8.

GENERAL DISCUSSION AND FUTURE PERSPECTIVES
The lack of knowledge surrounding the basis of cardiomyocyte dysfunction and heart failure susceptibility is a major roadblock to understand risk for heart failure and designing innovative strategies for therapy. Our understanding of cardiovascular biology has increased steadily in the past few decades, but discoveries were mainly driven by exploring in-vitro and in-vivo animal models because of the (obvious) limitations for human based cardiovascular research. This limitation has also hampered the development of new drugs and therapies, as not more than a third of the animal research can eventually be translated to humans. Innovation in genetic research has been happening at a much faster pace, and is taking the field of biology along. Our understanding of human genetics and heritability has exploded years after the discovery of the human genome and since then rapidly increasing together with the development of novel technologies for measuring (novel) genetic components in a more accurate, quicker and cheaper fashion. The quest for human heritability and genetic variants has also yielded many genetic associations for human traits in the last few years. Analyses of the genome wide variation opened up the possibility to study the genetic background of cardiovascular traits in (living) humans and, in turn, study biological mechanisms that may specifically be important for humans, for example in treatment of heart failure. This dissertation describes the associations of genetic variants with endophenotypes of heart failure, with special interest for blood biomarkers and electrocardiographic traits that reflect cardiac conduction, myocardial mass and myocardial repolarization. The associations were complemented with in-silico, in-vitro and in-vivo analyses to assess their relevance for cardiovascular biology.

Clinical insights

Human characteristics are established through evolution by DNA changes, resulting in genetic variations that make each individual unique. Most of those characteristics, which can be quantitative traits and common diseases, are complex. They are influenced by many different genetic variants throughout the genome and by many non-genetic factors as well. These genetic factors often do not have a single function but can be pleiotropic across many phenotypes with varying effect sizes and may also affect Mendelian traits, adding to the complexity of common traits. For example, it was recently shown in literature that the variant in KLKB1 identified in Chapter 2 also appeared to be associated with multiple other blood markers such as B-type Natriuretic Peptide, renin, aldosterone and uPAR levels; the variant in the ABO gene associated with galectin-3 (Chapter 3) has also been reported to be associated with coronary heart disease, IL-6 levels, Graves disease, Pancreatic cancer, among many others; and variants identified to influence electrocardiographic phenotypes (Chapters 4, 5, 6 and 7) are...
also associated among each other and phenotypes like lipid concentrations. In Chapter 5 reports on a novel genome wide association study of cardiac repolarization parameters of the electrocardiogram. Two of the 28 ST-T wave associated signals are also influencing susceptibility for Brugada syndrome, what is believed to be a rare Mendelian disorder. The latter observation suggests that common genetic variants, such as those identified in this thesis, may also be important for the susceptibility and perhaps expression of monogenic cardiac diseases, which could in turn lead to heart failure. Future efforts aimed at determining whether other genetic variants described in this dissertation also play a role in monogenic disease will be relevant for elucidating cardiovascular disease mechanisms.

The relationships between the various genotype-phenotype associations may be difficult to entangle, a consistent direction of effect in two phenotypes may imply a shared causal mechanism and genes, while the same genetic variant can also have opposite effects on another trait. For example, variants that increase QRS duration (part of QT interval) can decrease QT on the ECG and variants resulting in prolonging of QT interval could have higher risk of sudden cardiac death, opposite to what would be expected. The KLKB1 locus exemplifies a more simple and easy to understand effect as it seems to have a consistent effect across the peptides it is associated with and also has a suggestive, but directionally concordant, effect on left ventricular mass. More often, other nearby variants (across the LD spectrum) in the same genetic region are associated with different diseases and traits. It is important to dissect variants that co-localize from those that act independently which could be used to differentiate between shared or separate mechanisms. Understanding these pleiotropic relationships will be crucial for a comprehensive biological interpretation of these loci and to correctly understand complex human traits.

GWA studies have the potential to yield new insights for possible clinical applications, diagnosis, prevention, treatment and disease monitoring. A powerful approach to gain new clinical insights is via a Mendelian randomization approach, exploiting genetic pleiotropy among biological processes and disease. Mendelian randomization is a method that uses genetic variants as proxies for environmental exposures, with the aim to study its effect on another variable, such as a disease. The power of using genetic variants as proxy measures of an exposure lies in the ability to avoid confounding factors that could bias the environmental exposure due to reverse causation, e.g., disease processes that influence the exposure. One of the best examples may be the relationship between lipid levels and risk for cardiovascular risk. Higher plasma levels of HDL and lower LDL are typically associated with lower risk for myocardial infarction in epidemiological studies, however drugs that raise HDL have failed repeatedly, whereas LDL lowering drugs are known to work well. With hindsight sensitive drug trials of HDL raising drugs could have been avoided based on the results of Mendelian randomization studies on lipids. A genetic variant in the LIPG gene, or even multiple SNPs, specifically associated with HDL levels, show no association with risk for cardiovascular risk; whereas lower LDL levels have been confirmed as causal risk factor by a Mendelian randomization approach. This approach could also be applied to genetic variants that are described in this dissertation. Chapter 2 describes the association of four loci with cardiovascular disease and whether these peptides should be considered a target for therapy in heart failure. In addition, the other two variants that were identified in the kallikrein-kinin system (KLKB1 and F12) for MR-pro-ADM and CT-pro-ET-1 levels, should be used in a Mendelian randomization approach for kallikrein inhibition, thereby assessing the potential of plasma kallikrein as a novel therapeutic target. As described in the previous paragraph, the variant in KLKB1 also has a concordant effect on left ventricular mass, making it one of the most interesting findings of this dissertation with the potential to serve as a direct therapeutic target. Plasma kallikrein has previously been suggested as a potential drug target for interfering with hemotma expansion under conditions of hyperglycemia. Similarly, other variants, for example those influencing QRS-voltage of the ECG as proxy for cardiac hypertrophy (Chapter 6,7), should be evaluated on cardiac outcome in the general population or even in subpopulations such as heart failure patients or patients with inherited cardiomyopathies.

**BIOLICAL INSIGHTS**

Genome wide associations have the potential to discover novel biological knowledge relevant for human traits, as also described in this dissertation. The recent study on height has shown that by increasing the sample size well over 100,000’s of individuals, the relevance of the biological information of the associated loci also become clearer, suggesting that continued efforts will be effective to increase our understanding of cardiovascular biology. However, identifying all of the genetic components that define heritability of complex traits will remain difficult. In this dissertation I have mostly focused on variants that are common (frequency >5%) in the population. However, it is estimated that 30-50% of genetic variation is not tagged by commonly used genotyping chips which is suggesting that lower frequency variants may also be im-
portant for human (cardiac) traits. This is further symbolized by the estimation that each individual carries 60-70 novel genetic variants, so with the ever expanding population, most segregating variants will be rare. To detect the effects of rare variants, sample sizes of millions of individuals may be required to reliably detect the effects on human phenotypes under the assumption that their effect is similar to common variants. This complexity makes it very difficult and also very costly to identify all genetic variants underlying common human traits, unlike most Mendelian traits, where a single mutation with great effect often explains most of the occurrence of a single disease and its heritability. An alternative approach to increase statistical power by collecting many more individuals, would be to prioritize variants for replication, or even discovery, based on their functional enrichments in functional annotations and pathways with information from ENCODE\textsuperscript{18} and Roadmap epigenomics projects\textsuperscript{19} (Chapters 4-7) and tools like DEPICT (Chapter 5).

We, and others have shown, that ECG related variants reside preferentially in regulatory regions of the genome, such as enhancers of cardiac tissue (Chapter 4-7). Although prioritization strategies based on functional DNA elements have been suggested before\textsuperscript{20}, they have not been widely adopted yet. These strategies should be taken into consideration when designing new genome wide association studies of ECG traits as they could dramatically improve the power to detect novel genetic variants. Caution is warranted however as this method could lead to a bias towards the identification novel cardiac genes and biology, especially when incorporating information about genes function and their biological pathways. Also, non-regulatory variants in intergenic regions or functional annotations in tissue and cell types that have not been studied could point to novel biological mechanisms underlying the phenotype that we may not yet understand and should therefore not neglect.

The majority of potential causal variants identified in Chapter 2 and Chapter 3 change the amino acid sequence of its protein-products (coding SNPs), likely underlying the phenotype-genotype associations. However, a large proportion of GWAS loci do not overlap protein-coding genes (Chapter 4-7), making it difficult to understand the underlying mechanisms. Characterizing and understanding functional consequences of genetic variation in regulatory elements of the heart will be an important step towards elucidating the biology underlying human cardiovascular traits, in particular heart failure, but also complex traits in general. It gets increasingly clear that genetic variations within regulatory DNA regions could disrupt many different biological-genetic systems\textsuperscript{21,22} such as transcription factor binding sites (at enhancers, promoters or insulator sites), chromatin architecture, DNA methylation, alternative splicing, post-transcriptional processing, and microRNA targeting (see also Chapter 4-7). Most of the efforts are currently directed at understanding enhancer regions with disrupted transcription factor binding as a potential cause for differential gene expression\textsuperscript{21,23-25} which is also described in Chapter 5 and 6. Perhaps because this method is most easily understood and experimentally tested, or it could be the principle driver. The exact underlying mechanisms of genetic variations will need in-depth functional experiments. Low-to-high throughput screens like luciferase assays, EM-SAs, in-vivo LacZ, allele specific ChiP assays and 4C-Seq could be used to get a general overview of the mechanisms involved, but specifically tailored experiments for each locus will likely be required to fully understand the initial genotype-phenotype associations, perhaps by applying the emerging Crispr-cas9 technologies\textsuperscript{26,27} in relevant cell lines to observe the effects of variants in a controlled in vivo or in vitro environment.

Clearly, the methods for elucidating the mechanisms underlying genetic variation are still in their infancy and technologies to do so are rapidly evolving. There are several efforts ongoing to facilitate functional experiments as described above and interpret genetic variation underlying the epigenetic landscape of the genome. This dissertation relies on genetic variation data by generated genotyping arrays (in Prevent, Lifelines and others) that does not cover the complete picture of human genetic variation, which could alternatively be achieved by DNA sequencing. Therefore imputation (and consulting) of unobserved genotypes is necessary to increase the chance of identifying causal SNPs, possible by sequencing data generated by e.g. 1000 Genomes (Chapter 5), and the Dutch genome GoNL (Chapter 7), which are a significant step forward compared to the earliest genome wide association studies. Even more detailed SNP coverage of the human genome will become available through The Haplotype Reference Consortium\textsuperscript{28} that combines human genomes from up to 25 consortia (including 1000Genomes and GoNL), comprising 38,662 samples in one reference set. Such a large reference set will make it even more likely that the causal variants will be imputed. To interpret the genetic variation, it is also necessary to understand the role of the genetic region they reside in. Therefore, large consortia have been set up that are continuously generating data on the epigenetic landscape, gene and enhancer expression, such as: ENCODE\textsuperscript{29}, Roadmap Epigenomics\textsuperscript{30}, Blueprint Epigenome\textsuperscript{23}, FANTOM5\textsuperscript{31} and GTEX\textsuperscript{32}. These data sets are a valuable resource to interpret the genome, direct functional experiments, as shown in Chapter 4-6, and will continue to grow larger. Additional insights of human genetic variation will be facilitated by the Human Genome Variation Map\textsuperscript{33}, initiated as successor of the Human Genome Project. Until now, interpretation of DNA sequencing data (e.g. from epigenetic data sets and reference sets) relies on a single refer-
ence genome, which is insufficient for representing humanity as a whole. The Human Genome Variation Map project is developing a model of the genome to incorporate genetic variation, including the most complicated variants that are hard to align to a single (flat) genome. Together, these efforts will allow for better interpretation of genetic variation, epigenetic and expression datasets, and guide functional experiments to fully understand biological mechanisms underlying the associations identified in this dissertation mainly focused on cardiac function and proxies of left ventricular hypertrophy as endophenotypes of heart failure. Genetics have the potential to revolutionize medicine and point to novel therapeutic target with high accuracy, which is also outlined in the introduction of this dissertation. Understanding genetics, the underlying mechanisms and biology of human cardiomyocyte (dys)function, cardiac hypertrophy leading to heart failure may be essential to devise better (personalized) cardiovascular medicine.


32. UC Santa Cruz to lead effort to build a new map of human genetic variation. at <https://genomics.uoe.ucsc.edu/news/article/140>

Nederlandse genetica van cardiovasculaire kenmerken

Hart- en vaatziekten behoren wereldwijd tot één van de belangrijkste doodsoorzaken en worden een steeds groter probleem in de westelijke wereld door de alsnog ouder wordende populatie. Onze kennis over de pathofysiologie van menselijke hartcellen en over factoren die de gevoeligheid voor het krijgen van hartfalen bepalen is beperkt, wat het ontwikkelen van nieuwe therapeutische middelen bemoeilijkt. Bestaande therapieën voor hart- en vaatziekten zijn begrensd, in het bijzonder de behandeling van hartfalen, en er zijn in de afgelopen decennia weinig vernieuwende therapieën ontwikkeld. In tegenstelling tot de relatieve stagnatie van de ontwikkeling in hart- en vaatziektebehandelingen, is de kennis over het menselijke genoom explosief gestegen sinds het rond 2000 in kaart werd gebracht.

Rond 2006 zijn onderzoekers begonnen om systematisch uit te zoeken welke genetische variaties belangrijk zijn voor ziektes en karakteristieken (ook wel fenotypes genoemd) met behulp van genoomwijde associatie (GWA) scans. Menselijke eigenschappen zijn ontstaan door de evolutie heen, door middel van DNA veranderingen waardoor er genetische diversiteit ontstaat die ervaar zorgt dat elk individu uniek is. In de wereldwijd populatie zijn er miljoenen genetische variaties in het menselijke genoom en met behulp van een GWA scan is het mogelijk om te onderzoeken welke van deze variaties belangrijk zijn voor ziekten en eigenschappen. GWA scans hebben in het verleden al veel nieuwe biologische inzichten opgeleverd, genen geïdentificeerd die een doelwit zijn van al bestaande medicijnen, maar ook doelwitten aangewezen voor toekomstige middelen die ingezet kunnen worden, zoals toepassingen voor een te hoog cholesterol, type 2 diabetes, reuma en osteoporose. Het is echter lastig om associaties te vinden tussen genetische varianten en complexe ziektes zoals hartfalen omdat deze aandoening, net als veel andere hart- en vaat ziekten, vele oorzaken kent, waardoor het statistisch onderscheidingsvermogen relatief laag is. In dit proefschrift beschrijft ik het gebruik van zogenaamde 'endofenotypes' om nieuwe genetische varianten te identificeren die belangrijk zijn voor het hart- en vaatstelsel, in het bijzonder eigenschappen die gerelateerd zijn aan linkerlamerfunctie en cardiale...
geleiding. Endofenotypes liggen mogelijk dichter bij de functie van ons genoem dan de ziekten zelf, en kunnen in het ideale geval ook in gezonde mensen gemeten worden, wat meer statischonderscheidingsvermogen oplevert om nieuwe genetische varianten te vinden die belangrijk kunnen zijn voor het hart en vaatstelsel en ziektes zoals hartfalen. Het onderzoeken van cardiovasculaire risicofactoren stelt ons in staat om onze kennis van de biologie die onderliggend is aan hartfalen te vergroten en nieuwe aanknopingspunten te vinden voor nieuwe therapieën. Hoofdstuk 1 bevat een algemene introductie van dit proefschrift waarin het kader wordt gesteekt waarin de onderzoeksvragen worden gesteld. In de daarop volgende hoofdstukken worden verschillende genoomwijze associatie scans beschreven die ik, samen met vele andere onderzoekers, heb uitgevoerd.

Hoofdstuk 2 beschrijft vier genetische gebieden die we hebben geïdentificeerd in een GWA studie van op 2 peptides: midregional-proadrenomedullin en C-terminal-pro-endothelin-1, waarbij de eerste een vasodilator is en de laatste een vasococonstrictor. Deze stoffen zijn belangrijke voorspellers van hartfalen, linkerkamer hypertrofie en andere hart- en vaatziekten. Twee van de vier genetische gebieden KLKB1 en FXII, bleken geassocieerd te zijn met beide peptides en maken onderdeel uit van het kallikreine-kinine systeem, een belangrijk systeem dat de vaatdoorlaatbaarheid regelt. We hebben experimenteel aangetoond dat plasma kallikreine beide stoffen kan knippen, wat suggesteert dat dit gen ook betrokken is bij (in-)activatie van deze belangrijke peptides. KLKB1 en FXII kunnen belangrijke aanknopingspunten zijn voor toekomstige behandelingen van hart- en vaat ziektes, denk bijvoorbeeld aan een plasma kallikreine antagonist. Daarnaast hebben we ook een genetisch gebied gevonden dat specifiek is voor midregional-proadrenomedullin en een gebied dat specifiek is voor C-terminal-pro-endothelin-1. In Hoofdstuk 3 hebben we een GWA studie op galectin-3 uitgevoerd. Galectine-3 is een stof die gemeten wordt in het bloed waarvan verhoogde concentratie kan duiden op fibrosis van weefsel en verschillende ziektes, met name hart- en vaatziekten. De hoofdstukken 4, 5, 6 en 7 behandelen genetische studies naar electrocardiografische (ECG) parameters. Het doel van hoofdstuk 4 was om meer inzicht te krijgen in de genetische achtergrond van PR-interval. Eerdere genetische studies naar het PR-interval hebben meerdere gebieden op het genoom geïdentificeerd die belangrijk zijn voor een normale hartfunctie maar het is onduidelijk of deze de atria, de atrioventriculaire knoop of beide beïnvloeden. Het PR-interval beschrijft de geleiding van elektrische impulsen die door de boezems en atrioventriculaire knoop reizen, en kan worden opsplitst in P-top duur (atriale geleiding) en PR-segment duratie (geleiding door m.n de atrioventriculaire-knoop). Door te focussen op deze twee parameters zien we dat de meeste genetische varianten die geassocieerd zijn met het PR-interval hoofdzakelijk worden beïnvloed door PR-segment duur. Daarnaast hebben we ook nieuwe genetische varianten gevonden die specifiek zijn voor P-top duur en dus belangrijk kunnen zijn voor atriale biologie. Ook hebben we aangetoond dat de varianten die we hebben geïdentificeerd zijn verbreed in enigen om in-vito genetische factoren in cardiale weefels die duiden op activatie van genexpressie in het hart. De genen die bij de geassocieerde varianten laten kwamen hoger tot expressie in het atriale weefsel en de atrioventriculaire knoop dan de linkerkamer. In Hoofdstuk 5 hebben we gekeken naar de genetische achtergrond van cardiale repolarisatie zoals die wordt weergegeven door de ST-T golf op het elektrocardiogram. Voor deze analyse zijn in totaal 10 genotypes gedefinieerd tijdens de ST-T golf: 5 die de voltages tijdens het ST-segment beschrijven en 5 genotypes die voltages van de T-top beschrijven. De GWA studie is uitgevoerd binnen Prevent en Lifelines met externe validatie in onafhankelijke cohort studies met in totaal 37,977 individuen. Deze analyses hebben 28 genetische gebieden geïdentificeerd die relevant zijn voor cardiale biologie, waaronder ion kanalen, structuele eiwitten en andere genen die een rol spelen tijdens de ontwikkeling van het hart. We zagen verschillen in associatie tussen genetische varianten die waren geassocieerd met het ST gedeelte of de T-top, bijvoorbeeld kalium kana len zijn specifiek geassocieerd met voltages tijdens het ST-segment en niet met voltages van de T-top. Hoofdstukken 6 en 7 beschrijven beide de genetische achtergrond van het QRS-complex, zoals gemeten op het ECG. Een belangrijk verschijnsel bij hartfalen en andere hart- en vaat ziekten is cardiale hypertroefie, wat ontstaat door verdikking van de hartspier als gevolg van overbelasting. Het QRS-complex op het ECG wordt veroorzaakt door samentrekking van beide kamers en geeft nauwkeurig de grootte en massa van het hart weer. Om beter te begrijpen welke genetische factoren en biologische processen deze regelen hebben wij een GWA studie uitgevoerd in meer dan 55-duizend mensen van Europese afkomst, zoals beschreven in hoofdstuk 6. Daarnaast hebben we in vitro, in vivo en in silico analyses om de relevantie aan te tonen van de gevonden associaties en genetische gebieden. De resultaten van deze studie geven nieuwe biologische inzichten in cardiale hypertroefie. In Hoofdstuk 7 is de analyse op het QRS complex uitgebreid door te focussen op 27,039 mensen van overwegend Nederlandse afkomst. Door middel van GoNL, een genetische referentie set die specifiek is voor Nederlanders, konden we de associaties uit Hoofdstuk 6 gedetailleerder besturen en hebben we 8 nieuwe genotype-fenotype associaties geïdentificeerd.

Samenvattend kan gesteld worden dat de kennis van genetische varianten betrokken bij cardiovasculaire risicofactoren in de afgelopen jaren sterk is toegenomen. Deze bevindingen leveren niet alleen een belangrijke
Niek Verweij (1985) was born in Groningen. After finishing his secondary education (VWO, Zernike College) in 2003, Niek started to study Pharmacy and Medical Pharmaceutical Sciences at the University of Groningen. He carried out scientific research projects in diverse departments: pharmaceutical biology (RUG/FU, Berlin), toxicology/analytical chemistry (RUG), pharmaceutical technology (RUG) and at the department of pharmaceutical technology (RUG). For a short while, Niek continued working in the field of pharmaceutical technology as R&D scientist at Innocore Technologies, a company that develops drug delivery systems.

In 2011, Niek started his doctoral research at the Department of Cardiology (UMCG) where he was able to combine his interest in research with his affection for computers and programming. Niek has finished and initiated several research projects during the past 4 years, all related to one central objective: to identify genetic variants that influence endophenotypes of cardiovascular disease with special interest for electrocardiographic traits that reflect cardiac conduction, myocardial mass and myocardial repolarization. He complemented genetic associations by establishing biological mechanisms underlying the association using experimental work and computational biology approaches. As part of (inter)national consortia and groups he was responsible for the genetic analyses in the Prevend cohort (all traits) and Lifelines (electrocardiographic traits). In several consortia he was part of the core-analysis team providing bioinformatics support and drafting manuscripts, resulting in shared-first author publications in high tier journals.

Niek defends his dissertation titled “Genetics of Human Cardiovascular Traits” on May 27th 2015 in Groningen at the RUG. In June 2015 he will join the groups of Christopher Newton-Cheh (Massachusetts General Hospital of Harvard Medical School & Broad), and Laurie Boyer (Massachusetts Institute of Technology, Biology dept.) for one year to continue his research.


Submitted Manuscripts:


8. Alexander Teumer, Adrienne Tin, Rossella Sorice, Mathias Gorski, Nan Cher Yeo, … N Verweij, et al. “Genetic Variation of RAB38 and H57ST1 and Albuminuria in Diabetes”


*These authors contributed equally to the work.
I would also like to thank all individuals and families participating in the cohort studies.
Colophon

Niek Verweij
Genetics of human cardiovascular traits
PhD thesis – Department of Cardiology
University of Groningen, University Medical Center Groningen

Layout & Design
Groningen, Ramon Putman

Printing
Drukkerij de Marge


© Copyright 2015: Niek Verweij. All rights reserved.

No part of this book may be reproduced, stored in retrieval system, or transmitted in any form of by any means, without prior permission of the author.

Printing of this thesis was financially supported by: University of Groningen, Genzyme, Graduate School of Medical Sciences (GSMS), the Dutch Heart Foundation.