33403
18q12.2	34311659	rs879568	Duration	9.83E-08	57040	2.83E-02	12752	8.45E-09	69792
20p12.3	6460290	rs3929778	cornell	9.49E-07	53072	1.62E-03	12799	6.42E-09	65871
20q11.22	33540000	rs2025096	cornell	5.14E-08	56012	1.35E-04	12770	4.51E-11	68782
21q21.1	17126994	rs7283707	Leadsum	3.31E-08	45147	3.39E-02	12805	3.76E-09	57952
21q21.3	30154239	rs11700980	Leadsum	1.45E-07	42401	3.37E-02	12759	1.61E-08	55160

25. Table S25. Pearson correlation coefficients between QRS phenotypes

Pearson correlation coefficients between QRS phenotypes amongst LifeLines sample (in green) and SNP associations (-log10[*P*] in the European analysis, in blue).

	QRS-duration	12-lead sum product	Sokolow-Lyon product	Cornell product
QRS-duration		0.20	0.15	0.16
12-lead sum product	0.49		0.60	0.27
Sokolow-Lyon product	0.31	0.80		0.13
Cornell product	0.45	0.42	0.22	

26. Table S26. Chromatin data of Roadmap epigenomics project evaluated.

The number of each sample per experiment are indicated.

	Chromatin state					
Sample name	**H3K27ac**	**H3K27me3**	**H3K36me3**	**H3K4me1**	**H3K4me3**	**H3K9me3**
Adipose Nuclei	1	5	5	5	5	5
Adipose Tissue	1	0	0	0	0	0
Adrenal Gland	1	0	0	0	0	1
Adult Kidney	0	0	2	2	2	2
Adult Liver	0	2	3	3	3	3
Aorta	1	1	1	0	0	1
Bone Marrow Derived Mesenchymal Stem Cell Cultured Cells	0	4	4	4	4	4
Brain Anterior Caudate	1	2	2	2	2	2
Brain Cingulate Gyrus	1	1	2	2	2	2
Brain Hippocampus Middle	2	2	3	3	3	3
Brain Inferior Temporal Lobe	1	2	2	2	2	2
Brain Mid Frontal Lobe	1	1	2	2	2	2
Brain Substantia Nigra	0	2	2	2	2	2
Breast Luminal Epithelial Cells	0	1	1	1	0	1
Breast Myoepithelial Cells	0	2	2	2	2	2
CD19 Primary Cells	0	0	1	0	1	0
CD3 Primary Cells	0	1	1	0	1	0
CD34 Primary Cells	0	1	1	0	0	0
CD4 Memory Primary Cells	2	2	2	2	3	2
CD4 Naive Primary Cells	2	2	2	2	2	2
CD4+ CD25- CD45RA+ Naive Primary Cells	1	1	1	1	1	1
CD4+ CD25- CD45RO+ Memory Primary Cells	1	1	1	1	1	1
CD4+ CD25- IL17- PMA-Ionomycin stimulated MACS purified Th Primary Cells	1	1	1	1	1	1
CD4+ CD25- IL17+ PMA-Ionomcyin stimulated Th17	1	1	1	1	1	1

Primary Cells						
CD4+ CD25- Th Primary Cells	0	1	1	1	1	1
CD4+ CD25+ CD127- Treg Primary Cells	0	1	1	1	1	1
CD4+ CD25int CD127+ Tmem Primary Cells	1	1	1	1	1	1
CD8 Memory Primary Cells	1	2	2	2	2	2
CD8 Naive Primary Cells	2	3	3	3	3	3
Chondrocytes from Bone Marrow Derived Mesenchymal Stem Cell Cultured Cells	0	2	2	1	2	2
Colon Smooth Muscle	0	1	2	1	2	1
Colonic Mucosa	0	2	2	2	2	2
Duodenum Mucosa	0	2	2	2	2	0
Duodenum Smooth Muscle	1	1	1	1	1	1
Esophagus	1	1	1	1	1	1
Foetal Brain	0	4	3	4	3	4
Foetal Heart	0	1	1	0	0	1
Foetal Lung	0	1	1	2	0	1
Left Ventricle	2	2	2	2	2	2
Lung	1	0	0	1	1	1
Mesenchymal Stem Cell Derived Adipocyte Cultured Cells	0	4	5	5	5	5
Mobilized CD34 Primary Cells	3	8	7	6	6	7
Muscle Satellite Cultured Cells	0	3	3	3	3	3
Neurosphere Cultured Cells Cortex Derived	0	2	2	2	1	2
Neurosphere Cultured Cells Ganglionic Eminence Derived	0	2	2	2	1	2
Pancreas	1	1	1	1	1	1
Pancreatic Islets	0	1	1	1	1	0
Penis Foreskin Fibroblast Primary Cells	1	3	3	3	3	2
Penis Foreskin Keratinocyte Primary Cells	1	3	3	3	3	2
Penis Foreskin Melanocyte Primary Cells	1	3	3	3	3	3
Psoas Muscle	1	1	0	0	0	0
Rectal Mucosa	0	2	2	2	2	2
Rectal Smooth Muscle	0	1	1	1	1	1
Right Atrium	0	0	0	1	0	1

Right Ventricle	1	1	0	0	0	0
Sigmoid Colon	1	1	0	0	0	0
Skeletal Muscle	1	3	3	3	3	3
Small Intestine	1	1	1	0	0	0
Spleen	1	1	1	1	1	1
Stomach Mucosa	0	1	1	1	1	1
Stomach Smooth Muscle	1	1	2	1	2	2
Th17 Primary Cells	0	0	0	0	0	1
Treg Primary Cells	0	1	1	0	1	1

27. Table S27. Mammalian Phenotype (MP) identifiers of the 154 Mammalian Phenotypes queried.

Mammalian Phenotype Identifier					
MP:0000266	MP:0002953	MP:0004124	MP:0006268	MP:0010432	MP:0010578
MP:0000267	MP:0002972	MP:0004215	MP:0006321	MP:0010446	MP:0010579
MP:0000268	MP:0003137	MP:0004251	MP:0008022	MP:0010447	MP:0010580
MP:0000269	MP:0003141	MP:0004252	MP:0008772	MP:0010494	MP:0010592
MP:0000270	MP:0003210	MP:0004484	MP:0008788	MP:0010498	MP:0010599
MP:0000274	MP:0003221	MP:0004485	MP:0008823	MP:0010499	MP:0010612
MP:0000275	MP:0003222	MP:0004486	MP:0008824	MP:0010500	MP:0010630
MP:0000277	MP:0003223	MP:0004564	MP:0009328	MP:0010502	MP:0010631
MP:0000278	MP:0003393	MP:0004565	MP:0009382	MP:0010503	MP:0010632
MP:0000279	MP:0003394	MP:0004566	MP:0009416	MP:0010508	MP:0010633
MP:0000280	MP:0003567	MP:0004567	MP:0009418	MP:0010513	MP:0010634
MP:0000281	MP:0003898	MP:0004857	MP:0009863	MP:0010515	MP:0010636
MP:0000304	MP:0003915	MP:0004937	MP:0010392	MP:0010516	MP:0010638
MP:0001625	MP:0003916	MP:0005140	MP:0010393	MP:0010534	MP:0010640
MP:0001627	MP:0003921	MP:0005294	MP:0010394	MP:0010535	MP:0010655
MP:0002188	MP:0004032	MP:0005329	MP:0010402	MP:0010545	MP:0010656
MP:0002189	MP:0004056	MP:0005330	MP:0010412	MP:0010546	MP:0010724
MP:0002190	MP:0004057	MP:0005406	MP:0010413	MP:0010547	MP:0010725
MP:0002625	MP:0004058	MP:0005598	MP:0010414	MP:0010548	MP:0010754
MP:0002652	MP:0004060	MP:0005599	MP:0010415	MP:0010549	MP:0011264
MP:0002740	MP:0004067	MP:0005600	MP:0010416	MP:0010555	MP:0011388
MP:0002753	MP:0004084	MP:0005608	MP:0010417	MP:0010556	MP:0011390
MP:0002795	MP:0004086	MP:0006085	MP:0010418	MP:0010560	MP:0011394
MP:0002833	MP:0004116	MP:0006107	MP:0010419	MP:0010566	MP:0011395
MP:0002834	MP:0004117	MP:0006113	MP:0010420	MP:0010567	
MP:0002952	MP:0004123	MP:0006138	MP:0010421	MP:0010569	

Supplementary References

1. Levy, D. et al. Determinants of sensitivity and specificity of electrocardiographic criteria for left ventricular hypertrophy. *Circulation* **81**, 815-20 (1990).
2. Devereux, R.B., Koren, M.J., de Simone, G., Okin, P.M. & Kligfield, P. Methods for detection of left ventricular hypertrophy: application to hypertensive heart disease. *Eur Heart J* **14 Suppl D**, 8-15 (1993).
3. Okin, P.M. et al. Time-voltage QRS area of the 12-lead electrocardiogram: detection of left ventricular hypertrophy. *Hypertension* **31**, 937-42 (1998).
4. Kannel, W.B., Gordon, T. & Offutt, D. Left ventricular hypertrophy by electrocardiogram. Prevalence, incidence, and mortality in the Framingham study. *Ann Intern Med* **71**, 89-105 (1969).
5. Verdecchia, P. et al. Prognostic value of a new electrocardiographic method for diagnosis of left ventricular hypertrophy in essential hypertension. *J Am Coll Cardiol* **31**, 383-90 (1998).
6. Usoro, A.O., Bradford, N., Shah, A.J. & Soliman, E.Z. Risk of mortality in individuals with low QRS voltage and free of cardiovascular disease. *Am J Cardiol* **113**, 1514-7 (2014).
7. Kamath, S.A. et al. Low voltage on the electrocardiogram is a marker of disease severity and a risk factor for adverse outcomes in patients with heart failure due to systolic dysfunction. *Am Heart J* **152**, 355-61 (2006).
8. Sokolow, M. & Lyon, T.P. The ventricular complex in left ventricular hypertrophy as obtained by unipolar precordial and limb leads. *Am Heart J* **37**, 161-86 (1949).
9. Casale, P.N. et al. Electrocardiographic detection of left ventricular hypertrophy: development and prospective validation of improved criteria. *J Am Coll Cardiol* **6**, 572-80 (1985).
10. Siegel, R.J. & Roberts, W.C. Electrocardiographic observations in severe aortic valve stenosis: correlative necropsy study to clinical, hemodynamic,, and ECG variables demonstrating relation of 12-lead QRS amplitude to peak systolic transaortic pressure gradient. *Am Heart J* **103**, 210-21 (1982).
11. Molloy, T.J., Okin, P.M., Devereux, R.B. & Kligfield, P. Electrocardiographic detection of left ventricular hypertrophy by the simple QRS voltage-duration product. *J Am Coll Cardiol* **20**, 1180-6 (1992).
12. Casale, P.N., Devereux, R.B., Alonso, D.R., Campo, E. & Kligfield, P. Improved sex-specific criteria of left ventricular hypertrophy for clinical and computer interpretation of electrocardiograms: validation with autopsy findings. *Circulation* **75**, 565-72 (1987).
13. Mazzoleni, A., Curtin, M.E., Wolff, R., Reiner, L. & Somes, G. On the relationship between heart weights, fibrosis, and QRS duration. *J Electrocardiol* **8**, 233-6 (1975).
14. Hancock, E.W. et al. AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: part V: electrocardiogram changes associated with cardiac chamber hypertrophy: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society: endorsed by the International Society for Computerized Electrocardiology. *Circulation* **119**, e251-61 (2009).
15. Carlsson, M.B. et al. Left ventricular mass by 12-lead electrocardiogram in healthy subjects: comparison to cardiac magnetic resonance imaging. *J Electrocardiol* **39**, 67-72 (2006).
16. Okin, P.M., Roman, M.J., Devereux, R.B. & Kligfield, P. Electrocardiographic identification of increased left ventricular mass by simple voltage-duration products. *J Am Coll Cardiol* **25**, 417-23 (1995).
17. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
18. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
19. Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M.J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol* **32**, 381-5 (2008).

20. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78 (2007).
21. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* **44**, 369-75, S1-3 (2012).
22. Bernstein, B.E. et al. The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol* **28**, 1045-8 (2010).
23. Dunham, I. et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
24. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X.S. Identifying ChIP-seq enrichment using MACS. *Nat Protoc* **7**, 1728-40 (2012).
25. Kundaje, A. et al. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-30 (2015).
26. van den Boogaard, M. et al. Genetic variation in T-box binding element functionally affects SCN5A/SCN10A enhancer. *J Clin Invest* **122**, 2519-30 (2012).
27. May, D. et al. Large-scale discovery of enhancers from human heart tissue. *Nat Genet* **44**, 89-93 (2012).
28. He, A., Kong, S.W., Ma, Q. & Pu, W.T. Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. *Proc Natl Acad Sci U S A* **108**, 5632-7 (2011).
29. Blow, M.J. et al. ChIP-Seq identification of weakly conserved heart enhancers. *Nat Genet* **42**, 806-10 (2010).
30. Maurano, M.T. et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190-5 (2012).
31. Matys, V. et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res* **34**, D108-10 (2006).
32. Portales-Casamar, E. et al. JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res* **38**, D105-10 (2010).
33. Newburger, D.E. & Bulyk, M.L. UniPROBE: an online database of protein binding microarray data on protein-DNA interactions. *Nucleic Acids Res* **37**, D77-82 (2009).
34. Jolma, A. et al. DNA-binding specificities of human transcription factors. *Cell* **152**, 327-39 (2013).
35. Abecasis, G.R. et al. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061-73 (2010).
36. Grant, C.E., Bailey, T.L. & Noble, W.S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017-8 (2011).
37. Neph, S. et al. Circuitry and dynamics of human transcription factor regulatory networks. *Cell* **150**, 1274-86 (2012).
38. Hsiao, E.C. et al. Marking embryonic stem cells with a 2A self-cleaving peptide: a NKX2-5 emerald GFP BAC reporter. *PLoS One* **3**, e2532 (2008).
39. Wamstad, J.A. et al. Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. *Cell* **151**, 206-20 (2012).
40. Kattman, S.J. et al. Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell Stem Cell* **8**, 228-40 (2011).
41. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11**, R106 (2010).
42. Nix, D.A., Courdy, S.J. & Boucher, K.M. Empirical methods for controlling false positives and estimating confidence in ChIP-Seq peaks. *BMC Bioinformatics* **9**, 523 (2008).
43. Loh, Y.H. et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* **38**, 431-40 (2006).
44. van de Werken, H.J. et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat Methods* **9**, 969-72 (2012).
45. Kothary, R. et al. Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. *Development* **105**, 707-14 (1989).

46. Koopmann, T.T. et al. Genome-wide identification of expression quantitative trait loci (eQTLs) in human heart. *PLoS One* **9**, e97380 (2014).
47. Anders, S., Pyl, P.T. & Huber, W. *HTSeq - A Python framework to work with high-throughput sequencing data*, (2014).
48. Schulte, J.H. et al. Deep sequencing reveals differential expression of microRNAs in favorable versus unfavorable neuroblastoma. *Nucleic Acids Res* **38**, 5919-28 (2010).
49. Raychaudhuri, S. et al. Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet* **5**, e1000534 (2009).
50. Neely, G.G. et al. A global in vivo Drosophila RNAi screen identifies NOT3 as a conserved regulator of heart function. *Cell* **141**, 142-53 (2010).
51. Melkani, G.C., Bodmer, R., Ocorr, K. & Bernstein, S.I. The UNC-45 chaperone is critical for establishing myosin-based myofibrillar organization and cardiac contractility in the Drosophila heart model. *PLoS One* **6**, e22579 (2011).
52. Qian, L., Liu, J. & Bodmer, R. Slit and Robo control cardiac cell polarity and morphogenesis. *Curr Biol* **15**, 2271-8 (2005).
53. Monier, B., Astier, M., Semeriva, M. & Perrin, L. Steroid-dependent modification of Hox function drives myocyte reprogramming in the Drosophila heart. *Development* **132**, 5283-93 (2005).
54. Han, Z., Yi, P., Li, X. & Olson, E.N. Hand, an evolutionarily conserved bHLH transcription factor required for Drosophila cardiogenesis and hematopoiesis. *Development* **133**, 1175-82 (2006).
55. Herman, D.S. et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* **366**, 619-28 (2012).
56. Hu, Y. et al. An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics* **12**, 357 (2011).
57. Fink, M. et al. A new method for detection and quantification of heartbeat parameters in Drosophila, zebrafish, and embryonic mouse hearts. *Biotechniques* **46**, 101-13 (2009).
58. Sellin, J., Albrecht, S., Kolsch, V. & Paululat, A. Dynamics of heart differentiation, visualized utilizing heart enhancer elements of the Drosophila melanogaster bHLH transcription factor Hand. *Gene Expr Patterns* **6**, 360-75 (2006).
59. Dietzl, G. et al. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* **448**, 151-6 (2007).
60. Vogler, G. & Ocorr, K. Visualizing the beating heart in Drosophila. *J Vis Exp* (2009).
61. Ocorr, K., Fink, M., Cammarato, A., Bernstein, S. & Bodmer, R. Semi-automated Optical Heartbeat Analysis of small hearts. *J Vis Exp* (2009).
62. Alayari, N.N. et al. Fluorescent labeling of Drosophila heart structures. *J Vis Exp* (2009).
63. Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-5 (2012).
64. Pers, T.H. et al. Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* **6**, 5890 (2015).
65. Cvejic, A. et al. SMIM1 underlies the Vel blood group and influences red blood cell traits. *Nat Genet* **45**, 542-5 (2013).
66. Lage, K. et al. A human phenome-interactome network of protein complexes implicated in genetic disorders. *Nat Biotechnol* **25**, 309-16 (2007).
67. Blake, J.A., Bult, C.J., Eppig, J.T., Kadin, J.A. & Richardson, J.E. The Mouse Genome Database: integration of and access to knowledge about the laboratory mouse. *Nucleic Acids Res* **42**, D810-7 (2014).
68. Croft, D. et al. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res* **39**, D691-7 (2011).
69. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* **40**, D109-14 (2012).
70. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-9 (2000).

71. Wood, A.R. et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* **46**, 1173-86 (2014).
72. Geller, F. et al. Genome-wide association analyses identify variants in developmental genes associated with hypospadias. *Nat Genet* **46**, 957-63 (2014).
73. van der Valk, R.J. et al. A novel common variant in DCST2 is associated with length in early life and height in adulthood. *Hum Mol Genet* **24**, 1155-68 (2015).
74. Fried, L.P. et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* **1**, 263-76 (1991).
75. Pardo, L.M., MacKay, I., Oostra, B., van Duijn, C.M. & Aulchenko, Y.S. The effect of genetic drift in a young genetically isolated population. *Ann Hum Genet* **69**, 288-95 (2005).
76. Ferrucci, L. et al. Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. *J Am Geriatr Soc* **48**, 1618-25 (2000).
77. Melzer, D. et al. A genome-wide association study identifies protein quantitative trait loci (pQTLs). *PLoS Genet* **4**, e1000072 (2008).
78. Wichmann, H.E., Gieger, C. & Illig, T. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67 Suppl 1**, S26-30 (2005).
79. Holle, R., Happich, M., Lowel, H. & Wichmann, H.E. KORA--a research platform for population based health research. *Gesundheitswesen* **67 Suppl 1**, S19-25 (2005).
80. Chambers, J.C. et al. Common genetic variation near MC4R is associated with waist circumference and insulin resistance. *Nat Genet* **40**, 716-8 (2008).
81. Kooner, J.S. et al. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nat Genet* **40**, 149-51 (2008).
82. Michelucci, A. et al. Simultaneous assessment of electrocardiographic parameters for risk stratification: validation in healthy subjects. *Ital Heart J* **3**, 308-17 (2002).
83. Pattaro, C. et al. The genetic study of three population microisolates in South Tyrol (MICROS): study design and epidemiological perspectives. *BMC Med Genet* **8**, 29 (2007).
84. Shepherd, J. et al. The design of a prospective study of Pravastatin in the Elderly at Risk (PROSPER). PROSPER Study Group. PROspective Study of Pravastatin in the Elderly at Risk. *Am J Cardiol* **84**, 1192-7 (1999).
85. Shepherd, J. et al. Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *Lancet* **360**, 1623-30 (2002).
86. Pilia, G. et al. Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet* **2**, e132 (2006).
87. Raitakari, O.T. et al. Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol* **37**, 1220-6 (2008).
88. Estrada, K. et al. GRIMP: a web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics* **25**, 2750-2 (2009).
89. Sotoodehnia, N. et al. Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. *Nat Genet* **42**, 1068-76 (2010).
90. Verweij, N. et al. Genetic determinants of P wave duration and PR segment. *Circ Cardiovasc Genet* **7**, 475-81 (2014).
91. Arking, D.E. et al. Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. *Nat Genet* **46**, 826-36 (2014).
92. Newton-Cheh, C. et al. Common variants at ten loci influence QT interval duration in the QTGEN Study. *Nat Genet* **41**, 399-406 (2009).
93. den Hoed, M. et al. Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nat Genet* **45**, 621-31 (2013).
94. Eijgelsheim, M. et al. Genome-wide association analysis identifies multiple loci related to resting heart rate. *Hum Mol Genet* **19**, 3885-94 (2010).
95. Pfeufer, A. et al. Genome-wide association study of PR interval. *Nat Genet* **42**, 153-9 (2010).

96. Holm, H. et al. Several common variants modulate heart rate, PR interval and QRS duration. *Nat Genet* **42**, 117-22 (2010).
97. Li, Y. et al. A novel zinc-finger protein ZNF436 suppresses transcriptional activities of AP-1 and SRE. *Mol Biol Rep* **33**, 287-94 (2006).
98. Guan, K.L. et al. Growth suppression by p18, a p16INK4/MTS1- and p14INK4B/MTS2-related CDK6 inhibitor, correlates with wild-type pRb function. *Genes Dev* **8**, 2939-52 (1994).
99. Gambetta, K., Al-Ahdab, M.K., Ilbawi, M.N., Hassaniya, N. & Gupta, M. Transcription repression and blocks in cell cycle progression in hypoplastic left heart syndrome. *Am J Physiol Heart Circ Physiol* **294**, H2268-75 (2008).
100. Qian, F., Kruse, U., Lichter, P. & Sippel, A.E. Chromosomal localization of the four genes (NFIA, B, C, and X) for the human transcription factor nuclear factor I by FISH. *Genomics* **28**, 66-73 (1995).
101. Slupsky, J.R., Ohnishi, M., Carpenter, M.R. & Reithmeier, R.A. Characterization of cardiac calsequestrin. *Biochemistry* **26**, 6539-44 (1987).
102. di Barletta, M.R. et al. Clinical phenotype and functional characterization of CASQ2 mutations associated with catecholaminergic polymorphic ventricular tachycardia. *Circulation* **114**, 1012-9 (2006).
103. Breitbart, R.E. et al. A fourth human MEF2 transcription factor, hMEF2D, is an early marker of the myogenic lineage. *Development* **118**, 1095-106 (1993).
104. Ma, K., Chan, J.K., Zhu, G. & Wu, Z. Myocyte enhancer factor 2 acetylation by p300 enhances its DNA binding activity, transcriptional activity, and myogenic differentiation. *Mol Cell Biol* **25**, 3575-82 (2005).
105. Thierfelder, L. et al. Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. *Cell* **77**, 701-12 (1994).
106. Kamisago, M. et al. Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy. *N Engl J Med* **343**, 1688-96 (2000).
107. Peddy, S.B. et al. Infantile restrictive cardiomyopathy resulting from a mutation in the cardiac troponin T gene. *Pediatrics* **117**, 1830-3 (2006).
108. Castets, F. et al. A novel calmodulin-binding protein, belonging to the WD-repeat family, is localized in dendrites of a subset of CNS neurons. *J Cell Biol* **134**, 1051-62 (1996).
109. Gaillard, S., Bartoli, M., Castets, F. & Monneron, A. Striatin, a calmodulin-dependent scaffolding protein, directly binds caveolin-1. *FEBS Lett* **508**, 49-52 (2001).
110. Lange, S. et al. The kinase domain of titin controls muscle gene expression and protein turnover. *Science* **308**, 1599-603 (2005).
111. Hackman, P. et al. Tibial muscular dystrophy is a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin. *Am J Hum Genet* **71**, 492-500 (2002).
112. Satoh, M. et al. Structural analysis of the titin gene in hypertrophic cardiomyopathy: identification of a novel disease gene. *Biochem Biophys Res Commun* **262**, 411-7 (1999).
113. Gerull, B. et al. Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nat Genet* **30**, 201-4 (2002).
114. Rabert, D.K. et al. A tetrodotoxin-resistant voltage-gated sodium channel from human dorsal root ganglia, hPN3/SCN10A. *Pain* **78**, 107-14 (1998).
115. Yang, T. et al. Blocking Scn10a channels in heart reduces late sodium current and is antiarrhythmic. *Circ Res* **111**, 322-32 (2012).
116. Verkerk, A.O. et al. Functional Nav1.8 channels in intracardiac neurons: the link between SCN10A and cardiac electrophysiology. *Circ Res* **111**, 333-43 (2012).
117. Jung, E.H., Takeuchi, T., Nishino, K. & Itokawa, Y. Studies on the nature of thiamine pyrophosphate binding and dependency on divalent cations of transketolase from human erythrocytes. *Int J Biochem* **20**, 1255-9 (1988).
118. Gur, G. et al. LRIG1 restricts growth factor signaling by enhancing receptor ubiquitylation and degradation. *EMBO J* **23**, 3270-81 (2004).

119. Agrimi, G. et al. Identification of the human mitochondrial S-adenosylmethionine transporter: bacterial expression, reconstitution, functional characterization and tissue distribution. *Biochem J* **379**, 183-90 (2004).
120. Levy, C., Khaled, M. & Fisher, D.E. MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol Med* **12**, 406-14 (2006).
121. Hershey, C.L. & Fisher, D.E. Mitf and Tfe3: members of a b-HLH-ZIP transcription factor family essential for osteoclast development and function. *Bone* **34**, 689-96 (2004).
122. Kim, E.Y. et al. Enhanced desumoylation in murine hearts by overexpressed SENP2 leads to congenital heart defects and cardiac dysfunction. *J Mol Cell Cardiol* **52**, 638-49 (2012).
123. Brose, K. et al. Slit proteins bind Robo receptors and have an evolutionarily conserved role in repulsive axon guidance. *Cell* **96**, 795-806 (1999).
124. Cheng, Z. et al. Two novel HAND1 mutations in Chinese patients with ventricular septal defect. *Clin Chim Acta* **413**, 675-7 (2012).
125. McFadden, D.G. et al. The Hand1 and Hand2 transcription factors regulate expansion of the embryonic cardiac ventricles in a gene dosage-dependent manner. *Development* **132**, 189-201 (2005).
126. Harper, J.W., Adami, G.R., Wei, N., Keyomarsi, K. & Elledge, S.J. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* **75**, 805-16 (1993).
127. Sardiello, M. et al. A gene network regulating lysosomal biogenesis and function. *Science* **325**, 473-7 (2009).
128. Li, H.M. et al. Clearance of lysosomal glycogen accumulation by Transcription factor EB (TFEB) in muscle cells from lysosomal alpha-glucosidase deficient mice. *Biochem Biophys Res Commun* (2013).
129. Luo, W. et al. Targeted ablation of the phospholamban gene is associated with markedly enhanced myocardial contractility and loss of beta-agonist stimulation. *Circ Res* **75**, 401-9 (1994).
130. Akin, B.L. & Jones, L.R. Characterizing phospholamban to sarco(endo)plasmic reticulum Ca2+-ATPase 2a (SERCA2a) protein binding interactions in human cardiac sarcoplasmic reticulum vesicles using chemical cross-linking. *J Biol Chem* **287**, 7582-93 (2012).
131. Stennard, F.A. et al. Cardiac T-box factor Tbx20 directly interacts with Nkx2-5, GATA4, and GATA5 in regulation of gene expression in the developing heart. *Dev Biol* **262**, 206-24 (2003).
132. Kirk, E.P. et al. Mutations in cardiac T-box factor gene TBX20 are associated with diverse cardiac pathologies, including defects of septation and valvulogenesis and cardiomyopathy. *Am J Hum Genet* **81**, 280-91 (2007).
133. Cui, Y., Liao, Y.C. & Lo, S.H. Epidermal growth factor modulates tyrosine phosphorylation of a novel tensin family member, tensin3. *Mol Cancer Res* **2**, 225-32 (2004).
134. Razani, B. et al. Caveolin-1 regulates transforming growth factor (TGF)-beta/SMAD signaling through an interaction with the TGF-beta type I receptor. *J Biol Chem* **276**, 6727-38 (2001).
135. Garcia-Cardena, G., Fan, R., Stern, D.F., Liu, J. & Sessa, W.C. Endothelial nitric oxide synthase is regulated by tyrosine phosphorylation and interacts with caveolin-1. *J Biol Chem* **271**, 27237-40 (1996).
136. Park, D.S. et al. Caveolin-1/3 double-knockout mice are viable, but lack both muscle and non-muscle caveolae, and develop a severe cardiomyopathic phenotype. *Am J Pathol* **160**, 2207-17 (2002).
137. Tintignac, L.A. et al. Degradation of MyoD mediated by the SCF (MAFbx) ubiquitin ligase. *J Biol Chem* **280**, 2847-56 (2005).
138. Woodings, J.A., Sharp, S.J. & Machesky, L.M. MIM-B, a putative metastasis suppressor protein, binds to actin and to protein tyrosine phosphatase delta. *Biochem J* **371**, 463-71 (2003).
139. Ahn, V.E. et al. Structural basis of Wnt signaling inhibition by Dickkopf binding to LRP5/6. *Dev Cell* **21**, 862-73 (2011).
140. Janssens, B. et al. alphaT-catenin: a novel tissue-specific beta-catenin-binding protein mediating strong cell-cell adhesion. *J Cell Sci* **114**, 3177-88 (2001).

141. Li, J. et al. Loss of alphaT-catenin alters the hybrid adhering junctions in the heart and leads to dilated cardiomyopathy and ventricular arrhythmia following acute ischemia. *J Cell Sci* **125**, 1058-67 (2012).
142. Pagano, A. et al. Sec24 proteins and sorting at the endoplasmic reticulum. *J Biol Chem* **274**, 7833-40 (1999).
143. Beqqali, A. et al. CHAP is a newly identified Z-disc protein essential for heart and skeletal muscle function. *J Cell Sci* **123**, 1141-50 (2010).
144. Zhang, R. et al. Calmodulin kinase II inhibition protects against structural heart disease. *Nat Med* **11**, 409-17 (2005).
145. Flowerdew, S.E. & Burgoyne, R.D. A VAMP7/Vti1a SNARE complex distinguishes a non-conventional traffic route to the cell surface used by KChIP1 and Kv4 potassium channels. *Biochem J* **418**, 529-40 (2009).
146. Gautel, M., Zuffardi, O., Freiburg, A. & Labeit, S. Phosphorylation switches specific for the cardiac isoform of myosin binding protein-C: a modulator of cardiac contraction? *EMBO J* **14**, 1952-60 (1995).
147. Morello, F. et al. Liver X receptors alpha and beta regulate renin expression in vivo. *J Clin Invest* **115**, 1913-22 (2005).
148. Kuipers, I. et al. Activation of liver X receptors with T0901317 attenuates cardiac hypertrophy in vivo. *Eur J Heart Fail* **12**, 1042-50 (2010).
149. Schievella, A.R., Chen, J.H., Graham, J.R. & Lin, L.L. MADD, a novel death domain protein that interacts with the type 1 tumor necrosis factor receptor and activates mitogen-activated protein kinase. *J Biol Chem* **272**, 12069-75 (1997).
150. Meazza, R. et al. Expression of HOXC4 homeoprotein in the nucleus of activated human lymphocytes. *Blood* **85**, 2084-90 (1995).
151. Auvray, C. et al. HOXC4 homeoprotein efficiently expands human hematopoietic stem cells and triggers similar molecular alterations as HOXB4. *Haematologica* **97**, 168-78 (2012).
152. Arcioni, L. et al. The upstream region of the human homeobox gene HOX3D is a target for regulation by retinoic acid and HOX homeoproteins. *EMBO J* **11**, 265-77 (1992).
153. Simeone, A. et al. Two human homeobox genes, c1 and c8: structure analysis and expression in embryonic development. *Proc Natl Acad Sci U S A* **84**, 4914-8 (1987).
154. Yotov, W.V. & St-Arnaud, R. Differential splicing-in of a proline-rich exon converts alphaNAC into a muscle-specific transcription factor. *Genes Dev* **10**, 1763-72 (1996).
155. Li, H., Randall, W.R. & Du, S.J. skNAC (skeletal Naca), a muscle-specific isoform of Naca (nascent polypeptide-associated complex alpha), is required for myofibril organization. *FASEB J* **23**, 1988-2000 (2009).
156. Hotokezaka, Y. et al. alphaNAC depletion as an initiator of ER stress-induced apoptosis in hypoxia. *Cell Death Differ* **16**, 1505-14 (2009).
157. He, M., Wen, L., Campbell, C.E., Wu, J.Y. & Rao, Y. Transcription repression by Xenopus ET and its human ortholog TBX3, a gene involved in ulnar-mammary syndrome. *Proc Natl Acad Sci U S A* **96**, 10212-7 (1999).
158. Hoogaars, W.M. et al. Tbx3 controls the sinoatrial node gene program and imposes pacemaker function on the atria. *Genes Dev* **21**, 1098-112 (2007).
159. Bakker, M.L. et al. T-box transcription factor TBX3 reprogrammes mature cardiac myocytes into pacemaker-like cells. *Cardiovasc Res* **94**, 439-49 (2012).
160. Bamshad, M. et al. Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome. *Nat Genet* **16**, 311-5 (1997).
161. Spector, T.D. et al. Association between a variation in LRCH1 and knee osteoarthritis: a genome-wide single-nucleotide polymorphism association study using DNA pooling. *Arthritis Rheum* **54**, 524-32 (2006).
162. Roth, C., Schuierer, M., Gunther, K. & Buettner, R. Genomic structure and DNA binding properties of the human zinc finger transcriptional repressor AP-2rep (KLF12). *Genomics* **63**, 384-90 (2000).

163. Brewer, S., Feng, W., Huang, J., Sullivan, S. & Williams, T. Wnt1-Cre-mediated deletion of AP-2alpha causes multiple neural crest-related defects. *Dev Biol* **267**, 135-52 (2004).
164. Muller, F.U. et al. Transcription factor AP-2alpha triggers apoptosis in cardiac myocytes. *Cell Death Differ* **11**, 485-93 (2004).
165. Hosoda, T. et al. A novel myocyte-specific gene Midori promotes the differentiation of P19CL6 cells into cardiomyocytes. *J Biol Chem* **276**, 35978-89 (2001).
166. Van Sligtenhorst, I. et al. Cardiomyopathy in alpha-Kinase 3 (ALPK3)-Deficient Mice. *Vet Pathol* **49**, 131-41 (2012).
167. Baserga, R. The IGF-I receptor in cancer research. *Exp Cell Res* **253**, 1-6 (1999).
168. Moellendorf, S. et al. IGF-IR signaling attenuates the age-related decline of diastolic cardiac function. *Am J Physiol Endocrinol Metab* **303**, E213-22 (2012).
169. Huang, Y. et al. Igf Signaling is Required for Cardiomyocyte Proliferation during Zebrafish Heart Development and Regeneration. *PLoS One* **8**, e67266 (2013).
170. Abuzzahab, M.J. et al. IGF-I receptor mutations resulting in intrauterine and postnatal growth retardation. *N Engl J Med* **349**, 2211-22 (2003).
171. Hug, C. et al. T-cadherin is a receptor for hexameric and high-molecular-weight forms of Acrp30/adiponectin. *Proc Natl Acad Sci U S A* **101**, 10308-13 (2004).
172. Andreeva, A.V. et al. T-cadherin modulates endothelial barrier function. *J Cell Physiol* **223**, 94-102 (2010).
173. Kim, Y.D. et al. NSrp70 is a novel nuclear speckle-related protein that modulates alternative pre-mRNA splicing in vivo. *Nucleic Acids Res* **39**, 4300-14 (2011).
174. Weingarten, M.D., Lockwood, A.H., Hwo, S.Y. & Kirschner, M.W. A protein factor essential for microtubule assembly. *Proc Natl Acad Sci U S A* **72**, 1858-62 (1975).
175. Tan, T.Y. et al. Phenotypic expansion and further characterisation of the 17q21.31 microdeletion syndrome. *J Med Genet* **46**, 480-9 (2009).
176. Nakashima, S. Protein kinase C alpha (PKC alpha): regulation and biological function. *J Biochem* **132**, 669-75 (2002).
177. Braz, J.C. et al. PKC-alpha regulates cardiac contractility and propensity toward heart failure. *Nat Med* **10**, 248-54 (2004).
178. Berrueta, L. et al. The adenomatous polyposis coli-binding protein EB1 is associated with cytoplasmic and spindle microtubules. *Proc Natl Acad Sci U S A* **95**, 10596-601 (1998).
179. Kan, O.M. et al. Mammalian formin Fhod3 plays an essential role in cardiogenesis by organizing myofibrillogenesis. *Biol Open* **1**, 889-96 (2012).
180. Kanaya, H. et al. Fhos2, a novel formin-related actin-organizing protein, probably associates with the nestin intermediate filament. *Genes Cells* **10**, 665-78 (2005).
181. Minakuchi, M. et al. Identification and characterization of SEB, a novel protein that binds to the acute undifferentiated leukemia-associated protein SET. *Eur J Biochem* **268**, 1340-51 (2001).
182. Schinzel, A. & Giedion, A. A syndrome of severe midface retraction, multiple skull anomalies, clubfeet, and cardiac and renal malformations in sibs. *Am J Med Genet* **1**, 361-75 (1978).
183. Wozney, J.M. et al. Novel regulators of bone formation: molecular clones and activities. *Science* **242**, 1528-34 (1988).
184. Zhang, H. & Bradley, A. Mice deficient for BMP2 are nonviable and have defects in amnion/chorion and cardiac development. *Development* **122**, 2977-86 (1996).
185. Izumi, M. et al. Bone morphogenetic protein-2 inhibits serum deprivation-induced apoptosis of neonatal cardiac myocytes through activation of the Smad1 pathway. *J Biol Chem* **276**, 31133-41 (2001).
186. Desjardins, P.R., Burkman, J.M., Shrager, J.B., Allmond, L.A. & Stedman, H.H. Evolutionary implications of three novel members of the human sarcomeric myosin heavy chain gene family. *Mol Biol Evol* **19**, 375-93 (2002).
187. Warkman, A.S. et al. Developmental expression and cardiac transcriptional regulation of Myh7b, a third myosin heavy chain in the vertebrate heart. *Cytoskeleton (Hoboken)* **69**, 324-35 (2012).

188. Nagase, T., Kikuno, R., Ishikawa, K., Hirosawa, M. & Ohara, O. Prediction of the coding sequences of unidentified human genes. XVII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res* **7**, 143-50 (2000).
189. Esposito, T. et al. Digenic mutational inheritance of the integrin alpha 7 and the myosin heavy chain 7B genes causes congenital myopathy with left ventricular non-compact cardiomyopathy. *Orphanet J Rare Dis* **8**, 91 (2013).
190. Polekhina, G., Board, P.G., Gali, R.R., Rossjohn, J. & Parker, M.W. Molecular basis of glutathione synthetase deficiency and a rare gene permutation event. *EMBO J* **18**, 3204-13 (1999).
191. Shi, Z.Z. et al. Mutations in the glutathione synthetase gene cause 5-oxoprolinuria. *Nat Genet* **14**, 361-5 (1996).
192. Olivari, S., Galli, C., Alanen, H., Ruddock, L. & Molinari, M. A novel stress-induced EDEM variant regulating endoplasmic reticulum-associated glycoprotein degradation. *J Biol Chem* **280**, 2424-8 (2005).
193. Mast, S.W. et al. Human EDEM2, a novel homolog of family 47 glycosidases, is involved in ER-associated degradation of glycoproteins. *Glycobiology* **15**, 421-36 (2005).
194. Valero, R. et al. USP25, a novel gene encoding a deubiquitinating enzyme, is located in the gene-poor region 21q11.2. *Genomics* **62**, 395-405 (1999).
195. Valero, R. et al. Characterization of alternatively spliced products and tissue-specific isoforms of USP28 and USP25. *Genome Biol* **2**, RESEARCH0043 (2001).
196. Abbaszade, I. et al. Cloning and characterization of ADAMTS11, an aggrecanase from the ADAMTS family. *J Biol Chem* **274**, 23443-50 (1999).
197. Didangelos, A., Mayr, U., Monaco, C. & Mayr, M. Novel role of ADAMTS-5 protein in proteoglycan turnover and lipoprotein retention in atherosclerosis. *J Biol Chem* **287**, 19341-5 (2012).
198. Dupuis, L.E. et al. Altered versican cleavage in ADAMTS5 deficient mice; a novel etiology of myxomatous valve disease. *Dev Biol* **357**, 152-64 (2011).

Supplementary Figures

Fig. S1. Layout of study design.

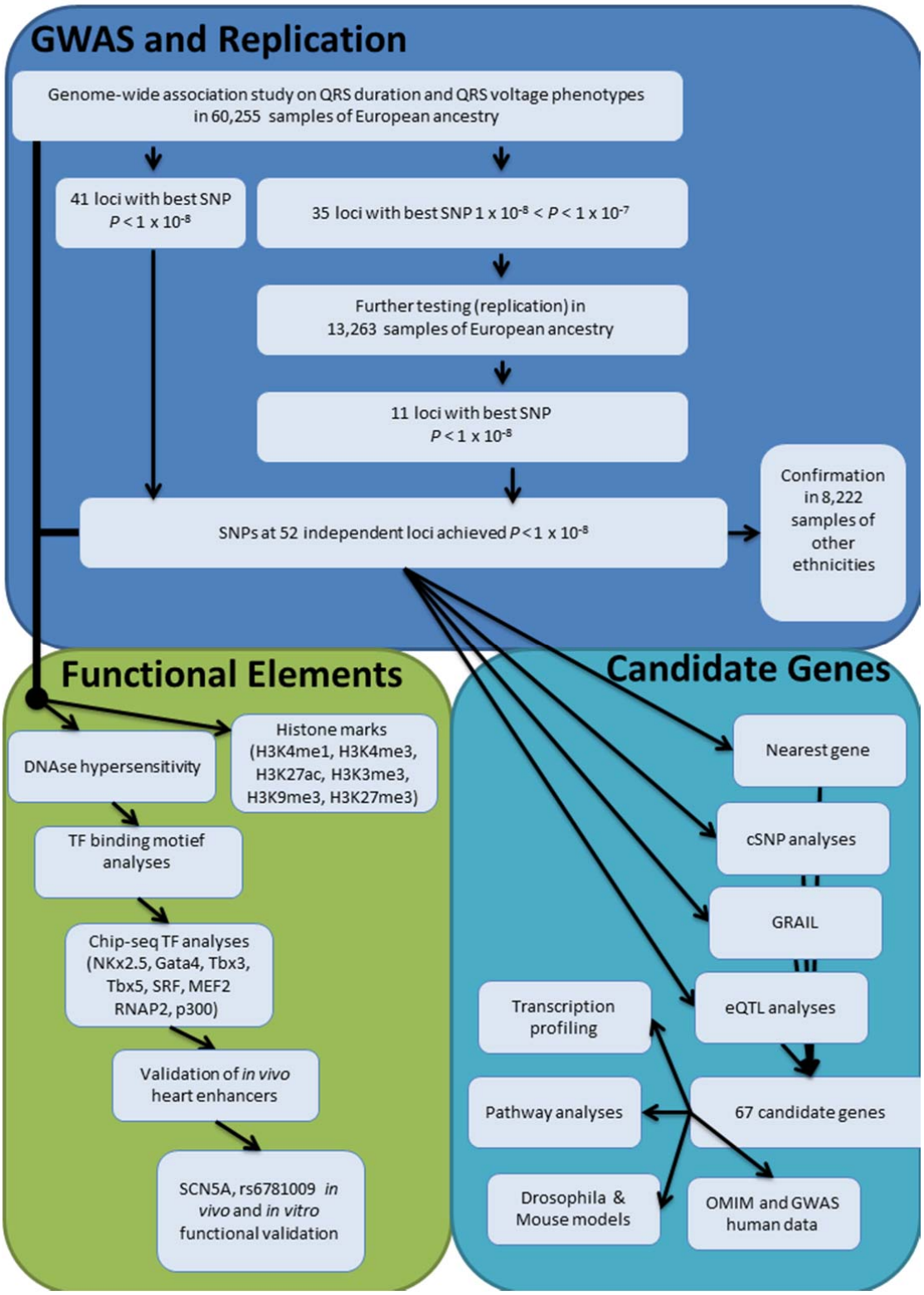
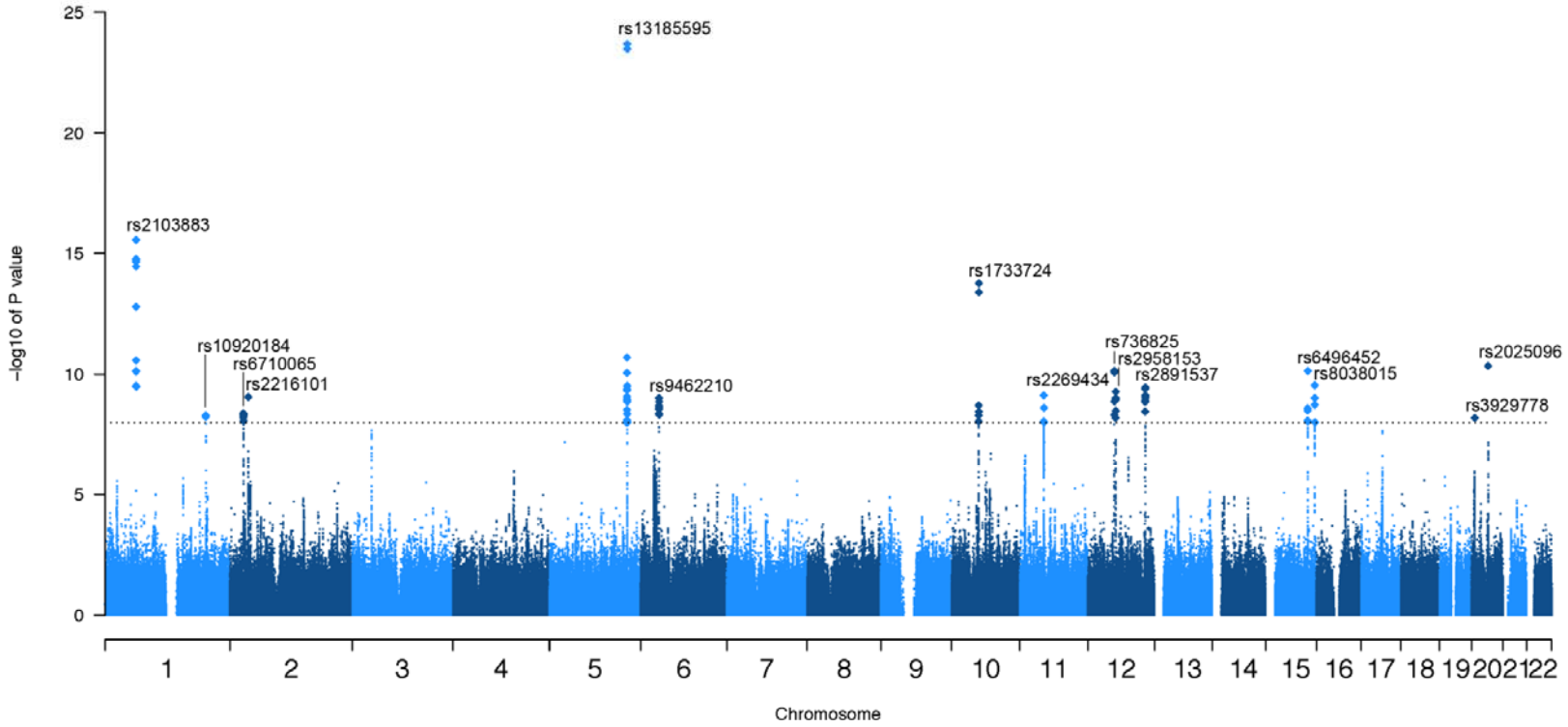


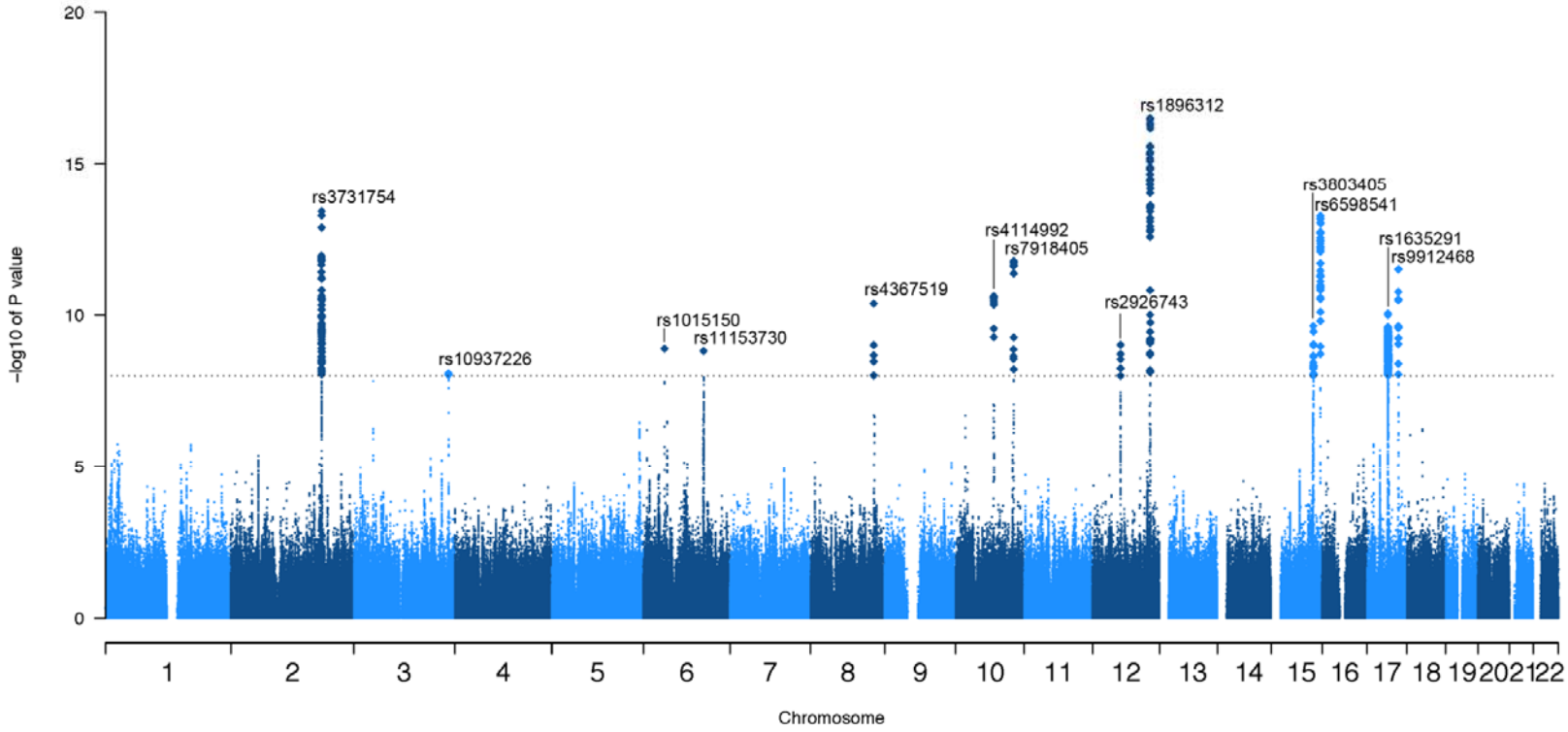
Fig. S2. Manhattan plots per QRS trait

SF2.1 to SF2.4: Manhattan plots showing the results for genome-wide association with QRS traits amongst Europeans. SNPs reaching genome-wide significance ($P<1\times10^{-8}$).

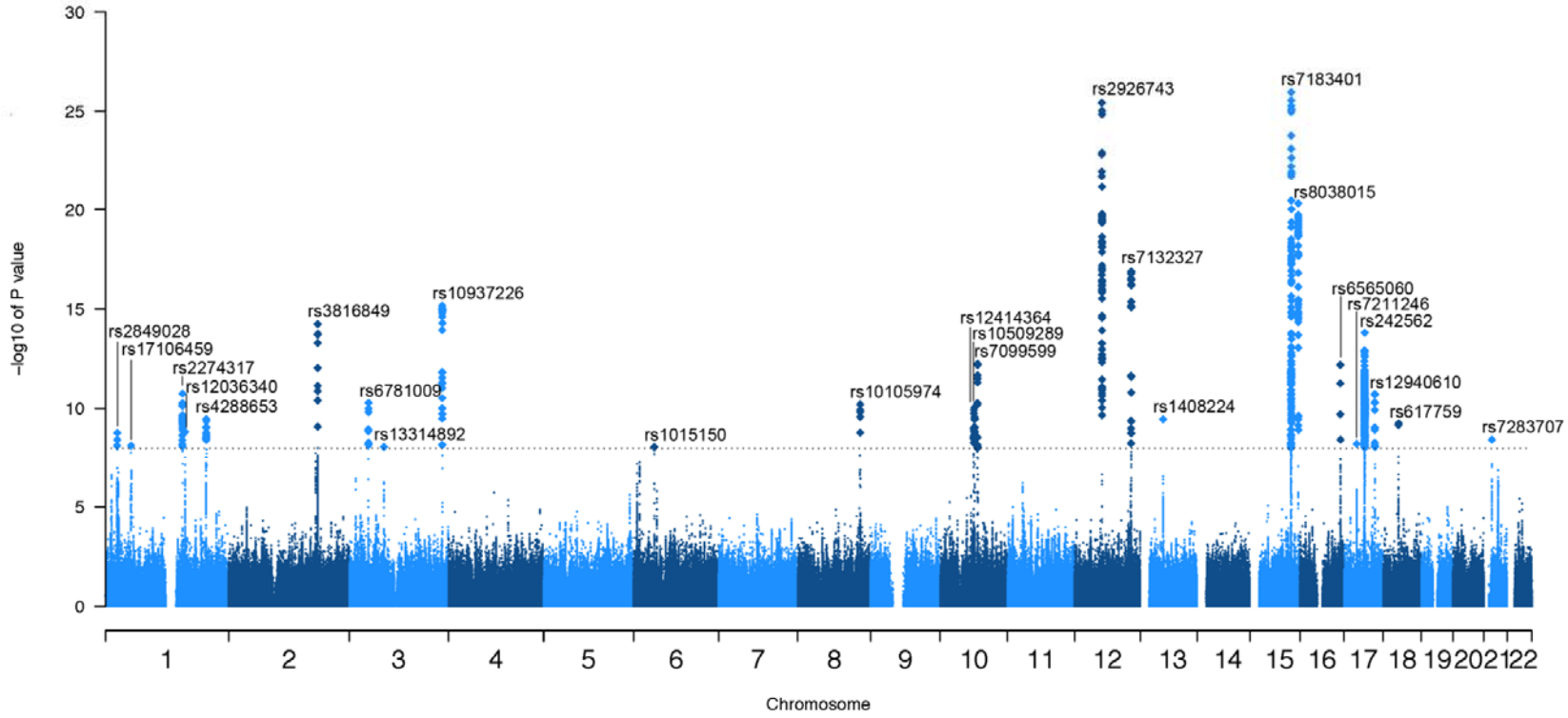
SF2.1: Cornell



SF2.2: Sokolow-Lyon



SF2.3: 12-leadsum



SF2.4: QRS duration

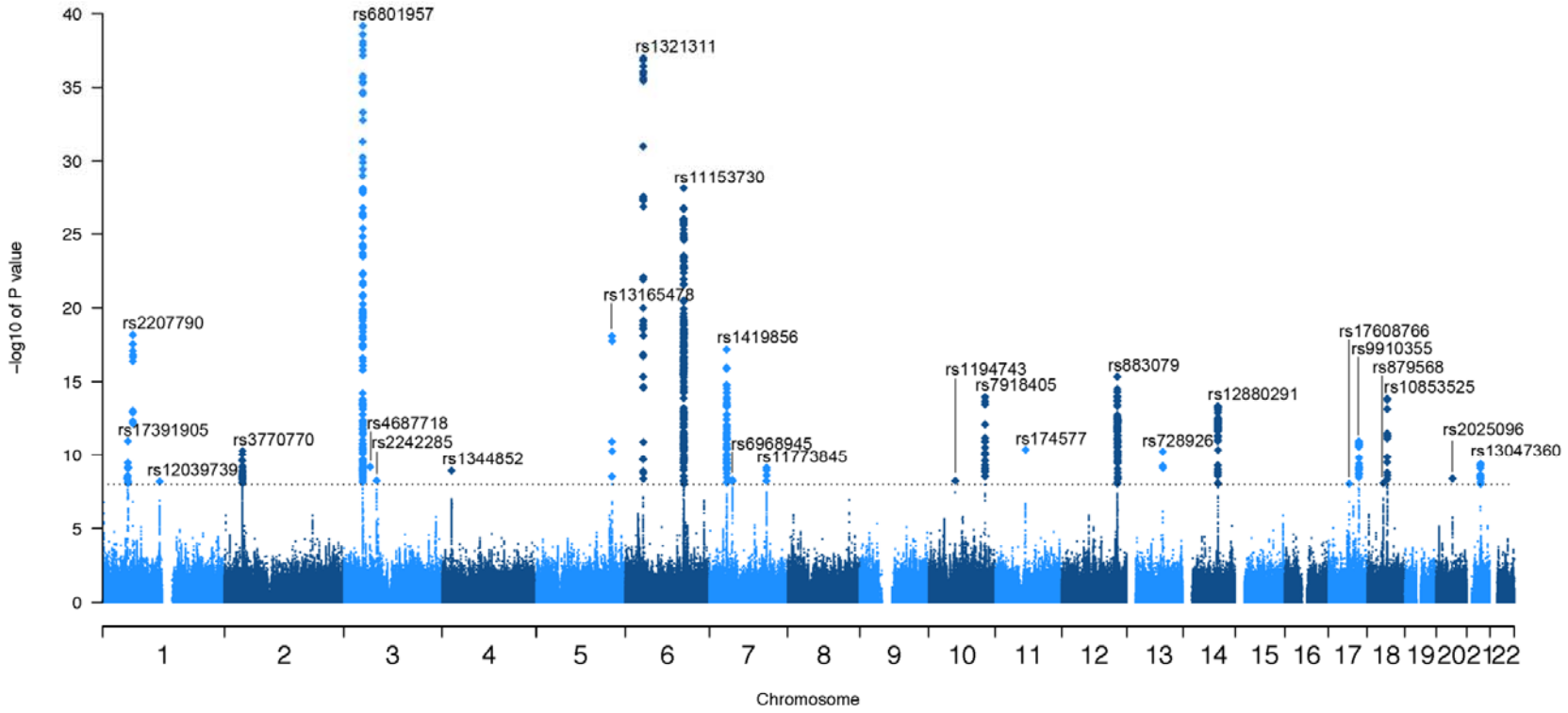
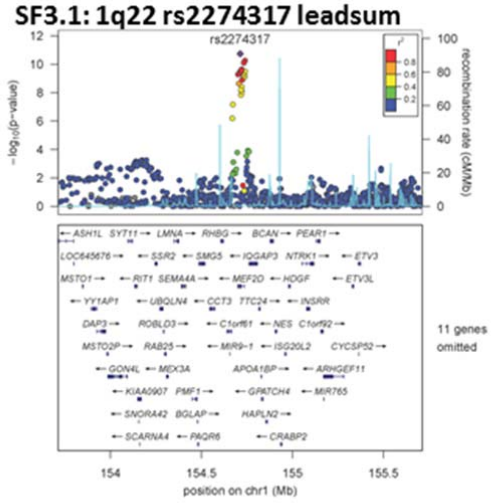


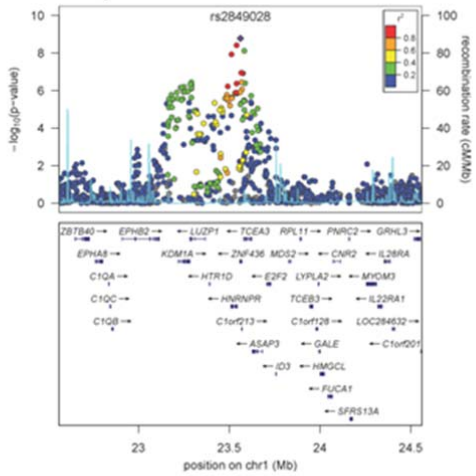
Fig. S3. Regional plots

SF3.1 to SF3.52: regional plots for the QRS trait phenotype sentinel SNPs. At each region pairwise LD with the sentinel SNP is indicated.

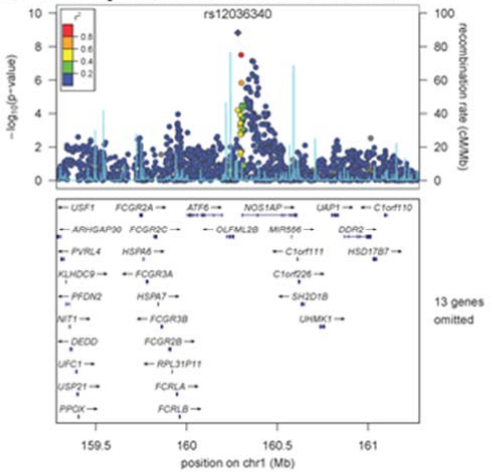
SF3.1: 1q22 rs2274317 leadsum



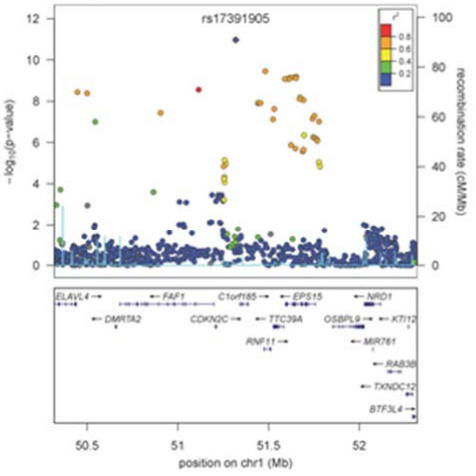
SF3.2: 1p36.12 rs2849028 leadsum



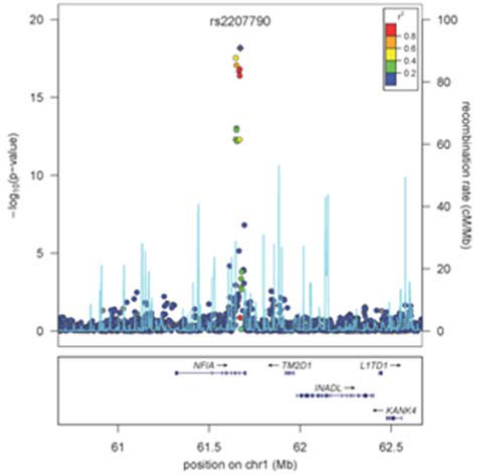
SF3.3: 1q23.3 rs12036340 leadsum



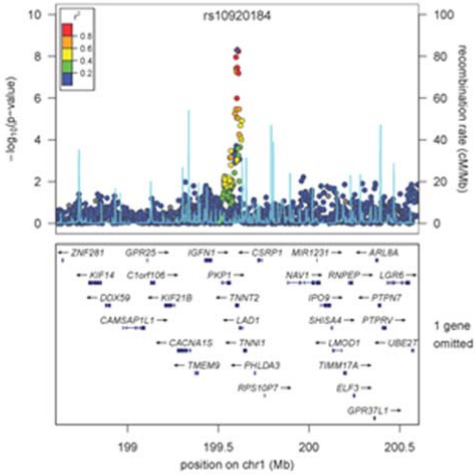
SF3.4: 1p32.3 rs17391905 duration



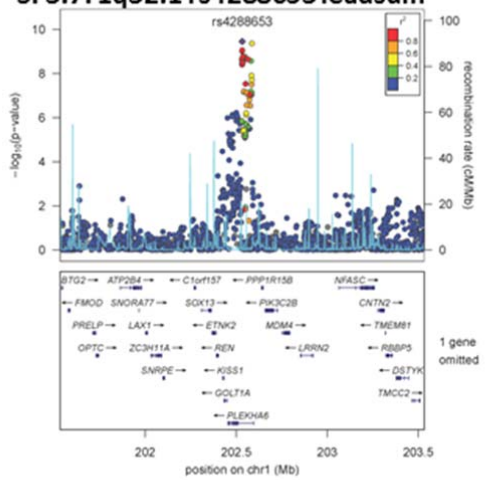
SF3.5: 1p31.3 rs2207790 duration



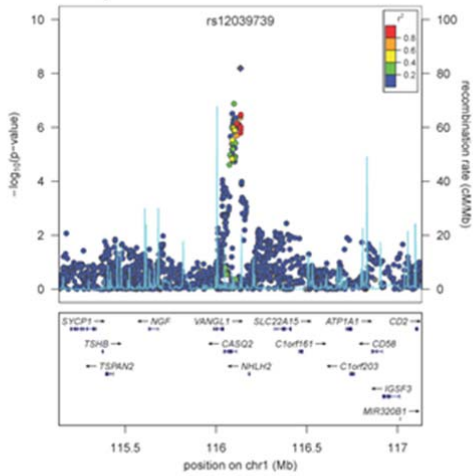
SF3.6: 1q32.1 rs10920184 cornell



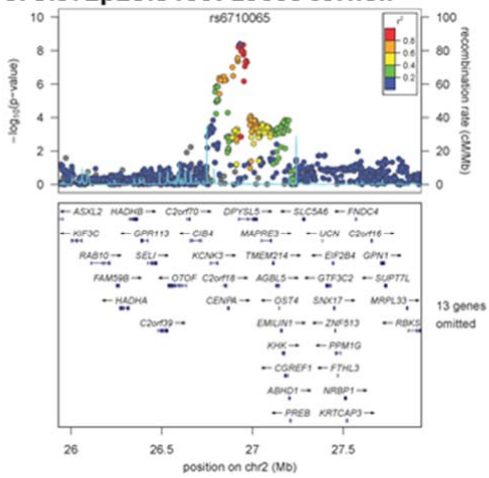
SF3.7: 1q32.1 rs4288653 leadsum



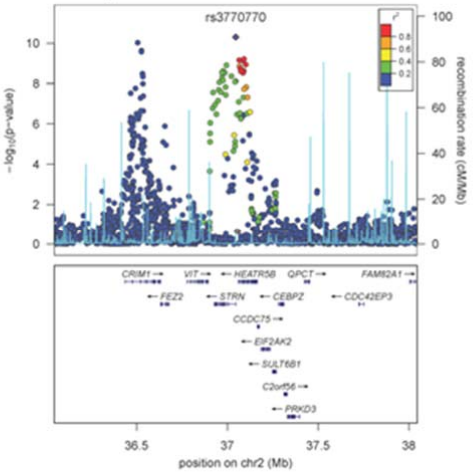
SF3.8: 1p13.1 rs12039739 duration



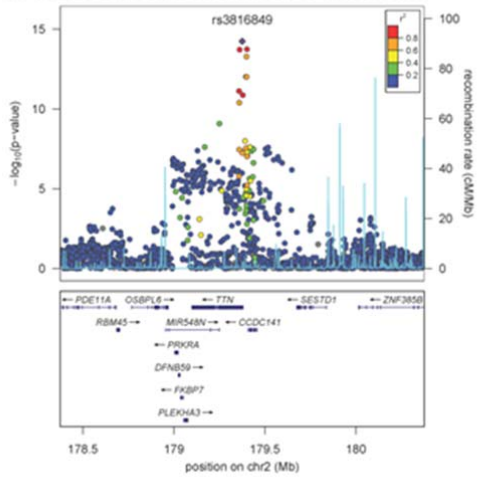
SF3.9: 2p23.3 rs6710065 cornell



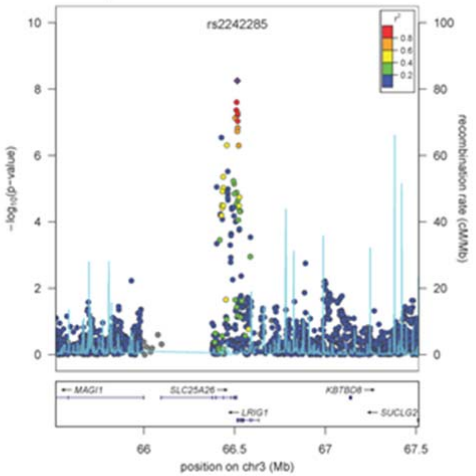
SF3.10: 2p22.2 rs3770770 duration



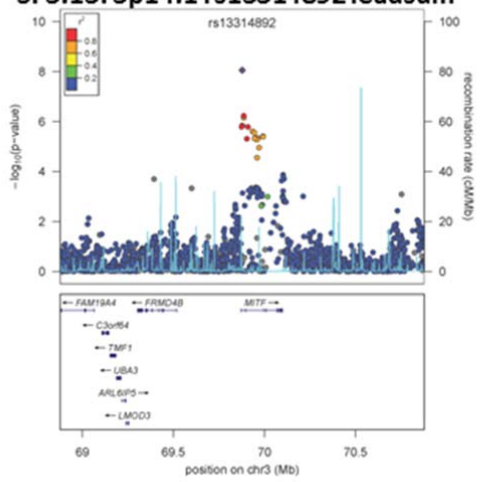
SF3.11: 2q31.2 rs3816849 leadsum



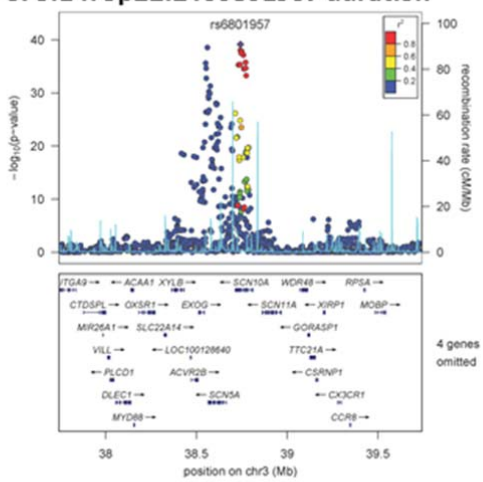
SF3.12: 3p14.1 rs2242285 duration



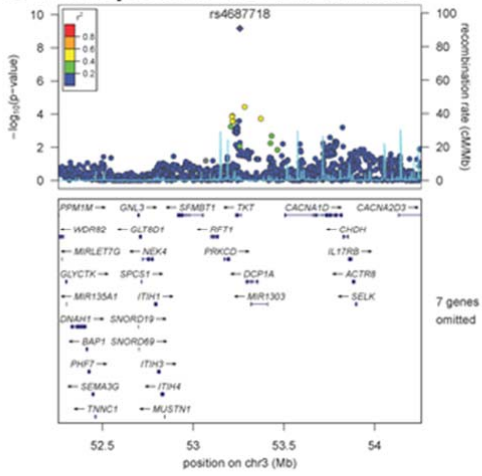
SF3.13: 3p14.1 rs13314892 leadsum



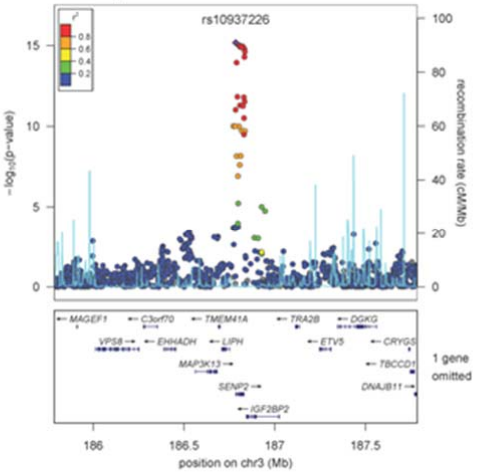
SF3.14: 3p22.2 rs6801957 duration



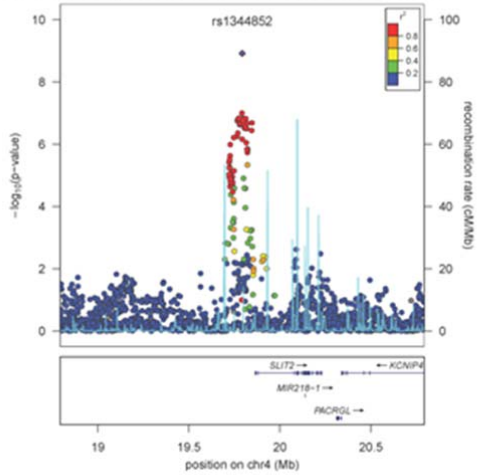
SF3.15: 3p21.1 rs4687718 duration



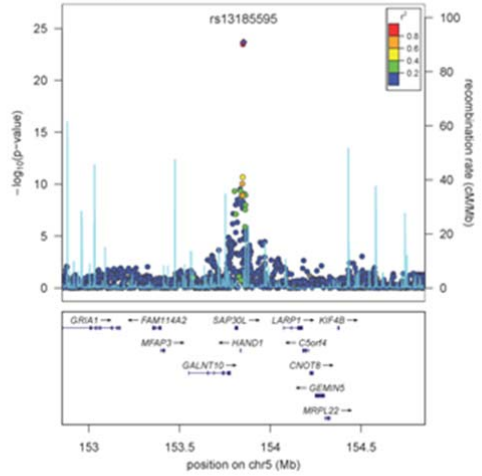
SF3.16: 3q27.2 rs10937226 leadsum



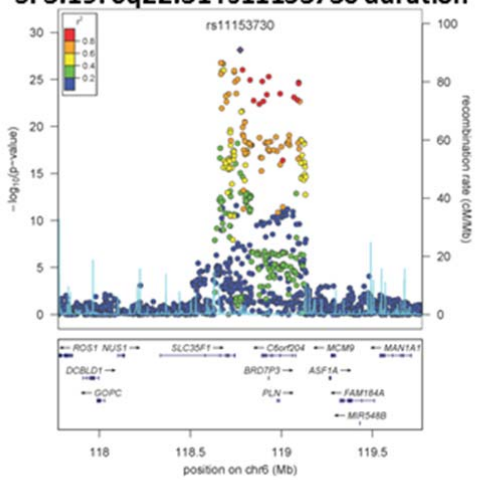
SF3.17: 4p15.31 rs1344852 duration



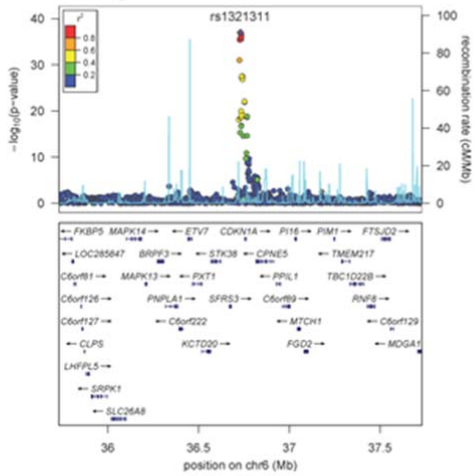
SF3.18: 5q33.2 rs13185595 cornell



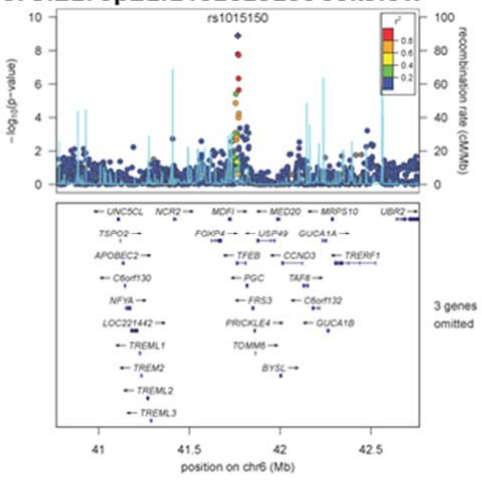
SF3.19: 6q22.31 rs11153730 duration



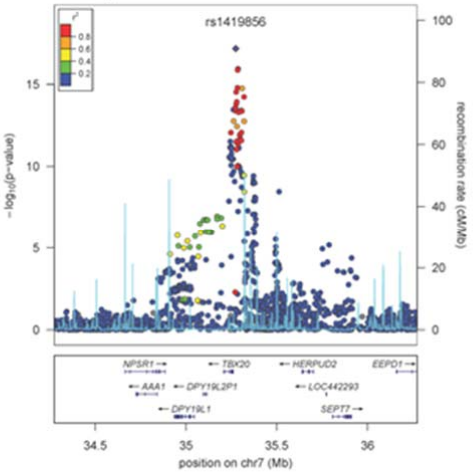
SF3.20: 6p21.2 rs1321311 duration



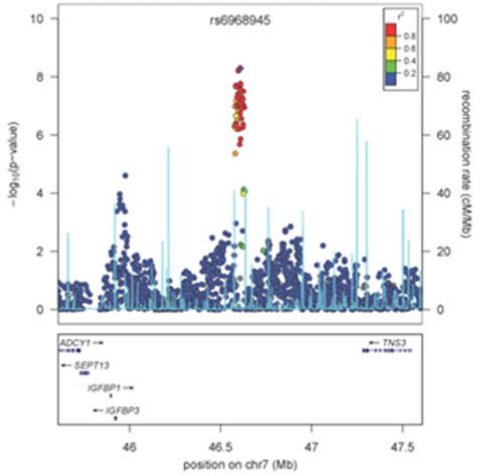
SF3.21: 6p21.1 rs1015150 sokolow



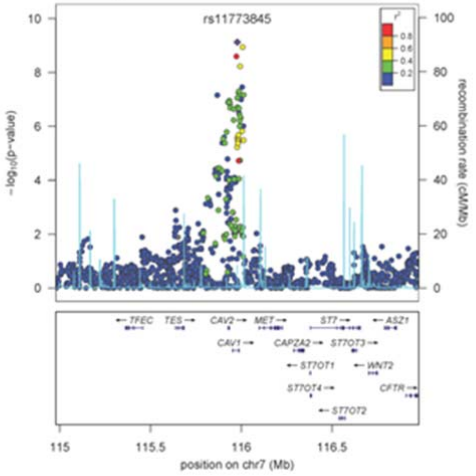
SF3.22: 7p14.2 rs1419856 duration



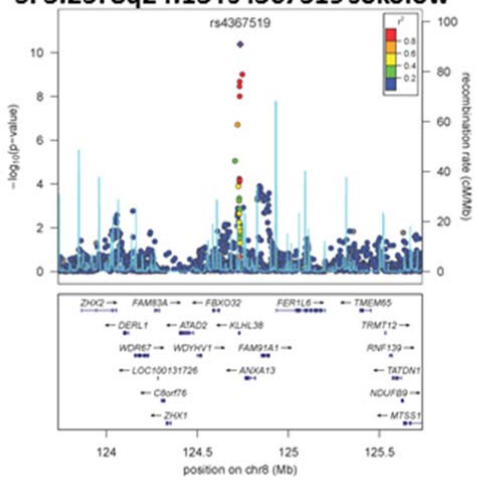
SF3.23: 7p12.3 rs6968945 duration



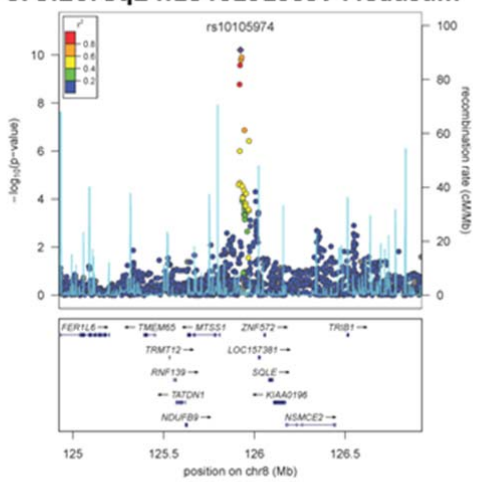
SF3.24: 7q31.2 rs11773845 duration



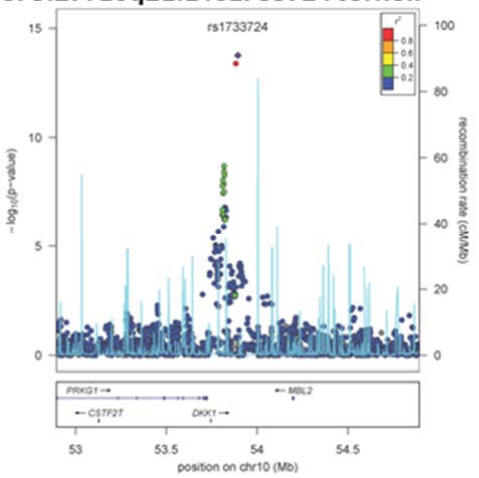
SF3.25: 8q24.13 rs4367519 sokolow



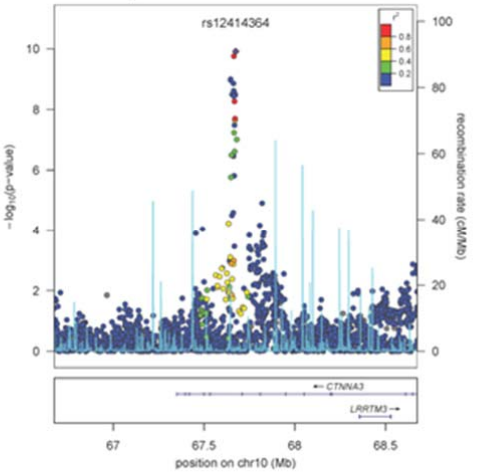
SF3.26: 8q24.13 rs10105974 leadsum



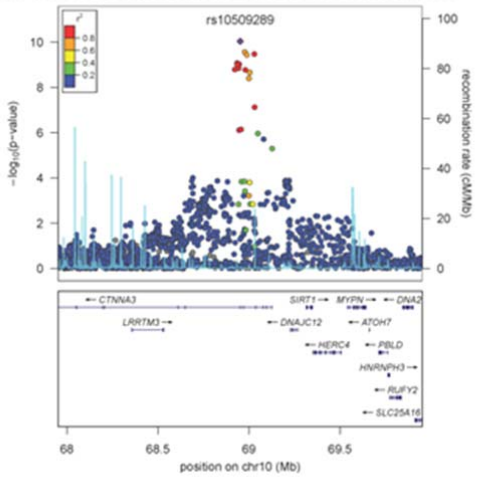
SF3.27: 10q21.1 rs1733724 cornell



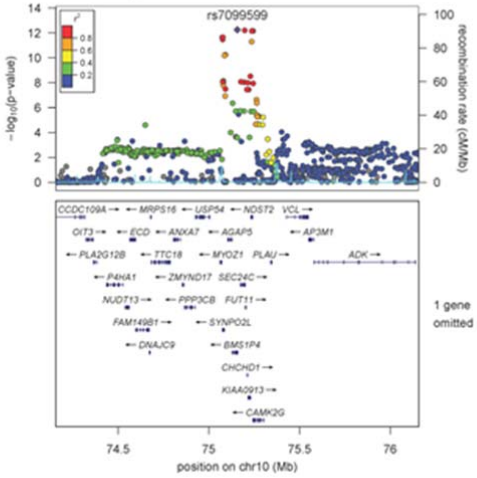
SF3.28: 10q21.3 rs12414364 leadsum



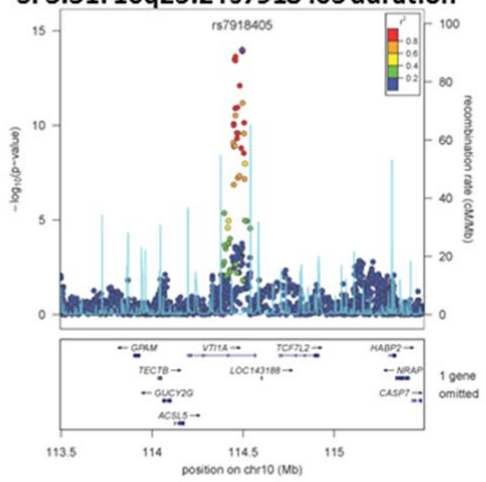
SF3.29: 10q21.3 rs10509289 leadsum



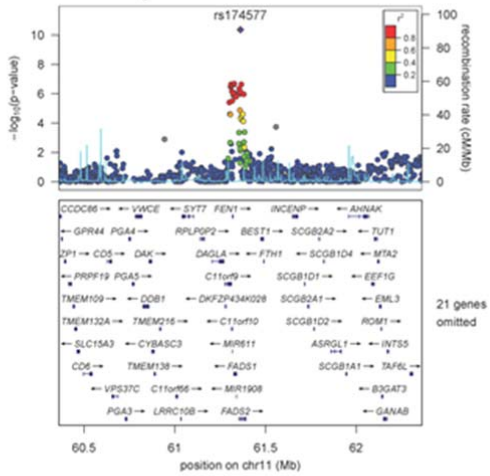
SF3.30: 10q22.2 rs7099599 leadsum



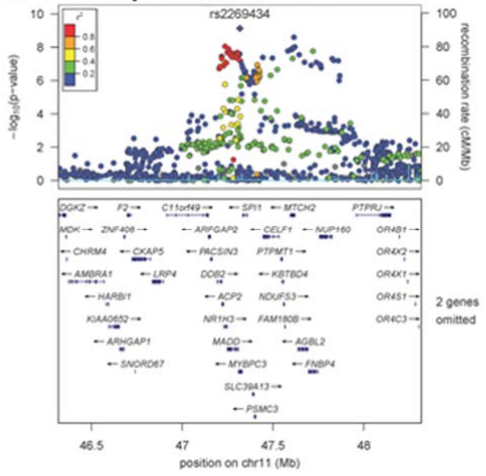
SF3.31: 10q25.2 rs7918405 duration



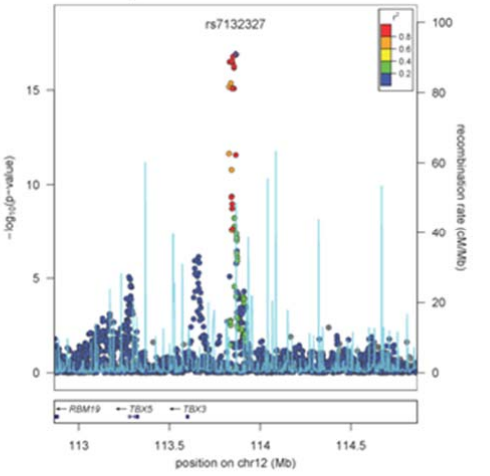
SF3.32: 11q12.2 rs174577 duration



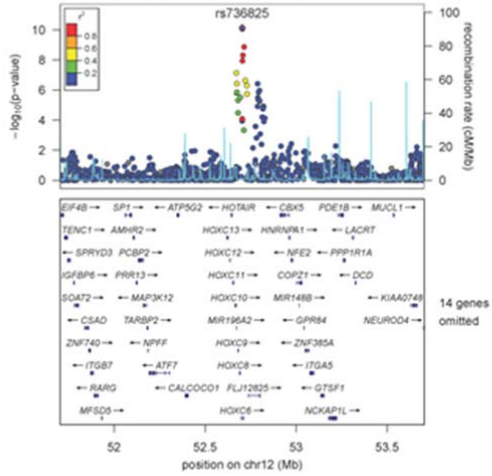
SF3.33: 11p11.2 rs2269434 cornell



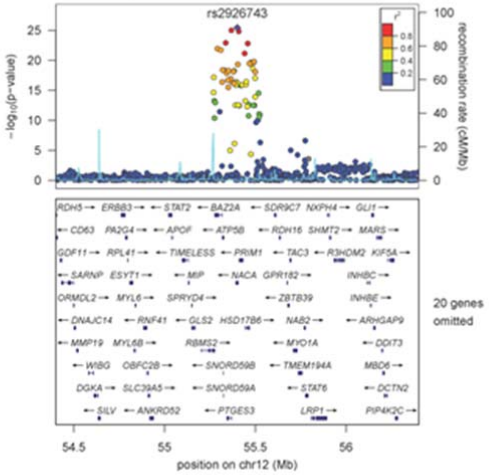
SF3.34: 12q24.21 rs7132327 leadsum



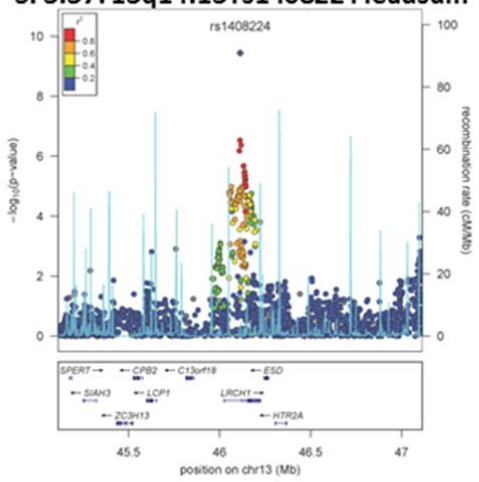
SF3.35: 12q13.13 rs736825 cornell



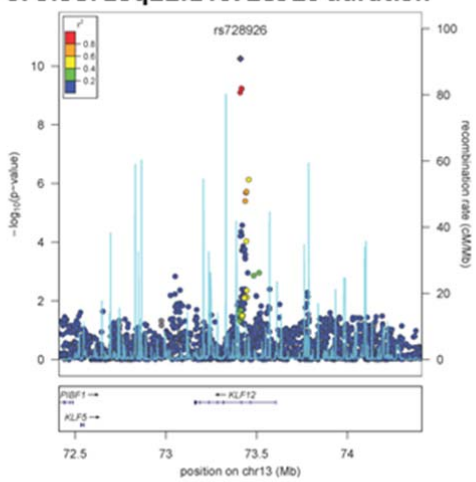
SF3.36: 12q13.3 rs2926743 leadsum



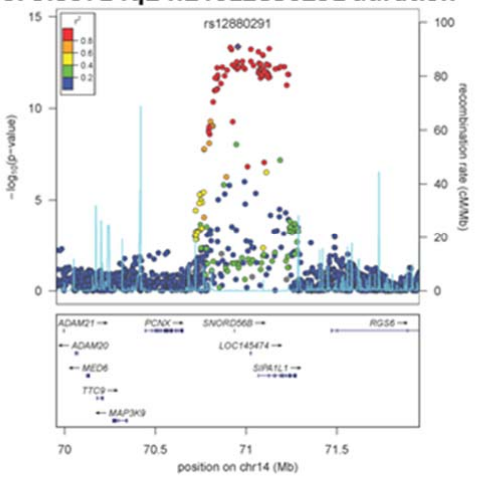
SF3.37: 13q14.13 rs1408224 leadsum



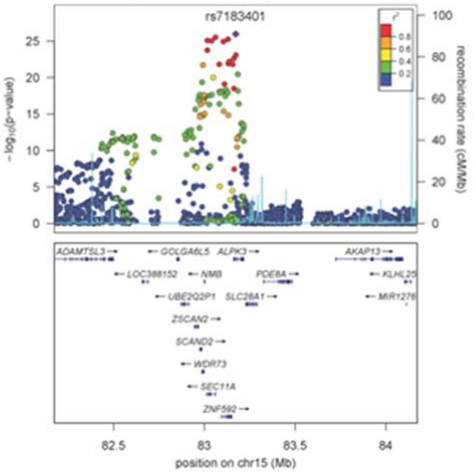
SF3.38: 13q22.1 rs728926 duration



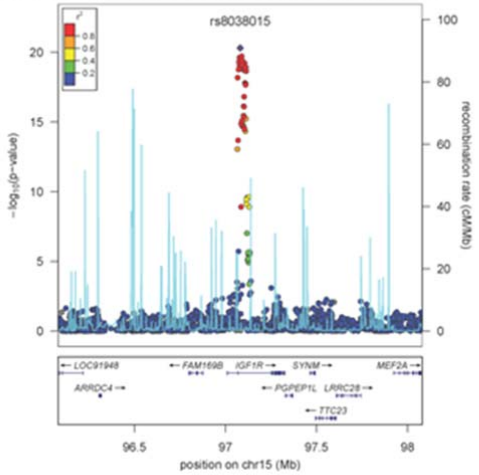
SF3.39: 14q24.2 rs12880291 duration



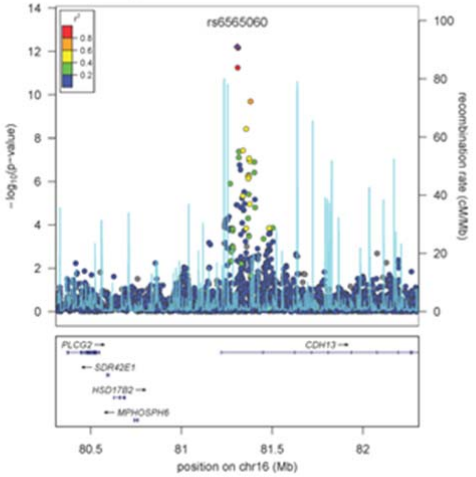
SF3.40: 15q25.3 rs7183401 leadsum



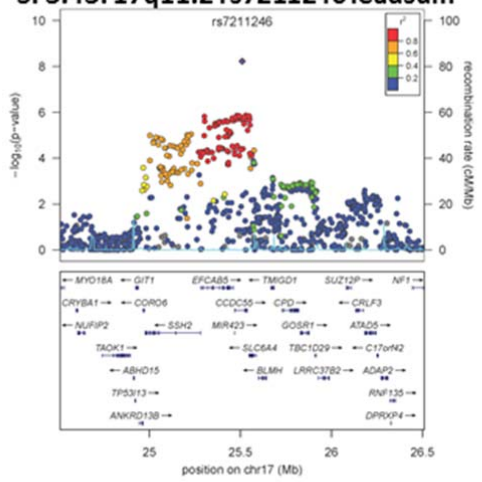
SF3.41: 15q26.3 rs8038015 leadsum



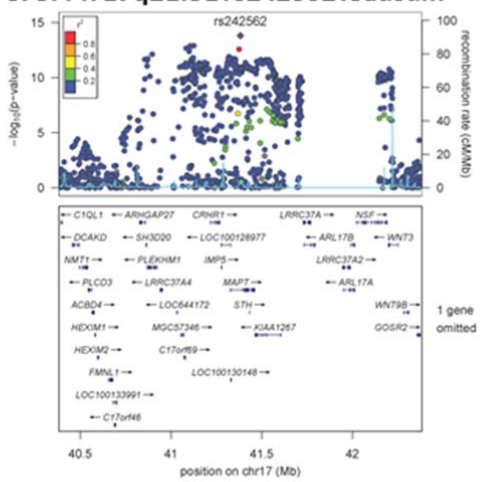
SF3.42: 16q23.3 rs6565060 leadsum



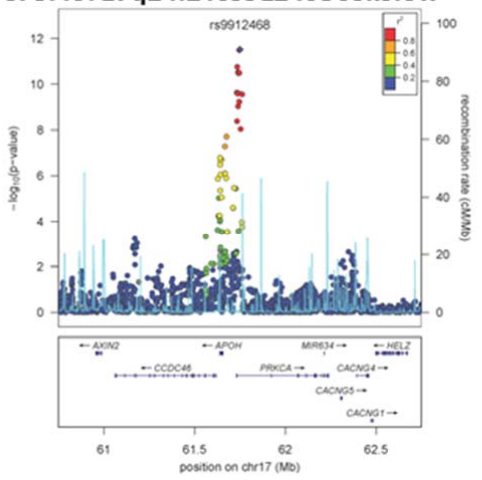
SF3.43: 17q11.2 rs7211246 leadsum



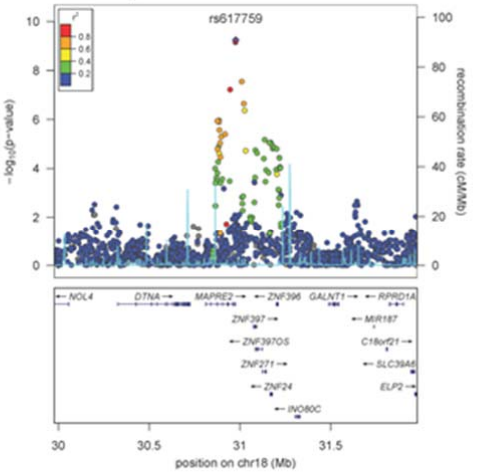
SF3.44: 17q21.31 rs242562 leadsum



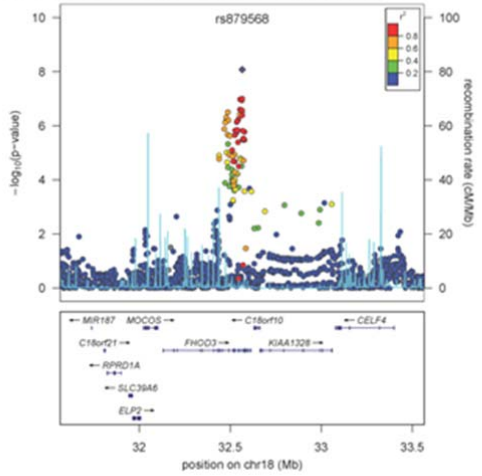
SF3.45: 17q24.2 rs9912468 sokolow



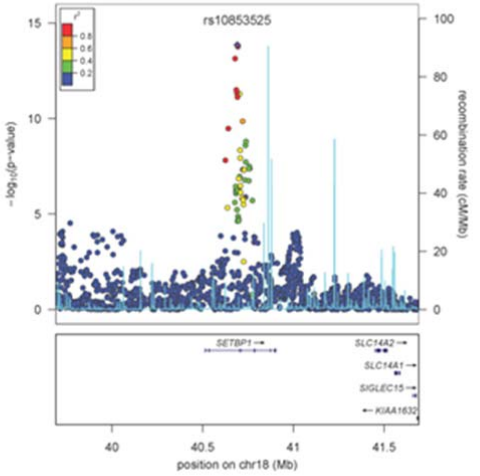
SF3.46: 18q12.2 rs617759 leadsum



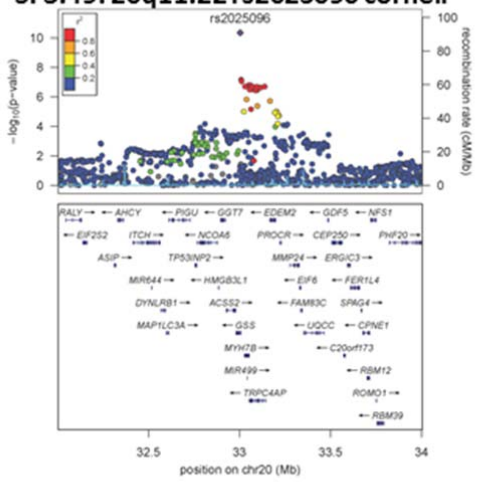
SF3.47: 18q12.2 rs879568 duration



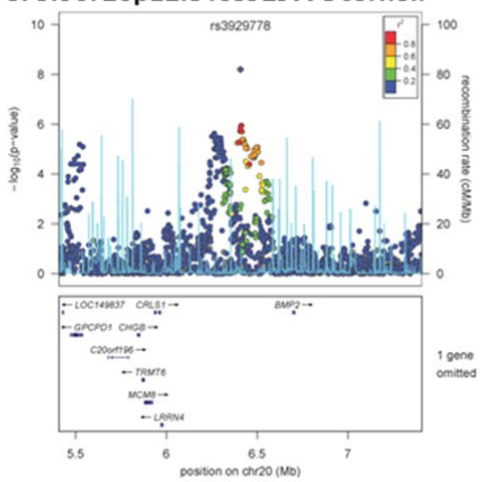
SF3.48: 18q12.3 rs10853525 duration



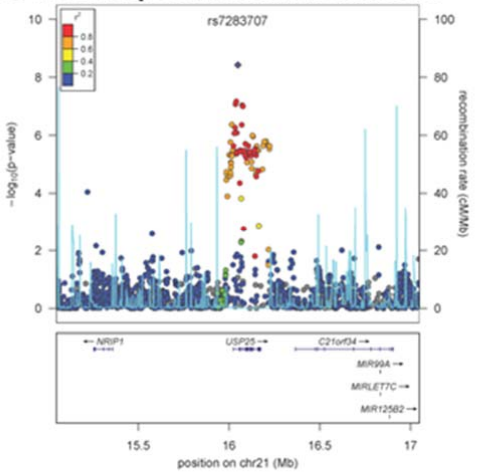
SF3.49: 20q11.22 rs2025096 cornell



SF3.50: 20p12.3 rs3929778 cornell



SF3.51: 21q21.1 rs7283707 leadsum



SF3.52: 21q21.3 rs13047360 duration

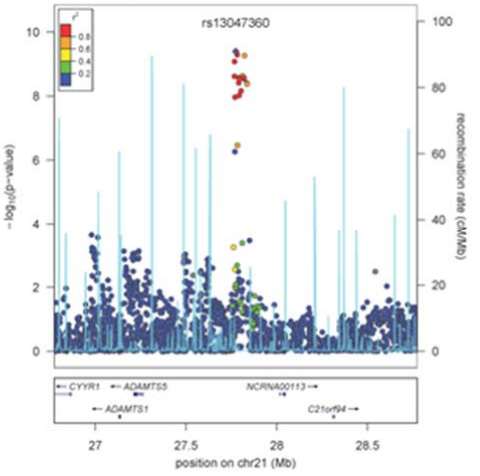


Fig. S4. Venn diagram on the overlap of genetic loci among the 4 QRS traits

Venn diagram shows the distribution of the 32 loci associated with a single trait and the 20 loci associated with two or more phenotypes. All locus-phenotype associations are also presented in table S6.

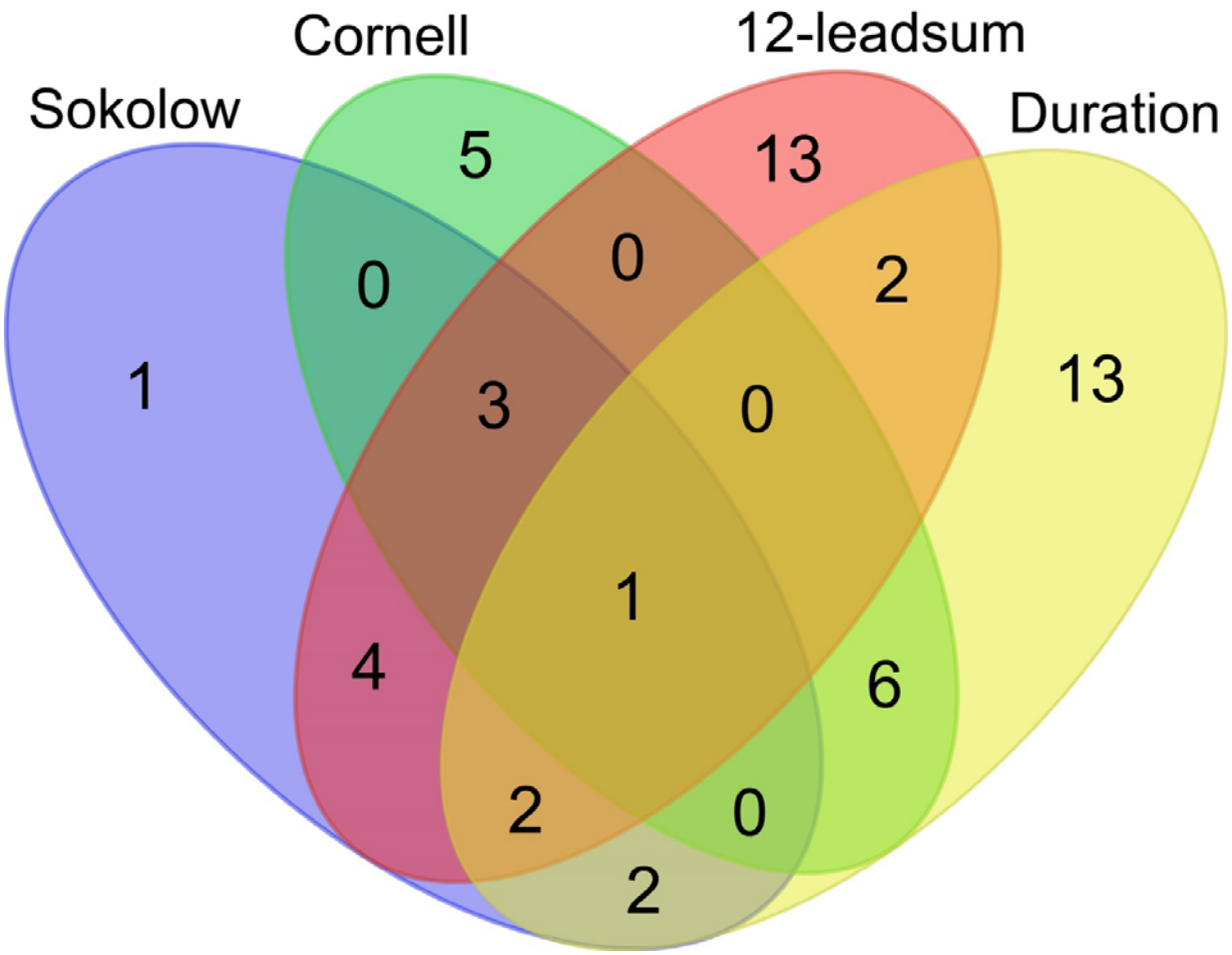


Fig. S5. Enrichment of chromatin states in human fetal heart tissue

To capture the greater complexity we performed an integrative analysis in an 15-state ChromHMM model representative of different functional regions of the genome. The left panel shows the enrichment of the 52 loci for the 15-state model using the five available core histone marks for human fetal heart tissue. The right panel shows the total number of the 52 loci overlapped by each feature.

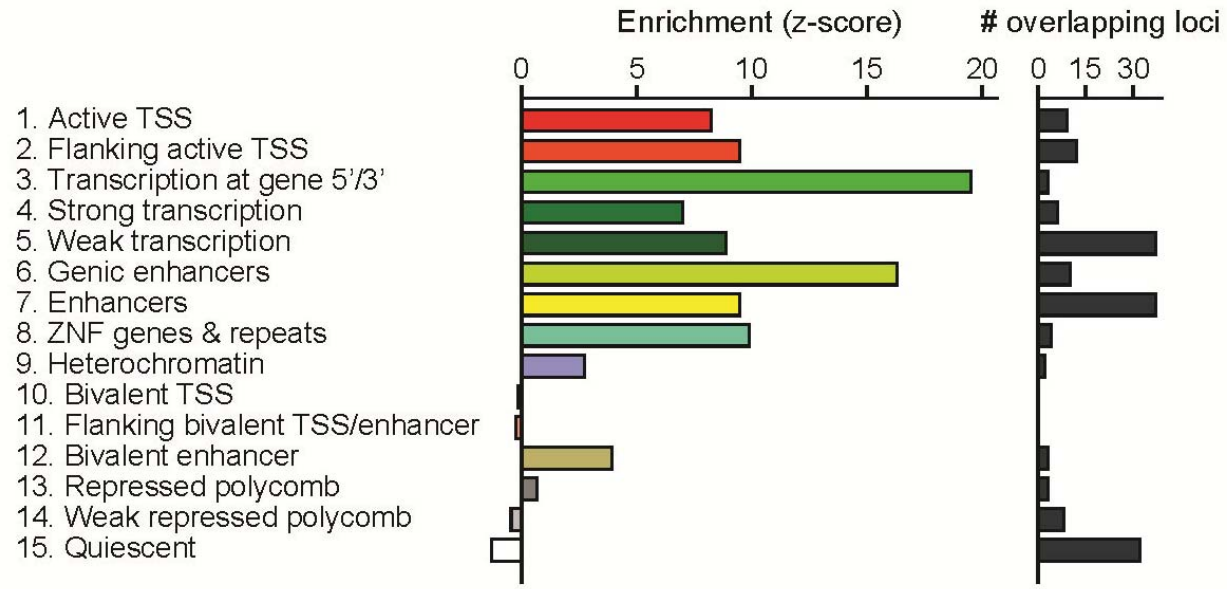


Fig. S6. Histone modifications during cardiomyocyte differentiation.

Enrichment of the 52 loci for histone modifications during cardiomyocyte differentiation (mouse). Enhancers are annotated by H3K4me1 peaks at least +/- 1kb away from an annotated TSS and designated as active or poised based on the presence (active) or absence (poised) of H3K27ac.

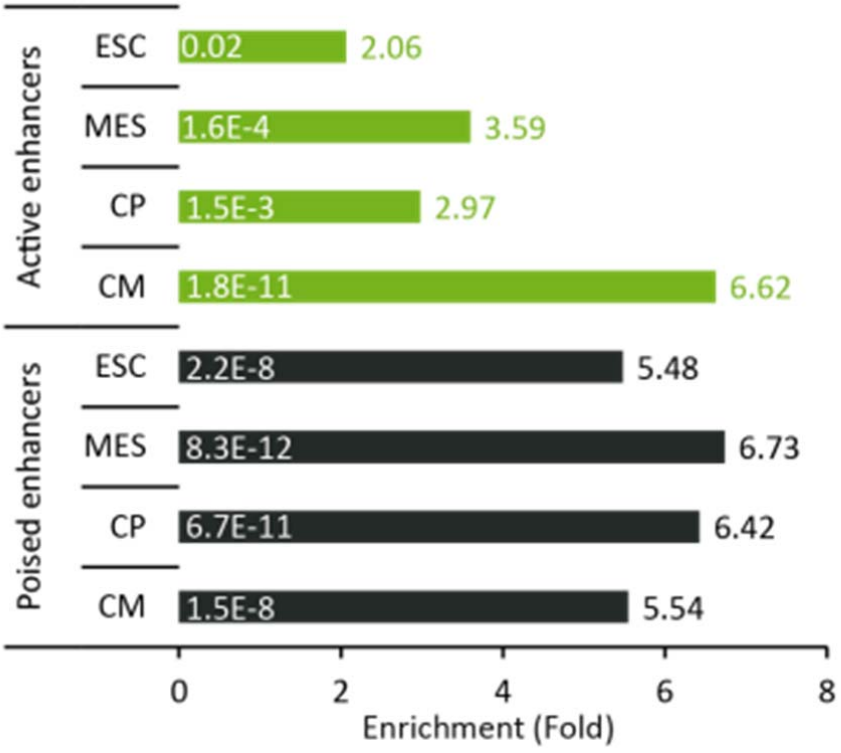


Fig. S7. Gene-expression data of candidate vs non-candidate genes.

Within heart and muscle tissue microarray-based data the 63 (of 67) available candidate genes are significantly more highly expressed as compared to non-candidate genes.

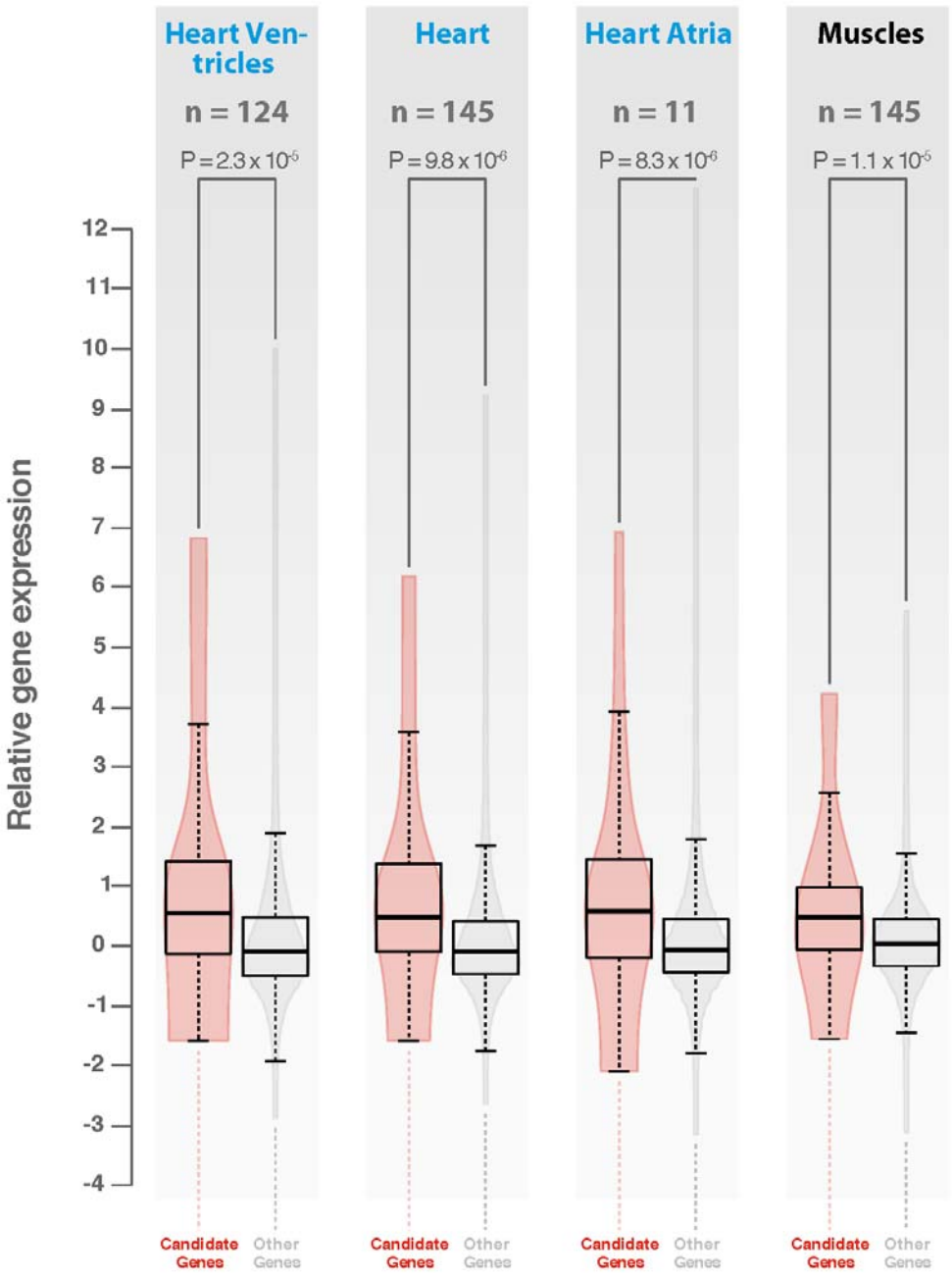


Fig. S8. Gene expression patterns of candidate genes across different tissues

Unsupervised hierarchical clustering of microarray-based expression levels of the candidate genes of 40 different tissues reveals that several genes are showing relatively high expression in heart and muscles.

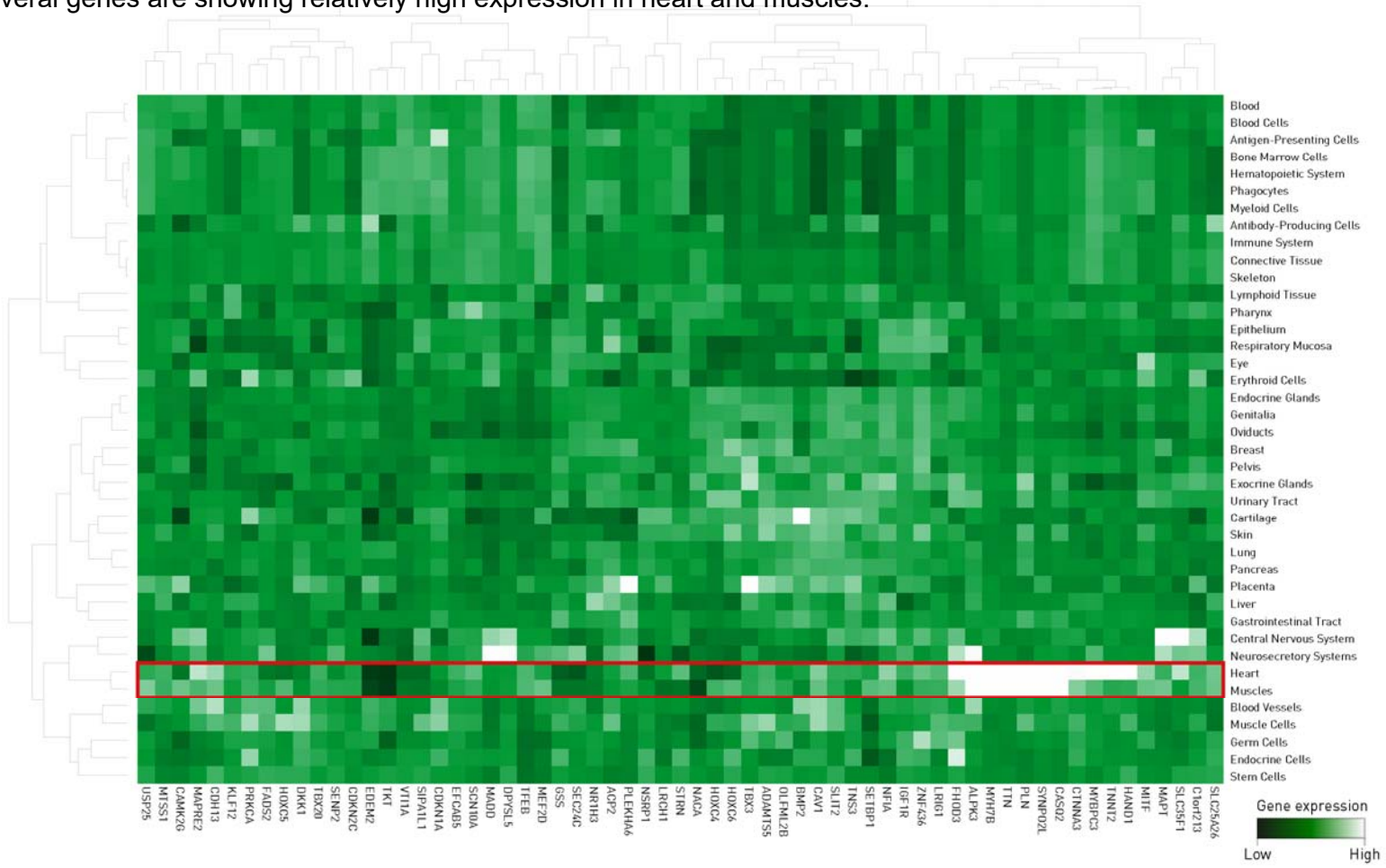


Fig. S9. Gene-expression during cardiomyocyte differentiation.

(a) The 54 (of 67) available candidate genes are highly expressed in RNA-seq data of cardiomyocytes, compared to non-candidate genes. (b) Unsupervised hierarchical clustering of RNA-seq based expression data of 54 candidate genes in 4 different cardiomyocyte (precursors) reveals that most of the genes are abundantly expressed in cardiomyocytes.

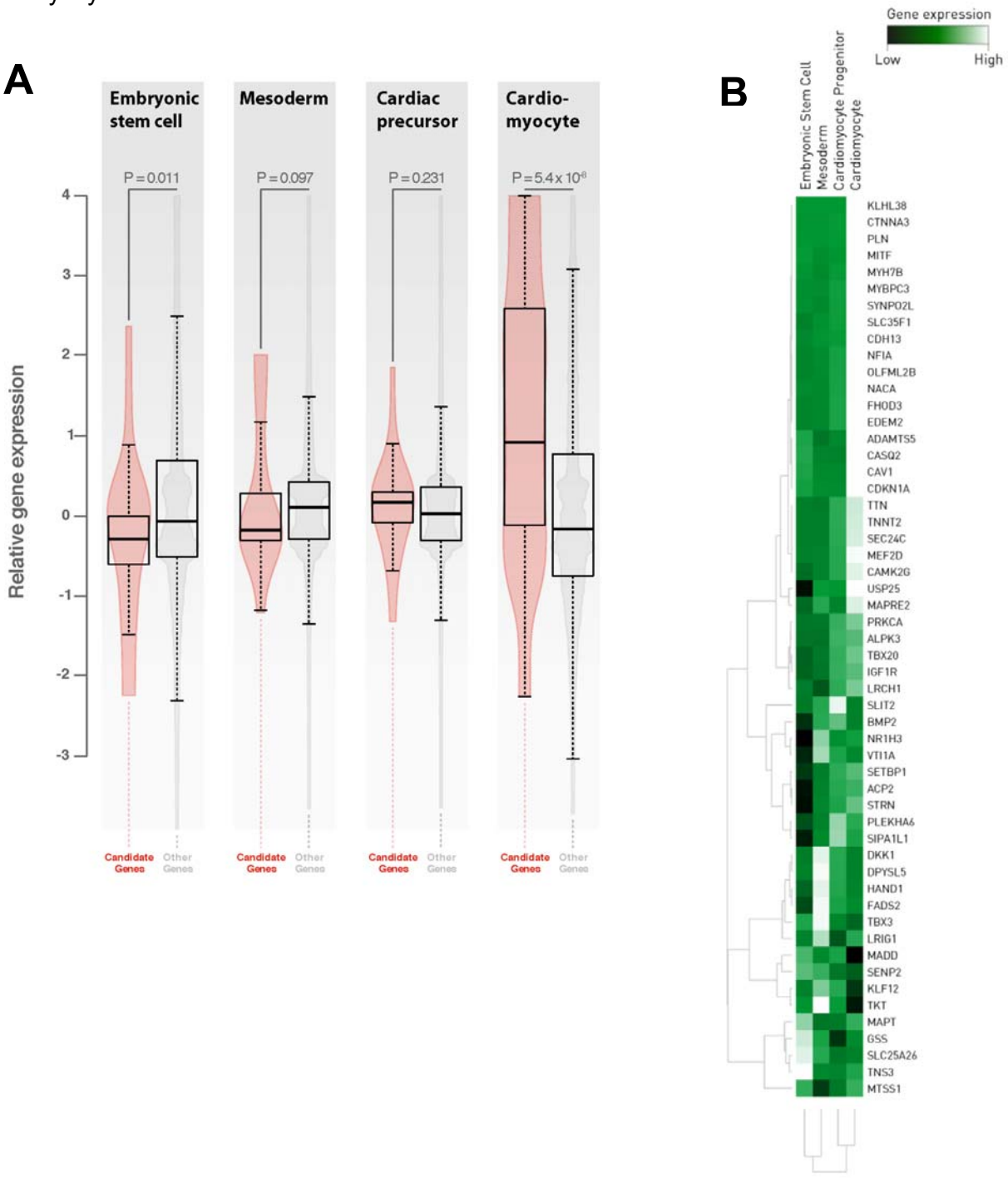


Fig. S10. Expected and observed fly and mouse models with cardiac phenotypes

(**a**) Expected results of 1M permutations of 67 genes lead on average to cardiac phenotypes in 3.97 (SD 1.93) D. Melanogaster flies. We observed 9 Drosophila models (orthologues of SLIT2, NR1H3, HAND1, MYH7B, TTN, SLC25A26, FHOD3, NACA, and STRN) with cardiac phenotypes for our 67 candidate genes ($P=1.84\times10^{-2}$, obtained by using a normal distribution approximation with the abovementioned mean and standard deviation. We find $P=1.69\times10^{-2}$ without applying a normal distribution approximation.) (**b**) Expected results of 1M permutations of 67 genes lead on average to cardiac phenotypes in 3.46 (SD 1.81) mice. We observed 18 mice with cardiac phenotypes for our 67 candidate genes ($P=3.4\times10^{-14}$ can only be obtained by using a normal approximation as the simulation did not report any event with 18 or more overlapping mouse orthologues.).

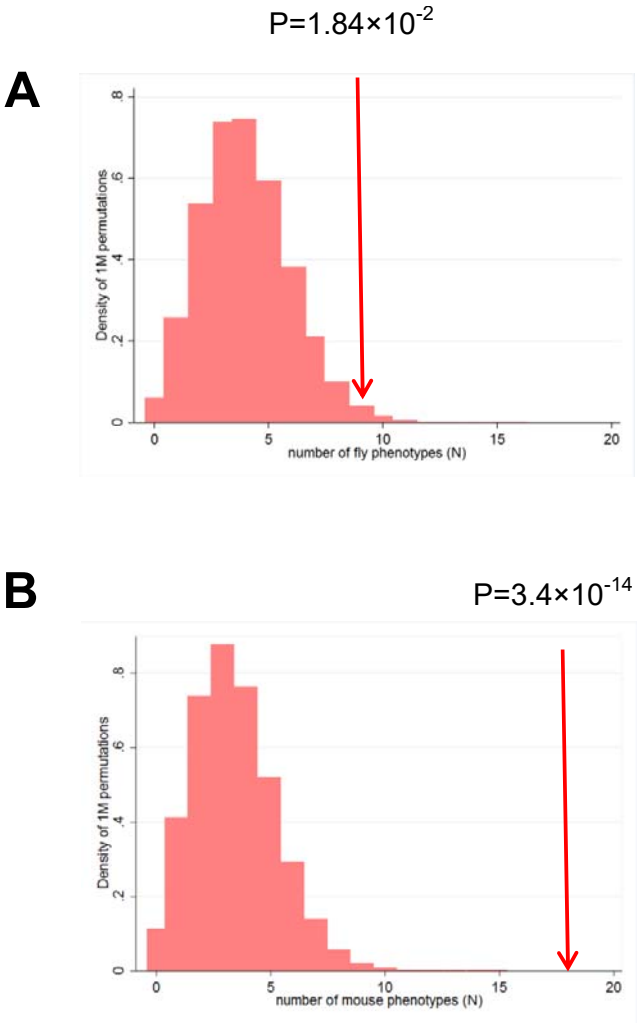


Fig. S11. eQTL and DNA functional data in the *NR1H3-MYBDPC3* region

Analysis of the 11p11.2 locus implicating NR1H3 as an additional candidate gene, next to MYBPC3, on the basis of strong eQTL, open chromatin and histone modifications. Diamond represents the lead SNP (rs2269434) in the region with in colour circles nearby SNPs and their LD to the sentinel and on the right y-axis the –log10(p-value) of association. In gray-shaded diamond and circles the eQTL association with NR1H3 with its corresponding p-value on the left y-axis. In the line-graph the tested functional elements are plotted for left ventricular tissue and foetal heart. At the bottom the genes are plotted.

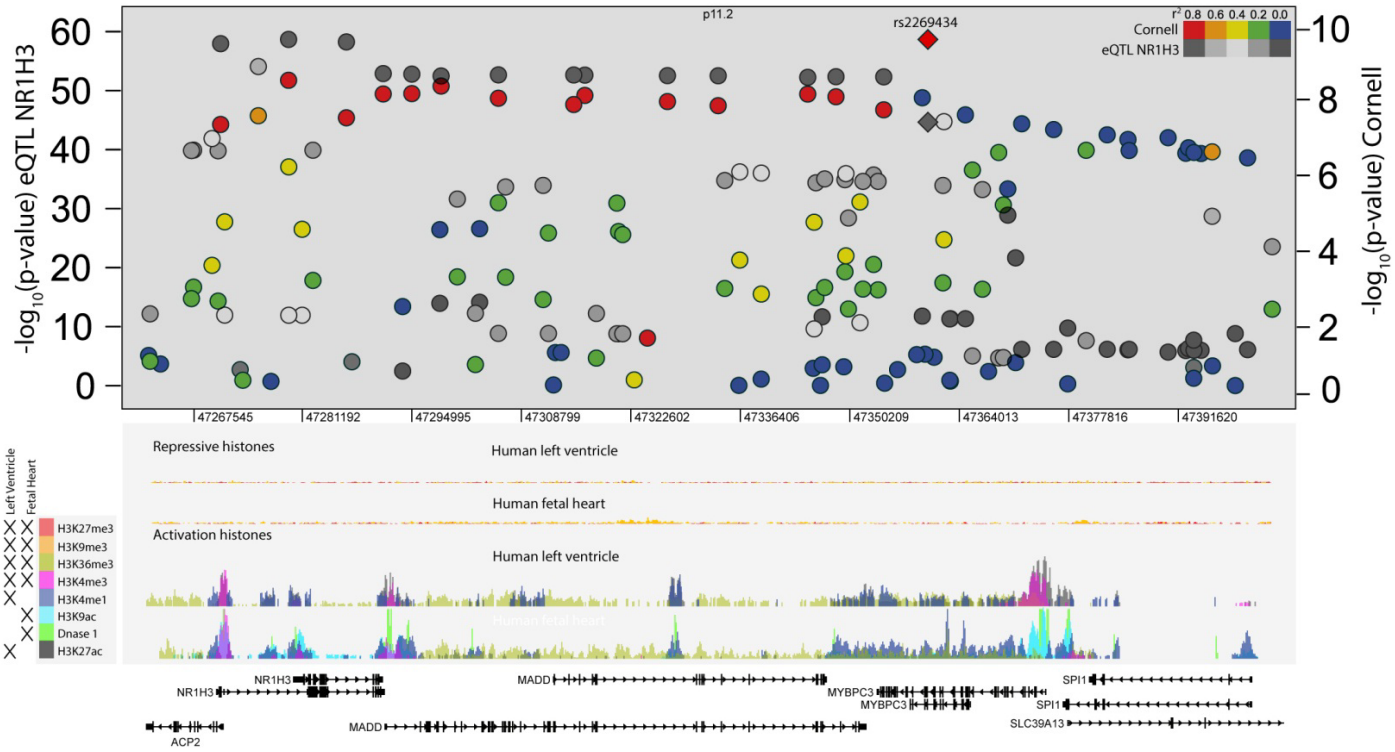


Fig. S12. DEPICT correlation structure within left ventricular dilatation meta-gene set

Example meta gene set “Left Ventricular Dilatation” consisting of 23 individual reconstituted gene sets. Gene sets are represented by nodes colored according to statistical significance, and similarities between them are indicated by edges scaled according to their correlation (only correlations with r>0.3 are shown).

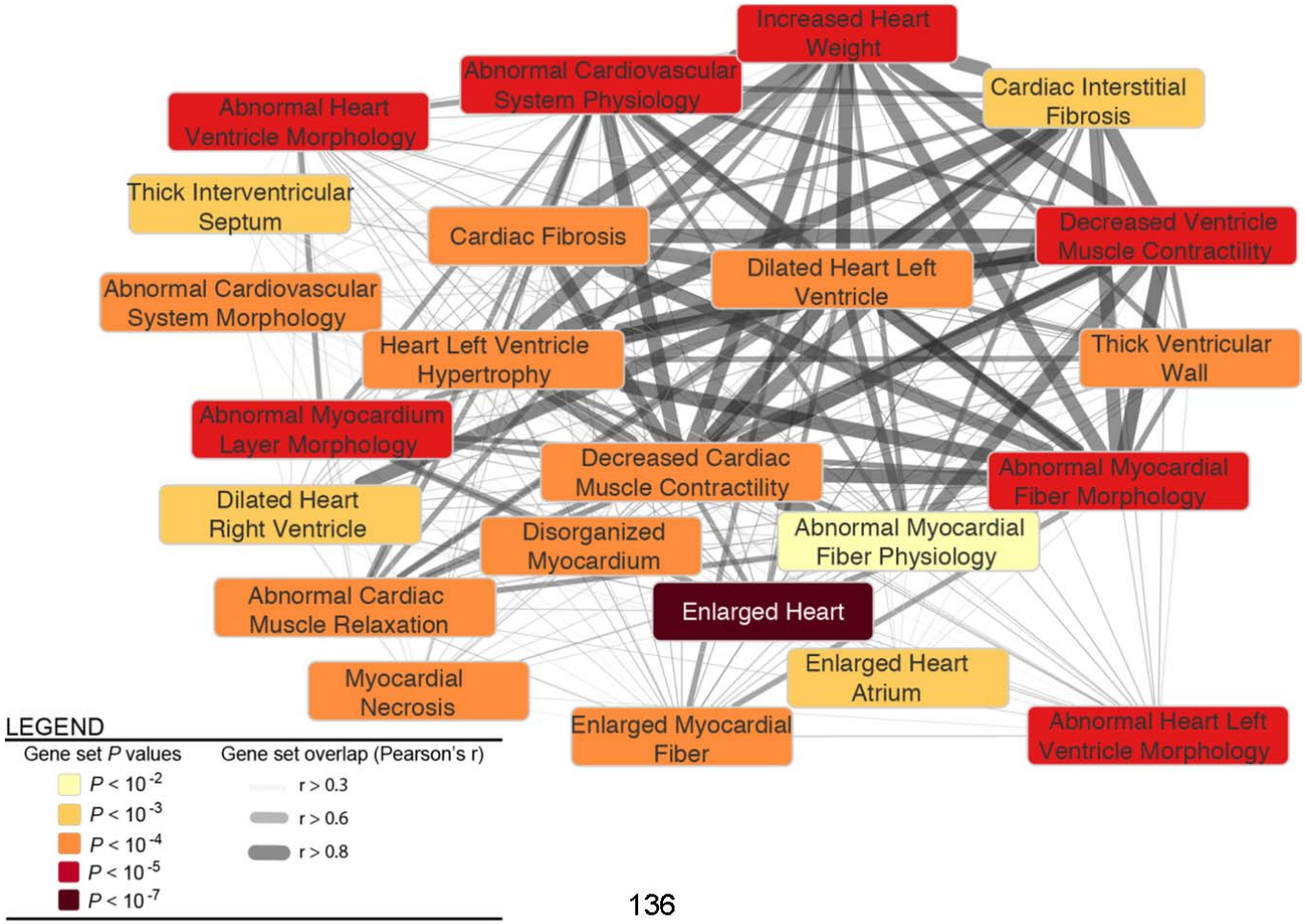


Fig. S13. Further examples of cardiac *in vivo* enhancers

Four cardiac in vivo enhancer correspond with our earlier work and have been published. For hs1912 (chr1:3251956-3256225) the embryo image is available[27]. We here provide representative embryo's for 3 additional elements which have not been shown before[23,27].

hs1932
chr15:97075263-97079751

hs1959
chr16:81325110-81327548

hs2078
chr12:52698505-52701712