Understanding the heritability of heart rhythm and conduction disorders

CLAUDIA TAMAR SILVA ALDANA

NOTA DE SALVEDAD DE RESPONSABILIDAD INSTITUCIONAL

"La Universidad del Rosario no se hace responsable por los conceptos emitidos por los investigadores en su trabajo, solo velará por el rigor científico, metodológico y ético del mismo en aras de la búsqueda de la verdad y la justicia"

Understanding the heritability of heart rhythm and conduction disorders

Promotoren

Prof.dr.ir. C.M van Duijn

Prof.dr. R. Willemsen

Prof.dr. C.M. Restrepo

Copromotor

Dr. A. Isaacs

Erasmus MC
Universitair Medisch Centrum Rotterdam

Universidad del Rosario

To my biggest love on the earth: my husband and daughters

CONTENTS

PUBLICATIONS AND MANUSCRIPTS BASED ON THE STUDIES DESCRIBED IN THIS THESIS

Chapter 2

Claudia Tamar Silva, Jan A. Kors, Ph.D.4, Najaf Amin, Abbas Dehghan, Jacqueline C.M. Witteman, Rob Willemsen, Ben A. Oostra, Cornelia M. Van Duijn, Aaron Isaacs.

Heretabilities, proportion of heritabilities explained by GWAS findings, and implications of cross-phenotype effects of PR interval.

Hum Genet. 2015 Nov; 134 (11-12):1211-9

Chapter 3

Claudia Tamar Silva, Irina V. Zorkoltseva, Najaf Amin, Ayşe Demirkan, Elisa van Leeuwen, Jan A. Kors, Marten van den Berg, Bruno H. Stricker, André G. Uitterlinden, Anatoly V. Kirichenko, Jacqueline C.M. Witteman, Rob Willemsen, Ben A. Oostra, Tatiana I. Axenovich, Cornelia M. van Duijn, Aaron Isaacs.

A combined linkage and exome sequencing analysis for ECG parameters in the Erasmus Rucphen Family study

Front. Genet. 7:190. doi: 10.3389/fgene.2016.00190

Chapter 4

Claudia Tamar Silva, Irina V. Zorkoltseva, Maartje N. Niemeijer, Marten E. van den Berg, Najaf Amin, Ayşe Demirkan, Elisa van Leeuwen, Adriana I. Iglesias, Laura B. Piñeros-Hernández, Carlos M. Restrepo, Jan A. Kors, Anatoly V. Kirichenko, Rob Willemsen, Ben A. Oostra, Bruno H. Stricker, André G. Uitterlinden, Tatiana I. Axenovich, Cornelia M. van Duijn and Aaron Isaacs A combined

linkage, microarray and exome analysis suggests MAP3K11 as a candidate gene for left ventricular hypertrophy

BMC Medical Genomics BMC series 2018 Mar 5;11(1):22. doi: 10.1186/s12920-018-0339-9.

Chapter 5

van der Harst P, van Setten J, Verweij N, Vogler G, Franke L, Maurano MT, Wang X, Mateo Leach I, Eijgelsheim M, Sotoodehnia N, Hayward C, Sorice R, Meirelles O, Lyytikäinen LP, Polašek O, Tanaka T, Arking DE, Ulivi S, Trompet S, Müller-Nurasyid M, Smith AV, Dörr M, Kerr KF, Magnani JW, Del Greco M F, Zhang W, Nolte IM, Silva CT, Padmanabhan S, Tragante V, Esko T, Abecasis GR, Adriaens ME, Andersen K, Barnett P, Bis JC, Bodmer R, Buckley BM, Campbell H, Cannon MV, Chakravarti A, Chen LY, Delitala A, Devereux RB, Doevendans PA, Dominiczak AF, Ferrucci L, Ford I, Gieger C, Harris TB, Haugen E, Heinig M, Hernandez DG, Hillege HL, Hirschhorn JN, Hofman A, Hubner N, Hwang SJ, Iorio A, Kähönen M, Kellis M, Kolcic I, Kooner IK, Kooner JS, Kors JA, Lakatta EG, Lage K, Launer LJ, Levy D, Lundby A, Macfarlane PW, May D, Meitinger T, Metspalu A, Nappo S, Naitza S, Neph S, Nord AS, Nutile T, Okin PM, Olsen JV, Oostra BA, Penninger JM, Pennacchio LA, Pers TH, Perz S, Peters A, Pinto YM, Pfeufer A, Pilia MG, Pramstaller PP, Prins BP, Raitakari OT, Raychaudhuri S, Rice KM, Rossin EJ, Rotter JI, Schafer S, Schlessinger D, Schmidt CO, Sehmi J, Silljé HH, Sinagra G, Sinner MF, Slowikowski K, Soliman EZ, Spector TD, Spiering W, Stamatoyannopoulos JA, Stolk RP, Strauch K, Tan ST, Tarasov KV, Trinh B, Uitterlinden AG, van den Boogaard M, van Duijn CM, van Gilst WH, Viikari JS, Visscher PM, Vitart V, Völker U, Waldenberger M, Weichenberger CX, Westra HJ, Wijmenga C, Wolffenbuttel BH, Yang J, Bezzina CR, Munroe PB, Snieder H, Wright AF, Rudan I, Boyer LA, Asselbergs FW, van Veldhuisen DJ, Stricker BH, Psaty BM, Ciullo M, Sanna S, Lehtimäki T, Wilson JF, Bandinelli S, Alonso A, Gasparini P, Jukema JW, Kääb S, Gudnason V, Felix SB, Heckbert SR, de Boer RA, Newton-Cheh C, Hicks AA, Chambers JC, Jamshidi Y, Visel A, Christoffels VM, Isaacs A, Samani NJ, de Bakker PI.

52 Genetic Loci Influencing Myocardial Mass.

J Am Coll Cardiol. 2016 Sep 27; 68 (13):1435-48. doi: 10.1016/j.jacc.2016.07.729.

Chapter 6

Claudia Tamar Silva, Herma van der Linde, Lies-Anne Severijnen, Jan A. Kors, Abbas Dehghan, Cornelia M. van Duijn, Aaron Isaacs, Rob Willemsen

Functional analyses of Arhgap24 in zebrafish, a gene previously associated with ECG variability.

Manuscript in preparation

CHAPTER 1: INTRODUCTION


ATTCGGTACCTCATAACGGCTAGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTT
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGGTACCTA

The heart is a muscular and a pacemaker organ that pumps blood through the blood vessels to provide the body with oxygen and nutrients, as well as the removal of metabolic wastes [1]. The human heart has four chambers: the upper right and left atria which drain blood through incoming cavae and pulmonary veins, respectively; and the lower right and left ventricles where blood is pumped through the pulmonary and aortic arteries, respectively. Pacemaker cells distributed along the sinoatrial node, the atrioventricular node and a conduction system that generates electrical impulses pulse determine the rhythm of contraction of the heart muscle [2].

Abnormalities of the heart rhythm or cardiac arrhythmias are characterized by conduction abnormalities that may lead to various conditions, including sudden cardiac death, atrial fibrillation, ventricular hypertrophy (LVH) among others. Sudden cardiac death is estimated to occur in 50 – 100 individuals per 100,000 per year in the United States (U.S.) and Europe [3]. Atrial fibrillation has become one of the most important public health problems and its prevalence is increasing due to our greater ability to treat chronic cardiac and non-cardiac diseases and aging of the populations [4]. These disorders impose high societal costs, both in terms of emotional well-being for patients and their relatives, and the financial burdens imposed on medical systems for patient care, medication, and surgery (such as pacemaker implantation).

In 1918, James B. Herrick advocated the use of the electrocardiogram (ECG) to diagnose myocardial infarction [5, 6]. Since then, the ECG has proven to be a key diagnostic tool for heart failure, arrhythmias, stress testing and cardiology consultation [7]. The ECG provides information on the depolarization and repolarization of myocardial tissue, reflecting electrical activity in the heart. Electrical activity abnormalities might indicate an evolving myocardial infarction, rhythm alterations, related pathology effects, cardiac exercise and rehabilitation, among other syndromes [8]. The overall rhythm of the heart and development of the heart muscle can be deducted from

the ECG (Figure 1) [9]. The most common parameters are the P wave, QRS interval, T wave and QT interval. The P wave reflects conduction of the cardiac impulse that is transmitted through the atria. The QRS complex amplitude is larger than the P wave and is produced by the ventricular contraction, after the ventricular myocardial cells depolarize [8, 10]. The T wave corresponds to the repolarization of the ventricle, while the QT interval depicts the time between the onset of ventricular depolarization and the end of ventricular repolarization, and the PR interval measures atrial and atrioventricular conduction from the sinoatrial node to the ventricular myocardium, primarily through the atrioventricular node [11]. The ECG can also be used to quantify LVH, a risk factor for cardiovascular morbidity and mortality [12]. More than 30 electrocardiographic indexes have been described for the diagnosis of LVH, the Sokolow-Lyon voltage index, the Cornell Voltage, Cornell product indexes, the Gubner index and the Romhilt-Estes score, with two different thresholds, are the indexes most commonly used [13]. Even though ECG has a low sensitivity for detection of LVH, the Sokolow-Lyon voltage together with the Cornell voltage duration product have been recommended as relevant parameters to access LVH according to the European Society of Hypertension. These guidelines are based on the LIFE study [14].

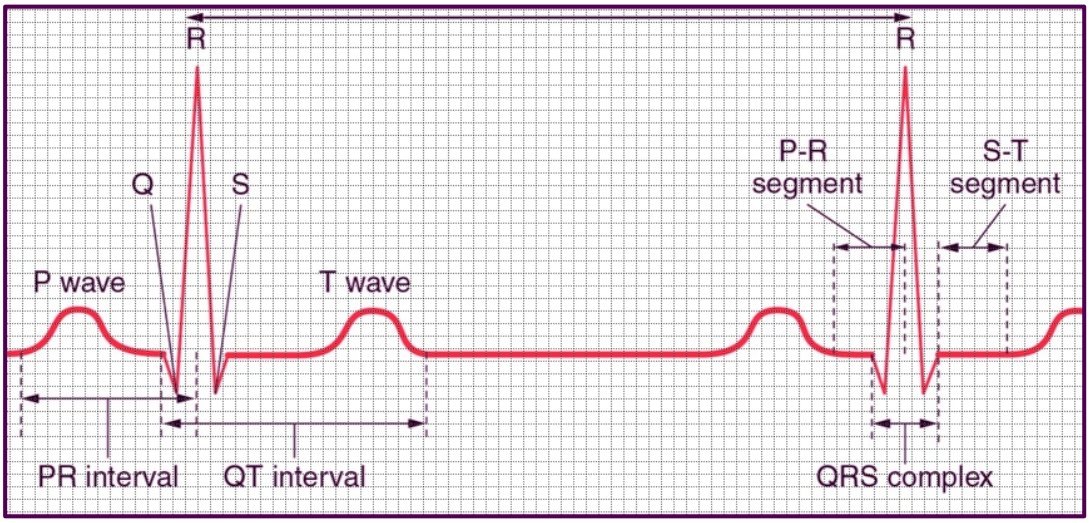


FIGURE 1. A typical electrocardiogram trace depicting various ECG intervals and waves observed in a normal ECG.

Heart rhythm disturbances Inheritance

It has been recognized for long, that genetics plays a key role in heart rhythm disorders, many of which have been linked to premature mortality [1]. Inherited heart rhythm disturbances, often also referred to as ion channelopathies, are a group of genetic conditions that can cause life-threatening arrhythmias. The most common are discussed in detail below: Brugada Syndrome, familial atrial fibrillation, catecholaminergic polymorphic ventricular tachycardia, long QT syndrome, progressive cardiac conduction defect and short QT Syndrome.

The Brugada syndrome was first described in 1992 and is characterized by an ST-segment elevation in the right precordial electrocardiogram that leads to a high incidence of sudden cardiac death in patients with structurally normal hearts. It affects 5 in 10,000 people worldwide [15] and is believed to cause up to 4-12% of cases of sudden cardiac death [16-19]. Clinically, there are three types of Brugada syndrome based on the electrophysiological classification: Type 1, characterized by a prominent ST-segment elevation ≥2 mm or 0.2 mV followed by a negative T-wave, with little or no isoelectric separation. Type 2 also has a high take-off ST-segment elevation, which gradually descends thereafter and is followed by a positive or biphasic T-wave that results in a saddle back configuration. Type 3 displays either a right precordial ST-segment elevation of <1 mm of

Table 1. Different genes involved in Brugada Syndrome (BrS)

BrS Subtype	OMIM	Nielsen	Fernández
BrS1	*SCN5A*	*SCN5A*	*SCN5A*
BrS2	*GPD1L*	*GPD1L*	*GPD1L*
BrS3	*CACNA1C*	*CACNA1C*	*CACNA1C*
BrS4	*CACNB2*	*CACNB2*	*CACNB2*
BrS5	*SCN1B*	*SCN1B*	*SCN1B*
BrS6	*KCNE3*	*KCNE3*	*KCNE3*
BrS7	*SCN3B*	*SCN3B*	*SCN3B*
BrS8	*HCN4*	*KCNH2*	*HCN4, KCNH2*
BrS9	*KCND3*	*KCNJ8*	*KCND3, KCNJ8*
BrS10	-	*CACNA2D1*	*CACNA2D1*
BrS11	-	*RANGRF*	*RANGRF*
BrS12	-	*KCNE5*	*KCNE5*
BrS13	-	*KCND3*	*KCND3*
BrS14	-	*HCN4*	*HCN4*
BrS15	-	*SLMAP*	*SLMAP*
BrS16	-	*TRPM4*	*TRPM4*
BrS17	-	*SCN2B*	*SCN2B*
-	-	-	*SCN10A*
-	-	-	*ABCC9*
-	-	-	*FGF12*
-	-	-	*HEY2*
-	-	-	*KCND2*
-	-	-	*PKP2*
-	-	-	*SEMA3A*

Three different classifications of BrS are shown in Table.1 OMIM considers 9 subtypes, Nielsen 17 subtypes and Fernandez described 24 genes involved in BrS without assigning specific subtypes to each gene. The genes shown in blue font

saddle back type, a coved type, or both [20].

Genetically, the OMIM database (http://www.omim.org/) reports nine types of Brugada syndrome (BrS1 – BrS9) based on mutations in nine different genes: *SCN5A*, *GPD1L*, *CACNA1C*, *CACNB2, SCN1B, KCNE3, SCN3B*, *HCN4* and *KCND3* (Table 1). Nielsen described 17 subtypes of Brugada syndrome (from BrS1 to BrS17) based on mutations in 17 genes. The first seven are identical to those reported in the OMIM database, but BrS8 and BrS9 in the Nielsen classification (*KCNH2* and *KCNJ8*) are different from those described in OMIM (*HCN4* and *KCND3*) (Table.1). In the Nielsen classification, mutations in *CACNA2D1, RANGRF, KCNE5*, *KCND3, SLMAP, TRPM4* and *SCN2B* characterize BrS10-BrS17 (Table 1) [21]. In 2017, seven additional genes were described by Fernández-Falgueras *et al* in *ABCC9, FGF12, HEY2, KCND2, PKP2, SCN10A* and *SEMA3A* (Table 1 and Figure 2) [21-23]. Even though these advances in dissecting the genetic causes of the Brugada syndrome, there is a large proportion of the patients with Brugada syndrome (60 – 70%) for whom the genetic variants responsible for this pathology remain to be discovered.

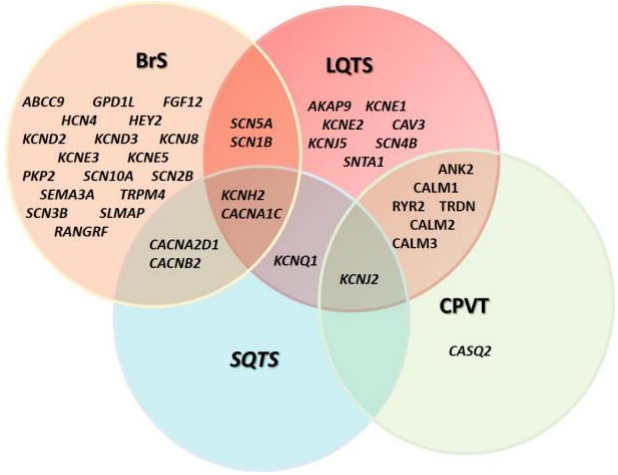


FIGURE 2. Genes related to Brugada syndrome, long QT syndrome, short QT syndrome and Catecholaminergic polymorphic ventricular tachycardia (adapted from Fernández-Falgueras *et al*) [23]. Among genes related to Brugada syndrome, long QT syndrome, short QT syndrome and Catecholaminergic polymorphic ventricular tachycardia, there are: ATP binding cassette subfamily C member 9 (ABCC9), glycerol-3-phosphate dehydrogenase 1 like (GPD1L), fibroblast growth factor 12 (FGF12), hyperpolarization activated cyclic nucleotide gated potassium channel 4 (HCN4), hes related family bHLH transcription factor with YRPW motif 2 (HEY2), potassium voltage-gated

channel subfamily D member 2 (KCND2), potassium voltage-gated channel subfamily D member 3 (KCND3), potassium voltage-gated channel subfamily J member 8 (KCNJ8), potassium voltage-gated channel subfamily E regulatory subunit 3 (KCNE3), potassium voltage-gated channel subfamily E regulatory subunit 5 (KCNE5), plakophilin 2 (PKP2), sodium voltage-gated channel alpha subunit 10 (SCN10A), sodium voltage-gated channel beta subunit 2 (SCN2B), semaphorin 3A (SEMA3A), transient receptor potential cation channel subfamily M member 4 (TRPM4), sodium voltage-gated channel beta subunit 3 (SCN3B), sarcolemma associated protein (SLMAP), RAN guanine nucleotide release factor (RANGRF), sodium voltage-gated channel alpha subunit 5 (SCN5A), sodium voltage-gated channel beta subunit 1(SCN1B), A-kinase anchoring protein 9 (AKAP9), potassium voltage-gated channel subfamily E regulatory subunit 1 (KCNE1), potassium voltage-gated channel subfamily E regulatory subunit 2 (KCNE2), caveolin 3 (CAV3), potassium voltage-gated channel subfamily J member 5 (KCNJ5), sodium voltage-gated channel beta subunit 4 (SCN4B), syntrophin alpha 1 (SNTA1), potassium voltage-gated channel subfamily H member 2 (KCNH2), calcium voltage-gated channel subunit alpha1 C (CACNA1C), ankyrin 2 (ANK2), calmodulin 1 (CALM1), ryanodine receptor 2 (RYR2), triadin (TRDN), calmodulin 2 (CALM2), calmodulin 3 (CALM3), calcium voltage-gated channel auxiliary subunit alpha2delta 1 (CACNA2D1), calcium voltage-gated channel auxiliary subunit beta 2 (CACNB2), potassium voltage-gated channel subfamily Q member 1 (KCNQ1), potassium voltage-gated channel subfamily J member 2 (KCNJ2), calsequestrin 2 (CASQ2).

Catecholaminergic polymorphic ventricular tachycardia are inherited cardiac channelopathies with an estimated prevalence of 1 in 10.000. Catecholaminergic polymorphic ventricular tachycardia (CPVT1 to CPVT5) are associated with mutations in respectively the ryanodine receptor 2 (*RYR2*), cancer susceptibility 2 (non-protein coding) (*CASC2*), trans-2,3-enoyl-CoA reductase like (*TECRL*), calmodulin 1 (*CALM1*) and triadin (*TRDN*) genes. Additionally, mutations in potassium voltage-gated channel subfamily J member 2 (*KCNJ2*), have been identified in patients with a CPVT-like phenotype [24, 25]. Two additional genes possibly involved in catecholaminergic polymorphic ventricular tachycardia are Ankyrin 2 (*ANK2*) and calmodulin 3 (*CALM3*) (Figure 2) [23]. Catecholaminergic polymorphic ventricular tachycardia is characterized by potentially life-threatening polymorphic ventricular tachycardias during exercise or emotional stress. There result in light-headedness, dizziness, syncope, and sudden death, in individuals without structural cardiac abnormalities [25, 26].

Pathologically, catecholaminergic polymorphic ventricular tachycardia are characterized by a dysregulation of intracellular calcium handling, and the subjacent molecular

mechanism includes dysfunction of the sarcoplasmic reticulum during exercise due to release of catecholamines related to intracellular calcium dysregulation. Calcium uptake is stimulated via beta-adrenergic input into the sarcoplasmic reticulum by increasing permeability to calcium in ryanodine receptor 2 (RYR2), a calcium channel. Sarcoplasmic reticulum calcium release, results in catecholamines and myocyte calcium loading, consequently increasing heart rate and the susceptibility to trigger ventricular tachycardia [25].

Long QT syndrome is a congenital disease with an estimated prevalence in 2009 of 1/2000 [27]. According to the portal for rare disease and orphan drugs (Orphanet), the prevalence in 2016 was about to 1/2500 (orphanet). This syndrome is characterized by prolongation of the QT interval, syncopal attacks due to ventricular arrhythmias, and an elevated risk of sudden cardiac death [28]. Syncope during exercise or high emotional states are usually the first symptoms. Strikingly, 50% of patients have the first cardiac event by the age of 12 years [29, 30]. Long QT syndrome is divided according to the underlying genetic substrate in long QT syndrome type1 to Long QT syndrome type15 [30, 31]. Diagnosis of Long QT syndrome according to the Schwartz score is based on: suggestive findings such as 1) prolongation of the corrected QT (QTc) interval bigger than 450ms (male) and 470ms (women) in the absence of specific conditions known to lengthen the interval, 2) 4-min recovery QT after exercise test ≥ 480ms, 3) torsades points, 4) lower heart rate for age and/or T-wave alterations on the ECG, and 5) a clinical history of syncope and/or congenital deafness [32, 33]. Nowadays, molecular genetic testing of one or more of the 15 genes known to be associated with long QT syndrome, confirms clinical assessment. Genes leading to Long QT syndrome type 1 to Long QT syndrome type15 are potassium voltage-gated channel subfamily Q member 1 (*KCNQ1*)*, KCNH2,* sodium voltage-gated channel alpha subunit 5 (*SCN5A*)*,* ankyrin-B (*ANKB*)*,* potassium voltage-gated channel subfamily E regulatory subunit 1 (*KCNE1*)*,* potassium voltage-gated channel subfamily E regulatory subunit 2 (*KCNE2*)*,* potassium voltage-

gated channel subfamily J member 2 (*KCNJ2*), calcium voltage-gated channel subunit alpha1 C (*CACNA1C*), caveolin 3 (*CAV3*), sodium voltage-gated channel beta subunit 4 (*SCN4B*), A-kinase anchoring protein 9 (*AKAP9*), potassium voltage-gated channel subfamily J member 5 (*KCNJ5*), syntrophin alpha 1 (*SNTA1*), calmodulin 1 (*CALM1*), and calmodulin 2 (*CALM2*) respectively [34]. Additional mutations have been described in other genes including sodium voltage-gated channel beta subunit 1 (*SCN1B*), *RYR2, TRDN* and *CALM3* (Figure 2) [23].

Long QT syndrome shows an autosomal dominant inheritance. This implies that individuals diagnosed with long QT syndrome usually have an affected parent, and that the risk of a child with long QT syndrome is 50%. However, a small proportion of the cases have *de novo* mutations. The mutational spectrum includes all type of mutations (missense, frameshift, nonsense, splice sites, deletions, and insertions), which are analyzed by different techniques like new generation sequencing, SNaPshot, whole exome sequencing and multiplex ligation-dependent probe amplification [35, 36]. More than 75% of the mutations are found in *KCNQ1, KCNH2* and *SCN5A* [37, 38] and the remaining genes represent only 5%. Approximately 20% of patients with long QT syndrome lack any of the known mutations [38]. This unknown mutations, could be uncovered through whole genome sequencing, looking for rare variants in unknown genes or regulatory regions.

Incomplete penetrance and variable expressivity have been described, conferring different risks in related individuals [39]. Recently, genetic factors have been described to be involved in disease modulation and clinical severity. Those factors are recognized as genetic modifier. The first variant described as genetic modifier influencing LQT was a single nucleotide polymorphism (SNP) in *KCNH2*-K897T, which modulates the clinical expression of a primary mutation for LQT2 in the same gene [34, 40]. There are variants in at least 18 genes involved in the pathophysiology of Long QT syndrome, three of those genes are genes with large effect on the phenotype (*KCNQ1, KCNH2*,

and *SCN5A*), and 15 with minor influence [41]. Further, SNPs that modulate Long QT syndrome phenotype have been described including polymorphisms in nitric oxide synthase 1 adaptor protein encoded by *NOS1AP*, this SNP in combination with *KCNQ1* (A341V) modulates occurrence of symptoms, with clinical severity and QT interval [34, 42].

Drugs may cause a prolonged QT interval as well, leading to some drug being taken of the market [43, 44]. These include drugs related to QT prolongation such as antiarrhythmic drugs (flecainide and amiodarone among others), and non-cardiac drugs as antidepressants like citalopram and antibiotics as erythromycin and fluoroquinolones [32]. Drug susceptibility can also be related to genetic variability. *NOS1AP*, is one of the strongest genes revealed by genome wide association studies (GWAs) related to QT interval and has a pharmacodynamic effect. NOS1P regulates the enzyme neuronal nitric oxide synthase (*nNOS*) and *nNOS* is a regulator of calcium levels [32]. Another gene known to influence pharmacodynamic susceptibility is *KCNH2*. Mutations in *KCNH2* are responsible for the congenital long QT syndrome type 2 and mutations in this gene have been *described in people* with prolonged QT interval induced by drugs [32]. Variation in pharmacokinetics response is due to genetics factors. Some polymorphisms in genes related to metabolism, absorption, distribution and drug elimination are responsible for these differences. Among these genes of the cytochrome P450 (CYP) like system such as cytochrome P450 family 2 subfamily B member 6 (*CYP2B6*), cytochrome P450 family 2 subfamily C member 9 (*CYP2C9*), cytochrome P450, family 19 (*CYP19*), cytochrome P450 family 2 subfamily D member 6 (C*YP2D6*), cytochrome P450 family 3 subfamily A member 4 (*CYP3A4*), which encode for proteins involved in drug metabolism in the liver have been described related to drug-induced QT interval prolongation [32].

Short QT Syndrome. Until 2014, approximately 100 short QT syndrome patients were reported in the literature [62]. Short QT Syndrome is a rare disease, with debated diagnostic

criteria and a cutoff value is not fully established [63]. It has an estimated prevalence that is lower than 1 in 10.000 [23], and is defined by: a QTc interval ≤340 ms or a QTc interval between 341 ms and 360 ms and additionally, one or more of the following factors: family history of short QT syndrome, family history of unexplained cardiac arrest at 40 years of age or younger, history of cardiac arrest or syncope, or the presence of a disease-causing disease mutation. Mazzanti et al proposed that those with a Short QT syndrome interval ≤360 ms should be classified as suspected patients [62].

Several groups have reported a relationship to sudden cardiac death morbidity. In 1993, Algra et al showed a 2-fold higher risk of arrhythmias and sudden cardiac death in people with short QT interval. In 2000 Gussak et al showed the relationship among short QT syndrome and sudden cardiac death [61, 64, 65]. Guzzak et al described two cases with Short QT syndrome and spontaneous atrial fibrillation, but it was not until 2003 when a new autosomal dominant sort QT syndrome was reported [66] based on seven patients with short QT interval and syncope, palpitations and sudden cardiac death [67]. The age of onset ranged between infancy and old age, and 25% to 33% are presented with cardiac arrest, and 15% of cases are presented with syncope. Other minor events at the clinical expression described involve palpitations and/or dizziness.

So far, six types of short QT syndrome have been described according to the underlying genes. *KCNH2, KCNQ1, KCNJ2, CACNA1C, CACNB2, CACNA2D1* are related to Short QT syndrome type 1 to Short QT syndrome type 6, respectively (Figure 2). These mutations lead to loss of normal rectification of the electrical current at plateau voltages, and consequently an increase of the rapid activating current potassium channel (IKr). Since ventricular action potentials are directly related to the duration of the QT interval, an action potential shortening produced by a shortening of the refractory period creates an increased ventricular and atrial susceptibility to premature stimulation [67]. *KCNQ1* mutations have been studied in detail and these studies demonstrated

that V141M abolishes pacemaker activity of the sinoatrial node and shortens the action potential duration of human ventricular myocytes [67-69]. Collectively, these changes cause a significant increase in the inwardly rectifying potassium current [70].

Figure 2 shows that in addition to overlap in genes involved in various disorders, mutations in genes encoding for potassium channels or their subunits (Table 1) are the predominant gene family involved in these syndromes. They are the largest group of ion channels in the human heart, and consequently these channels contribute to distinct phases of action potential, and consequently with cardiomyocyte repolarization [71]. Mutations in genes encoding these proteins are related to Brugada syndrome, atrial fibrillation, long and short QT syndrome as described in Table 1 [71].

Gene Name	**Symbol**	**Phenotype**
POTASSIUM CHANNEL, VOLTAGE-GATED, KQT-LIKE SUBFAMILY, MEMBER 1	*KCNQ1*	Atrial fibrillation, long QT syndrome 1, short QT syndrome 2
POTASSIUM CHANNEL, VOLTAGE-GATED, SUBFAMILY H, MEMBER 2	*KCNH2*	Long QT syndrome 2, short QT syndrome 1, Brugada syndrome 8
POTASSIUM CHANNEL, VOLTAGE-GATED, SHAKER-RELATED SUBFAMILY, MEMBER 5	*KCNA5*	Atrial fibrillation
POTASSIUM CHANNEL, INWARDLY RECTIFYING, SUBFAMILY J, MEMBER 5	*KCNJ5*	Long QT syndrome 13
POTASSIUM CHANNEL, VOLTAGE-GATED, ISK-RELATED SUBFAMILY, MEMBER 1	*KCNE1*	Long QT syndrome 5
POTASSIUM CHANNEL, INWARDLY RECTIFYING, SUBFAMILY J, MEMBER 8	*KCNJ8*	Brugada syndrome 9
POTASSIUM CHANNEL, INWARDLY RECTIFYING, SUBFAMILY J, MEMBER 2	*KCNJ2*	Atrial fibrillation, short QT syndrome 3, long QT syndrome 7
POTASSIUM CHANNEL, VOLTAGE-GATED, ISK-RELATED SUBFAMILY, MEMBER 3	*KCNE3*	Brugada syndrome 6

Gene Name	Symbol	Phenotype
POTASSIUM CHANNEL, VOLTAGE-GATED, ISK-RELATED SUBFAMILY, MEMBER 2	*KCNE2*	Atrial fibrillation, long QT syndrome 6
POTASSIUM VOLTAGE-GATED CHANNEL, SHAL-RELATED SUBFAMILY, MEMBER 3	*KCND3*	Brugada syndrome 13
Potassium Voltage-Gated Channel Subfamily E Regulatory Subunit 5	*KCNE5*	Brugada syndrome 5

TABLE 1. Potassium channels related to channelopaties

Familial atrial fibrillation**.** Atrial fibrillation is characterized by a fast an irregular heartbeat due to an uncoordinated electrical activity in the heart's atria. It is the most prevalent supraventricular sustained arrhythmia affecting nearly 33.5 million people worldwide and the number of affected individuals by this pathological condition is increasing over time and has doubled since 2010. Atrial fibrillation is associated with an increased risk of stroke, sudden death, heart failure, dementia, and mortality. One of the largest population-based cohort from the UK Clinical Practice Research Datalink showed that the incidence of atrial fibrillation has increased from 5.9/1000 person-year in 2001 to 6.9/1000 person-year in 2013 [45]. There are several risk factors for atrial fibrillation such as the use of Ivabradine for treatment of heart failure, diastolic dysfunction, and hemodialysis, among others [45]. Hemodialysis itself, in patients with an implanted pacemaker or defibrillator may trigger atrial fibrillation, with a prevalence ranging among 13%-23% [45-47]. This could be explained by two different pathways, the first one is related to intravascular volume reduction, causing liberation of catecholamine and sympathetic activation. The second one is related to transmembranous fluxes of electrolytes, especially potassium which is produced during hemodialysis, suggesting an association with the concentration of potassium [45, 46].

For long it has been recognized that there is a strong genetic component determining the risk of atrial fibrillation [48]. Various studies showed that family members have an increased

relative risk of atrial fibrillation compared to the general population [49-51]. Genetic variants have been associated with atrial fibrillation and involve genes encoding signaling molecules, potassium channel, proteins involved in cardiac polarization and repolarization, cardiac gap junctions proteins, transcription factors, and sodium channels: paired like homeodomain 2 (*PITX2*), zinc finger homeobox 3 (*ZFHX3*), potassium calcium-activated channel subfamily N member 3 (*KCNN3*), caveolin 1/caveolin2 (*CAV1/CAV2*), paired related homeobox 1 (*PRRX1*), spectrin repeat containing nuclear envelope protein 2 (*SYNE2*), chromosome 9 open reading frame 3 (*C9orf3*), *HCN4,* synaptopodin 2 like (*SYNPO2L*), *KCNQ1,* potassium voltage-gated channel subfamily E regulatory subunit 1 to 5(*KCNE1-5*), *SCN5A,* T-box 5 (*TBX5*), sodium voltage-gated cannel geta subunit 1 to 4 (*SCN1B-4*), nucleoporin 155 (*NUP155*), natriuretic peptide A (*NPPA*), GATA binding protein 4 (*GATA4*), GATA binding protein 6 (*GATA6*), lamin A/C (*LMNA*), gremlin 2, DAN family BMP antagonist (*GREM2*), gap junction protein alpha 1 (*GJA1*), gap junction protein alpha 5 (*GJA5*), *KCNA5, KCNJ2, ABCC9, PRRX1, KCND3, KCNH2, KCNJ8* and NK2 homeobox 5 (*NKX2-5*) [52, 53]. Atrial fibrillation diagnostic testing may include, ECG, echocardiogram and a chest X-ray. Despite its clinical relevance, treatments have low efficacy, due to poor understanding of atrial fibrillation pathophysiology, which makes clinical control more difficult to reach. Inter individual variability and complex genetic inheritance are part of the heterogeneous nature of atrial fibrillation.

Progressive cardiac conduction defect. It is a common genetic disease that occurs in adults, and appears typically in the fifth decade of life. More than 50 families presenting this pathological condition have been described in the literature. This disease affects the His-Purkinje system and is characterized by a progressive slowing of cardiac conduction and prolongation of QRS complex, leading to the atrioventricular block. The disease, also called as Lenègre or Lev disease, is either asymptomatic or manifests as dyspnea, dizziness, syncope, abdominal pain, heart failure or sudden death [54]. Currently, therapeutic strategies for progressive cardiac conduction

defects involve the implantation of a permanent pacemaker. In patients who receive a pacemaker implantation, the prognosis is excellent and their life expectancy is very close to that of the general population, except in those with *LMNA* mutations that can lead to ventricular tachycardia and sudden cardiac death. In this population, cardioverter defibrillator implantation is recommended in case of severe cardiac conduction defect. Progressive cardiac conduction defect is an autosomal dominant inherited disease, mutations in *SCN5A, SCN1B, TRPM4, NKX2.5, TBX5* and recently potassium two pore domain channel subfamily K member 17 (*KCNK17*) have been described [55, 56]. *SCN5A* mutations are related to several cardiac diseases, including lethal arrhythmias, long QT syndrome type 3, early-onset lone atrial fibrillation, dilated cardiomyopathy, Brugada syndrome, and channelopathies. Phenotypic variability of *SCN5A* mutation carriers is called overlap syndrome. Patients suffering this pathological condition display overlapping clinical manifestations of the different *SCN5A*-related syndromes [54, 57]. The phenotypic difference is an unclear phenomenon, and could be related to either a gain or loss of function of the channel. Long QT syndrome type 3 is related to gain of function of *SCN5A,* whereas Brugada syndrome is caused by loss of function [58]. Mutations in *SCN1B* have been described in families with alteration of the conduction system. *SCN1B*, encodes beta1 subunit of the voltage-gated sodium channel, this beta-subunit interact with the cardiac sodium channel protein Nav1.5 [54]. *TRPM4* is involved in the pathogenesis of conduction disorders through gain-of-function mutations. Mutant TRPM4 channels produce a higher voltage than their wild-type counterparts, leading to a cell membrane depolarization [54]. Kruse M *et al* proposed that this is related to deSUMOylation intensity, which may impair endocytosis and stabilize the mutant channels at the level of the cell surface [59]. The gain of function has been related to altered deSUMOylation, which leads to a depolarization of the membrane due to the mutant channels [59]. Gain of function mutation in the gene encoding potassium channel TASK-4 results in an increase in voltage amplitude, membrane

hyperpolarization and slow conductivity [55]. The *KCNK17* missense mutation (G88R) implicated in progressive cardiac conduction defects was described in a single patient. Additionally, other genes related to progressive cardiac conduction defects are associated with congenital heart disease, including transcription factors related to endocardial cushion remodeling, conduction system development, and cardiac chamber formation like *NKX2.5* and *TBX5* [60, 61]. *NKX2.5* encodes a cardiac-specific homeobox transcription factor, which could harbor a large number of mutations related to different congenital heart phenotypes.

From Mendelian genetics to complex genetics

There are ongoing efforts to screen for mutations in high-risk families to prevent sudden cardiac death and atrial fibrillation at an early stage to ensure therapeutical interventions to prevent morbidity and mortality. Like familial forms of dyslipidemia, screening programs approach relatives of patients systematically and invite them to participate for clinical and genetic evaluations. This type of cascade screening is controversial as such screening programs may undermine the autonomy of relatives, who may feel obliged to participate [62]. However, in families with known mutations, cascade screening may be extremely effective and successful in preventing morbidity and mortality. Nevertheless, for many families the genetic cause of disease is still not understood and there is an urgent need to search for rare variants explaining the disease in these families. The classical approach to find rare variants with large effects is to conduct genome wide studies covering the full genome in families and analyze the data statistically using linkage analysis. In linkage analyses, co-segregation of DNA markers with the disease is assessed (Figure 3; left side). The rationale of these analyses is that co-segregation occurs not only for the disease mutation, but also for genetic variants in linkage disequilibrium with the mutation underpinning the disease. When two variants are located close together in a chromosome, it is

unlikely that a new mutation occurs and therefore these variants are likely passed on jointly from one generation to the next in a family.

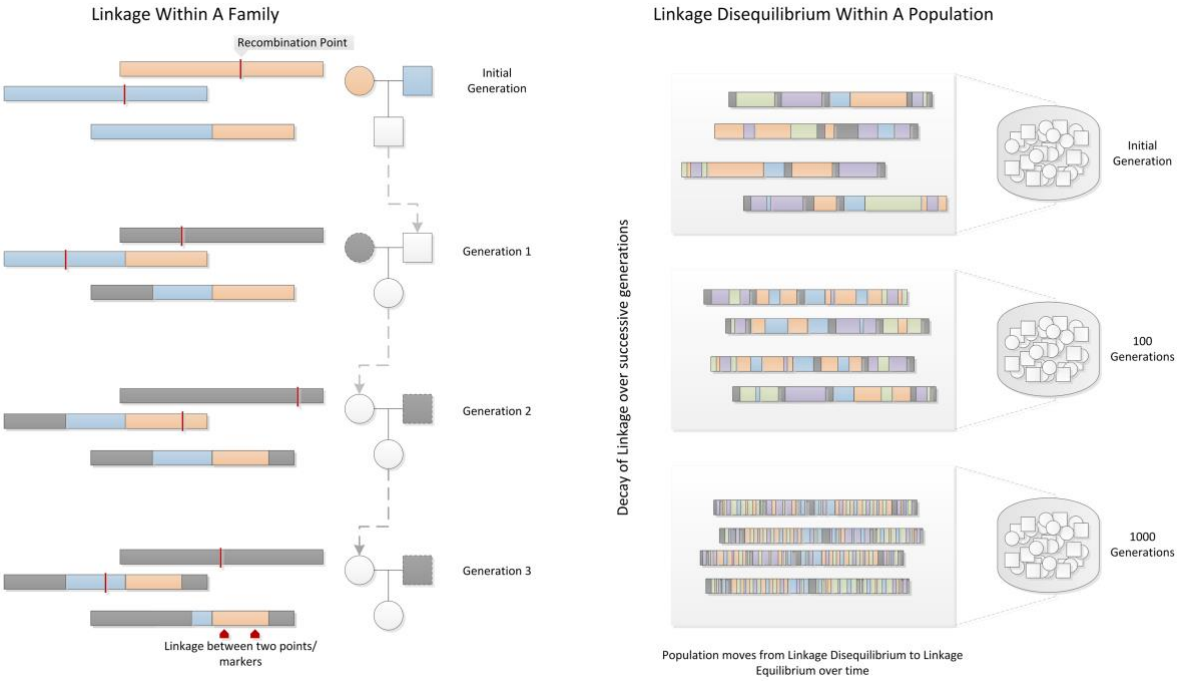


FIGURE 3. Linkage and linkage Disequilibrium

Within a family, linkage occurs when two genetic markers (points on a chromosome) remain linked on a chromosome rather than being broken apart by recombination events during meiosis, shown as red lines. In a population, contiguous stretches of founder chromosomes from the initial generation are sequentially reduced in size by recombination events. Over time, a pair of markers or points on a chromosome in the population move from linkage disequilibrium to linkage equilibrium, as recombination events eventually occur between every possible point on the chromosome. Source: Bush and Moore [63].

Without any doubt, within the general population, rare variants that convey an elevated risk of disease occur and may explain part of the disease [12, 64, 65]. However, in a substantial number of patients, the genetic architecture of conduction disorders appears to be more complex, involving the interplay of multiple genes and non-genetic risk factors. The effect of a single common variant on disease risk for an individual may be small. However, the additive effect of many of those common low risk variants may be substantial, depending on the combination of the genetic risk factors a person carries and their effects on the disease. Figure 4 shows that life time

risks for those carrying over 25 risk variants may increase up to 60%. Common variants implicated in a disease can be discovered by the same principle as linkage analyses, i.e., the assumption that only loci close to the disease locus are segregating together in the population. However, if we are dealing with very distantly related or even unrelated subjects, linkage analyses fail. Association analyses has proven to be a powerful approach to discover these genes of minor effect in unrelated persons. In the past decade, many of such genes have been identified by association analyses [66]. The basic rationale of association is that genes causally related to a disease should be found more often in cases than in controls. However, since recombination between two genes that are close together on a chromosome is unlikely also in unrelated subjects from the same population (linkage disequilibrium), genetic variants in the nearby of the causal variant will also be found more often in affected persons than in unaffected subjects. This phenomenon will result in association of the disease to genetic variants near the causal variant [63]. The effect of a single variant is small but as each of us may carry a substantial number of low risk variants for disorders, the impact of the genes may be substantial (see figure 4) and may increase the risk of disease 6 fold (up to 60%) depending on the effect of the variants carried by a person.

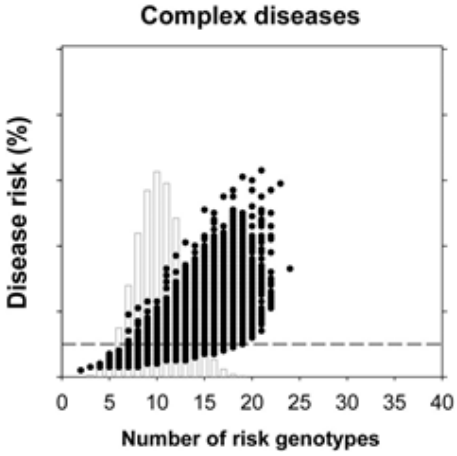


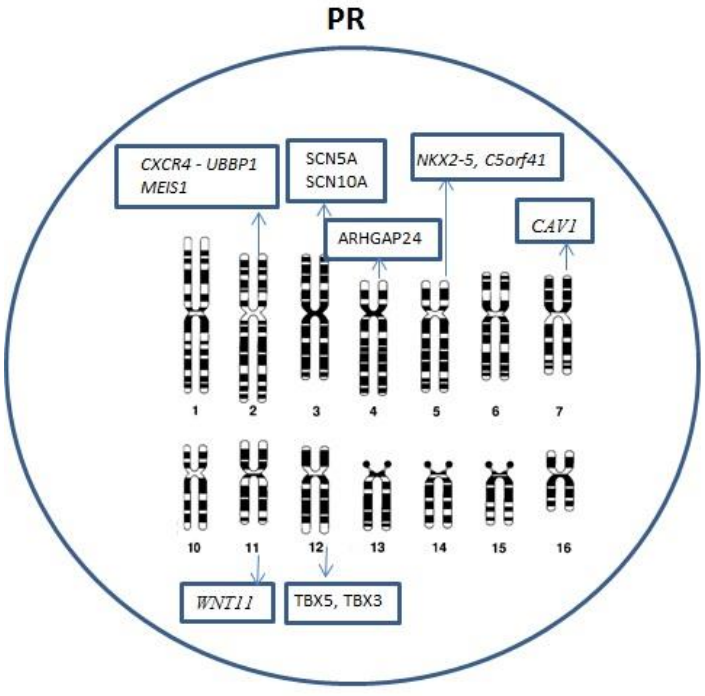
FIGURE 4. Disease risks when adding up the effects of multiple low risk variants

Disease risks for the complex diseases example were based on simulated data assuming a population risk of disease of 10% (dashed line), frequencies of the risk genotypes varying between 1 and 60% and odds ratios varying from 1.05 to 2.0. The bars in the scatterplot represent the

frequency distribution of the number of risk genotypes. The example and the simulation strategy have been described previously [67].

GWAs has proven to be a powerful approach to discover common but of small effect risk variants. Large scale studies of the various ECG parameters have brought to surface a large number of genetic risk variants. ECG parameters in the general population show non-mendelian inheritance patterns and are most likely explained by the additive effect of common variants [3].

By GWAS, more than 120 loci involved in ECG variability have been uncovered (Figure 5). These loci are related with the PR, QRS and QT intervals (Figure 5) [67-78]. Also, for common variants, there are several genes with different phenotypic effects over different hereditary diseases, e.g. *KCN5A and KNNJ2*, among others. Indeed, in addition to these highly penetrant, rare mutations, recent evidence suggests that combinations of common variants can also lead to conditions that emulate rare Mendelian disorders, such as Brugada syndrome [67, 68].



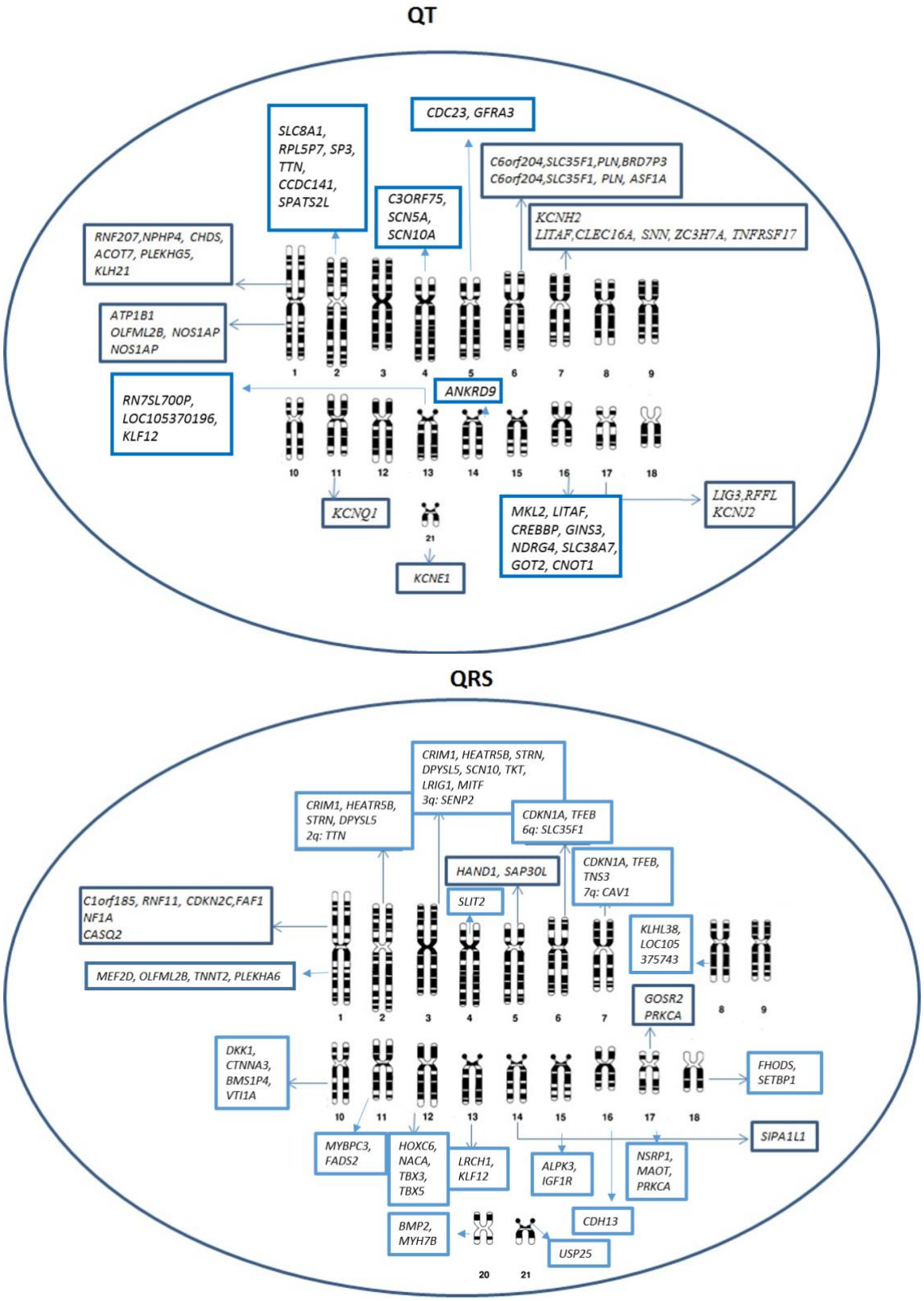


FIGURE 5. Loci associated by GWAs with each ECG interval: PR, QRS and QTCandidate genes identified by GWAs for PR, QT and QRS interval and their chromosomal location

Scope of this thesis

The scope of this thesis is to understand the heritability of heart rhythm and conduction disorders. Heritability is the portion of the phenotypic variability explained by genetic components. Several studies estimated a high heritability for RR interval (40% - 98%) and moderate heritability's for QT/QTc (25% - 67%), PR (34% - 46%), and QRS (33% - 43%) [69-73, 79-83]. To date, no studies have directly estimated the extent to which the GWAs loci explain the heritability. In the chapter 2 of this thesis, I present a heritability study of the various ECG parameters in the Erasmus Rucphen Family study (ERF). ERF is a family based study, a cohort derived from a region in the Southwest of the Netherlands. In the ERF study, we addressed the following question: what is the extent of heritability that can be explained by GWAs findings up to date?

In this thesis, I also aimed to discover new loci that may explain the heritability of heart rhythm and conduction disorders. My first aim was to discover new rare variants with large effects. To this end, we conducted linkage analyses of several ECG parameters including classical parameters QT, QRS and PR for sudden cardiac death (chapter 3) and LVH (chapter 4) in the ERF study. We combined the linkage analyses with association studies of the region. Association is not only powerful to detect common variants with small effects but can also be used for detecting rare variants with modest effects under a linkage peak [84]. Chapter 5 and 6 present 2 GWA studies. Chapter 5 involves the findings of a meta-analysis of GWAS. We identified 52 genomic loci, associated to 4 QRS traits providing new knowledge into genes and pathways related to myocardial mass. Chapter 6, involves a study using the genome of the Netherlands as a basis to identify new genes involved in conduct disorders. In chapter 7, a functional study of the *ARHGAP24* gene is presented including a search for rare variants in this gene associated to ECG parameters. I explored the cellular function of *ARHGAP24* in heart development using a

knockdown strategy with morpholino antisense oligonucleotides in zebrafish. Finally, the findings of the thesis are discussed in chapter 8.

References

1. Zelis, R., et al., *Cardiocirculatory dynamics in the normal and failing heart.* Annu Rev Physiol, 1981. **43**: p. 455-76.
2. R.C, R.J.S.a.R.R., *The normal heart.* American Heart Journal, 1942. **23**(4): p. 455-467.
3. Fishman, G.I., et al., *Sudden cardiac death prediction and prevention: report from a National Heart, Lung, and Blood Institute and Heart Rhythm Society Workshop.* Circulation, 2010. **122**(22): p. 2335-48.
4. Zoni-Berisso, M., et al., *Epidemiology of atrial fibrillation: European perspective.* Clin Epidemiol, 2014. **6**: p. 213-20.
5. Go, A.S., et al., *Heart disease and stroke statistics--2014 update: a report from the American Heart Association.* Circulation, 2014. **129**(3): p. e28-e292.
6. Dalen, J.E., et al., *The Epidemic of the 20 Century: Coronary Heart Disease.* Am J Med, 2014.
7. King, M., J. Kingery, and B. Casey, *Diagnosis and evaluation of heart failure.* Am Fam Physician, 2012. **85**(12): p. 1161-8.
8. Reilly, R.B. and T.C. Lee, *Electrograms (ECG, EEG, EMG, EOG).* Technol Health Care, 2010. **18**(6): p. 443-58.
9. Geselowitz, D.B., P.H. Langner, Jr., and F.T. Mansure, *Further studies on the first derivative of the electrocardiogram, including instruments available for clinical use.* Am Heart J, 1962. **64**: p. 805-14.
10. Greene, K.R., *The ECG waveform.* Baillieres Clin Obstet Gynaecol, 1987. **1**(1): p. 131-55.
11. Davey, P., *A new physiological method for heart rate correction of the QT interval.* Heart, 1999. **82**(2): p. 183-6.
12. Schillaci, G., F. Battista, and G. Pucci, *A review of the role of electrocardiography in the diagnosis of left ventricular hypertrophy in hypertension.* J Electrocardiol, 2012. **45**(6): p. 617-23.
13. Pewsner, D., et al., *Accuracy of electrocardiography in diagnosis of left ventricular hypertrophy in arterial hypertension: systematic review.* BMJ, 2007. **335**(7622): p. 711.
14. European Society of Hypertension-European Society of Cardiology Guidelines, C., *2003 European Society of Hypertension-European Society of Cardiology guidelines for the management of arterial hypertension.* J Hypertens, 2003. **21**(6): p. 1011-53.
15. Reference, G.H., *Brugda Syndrome.* 2015.
16. Gourraud, J.B., et al., *The Brugada Syndrome: A Rare Arrhythmia Disorder with Complex Inheritance.* Front Cardiovasc Med, 2016. **3**: p. 9.
17. Brugada, P. and J. Brugada, *Right bundle branch block, persistent ST segment elevation and sudden cardiac death: a distinct clinical and electrocardiographic syndrome. A multicenter report.* J Am Coll Cardiol, 1992. **20**(6): p. 1391-6.
18. Juang, J.M. and S.K. Huang, *Brugada syndrome--an under-recognized electrical disease in patients with sudden cardiac death.* Cardiology, 2004. **101**(4): p. 157-69.
19. Papadakis, M., et al., *Sudden cardiac death with autopsy findings of uncertain significance: potential for erroneous interpretation.* Circ Arrhythm Electrophysiol, 2013. **6**(3): p. 588-96.
20. Wilde, A.A., et al., *Proposed diagnostic criteria for the Brugada syndrome: consensus report.* Circulation, 2002. **106**(19): p. 2514-9.
21. Nielsen, M.W., et al., *The genetic component of Brugada syndrome.* Front Physiol, 2013. **4**: p. 179.
22. Brugada, R., et al., *Brugada syndrome.* Methodist Debakey Cardiovasc J, 2014. **10**(1): p. 25-8.

23. Fernandez-Falgueras, A., et al., *Cardiac Channelopathies and Sudden Death: Recent Clinical and Genetic Advances.* Biology (Basel), 2017. **6**(1).
24. Rios, E., et al., *The couplonopathies: A comparative approach to a class of diseases of skeletal and cardiac muscle.* J Gen Physiol, 2015. **145**(6): p. 459-74.
25. Faggioni, M., C. van der Werf, and B.C. Knollmann, *Sinus node dysfunction in catecholaminergic polymorphic ventricular tachycardia: risk factor and potential therapeutic target?* Trends Cardiovasc Med, 2014. **24**(7): p. 273-8.
26. Napolitano, C. and S.G. Priori, *Diagnosis and treatment of catecholaminergic polymorphic ventricular tachycardia.* Heart Rhythm, 2007. **4**(5): p. 675-8.
27. Schwartz, P.J., et al., *Prevalence of the congenital long-QT syndrome.* Circulation, 2009. **120**(18): p. 1761-7.
28. Jongbloed, R.J., et al., *Novel KCNQ1 and HERG missense mutations in Dutch long-QT families.* Hum Mutat, 1999. **13**(4): p. 301-10.
29. Moss, A.J., et al., *The long QT syndrome. Prospective longitudinal study of 328 families.* Circulation, 1991. **84**(3): p. 1136-44.
30. Mizusawa, Y., M. Horie, and A.A. Wilde, *Genetic and clinical advances in congenital long QT syndrome.* Circ J, 2014. **78**(12): p. 2827-33.
31. Goldenberg, I., W. Zareba, and A.J. Moss, *Long QT Syndrome.* Curr Probl Cardiol, 2008. **33**(11): p. 629-94.
32. Niemeijer, M.N., et al., *Pharmacogenetics of Drug-Induced QT Interval Prolongation: An Update.* Drug Saf, 2015. **38**(10): p. 855-67.
33. Schwartz, P.J. and L. Crotti, *QTc behavior during exercise and genetic testing for the long-QT syndrome.* Circulation, 2011. **124**(20): p. 2181-4.
34. Schwartz, P.J., L. Crotti, and R. Insolia, *Long-QT syndrome: from genetics to management.* Circ Arrhythm Electrophysiol, 2012. **5**(4): p. 868-77.
35. J, E., et al., *Mutation analysis for the detection of long QT-syndrome (LQTS) associated SNPs.* Int J Legal Med, 2016.
36. Williams, V.S., et al., *Multiplex ligation-dependent probe amplification copy number variant analysis in patients with acquired long QT syndrome.* Europace, 2015. **17**(4): p. 635-41.
37. Zumhagen, S., et al., *Inherited long QT syndrome: clinical manifestation, genetic diagnostics, and therapy.* Herzschrittmacherther Elektrophysiol, 2012. **23**(3): p. 211-9.
38. Schwartz, P.J., et al., *Impact of genetics on the clinical management of channelopathies.* J Am Coll Cardiol, 2013. **62**(3): p. 169-80.
39. Priori, S.G., C. Napolitano, and P.J. Schwartz, *Low penetrance in the long-QT syndrome: clinical impact.* Circulation, 1999. **99**(4): p. 529-33.
40. Crotti, L., et al., *KCNH2-K897T is a genetic modifier of latent congenital long-QT syndrome.* Circulation, 2005. **112**(9): p. 1251-8.
41. Purvis, I.W. and I.R. Franklin, *Major genes and QTL influencing wool production and quality: a review.* Genet Sel Evol, 2005. **37 Suppl 1**: p. S97-107.
42. Crotti, L., et al., *NOS1AP is a genetic modifier of the long-QT syndrome.* Circulation, 2009. **120**(17): p. 1657-63.
43. Shah, R.R., *Drug-induced prolongation of the QT interval: why the regulatory concern?* Fundam Clin Pharmacol, 2002. **16**(2): p. 119-24.
44. Shah, R.R., *Drug-induced QT interval prolongation: does ethnicity of the thorough QT study population matter?* Br J Clin Pharmacol, 2013. **75**(2): p. 347-58.
45. Jawad-Ul-Qamar, M. and P. Kirchhof, *Almanac 2015: atrial fibrillation research in Heart.* Heart, 2016. **102**(8): p. 573-80.

46. Buiten, M.S., et al., *The dialysis procedure as a trigger for atrial fibrillation: new insights in the development of atrial fibrillation in dialysis patients.* Heart, 2014. **100**(9): p. 685-90.
47. Soliman, E.Z., et al., *Chronic kidney disease and prevalent atrial fibrillation: the Chronic Renal Insufficiency Cohort (CRIC).* Am Heart J, 2010. **159**(6): p. 1102-7.
48. Fox, C.S., et al., *Parental atrial fibrillation as a risk factor for atrial fibrillation in offspring.* JAMA, 2004. **291**(23): p. 2851-5.
49. Ellinor, P.T., et al., *Familial aggregation in lone atrial fibrillation.* Hum Genet, 2005. **118**(2): p. 179-84.
50. Arnar, D.O., et al., *Familial aggregation of atrial fibrillation in Iceland.* Eur Heart J, 2006. **27**(6): p. 708-12.
51. Gundlund, A., et al., *Familial clustering and subsequent incidence of atrial fibrillation among first-degree relatives in Denmark.* Europace, 2016. **18**(5): p. 658-64.
52. Olesen, M.S., et al., *Atrial fibrillation: the role of common and rare genetic variants.* Eur J Hum Genet, 2014. **22**(3): p. 297-306.
53. Mahida, S., *Transcription factors and atrial fibrillation.* Cardiovasc Res, 2014. **101**(2): p. 194-202.
54. Baruteau, A.E., V. Probst, and H. Abriel, *Inherited progressive cardiac conduction disorders.* Curr Opin Cardiol, 2015. **30**(1): p. 33-9.
55. Friedrich, C., et al., *Gain-of-function mutation in TASK-4 channels and severe cardiac conduction disorder.* EMBO Mol Med, 2014. **6**(7): p. 937-51.
56. Gourraud, J.B., et al., *Identification of a strong genetic background for progressive cardiac conduction defect by epidemiological approach.* Heart, 2012. **98**(17): p. 1305-10.
57. Remme, C.A., A.A. Wilde, and C.R. Bezzina, *Cardiac sodium channel overlap syndromes: different faces of SCN5A mutations.* Trends Cardiovasc Med, 2008. **18**(3): p. 78-87.
58. Keller, D.I., L. Carrier, and K. Schwartz, *Genetics of familial cardiomyopathies and arrhythmias.* Swiss Med Wkly, 2002. **132**(29-30): p. 401-7.
59. Kruse, M., et al., *Impaired endocytosis of the ion channel TRPM4 is associated with human progressive familial heart block type I.* J Clin Invest, 2009. **119**(9): p. 2737-44.
60. McCulley, D.J. and B.L. Black, *Transcription factor pathways and congenital heart disease.* Curr Top Dev Biol, 2012. **100**: p. 253-77.
61. Algra, A., et al., *QTc prolongation measured by standard 12-lead electrocardiography is an independent risk factor for sudden death due to cardiac arrest.* Circulation, 1991. **83**(6): p. 1888-94.
62. de Wert, G., *Cascade screening: whose information is it anyway?* Eur J Hum Genet, 2005. **13**(4): p. 397-8.
63. Bush, W.S. and J.H. Moore, *Chapter 11: Genome-wide association studies.* PLoS Comput Biol, 2012. **8**(12): p. e1002822.
64. Shah, S., et al., *Four genetic loci influencing electrocardiographic indices of left ventricular hypertrophy.* Circ Cardiovasc Genet, 2011. **4**(6): p. 626-35.
65. Schillaci, G., et al., *Improved electrocardiographic diagnosis of left ventricular hypertrophy.* Am J Cardiol, 1994. **74**(7): p. 714-9.
66. Visscher, P.M., et al., *10 Years of GWAS Discovery: Biology, Function, and Translation.* Am J Hum Genet, 2017. **101**(1): p. 5-22.
67. Janssens, A.C. and C.M. van Duijn, *Genome-based prediction of common diseases: advances and prospects.* Hum Mol Genet, 2008. **17**(R2): p. R166-73.
68. Bezzina, C.R., et al., *Common variants at SCN5A-SCN10A and HEY2 are associated with Brugada syndrome, a rare disease with high risk of sudden cardiac death.* Nat Genet, 2013. **45**(9): p. 1044-9.

69. Smith, J.G., et al., *Genome-wide association study of electrocardiographic conduction measures in an isolated founder population: Kosrae.* Heart Rhythm, 2009. **6**(5): p. 634-41.
70. Russell, M.W., et al., *Heritability of ECG measurements in adult male twins.* J Electrocardiol, 1998. **30 Suppl**: p. 64-8.
71. Dalageorgou, C., et al., *Heritability of QT interval: how much is explained by genes for resting heart rate?* J Cardiovasc Electrophysiol, 2008. **19**(4): p. 386-91.
72. Havlik, R.J., et al., *Variability of heart rate, P-R, QRS and Q-T durations in twins.* J Electrocardiol, 1980. **13**(1): p. 45-8.
73. Haarmark, C., et al., *Heritability of Tpeak-Tend interval and T-wave amplitude: a twin study.* Circ Cardiovasc Genet, 2011. **4**(5): p. 516-22.
74. van der Harst, P., et al., *52 Genetic Loci Influencing Myocardial Mass.* J Am Coll Cardiol, 2016. **68**(13): p. 1435-48.
75. Verweij, N., et al., *Twenty-eight genetic loci associated with ST-T-wave amplitudes of the electrocardiogram.* Hum Mol Genet, 2016. **25**(10): p. 2093-2103.
76. Arking, D.E., et al., *Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization.* Nat Genet, 2014. **46**(8): p. 826-36.
77. Marroni, F., et al., *A genome-wide association scan of RR and QT interval duration in 3 European genetically isolated populations: the EUROSPAN project.* Circ Cardiovasc Genet, 2009. **2**(4): p. 322-8.
78. van der Harst, P., et al., *52 Genetic Loci Influencing Myocardial Mass.* J Am Coll Cardiol, 2016. **68**(13): p. 1435-1448.
79. Im, S.W., et al., *Analysis of genetic and non-genetic factors that affect the QTc interval in a Mongolian population: the GENDISCAN study.* Exp Mol Med, 2009. **41**(11): p. 841-8.
80. Holm, H., et al., *Several common variants modulate heart rate, PR interval and QRS duration.* Nat Genet, 2010. **42**(2): p. 117-22.
81. Kolder, I.C., M.W. Tanck, and C.R. Bezzina, *Common genetic variation modulating cardiac ECG parameters and susceptibility to sudden cardiac death.* J Mol Cell Cardiol, 2012. **52**(3): p. 620-9.
82. Newton-Cheh, C., et al., *Common variants at ten loci influence QT interval duration in the QTGEN Study.* Nat Genet, 2009. **41**(4): p. 399-406.
83. Mutikainen, S., et al., *Genetic influences on resting electrocardiographic variables in older women: a twin study.* Ann Noninvasive Electrocardiol, 2009. **14**(1): p. 57-64.
84. Amin, N., et al., *Refining genome-wide linkage intervals using a meta-analysis of genome-wide association studies identifies loci influencing personality dimensions.* Eur J Hum Genet, 2013. **21**(8): p. 876-82.

CHAPTER 2


ATTCGGTACCTCATAACGGCTAGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTT
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGGTACCTA

CHAPTER 2

Heritabilities, proportion of heritabilities explained by GWAS findings, and implications of cross-phenotype effects of PR interval.

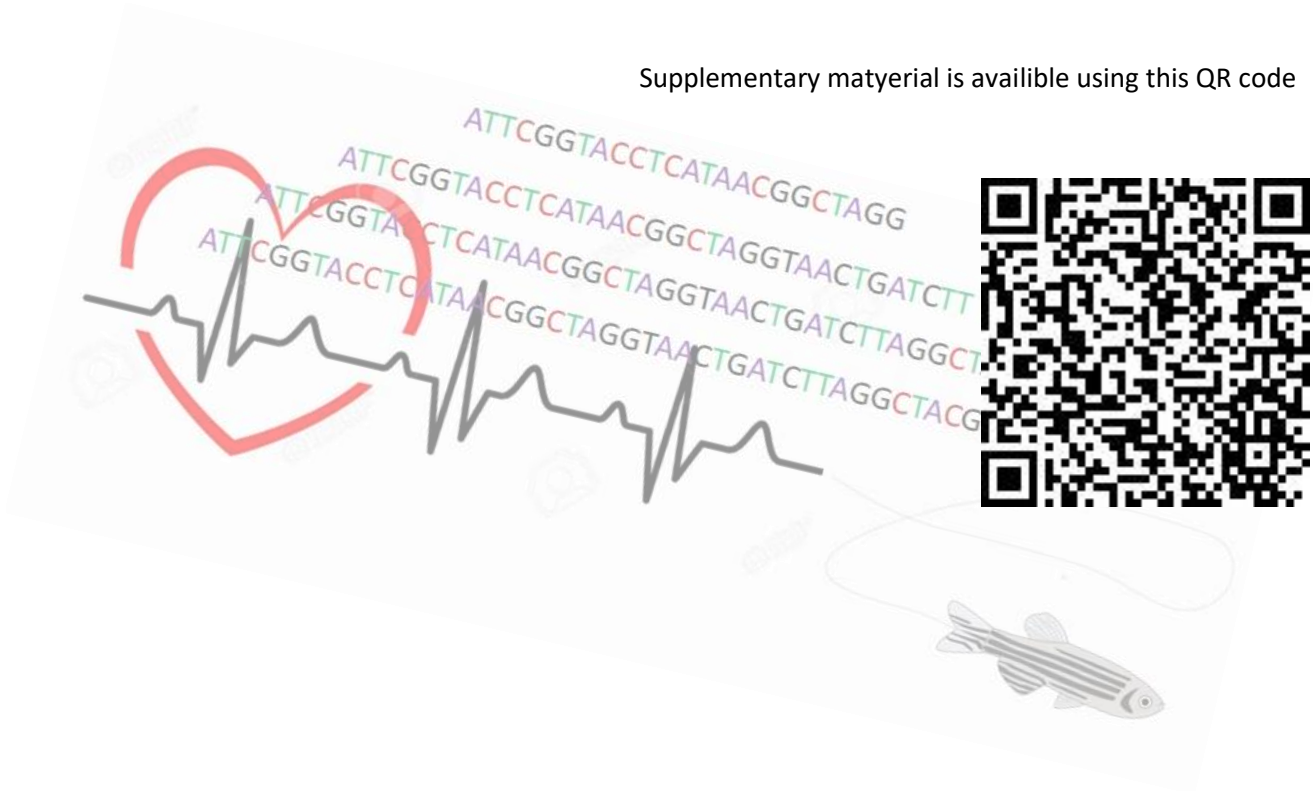
Claudia Tamar Silva, Jan A. Kors, Ph.D.4, Najaf Amin, Abbas Dehghan, Jacqueline C.M. Witteman, Rob Willemsen, Ben A. Oostra, Cornelia M. Van Duijn, Aaron Isaacs.

Hum Genet. 2015 Nov; 134(11-12):1211-9.

The supplemental information for this paper is available in Chapter 7.1.

doi: 10.1007/s00439-015-1595-9. Epub 2015 Sep 18. PMID:2638555

Supplementary matyerial is availible using this QR code



Abstract

Electrocardiogram (ECG) measurements are a powerful tool for evaluating cardiac function and are widely used for the diagnosis and prediction of a variety of conditions, including myocardial infarction, cardiac arrhythmias, and sudden cardiac death. Recently, genome-wide association studies (GWASs) identified a large number of genes related to ECG parameter variability, specifically for the QT, QRS, and PR intervals. The aims of this study were to establish the heritability of ECG traits, including indices of left ventricular hypertrophy, and to directly assess the proportion of those heritabilities explained by GWAS variants. These analyses were conducted in a large, Dutch family-based cohort study, the Erasmus Rucphen Family study using variance component methods implemented in the SOLAR (Sequential Oligogenic Linkage Analysis Routines) software package. Heritability estimates ranged from 34 % for QRS and Cornell voltage product to 49 % for 12-lead sum. Trait-specific GWAS findings for each trait explained a fraction of their heritability (17 % for QRS, 4 % for QT, 2 % for PR, 3 % for Sokolow–Lyon index, and 4 % for 12-lead sum). The inclusion of all ECG-associated single nucleotide polymorphisms explained an additional 6 % of the heritability of PR. In conclusion, this study shows that, although GWAS explain a portion of ECG trait variability, a large amount of heritability remains to be explained. In addition, larger GWAS for PR are likely to detect loci already identified, particularly those observed for QRS and 12-lead sum.

Chapter 2

FOR EDITORIAL REASONS, THIS CHAPTER IS ONLY AVAILABLE IN THE JOURNAL OR

THROUGH PUBMED (Free Access)

Chapter 3


ATTCGGTACCTCATAACGGCTAGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTT
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGGTACCTA

Chapter 3

A combined linkage and exome sequencing analysis for ECG parameters in the Erasmus Rucphen Family study

Claudia Tamar Silva, Irina V. Zorkoltseva, Najaf Amin, Ayşe Demirkan, Elisa van Leeuwen, Jan A. Kors, Marten van den Berg, Bruno H. Stricker, André G. Uitterlinden, Anatoly V. Kirichenko, Jacqueline C.M. Witteman, Rob Willemsen, Ben A. Oostra, Tatiana I. Axenovich, Cornelia M. van Duijn*, Aaron Isaacs*.

* These authors contributed equally to this work.

Front Genet. 2016 Nov 8; 7:190. eCollection 2016.

PMID: 27877193

Supplementary matyerial is availible using this QR code

Abstract

Electrocardiogram (ECG) measurements play a key role in the diagnosis and prediction of cardiac arrhythmias and sudden cardiac death. ECG parameters, such as the PR, QRS, and QT intervals, are known to be heritable and genome-wide association studies of these phenotypes have been successful in identifying common variants; however, a large proportion of the genetic variability of these traits remains to be elucidated. The aim of this study was to discover loci potentially harboring rare variants utilizing variance component linkage analysis in 1547 individuals from a large family-based study, the Erasmus Rucphen Family Study (ERF). Linked regions were further explored using exome sequencing. Five suggestive linkage peaks were identified: two for QT interval (1q24, LOD = 2.63; 2q34, LOD = 2.05), one for QRS interval (1p35, LOD = 2.52) and two for PR interval (9p22, LOD = 2.20; 14q11, LOD = 2.29). Fine-mapping using exome sequence data identified a C > G missense variant (c.713C > G, p.Ser238Cys) in the *FCRL2* gene associated with QT (rs74608430; $\boldsymbol{P}$ = 2.8 × 10^{-4}, minor allele frequency = 0.019). Heritability analysis demonstrated that the SNP explained 2.42% of the trait's genetic variability in ERF ($\boldsymbol{P}$ = 0.02). Pathway analysis suggested that the gene is involved in cytosolic Ca2+ levels ($\boldsymbol{P}$ = 3.3 × 10^{-3}) and AMPK stimulated fatty acid oxidation in muscle ($\boldsymbol{P}$ = 4.1 × 10^{-3}). Look-ups in bioinformatics resources showed that expression of *FCRL2* is associated with *ARHGAP24* and *SETBP1* expression. This finding was not replicated in the Rotterdam study. Combining the bioinformatics information with the association and linkage analyses, *FCRL2* emerges as a strong candidate gene for QT interval.

Keywords: genetics, epidemiology, electrocardiography, linkage, exome

Introduction

The electrocardiogram (ECG) is an important tool for diagnosing, monitoring and evaluating risk in patients with cardiovascular disease (CVD; [1, 2]. ECG measurements, such as PR interval, QRS complex duration, and QT interval, are used for the diagnosis and prediction of cardiac arrhythmias and sudden cardiac death (SCD; [3]. Myocardial depolarization and repolarization time are measured by the QT interval: the time between the onset of the QRS complex and the end of the T wave. QT shortening or prolongation has been associated with an increased risk for arrhythmias and SCD [4]. PR interval and QRS duration are measures of cardiac conduction time; QRS duration reflects conduction through the ventricular myocardium, while PR interval measures atrial and atrioventricular conduction from the sinoatrial node to the ventricular myocardium, primarily through the atrioventricular node [5, 6].

There are significant genetic contributions to ECG measurements; genome-wide association studies (GWAS) identified at least 71 common variants associated with their variability [7-14]. A number of these associations were established in loci containing genes that encode proteins with previously known roles in heart development and function, such as cardiac transcription factors; sodium, calcium, and potassium ion channels; genes with a role in myocardial electrophysiology; and others involved in the conduction of electrical impulses [3]. These include *ARHGAP24, SETBP1, LRIG1, CREBBP, MEIS1*. *TBX20*, and *TBX5*. Some ion channel encoding genes, such as *SCN5A, HERG, KCNE1*, and *KCNE2*, have been associated with long QT syndrome (LQTS; [15], atrial fibrillation (AF) and

Brugada Syndrome [16]. Collectively, however, these loci explain only modest proportions of phenotypic variability; GWAS SNPs specific for each trait account for limited trait heritability (17% for QRS, 4% for QT, and 2% for PR) [17].

Genome-wide association studies generally interrogate only common variants, typically of small effect. Families, in addition to being robust against population stratification, may be enriched for less frequent variants, which can potentially be identified by linkage and fine mapping. The aim of this study, therefore, was to discover less frequent variants using linkage analysis in a large family-based study, the Erasmus Rucphen Family Study (ERF).

Methods

Study population

The ERF study, which is a part of the Genetic Research in Isolated Populations (GRIP) Program, is a family-based study including over 3000 participants descendant from 22 couples that lived in the Rucphen region in the southwest Netherlands in the 19th century [18]. All descendants of those couples were invited to visit the clinical research center in the region, where they were examined in person [19]. Interviews at the time of blood sampling were performed by medical practitioners and included questions on a broad range of topics, including current medication use and medical history [20]. Height and weight were measured with the participant in light underclothing and body mass index (kg/m2) was computed. Blood pressure (BP) was measured twice on the right arm in a sitting position after at least five minutes rest, using an automated device (OMRON 711, Omron Healthcare, Bannockburn, IL, USA). The average of the two measures was used for analysis. Hypertension was defined through the use of antihypertensive medication and/or through the assessment of BP measurements according to the World Health Organization [21] guidelines (individuals with BP ≥ 140/90 mmHg should be regarded as hypertensive). The Medical Ethics

Committee of the Erasmus University Medical Center appro[22]ved the ERF study protocol and all participants, or their legal representatives, provided written informed consent.

ECG measurement and interpretation

Examinations included 10 s 12-lead ECG measurements, recorded with an ACTA-ECG (Esaote, Florence, Italy) with a sampling frequency of 500 Hz. Digital measurements of the ECG parameters were made using the Modular ECG Analysis System (MEANS; [22]. Briefly, MEANS operates on multiple simultaneously recorded leads, which are transformed to a detection function that brings out the QRS complex and the other parts of the signal. MEANS determines common onsets and offsets for all 12 leads together on one representative averaged beat, with the use of template matching techniques. The measurement and diagnostic performance of MEANS have been extensively evaluated, both by the developers and by others [22-26]. The MEANS criteria for MI are mainly based on pathological Q waves, QR ratio, and R-wave progression [27]. A cardiologist, specialized in ECG methodology, ascertained the final diagnosis of MI. QT interval was corrected for heart rate using Bazett's formula in all analyses [28].

Genotyping and statistical analyses of the linkage study

Illumina's HumanHap6k Genotyping BeadChip (6K Illumina Linkage IV Panels^{R}) was used for genotyping for the linkage analyses. All genotyping procedures were performed according to the manufacturer's protocols. Only markers with minor allele frequency (MAF) > 0.05 were selected for further analysis. Genotyping errors leading to Mendelian inconsistencies were detected using PedCheck [29]. Unlikely double recombination events were detected using MERLIN [30]. All observed Mendelian errors were eliminated from the data. A total of 5250 autosomal SNPs with a call rate greater than 95% were included in the linkage analyses. All traits were adjusted for age, sex, BMI and height and inverse-normal transformation of ranks was applied before analysis. One thousand five hundred and forty-seven people with complete ECG, covariate, and genotype data

were included in the initial analysis. Variance component multipoint linkage was performed using the –vc option in the MERLIN v.1.0.1 software [30, 31]. This program calculates exact IBD sharing probabilities using the Lander-Green algorithm, requiring restriction of pedigree size. Because of this, the large single ERF pedigree with multiple loops was split into non-overlapping fragments of no more than 18 bits with the help of the PedSTR program [32]. Final variance component two-point linkage analysis for the identified *FCRL2* variant (rs74608430) was performed using Merlin in one large, single pedigree.

Regions of interest with LOD > 1.9 were selected for further study [33]. Borders of the linkage areas were defined as LOD score minus 2 support intervals (LOD-2 SI) around the linkage peaks. Genes within the LOD-2 SI were annotated using SCAN (SNP and CNV Annotation Database[1]).

Exome-sequencing

Exomes for 1336 individuals from ERF were sequenced at the Center for Biomics, Department of Cell Biology, Erasmus MC, the Netherlands, using the Agilent V4 capture kit on an Illumina HiSeq2000 sequencer using the TruSeq Version 3 protocol. Mean depth base was 74.23x (median = 57x) and mean depth region was 65.26x (median = 52.87x). The sequence reads were aligned to the human genome build 19 (hg19) using BWA and the NARWHAL pipeline [34, 35]. The aligned reads were processed further using the IndelRealigner, MarkDuplicates, and TableRecalibration tools from the Genome Analysis Toolkit (GATK) and Picard[2] to remove systematic biases and to recalibrate the PHRED quality scores in the alignments. Genetic variants were called using the Unified Genotyper tool of the GATK. About 1.4 million Single Nucleotide Variants (SNVs) were called and, after removing the low quality variants (QUAL < 150), we retrieved 577,703 SNVs in 1,309 individuals. Linear regression analyses, with SNVs in an additive model, were conducted on ECG measures, adjusted for age, sex, BMI, and height. To reduce the burden of multiple testing,

we assessed only damaging variants in the LOD-2 SI; we found 324 such variants for QT, 52 for QRS and 61 for PR. We employed a Bonferroni correction for the number of deleterious mutations selected for each trait (QT: $\boldsymbol{P} = 1.5 \times 10^{-4}$, QRS: $\boldsymbol{P} = 9.6 \times 10^{-4}$, and PR: $\boldsymbol{P} = 8.2 \times 10^{-4}$). The proportion of trait variance explained by the SNP was calculated using the Merlin software [30].

Replication

We sought to replicate our findings in the Rotterdam Study (RS) cohort. The RS is an ongoing prospective cohort study conducted since 1990 in the city of Rotterdam in The Netherlands [36]. The Illumina Exome BeadChip array ("exome chip") was developed through a large international initiative to efficiently study coding variants spanning the genome. The v1.0 array contains 247,870 variants, which were genotyped in 3,183 individuals from the RS population. Calling for this sample, and numerous others, was done centrally (in total, 62,267 samples). After rigorous quality control and exclusion of variants that were monomorphic or too rare to analyze, the final dataset consisted of 108,678 polymorphic variants in 3,163 individuals.

Bioinformatics analysis

To predict the functionality of genetic variants, and for comparison to BWA and NARWHAL, annotations were also performed using the dbNSFP (database of human non-synonymous SNPs and their functional predictions[3] and Seattle[4] databases. These databases gave functional prediction results from four different programs (PolyPhen-2, SIFT, MutationTaster, and LRT) [37-40], in addition to gene and variant annotations. Genes containing nominally significant variants (Table 2) were analyzed using Ingenuity Pathway Analysis (IPA; Ingenuity systems Inc, Redwood city, CA, USA). Several IPA modules were implemented: the "core analysis" was used to assess pathways, relationships, and mechanisms relevant to the dataset; the "upstream regulator analysis" was implemented to identify molecules (including microRNA and transcription factors) that may affect expression levels; and the "downstream effects analysis" was utilized to predict

downstream biological processes that are increased or decreased[5]. The GEO2R[6] tool was used to analyse microarray-based expression data in the GEO database (GEO Accession numbers: GSE2240 and GSE41177). The Gene Network tool[7] was used to describe co-expression networks and to assess potential functional effects of identified genes.

Table 2. Selection of the coding variants

Trait	Locus	Variants in the coding region					Observations <=5%			Predicted to be damaging			
		Synonymous	Missense	Stop	Splice	Total	Missense	Stop	Splice	Missense	Stop	Splice	Genes
QT	1	3110	5089	117	36	8353	660	0	4	207	2	0	*DENND2C, RWDD3, FCER1A, GPR25, CD1C, OMA1, LIX1L, LRRC8B,TPR, HOOK1, GTF2B, TXNIP, DDR2, CNN3, RBM15, BCL9, IVNS1ABP TNN, CEPT1, ACOT11, SARS, VAV3, TOMM40L, GABPB2, RFX5, ETV3L, APOBEC4, KIAA1614 ASPM,SPRR3, CEP350, C1orf168 COL24A1, SEMA6C, C1orf49, CACNA1S, IVL, VSIG8, EDEM3, HMCN1, TBX19, GLRX2, IFI16, PODN,INADL, MPL, HYI, CAPZA1, AMIGO1, HCN3, RTCD1, OR10J1, FLG, DMRTB1,SPTA1, HFM1,CFHR2, FCRL2, NCF2, CHIA, RBMXL1,C8A, SGIP1, FMO4, GBP1, CELSR2,ODF2L, PEAR1, FCRL1, SLC44A5, UROD,MOBKL2C, LRRC7, LRRC8C,IPO9,PRPF38B, MSH4, KIFAP3, LAMC2, PAQR6, ZNF687,MIER1, SMG7, TMEM61, ALX3,FAM189B, PDE4DIP, ATPAF1,C1orf50, PRRC2C, ZNF281, IGSF3, CRCT1, UQCRH, SLC27A3, NPHS2, PKLR, ATP1A4,TMEM125, TNR, , OVGP1, SHCBP1L, UHMK1, B4GALT2 RNF220,PIAS3, KIF2C, TARS2, TMEM59,PIGK,CMPK1, PIK3R3, METTL11B CITED4,EFCAB7,TTF2, AXDND1, DDX20, IGSF9, LEPRE1,ADAMTSL4 WDR77, GNAT2, GPSM2, PPM1J, ABCA4, EXTL2, AP4B1, HIVEP3, UBQLN4, POLR3C, NEGR1, TBX15, GBP6, KIAA1324 DPYD,F5, GJA5,CYP4A22, HENMT1,MRPL37,TDRD5, ZBTB7B, SPATA6, FCRLB, ABL2, ZFYVE9, LAMC1, RHBG, DUSP12, ZYG11A, WDR3, FAAH, C1orf106 HSD3B1, CTSS, TRIM45, ALG6, ACP6, PRUNE, TRIM46, AGL, MAGI3, C1orf27, AL359075.1 SLC5A9,*

												EBNA1BP2, COL11A1, FGGY, AMPD1, FAM63A, GLT25D2, DMRTA2, EVI5, DPT, OR6P1
	2	1662	2444	40	17	4165	328	0	2	113	2	0	*CRYGA, TTN, ARMC9, GTF3C3, ADAM23, ZFAND2B, PER2, COL6A3, TNS1, PAX3, HDAC4, OBSL1, CAPN10, IGFBP5, TMEM198, ESPNL, SPAG16, COL4A3, ANKAR, NEUROD1, NOP58, DNAH7, IQCA1, CCDC141, KIF1A, CASP10, SSFA2, CRYGC, ECEL1, AP1S3, COL5A2, NDUFS1, ATF2, STK36, UNC80, ABCB6, KIAA1486, ANKMY1, C2orf67, PLEKHM3, CNPPD1, ALPP, EFHD1, ZSWIM2, C2orf62, AQP12B, WIPF1, PDE11A, GLB1L, CCDC150, DGKD, SERPINE2, ABCA12, ITGAV, IDH1, SPHKAP, FN1, CDK15, GPR35, WNT10A, CYP27A1, ACSL3, ANKZF1, DNAJC10, FBXO36, STK16, MYO1B, KLHL30, PIKFYVE, DES, ASNSD1*
QRS	1	1057	1446	25	17	2546	152	1	4	51	1	0	*OTUD3, PHC2, SYF2, DHDDS, EPB41, NBPF3, ZBTB40, COL16A1, RAP1GAP, C1orf38, EPHA10, MACF1, PADI4, LDLRAP1, RCC2, AK2, SEPN1, TMCO2, HSPG2, MAP3K6, TMCO4, CCDC28B, TMEM234 GRHL3, ALDH4A1, GJB4, MAN1C1, SERINC2, E2F2, MUL1, PHACTR4, MYOM3, SRRM1, RLF, TINAGL1, KIAA0319L, C1orf94, C1orf63, UBXN11, USP48*
PR	9	375	656	8	5	1053	96	0	0	29	0	0	*DENND4C, CA9, FRMPD1, PLIN2, CCIN, IFT74, UBAP1, IFNA10, RECK, UNC13B, GRHPR, KIAA1045, FREM1, OR2S2, IFNA14, FAM154A, KIAA1797, RGP1, ALDH1B1, NOL6, (GALT; GALT; RP11-195F19.29), PTPLAD2, DDX58*
	14	440	792	24	6	1276	86	0	1	31	1	0	*HEATR5A, RABGGTA, LRRC16B, RBM23, CMA1, SUPT16H, MMP14, PARP2, CEBPE,*

	OR4K1, PRKD1, LRRC16B, MYH6, PSMB11, HEATR5A, LRP10, LRRC16B, TTC5, OR10G3, OR4N5, MYH6, TEP1, SDR39U1, TEP1, SLC7A7, LRP10, TEP1, ADCY4, (AL163636.2;AL163636.2;AL163636.2;RNASE4;RNASE4; RNASE4), PCK2, ARHGEF40, KLHL33

Table 2. Selection of the coding variants

Results

Table 1 shows the characteristics of the participants included in the discovery linkage analyses and exome sequencing, as well as the exome chip replication sample. There were no significant differences between the largely overlapping linkage and exome sequence groups. The replication sample was considerably older, and was characterized by increased frequency of hypertension (and BP differences), increased PR interval and decreased QT interval compared to the discovery samples. The three ECG traits studied (the QT, QRS, and PR intervals) demonstrated only modest pair-wise correlations in the discovery dataset (Supplementary Table 1).

	Linkage Studies			Exome-sequence		
	ERF = 1860			ERF = 1309		
	Mean (S.D.)	Minimum	Maximum	Mean (S.D.)	Minimum	Maximum
Males (n, %)	775 (42%)			509 (40%)		
Age (years)	46.4 (13.8)	16.6	85.3	47.7 (14.1)	18.2	86.1
BMI (kg/m2)	26.6 (4.6)	15.5	61.8	26.6 (4.4)	15.5	61.8
Height (cm)	167.4 (9.1)	143.6	196.5	166.8 (9.1)	141.0	196.5
Weight (kg)	75.9 (15.1)	41.9	161.0	74.29 (14.5)	42.1	161.0
SBP (mm Hg)	137.7 (19.1)	85.5	217.0	138.3 (19.6)	85.5	239.0
DBP (mm Hg)	79.6 (9.7)	54.5	120.0	79.5 (9.7)	53.5	127.5
Hypertension	766 (41.1%)			549 (43%)		
PR	152 (22.4)	92	308	152.8 (22.4)	96.0	308.0
QT	403.1 (22.4)	336.0	531.0	403.8 (22.0)	336.0	531.0
QRS	96.8 (9.9)	68.0	120.0	96.7 (9.9)	68.0	120.0

Table 1. Descriptive statistics of the study population

Supplementary Table 2 shows the linkage results for the ECG traits, which yielded a total of five regions with suggestive LOD scores (LOD > 1.9). QT was suggestively linked to two regions, on chromosome 1 (LOD = 2.63) and on chromosome 2 (LOD = 2.05). A suggestive LOD score for QRS was observed on chromosome 1 (LOD = 2.52) and, for PR, two suggestive regions were located on chromosomes 9 and 14 with LOD scores of 2.20 and 2.29, respectively (Supplementary Table 2). Plots of the linked regions are shown in Figure 1.

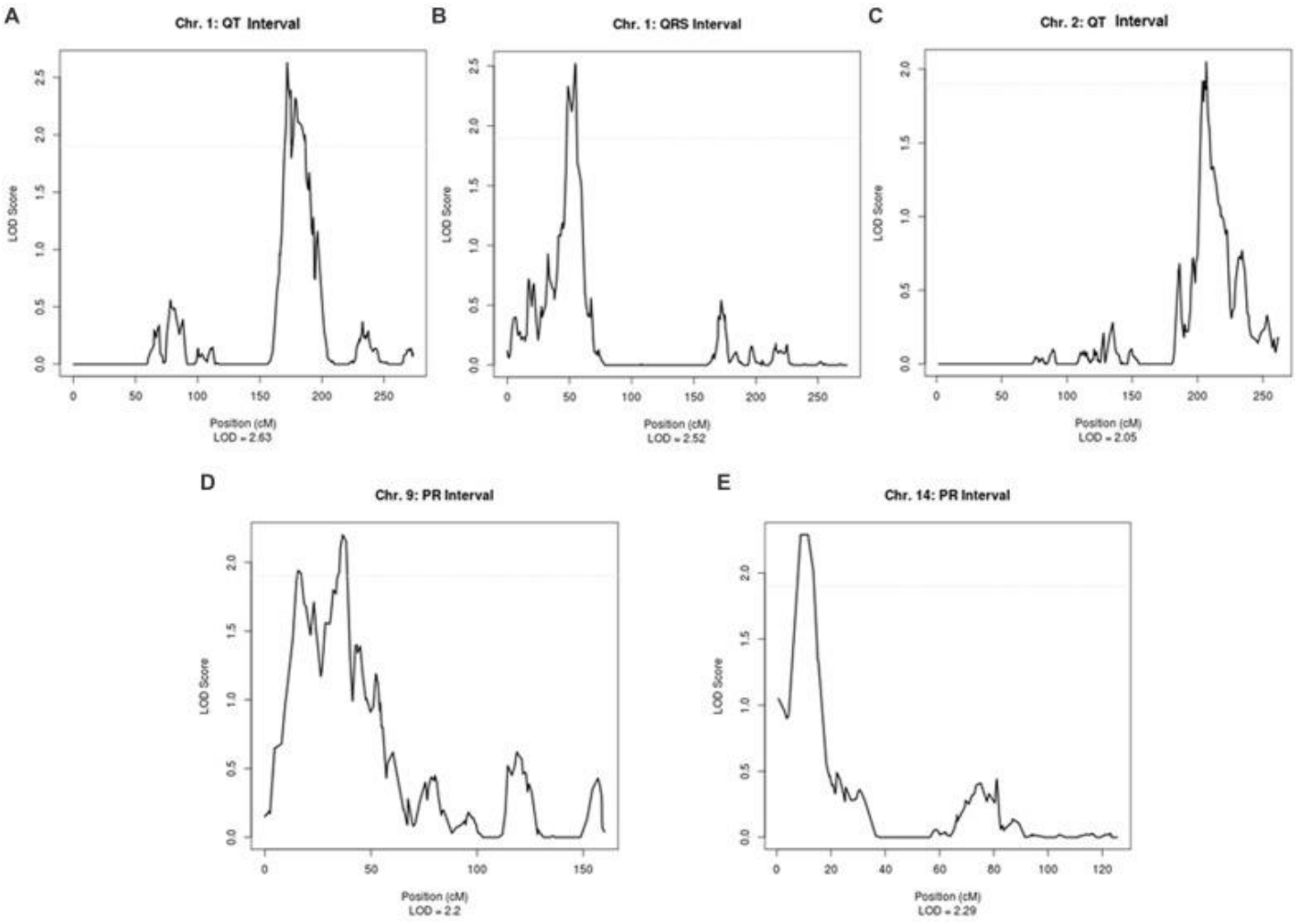


Figure 1. Linkage peaks for ECG traits

Our analysis of coding variants in these linked regions revealed 55,050 variants in coding regions of genes under the peaks, as described in Table 2. Of these mutations, 1334 had a frequency less than or equal to 5%, 437 were predicted to be damaging by at least two of the prediction software packages used, and six were nonsense variants. By linkage peak, there were 207 missense damaging mutations and two nonsense mutations on 14 and 113 missense damaging mutations and two nonsense mutations on 2q32 for QT; 51 missense mutations and one nonsense mutation on 1p36 for QRS; and 29 missense mutations on 9q21 and 31 missense mutations and one nonsense mutation on 14q12 for PR. In total, 21 variants had nominal regression ***P***-values less than 0.05 (the smallest ***P***-values under each linkage peak were ***P*** = 2.8 × 10-^{4} for QT on chromosome 1, ***P*** = 2.3 × 10^{-2} for QT on chromosome 2, ***P*** = 2.6 × 10^{-2} for QRS on chromosome 1, ***P*** = 1.9 × 10^{-2} for PR on chromosome 9, and ***P*** = 1.9 × 10^{-2} for PR on chromosome 14) without reaching the significance levels needed to account for multiple comparisons (Supplementary Table 3). Looking for known genes under the linkage peaks (Supplementary Table 4), we found two variants previously related to heart failure, *TTN* (rs72648923; P = 5.5 × 10-2, MAF = 1.4 × 10-2) and *HSD3B1* (P = 3.9 × 10-2 MAF = 1.1 × 10-2). Neither achieved statistical significance after Bonferroni correction, although both genes were marginally associated with QT. Only a single variant, a C > G (Ser > Cys) variant in *FCRL2* (rs74608430; ***P*** = 2.8 × 10-4, MAF = 1.9 × 10^{-2}), approached the Bonferroni threshold for multiple-testing (***P*** = 1.5 × 10^{-4}). This variant, under the linkage peak on chromosome 1q23.1 for QT, is highly conserved (scorePhastCons = 0.998) and also predicted by PolyPhen-2 to be damaging (0.999). In the whole ERF population, rs74608430 explained 2.42% of the heritability of QT (reducing the LOD to 1.1; h^2 = 0.87%; ***P*** = 0.02). This finding was not replicated in the RS (***P*** = 0.12, β = 0.14). A sequence kernel association test analysis of the gene also failed to achieve significance in the replication sample (***P*** = 0.44).

Not much is known about the function of *FCRL2*. Among the functions predicted by Gene Network are the regulation of cytosolic Ca^{2+} levels (***P*** = 3.3 × 10^{-3}) and AMPK stimulated fatty acid oxidation in muscle (***P*** = 4.1 × 10^{-3}). In the GEO database, *FCRL2* expression was higher in AF [41, 42]. Supplementary Figure

1A shows the genes co-expressed with *FCRL2*, according to Gene Network. Two genes that have been associated with ECG outcomes by GWAS emerge: *ARHGAP24*, associated with PR, and *SETBP1*, associated with QRS [12-14]. In the chromosome 1 region linked to QT, looking for co-expression, we found correlations between *DMRTA2*. *CEP350*, and *MPL* with genes previously associated with ECG traits: *DMRTA2* is co-expressed with *LRIG1*, a QRS associated gene (Supplementary Figure 1B); *MPL* is in a module with *MEIS1*, associated with PR (Supplementary Figure 1C); and *CEP350* interacts with *CREBBP*, associated with QT (Supplementary Figure 1D). These three genes are not in linkage disequilibrium with each other. At the chromosome 2q34 locus linked with QT, a heart failure gene, *TTN,* was under the linkage peak. According to Gene Network analysis, expression of TTN is related to expression of three previously known QT genes (*ATP1B, TCEA3*, and *PLN*) and two QRS and PR associated genes (*TBX20* and *TBX5*) (Supplementary Figure 1E) [8, 12-14]. Additionally, *SPHKAP*, on chromosome 2 under the QT linkage peak, is co-expressed with *TBX5* (Supplementary Figure 1F).

Discussion

Linkage analysis is an important tool for the identification of genomic regions influencing trait variability. The role of *TPM1* mutations with sudden death is a clear example of a locus discovered by linkage analysis [43, 44]. The advantages of family studies include control of heterogeneity and population stratification [43, 45]. We performed a linkage study on ECG measurements and identified five suggestive regions (1p35.1, 1q24.2, 2q34, 9p22.2, 14q11.2). Rare variant analysis in these regions uncovered two genes related to heart failure, *TTN* (P = 5.5 × 10-2) and *HSD3B1* (P = 3.9 × 10-2) and one gene with unknown cardiac function *FCRL2* (P = 2.8 × 10-4). None of them reaches statistical significance level after correction for multiple comparisons.

This study was conducted in a large, well-characterized family-based cohort, ascertained on the basis of genealogy and not phenotype. Multiple levels of genetic data, including a linkage panel and exome sequence data, provided a powerful dataset for identifying variants that may not be easily discovered with GWAS. Unfortunately, exome data was not available in the whole cohort, which could limit our ability to identify causal variants. Additionally, the sequence data did not include extra-genic or intronic variants that may be responsible for the observed linkage peaks.

Our analysis of rare coding variants in these linkage regions revealed 55,050 variants in coding regions. One thousand three hundred and thirty-four of these mutations had a frequency less than or equal to 5% and 437 were predicted to be damaging; none reached the significance threshold accounting for multiple comparisons. These variants spanned genes, including *TTN* and HSD3B1, which have been previously related to CVDs. HSD3B, a gene on chromosome 1 (1p13.1), has two isoforms (HSD3B1 and HSD3B2) that were found to be associated with an increase in plasma aldosterone [46]. Changes in circulating aldosterone levels can modulate BP and hypertrophy (HT). A genome wide linkage analysis revealed that *HSD3B1* is a locus for BP variation [46].

Another interesting gene covered by these variants was *TTN*; this gene encodes a sarcomeric protein named Titin, with a crucial role in sarcomeric structural integrity and muscle elasticity. Mutations in *TTN* have been shown to cause heart failure in humans. Additionally, mouse models with *TTN* mutations exhibit weak heart contractility and heart failure [47-49] and hearts of mutant embryos displayed weak spontaneous contraction [49]. Additionally, the *TTN* network includes three QT associated genes, *ATP1B, TCEA3*, and *PLN*. *TBX320*, a QRS associated gene; and TBX5 (a QRS and QT associated gene).

We also identified a less frequent C > G missense variant (rs74608430) in the *FCRL2* gene under the linkage peak on chromosome 1p23.1. This variant explains 2.42% (h2 = 0.87%, P = 0.02) of the total genetic variance of QT (h2 = 36%) in the ERF population. *FCRL2* has not been previously described with respect to cardiac function. Bioinformatics resources, however, showed that *FCRL2* expression is associated with *ARHGAP24* and *SETBP1* expression, two genes implicated in ECG variability by GWAS. This suggests that *FCRL2* may be relevant for heart function. *FCRL2* is expressed mostly in liver, heart, testis and kidney8. Gene Network predicts that it may be relevant for cytosolic Ca^{2+} levels and *AMPK* stimulated fatty acid oxidation in muscle. These are plausible pathways for QT function. This finding for rs74608430, however, was not replicated in the RS, in which the MAF was 2.9×10^{-2}. The absence of replication could be related to environmental differences influencing complex gene-environment interactions between these two study groups [50]. Another plausible

explanation is that, due to longer stretches of linkage disequilibrium in the family-based ERF sample, rs74608430 is tagging another variant in ERF and this is not the case in the general population.

Further, Ingenuity analysis revealed that *FCRL2* is correlated with some microRNAs (such as miR-1263, miR337-5p, miR-4699-3p, miR518e-3p, miR-507, miR3689a-5p, miR-507, miR-3622a-5p, miR-450b-5p, miR-4720-3p, and miR-1253). Among these, miR-337-5p is known to be differentially expressed in patients with valvular heart disease and patients with chronic AF [51]. This is consistent with the GEO database at NCBI9, which suggests that *FCRL2* is upregulated in patients with AF and dilated cardiomyopathy. In summary, the bioinformatics data available for this gene supports the hypothesis that *FCRL2* may be involved in heart function, and, specifically, related to ECG variability.

Additional interesting genes have been uncovered under the linkage peaks. First, the PR linkage peak on chromosome 14 contains damaging variants in the alpha and beta subunits of cardiac myosin MYH6 and MYH7. Previous studies showed that genetic variants in these two genes have been found in hypertrophic cardiomyopathy [52-56], dilated cardiomyopathy [56, 57] and atrial septal defect [58]. Second, we found *TNNT2* under the linkage peak on chromosome 1 for QT, which harbors known mutations underlying hypertrophic cardiomyopathy [59] and familial dilated cardiomyopathy [57].

No explanatory variants were found for the other loci, for which there are a number of potential explanations. Linkage peaks are not precise in highlighting the location of the causal variant; even the region of interest cannot be easily pinpointed. Additionally, we did not take into account alternative forms of genetic variation, such as structural and copy number variations (CNVs) or repeats in the linkage regions. Lastly, causal rare variants may be located outside the coding sequence, which we did not include in our sequencing analyses.

Conclusion

Although the combination of linkage and exome sequencing did not lead to the identification of a causal variant, suggestive linkage regions contain a number of plausible candidate genes, including *FCRL2*. *TTN, MYH6, MYH7, TNNT2*, and *HSD321.* Further analysis will need to be performed to demonstrate the involvement of these proteins in ECG measurements. We could not explain these with exonic sequence variants, so they will require more extensive follow-up, but provide potentially important indicators of the location of variation influencing ECG.

Author Contribution

CS: Formal analysis, writing – original draft preparation; IZ: Formal analysis; NA: Formal analysis; AD: Formal analysis; EvL: Formal analysis; JK: Formal analysis, investigation, software; MvB: Formal analysis; BS: Investigation, resources; AU: Investigation, resources; AK: Formal analysis, software; JW: Investigation, resources; RW: Writing – original draft preparation, supervision; BO: Investigation, resources; TA: Formal analysis, supervision; CvD: Conceptualization, formal analysis, investigation, resources, writing – original draft preparation, supervision; AI: Conceptualization, formal analysis, writing – original draft preparation, supervision.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The Erasmus Rucphen Family (ERF) study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP Grant No. 018947 (LSHG-CT-2006- 01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007–2013)/Grant Agreement HEALTH-F4- 2007-201413 by the European Commission under the programme

"Quality of Life and Management of the Living Resources" of 5th Framework Programme (No. QLG2-CT-2002-01254). High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organization for Scientific Research and the Russian Foundation for Basic Research (NWO- RFBR 047.017.043), and Russian Federal Agency of Scientific Organizations projects VI.53.2.2 and 0324-2015-0003. Exome sequencing analysis in ERF was supported by the Netherlands Organization for the Health Research and Development grant for Project No. 91111025. This work was also supported by the Federal Agency of Scientific Organizations #0324-2015-0003 (IZ, AK and TA). The Rotterdam Study is funded by Erasmus Medical Center (MC) and Erasmus University, Rotterdam; Netherlands Organization for the Health Research and Development; the Research Institute for Diseases in the Elderly; the Ministry of Education, Culture and Science; the Ministry for Health, Welfare and Sports; the European Commission; and the Municipality of Rotterdam.

URLs

[1]http://www.scandb.org

[2]http://picard.sourceforge.net

[3]http://varianttools.sourceforge.net/Annotation/DbNSFP

[4]http://snp.gs.washington.edu/SeattleSeqAnnotation/

[5]https://www.ingenuity.com/wp-content/themes/ingenuity-qiagen/pdf/ipa/ipa_datasheet.pdf

[6]http://www.ncbi.nlm.nih.gov/geo/geo2r/

[7]http://genenetwork.nl:8080/GeneNetwork

[8]http://www.bioinfo.mochsl.org.br/miriad/gene/FCRL2/

[9]http://www.ncbi.nlm.nih.gov/geo/

References

1. Lin, W.H., H. Zhang, and Y.T. Zhang, *Investigation on cardiovascular risk prediction using physiological parameters.* Comput Math Methods Med, 2013. **2013**: p. 272691.
2. Pelto, H., D. Owens, and J. Drezner, *Electrocardiographic findings suggestive of cardiomyopathy: what to look for and what to do next.* Curr Sports Med Rep, 2013. **12**(2): p. 77-85.
3. Kolder, I.C., M.W. Tanck, and C.R. Bezzina, *Common genetic variation modulating cardiac ECG parameters and susceptibility to sudden cardiac death.* J Mol Cell Cardiol, 2012. **52**(3): p. 620-9.
4. Newton-Cheh, C. and R. Shah, *Genetic determinants of QT interval variation and sudden cardiac death.* Curr Opin Genet Dev, 2007. **17**(3): p. 213-21.
5. Cheng, M., et al., *Electrocardiographic PR prolongation and atrial fibrillation risk: a meta-analysis of prospective cohort studies.* J Cardiovasc Electrophysiol, 2015. **26**(1): p. 36-41.
6. Mozos, I. and A. Caraba, *Electrocardiographic Predictors of Cardiovascular Mortality.* Dis Markers, 2015. **2015**: p. 727401.
7. Arking, D.E., et al., *A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization.* Nat Genet, 2006. **38**(6): p. 644-51.
8. Arking, D.E., et al., *Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization.* Nat Genet, 2014. **46**(8): p. 826-36.
9. Newton-Cheh, C., et al., *Genome-wide association study of electrocardiographic and heart rate variability traits: the Framingham Heart Study.* BMC Med Genet, 2007. **8 Suppl 1**: p. S7.
10. Newton-Cheh, C., et al., *Common variants at ten loci influence QT interval duration in the QTGEN Study.* Nat Genet, 2009. **41**(4): p. 399-406.
11. Pfeufer, A., et al., *Common variants at ten loci modulate the QT interval duration in the QTSCD Study.* Nat Genet, 2009. **41**(4): p. 407-14.
12. Pfeufer, A., et al., *Genome-wide association study of PR interval.* Nat Genet, 2010. **42**(2): p. 153-9.
13. Holm, H., et al., *Several common variants modulate heart rate, PR interval and QRS duration.* Nat Genet, 2010. **42**(2): p. 117-22.
14. Sotoodehnia, N., et al., *Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction.* Nat Genet, 2010. **42**(12): p. 1068-76.
15. Tristani-Firouzi, M., et al., *Molecular biology of K(+) channels and their role in cardiac arrhythmias.* Am J Med, 2001. **110**(1): p. 50-9.
16. Hedley, P.L., et al., *The KCNE genes in hypertrophic cardiomyopathy: a candidate gene study.* J Negat Results Biomed, 2011. **10**: p. 12.
17. Silva, C.T., et al., *Heritabilities, proportions of heritabilities explained by GWAS findings, and implications of cross-phenotype effects on PR interval.* Hum Genet, 2015. **134**(11-12): p. 1211-9.
18. Pardo, L.M., et al., *The effect of genetic drift in a young genetically isolated population.* Ann Hum Genet, 2005. **69**(Pt 3): p. 288-95.
19. Aulchenko, Y.S., et al., *Linkage disequilibrium in young genetically isolated Dutch population.* Eur J Hum Genet, 2004. **12**(7): p. 527-34.
20. Sayed-Tabatabaei, F.A., et al., *Heritability of the function and structure of the arterial wall: findings of the Erasmus Rucphen Family (ERF) study.* Stroke, 2005. **36**(11): p. 2351-6.
21. *1999 World Health Organization-International Society of Hypertension Guidelines for the Management of Hypertension. Guidelines Subcommittee.* J Hypertens, 1999. **17**(2): p. 151-83.
22. van Bemmel, J.H., J.A. Kors, and G. van Herpen, *Methodology of the modular ECG analysis system MEANS.* Methods Inf Med, 1990. **29**(4): p. 346-53.
23. Willems, J.L., et al., *A reference data base for multilead electrocardiographic computer measurement programs.* J Am Coll Cardiol, 1987. **10**(6): p. 1313-21.
24. Willems, J.L., et al., *The diagnostic performance of computer programs for the interpretation of electrocardiograms.* N Engl J Med, 1991. **325**(25): p. 1767-73.

25. de Bruyne, M.C., et al., *Diagnostic interpretation of electrocardiograms in population-based research: computer program research physicians, or cardiologists?* J Clin Epidemiol, 1997. **50**(8): p. 947-52.
26. Eijgelsheim, M., et al., *Identification of a common variant at the NOS1AP locus strongly associated to QT-interval duration.* Hum Mol Genet, 2009. **18**(2): p. 347-57.
27. Leening, M.J., et al., *Unrecognised myocardial infarction and long-term risk of heart failure in the elderly: the Rotterdam Study.* Heart, 2010. **96**(18): p. 1458-62.
28. Funck-Brentano, C. and P. Jaillon, *Rate-corrected QT interval: techniques and limitations.* Am J Cardiol, 1993. **72**(6): p. 17b-22b.
29. O'Connell, J.R. and D.E. Weeks, *PedCheck: a program for identification of genotype incompatibilities in linkage analysis.* Am J Hum Genet, 1998. **63**(1): p. 259-66.
30. Abecasis, G.R., et al., *Merlin--rapid analysis of dense genetic maps using sparse gene flow trees.* Nat Genet, 2002. **30**(1): p. 97-101.
31. Gudbjartsson, D.F., et al., *Allegro, a new computer program for multipoint linkage analysis.* Nat Genet, 2000. **25**(1): p. 12-3.
32. Kirichenko, A.V., et al., *PedStr software for cutting large pedigrees for haplotyping, IBD computation and multipoint linkage analysis.* Ann Hum Genet, 2009. **73**(Pt 5): p. 527-31.
33. Lander, E. and L. Kruglyak, *Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results.* Nat Genet, 1995. **11**(3): p. 241-7.
34. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.
35. Brouwer, R.W., et al., *NARWHAL, a primary analysis pipeline for NGS data.* Bioinformatics, 2012. **28**(2): p. 284-5.
36. Hofman, A., et al., *The Rotterdam Study: 2014 objectives and design update.* Eur J Epidemiol, 2013. **28**(11): p. 889-926.
37. Chun, S. and J.C. Fay, *Identification of deleterious mutations within three human genomes.* Genome Res, 2009. **19**(9): p. 1553-61.
38. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
39. Schwarz, J.M., et al., *MutationTaster evaluates disease-causing potential of sequence alterations.* Nat Methods, 2010. **7**(8): p. 575-6.
40. Vaser, R., et al., *SIFT missense predictions for genomes.* Nat Protoc, 2016. **11**(1): p. 1-9.
41. Barth, A.S., et al., *Reprogramming of the human atrial transcriptome in permanent atrial fibrillation: expression of a ventricular-like genomic signature.* Circ Res, 2005. **96**(9): p. 1022-9.
42. Yeh, Y.H., et al., *Region-specific gene expression profiles in the left atria of patients with valvular atrial fibrillation.* Heart Rhythm, 2013. **10**(3): p. 383-91.
43. Ott, J., Y. Kamatani, and M. Lathrop, *Family-based designs for genome-wide association studies.* Nat Rev Genet, 2011. **12**(7): p. 465-74.
44. Mango, R., et al., *Next Generation Sequencing and Linkage Analysis for the Molecular Diagnosis of a Novel Overlapping Syndrome Characterized by Hypertrophic Cardiomyopathy and Typical Electrical Instability of Brugada Syndrome.* Circ J, 2016. **80**(4): p. 938-49.
45. Reich, D., et al., *A whole-genome admixture scan finds a candidate locus for multiple sclerosis susceptibility.* Nat Genet, 2005. **37**(10): p. 1113-8.
46. Shimodaira, M., et al., *Association of HSD3B1 and HSD3B2 gene polymorphisms with essential hypertension, aldosterone level, and left ventricular structure.* Eur J Endocrinol, 2010. **163**(4): p. 671-80.
47. Gerull, B., et al., *Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy.* Nat Genet, 2002. **30**(2): p. 201-4.
48. Xu, X., et al., *Cardiomyopathy in zebrafish due to mutation in an alternatively spliced exon of titin.* Nat Genet, 2002. **30**(2): p. 205-9.
49. May, S.R., et al., *A Titin mutation defines roles for circulation in endothelial morphogenesis.* Dev Biol, 2004. **270**(1): p. 31-46.

50. Teo, Y.Y., et al., *Genome-wide comparisons of variation in linkage disequilibrium.* Genome Res, 2009. **19**(10): p. 1849-60.
51. Cooley, N., et al., *Influence of atrial fibrillation on microRNA expression profiles in left and right atria from patients with valvular heart disease.* Physiol Genomics, 2012. **44**(3): p. 211-9.
52. Anan, R., et al., *Prognostic implications of novel beta cardiac myosin heavy chain gene mutations that cause familial hypertrophic cardiomyopathy.* J Clin Invest, 1994. **93**(1): p. 280-5.
53. Lankford, E.B., et al., *Abnormal contractile properties of muscle fibers expressing beta-myosin heavy chain gene mutations in patients with hypertrophic cardiomyopathy.* J Clin Invest, 1995. **95**(3): p. 1409-14.
54. Rayment, I., et al., *Structural interpretation of the mutations in the beta-cardiac myosin that have been implicated in familial hypertrophic cardiomyopathy.* Proc Natl Acad Sci U S A, 1995. **92**(9): p. 3864-8.
55. Niimura, H., et al., *Sarcomere protein gene mutations in hypertrophic cardiomyopathy of the elderly.* Circulation, 2002. **105**(4): p. 446-51.
56. Carniel, E., et al., *Alpha-myosin heavy chain: a sarcomeric gene associated with dilated and hypertrophic phenotypes of cardiomyopathy.* Circulation, 2005. **112**(1): p. 54-9.
57. Kamisago, M., et al., *Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy.* N Engl J Med, 2000. **343**(23): p. 1688-96.
58. Ching, Y.H., et al., *Mutation in myosin heavy chain 6 causes atrial septal defect.* Nat Genet, 2005. **37**(4): p. 423-8.
59. Thierfelder, L., et al., *Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere.* Cell, 1994. **77**(5): p. 701-12.

Chapter 4


ATTCGGTACCTCATAACGGCTAGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTT
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGGTACCTA

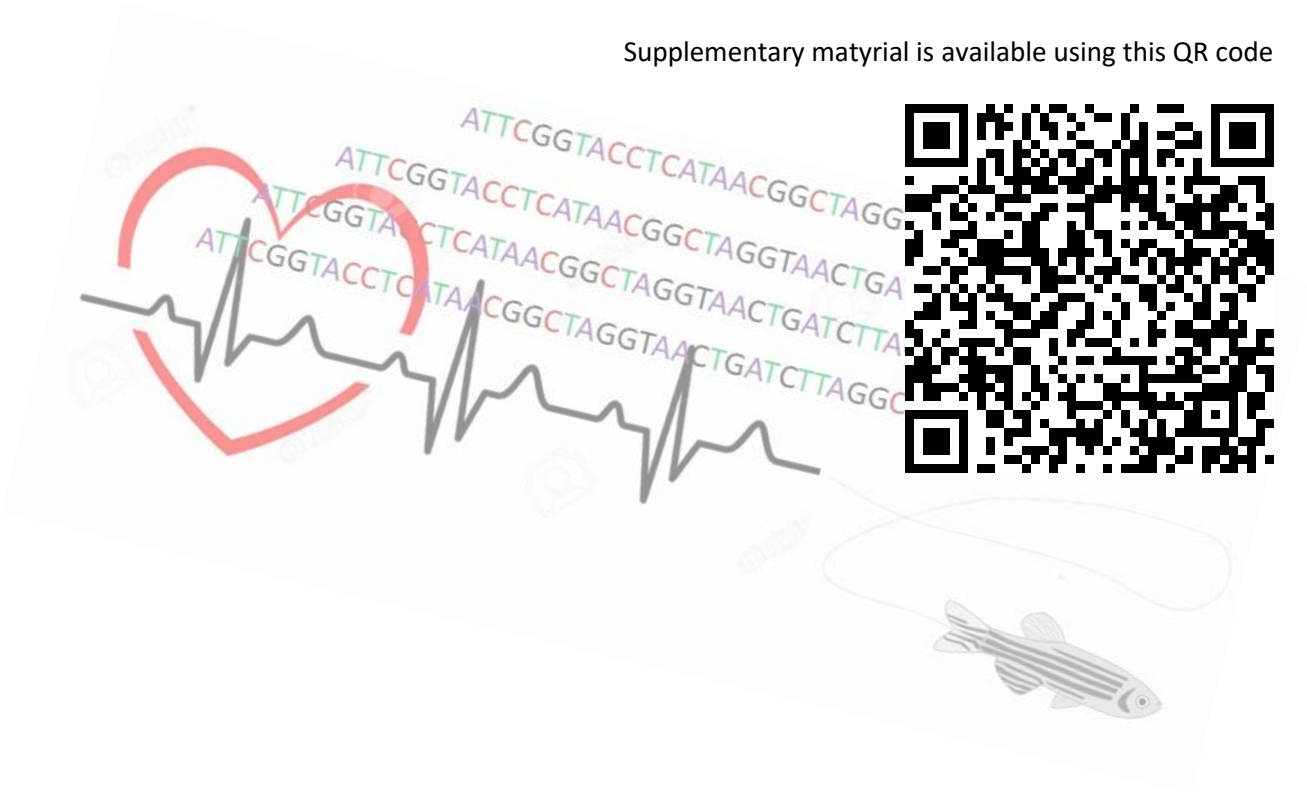
Chapter 4

A combined linkage, microarray and exome analysis suggests MAP3K11 as a candidate gene for left ventricular hypertrophy

Claudia Tamar Silva, Irina V. Zorkoltseva, Maartje N. Niemeijer, Marten E. van den Berg, Najaf Amin, Ayşe Demirkan, Elisa van Leeuwen, Adriana I. Iglesias, Laura B. Piñeros-Hernández, Carlos M. Restrepo, Jan A. Kors, Anatoly V. Kirichenko, Rob Willemsen, Ben A. Oostra, Bruno H. Stricker, André G. Uitterlinden, Tatiana I. Axenovich, Cornelia M. van Duijn and Aaron Isaacs.

BMC Medical GenomicsBMC series – 201811:22. https://doi.org/10.1186/s12920-018-0339-9

Supplementary matyrial is available using this QR code



Abstract

Background. Electrocardiographic measures of left ventricular hypertrophy (LVH) are used as predictors of cardiovascular risk. We combined linkage and association analyses to discover novel rare genetic variants involved in three such measures and two principal components derived from them.

Methods. The study was conducted among participants from the Erasmus Rucphen Family Study (ERF), a Dutch family-based sample from the southwestern Netherlands. Variance components linkage analyses were performed using Merlin. Regions of interest (LOD > 1.9) were fine-mapped using microarray and exome sequence data.

Results. We observed one significant LOD score for the second principal component on chromosome 15 (LOD score = 3.01) and 12 suggestive LOD scores. Several loci contained variants identified in GWAS for these traits; however, these did not explain the linkage peaks, nor did other common variants. Exome sequence data identified two associated variants after multiple testing corrections were applied.

Conclusions. We did not find common SNPs explaining these linkage signals. Exome sequencing uncovered a relatively rare variant in MAPK3K11 on chromosome 11 (MAF = 0.01) that helped account for the suggestive linkage peak observed for the first principal component. Conditional analysis revealed a drop in LOD from 2.01 to 0.88 for MAP3K11, suggesting that this variant may partially explain the linkage signal at this chromosomal location. MAP3K11 is related to the JNK pathway and is a pro-apoptotic kinase that plays an important role in the induction of cardiomyocyte apoptosis in various pathologies, including LVH.

Background

Left ventricular hypertrophy (LVH) is a predictor of increased cardiovascular morbidity and mortality [1]. Those with LVH have a 2-fold increased risk of adverse events, particularly ischemic heart disease and chronic heart failure [2, 3]. Increased left ventricular mass maintains cardiac pump performance in response to cardiovascular insults, such as coronary heart disease [3, 4]. Risk factors for LVH are elevated systolic blood pressure, obesity, hypertension, insulin resistance, valvular heart disease and advanced age, among others [2, 5, 6]. LVH proxy measurements can be assessed through noninvasive methods, such as echocardiography and magnetic resonance imaging, however, electrocardiographic measurements are the most used worldwide [7]. LVH proxy measurements include calculations of the Sokolow-Lyon index (SL), the Cornell voltage product (CV) and the 12-lead sum QRS product (12LS). Several studies have demonstrated that genetic factors influence electrocardiographic and echocardiographic measures of LVH [2, 4, 5, 8, 9]. We recently demonstrated that these measures contain a substantial heritable component (SL = 0.46, 12LS = 0.49 and CV = 0.34) [10].

Genome-wide linkage analyses, candidate gene association studies, genome-wide association studies (GWAS) and gene mapping have been conducted to identify genes influencing LVH. In the first GWAS of these traits, two loci, PTGES3 and NMB, reached genome-wide significance. IGF1R and SCN5A were identified and replicated without reaching genome-wide significance [5]. Recently, an expanded GWAS detected a number of novel loci influencing CV, SL, and 12LS [11]. Among these were 32 loci containing genes with known cardiac function, coding for cardiac sarcomere components or related to cardiac myocyte function. Evidence for linkage of echocardiographic LV mass to chromosome 5 (LOD score = 1.6) and electrocardiographic LV mass to chromosome 7 (LOD score = 1.67) [8] and chromosome 12 (LOD score = 2.19 and 3.11) [8, 12]

were reported in linkage studies, with the strongest evidence for chromosome 12 [3]. As is the case for other complex outcomes, most candidate genes studies have not been replicated and do not reach genome-wide significance [3].

Exome sequencing has been successfully used for Mendelian disorders [13]. More recently, this technology has been extended to the analysis of non-Mendelian diseases and complex traits, as rare variants with large effects can contribute to the heritability of common traits. The aim of this study was to discover rare variants by linkage analysis in a large family-based study, the Erasmus Rucphen Family (ERF) study. Linked regions were fine-mapped in detail using microarray data and exome sequencing.

Methods

Study population

The ERF study is a family-based study including over 3000 participants descendant from 22 couples that lived in the Rucphen region in the southwest Netherlands in the nineteenth century [14]. All descendants of those couples were invited to visit the clinical research center in the region where they were examined in person [15]. Interviews at the time of blood sampling were performed by medical practitioners and included questions on current medication use and medical history [16]. Additionally, participants were asked to bring their current medications with them to the study center; these were cross-referenced with general practitioner and pharmacy records. Height and weight were measured with the participant in light underclothing and body mass index (kg/m2) was computed. Blood pressure was measured twice on the right arm in a sitting position after at least 5 min rest, using an automated device (OMRON 711, Omron Healthcare, Bannockburn, IL, USA). The average of the two measures was used for analysis. Hypertension status was identified

through the use of antihypertensive medication and/or through the assessment of blood pressure measurements according to the guidelines of the World Health Organization [17]. The Medical Ethics Committee of the Erasmus University Medical Center approved the ERF study protocol and all participants, or their legal representatives, provided written informed consent.

ECG interpretation and measurements

Examinations included 12-lead ECG measurements. A 10 s 12-lead ECG (on average, 8 to 10 beats) was recorded with an ACTA-ECG electrocardiograph (Esaote, Florence, Italy) with a sampling frequency of 500 Hz. Digital measurements of the ECG parameters were made using the Modular ECG Analysis System (MEANS) [18, 19]. Briefly, MEANS operates on multiple simultaneously recorded leads, which are transformed to a detection function that brings out the QRS complex and the other parts of the signal. MEANS determines common onsets and offsets for all 12 leads together on one representative averaged beat, with the use of template matching techniques. The measurement and diagnostic performance of MEANS has been extensively evaluated, both by the developers and by others [19-22]. The MEANS criteria for MI are mainly based on pathological Q waves, QR ratio, and R-wave progression [20]. A cardiologist, specialized in ECG methodology, ascertained the final diagnosis of MI.

MEANS was used to measure QRS complex duration and the three LVH proxies. Sokolow-Lyon was defined as the sum of the S wave in V1 plus the R wave in V5 or V6, Cornell as the sum of R in aVL and S in V3, and 12-lead as the sum of R to S in all 12 leads; these three voltages were then multiplied by QRS duration to obtain voltage-duration products as an approximation of the area under the QRS complex [21-23]. Principal component (PC) analysis was applied to the three original measurements (SL, 12LS and CV) to capture the correlation structure between traits. Two

PCs, PC1 and PC2, captured more than 94% of the total variance and were also assessed as phenotypes in these analyses. All traits were adjusted for sex, age, BMI and height and the residuals were rank transformed prior to analysis.

Genotyping and statistical analysis of the linkage study

Illumina's HumanHap6k Genotyping BeadChip (6 K Illumina Linkage IV Panels®) was used for genotyping for the linkage analyses. All genotyping procedures were performed according to the manufacturer's protocols. Only markers with a minor allele frequency (MAF) > 0.05 were selected for further analysis. Genotyping errors leading to Mendelian inconsistencies were detected using PedCheck [24]. Unlikely double recombination events were detected using MERLIN [25]. All detected errors were eliminated from the data. A total of 5250 autosomal SNPs with a call rate greater than 95% were utilized for the linkage analyses. Among the 2385 individuals who were phenotyped for LVH measures, 1860 people also had genotype data and were included in the linkage study. Variance component multipoint linkage was performed using the --vc option in MERLIN v.1.0.1 [25, 26]. This program calculates exact IBD sharing probabilities using the Lander-Green algorithm, requiring restrictions on pedigree size. Because of this, the single ERF pedigree with multiple loops was split into non-overlapping fragments of no more than 18 bits with the help of the PedSTR program [27].

Regions of interest with LOD > 1.9 were selected for further analysis. Borders of the linkage areas were defined as LOD score minus 2 support intervals (LOD-2 SI) around the linkage peaks. Genes within the LOD-2 SI were annotated using SCAN (SNP and CNV Annotation Database).

Genotyping and statistical analysis of the association study

Of 2385 phenotyped people, dense genotypes were available for 2128 subjects, typed on 3 different genotyping platforms (Illumina 318 K, Illumina 370 K and Affymetrix 250 K), which were merged first (median number of quality controlled SNPs per individual = 325,500) and then ~ 2.54 million SNPs were imputed using MACH (v1.0.16) [28, 29], with the HapMap build 36 (release 22) CEU population as reference. Within each genotyping batch, only SNPs with a call rate > 98%, MAF > 1% and Hardy-Weinberg Equilibrium P-value > 10− 6 were used for imputations. To account for relatedness, a genomic kinship matrix was computed in GenABEL [30]. This matrix was then incorporated into linear mixed-effects regression models, as implemented in ProbABEL [31], which were used to assess the association of variants in the LOD-2 SI with the LVH phenotypes. P-values were adjusted with the FDR-based q-value technique [32].

Exome sequencing

The exomes of 1336 individual from the ERF population were sequenced “in-house” at the Center for Biomics of the Department of Cell Biology of the Erasmus MC, the Netherlands, using the Agilent version V4 capture kit on an Illumina HiSeq 2000 sequencer using the TruSeq Version 3 protocol. Mean depth base was 74.23× (median = 57×) and mean depth region was 65.26× (median = 52.87×). The sequence reads were aligned to the human genome build 19 (hg19) using BWA and the NARWHAL pipeline [33, 34]. The aligned reads were processed further using the IndelRealigner, MarkDuplicates and TableRecalibration tools from the Genome Analysis Toolkit (GATK) and Picard (http://broadinstitute.github.io/picard/) to remove systematic biases and to recalibrate the PHRED quality scores in the alignments. Genetic variants were called using the Unified Genotyper tool of the GATK. About 1.4 million Single Nucleotide Variants (SNVs) were

called and, after removing the low quality variants (QUAL < 150), we retrieved 577,703 SNVs in 1309 individuals. ECG and covariate data were available for 1072 of these samples. Further, for comparison and to predict the functionality of the variants, annotations were also performed using the dbNSFP (database of human non-synonymous SNPs and their functional predictions, http://varianttools.sourceforge.net/Annotation/DbNSFP) and Seattle (http://snp.gs.washington.edu/SeattleSeqAnnotation138/) databases. These databases gave functional prediction results from four different programs, PolyPhen-2, SIFT, MutationTaster and LRT, apart from gene and variant annotations.

We employed a Bonferroni correction for the number of deleterious mutations selected for each trait to correct for multiple comparisons in the exome data: 101 for SL (P-value = $4.9 \times 10-4$), 98 for CV (P-value = $5.1 \times 10-4$) and 60 for 12 LS (P-value = $8.3 \times 10-4$). For the PCs, the numbers were 141 for PC1 (P-value = $3.5 \times 10-4$) and 71 for PC2 (P-value = $7.0 \times 10-4$).

Replication

Four SNPs (rs139580877, rs138968470, rs35996030 and rs142551296) were selected for replication in the Rotterdam Study (RS). The Rotterdam Study is a prospective cohort study ongoing since 1990 in the city of Rotterdam in the Netherlands [35].

Exomes from 1764 individuals from the RS population were sequenced at an average depth of 20× using the Nimblegen SeqCap EZ V2 capture kit on an Illumina HiSeq 2000 sequencer and the TrueSeq Version 3 protocol. The sequence reads were aligned to hg19 using BWA. Subsequently, the aligned reads were processed further using Picard, SAMtools and GATK. Genetic variants were called using the Unified Genotyper Tool from GATK. Samples with low concordance

to genotyping array (< 95%), low transition/transversion ratio (< 2.3), high heterozygote to homozygote ratio (> 2.0) and low call rate (< 80%) were removed from the data. SNVs with a low call rate (< 90%) and out of HWE (P-value < 10− 6) were also removed from the data. The final dataset consisted of 635,814 SNVs in 1450 individuals with complete phenotype and covariate data.

One SNP, rs139580877, was not available in the Rotterdam Study exome data. This variant was imputed using the GIANT 1000 Genomes Phase I Version 3 All reference panel, as previously described [36]. In brief, after filtering SNPs genotyped with the Illumina v3 Infinium II HumanHap550 microarray for deviations from Hardy-Weinberg proportions ($P < 1 \times 10-6$), call rate (< 98%), MAF (< 0.01), and Mendelian errors (> 100), MACH was used to perform the imputations.

Results

Table 1 shows characteristics of the participants in the LVH linkage, microarray, and exome sequence analyses. The proportion of LVH cases for each proxy measure was determined using published cut-off values [37, 38]. There were no significant differences between these overlapping groups. Table 2 shows the correlation between the traits (r = 0.76 in the adjusted model for SL and 12LS, 0.17 between SL and CV, and 0.48 for CV and 12LS). Table 3 shows the loadings of the three LVH proxies (SL, CV, 12LS) to the two PCs that were constructed. PC1 predominantly captured SL and 12LS, while PC2 correlated strongly with CV and moderately with SL. Table 4 shows the linkage results for the LVH proxy measures, which yielded a total of seven regions with suggestive LOD scores (LOD > 1.9). SL was linked to three regions, with the highest LOD score for chromosome 20 (LOD = 2.64) and two additional regions on chromosomes 4 (LOD = 2.14) and 15 (LOD = 1.92).

Suggestive LOD scores for CV were seen on chromosomes 1 (LOD = 2.4) and 6 (LOD = 2.17). There was suggestive linkage of 12LS to chromosomes 5 (LOD = 2.18) and 20 (LOD = 2.12). Linkage results for the principal component analysis of the LVH measures showed one significant LOD score for PC2 on 15q11.2 (LOD = 3.01). This region was also linked to SL (LOD = 1.92). Two regions were suggestively linked to PC1: 11q13.4 (LOD = 2.01) and 20p12.1 (LOD = 2.83), which was also linked to SL and 12LS. For PC2, there were three suggestive linkage results, for chromosomes 6 (LOD = 2.09), 9 (LOD = 2.35) and 22 (LOD = 1.99). The chromosome 6 region was also linked to CV. Plots showing the linked regions by chromosome are provided in Fig. 1. Table 5 shows the top common variant microarray-based association signals under the LVH trait linkage peaks, including P-values and MAF for each SNP. None achieved statistical significance after correction for multiple comparisons.

	LVH GWAS			Linkage Studies			Exon sequence		
	n = 2128			n = 1860			n = 1309		
	Mean (S.D.)	Minimum	Maximum	Mean (S.D.)	Minimum	Maximum	Mean (S.D.)	Minimum	Maximum
Males	899 (42%)			775 (42%)			408 (38%)		
Age (y)	47.0 (13.82)	16.6	85.3	46.5 (13.79)	16.6	85.3	46.51 (13.7)	18.7	81.0
BMI (kg/m^2)	26.7 (4.57)	15.5	61.8	26.7 (4.58)	15.5	61.8	26.4 (4.3)	15.5	61.8
Height (cm)	167.6 (9.31)	139.3	196.5	167.4 (9.19)	143.6	196.5	166.7 (9.0)	143.6	196.5
Weight (kg)	75.1 (15.16)	41.9	161.0	74.9 (15.5)	41.9	161	73.6 (14.3)	42.1	161.0
SBP (mm Hg)	138.4 (19.5)	85.5	222.0	137.7 (19.1)	85.5	217.0	137.0 (18.7)	85.5	216.0
DBP (mm Hg)	79.9 (9.8)	53.5	124.0	79.7 (9.7)	54.5	120.0	79.1 (9.6)	53.5	120.0
Hypertension	913 (43%)			766 (42%)			549 (51%)		
SL	2344 (690.6)	884.0	5288.0	2341 (690.6)	884	52.9	2319 (659.0)	967	5288.0
CV	1173.5 (505.1)	93.1	4126.1	1170.0 (497.3)	93.1	3952.8	1151.6 (659.0)	155.8	3853.0
12LS	13,862 (3812.3)	4993	39250	13,805 (3767.8)	49.9	39.2	13,610.0 (3628.7)	5485.0	36,364
LVH (SL)	138 (6.5%)			120 (6.4%)			66 (6.2%)		
LVH (CV)	41 (1.9%)			32 (1.7%)			20 (1.9%)		
LVH (12LS)	176 (8.3%)			147 (7.9%)			76 (7.1%)		

Table 1. Descriptive statistics of the Erasmus Rucphen Family (ERF) study population
Values presented are mean (standard devition) or n (%)
BMI: Body Mass Index, SBP: Systolic blood pressure, DBP: Diastolic blood pressure, SL: Sokolow-Lyon index, CV: Cornell product, 12LS: 12-lead sum product.

	Unadjusted	**Adjusted**
SL – 12LS	0.80	0.76
SL – CV	0.29	0.17
CV – 12LS	0.56	0.48

Table 2. Pearson's correlations between LVH proxy measures
SL: Sokolow-Lyon; CV: Cornell Voltage product; 12LS: twelve-lead sum product; PC1: first principal component; PC2: second principal component.
Adjusted model included age, sex, body-mass index, and height.

	Principal Component	
	PC1	PC2
SL	0.**84**	-0.48
CV	0.61	0.**78**
12LS	**0.95**	-0.08

Table 3. PC Loadings for LVH Proxies
SL: Sokolow-Lyon; CV: Cornell Voltage product; 12LS: twelve-lead sum product; PC1: first principal component; PC2: second principal component.
Adjusted model included age, sex, body-mass index, and height.

Trait	N	Chr.	SNP	Position (cM)	LOD_{MAX}
SL	1860	4	rs1032328	144.46	2.14
SL	1860	15	rs290370	112.3	1.92
SL	1860	20	rs204115	38.11	2.64
CV	1860	1	rs6619	59.63	2.4
CV	1860	6	rs2040431	108.31	2.17
12LS	1860	5	rs1442470	42.3	2.18
12LS	1860	20	rs466243	40.7	2.12
PC1	1860	11	rs1530354	65.21	2.01
PC1	1860	20	rs2077147	45.09	2.83
PC2	1860	6	rs1391503	99.69	2.09
PC2	1860	9	rs748530	40.22	2.35
PC2	1860	15	rs1562203	0	3.01
PC2	1860	22	rs138383	46.89	1.99

Table 4. Results of linkage analyses
PC1: first principal component; PC2: second principal component; N: sample size; Chr.: chromosome; LOD_{MAX}: LOD score at SNP.
Model adjusted for age, sex, body-mass index, and height.

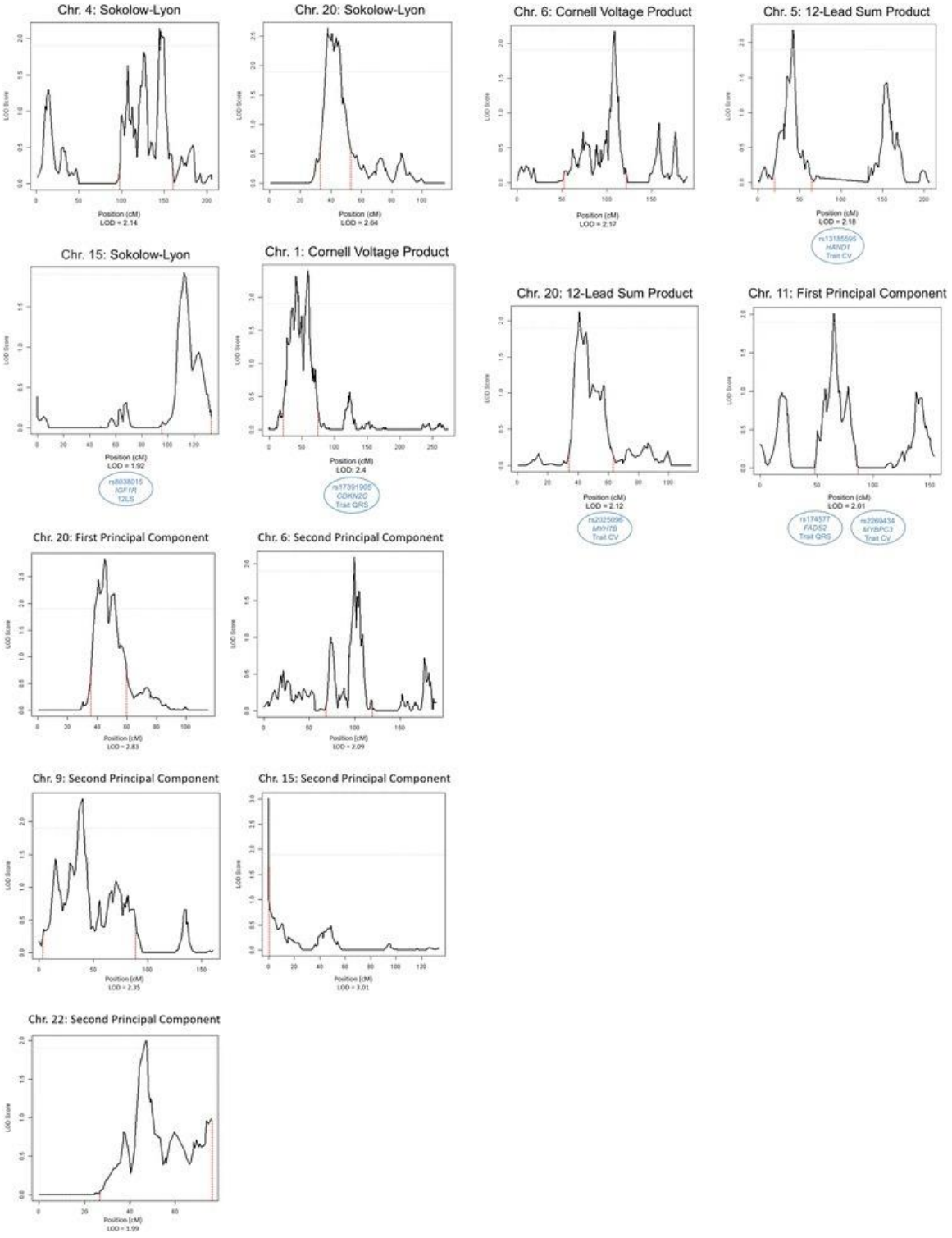


Figure 1. Linkage peaks for the LVH proxy measures. Plots depicting the linked regions by trait and chromosome. The grey dashed horizontal line indicates the threshold for suggestive linkage. The red dashed vertical lines show the borders of the LOD score minus 2 support intervals (LOD-2 SI). The blue circles contain SNPs identified in previous GWAS for these traits in the LOD-2 S

Outcome	Region	SNP	MAF	Gene	P-value	Q-value
SL	4q26	rs6839953	0.27	*TRAM1L1*	$1.34x10^{-4}$	0.47
SL	15q26.2	rs11074275	0.48	*MCTP2*	$4.27x10^{-4}$	0.79
SL	20p12.1	rs721243	0.19	*ISM1*	$7.37x10^{-5}$	0.15
CV	1p35.1	rs16835131	0.06	*SYNC*	$1.35x10^{-5}$	0.35
CV	6q15	rs10944412	0.27	*RNGTT*	$4.60x10^{-5}$	0.93
12LS	5p15.2	rs2589661	0.10	*ROPN1L*	$1.26x10^{-4}$	0.46
12LS	20p11.23	rs6106235	0.18	*C20orf26*	$1.69x10^{-5}$	0.09
PC1	11q12.2	rs1790325	0.04	*FADS1*	$2.85x10^{-5}$	0.08
PC1	20p12.1	rs13036282	0.005	*SPTLC3*	$2.30x10^{-4}$	0.63
PC2	6q16.3	rs1475922	0.06	*GRIK2*	$1.64x10^{-4}$	0.94
PC2	9p24.1	rs10975939	0.003	*KDM4C*	$4.67x10^{-4}$	1.00
PC2	15q11.2	rs8043191	0.03	*CYFIP1*	$5.95x10^{-3}$	0.52
PC2	22q13.33	rs2688089	0.45	*C22orf34*	$7.02x10^{-5}$	0.56

Table 5. Top association signal under LVH trait linkage peaks
SL: Sokolow-Lyon; CV: Cornell Voltage product; 12LS: twelve-lead sum product; PC1: first principal component; PC2: second principal component; MAF: minor allele frequency.

Variants in the coding sequence

The results of the search for less frequent exonic variants are summarized in Additional file 1: Table S1. We focused on relatively rare (frequency < 5%) missense variants predicted to be deleterious by at least two of the prediction algorithms used and non-sense variants. This selection yielded 471 variants in 356 genes in the 13 linkage intervals (LOD-2 SI), which we analysed with respect to the LVH proxy measures and PCs. Additional file 1: Table S2 shows the results with a nominal P-value ≤0.05 after regressing out the effects of age, BMI, height and sex. This effort uncovered an A > G variation (rs139580877) in the SPEF2 gene on 5p13.2, which was significantly associated with 12LS when adjusted for multiple testing (P-value = 4.2 × 10− 4). This variant, with 108 carriers in ERF, is predicted to be probably damaging by PolyPhen-2 with a score of 0.972 and as deleterious by SIFT with a score ranging between 0.02 and 0.03. It is a missense variant, among more than 2000 described for this gene. In the principal components analysis, rs138968470, on 11q13.1 in the MAP3K11 gene, was associated with PC1 adjusted for multiple testing (P-value = 3.5 × 10− 4). SKAT-O and burden tests provided some supporting evidence for

the association of this gene with LVH proxy measures (Additional file 1: Table S3). Additionally, at the SL chromosome 4 locus, we identified a C > G variation (rs142551296) in PRSS12 that approached significance (P-value = 8.4 × 10− 4). A second, more common intragenic variant inside PRSS12 was nominally associated (rs35996030; P-value = 0.04). We re-ran the linkage analyses conditioning on these variants to see if they explained the observed linkage signals. For PC1, the LOD score in the 11q13.4 linkage region dropped in the conditional analysis (from 2.01 to 0.88), suggesting that the associated variant (rs138968470), or neighbouring variants in linkage disequilibrium (LD), explained the linkage signal. This variant also showed evidence of association with the two traits (12LS and SL) underlying PC1 (P-value = 3.0 × 10− 4 and P-value = 1.2 × 10− 3, respectively). Using Gene Network (http://genenetwork.nl/gene/ENSG00000173327), to perform in-depth analyses of the expression of MAP3K11, demonstrated that its expression is strongly linked to rho signalling (ARGHGEF15, ARHGDIA) (Fig. 2).

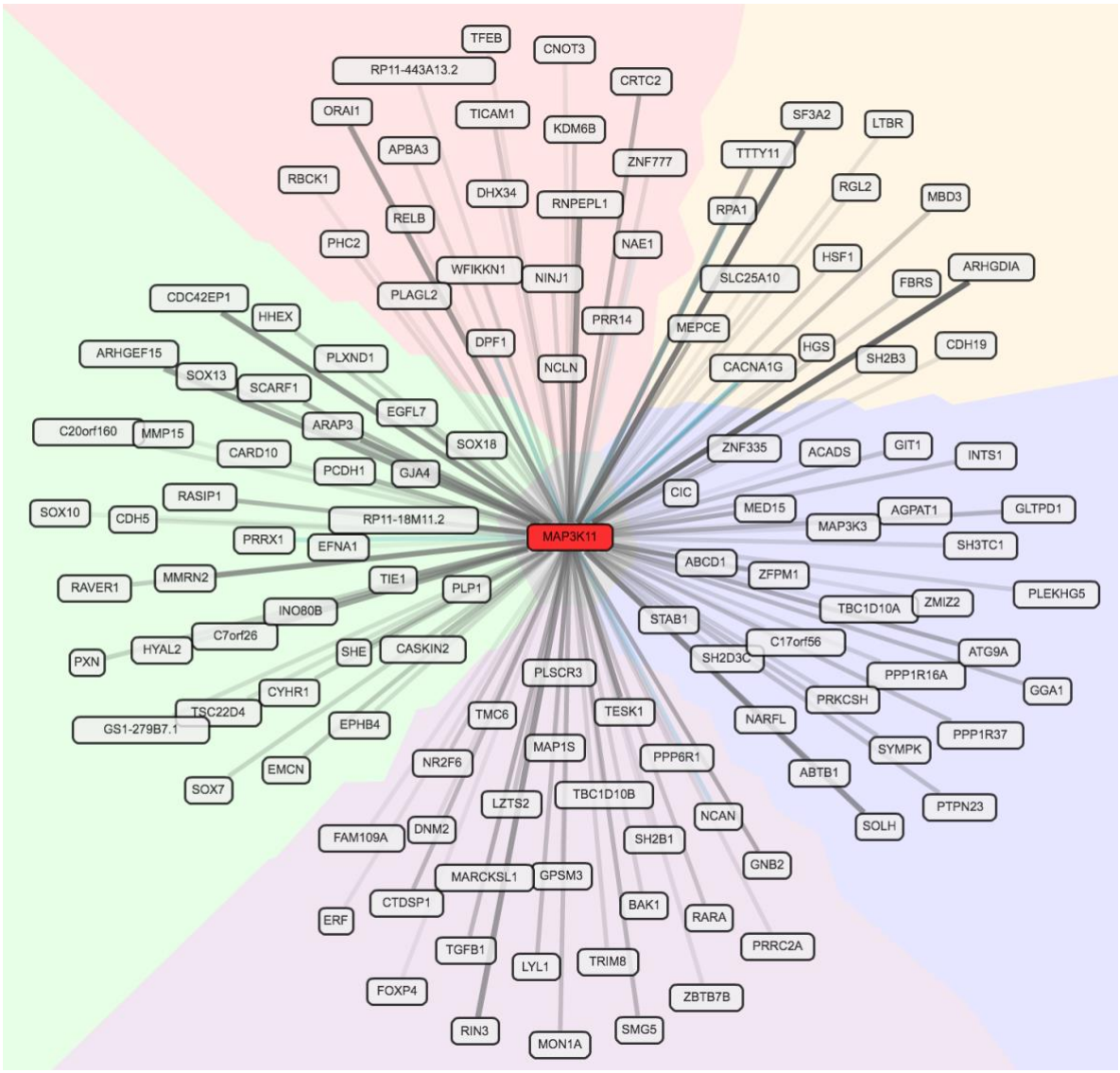


Figure 2. MAP3K11 gene network interactions.
Looking for interactions for MAP3K11, we searched Gene Network (http://genenetwork.nl/gene/ENSG00000173327). One hundred twenty-nine gene-gene interactions are shown

Five of the linkage peaks contained loci recently identified in GWAS studies [5, 11]. To determine if the linkage signals were a result of those common variants, linkage was performed a second time, conditioned on the GWAS index SNPs. These analyses demonstrated that the observed peaks were not explained by the GWAS SNPs, although the estimates fluctuated somewhat, likely as a result of smaller sample sizes (Additional file 1: Table S4).

Replication

Summary statistics for the Rotterdam Study sample are provided in Additional file 1: Table S5. The variant rs139580877 was imputed, using the 1000 Genomes reference panel; the imputation quality score (MACH RSQ) for this variant was 0.65, with a minor allele frequency of 0.008. The effect estimate for 12LS was essentially zero, and therefore, did not replicate the ERF findings (Additional file 1: Table S6). The other variants of interest, rs35996030, rs138968470 and rs142551296, were directly genotyped in a subset of the Rotterdam Study cohort (n = 1450). There was no evidence of association for any of these variants in the Rotterdam Study.

Discussion

We performed a linkage study on LVH proxy measurements, and PCs, and identified one significant locus (15q11.2) and 10 suggestive regions (1p34, 4q31, 5p14, 6q15, 6q21, 9p21, 11q13.4, 15q25, 20p12, 22q13). Exome variant analysis in these regions uncovered a missense coding variation in MAP3K11 on 11q134 for PC1; the MAP3K11 variant substantially decreased the LOD score for this peak. The 24 carriers of this missense mutation clustered into five pedigrees in the ERF population (Additional file 1: Figure S2).

Genetic variants discovered by GWAS, based on individual single-nucleotide polymorphisms (SNPs), explain only a small proportion of the heritability of complex traits [10, 39, 40]; we found variants with larger effect sizes compared to the ones found with GWAS. Our analysis of rare coding variants in these linkage regions revealed a variant, rs138968470 on 11q13.1 in the MAP3K11 gene, associated with PC1. Conditional linkage analysis, including the MAP3K11 variant, reduced the LOD score (from 2.01 to 0.88), suggesting that this variant largely explained the linkage signal at this chromosomal location. The SNP is located in the first exon of a gene encoding a protein that belongs to the serine/threonine kinase family of mitogen-activated protein kinases. MAP3K11 (also known as Mixed Lineage Kinase 3 (MLK3)) [34], works as a positive regulator of the c-Jun N-terminal kinase (JNK) signalling pathway [41]. MAP3K11 has a CDC42 and Rac interacting proteins binding domain (CRIB); autophosphorylation of MAP3K11 and the induction of JNK is mediated through this CRIB domain bound to Cdc42/Rac/GTP [42]. JNK, an important member of the mitogen-activated protein kinase family (MAPK), is a pro-apoptotic kinase that plays an important role in the induction of cardiomyocyte apoptosis in various pathologies [43]. Apoptosis increases with LVH, a critical mechanism that mediates the transition from compensated hypertrophy to heart failure [44]. In this way, a damaging mutation in MAP3K11 may be related to regulation of JNK and the subsequent JNK controlled pathway.

The other significant missense variant was rs139580877, located on 5p14. This variant is in exon 9 of the gene that encodes the sperm flagellar protein (SPEF2), which has been postulated to play an important role in spermatogenesis and flagellar assembly [45]. This SNP was not found to be responsible for the linkage signal in the region, despite its strong association. The association with this relatively common variant (MAF = 0.015) could not be confirmed in the Rotterdam Study. One additional finding was studied further: a C/G variant (rs142551296) in the PRSS12 gene, underlying the SL locus on chromosome 4, which approached significance (P-value = $8.4 \times 10-4$),

but did not replicate in the Rotterdam Study. Absence of replication could be related to imputation quality for rs139580877 and the low number of carriers for the other SNPs (Additional file 1: Table S4).

A number of the linkage peaks contained SNPs identified in a large GWAS of these traits. Linkage analysis, conditioned on the index SNPs from the GWAS, did not significantly alter the linkage results. This suggests that the linkage peaks were not driven by the common variants identified in the GWAS.

No explanatory variants were found for most of the loci (suggestively) linked to LVH, for which there are a number of potential explanations. Linkage peaks are not precise in highlighting the location of the causal variant; even the region of interest cannot be easily pinpointed. Additionally, we did not take into account alternative mechanisms, such as structural and copy number variations (CNVs) or repeats in the linkage regions. Lastly, causal rare variants may be located outside the coding sequence, which we did not include in our sequencing analyses.

Conclusions

In conclusion, 13 loci were identified for ECG LVH proxy measures and PCs using linkage analysis in a large pedigree; these were subsequently fine-mapped with microarray and exome sequence data. Common variation from the microarrays did not explain these peaks. The exome data, though, suggested the involvement of MAP3K11 (11q13) in LVH through the regulation of JNK. However, we cannot exclude the presence of other variants that are in linkage disequilibrium with the MAP3K11 variant (rs138968470) that might explain the observed association.

Further analysis will need to be performed to demonstrate the involvement of this protein in LVH. A number of other suggestively linked peaks were determined. We could not explain these with microarray or exonic sequence variants at present, asking for more extensive follow-up outside the coding regions.

Acknowledgements

We are grateful to all study participants 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. Najaf Amin is supported by the Hersenstichting Nederland (the Netherlands Brain Foundation) (Project Number F2013(1)-28).

Rotterdam study: We are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

We thank Mr. Pascal Arp, Ms. Mila Jhamai and Mr. Marijn Verkerk for their help in creating the RS-Exome-Sequencing database.

Funding

The Erasmus Rucphen Family (ERF) study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP Grant No. 018947 (LSHG-CT-2006- 01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007–2013)/Grant Agreement HEALTH-F4–2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (No. QLG2-CT-2002-01254). High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organization for Scientific Research and the

Russian Foundation for Basic Research (NWO- RFBR 047.017.043), and Russian Federal Agency of Scientific Organizations projects VI.53.2.2 and 0324–2015-0003. Exome sequencing analysis in ERF was supported by the Netherlands Organization for the Health Research and Development grant for Project No. 91111025. This work was also supported by the Federal Agency of Scientific Organizations #0324–2015-0003 (IZ, AK and TA). Finally, this work was further funded by the European Union's Horizon 2020 research and innovation programme as part of the Common mechanisms and pathways in Stroke and Alzheimer's disease (CoSTREAM) project (www.costream.eu, grant agreement No 667375) and the European Union's Horizon 2020 research and innovation programme Marie Skłodowska-Curie Research and Innovation Staff Exchange (RISE) under the grant agreement No 645740 as part of the Personalized pREvention of Chronic DIseases (PRECeDI) project.

The Rotterdam Study is funded by Erasmus Medical Center (MC) and Erasmus University, Rotterdam; Netherlands Organization for the Health Research and Development; the Research Institute for Diseases in the Elderly; the Ministry of Education, Culture and Science; the Ministry for Health, Welfare and Sports; the European Commission; and the Municipality of Rotterdam. The Exome-Sequencing dataset was funded by the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) sponsored Netherlands Consortium for Healthy Aging (NCHA; Project No. 050–060-810), by the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and by the and by a Complementation Project of the Biobanking and Biomolecular Research Infrastructure Netherlands (BBMRI-NL; www.bbmri.nl; Project Number CP2010–41).

Availability of data and materials

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to them containing information that could compromise research participant privacy/consent.

Abbreviations

12LS 12-lead sum QRS product

BMI Body mass index

CRIB CDC42 and Rac interacting proteins binding domain

CV Cornell voltage product

ECG Electrocardiogram

ERF Erasmus Rucphen Family Study

GATK Genome Analysis Toolkit

GWAS Genome Wide Association Studies

HWE Hardy Weinberg Equilibrium

JNK c-Jun N-terminal kinase

LD Linkage disequilibrium

LVH Left ventricular hypertrophy

MAF Minor allele frequency

MEANS Modular ECG Analysis System

MI Myocardial infarction

PC Principal component

SL Sokolow-Lyon index

SNP Single-nucleotide polymporphisms

SNVs Single nucleotide variants

MAPK Mitogen activated protein kinase family

SPEF2 Sperm flagellar protein

CNV Copy number variations

SBP Systolic blood pressure

DBP Diastolic blood pressure

dbNSFP Database of human non-synonymous SNPs and their functional predictions

Authors' contributions

Formal Analysis: CTS, IVZ, MN, MEvdB, NA, AD, EvL, AIG, LBP, JAK, AVK, TIA, CMvD, AI. Writing – Original Draft Preparation: CTS, RW, CMvD, AI. Investigation and Resources: CMR, BAO, BHS, AGU, CMvD. Software: AVK. Supervision: CMR, RW, TIA, CMvD, AI. Conceptualization: CMvD, AI. Writing – Review & Editing: ALL AUTHORS. All authors read and approved the final manuscript.

Notes

Ethics approval and consent to participate

The Medical Ethics Committee of the Erasmus University Medical Center approved the ERF study protocol and all participants, or their legal representatives, provided written informed consent. Approval MEC 213.57512002.

Consent for publication

Not applicable.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Maps and institutional affiliations.

Footnotes

Electronic supplementary material

The online version of this article (10.1186/s12920-018-0339-9) contains supplementary material, which is available to authorized users.

References

1. Benjamin, E.J. and D. Levy, *Why is left ventricular hypertrophy so predictive of morbidity and mortality?* Am J Med Sci, 1999. **317**(3): p. 168-75.
2. Arnett, D.K., L. de las Fuentes, and U. Broeckel, *Genes for left ventricular hypertrophy.* Curr Hypertens Rep, 2004. **6**(1): p. 36-41.
3. Bella, J.N. and H.H. Goring, *Genetic epidemiology of left ventricular hypertrophy.* Am J Cardiovasc Dis, 2012. **2**(4): p. 267-78.
4. Arnett, D.K., et al., *Novel genetic variants contributing to left ventricular hypertrophy: the HyperGEN study.* J Hypertens, 2009. **27**(8): p. 1585-93.
5. Shah, S., et al., *Four genetic loci influencing electrocardiographic indices of left ventricular hypertrophy.* Circ Cardiovasc Genet, 2011. **4**(6): p. 626-35.
6. Post, W.S., et al., *Heritability of left ventricular mass: the Framingham Heart Study.* Hypertension, 1997. **30**(5): p. 1025-8.
7. Foppa, M., B.B. Duncan, and L.E. Rohde, *Echocardiography-based left ventricular mass estimation. How should we define hypertrophy?* Cardiovasc Ultrasound, 2005. **3**: p. 17.
8. Mayosi, B.M., et al., *Genome-wide linkage analysis of electrocardiographic and echocardiographic left ventricular hypertrophy in families with hypertension.* Eur Heart J, 2008. **29**(4): p. 525-30.
9. Mutikainen, S., et al., *Genetic influences on resting electrocardiographic variables in older women: a twin study.* Ann Noninvasive Electrocardiol, 2009. **14**(1): p. 57-64.
10. Silva, C.T., et al., *Heritabilities, proportions of heritabilities explained by GWAS findings, and implications of cross-phenotype effects on PR interval.* Hum Genet, 2015. **134**(11-12): p. 1211-9.
11. van der Harst, P., et al., *52 Genetic Loci Influencing Myocardial Mass.* J Am Coll Cardiol, 2016. **68**(13): p. 1435-1448.
12. Wang, L., et al., *Novel quantitative trait locus is mapped to chromosome 12p11 for left ventricular mass in Dominican families: the Family Study of Stroke Risk and Carotid Atherosclerosis.* BMC Med Genet, 2009. **10**: p. 74.
13. Zhi, D., et al., *Whole-exome sequencing and an iPSC-derived cardiomyocyte model provides a powerful platform for gene discovery in left ventricular hypertrophy.* Front Genet, 2012. **3**: p. 92.
14. Pardo, L.M., et al., *The effect of genetic drift in a young genetically isolated population.* Ann Hum Genet, 2005. **69**(Pt 3): p. 288-95.
15. Aulchenko, Y.S., et al., *Linkage disequilibrium in young genetically isolated Dutch population.* Eur J Hum Genet, 2004. **12**(7): p. 527-34.
16. Sayed-Tabatabaei, F.A., et al., *Heritability of the function and structure of the arterial wall: findings of the Erasmus Rucphen Family (ERF) study.* Stroke, 2005. **36**(11): p. 2351-6.
17. *1999 World Health Organization-International Society of Hypertension Guidelines for the Management of Hypertension. Guidelines Subcommittee.* J Hypertens, 1999. **17**(2): p. 151-83.
18. Eijgelsheim, M., et al., *Identification of a common variant at the NOS1AP locus strongly associated to QT-interval duration.* Hum Mol Genet, 2009. **18**(2): p. 347-57.
19. van Bemmel, J.H., J.A. Kors, and G. van Herpen, *Methodology of the modular ECG analysis system MEANS.* Methods Inf Med, 1990. **29**(4): p. 346-53.
20. Willems, J.L., et al., *The diagnostic performance of computer programs for the interpretation of electrocardiograms.* N Engl J Med, 1991. **325**(25): p. 1767-73.

21. Willems, J.L., et al., *A reference data base for multilead electrocardiographic computer measurement programs.* J Am Coll Cardiol, 1987. **10**(6): p. 1313-21.
22. de Bruyne, M.C., et al., *Diagnostic interpretation of electrocardiograms in population-based research: computer program research physicians, or cardiologists?* J Clin Epidemiol, 1997. **50**(8): p. 947-52.
23. Leening, M.J., et al., *Unrecognised myocardial infarction and long-term risk of heart failure in the elderly: the Rotterdam Study.* Heart, 2010. **96**(18): p. 1458-62.
24. O'Connell, J.R. and D.E. Weeks, *PedCheck: a program for identification of genotype incompatibilities in linkage analysis.* Am J Hum Genet, 1998. **63**(1): p. 259-66.
25. Abecasis, G.R., et al., *Merlin--rapid analysis of dense genetic maps using sparse gene flow trees.* Nat Genet, 2002. **30**(1): p. 97-101.
26. Gudbjartsson, D.F., et al., *Allegro, a new computer program for multipoint linkage analysis.* Nat Genet, 2000. **25**(1): p. 12-3.
27. Kirichenko, A.V., et al., *PedStr software for cutting large pedigrees for haplotyping, IBD computation and multipoint linkage analysis.* Ann Hum Genet, 2009. **73**(Pt 5): p. 527-31.
28. Nothnagel, M., et al., *A comprehensive evaluation of SNP genotype imputation.* Hum Genet, 2009. **125**(2): p. 163-71.
29. Li, Y., et al., *MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes.* Genet Epidemiol, 2010. **34**(8): p. 816-34.
30. Aulchenko, Y.S., et al., *GenABEL: an R library for genome-wide association analysis.* Bioinformatics, 2007. **23**(10): p. 1294-6.
31. Aulchenko, Y.S., M.V. Struchalin, and C.M. van Duijn, *ProbABEL package for genome-wide association analysis of imputed data.* BMC Bioinformatics, 2010. **11**: p. 134.
32. Storey, J.D. and R. Tibshirani, *Statistical significance for genomewide studies.* Proc Natl Acad Sci U S A, 2003. **100**(16): p. 9440-5.
33. Brouwer, R.W., et al., *NARWHAL, a primary analysis pipeline for NGS data.* Bioinformatics, 2012. **28**(2): p. 284-5.
34. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.
35. Hofman, A., et al., *The Rotterdam Study: 2014 objectives and design update.* Eur J Epidemiol, 2013. **28**(11): p. 889-926.
36. Gorski, M., et al., *1000 Genomes-based meta-analysis identifies 10 novel loci for kidney function.* Sci Rep, 2017. **7**: p. 45040.
37. Barrios, V., et al., *[Computerized interpretation of the electrocardiogram in the diagnosis of left ventricular hypertrophy. The ELECTROPRES project].* Rev Clin Esp, 2011. **211**(8): p. 391-9.
38. Schillaci, G., F. Battista, and G. Pucci, *A review of the role of electrocardiography in the diagnosis of left ventricular hypertrophy in hypertension.* J Electrocardiol, 2012. **45**(6): p. 617-23.
39. Fridley, B.L. and J.M. Biernacka, *Gene set analysis of SNP data: benefits, challenges, and future directions.* Eur J Hum Genet, 2011. **19**(8): p. 837-43.
40. Eichler, E.E., et al., *Missing heritability and strategies for finding the underlying causes of complex disease.* Nat Rev Genet, 2010. **11**(6): p. 446-50.
41. Chadee, D.N. and J.M. Kyriakis, *MLK3 is required for mitogen activation of B-Raf, ERK and cell proliferation.* Nat Cell Biol, 2004. **6**(8): p. 770-6.
42. Du, Y., et al., *Cdc42 induces activation loop phosphorylation and membrane targeting of mixed lineage kinase 3.* J Biol Chem, 2005. **280**(52): p. 42984-93.

43. Rui, T. and Q. Tang, *IL-33 attenuates anoxia/reoxygenation-induced cardiomyocyte apoptosis by inhibition of PKCbeta/JNK pathway.* PLoS One, 2013. **8**(2): p. e56089.
44. Gelpi, R.J., et al., *Apoptosis in severe, compensated pressure overload predominates in nonmyocytes and is related to the hypertrophy but not function.* Am J Physiol Heart Circ Physiol, 2011. **300**(3): p. H1062-8.
45. Sironen, A., et al., *Loss of SPEF2 function in mice results in spermatogenesis defects and primary ciliary dyskinesia.* Biol Reprod, 2011. **85**(4): p. 690-701.

Chapter 5


ATTCGGTACCTCATAACGGCTAGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTT
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGGTACCTA

Chapter 5

52 Genetic Loci Influencing Myocardial Mass

van der Harst P, van Setten J, Verweij N, Vogler G, Franke L, Maurano MT, Wang X, Mateo Leach I, Eijgelsheim M, Sotoodehnia N, Hayward C, Sorice R, Meirelles O, Lyytikäinen LP, Polašek O, Tanaka T, Arking DE, Ulivi S, Trompet S, Müller-Nurasyid M, Smith AV, Dörr M, Kerr KF, Magnani JW, Del Greco M F, Zhang W, Nolte IM, Silva CT, Padmanabhan S, Tragante V, Esko T, Abecasis GR, Adriaens ME, Andersen K, Barnett P, Bis JC, Bodmer R, Buckley BM, Campbell H, Cannon MV, Chakravarti A, Chen LY, Delitala A, Devereux RB, Doevendans PA, Dominiczak AF, Ferrucci L, Ford I, Gieger C, Harris TB, Haugen E, Heinig M, Hernandez DG, Hillege HL, Hirschhorn JN, Hofman A, Hubner N, Hwang SJ, Iorio A, Kähönen M, Kellis M, Kolcic I, Kooner IK, Kooner JS, Kors JA, Lakatta EG, Lage K, Launer LJ, Levy D, Lundby A, Macfarlane PW, May D, Meitinger T, Metspalu A, Nappo S, Naitza S, Neph S, Nord AS, Nutile T, Okin PM, Olsen JV, Oostra BA, Penninger JM, Pennacchio LA, Pers TH, Perz S, Peters A, Pinto YM, Pfeufer A, Pilia MG, Pramstaller PP, Prins BP, Raitakari OT, Raychaudhuri S, Rice KM, Rossin EJ, Rotter JI, Schafer S, Schlessinger D, Schmidt CO, Sehmi J, Silljé HH, Sinagra G, Sinner MF, Slowikowski K, Soliman EZ, Spector TD, Spiering W, Stamatoyannopoulos JA, Stolk RP, Strauch K, Tan ST, Tarasov KV, Trinh B, Uitterlinden AG, van den Boogaard M, van Duijn CM, van Gilst WH, Viikari JS, Visscher PM, Vitart V, Völker U, Waldenberger M, Weichenberger CX, Westra HJ, Wijmenga C, Wolffenbuttel BH, Yang J, Bezzina CR, Munroe PB, Snieder H, Wright AF, Rudan I, Boyer LA, Asselbergs FW, van Veldhuisen DJ, Stricker BH, Psaty BM, Ciullo M, Sanna S, Lehtimäki T, Wilson JF, Bandinelli S, Alonso A, Gasparini P, Jukema JW, Kääb S, Gudnason V, Felix SB, Heckbert SR, de Boer RA, Newton-Cheh C, Hicks AA, Chambers JC, Jamshidi Y, Visel A, Christoffels VM, Isaacs A, Samani NJ, de Bakker PI.

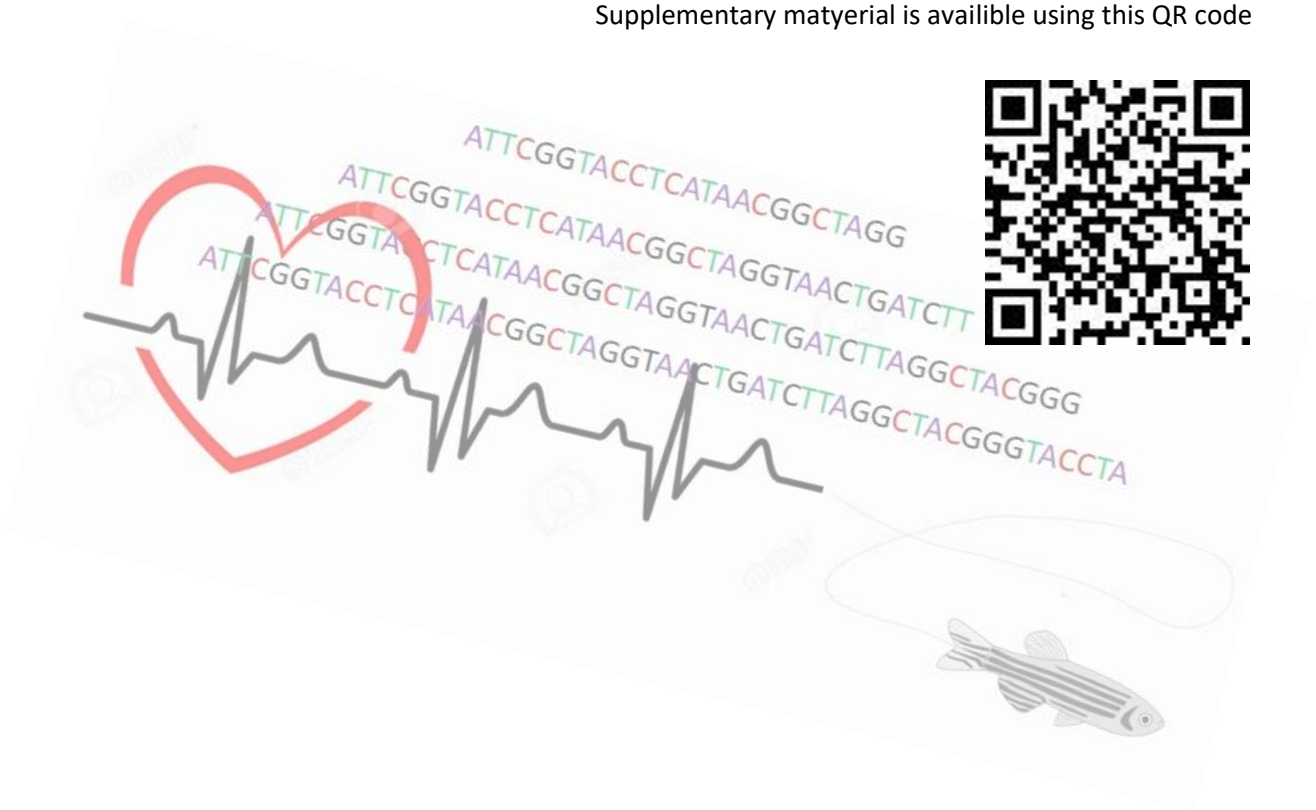
Chapter 5

52 Genetic Loci Influencing Myocardial Mass

J Am Coll Cardiol. 2016 Sep 27; 68 (13):1435-48. doi: 10.1016/j.jacc.2016.07.729.

PMID:27659466

Supplementary matyerial is availible using this QR code



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.

Objectives. This meta-analysis sought 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, that are reliably associated with 1 or more QRS phenotypes at $p < 1 \times 10^{-8}$. 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 provided evidence that these loci play a role in cardiac hypertrophy. We further highlighted 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.

Key Words. Electrocardiogram; genetic association study; heart failure; left ventricular hypertrophy; QRS

Abbreviations and Acronyms. DHS, deoxyribonuclease hypersensitivity sites; ECG, electrocardiogram; eQTL, expression quantitative trait locus; GWAS, genome-wide association study; LD, linkage disequilibrium; RNAi, ribonucleic acid interference; SNP, single nucleotide polymorphism; TF, transcription factor

FOR EDITORIAL REASONS, THIS CHAPTER IS ONLY AVAILABLE IN THE JOURNAL OR THROUGH PUBMED. (Free Access)

CHAPTER 6


ATTCGGTACCTCATAACGGCTAGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTT
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGGTACCTA

CHAPTER 6

Arhgap24 a suspicious gene involved in heart development.

Claudia Tamar Silva, Herma van der Linde, Lies-Anne Severijnen, Jan A. Kors, Abbas Dehghan,

Cornelia M. van Duijn, Aaron Isaacs, Rob Willemsen

Manustript in preparation

Abbreviations. Genome Wide Association Studies (GWAS), electrocardiogram (ECG), Erasmus Rucphen Family study (ERF), Minor Allele frequency (MAF), splice blocking morpholinos (SB MO), cardiac myosin light chain 2 gene (cmlc2), enhanced green fluorescent protein (EGFP), hours post fertilization (hpf), complementary DNA (cDNA), Genotype-Tissue expression (GTex), Modular ECG Analysis System (MEANS), wildtype (wt), Sokolow-Lyon (SL), Cornell Voltage (CV), 12-lead sum (12LS), Myocardial Infarction (MI), Left Ventricualr Hypertrophy (LVH), Genotupe-Tissue expression (GTex), Single Nucleotide Variants (SNV), Genome Analysis Toolkit (GATK).

Abstract. *ARHGAP24* is a gene previously associated with PR interval, but functional variants or experiments supporting its role in cardiac development and function are lacking**.** The aim of this project was to establish the normal cellular function of *ARHGAP24* related to heart development and function, using a morpholino knockdown strategy in zebrafish. Knockdown of this gene in zebrafish showed heart abnormalities and a reduction in heart rate in morphants. Additionally, we performed exon sequence analyses of *ARHGAP24* in an isolated Dutch population. The sequencing data revealed six damaging variants predicted by polyPhen and CADD among thirty-nine variants. Analysis of these damaging mutations showed that rs144785317 influenced the QT and QRS. Nominal evidence for association to QT was found for rs35521695, a missense mutation at the exon-intron border. The most strongly associated variant was rs61758879, a missense mutation, associated to 12LS. Our findings support a functional role for *ARHGAP24* in normal heart function.

Key words. Electrocardiogram, zebrafish, ARHGAP24, GWAS, ERF.

Introduction

The electrocardiogram (ECG) is an important tool for evaluating the cardiac conduction system. The measurements obtained from the ECG include RR interval, PR interval, QRS duration and QT interval, representing various aspects of the conduction system. Each of these have been found to be predictive of cardiovascular events [1, 2]. Family studies have demonstrated that these measurements have a substantial genetic component, with heritabilities that range between 34 and 40% [3].

Genome-wide association studies (GWAS) have identified ~ 75 genes which contribute to ECG trait variability [1, 2, 4-12]. Among these,ARHGAP24 gene is a member of the *ARHGAP* family, which encodes for a negative regulator of Rho GTPases and has been implicated in actin remodelling, cell polarity and cell migration [13]. A significant association between prolongation of the PR interval and common intronic variant rs7660702 in *ARHGAP24* was reported. Because their role in differentiation and development, this gene is a promising gene with unknown function that might be related to cardiac development. However, the strongly associated, rs76922808 (MAF=0.32) variant is an intronic variant that does not have any known functional effects on the protein, asking for more detailedfunctional analyses. Another important challenge in the “post GWAS era” is to validate the pathogenicity of *ARHGAP24* in animal models. Functional analysis using zebrafish as an animal model is advantageous for heart studies because embryonic transparency allows for the easy assessment of heart, and other, developmental abnormalities.

The aim of this project was to establish the normal cellular function of *ARHGAP24* in early heart development, using a morpholino knockdown strategy in zebrafish, and exome sequence analysis in a family based cohort (ERF) population was conducted to establish the role of less frequent coding variants.

Experimental Section

Zebrafish strains and husbandry

The zebrafish (Danio rerio, Hamilton 1822) strains used for this work were transgenic zebrafish carrying a cardiac-specific promoter containing the upstream sequence of the zebrafish cardiac myosin light chain 2 gene (cmlc2) and a reporter enhanced green fluorescent protein (EGFP) [14, 15] and the control AB line. Adults were maintained at 28˚C on a 14 hour-light/10 hour-dark cycle. Embryos were collected from natural mating and raised in system water containing methylene blue at 28˚C. Developmental stages were determined according to Kimmel [16]. All procedures and conditions were in accordance with Dutch animal welfare legislation. The animal protocols used in this work were evaluated and approved by the Institutional Review Board for experimental animals of Erasmus MC, Rotterdam (DEC Nr. EMC 2088 (140-10-09); October 18th, 2012). They are in accordance with FELASA and ARRIVA guidelines and the European law for Laboratory Animal Experimentation.

Zebrafish Arhgap24 gene

The protein and gene sequences of Arhgap24 from human and zebrafish were taken from the Ensemble genome browser (accession numbers ENSDARG0000010097). Three isoforms protein coding are reported, arhgap24-003 (ENSDART00000172124.1), which encodes the same peptide as arhgap24-201 (ENSDAR00000170710.1) and arhgap-001 (ENSDART00000137809). Sequence alignment is shown in Annex 1, which revealed two distinct isoforms. Alignment with human isoforms is shown in Annex 2. Sequencing of the first part of the zebrafish gene was carried out in order to facilitate morpholino design.

Morpholino

Two non-overlapping morpholino antisense oligonucleotides were used; both morpholinos were designed as splice blocking morpholinos (SB MO). One morpholino was designed over the splice acceptor site of intron 2 (3 in the other isoform) and the second one was designed over the splice donor site of intron 3 (4 in the other isoform). Both morpholinos were obtained from Gene-Tools (Philomath, OR, USA): ATCCCTGAAACACAAGCACACAGGA SB MO GTGCATTAAGAGCAAGTACCAGTCA SB MO2. Morpholinos were reconstituted in distilled water and further diluted in Danieau buffer and Phenol Red (Sigma Chemical o., St Louis, MO, USA) solution. Injections were carried out using eggs at the one or two cell stage, into the yolk sac, using a Pneumatic PicoPump (World Precision Instruments, Berlin, Germany). Morpholino titration was performed using different concentrations of each morpholino individually and in combination. Efficient knockdown and minimum off target effects was established at 4 ng MO of each MOs, we use it together. Injected embryos were incubated at 28.5˚C, after 24, 48 and 72 hours post fertilization (hpf) morphants were collected. Knockdown efficiency was confirmed through quantitative RT-PCR (q RT-PCR), using delta delta CT method. Succinate dehydrogenase complex flavoprotein subunit A (*Sdha)*, was use as housekeeping gene.

Preparation of RNA from zebrafish and quantitative polymerase chain reaction

Total RNA was isolated from 50 uninjected and 50 morpholino injected embryos at 24, 48 and 72 hpf using RNA bee (Tri-Test.inc). Complementary DNA (cDNA) was obteined using the SuperScript First-Strand synthesis system for RT-PCR (Invitrogen, California USA) according to manufacturer's instructions. To measure mRNA levels, q RT-PCR on cDNA samples was carried out using SYBR® Select Master Mix for the CFX96 qPCR detection system. Primers used for q RT-PCR were designed

using Primer3Plus [17, 18]. Primers for the *sdha* reference gene were designed using Primer Express software (version 2.0.0).

P53 Coinjection

Although morpholino injection induces sequence-specific gene knockdown in multiple systems, it has been reported that sometimes off-target effects may occur through P53-induced apoptosis. Coinjections of *Arhgap24* SB MO with *p53* MO is a tool to evaluate off-target effects due to P53 apoptosis [19]. Embryos were injected simultaneously with 4 ng of both *Arhgap24* SB MO and 4 ng of *p53* MO in each embryo. Phenotype was assessed after 24, 48 and 72 hpf.

Microscopy, heartbeat and histology

Wild type embryos and morphants were analyzed *in vivo* at 24, 48 and 72 hpf using fluorescence microscopy (Leica MZ16FA). Heartbeat was counted for 30 seconds at 48 hpf. Comparisons of the heart rate difference between morphants and WT was evaluated in order to establish the effect of the knockdown on heart rhythm. For histology, embryos were fixed in 4% paraformaldehyde at 4˚C overnight, embedded in paraffin using standard procedures and 6 µm sections were cut. Subsequently, hematoxilin-eosin staining of the section was carried out using a standard protocol.

Study population

The human component of the study was performed in the Erasmus Rucphen Family study (ERF), a cohort derived from a region in the southwest of the Netherlands. The population was established in the middle of the 18th century by a limited number of founders, has experienced minimal immigration and emigration, and has exponentially increased in size in the last century. The ERF study was instituted in this population to unravel genes underlying quantitative trait variation in humans [20]. Since the population was sampled on the basis of genealogy, and not on a specific phenotype, the chances of findings confounded by disease status or co-morbidity are reduced.

Medical practitioners performed interviews at the time of blood sampling. The Medical Ethics Committee of the Erasmus University Medical Center approved the ERF study protocol and all participants, or their legal representatives, provided written informed consent.

ECG interpretation and measurement

Examinations included 12-lead ECG measurements. A 10-second 12-lead ECG (on average, 8 to 10 beats) was recorded with an ACTA-ECG electrocardiograph (Esaote, Florence, Italy) with a sampling frequency of 500 Hz. Digital measurements of the ECG parameters were made using the Modular ECG Analysis System (MEANS) [21]. Briefly, MEANS operates on multiple simultaneously recorded leads, which are transformed to a detection function that brings out the QRS complex and the other parts of the signal. MEANS determines common onsets and offsets for all 12 leads together on one representative averaged beat, with the use of template matching techniques. The measurement and diagnostic performance of MEANS has been extensively evaluated, both by the developers and by others [22, 23]. The MEANS criteria for myocardial infarction (MI) are mainly based on pathological Q waves, QR ratio, and R-wave progression [21]. A cardiologist, specialized in ECG methodology, ascertained the final diagnosis of MI.

MEANS was used to measure several ECG parameters (QRS, PR, and QT) and the three LVH proxies (SL, CV, and 12LS). Sokolow–Lyon was defined as the sum of the S wave in V1 plus the R wave in V5 or V6, Cornell as the sum of R in aVL and the S in V3, and 12-lead as the sum of R–S in all 12 leads; these three voltages were then multiplied by QRS duration to obtain voltage-duration products as an approximation of the area under the QRS complex [24-26]. QT interval was adjusted for heart rate using Bazett's formula [27, 28]. All traits were adjusted for sex, age, BMI, height and heart rate (with the exception of QT), and rank transformed prior to analysis.

Statistical analysis

Individuals were excluded from analysis if their ECG showed evidence of atrial fibrillation, myocardial infarction, left or right bundle branch block, or atrioventricular block. Additional exclusion criteria consisted of pacemaker implantation, Wolff-Parkinson-White syndrome, pregnancy, and use of Type I or III anti-arrythmic medications or digoxin, which may shorten the QT interval [5]. Individuals with QRS > 120 ms were excluded from the QRS, QT and LVH proxy analyses. Those with PR ≥ 320 ms or ≤ 80 ms were excluded from the PR analyses. Those with QRS axis > 90 or < -30 were excluded from the LVH proxy analyses. These exclusions were implemented to keep our data consistent with the GWAS. Difference among injected and and WT was done by Mann-Whitney, Wilcoxon and TTest. Comparison among phenotype positive and WT was done using TTest.

Exome sequencing

Exomes of 1309 individuals from the ERF Study were sequenced at the Center for Biomics of the Department of Cell Biology of the Erasmus MC, the Netherlands, using the Agilent V4 capture kit on an Illumina HiSeq2000 sequencer using the TruSeq Version 3 protocol. Mean depth base was 74.23x (median = 57x) and mean depth region was 65.26x (median = 52.87x). The sequence reads were aligned to the human genome build 19 (hg19) using BWA and the NARWHAL pipeline [29, 30]. The aligned reads were processed further using the IndelRealigner, MarkDuplicates and TableRecalibration tools from the Genome Analysis Toolkit (GATK) and Picard (http://picard.sourceforge.net) to remove systematic biases and to recalibrate the PHRED quality scores in the alignments. Genetic variants were called using the Unified Genotyper tool of the GATK. About 1.4 million Single Nucleotide Variants (SNVs) were called and, after removing the low quality variants (QUAL < 150), we retrieved 577,703 SNVs in 1,309 individuals. Further, for comparison and to predict the functionality of the variants, annotations were also performed

using the dbNSFP (database of human non-synonymous SNPs and their functional predictions http://varianttools.sourceforge.net/Annotation/DbNSFP) and Seattle (http://snp.gs.washington.edu/SeattleSeqAnnotation131/) databases. These databases gave functional prediction results from four different programs (polyPhen2, SIFT, MutationTaster and LRT), apart from gene and variant annotations.

eQTL Analyses

We used the Genotype-Tissue expression (GTex) project database (http://www.gtexportal.org/home/) to examine whether interesting variants had *cis* eQTL effects.

Results

Identification and characterization of zebrafish Arhgap24

A search for *Arhgap24* orthologues in zebrafish through the Ensemble database revealed three reported isoforms, due to an alternative first exon. Alignment of the human and zebrafish Arhgap24 proteins showed homologies of 65% homology in the nucleotide sequence and 72% in the amino acid sequence (Annex 1).

Figure 1. arhgap24 zebrafish isoforms
ENSDART00000170710.1 and ENSDART00000172124.1 have the same coding sequence, and share sequence with ENSDAR00000158836.1 as is shown in the figure since the aa 92 of the first two isoforms.

Knockdown efficiency

After determining primer efficiency, we performed q RT-PCR to measure in time the constitutive *Arhgap24* expression in the wt and cmlc2:GFP transgenic embryos. *Arhgap24* expression decreased significantly over time in both the wt and cmlc2:GFP embryos, being at the highest level at 24 hpf for both lines (Figure 2). All time periods were compared with expression levels in embryos at 24 hpf. Expression of *Arhgap24* in cmlc2:GFP transgenic embryos, decreased at 48 and 72 hpf, compared with the expression at 24 hpf, significant difference was found between WT and morphants, we use Mann-Whitney and we get a signitivan difference among the expression of both groups .

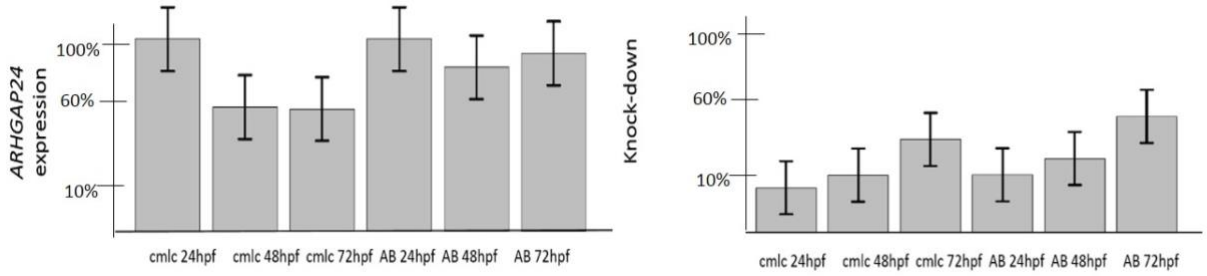


Figure 2. Knockdown of arhgap24 in zebrafish
Difference among expression of WT and MOs shows significant difference among these two groups. Statistical analysis givis a P-value of 0.0000.

Knockdown efficiency was confirmed by q RT-PCR in morphants at different time points. We performed injections with of a mix of both MOs and morphants were collected after 24, 48 and 72 hours. Figure 1 shows the q RT-PCR results, revealing a knockdown efficacy in the cmcl2:GFP reporter line of 87%, 83.6% and 73.4% at 24 hrs, 48 hrs and 72 hrs, respectively (Table 1). Similar results were obtained for the AB line. Each injected group was compared with control WT embryos without injection.

Zebrafish line	Knockdown (%)
cmcl2:GFP 24	0
24_inj	87
48 inj	83
72 inj	73
AB 24	0
AB 24 inj	83
AB 48 inj	78
AB 72 inj	66

Table 1. Knockdown efficacy. In this table is showed knockdown percentage.

Microscopy, heart beating and histology

This efficiency of knockdown is consistent with heart abnormalities seen in the morphants; in 50% of the injected embryos, we observed cardiac defects, as shown in Figure 3. Figure 3A shows a wild type embryo, while Figure 3B depicts a morphant lacking a cardiac loop. Additionally, these

embryos suffer from cardiac edema. Figure 3C shows a morphant with a reverse loop, while the morphant in Figure 3D shows an enlarged atrium. Differences in cardiac morphology were also confirmed by histological analysis, as illustrated in Figure 4. The panel on the left shows a WT embryo heart at 3dpf. The right depicts a morphant at 3 dpf with a dilated heart, loop defect and cardiac edema.

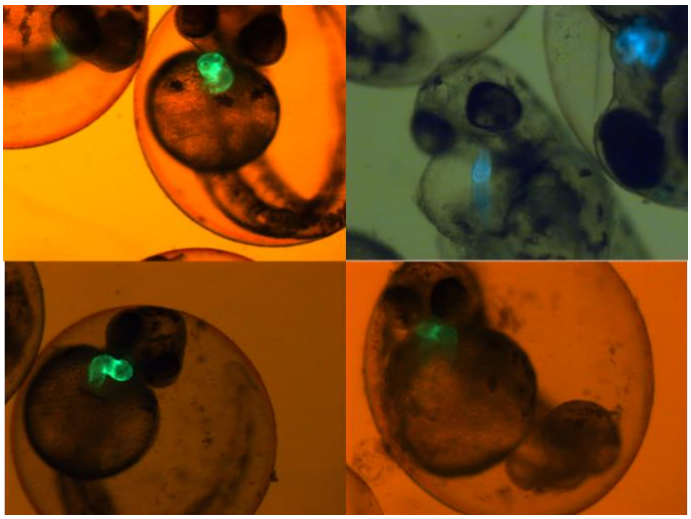


Figure 3. Cardiac abnormalities in ARHGAP24 MO Embryos at 48 hpf.3A: Wild-type, 3B: heart without cardiac loop, 3C: Embryo with reverse loops and 3D: Embryo with heart with big atrium.

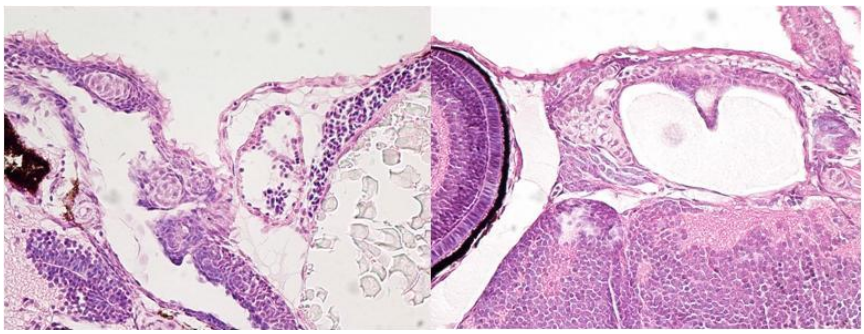


Figure 4. Cardiac changes after MO injection
The panel on the left shows a WT embryo heart at 3dpf. The right depicts a morphant at 3 dpf with a dilated heart, loop defect and cardiac edema.

Next, we classified morphants according to their phenotype as phenotype positive (reverse loop, big chambers, no loop, etc) and phenotype negative (normal heart). Heart rate was assessed in morphants that were phenotype positive, phenotype negative or wild type. Heart rate was significantly different between morphants (all injected embryos) and the control group (P = $4.856x10^{-5}$), as illustrated in Figure 5.

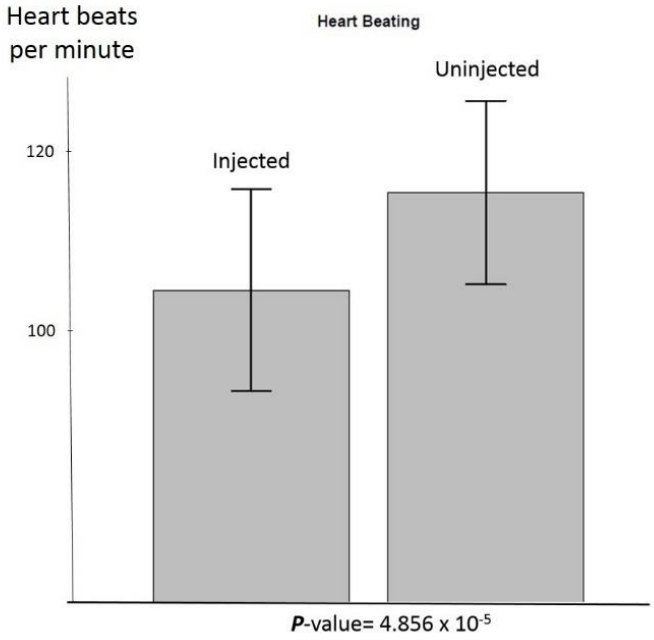


Figure 5. Heart Beat counts in WT and Morphants of 48 hpf.
Box plot that represents the significant differences between heart beating in WT embryos and morphants (injected).

ARHGAP24 down-regulated genes

ARHGAP24 is known to be involved in the regulation of two rhoGTPases, Cdc42 and Rac1 [31]. For this reason, we explored expression levels of these two genes after MO injection. We performed a q RT-PCR for these two downstream genes in the *ARHGAP24* pathway, however, we did not observe significant differences between WT and morphants at 1 and 2dpf (Data not shown).

P53 co-injection

To analyze off-target effects of the MOs we performed co-injections of our SB-MO with P53. We observed the same phenotype between the morphants co-injected with P53 and the morphants injected with the morpholino alone. P53 injected embryos did not exhibit differences compared with the control group, as depicted in figure 6.

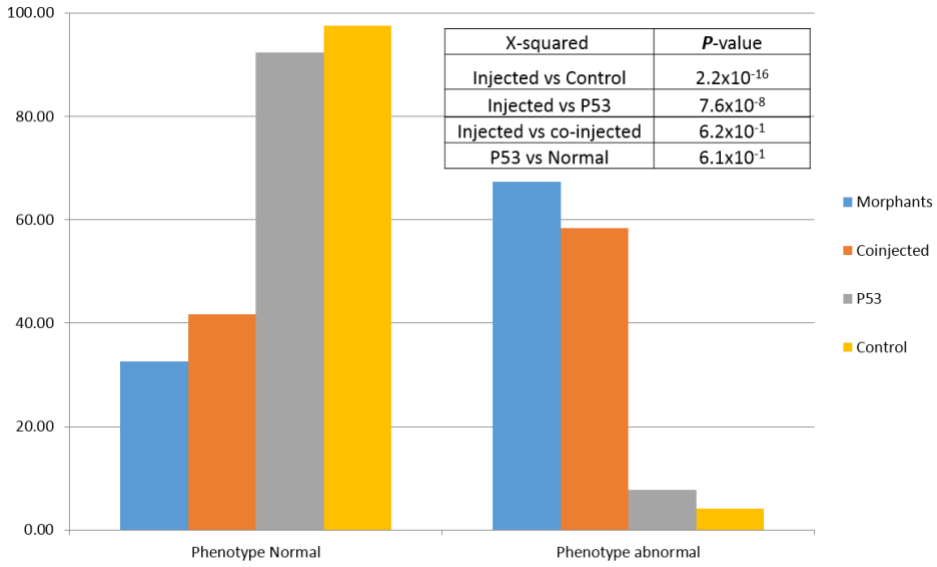


Figure 6. P53 rescue experiment**.**
After co-injection with P53 the injected embryos do not show a cardiac phenotype

Exome sequencing of the ARHGAP24 region

To determine potential rare variants in *ARHGAP24*, we explored exonic variants in 1309 individuals from the ERF study. All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Medical Ethics Committee of Erasmus MC. We found thirty nine variants, among these six damaging variants, as predicted by polyPhen-2 and CADD: rs144785317, rs61758879, rs35521695, rs147870358, 4:86643074 and 4:86916568 (Table 2). T-test analysis of these damaging mutations revealed that rs144785317 in codon 67 (G→E) associated the QT and QRS intervals (nominal ***P*** = 0.04 & 0.045). Prediction of the consequences of rs144785317 using

mutation taster[32] included loss of the PH domain of the protein and an altered splice site, which PolyPhen-2 predicted as probably damaging (score = 1). Also for rs35521695 a nominal significant association to QT was found. This variant results in a missense mutation (P → A) in codon 417 affecting an exon-intron border, which PolyPhen-2 predicts as probably damaging (score = 0.996). The most strongly associated variant is rs61758879, which is associated to 12LS ($p=6.2 \times 10^{-5}$). This variant involved a missense mutation (R → L). Analyses for the remaining variants did not provide any evidence of association.

Variant	Alleles	MAF	polyPhen	Carriers (n)	CADD	NAP*	*P*-value
				9	27.70	QT	0.04
rs144785317	A/G	3.4×10^{-3}	1			QRS	0.05
Chr4:86643074	C/T	3.8×10^{-4}	0.996	1	23.80	-	-
rs61758879	G/T	4.5×10^{-3}	0.996	12	11.09	-	-
rs35521695	C/G	5.0×10^{-2}	0.429	132	22.80	QT	0.01
rs147870358	C/T	1.0×10^{-3}	0.909	3	23.60	-	-
Chr4: 86916568	A/C	7.6×10^{-4}	0.682	2	9.67	-	-

Table 2. ARHGAP24variants predicted to be possibly damaging (0.15-0.85) and probably damaging (>0.85) by polyPhen in the ERF Study. *NAP: Nominally associated Phenotype

eQTL analyses

No significant cis or trans eQTLs effects for rs7660702, rs7692808, rs144785317, rs35521695, rs61758879, rs35521695, rs142672228 or rs147870358 were found in the GTEx project database (http://www.gtexportal.org/home/). Even though checking co-expressed genes (http://coxpresdb.jp/), we found two interesting genes *CAV1* and *CXCR4* (Figure 7). *CAV1* encodes for Caveolae Protein, a component of plasma membrane involved in cytoesqueleton remodeling, *CXR4* encodes for a CXC chemokine receptor specific for stromal cell-derived factor-1.

Figure 7. ARHGAP 24 co-expressed genes

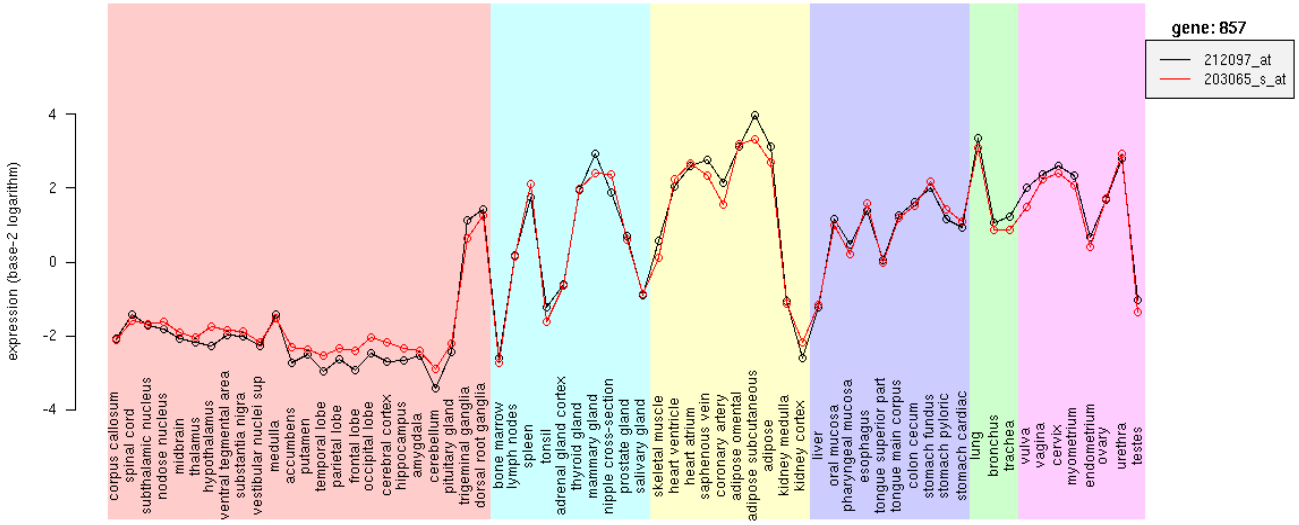


Figure 7A. CAV1 co-expression

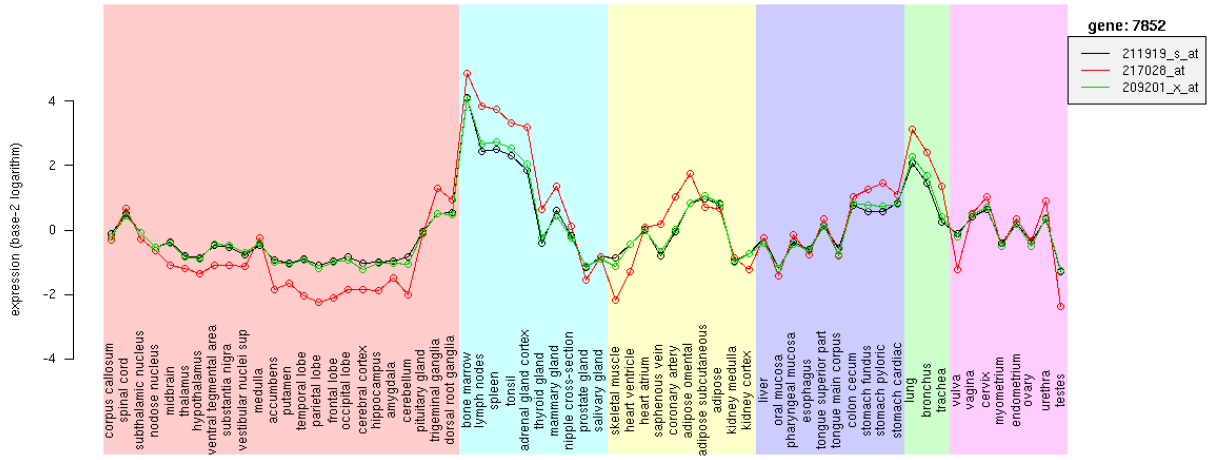


Figure 7B. *CXCR4* co-expression

Discussion

In the present study, we manipulated the expression of *Arhgap24* using a MO antisense strategy to examine the contribution of transient Arhgap24 protein knockdown to early cardiac development and function. In wild type embryos, expression of this gene starts at an early embryonic stage (24 hpf) and is reduced to approximately 60% after 48 and 72 hpf. This expression pattern suggests a function during very early development for *Arhgap24*. The results of our study reveals that knockdown of *Arhgap24* is related to abnormal heart development and function. Morphants, with a knockdown efficacy of approximately 85%, exhibited significant heart abnormalities and heart rate reduction. Our finding that heart rate reductions occurred in fish without observable heart abnormalities corroborates with the findings of GWAS that *Arhgap24* is associated with a mild phenotype. Histological analysis of zebrafish heart in embryos with cardiac defects exhibits dilated hearts compared with wild type embryos. Although the phenotype is mild, our results support a role of *Arhgap24* in normal cardiac development and function.

Rho GTPases, including Rac1 and Cdc42, comprise a major branch of the Ras superfamily of small GTPases, and Rho GTPase function has been implicated in cancer progression due to their function in cell migration, growth, proliferation, survival and angiogenesis [33]. *ARHGAP24* is a negative regulator of Rho GTPases, particularly Rac1 and Cdc42. We did not observe differences in the expression levels of either Cdc42 or Rac1 in embryos of 1 and 2dpf. This does not exclude a role for *ARHGAP24* as regulator of Cdc42 and Rac1 later in life. Although our observations point to a function for *Arhgap24* during heart development, further studies are needed, including electrocardiograms of morphants and wild type embryos, to confirm electrocardiographic changes in the morphants.

Also our sequence analyses supports the hypothesis that *ARHGAP24* is functionally involved in cardiac function. Exon sequencing analysis in the ERF population revealed a nominally

significant association between rs35521695 to QT and rs144785317 for QT and QRS. Both variants are related to altered splice sites and are probably damaging in the case of rs144785317. The rs35521695 variant at codon 414 is predicted to be possibly damaging. These two mutations in ERF are related to a loss of function of the protein. The third and most strongly associated variant is rs61758879, which is associated to 12LS ($p=6.2 \times 10^{-5}$. This variant involved a missense mutation (R → L). These findings extend the GWAS analysis that identified a common intronic variant.

Conclusions

Our experiments in zebrafish show that *Arhgap24* knock down affects early cardiac development and function. Histological evidence of dilated hearts confirms the presence of abnormalities in morphant hearts. Additional support is provided by the effects of three damaging *ARHGAP24* genetic variants on ECG in the ERF population. Our experimental studies in zebrafish and observational studies in humans suggest that the *ARHGAP24* gene is involved in normal cardiac development and function.

Conflicts of interest. None

Author Contributions.

Claudia Tamar Silva, Herma van der Linde, Lies-Anne Severijnen, and Rob Willemsen were involved in experimental part and analyses. Claudia Tamar Silva, Rob Willemsen, Aaron Isaacs, Cornelia van Duijn were involved in the design of the study and writing the paper. Jon A kors, Abas Dehghan were involved in ECG interpretation, measurement and analysis.

Acknowledgements. The authors are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, Jeannette Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection.

Funding.

The Erasmus Rucphen Family (ERF) study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP Grant No. 018947 (LSHG-CT-2006- 01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007–2013)/Grant Agreement HEALTH-F4- 2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (No. QLG2-CT-2002-01254). High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organization for Scientific Research and the Russian Foundation for Basic Research (NWO- RFBR 047.017.043), and Russian Federal Agency of Scientific Organizations projects VI.53.2.2 and 0324-2015-0003. Exome sequencing analysis in ERF was supported by the Netherlands Organization for the Health Research and Development grant for Project No. 91111025. This work was also supported by the Federal Agency of Scientific Organizations #0324-2015-0003 (IZ, AK and TA). The Rotterdam Study is funded by Erasmus Medical Center (MC) and Erasmus University, Rotterdam; Netherlands Organization for the Health Research and Development; the Research Institute for Diseases in the Elderly; the Ministry of Education, Culture and Science; the Ministry for Health, Welfare and Sports; the European Commission; and the Municipality of Rotterdam.

Disclosures

Ethical approval and consent to participate. The Medical Ethics Committee of the Erasmus University Medical Center approved the ERF study protocol and all participants, or their legal representatives, provided written informed consent. Approval MEC 213.57512002.

Consent for publication. Not applicable.

References

1. Holm, H., et al., *Several common variants modulate heart rate, PR interval and QRS duration.* Nat Genet, 2010. **42**(2): p. 117-22.
2. Newton-Cheh, C., et al., *Genome-wide association study of electrocardiographic and heart rate variability traits: the Framingham Heart Study.* BMC Med Genet, 2007. **8 Suppl 1**: p. S7.
3. Silva, C.T., et al., *Heritabilities, proportions of heritabilities explained by GWAS findings, and implications of cross-phenotype effects on PR interval.* Hum Genet, 2015.
4. Arking, D.E., et al., *Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization.* Nat Genet, 2014. **46**(8): p. 826-36.
5. Eijgelsheim, M., et al., *Identification of a common variant at the NOS1AP locus strongly associated to QT-interval duration.* Hum Mol Genet, 2009. **18**(2): p. 347-57.
6. Newton-Cheh, C., et al., *Common variants at ten loci influence QT interval duration in the QTGEN Study.* Nat Genet, 2009. **41**(4): p. 399-406.
7. Nolte, I.M., et al., *Common genetic variation near the phospholamban gene is associated with cardiac repolarisation: meta-analysis of three genome-wide association studies.* PLoS One, 2009. **4**(7): p. e6138.
8. Pfeufer, A., et al., *Genome-wide association study of PR interval.* Nat Genet, 2010. **42**(2): p. 153-9.
9. Pfeufer, A., et al., *Common variants at ten loci modulate the QT interval duration in the QTSCD Study.* Nat Genet, 2009. **41**(4): p. 407-14.
10. Sotoodehnia, N., et al., *Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction.* Nat Genet, 2010. **42**(12): p. 1068-76.
11. van der Harst, P., et al., *52 Genetic Loci Influencing Myocardial Mass.* J Am Coll Cardiol, 2016. **68**(13): p. 1435-48.
12. Verweij, N., et al., *Twenty-eight genetic loci associated with ST-T-wave amplitudes of the electrocardiogram.* Hum Mol Genet, 2016. **25**(10): p. 2093-2103.
13. Xu, G., et al., *ARHGAP24 inhibits cell cycle progression, induces apoptosis and suppresses invasion in renal cell carcinoma.* Oncotarget, 2016. **7**(32): p. 51829-51839.
14. Huang, C.J., et al., *Conditional expression of a myocardium-specific transgene in zebrafish transgenic lines.* Dev Dyn, 2005. **233**(4): p. 1294-303.
15. Ho, Y.L., et al., *In vivo assessment of cardiac morphology and function in heart-specific green fluorescent zebrafish.* J Formos Med Assoc, 2007. **106**(3): p. 181-6.
16. Kimmel, C.B., et al., *Stages of embryonic development of the zebrafish.* Dev Dyn, 1995. **203**(3): p. 253-310.
17. Untergasser, A., et al., *Primer3--new capabilities and interfaces.* Nucleic Acids Res, 2012. **40**(15): p. e115.
18. Koressaar, T. and M. Remm, *Enhancements and modifications of primer design program Primer3.* Bioinformatics, 2007. **23**(10): p. 1289-91.
19. Robu, M.E., et al., *p53 activation by knockdown technologies.* PLoS Genet, 2007. **3**(5): p. e78.
20. Pardo, L.M., et al., *The effect of genetic drift in a young genetically isolated population.* Ann Hum Genet, 2005. **69**(Pt 3): p. 288-95.
21. van Bemmel, J.H., J.A. Kors, and G. van Herpen, *Methodology of the modular ECG analysis system MEANS.* Methods Inf Med, 1990. **29**(4): p. 346-53.

22. de Bruyne, M.C., et al., *Diagnostic interpretation of electrocardiograms in population-based research: computer program research physicians, or cardiologists?* J Clin Epidemiol, 1997. **50**(8): p. 947-52.
23. Willems, J.L., et al., *A reference data base for multilead electrocardiographic computer measurement programs.* J Am Coll Cardiol, 1987. **10**(6): p. 1313-21.
24. Casale, P.N., et al., *Electrocardiographic detection of left ventricular hypertrophy: development and prospective validation of improved criteria.* J Am Coll Cardiol, 1985. **6**(3): p. 572-80.
25. Siegel, R.J. and W.C. Roberts, *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, 1982. **103**(2): p. 210-21.
26. Sokolow, M. and T.P. Lyon, *The ventricular complex in right ventricular hypertrophy as obtained by unipolar precordial and limb leads.* Am Heart J, 1949. **38**(2): p. 273-94.
27. Bazett, H.C., *The time relations of the blood-pressure changes after excision of the adrenal glands, with some observations on blood volume changes.* J Physiol, 1920. **53**(5): p. 320-39.
28. Roguin, A., *Henry Cuthbert Bazett (1885-1950)--the man behind the QT interval correction formula.* Pacing Clin Electrophysiol, 2011. **34**(3): p. 384-8.
29. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.
30. Brouwer, R.W., et al., *NARWHAL, a primary analysis pipeline for NGS data.* Bioinformatics, 2012. **28**(2): p. 284-5.
31. Akilesh, S., et al., *Arhgap24 inactivates Rac1 in mouse podocytes, and a mutant form is associated with familial focal segmental glomerulosclerosis.* J Clin Invest, 2011. **121**(10): p. 4127-37.
32. Schwarz, J.M., et al., *MutationTaster evaluates disease-causing potential of sequence alterations.* Nat Methods, 2010. **7**(8): p. 575-6.
33. Roberts, P.J., et al., *Rho Family GTPase modification and dependence on CAAX motif-signaled posttranslational modification.* J Biol Chem, 2008. **283**(37): p. 25150-63.

Chapter 7


ATTCGGTACCTCATAACGGCTAGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTT
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGGTACCTA

Chapter 7

General discussion and Summary

Chapter 7.1

General Discussion

General discussion

The overarching aim of this thesis is to outline the genetic foundations shaping the heart rhythm evaluated as a complex phenotype. In this vein, delineating the phenotypic variance attributed to genetic effects (heritability), represents a fundamental task. To achieve this end, we employed modeling from different branches of the genetic epidemiology (defined as the medical science that outlines the causes of disease in aggregates of biological relatives) such as linkage analysis [1] (evaluation of traits co-segregation in pairs of relatives) and case-control studies (evaluation of linkage disequilibrium using genome-wide association studies -GWAS). Finally, we also designed functional studies using animal models (zebrafish), to evaluate the potential functional effect of those variants harbored in genes that resulted linked and/or associated to cardiophysiological processes and underpinning the genetic architecture of the heart rhythm and its pathological counterpart (conduction disorders).

Thereafter, we will discuss our most relevant findings, challenges, clinical implications, and future directions of the genetic epidemiological studies of the heart rhythm and conduction disorders.

How much of the heart rhythm phenotype variance in healthy individuals can be explained by genetic effects and apportioned to polymorphic markers?

This question was tackled by proposing the hypothesis that a significant part of the heart rhythm phenotype variance was explained by the effects of genetic variants harbored in genes implicated in cardiac electrophysiological pathways. To contrast this hypothesis, we evaluated the heritability underpinning the electrophysiology of the heart rhythm in a cohort of individuals ascertained from multigenerational and extended family belonging to a genetic isolate (the ERF cohort). This

extended cohort, that originated from 22 families from the 18th century in the southwest of Netherlands, exhibit extraordinary unique features that increase the power to detect the transmission of genetic effects cosegregating with conduction traits (ECG and LVH) and the exceptional possibility of detecting traits cosegregating (linkage) with genomic regions.

This constituted an original and unique approach since previous studies estimated heart rhythm parameters heritability mainly in twin or extended family studies (Table 1), and these studies did not evaluate cosegregation with genetic markers. Overwhelming evidence that accumulated from the last few decades has demonstrated that the use of extended and multigenerational studies ascertained from especial populations *i.e.,* genetic isolates and populations resulting from recent effects of admixture are extremely powerful to dissect genetics from environmental and to extract these components from random noise [2, 3]. We showed that there is a small proportion of heritability that can be explained and apportioned to SNPs: 4% for QT, 17% for QRS, 2% for PR and 4% for twelve-lead sum (12LS). Interestingly, we did not find any loci explaining heritability for the Cornell voltage (CV) product, and for the Sokolow-Lyon index (SL). Additionally, the inclusion of all ECG-associated SNPs further explained an additional proportion of the PR heritability (6%), suggesting that the presence of substantial cross traits effects may occur.

Study population	n	Heritabilities measures	Reference
Adult male twins	251 pairs	RR 77%	[4]
Female Twin	372 pairs	HR 55%, QT 60%, QTc 50%	[5]
Twin study	355 pairs	HR 54%, 34%	[6]
Family study	2909 individuals from 847 families	PR 34%, QRS 43%, QTc 40%, HR 34%	[7]
Cohort study	1962 cohort participants	QT 35%, 37% QT peak interval, 25% JT	[8]
Cohort study	4660 cohort participants	QT 41%, 40%RR	[9]
Twin study	446 monozygotic and 365 dizogotic twins	QT 67%, RR 55%, QTc 42%	[10]
Isolated population	1080 individuals	QT 31%	[11]
Isolated population	1064 individuals	PR 34%, PR segment 31%, Pwave 17%, QRS 3%	[12]

Table 1. Previous twin studies revealed different percentages of heritabilities.

The estimates of ECG traits heritability in our cohort contrast significantly from previously family or twin studies *e.g.* QT (31-67%), QRS (3-43%) and PR (around 34%) (Table 1). I think that these, apparently discordant results, can be explained by the different approach and populations used by the different studies, highest measurements were obtained through twin studies, it has been described that twin inflate heritability estimates, because the equal environment assumption [13-15] Since individuals of ERF cohort did not show any heart-related rare condition, they are

quite suitable to be evaluated by whole genome scanning data represented by common variants (which are expected to explain the phenotypic variance of universal traits like ECG and LVH in healthy individuals. An additional and important point is that results of heritability of ECG traits obtained in our study are lower than those reported by earlier studies. This could be explained by their heritability estimation, the fact that distant relatives were included, which consequently affected the likelihood of sharing the same environment. The latter is in complete contrast with studies involving twins and parent-offspring pairs where the heterogeneity of the environment is lower.

It is important to discuss that in 2012, when this project started, there were no other studies employing similar approaches as those reported in here. In fact, only in 2017, Nolte et al, described the heritability of ECG traits using classical twin modeling versus heritability estimation using a SNP panel [16]. This study estimated heritabilities of 55% for PR interval and 53% for QRS and QTc intervals using classical monozygotic twins heritability, estimation with SNP inclusion the heritability estimates were 26% for PR, 23% for QRS and 28% for QTc intervals, which corroborated those heritabilities obtained in twin modeling are inflated, and supporting the robust basis for future studies exploring genetic variants responsible of cardiac conduction traits. Contrary to our findings, and in general to other studies that show figures lower than 10%, these findings support the idea that common SNPs used in their study explain a big portion of the heritability, which implies that there are a number of SNPs that remain to be found in our cohort. Another potential explanation is that the panel of common SNPs used in the study of Nolte et al., are not representative of the genomic variation that characterizes the genetic isolate of our ERF study This issue was recently addressed by Speed et al [17] in which a more accurate model was derived empirically to describe how heritability varies with minor allele frequency, linkage disequilibrium, and genotype certainty, indicating that variation of gene frequencies throughout

populations might affect the heritabilities estimation. In this vein, a study in a plant model suggest that this strong dependency of allelic average effects on genetic background implies that epistasis is a major determinant of the additive genetic variance, and thus, the population's ability to respond to natural selection, a factor that has not been considered in the equation [18]. On the other side, it is important to mention that almost every method of heritability estimation relies on the use of a set of SNPs acting in an additive fashion while Nolte et al use a complete array, I used a subset of SNPs previously associated with ECG traits, for PR i.e. we included 9 SNPs, there are new uncovered genes that we did not included in this study (table 2). However, there are many examples of genetic effects acting in a non-linear way and shaping epistatic interactions that does not follow a Gaussian distribution [14]. Finally, the same argument of significant effects of environmental factors underpinning differential risk in distant biological relatives is valid to explain discordant estimates of heritabilities.

Trait	SNP	Gen	Region	Ref
QT	rs11153730	*SLC35F1, PLN*	6 q22.31	[19, 20]
	rs11756438	*c6orf204, SLC35F1, PLN, ASF1A*	6q22.31	[21]
	rs12029454	*NOS1AP*	1q23.3	
	rs2478333	*RN7SL700P, LOC105370196*	13q14.2	[22]
	rs846111	*RNF207*	1p36.31	[19]
	rs12997023	*SLC8A1*	2p22.1	
	rs6793245	*SCN5A, SCN10A*	3p22.2	
	rs3807375	*KCNH2*	7q36.1	
	rs7122937	*KCNQ1*	11p15.5	
	rs735951	*LITAF*	16p13.13	
	rs246196	*CNOT1*	16q21	
	rs105253	*LIG3*	17q12	

Trait	SNP	Gen	Region	Ref
	6			
	rs1396515	*KCNJ2*	17q24.3	
	rs10919070	*ATP1B1*	1q24.2	
	rs2298632	*TCEA3*	1p36.12	
	rs938291	*RPL5P7, SP3*	2q31.1	
	rs7561149	*TTN, CCDC141*	2q31.2	
	rs295140	*SPATS2L*	2q33.1	
	rs17784882	*C3ORF75*	3p21.31	
	rs2363719	*SLC4A4*	4q13.3	
	rs3857067	*SMARCAD1*	4q22.2	
	rs10040989	*CDC23, GFRA3*	5q31.2	
	rs7765828	*GMPR*	6p22.3	
	rs9920	*CAV1*	7q31.2	
	rs16936870	*NCOA2*	8q13.3	
	rs11779860	*LAPTM4B*	8q22.1	
	rs1961102	*AZIN1*	8q22.3	
	rs2485376	*GBF1*	10q24.32	
	rs174583	*FADS1, FADS2, FADS3*	11q12.2	
	rs3026445	*ATP2A2*	12q24.11	
	rs728926	*KLF12*	13q22.1	
	rs2273905	*ANKRD9*	14q32.31	
	rs3105593	*USP50, TRPM7*	15q21.2	
	rs1296720	*CREBBP*	16p13.3	
	rs246185	*MKL2*	16p13.12	
	rs9892651	*PRKCA*	17q24.2	
	rs727957	*KCNE1*	21q22.12	[

Trait	SNP	Gen	Region	Ref
				23]
QRS /SL/CV	rs1733724	*DKK1*	10 q21.1	[24]
	rs1886512	*KLF12*	13 q22.1	
	rs17608766	*GOSR2*	17 q21.32	
	rs9912468	*PRKCA*	17 q24.2	
	rs11848785	*SIPA1L1*	14 q24.2	
	rs17391905	*CDKN2C*	1 p32.3	
	rs9851724	*SCN5A, SCN10A*	3 p22.2	
	rs2242285	*LRIG1, SLC25A26*	3 p14.1	
	rs9470361	*CDKN1A*	6 p21.2	
	rs9436640	*NFIA*	1 p31.3	
	rs4074536	*CASQ2*	1 p13.1	
	rs7562790	*CDKN1A*	6 p21.2	
	rs17020136	*HEATR5B*	2 p22.2	
	rs4687718	*TKT*	3 p21.1	
	rs13165478	*HAND1*	5 q33.2	
	rs1362212	*TBX20*	7 p14.2	
	rs10850409	*MED13, TBX3*	12 q24.21	
	rs7784776	*IGFBP3*	7 p12.3	
	rs7342028	*VTI1A*	10 q25.2	
	rs3825214	*TBX5*	12 q24.21	[23]
PR	rs11708996	*SCN5A*	3p22.2	[25]
	rs11897119	*MEIS1*	2p14	

Trait	SNP	Gen	Region	Ref
	rs1896312	*TBX5/TBX3*	12q24.21	
	rs251253	*NKX2-5*	5q35.1	
	rs3807989	*CAV1/CAV2*	7q31.2	
	rs4944092	*WNT11*	11q13.5	
	rs6800541	*SCN10A*	3p22.2	
	Rs11047543	*SOX5*	12p12.1	
	rs882300	*ARHGAP24*	4q21.23	
12LS	rs2290893	*PTGES3*	12q13.3	[26]
	rs2292462	*NMB*	15q25.2	
	rs1733724	*DKK1*	10q21.1	[24]
	rs7342028	*VTI1A*	10 q25.2	
	rs3825214	*TBX5*	12 q24.21	
	rs1886512	*KLF12*	13 q22.1	
	rs17608766	*GOSR2*	17 q21.32	
	rs17391905	*CDKN2C*	1 p32.3	
	rs9851724	*SCN5A, SCN10A*	3 p22.2	
	rs2242285	*LRIG1, SLC25A26*	3 p14.1	
	rs9470361	*CDKN1A*	6 p21.2	
	rs11848785	*SIPA1L1*	14 q24.2	
	rs9912468	*PRKCA*	17 q24.2	
	rs9436640	*NFIA*	1 p31.3	
	rs4074536	*CASQ2*	1 p13.1	
	rs7562790	*CDKN1A*	6 p21.2	
	rs170201	*HEATR5B*	2 p22.2	

Trait	SNP	Gen	Region	Ref
	36			
	rs4687718	*TKT*	3 p21.1	
	rs13165478	*HAND1*	5 q33.2	
	rs1362212	*TBX20*	7 p14.2	
	rs10850409	*MED13, TBX3*	12 q24.21	
	rs7784776	*IGFBP3*	7 p12.3	

Table 2. SNPs used for heritability analysis

Following the publication of our manuscript on heritabilities, we participated in several studies where many other new genomic variants were reported associated to ECG traits [27-31] (Table 3, to be constructed). For example, Bihlmeyer et al used exomeChip analysis (in which functional variants are overrepresented when compared to neutral ones), and described 10 loci modulating QT and JT intervals duration [29]. Six of these loci were associated to QT: *PM20D1, SLC4A3, CASR, NRAP, ZNF37A* and *GOS2* and four with JT interval: *SENP2, SLC12A7, CNKN1A* and *NACA* (Table 3, to be constructed). Moreover, their analyses showed that some of those genes are involved in the generation of physical force of contraction inside the cardiomyocytes, and also in electrical conduction. In another study, J. van Setten et al showed seven new loci associated. Three of those loci are associated to PR (*KCDN3, NR3C1* and *PLN*), other three associated with QT (*KCNE1*, *SGIP1,* and *NFKB1*), and 1 associated with QRS (*ATP2A2*) [28] (Table 3, to be constructed). Verweij et al reported 28 genome-wide significant loci associated to ST-T wave amplitudes [30] (Table 3) and van den Berg et al described 5 novel heart rate loci *KIAA1755*, *C10orf71, DALDR3, TESK2* and *MAPK8* (Table 3, to be constructed). In chapter 5 (van der Harst et al), we described 52 SNPs related to genes influencing myocardial mass [31] (Table 3, to be constructed).

Trait	SNP	Chr	Closest Gene	Ref
PR	rs4648819	1	*SKI*	Primma submitted paper to nature communication. Van Settern et al.
	rs7538988	1	*EPS15*	
	rs12127701	1	*MYBPHL*	
	rs11264339	1	*KRTCAP2*	
	rs397637	1	*OBSCN*	
	rs3856447	2	*ID2*	
	rs2732860	2	*TMEM182*	
	rs13018106	2	*FIGN*	
	rs922984	2	*TTN*	
	rs9826413	3	*EOMES*	
	rs900669	3	*FRMD4B*	
	rs13087058	3	*PDZRN3*	
	rs16858828	3	*PHLDB2*	
	rs6441111	3	*CCNL1*	
	rs7638853	3	*SENP2*	
	rs17446418	4	*CAMK2D*	
	rs3733409	4	*FAT1*	
	rs7729395	5	*PAM*	
	rs11763856	7	*TBX20 / HERPUD2*	
	rs2129561	7	*MKLN1*	
	rs881301	8	*FGFR1*	
	rs12678719	8	*ZFPM2*	
	rs12359272	10	*ALDH18A1 SORBS1*	
	rs12257568	10	*SH3PXD2A OBFC1*	
	rs1372797	11	*NAV2*	
	rs11067773	12	*MED13L*	
	rs718426	13	*EFHA1*	
	rs2585897	13	*XPO4*	
	rs9590974	13	*LRCH1*	
	rs11465506	14	*IL25 / MYH6*	

Trait	SNP	Chr	Closest Gene	Ref
	rs4901308	14	*FERMT2*	
	rs17767398	14	*SNORD56B*	
	rs904974	15	*TLE3*	
	rs1984481	17	*MYOCD*	
	rs75013985	1	*KCND3*	van Setten J. EJHG submitted paper
	rs17287745	5	*NR3C1 / ARHGAP26*	
	rs74640693	6	*PLN / SLC35F1*	
SL	41767282	6p21.1	*TFEB*	[31]
	124735610	8q24.13	*KLHL38*	
	61748819	17q24.2	*PRKCA*	
Leadsum	rs2849028	1p36.12	*ZNF436*	
	rs2274317	1q22	*MEF2D*	
	rs12036340	1q23.3	*OLFML2B*	
	rs4288653	1q32.1	*PLEKHA6*	
	rs3816849	2q31.2	*TTN*	
	rs13314892	3p14.1	*MITF*	
	rs10937226	3q27.2	*SENP2*	
	rs10105974	8q24.13	*LOC105375743*	
	rs12414364	10q21.3	*CTNNA3*	
	rs10509289	10q21.3	*CTNNA3*	
	rs7099599	10q22.2	*BMS1P4*	
	rs2926743	12q13.3	*NACA*	
	rs7132327	12q24.21	*TBX3*	
	rs1408224	13q14.13	*LRCH1*	

Trait	SNP	Chr	Closest Gene	Ref
	rs7183401	15q25.3	*ALPK3*	
	rs8038015	15q26.3	*IGF1R*	
	rs6565060	16q23.3	*CDH13*	
	rs7211246	17q11.2	*NSRP1*	
	rs242562	17q21.31	*MAPT*	
	rs617759	18q12.2	*MAPRE2*	
	rs7283707	21q21.1	*USP25*	
Cornell	rs10920184	1q32.1	*TNNT2*	
	rs6710065	2p23.3	*DPYSL5*	
	rs13185595	5q33.2	*HAND1*	
	rs1733724	10q21.1	*DKK1*	
	rs2269434	11p11.2	*MYBPC3*	
	rs736825	12q13.13	*HOXC6*	
	rs3929778	20p12.3	*BMP2*	
	rs2025096	20q11.22	*MYH7B*	
QRS	rs17391905	1p32.3	*CDKN2C*	
	rs2207790	1p31.3	*NFIA*	
	rs12039739	1p13.1	*CASQ2*	
	rs3770770	2p22.2	*STRN*	
	rs6801957	3p22.2	*SCN10A*	
	rs4687718	3p21.1	*TKT*	
	rs2242285	3p14.1	*LRIG1*	
	rs1344852	4p1	*SLIT2*	

Trait	SNP	Chr	Closest Gene	Ref
		5.31		
	rs1321311	6p21.2	*CDKN1A*	
	rs11153730	6q22.31	*SLC35F1*	
	rs1419856	7p14.2	*TBX20*	
	rs6968945	7p12.3	*TNS3*	
	rs11773845	7q31.2	*CAV1*	
	rs7918405	10q25.2	*VTI1A*	
	rs174577	11q12.2	*FADS2*	
	rs728926	13q22.1	*KLF12*	
	rs12880291	14q24.2	*SIPA1L1*	
	rs879568	18q12.2	*FHOD3*	
	rs10853525	18q12.3	*SETBP1*	
	rs28637922	12	*ATP2A2 / ANAPC7*	van Setten J. EJHG
QT	rs6588213	1	*SGIP1*	
	rs11097788	4	*NFKB1*	van Setten J. EJHG
	rs1805128	21	*KCNE1*	
HR	rs17853159	1	*TESK2*	[32]
	rs3087866	3	*DALRD3*	
	rs1635852	7	*JAZF1*	
	rs10857472	10	*C10orf71*	
	rs3793706	10	*SEC31B*	
	rs260505	1p36.33	*SKIn*	
	rs2072944	1p36.12	*LUZP1, KDM1A, WNT4*	
	rs2298632	1p36.12	*TCEA3*	
	rs2207792	1p31.3	*NFIA*	

Trait	SNP	Chr	Closest Gene	Ref
	rs12145374	1p13.2	*KCND3, FAM212B*	
	rs10908505	1q22	*MEF2D*	
	rs12567315	1q23.3	*NOS1AP*	
	rs545833	1q24.2	*DPT*	
	rs7576036	2p15	*XPO1*	
	rs1866666	2q33.1	*PLCL1,, MARS2, RFTN2, MOB4*	
	rs4684185	3p25.1	*LSM3, TMEM43*	
	rs7638909	3p22.2	*SCN5A, ACVR2B*	
	rs6801957	3p22.2	*SCN10A, SCN5A, ACVR2B*	
	rs7756236	6p21.31	*CDKN1A*	
	rs210966	6q22.2	*ROS1, VGLL2*	
	rs9388451	6q22.31	*HEY2*	
	rs1458942	8p23.1	*TNKS, SGK223, XKR6, PPP1R3B*	
	rs7011924	8p23.1	*DEFB136, NEIL2*	
	rs2286582	12p13.32	*GALNT8*	
	rs10842350	12p12.1	*SOX5*	
	rs10850409	12q24.21	*TBX3*	
	rs728926	13q22.1	*KLF12*	
	rs7174918	15q26.3	*IGF1R*	
	rs7192150	16p13.3	*LMF1, SOX8*	
	rs735951	16p13.13	*LITAF*	
	rs4784939	16q21	*GINS3*	
	rs8057901	16q	*NDRG4*	

Trait	SNP	Chr	Closest Gene	Ref
		21		
	rs8083566	18q12.1	*CDH2*	
	rs11673003	19q13.33	*KCNA7, NTF4, GYS1, HRC*	
	rs6087666	20q11.22	*TRPC4AP, EDEM2, MYH7B, NCOA*	
	rs6088738	20q11.22	*EDEM2, PROCR, MYH7B, NCOA*	
	rs11907908	20q13.13	*ZNFX1 (−AS1), STAU1*	
	rs6019750	20q13.13	*KCNB1, STAU1*	

Table 3. New uncovered genes

How do common variants associated to myocardial mass influence heart rhythm?

There are common variants that were uncovered through association studies, like GWAs, which could expand our knowledge about the genetic component of ECG intervals. The cardiac ventricle muscular contraction, caused by cardiac repolarization, is represented in the EKG by the QRS interval. In chapter 5, I showed 52 loci associated to myocardial mass, which indirectly is a major lead to understand those genetic factors influencing the QRS complex. Further, it is valid to extrapolate that the dissection of these genes, related with heart function, could be useful to predict, preclinical, clinical, follow up and natural history of the cardiovascular disease.

We revealed genome-wide significant loci associated with cardiac repolarization improving the knowledge of ECG architecture [8]. Thus far, we uncover 28 loci associated to the ST-T-wave interval. Following this findings, in an additional paper, we studied heart rate and performed a meta-analysis of 104,452 individuals of European-ancestry using a exome chip and validated our results, by replicating them in a set of independent samples. This meta-analysis revealed 5 new

heart rate loci *(TESK2, DALRD3, JAZF1, Z10orf71, and SEC31B).* Four of these loci were validated in our study and also recently published in the UK biobank study *(RNF207, SCN10A, 5p13.3 and KDELR3).* There was another locus, reported with a new secondary signal at previously reported *KIAA1755* locus. We did not find rare SNV associations with HR, suggesting that we need larger sample sizes to reach enough power to detect rare variants [52]. Furthermore, we uncovered additional loci associated to QT, PR and QRS traits (Table 3). We revealed new associated genes related to ECG traits and consequently related to heart rhythm function. Thus, we included these genes together with the previously described ones to perform pathway analysis and ontogenetic enrichment under the hypothesis that these loci must be overrepresented in biochemical, cell and genetic processes configuring networks. We found that these genes are significantly overrepresented in pathways/processes involved in cardiovascular pathology such as: vascular fistula, cardiac arrhythmias, cardiomyopathies, cardiovascular disease, heart disease, and cardiovascular abnormalities. Gene ontology revealed that genes like *MEF2D, TBX3, MEF2, IGF1R, LUZP1, HEY2,* and *NDRG4* are involved in regulation of heart rate and heart development, giving us an additional support of the relevance of our findings.

Are rare variants related to heart rhythm?

The occurrence of rare mutations, and the large amount of heritability that is not explained by common variation, motivates the need for both mutation screening and alternative approaches to genome wide association studies that focuses on common variants. Some of the best approaches to elucidate these rare mutations are whole exome or genome capture, and next generation sequencing. Given that the power of this type of studies is limited because of the rareness of the phenotypes and of the genotypes, we followed a classical approach applying genetic linkage analysis to families segregating specific ECG phenotypes as outlined by Amin et al [33]. These approaches focus in the Mendelian effect of these rare variants, a phenomenon that might be

unique and crucial to define real targets for genetic-engineering interventions and for the development of new medications.

In this vein, using linkage analyses of classical ECG outcomes (chapter 3), we found suggestive peaks of linkage underpinning the QT interval (1q24, LOD = 2.63; 2q34, LOD = 2.05), QRS interval (1p35, LOD = 2.52) and PR interval (9p22, LOD = 2.20; 14q11, LOD = 2.29). [34, 35]. Fine-mapping of these suggestive regions using exome sequence and microarray high resolution genotyping identified a rare variant (minor allele frequency = 0.0186) harbored in *FCRL2* locus linked (LOD = 2.63) and associated to QT (***P***=0.024) and explaining 0.83% of the variance of the QT in ERF. *In silico* bioinformatic analyses showed that levels of expression of *FCRL2* are associated to *ARHGAP24* and *SETBP1* expression, two genes previously identified in GWAS associated to PR and QRS intervals (Chapter 4) [23, 25, 31, 36].

In order to identify new rare variants implicated in left ventricular hypertrophy (LVH) as defined by ECG parameters, we combined bioinformatics analyses with our association and linkage results. We performed principal components (PCs) analyses of the LVH traits that capture such effects. The linkage study of LVH proxy PCs measurements identified one significant locus (15q11.2-LOD=3.01) and 12 suggestive regions (1p34-LOD=2.4, 4q31-LOD=2.14, 5p14-LOD=2.18, 6q15-LOD=2.17, 9p21-LOD=2.35, 11q13.4-LOD=2.01, 15q25-LOD=1.92, 20p12.1=LOD=2.634 for SL and 2.83 for PC1, 20p11.23-LOD=2.12, 22q13-LOD=1.99).

Rare variant analyses in these regions uncovered a missense coding variation harbored in *MAP3K11* gene. This *MAP3K11* variant substantially decreased the LOD score for this PC1 linkage peak. Conditional analysis revealed a drop from 2.8 to 0.8 for *MAP3K11* suggesting that this variant explains a large proportion of this chromosome linkage signal. The Principal component 1 is mainly determined by the ECG parameter 12LS and SL. The *MAPK11* variant also showed evidence of association with the two traits: the ***P***-value for 12 LS was 3.0×10^{-4} and 1.2×10^{-3} for

SL. *MAPK11* is related with JNK pathway, which is a pro-apoptotic kinase that plays important roles in the induction of cardiomyocyte apoptosis in various pathologies including LVH.

Are common genes under linkage peaks?

Of those previously reported loci associated with the QRS interval, *CASQ2, CDKN2, NFIA,* and *TRIM63,* are under QRS linkage peak in 1p35 [31, 36]. Under this same linkage region, we found *KCND3* and *MFSD2A* associated with PR and *SGIP1-TCTEX1D1,* reported associated with the QT interval [27, 28, 37, 38]. Additionally, there is a gene previously associated with resting heart rate*, RNF220* [39].

Under the QT linkage peaks in 1q24 and 2q34, there are several genes previously associated with QT*: NOS1AP, OLFML2B, SH2D1B, DPT, SLC19A2, ATP1B1, OSBPL6, TTN* and *CCDC141* [12, 19, 40-42]. Recently, it was reported a new gene associated to the QT interval harbored in 2q35: *SCL4A*3 [29]. Other genes harbored in these two regions are: *MEF2D, HMCN1, WIPF1, CCDC141, PDE11A, SPEG* and *VWC2L*, which are associated to QRS, heart rate and the PR interval [19, 31, 38, 39]. Finally, inside the PR linkage peak, harbored in 14q11 are three previously reported genes: *MYH6, NUBPL* and *ARHGEF40* [23, 39, 43] (Table4).

As a whole these results show that our described linkage regions contained ECG associated genes, in a fashion that significantly and conspicuously differs from randomness, supporting these areas as regions that contain candidate/causal genomic variants.

Gene	Associated trait	Chromosome position	Linkage trait	Reference
CASQ2	QRS duration	1p13.1	QRS	[31]
KCND3	PR	1p13.2	QRS	[28]
NFIA	QRS duration	1p31.3	QRS	[31, 36]
SGIP1-TCTEX1D1	QT	1p31.3	QRS	[28]
CDKN2C	QRS	1p32.3	QRS	[36]
RNF220	Resting heart rate	1p34.1	QRS	[39]
MFSD2A	PR interval	1p34.2	QRS	[38]
TRIM63	QRS complex (12-leadsum)	1p36.11	QRS	[31]
TCEA3	QT interval	1p36.12	QRS	[19]
MEF2D	Resting heart rate	1q22	QT	[39]
NOS1AP	QT	1q23.3	QT	[19]
OLFML2B	QT interval	1q23.3	QT	[19]
SH2D1B	QT interval	1q23.3	QT	[19]
ATP1B1	QT	1q24.2	QT	[19]
DPT	QT interval	1q24.2	QT	[19]
SLC19A2	QT interval	1q24.2	QT	[19]
HMCN1	QRS duration	1q31.1	QT	[38]
WIPF1	QRS complex (12-leadsum)	2q31.1	QT	[31]
CCDC141	Heart rate	2q31.2	QT	[41]
CCDC141	QT	2q31.2	QT	[19]
OSBPL6	QT interval	2q31.2	QT	[44]
PD	Heart rate	2q31.	QT	[41]

E11A		2			
TTN	QT	2	2q31.2	QT	[19]
VWC2L	PR interval		2q34	QT	[38]
SPEG	Resting heart rate		2q35	QT	[39]
SLC4A3	QT		2q35	QT	[29]
ARHGEF40	Resting heart rate	.2	14q11	PR	[39]
MYH6	Electrocardiographic traits	.2	14q11	PR	[23]

Table 4. Associated genes inside linkage ECG regions

I decided to search for these 52 SNPs related to genes influencing myocardial mass in the linkage regions described in chapter 4 and related to LVH proxy measurements. In Table 5 and Figure 2, I show these intersected SNPs. We found that rs17391905, associated to the QRS interval, is inside the CV linkage region highlighted by the SNP rs6619. The rs17391905 variant is located in the neighbourhood of the *CDKN2C* gene. The CV linkage peak highlighted by the rs14442470 SNP contains the rs13185595 SNP that is anchored upstream of the *HAND1* gene that is associated with CV. Other linkage region, a PC1 with the highest LOD score represented by variation at the rs1530354 SNP, contains the rs2269434 SNP, an intronic variant inside the *MYBPC3*, which is associated to CV. Genomic variation harboured in the *MYBPC3* gene is associated with a causal relationship to to cardiomyopathy. Inside the same CV region is harboured the rs174577 SNP, an intronic variant inside the *FADS2* gene that turns out to be associated to the QRS interval duration. The SL peak, highlighted by the rs290370, contains the rs8038015 SNP, an intronic variant inside the *IGF1R* gene which is associated to 12LS. Finally, the peak highlighted by the rs466243 SNP contains the rs2025096 that is anchored in the neighbourhood of the *MYH7B* gene. Mutations in *MYH7B* have been recently linked to left ventricular non-compaction cardiomyopathy [45].

SNP	Linkage region trait	SNP inside linkage region	Trait	Closer gene
rs6619	CV	rs17391905	QRS	*CDKN2C*
rs14442470	12LS	rs13185595	CV	*HAND1*
rs1530354	PC1	rs2269434	CV	*MYBPC3*
		rs174577	QRS	*FADS2*
rs290370	SL	rs8038015	12LS	*IGF1R*
rs466243	12LS	rs2025096	CV	*MYH7B*

Table 5. Intersectional SNPs among 52 loci influencing myocardial mass and linkage regions associated to LVH.

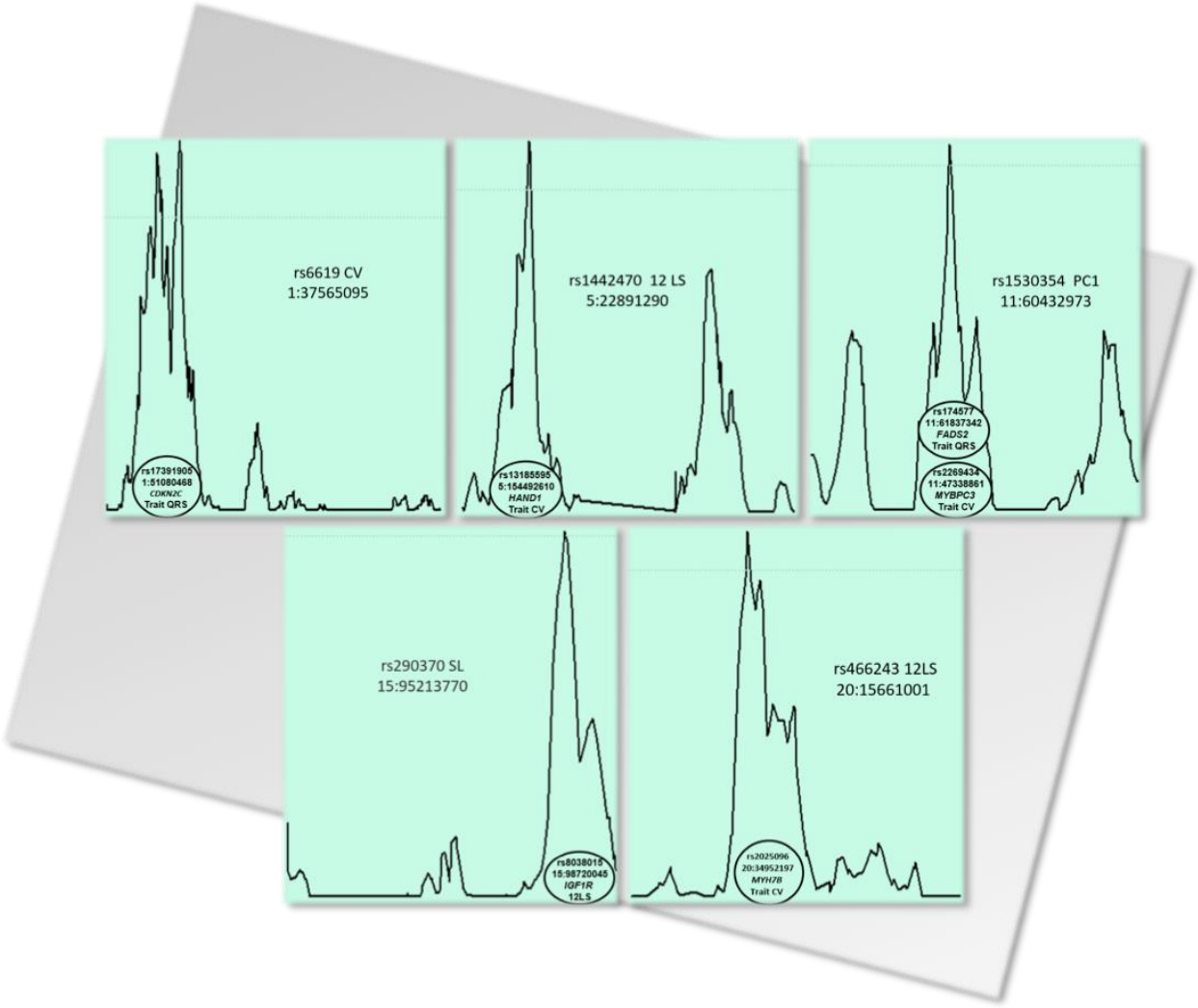


Figure 2. LVH linkage regions overlapping with QRS associated SNPs

In this figure are shown five of our thirteen LVH linkage regions. In circles six SNPs of our 52 loci influencing myocardial mass inside previous linkage reported regions.

Deviations of the ST-T wave amplitude can be suggestive to different heart abnormalities. In two additional studies, we revealed 28 genome-wide significant loci explaining an important phenotypic variance of the ST-T wave amplitudes. ST-segment and the adjacent T wave revealing that quantitative endophenotypes underpinning cardiac repolarization might be related to repolarization abnormalities. The KCND3 gene gives the strongest signal of association. The KCND3 gene encodes Kv4.3 a member of voltage-gated potassium channels and it, has been related to atrial fibrillation, heart failure and P-wave duration [27, 28, 37].

As a whole, these findings reported by other studies are consistent with those evidently remarked under our linkage peaks and they must contain true causal variants underpinning cardiac traits. Future analysis of extended families from different populations will help to dissect the best these variants/genes.

How animal models can be used to define the functional role of associated variants?

In chapter 5, we used drosophila and mouse models in an attempt of understanding the role of four genes *i.e CG4743/SLC25A26, Fhos/FHOD3, Cka/STRN, and NACα/NACA* that were associated with QRS and leadsum Using RNAi we knocked down the function of these genes, specifically within the heart of Drosophila, and found the development of severe cardiac malformations, which confirm the critical role of these genes and eventually point out to the fact that variants highlighted by the positive findings of our GWAS studies play a major role in causing EKG traits(table 6).

Gen	Phenotype in D.melanogaster
Previously described	
Mhc/MYH7B[46]	
Slit/SLIT2[47]	
EcR/NR1H[48]	Cardiac abnormalities
Hand/HAND1[49]	Cardiac Genesis
TTN[50]	Human cardiomyopathy
Functional analysis chapter 5	
Hand/HAND1	Reduced cardiac heart rate without heart abnormalities
Cka/STRN	Reduced cardiac heart rate with reduction in diastolic diameters and contractility
NACalfa/NACA	Complete loss of cardiac tissue beginning at eclosion
CG4743/SLC25A26	
Fhos/FHOD3	Without cardiac phenotype

Table 6. *D. melanogaster* functional analysis

In chapter 6, we describe the effect of one novel PR intronic related locus, *ARHGAP24*, performing a knockdown strategy in zebrafish as a model. *ARHGAP24* is one of the genes associated to the PR interval ($2.5x10^{-17}$) in genome wide association studies (Ref). The gene is a negative regulator of Rho GTPases implicated in chromatin remodeling, cell polarity, and cell migration. The role of this gene in the heart function is unknown. We showed that the knockdown of *arhgap24* in zebrfafish morphant embryos developed heart abnormalities when compared to control zebrafish and suggested that this gene is a major player during cardiac development. Also, in chapter 6, we describe a missense mutation (P → A) at codon 417 *ARHGAP24* that is associated to QT interval with nominal significance. A second variant at codon 67 (G→E) is marginally

associated to QT and QRS, and a third to LVH. Even though more studies are granted, these findings might suggest that there are cross trait effects in *ARHGAP24* (pleiotropy).

Functional analysis are one of the tools to demonstrate the function of a gene, with these analysis we showed the function of associated genes, probing their relation with heart function and heart development, in the future these associated genes could be helpful in prognosis of heart abnormalities.

How epigenetic explains heart rhythm variability?

Other approximation to establish genes function are in-silico analysis to know associated genes relation with epigenetic important regions. Epigenetics is a field related to gene expression, control, and modifications. Epigenetic mechanisms include histone modifications, DNA methylation, and RNA interactions. Cis-regulatory elements such as promoters or enhancers susceptible to epigenetic modifications are marked by DNase I hypersensitive sites (DHSs). Among our findings described in **Chapter 5**, we found that 42 of 52 sentinel SNPs were in DHSs. Additionally, we found 22 of 52 in DHSs in human fetal heart tissue, 11 of them are related to transcription factor recognition sites. These findings are important, because polymorphism in these sites could modify gene expression of genes related to heart development and function.

Furthermore, we also found that some of those genes play important roles as transcription factors (for instance *HEY2, MEF2D, SOX5* and *SOX8* of SOX family) binding to active enhancers and promoters.

Clinical implications and further research

As described previously, cardiac conduction abnormalities lead to various conditions, including sudden cardiac death (SDC), atrial fibrillation (AF), ventricular hypertrophy, and sick sinus syndrome, among others. SCD is estimated to occur in between 50-100 individuals per 100,000 per annum in the U.S. and the European populations [51], while the prevalence of AF in the European

Union is expected to double to 17.6 million cases per year, by 2060 [52]. These illnesses, therefore, impose large (and increasing) burdens on their societies. Genetics has been essential in advancing our knowledge of cardiac conduction disorders over the last two decades. First, these studies helped to clarify the physiological underpinnings of conduction and identified new pathways. Thus the results of genomic research provides new target genes and proteins for pharmaceutical research, and potentially useful for prediction of genetically transmitted cardiovascular disease.

Ultimately, increasing our knowledge of genetic variants influencing conduction disorders will improve molecular diagnosis and clinical risk prediction. The translation of these findings to the clinical setting is expected to occur soon. This thesis described that the GWAS loci identified up-to-date explain less than 20% of the heritability in the various ECG parameters. Yet, new pathways have been uncovered and there is no evidence that GWAS has reached its limits. By increasing sample size and marker density in ECG research, new genes have been identified. Particularly, for LVH related parameters a very small percentage of the heritability is explained. Indeed, the sample size of studies of LVH related traits have been small.

This thesis also shows that there may be new rare variants (minor allele frequency <0.05) involved in ECG outcomes that can be identified in family-based studies such as Erasmus Rucphen Family using a combined linkage and association approach. I expect these results would have clinical utility in both the short and long-term future. Although genetic testing does not currently perform well for risk stratification, as the number of known variants increases, genetic testing will enhance our ability to discriminate those at higher risk for conduction disorders. The incorporation of rare(r) variants should dramatically improve the utility of risk prediction profiles in specific families in which these variants segregate. This will open opportunities for improved personalized medicine in which the preventive strategies are tailored towards (rare) family specific causes of disease. One avenue of prevention may be cascade screening in families, as is at present

conducted for familial forms of dyslipidemia (references). In cascade screening all relative of carriers of a mutation are invited for genetic screening and tailored personalized prevention is offered to carriers with the family. Such families may be extended to 5-10 generations and involve hundreds of relatives.

Another avenue that may improve risk prediction is to model gene-interactions. We have not addressed this issue in this thesis. Gene interactions may explain part of the missing heritability. Studies of gene interactions have been hampered by low statistical power. This concerns both gene-gene as well as gen-environment interactions. Part of the problem is that effects are small for variants identified by GWAS, making it difficult to discriminate one small effect from another smaller effect. A new and more powerful avenue for interaction studies may be the use of risk scores, in which the effects of multiple genes within and over different biological pathways are captured. An interesting question to address will be whether environmental risk factors, such as smoking, could interact with the genes representing a single pathway or rather with a general risk score representing all pathways. In the latter case, the interaction is more likely to occur downstream from the disease pathway.

One major problem in complex genetic research is to determine which variants are causally related to the disease. In this thesis we used a functional approach implemented in an animal models, the zebra fish, mouse and drosophila. Although this is a straightforward experimental model, there is the need to increase the throughput of these experiments to speed up translational research.

Finally, we can conclude that in this thesis we get a new approximation for heritability analysis, uncovering the proportion of variability of ECG traits due to genes. Although we do not find significant genes under linkage peaks, we stablished to new candidate genes for ECG and LVH traits: *FCRL2 and MAP3K11*, it is necessary to perform new studies to determine the

relation of these two genes with heart rhythm, which could be useful for prognosis and risk determination. Common variants uncover by GWAs give us clues about candidate genes, functional analysis support our findings, we can conclude that our uncovered genes has strong evidence of association and are involved in heart rhythm, future studies has to be conducted in other populations.

References

1. Morton, N.E., *Outline of Genetic Epidemiology*. 1982: Karger.
2. Palacio, J.D., et al., *Attention-deficit/hyperactivity disorder and comorbidities in 18 Paisa Colombian multigenerational families.* J Am Acad Child Adolesc Psychiatry, 2004. **43**(12): p. 1506-15.
3. Wallis, D., et al., *Polymorphisms in the neural nicotinic acetylcholine receptor alpha4 subunit (CHRNA4) are associated with ADHD in a genetic isolate.* Atten Defic Hyperact Disord, 2009. **1**(1): p. 19-24.
4. Russell, M.W., et al., *Heritability of ECG measurements in adult male twins.* J Electrocardiol, 1998. **30 Suppl**: p. 64-8.
5. Dalageorgou, C., et al., *Heritability of QT interval: how much is explained by genes for resting heart rate?* J Cardiovasc Electrophysiol, 2008. **19**(4): p. 386-91.
6. Havlik, R.J., et al., *Variability of heart rate, P-R, QRS and Q-T durations in twins.* J Electrocardiol, 1980. **13**(1): p. 45-8.
7. Li, J., et al., *Familial aggregation and heritability of electrocardiographic intervals and heart rate in a rural Chinese population.* Ann Noninvasive Electrocardiol, 2009. **14**(2): p. 147-52.
8. Newton-Cheh, C., et al., *QT interval is a heritable quantitative trait with evidence of linkage to chromosome 3 in a genome-wide linkage analysis: The Framingham Heart Study.* Heart Rhythm, 2005. **2**(3): p. 277-84.
9. Akylbekova, E.L., et al., *Clinical correlates and heritability of QT interval duration in blacks: the Jackson Heart Study.* Circ Arrhythm Electrophysiol, 2009. **2**(4): p. 427-32.
10. Haarmark, C., et al., *Heritability of Tpeak-Tend interval and T-wave amplitude: a twin study.* Circ Cardiovasc Genet, 2011. **4**(5): p. 516-22.
11. Im, S.W., et al., *Analysis of genetic and non-genetic factors that affect the QTc interval in a Mongolian population: the GENDISCAN study.* Exp Mol Med, 2009. **41**(11): p. 841-8.
12. Smith, J.G., et al., *Genome-wide association study of electrocardiographic conduction measures in an isolated founder population: Kosrae.* Heart Rhythm, 2009. **6**(5): p. 634-41.
13. Maher, B., *Personal genomes: The case of the missing heritability.* Nature, 2008. **456**(7218): p. 18-21.
14. Zuk, O., et al., *The mystery of missing heritability: Genetic interactions create phantom heritability.* Proc Natl Acad Sci U S A, 2012. **109**(4): p. 1193-8.
15. Nolte, I.M., et al., *A Comparison of Heritability Estimates by Classical Twin Modeling and Based on Genome-Wide Genetic Relatedness for Cardiac Conduction Traits.* Twin Res Hum Genet, 2017. **20**(6): p. 489-498.
16. Nolte, I.M., et al., *A Comparison of Heritability Estimates by Classical Twin Modeling and Based on Genome-Wide Genetic Relatedness for Cardiac Conduction Traits.* Twin Res Hum Genet, 2017: p. 1-10.
17. Speed, D., et al., *Reevaluation of SNP heritability in complex human traits.* Nat Genet, 2017. **49**(7): p. 986-992.
18. Monnahan, P.J. and J.K. Kelly, *Epistasis Is a Major Determinant of the Additive Genetic Variance in Mimulus guttatus.* PLoS Genet, 2015. **11**(5): p. e1005201.
19. Arking, D.E., et al., *Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization.* Nat Genet, 2014. **46**(8): p. 826-36.
20. Nolte, I.M., et al., *Common genetic variation near the phospholamban gene is associated with cardiac repolarisation: meta-analysis of three genome-wide association studies.* PLoS One, 2009. **4**(7): p. e6138.

21. Newton-Cheh, C., et al., *Common variants at ten loci influence QT interval duration in the QTGEN Study.* Nat Genet, 2009. **41**(4): p. 399-406.
22. Marroni, F., et al., *A genome-wide association scan of RR and QT interval duration in 3 European genetically isolated populations: the EUROSPAN project.* Circ Cardiovasc Genet, 2009. **2**(4): p. 322-8.
23. Holm, H., et al., *Several common variants modulate heart rate, PR interval and QRS duration.* Nat Genet, 2010. **42**(2): p. 117-22.
24. Sotoodehnia, N., et al., *Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction.* Nat Genet, 2010. **42**(12): p. 1068-76.
25. Pfeufer, A., et al., *Genome-wide association study of PR interval.* Nat Genet, 2010. **42**(2): p. 153-9.
26. Shah, S., et al., *Four genetic loci influencing electrocardiographic indices of left ventricular hypertrophy.* Circ Cardiovasc Genet, 2011. **4**(6): p. 626-35.
27. Christophersen, I.E., et al., *Fifteen Genetic Loci Associated With the Electrocardiographic P Wave.* Circ Cardiovasc Genet, 2017. **10**(4).
28. Jessica van Setten, N.V., Hamdi Mbarek, Marieke Niemeijer, Stella Trompet, Dan E. Arking, Jennifer A. Brody, Ilaria Gandin, Niels Grarup, Leanne M. Hall, Leo-Pekka Lyytikäinen, Hao Mei, Martina Müller-Nurasyid, Bram P. Prins, Antonietta Robino, Albert V. Smith, Helen Warren, Folkert W. Asselbergs, Dorret I. Boomsma, Mark Caulfield, Mark Eijgelsheim, Ian Ford, Torben Hansen, Tamara B. Harris, Susan R. Heckbert, Jouke-Jan Hottenga, Anita Iorio, Jan A. Kors, Allan Linneberg, Peter MacFarlane, Thomas Meitinger, Christopher P. Nelson, Olli T. Raitakari, Claudia T. Silva Aldana, Gianfranco Sinagra, Moritz Sinner, Elsayed Z. Soliman, Andre Uitterlinden, Cornelia M. van Duijn, Melanie Waldenberger, Alvaro Alonso, Paolo Gasparini, Vilmundur Gudnason, Yalda Jamshidi, Stefan Kääb, Jørgen K Kanters, Terho Lehtimäki, Patricia Munroe, Annette Peters, Nilesh J. Samani, Nona Sotoodehnia, Sheila Ulivi, James G. Wilson, Eco J.C. de Geus, J. Wouter Jukema, Bruno Stricker, Pim van der Harst, Paul I.W. de Bakker, Aaron Isaacs, *Genome of the Netherlands imputation identifies seven new loci for quantitative ECG traits in a meta-analysis of 30,000 samples.* Human Molecular Genetics - submitted, 2017.
29. Nathan A Bihlmeyer, J.A.B., Albert Vernon Smith, Helen R Warren, Honghuang Lin, Aaron Isaacs, Ching-Ti Liu, Jonathan Marten, Farid Radmanesh, Leanne M Hall, Niels Grarup, Hao Mei, Martina Müller-Nurasyid, Jennifer E Huffman, Niek Verweij, Xiuqing Guo, Jie Yao Ruifang Li-Gao, Marten van den Berg, Stefan Weiss, Bram P Prins, Jessica van Setten, Jeffrey Haessler, Leo-Pekka Lyytikäinen, Man Li, Alvaro Alonso, Elsayed Z Soliman, Joshua C Bis, Tom Austin, Yii-Der Ida Chen, Bruce M Psaty, Tamara B Harrris, Lenore J Launer, Sandosh Padmanabhan, Anna Dominiczak, Paul L Huang, Zhijun Xie, Patrick T Ellinor, Jan A Kors, Archie Campbell, Alison D Murray, Christopher P Nelson, Martin D Tobin, Jette Bork-Jensen, Torben Hansen, Oluf Pedersen, Allan Linneberg, Moritz F. Sinner, Annette Peters, Melanie Waldenberger, Thomas Meitinger, Siegfried Perz, Ivana Kolcic, Igor Rudan, Rudolf A de Boer, Peter van der Meer, Henry J Lin, Kent D Taylor, Renée de Mutsert, Stella Trompet, J Wouter Jukema, Arie C Maan, Bruno H C Stricker, Fernando Rivadeneira, André Uitterlinden, Uwe Völker, Georg Homuth, Henry Völzke, Stephan B Felix, Massimo Mangino, Timothy D Spector, Michiel L Bots, Marco Perez, Olli T Raitakari, Mika Kähönen, Nina Mononen, Vilmundur Gudnason, Patricia B Munroe, Steven A Lubitz, Cornelia M van Duijn, Christopher H Newton-Cheh, Caroline Hayward, Jonathan Rosand, Nilesh J Samani, Jørgen K. Kanters, James G. Wilson, Stefan Kääb2, Ozren Polasek, Pim van der Harst, Susan R Heckbert, Jerome I Rotter, Dennis O Mook-Kanamori, Mark Eijgelsheim, Marcus Dörr, Yalda Jamshidi, Folkert W Asselbergs, Charles Kooperberg, Terho Lehtimäki, Dan E Arking,

Nona Sotoodehnia, *ExomeChip-wide analysis of 95,626 individuals identifies 10 novel loci associated with QT and JT intervals.* Circulation: CV genetics - submitted, 2017.
30. Verweij, N., et al., *Twenty-eight genetic loci associated with ST-T-wave amplitudes of the electrocardiogram.* Hum Mol Genet, 2016. **25**(10): p. 2093-2103.
31. van der Harst, P., et al., *52 Genetic Loci Influencing Myocardial Mass.* J Am Coll Cardiol, 2016. **68**(13): p. 1435-48.
32. van den Berg, M.E., et al., *Discovery of novel heart rate-associated loci using the Exome Chip.* Hum Mol Genet, 2017. **26**(12): p. 2346-2363.
33. Amin, N., et al., *Refining genome-wide linkage intervals using a meta-analysis of genome-wide association studies identifies loci influencing personality dimensions.* Eur J Hum Genet, 2013. **21**(8): p. 876-82.
34. Gudbjartsson, D.F., et al., *Allegro, a new computer program for multipoint linkage analysis.* Nat Genet, 2000. **25**(1): p. 12-3.
35. Abecasis, G.R., et al., *Merlin--rapid analysis of dense genetic maps using sparse gene flow trees.* Nat Genet, 2002. **30**(1): p. 97-101.
36. Evans, D.S., et al., *Fine-mapping, novel loci identification, and SNP association transferability in a genome-wide association study of QRS duration in African Americans.* Hum Mol Genet, 2016. **25**(19): p. 4350-4368.
37. Verweij, N., et al., *Genetic determinants of P wave duration and PR segment.* Circ Cardiovasc Genet, 2014. **7**(4): p. 475-81.
38. Jeff, J.M., et al., *Generalization of variants identified by genome-wide association studies for electrocardiographic traits in African Americans.* Ann Hum Genet, 2013. **77**(4): p. 321-32.
39. Eppinga, R.N., et al., *Identification of genomic loci associated with resting heart rate and shared genetic predictors with all-cause mortality.* Nat Genet, 2016. **48**(12): p. 1557-1563.
40. Kim, J.W., et al., *A common variant in SLC8A1 is associated with the duration of the electrocardiographic QT interval.* Am J Hum Genet, 2012. **91**(1): p. 180-4.
41. den Hoed, M., et al., *Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders.* Nat Genet, 2013. **45**(6): p. 621-31.
42. Sano, M., et al., *Genome-wide association study of electrocardiographic parameters identifies a new association for PR interval and confirms previously reported associations.* Hum Mol Genet, 2014. **23**(24): p. 6668-76.
43. Volpi, S., et al., *Whole genome association study identifies polymorphisms associated with QT prolongation during iloperidone treatment of schizophrenia.* Mol Psychiatry, 2009. **14**(11): p. 1024-31.
44. Smith, J.G., et al., *Impact of ancestry and common genetic variants on QT interval in African Americans.* Circ Cardiovasc Genet, 2012. **5**(6): p. 647-55.
45. 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, 2013. **8**: p. 91.
46. Melkani, G.C., et al., *The UNC-45 chaperone is critical for establishing myosin-based myofibrillar organization and cardiac contractility in the Drosophila heart model.* PLoS One, 2011. **6**(7): p. e22579.
47. Qian, L., J. Liu, and R. Bodmer, *Slit and Robo control cardiac cell polarity and morphogenesis.* Curr Biol, 2005. **15**(24): p. 2271-8.
48. Monier, B., et al., *Steroid-dependent modification of Hox function drives myocyte reprogramming in the Drosophila heart.* Development, 2005. **132**(23): p. 5283-93.

49. Han, Z., et al., *Hand, an evolutionarily conserved bHLH transcription factor required for Drosophila cardiogenesis and hematopoiesis.* Development, 2006. **133**(6): p. 1175-82.
50. Hu, Y., et al., *An integrative approach to ortholog prediction for disease-focused and other functional studies.* BMC Bioinformatics, 2011. **12**: p. 357.
51. Deo, R. and C.M. Albert, *Epidemiology and genetics of sudden cardiac death.* Circulation, 2012. **125**(4): p. 620-37.
52. Krijthe, B.P., et al., *Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060.* Eur Heart J, 2013. **34**(35): p. 2746-51.

Chapter 7.2

Summary / Samenvatting

Summary

The electrocardiogram (ECG) is a tool to obtain evidence for arrhythmias, which are related with cardiac conduction system abnormalities, myocardial ischemias and cardiac response to drugs. Additionally, ECG measurements are useful in prediction of the outcomes of patients with heart rythm disease, and prediction of cardiovascular mortality in healthy subjects.

Significant genetic contribution to ECG measurements have been established, at least 58 loci have been associated with ECG measurements variability. The aim of this study was to discover rare and common variants by linkage analysis in a large family-based study the Erasmus Rucphen family (ERF) study, moreover we did a heritability estimation and finally we perform a functional analysis of ARHGAP24 a gene previously associated with ECG variability. We search the linked regions in detail using exon sequencing.

To find common variants underlying the linkage peaks we performed association analysis within the linkage regions using the SNP data from ERF. As common variations did not explain the linkage peaks; we next explored the hypothesis whether the linkage is explained by rare exonic variants in these regions. This effort does not uncover any significant variation. We establish heritability of ECG measurements: 37% for PR and 33% for QT and QRS.

Looking for common variants, we performed a genome wide association study (GWAs) for myocardial mass, and we found 32 novel loci, among 52 genomic loci, associated with this trait. Knockdown studies in *Drosophila*, let us validate some of our findings, since we found specific cardiac defects.

Finally, we perform a functional analysis, using a morpholino strategy; we depleted ARHGAP24 expression in zebrafish. Zebrafish embryos exhibit heart abnormalities, cardiac edema and heart beating reduction. We can conclude that ARHGAP24 is related with heart development, since its knockdown induces changes in heart zebrafish morphology and function.

Samenvatting

Chapter 8

Chapter 8

Appendix


ATTCGGTACCTCATAACGGCTAGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTT
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGG
ATTCGGTACCTCATAACGGCTAGGTAACTGATCTTAGGCTACGGGTACCTA

Chapter 8.1

Acknowledgements/Agradecimientos

Acknowledgements/Agradecimientos

In the last eight years, I knew amazing people, people that in one or other way help me to grow, help me to learn and help me to live this PhD experience. So many things lived: Been a student again, leave my country, laboratory experiments, learn to deal with R or not get loss all the time in Rotterdam, among funny and painful memories. I want to say thank you, I hope not to forget anyone, the list of people is huge, especially because this PhD starts as a dream, dreamed by three people: Monica Ortegón, Carlos M Restrepo and Carlos Trillos. Thanks for this ERACOL dream, thanks for this amazing opportunity.

I would like to express my gratitude to my promoters starting with Cornelia van Duijn. Dear Cornelia, thanks for accept me in your group, I learn so many things from you, if is valid to say I admire you. I admire your intelligence and good memory. Thanks a lot for your support and concern. Secondly, to Rob, you teach me and guide in a lot of things, even how to skate, thanks for all your support with my zebrafish project and in my thesis. It is difficult to express my appreciation for both of you.

Secondly to my co-promoter: Aaron, your invaluable support, especially by the ending of my first year when you really worried about me, helps me with a new project canalizing my knowledge and interest. My appreciation for your concern and patience. I remembered one day, that you were really upset because I did not understand what you were saying to me; then, you set in my desk and explained me step by step how to do it.

And of course, to my PhD director in Colombia, Doctor Carlos M Restrepo. Dr Restrepo you arrived one day, with the idea that ERACOL was my opportunity to do my PhD. You told me: “you

have to apply". Since you talk with me, I was thinking, my boss it´s totally crazy, how can I leave my family, my work, my place, everything? And then I arrived home and my family support your idea, and I decided to apply to ERACOL scholarship. You put me there, you believe in me. Thanks a lot, for your dream, that make me dream enough to take my bags and fly to the Netherlands.

Special thanks to the members of my doctoral committee doctor Juan Manuel Anaya, Ben Oostra, Rob Willemsen and …. I really appreciate your effort, time and energy invested in my path towards promotion and my PhD candidature, using your time, to read my work.

Thanks a lot, to Najaf Amin, Najaf you are a great person, a great teacher, thanks for your expert advice. Najaf, you guide me and support in some sad moments for me, thanks a lot. Thanks a lot, to Tatyana Asenovich and Irina Zorkoltseva, for their valuable contributions to the work of this thesis. Thanks to Jan Kors, Marieke Niemeyer and Bruno Stricker for your contributions and help.

To all my coauthors, specially to Pim vab der Harst, Jessica van Setten, Nona Sotoodehnia, Dan E Arking and Paul I.W. de Bakker for all your support to my work.

To the participants of the Erasmus Rucphen Family Study and the Rotterdam Study without you, it was not possible to do this research, you invest your time and agreed to share their personal information, thank you.

To college of Genepi office: Sven, Elisa, Sara, Carla, Ayse, Maarten, Lennart, Nati, Linda, Maksim, Maaike and Dina. Maaike, thanks for all your help. To Lab people: Widagdo, Karina, Marialuisa, Ronald, Bianca, Ragendra, Tianna, Maria, Andreea, Renata, Francesca, Sotiris, Vanessa. Especial thanks to Herma and Lies-Anne. Herma and Lies-Anne, you always support me, thanks for your kindness and appreciation. All of you people make my days a learning, funny experience, arrive to the lab was a enjoyable experience, it was a pleasure to work at your side.

Mark, Ilona and Luba, thanks for your hospitality, your invitation, your kindness.

There were people around my youngest daughter when she was totally by herself in Bogota: Cami, Pamu, Pachu and of course my favorite cousin Magda and her lovely family (Gerardo, Fede y mi negra Hermosa): Ustedes simplemente son lo mejor de lo mejor y los adoro. Gracias por haber estado para Cata cuando nadie más estuvo. Cami “mi hija adoptiva” gracias por tu amor, siempre tendrás un espacio especial en mi corazón, Pamu mi sobrina hermosa, te quiero. Pachú tú al lado de Cata fiel y firme.

To my trip adeventure cowrkers and parthers of this amazing PhD, for the share tears and laughing. ERACOL people thanks: Ma Jo, mi peque, Dani, Juana la loca, Jairo, Marcela Gálvez, Stephanny, Juan Guillermo, Dianita. A July Benavides, Mónia Niño, Jorge Cárdenas, Jorge Rugeles, Calvache, Iván: gracias por nuestras largas jornadas de estudio y por todos los momentos compartidos. Ustedes hicieron el inicio de esta aventura mucho más llevadera. Mile and Nelson, thanks for sharing with us your special friendship, thanks a lot, for receive us in your place, for your kindness and all your support. Andre y Jorgito you are amazing. Thanks for Christmas eves at your place, for your lovely an especial friendship. Alexandra, gracias. Especial agradecimiento a Jose Andrés Calvache, no sé cómo hubiera sobrevivido a data analysis y a biostatistical methods II sin tu ayuda. Angie y Moni gracias por todo su cariño y hospitalidad. Cesar, gracias por todo el apoyo, el cariño y todos aquellos momentos compartidos. Mónica Pinilla, Andrés Patiño, Catherine Cantor, Maria Carolina Medina, Ingrid Leal, gracias.

Lucas, no solo te convertiste en mi gran amiga, compañera de batallas, sino que aprendí mucho de ti, aun estando lejos, un mensaje y sin importar tu cansancio o tus ocupaciones siempre tuviste tiempo para mí y mis innumerables preguntas.

Ana Catalina, my lovely daughter. At the beginning you get up at 4.00 AM to talk with me, your support living alone and been an extraordinary human being and a marvelous daughter. When I hug you for the first time after one year and a half I cannot believe that I was hugging you at the other side of the Ocean. Ana Cata, THANKS, with all my heart thanks, thanks for helping me in everything.

Astrid Vrakking, Thanks for your support and kindness.

Mama gallina, tú sabes lo especial que eres para mí, gracias por todo tun cariño.

Colombian'office people: Dorita y Noris, ustedes son seres muy especiales y siempre las llevaré en mi corazón. Gracias por su compañía, por su cariño, por sus concejos, por su amistad. Dorita, te admiro profundamente, Noris eres un ser humano absolutamente maravillosa.

To Mauricio Arcos-Burgos and Claudio Mastronardi, you believe in me, support me and IMT is the best team ever. You are great and amazing, thanks for everything.

To Rosario University, its directives, especially Dr Hans Peter Knudsen and Leonardo Palacios, thanks for all the support, for my scholarship, my license. Thanks Dr Gustavo Quintero, thanks Universidad del Rosario.

To my master students, because all days I learned a lot from you: Andreita, Ana María, Caro sierra, Caro Carlosama, Jessi, Rossi, Sandrita, Alejo, Ma del Socorro, Danyela, Daniela, Olguita, Constanza, Mauricio, Julian, Rodrigo, Indira. To my special students of the school of medicine: Lau y Sonia, gracias por su cariño. A mis mentis que me inspiran día a día a ser una mejor persona.

To my family, specially to my mom, porque me enseñaste a soñar, a creer en mí y a luchar por lo que creo. Gracias por ese libro que pusiste en mis manos cuando tenía 11 o 12 acerca de la vida de madam Curie, creo que eso ayudó a inspirar mis sueños y a volar lejos. Gracias por ser la mejor mamá del mundo, Gracias por querer hablar conmigo todos los días, por contar los días para mi regreso y por ese abrazo interminable cuando nos volvimos a ver. Te amo!!! María Fabiola

Nusbaum, no hay palabras suficientes para decirte gracias, creíste en mí, extendiste tus manos llenas de amor y generosidad. Aún me recuerdo llorando por separme de ti cuando apenas era una niña y ese vincula de hermanas, ha trascendido el tiempo y el espacio. John, thanks for your kideness and love. Gracias. A Pacho “mi Viejo curare” por motivarme a hacer ciencia, por ser un gran hermano, por tu sonrisa y tu cariño. Arturo, eres un gran hermano y te adoro, gracias por tu amor. Gracias Gloria, Yima, loquito, Oscar, Victor, ustedes de una u otra manera han aportado grandes cosas a mi vida, gracias. Thanks to Alma, Santa Lolis, Marly, Olguita, Rubi, To my nephiews, and nices, specially Pachu, Papu, Camilo, Pamu to my favorit sicologists (Arturo & MMer), Alejo, thanks for your love. Alex, thanks for your calls. Each of you support me in diverse ways, but the most important was your love.

Caro, gracias por ser tan especial, gracias por haber llegado a mi vida y ser parte de ella. Eres una persona que me ha llenado de cariño y amor. Gracias por acompañar a Cata, gracias por tu amor. Te quiero hija.

And you, Alberto Castellanos, dejaste todo por mí, empacaste la maleta llena de sueños, por ir tras de mis sueños, te convertiste en el mayor apoyo que una mujer sueña tener. Tú eres mi mejor amigo, mi esposo, mi compañero, mi amante, gracias mi amor por tu apoyo, por tu amor, gracias por creer en mí, en mis sueños. Gracias por tomar mi mano y embarcarte conmigo al otro lado del mundo para vivir esta aventura loca llamada doctorado.

Chapter 8.2
About the Author

About the Author

Claudia Tamar Silva Aldana was born in Neiva, Colombia, on March 13, 1967. The younguest of 5 brothers and 4 daughters, Arturo Silva was her fathers' name and Fabiola Aldana de Silva is her moms' name. She is married with Alberto Castellanos and she has two daughters: Carolina and Catalina. She completed her school (pre – university) education at the Cardenal Sancha in Bogotá, Colombia and started her degree in Biology at Pedagógica Nacional University, Bogotá, Colombia.

In 1992 she obtained her title and in 1993 she started her magister in Biology with an emphasis in Genetics that she finished in 1998. During 1997, she started to work as a student in Rosario University because her magister thesis and in 1998 she stars work as a professor there.

In 2005 she was promoted to assistant professor and in 2008 to principal professor.

During 2010 she won a scholarship to do her PhD in The Netherlands. As part of the academic program in 2011 she obtained her second Master of Science in Health Sciences, specialisation Genetic Epidemiology in Erasmus University, Rotterdam, the Netherlands.

She started the work presented in this PhD thesis under supervision of Cornelia van Dujin (promoter) at the Genetic Epidemiology Unit, Deparment of Epidimiology, Erasmus Medical center Rotterdam, and Rob Willemsen, at clinical genetics department at Erasmus MC, the Netherlands, as her co-promoter she worked with Dr. Aaron Isaacs. In Colombia her director was Carlos Martin Restrepo, Rosario University, Bogotá, Colombia.

Currently she works as Principal Professor at Universidad del Rosario in Bogotá, Colombia.

Chapter 8.3

List of publications

List of publications

- RESTREPO C., CORREAL MC., GONZALES A., LOMBO T., GOMMEZ Y., SILVA CT., IZQUIERDO I., Mutaciones en el gen de la Distrofia Muscular Ligada al sexo en una población colombiana. Correlación clínico molecular. Crónica Científica 1.997.

- **Silva CT.,** Restrepo CM., Gómez Y., Correal MC., Izquierdo A., Lombo T., Gonzales A., Hernández P. Análisis de las deleciones del gen de la Distrofina en 28 pacientes con Distrofia Muscular de Duchenne (DMD) y Becker (DMB). Boletín Informativo Sociedad Colombiana de Genética (Memorias). Medellín Junio 1.999. Conferencia III Congreso Colombiano de Genética. Medellín Nov. 1.998.

- N Contreras, **CT Silva**, CM Restrepo. Posible Asociación entre fragilidad cromosómica y aborto recurrente. Memorias IV Congreso Colombiano de Genética. Popayán, 23-25 Febrero 2.000.

- Contreras N., **Silva C**., Mateus H., Restrepo CM. Informe de un caso con dup (21q) en una niña con diagnóstico clínico de Síndrome Down. EN: II Congreso Internacional V Congreso Colombiano de Genética. Genética en el siglo XXI: Avances y desafíos. Acta Biológica Colombiana. Vol. 6, No. 2, 2.001. pp. 68

- Hernández P., Gómez Y., **Silva CT.,** Restrepo CM., Identificación de portadoras de Distrofia Muscular de Duchenne y Becker (DMD/DMB) mediante análisis de dosis génica y polimorfismos de DNA. EN: Acta Biológica Colombiana. Vol. 6 No. 2. 2.001 pp. 12.

- Restrepo CM, Pineda L., Gómez Y., Gonzales A., **Silva CT.,** Villalobos MC., Morales A., Barrera-Saldaña H. Mutaciones en el gen regulador de la conductancia transmembranal de la conductancia transmembranal de la fib0rosis Quística en tres países lationoamericanos. EN: Acta Biológica Colombiana. Vol. 6 No. 2, 2.001. pp.11.

- Restrepo CM, **Silva CT**., Clonación: desde la oveja Dolly hasta el hombre. Qué podemos esperar? EN: Controversias en Ginecología y Obstetricia. Vol. 8 No. 41. Agosto 2.001. pp 144-150.

- Contreras NC., **Silva CT**., Restrepo CM., Asociación entre fragilidad cromosómica y aborto recurrente. EN: Ginecología y Obstetricia. Vol 8 No.44 Noviembre 2.001.

- Restrepo CM., **Silva CT**., Fonseca D., Gómez Y., Torres L., Guío S., Gutierrez C., Identificación de portadoras de hemofilia A para asesoramiento genético. EN: Controversias en Ginecología y Obstetricia. Vol. 11 No. 58. Marzo 2.003. pp 1105-1112.

- **Silva C**, Contreras N, Valenzuela F., Restrepo CM., Estudio Citogenético en líquido Amniótico para la detección de anomalías cromosómicas del feto. En: Controversias en Ginecología y Obstetricia. Vol. 11 No. 61 Junio 2.003. pp. 1227-1241.

- Mateus H, Fonseca D, **Silva C**, Contreras N, Restrepo C. El *virus del papiloma humano* una visión global del principal factor de riesgo para cáncer de cerviz. Controversias en Ginecología y Obstetricia. Vol. 13 Nª 72 del 2.004

- Fonseca D, Mateus H, **Silva CT**, Contreras NC, Payan C, Restrepo CM. Aspectos Clínicos-Moleculares de la Hiperplasia Suprarrenal Congénita. Controversias En Ginecología y Obstetricia. v.13, n.73, p.1857 - 1870, 2004

- **Silva C**, Fonseca D, Mateus H, Contreras N, Restrepo C. Deleciones en el gen de la distrofina en 62 familias colombinas, correlación genotipo-fenotipo para Análisis la distrofia muscular de Duchenne y Becker. Colombia Médica. 35(4):2004. Oct-Dic. http://colombiamedica.univalle.edu.co/Vol35No4/PDF/DELECION.PDF

- Payan C., Mateus H., Fonseca D., **Silva CT**, Contreras N., Restrepo CM., Falla ovárica prematura y Síndrome X-Frágil. Controversias En Ginecología y Obstetricia. v.13, n.74-75, p.1899 - 1903, 2005

- **Silva CT.,** Fonseca D., Contreras N., Mateus H., Payán C., Restrepo CM., Mecanismos Moleculares del Cáncer de Seno. Controversias En Ginecología y Obstetricia. v.14, n.76, p.2000 - 2006, 2004.

- Fonseca Dora, **Silva Claudia**, Gutierrez Andrés, Coll Mauricio, Arteaga Clara, Malo Gustavo, Giraldo Alejandro. Identificación de mutaciones en el gen de la 21 de hidroxilasa en pacientes afectados con Hiperplasia Suprarrenal Congénita. Biomédica. 25 (2): Junio 2005.http://redalyc.uaemex.mx/redalyc/pdf/843/84325209.pdf

- Fonseca Dora, Mateus Heidi, **Silva Claudia**, Contreras Nora, Restrepo Carlos, Deficiencia de glucosa 6-fosfato deshidrogenasa. Acta Médica Colombiana Vol. 30 No.2. Abril-junio 2.005: 59-64. http://www.scielo.org.co/pdf/amc/v30n2/v30n2a5.pdf

- **Silva Claudia**, Fonseca Dora, Mateus Heidi, Contreras Nora, Restrepo Carlos. Distrofia Muscular de Duchenne y Becker: Una visión molecular. Acta Médica Colombiana Vol. 30 No.3. Julio-Septiembre 2.005: 112-116. http://www.scielo.org.co/pdf/amc/v30n3/v30n3a5.pdf

- Mateus Heidi, Fonseca Dora, **Silva Claudia**, Contreras Nora. ADN fetal en sangre materna: implicaciones para el diagnóstico prenatal. Controversias en Ginecología y Obstetricia. Vol 14 No.79. 2132-2139. 2.005.

- Fonseca Dora, Gutiérrez Andrés, Mateus Heidi, **Silva Claudia**, Contreras Nora, Giraldo Alejandro. Análisis de muestras de orina para la detección molecular de enfermedades infecciosas. Aplicación en la identificación de citomegalovirus humano. Rev. Cienc. Salud. Bogotá. Vol. 3 No.2136-147. Julio-Diciembre de 2.005.http://redalyc.uaemex.mx/redalyc/src/inicio/ArtPdfRed.jsp?iCve=56230204

- JAIME SIMBAQUEBA, **SILVA ALDANA**, CLAUDIA TAMAR; FONSECA MENDOZA, DORA JANETH; MATEUS ARBELAEZ, HEIDI ELIANA; RESTREPO FERNÁNDEZ, CARLOS MARTÍN. Identificación de deleciones en pacientes colombianos con Distrofia Muscular de Ducehnne y Becker. IV congreso Internacional y VII congreso colombiano de Genética. EN: Revista de la Facultad de salud Universidad Industrial de Santander. Vol 38 (1).

- FONSECA MENDOZA, DORA JANETH, **SILVA ALDANA CLAUDIA TAMAR**; MATEUS ARBELAEZ, HEIDI ELIANA; RESTREPO FERNÁNDEZ, CARLOS MARTÍN. Análisis de dosis génica y construcción de haplotipos para la identificación de portadoras de mutaciones en el gen de la Distrofina. II encuentro de Investigadores Universidad del Rosario.

- Rodríguez G, Pineda AC, **Silva CT**, Rios DI, Amaya JC, Lagos A, Mateus J. Complejo *Burkholderia cepacea*: Reporte de un brote en una institución de tercer nivel de Bogotá, Colombia, año 2005. Saludarte. Vol 5 (1); p 58. Octubre 2006.

- Bermeo Sandra Milena, **Silva Claudia Tamar**, Fonseca Dora Janeth, Restrepo Carlos Martín. Hemofilia: diagnóstico molecular y alternativas de tratamiento. Revista Colombia Médica. Vol 38 No. 3, 2007 (Julio-Septiembre). http://www.scielo.org.co/pdf/cm/v38n3/v38n3a13.pdf
-

- **Silva CT**. Mateus H., Fonseca D., Simbaqueba J. Identificación de deleciones en exones situados fuera del Hot spot distal del gen de la Distrofina en pacientes afectados con distrofia Muscular de *Duchene*. Cuadernos de Medicina en Investigación y Salud. Dic 2007. Vol 1 No. 3. 206-212.

- Fonseca Dora, **Silva Claudia T,** Mateus Heidi. Detección de portadoras de distrofia muscular de Duchenne en familias colombianas mediante análisis de microsatélites. Revista Colombia Médica. Vol. 39 No2. (Julio-Septiembre) 2008. http://colombiamedica.univalle.edu.co/Vol39No2Supl2/htmlv39n2s2/v39n2s2a2.pdf
-

- Fonseca D., **Silva CT.,** Mateus H., Restrepo CM., Identificación de deleciones en portadoras de distrofia muscular de Duchenne. Revista Acta Médica Colombiana, Vol. 33 No. 2 (abril-junio) de 2008. 63-67. http://www.scielo.org.co/pdf/amc/v33n2/v33n2a4.pdf

- **Silva CT.,** Contreras NC., Fonseca DJ., utilidad de la citogenética en la medicina actual, una visión histórica y de aplicación. Revista Acta Médica Colombiana. Vol 33 No. 4 (Oct-Diciembre) de 2008. 309-316. http://www.scielo.org.co/pdf/amc/v33n4/v33n4a9.pdf

- Fonseca Dora J., Mateus H.E., Contreras N., Sánchez R., Herrera T., **Silva A Claudia T**., Análisis de deleciones en 15 exones situados detnro y fuera del hot spot mutacional del gen de la distrofina en pacientes con distrofia muscular de Duchenne. Revista Ciencias de la salud. Bogotá (Colombia). V / No. 2 pp5-66 Mayo-agosto. 2009.

- HEIDI ELIANA MATEUS ARBELAEZ, **CLAUDIA SILVA ALDANA,** NORA CONTRERAS BRAVO, SANDRA OSPINA LAGOS, DORA FONSECA, "Análisis Clínico y Molecular de una Paciente con Pentasomía del Cromosoma X". En: Acta Biológica Colombiana ISSN: 0120-548X ed: Facultad De Ciencias Universidad Nacional. v.15 fasc.2 p.61 - 72 ,2010.

- CONTRERAS BRAVO NORA CONSTANZA, **SILVA ALDANA CLAUDIA,** MATEUS ARBELAEZ HEIDI ELIANA. "Correlación genotipo-fenotipo y análisis molecular en pacientes con síndrome Down". Revista ciencias de la salud. 2012. 10(3) 295-305.

- Dora Janeth Fonseca Mendoza, Heidi Eliana Mateus Arbelaez y **Claudia Tamar Silva**. PERDIDA DE HETEROCIGOCIDAD E IDENTIFICACION DE PORTADORAS DE DISTROFIA MUSCULAR DE DUCHENNE: UN CASO FAMILIAR CON EVENTO DE RECOMBINACION. Revista Ciencias de la salud. Vol 10 No. 1. Bogotá jan/abril. 2012.

- **Silva CT.,** Kors JA., Amin N., Dehghan A., Witteman J., Oostra B., Van Duijn C., Isaacs A., Heritabilities proportions of heritabilities explained by GWAS findings, and implications of cross-phenotype effects on PR interval. Hum Genet. 2015 Nov;134(11-12):1211-9.

- Verweij N, Mateo Leach I, Isaacs A, Arking DE, Bis JC, Pers TH, Van Den Berg ME, Lyytikäinen LP, Barnett P, Wang X; LifeLines Cohort Study., Soliman EZ, Van Duijn CM, Kähönen M, Van Veldhuisen DJ, Kors JA, Raitakari OT, **Silva CT**, Lehtimäki T, Hillege HL, Hirschhorn JN, Boyer LA, Van Gilst WH, Alonso A, Sotoodehnia N, Eijgelsheim M, De Boer RA, De Bakker PI, Franke L, Van Der Harst P. Twenty-eight genetic loci associated with ST-T-wave amplitudes of the electrocardiogram. Hum Mol Genet. 2016 May 15;25(10):2093-2103.

- van der Harst P, van Setten J, Verweij N, Vogler G, Franke L, Maurano MT, Wang X, Mateo Leach I, Eijgelsheim M, Sotoodehnia N, Hayward C, Sorice R, Meirelles O, Lyytikäinen LP, Polašek O, Tanaka T, Arking DE, Ulivi S, Trompet S, Müller-Nurasyid M, Smith AV, Dörr M, Kerr KF, Magnani JW, Del Greco M F, Zhang W, Nolte IM, **Silva CT,** Padmanabhan S, Tragante V, Esko T, Abecasis GR, Adriaens ME, Andersen K, Barnett P, Bis JC, Bodmer R, Buckley BM, Campbell H, Cannon MV, Chakravarti A, Chen LY, Delitala A, Devereux RB, Doevendans PA, Dominiczak AF, Ferrucci L, Ford I, Gieger C, Harris TB, Haugen E, Heinig M,

Hernandez DG, Hillege HL, Hirschhorn JN, Hofman A, Hubner N, Hwang SJ, Iorio A, Kähönen M, Kellis M, Kolcic I, Kooner IK, Kooner JS, Kors JA, Lakatta EG, Lage K, Launer LJ, Levy D, Lundby A, Macfarlane PW, May D, Meitinger T, Metspalu A, Nappo S, Naitza S, Neph S, Nord AS, Nutile T, Okin PM, Olsen JV, Oostra BA, Penninger JM, Pennacchio LA, Pers TH, Perz S, Peters A, Pinto YM, Pfeufer A, Pilia MG, Pramstaller PP, Prins BP, Raitakari OT, Raychaudhuri S, Rice KM, Rossin EJ, Rotter JI, Schafer S, Schlessinger D, Schmidt CO, Sehmi J, Silljé HH, Sinagra G, Sinner MF, Slowikowski K, Soliman EZ, Spector TD, Spiering W, Stamatoyannopoulos JA, Stolk RP, Strauch K, Tan ST, Tarasov KV, Trinh B, Uitterlinden AG, van den Boogaard M, van Duijn CM, van Gilst WH, Viikari JS, Visscher PM, Vitart V, Völker U, Waldenberger M, Weichenberger CX, Westra HJ, Wijmenga C, Wolffenbuttel BH, Yang J, Bezzina CR, Munroe PB, Snieder H, Wright AF, Rudan I, Boyer LA, Asselbergs FW, van Veldhuisen DJ, Stricker BH, Psaty BM, Ciullo M, Sanna S, Lehtimäki T, Wilson JF, Bandinelli S, Alonso A, Gasparini P, Jukema JW, Kääb S, Gudnason V, Felix SB, Heckbert SR, de Boer RA, Newton-Cheh C, Hicks AA, Chambers JC, Jamshidi Y, Visel A, Christoffels VM, Isaacs A, Samani NJ, de Bakker PI. 52 Genetic Loci Influencing Myocardial Mass. J Am Coll Cardiol. 2016 Sep 27;68 (13):1435-48. doi: 10.1016/j.jacc.2016.07.729.

- **Silva CT**, Zorkoltseva IV, Amin N, Demirkan A, van Leeuwen EM, Kors JA, van den Berg M, Stricker BH, Uitterlinden AG, Kirichenko AV, Witteman JC, Willemsen R, Oostra BA, Axenovich TI, van Duijn CM, Isaacs A. A Combined Linkage and Exome Sequencing Analysis for Electrocardiogram Parameters in the Erasmus Rucphen Family Study. Front Genet. 2016 Nov 8;7:190.

- Marten E. van den Berg, Helen R. Warren, Claudia P. Cabrera, Niek Verweij, Borbala Mifsud, Jeffrey Haessler, Nathan A. Bihlmeyer, Yi-Ping Fu, Stefan Weiss, Henry J. Lin, Niels Grarup, Ruifang Li-Gao, Giorgio Pistis, Nabi Shah, Jennifer A. Brody, Martina Mu¨ ller-Nurasyid, Honghuang Lin, Hao Mei, Albert V. Smith, Leo-Pekka Lyytikainen, Leanne M. Hall, Jessica van Setten, Stella Trompet, Bram P. Prins, Aaron Isaacs, Farid Radmanesh, Jonathan Marten, Aiman Entwistle, Jan A. Kors, Claudia T. Silva, Alvaro Alonso, Joshua C. Bis, Rudolf de Boer, Hugoline G. de Haan, Rene´e de Mutsert, George Dedoussis, Anna F. Dominiczak, Alex S. F. Doney, Patrick T. Ellinor, Ruben N. Eppinga, Stephan B. Felix, Xiuqing Guo, Yanick Hagemeijer, Torben Hansen, Tamara B. Harris, Susan R. Heckbert, Paul L. Huang, Shih-Jen Hwang, Mika Kahonen, Jørgen K. Kanters, Ivana Kolcic, Lenore J. Launer, Man Li, Jie Yao, Allan Linneberg, Simin Liu, Peter W. Macfarlane, Massimo Mangino, Andrew D. Morris, Antonella Mulas, Alison D. Murray, Christopher P. Nelson, Marco Orru, Sandosh Padmanabhan, Annette Peters, David J. Porteous, Neil Poulter, Bruce M. Psaty, Lihong Qi, Olli T. Raitakari, Fernando Rivadeneira, Carolina Roselli, Igor Rudan, Naveed Sattar, Peter Sever, Moritz F. Sinner, Elsayed Z. Soliman, Timothy D. Spector, Alice V. Stanton, Kathleen E. Stirrups, Kent D. Taylor, Martin D. Tobin, Andre´ Uitterlinden, Ilonca Vaartjes, Arno W. Hoes, Peter van der Meer, Uwe Volker, Melanie Waldenberger, Zhijun Xie, Magdalena Zoledziewska, Andrew Tinker, Ozren Polasek, Jonathan Rosand, Yalda Jamshidi, Cornelia M. van Duijn, Eleftheria Zeggini, J. Wouter Jukema, Folkert W. Asselbergs, Nilesh J. Samani, Terho Lehtimaki, Vilmundur Gudnason, James Wilson, Steven A. Lubitz, Stefan Ka€ab, Nona Sotoodehnia, Mark J. Caulfield, Colin N. A. Palmer, Serena Sanna, Dennis O. Mook-Kanamori, Panos Deloukas, Oluf Pedersen, Jerome I. Rotter, Marcus Dorr, Chris J. O'Donnell9, Caroline Hayward, Dan E. Arking, Charles Kooperberg,

Pim van der Harst, Mark Eijgelsheim, Bruno H. Stricker and Patricia B. Munroe. Discovery of novel heart rate-associated loci using exome chip, Human Molecular Genetics. Human Molecular Genetics, 2017, Vol. 26, No. 12: 234-63. PMID: 28379579. Doi: 10.1093/hmg/ddx113

- **Silva CT,** Zorkoltseva IV, Niemeijer MN, van den Berg ME, Amin N, Demirkan A, van Leeuwen E, Iglesias AI, Piñeros-Hernández LB, Restrepo CM, Kors JA, Kirichenko AV, Willemsen R, Oostra BA, Stricker BH, Uitterlinden AG, Axenovich TI, van Duijn CM, Isaacs A. A combined linkage, microarray and exome analysis suggests MAP3K11 as a candidate gene for left ventricular hypertrophy. BMC Med Genomics. 2018 Mar 5;11(1):22. doi: 10.1186/s12920-018-0339-9

- Cabrera R, Miranda-Fernández MC, Huertas-Quiñones VM, Carreño M, Pineda I, Restrepo CM, **Silva CT**, Quero R, Cano JD, Manrique DC, Camacho C, Tabares S, García A, Sandoval N, Moreno Medina KJ, Dennis Verano RJ. Identification of clinically relevant phenotypes in patients with Ebstein anomaly. Clin Cardiol. 2018 Mar;41(3):343-348. doi: 10.1002/clc.22870. Epub 2018 Mar 22. PMID: 29569399

- Miranda-Fernández MC, Ramírez-Oyaga S, Restrepo CM, Huertas-Quiñones VM, Barrera-Castañeda M, Quero R, Hernández-Toro CJ, **Tamar Silva C**, Laissue P, Cabrera R. Identification of a New Candidate Locus for Ebstein Anomaly in 1p36.2. Mol Syndromol. 2018 May;9(3):164-169. doi: 10.1159/000488820. Epub 2018 Apr 28. PMID: 29928183

Accepted papers:

- Jessica van Setten, Jennifer A. Brody, Yalda Jamshidi, Brenton R. Swenson, Anne M. Butler, Harry Campbell, Fabiola M. Del Greco, Daniel S. Evans, Quince Gibson, Daniel F. Gudbjartsson, Kathleen F. Kerr, Bouwe P. Krijthe, Leo-Pekka Lyytikainen, Christian Muller, Martina Muller-Nurasyid, Ilja M. Nolte, Sandosh Padmanabhan, Marylyn D. Ritchie, Antonietta Robino, Albert V. Smith, Maristella Steri, Toshiko Tanaka, Alexander Teumer, Stella Trompet, Sheila Ulivi, Niek Verweij, Xiaoyan Yin, David O. Arnar, Folkert W. Asselbergs, Joel S. Bader, John Barnard, Josh Bis, Stefan Blankenberg, Eric Boerwinkle, Yuki Bradford, Brendan M. Buckley, Mina K. Chung, Dana Crawford, Marcel den Hoed, Josh Denny, Anna F. Dominiczak, Georg B. Ehret, Mark Eijgelsheim, Patrick T. Ellinor, Stephan B. Felix, Oscar H. Franco, Lude Franke, Tamara B. Harris, Hilma Holm, Gandin Ilaria, Annamaria Iorio, Mika Kahonen, Ivana Kolcic, Jan A. Kors, Edward G. Lakatta, Lenore J. Launer, Honghuang Lin, Henry J. Lin, Ruth J.F. Loos, Steven A. Lubitz, Peter W. Macfarlane, Jared W. Magnani, Irene Mateo Leach, Thomas Meitinger, Braxton D. Mitchell, Thomas Munzel, George J. Papanicolaou, Annette Peters, Arne Pfeufer, Peter P. Pramstaller, Olli T. Raitakari, Jerome I. Rotter, Igor Rudan, Nilesh J. Samani, David Schlessinger, Claudia T. Silva Aldana, Moritz F. Sinner, Jonathan D. Smith, Harold Snieder, Elsayed Z. Soliman, Timothy D. Spector, David J. Stott, Konstantin Strauch, Kirill V. Tarasov, Andre G.

Uitterlinden, David R. van Wagoner, Uwe Volker, Henry Volzke, Melanie Waldenberger, Harm Jan Westra, Philipp S. Wild, Tanja Zeller, Alvaro Alonso, Christy L. Avery, Stefania Bandinelli, Emelia J. Benjamin, Francesco Cucca, Marcus Dorr, Luigi Ferrucci, Paolo Gasparini, Vilmundur Gudnason, Caroline Hayward, Susan R. Heckbert, Andrew A. Hicks, J. Wouter Jukema, Stefan Kaab, Terho Lehtimaki, Yongmei Liu, Patricia B. Munroe, Afshin Parsa, Ozren Polasek, Bruce M. Psaty, Dan M. Roden, Renate B. Schnabel, Gianfranco Sinagra, Kari Stefansson, Bruno H. Stricker, Pim van der Harst, Cornelia M. van Duijn, James F. Wilson, Sina Gharib, Paul I.W. de Bakker, Aaron Isaacs, Dan E. Arking, Nona Sotoodehnia. PR interval genome-wide association meta-analysis identifies 50 loci associated with atrial and atrioventricular electrical activity. Nature Communication. DOI: 10.1038/s41467-018-04766-9.

- Jessica van Setten, Niek Verweij, Hamdi Mbarek, Marieke Niemeijer, Stella Trompet, Dan E. Arking, Jennifer A. Brody, Ilaria Gandin, Niels Grarup, Leanne M. Hall, Leo-Pekka Lyytikäinen, Hao Mei, Martina Müller-Nurasyid, Bram P. Prins, Antonietta Robino, Albert V. Smith2, Helen Warren, Folkert W. Asselbergs, Dorret I. Boomsma4, Mark Caulfield, Mark Eijgelsheim, Ian Ford, Torben Hansen, Tamara B. Harris, Susan R. Heckbert, Jouke-Jan Hottenga, Anita Iorio, Jan A. Kors, Allan Linneberg, Peter MacFarlane, Thomas Meitinger, Christopher P. Nelson, Olli T. Raitakari, Claudia T. Silva Aldana, Gianfranco Sinagra, Moritz Sinner, Elsayed Z. Soliman, Andre Uitterlinden, Cornelia M. van Duijn, Melanie Waldenberger, Alvaro Alonso, Paolo Gasparini, Vilmundur Gudnason, Yalda Jamshidi, Stefan Kääb, Jørgen K Kanters, Terho Lehtimäki, Patricia Munroe, Annette Peters, Nilesh J. Samani, Nona Sotoodehnia, Sheila Ulivi, James G. Wilson, Eco J.C. de Geus, J. Wouter Jukema, Bruno Stricker, Pim van der Harst, Paul I.W. de Bakker, Aaron Isaacs, GWAS meta-analysis de 30,000 samples identifies seven novel loci for quantitative ECG traits. European Journal of human genetics.

Chapter 8.4

PHD portafolio summary

Name PhD student:	Claudia Tamar Silva Aldana
Erasmus MC Department	Epidemiology / Clinical Genetics
Universidad del Rosario Department	Genética
Research School	Netherlands Institute for Health Sciences (Nihes) Doctorado en ciencias Biomédicas Universidad del Rosario
PhD period	Augost 2010- Augost 2018
Promotor (s)	Prof.dr.ir C.M van Duijn, Prof.dr. R Willemsen,Prf. dr. Carlos M Restrepo
Co-promoter	Porf. Dr Aaron Isaacs

PhD training

	Year	Workload (Hours/ECTS/créditos)
In-depth courses		
NIHES Master of Science in Health Sciences	2010-2011	
		4.3
Study design		
Classical Methods for Data-analysis		5.7
Modern Statistical Methods		4.3
Genetic-Epidemiologic Research Methods		5.7
SNP´s and Human Diseases		1.4
Psychiatric Epidemiology		1.1
Courses for the Quantitative Researcher		1.4
Introduction to Clinical and public Health Genomics		1.9
European Human Genetics Conference 2011		1.1
Advances in Genome-Wide Association Studies		1.4
Family-based Genetic Analysis		1.4
Introduction to Medical Writing		1.1
Working with SPSS for Windows		0.15
Summer Course English		1.4
Development Research Proposal		2.5
Oral Research Presentation		1.4
Research Period		29.2
Universidad del Rosario	2013-2014	

Bioética
Seminario de Ciencia y Tecnología
Thesis seminars
Thesis I, II, III and special

Congress presentations

Genetics of ECG Traits: An overview. XLVIII CONGRESO NACIONAL DE CIENCIAS BIOLÓGICAS	2013	24
XIII Congreso Colombiano de Genética	2014	24

Oral presentations at lab meetings

Smoking and methylation levels	2011
Epigenetics	2012
Zebrafish preliminary results	2012
Zebrafish an update	2013