DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
186

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
186

KATRIN KEPP

Genes involved in cardiovascular traits: detection of genetic variation in Estonian and Czech populations

TARTU UNIVERSITY
PRESS

Institute of Molecular and Cell Biology, University of Tartu, Estonia

Dissertation is accepted for the commencement of the degree of Doctor of Philosphy (in molecular diacgnostics) on 05.10.2010 by the Council of the Institute of Molecular and Cell Biology, University of Tartu

Supervisor:	Prof. Maris Laan, PhD Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Estonia
Opponent:	Prof. Kimmo Kontula, MD, PhD, Department of Medicine, Vice-Rector of University of Helsinki; Chief Physician, Unit of Internal Medicine, Helsinki University Central Hospital, Finland
Commencement:	Room No 217, 23 Riia Str., Tartu, on November 19th 2010, at 12.00

The publication of this dissertation is granted by the University of Tartu

ISSN 1024–6479
ISBN 978–9949–19–490–2 (trükis)
ISBN 978–9949–19–491–9 (PDF)

Tartu Ülikooli Kirjastus
www.tyk.ee
Tellimus nr 597

TABLE OF CONTENTS

LIST OF ORGINAL PUBLICATIONS

This thesis is based on the following original articles referred to in the text by their Roman numerals:

I Sõber S, Org E, **Kepp K**, Juhanson P, Eyheramendy S, Gieger C, Lichtner P, Klopp N, Veldre G, Viigimaa M, Döring A; Kooperative Gesundheitsforschung in der Region Augsburg Study, Putku M, Kelgo P; HYPertension in ESTonia Study, Shaw-Hawkins S, Howard P, Onipinla A, Dobson RJ, Newhouse SJ, Brown M, Dominiczak A, Connell J, Samani N, Farrall M; MRC British Genetics of Hypertension Study, Caulfield MJ, Munroe PB, Illig T, Wichmann HE, Meitinger T, Laan M. Targeting 160 candidate genes for blood pressure regulation with a genome-wide genotyping array. PLoS One. 2009; 4(6), e6034.

II **Kepp K**, Juhanson P, Kozich V, Ots M, Viigimaa M, Laan M. Resequencing *PNMT* in European hypertensive and normotensive individuals: no common susceptibilily variants for hypertension and purifying selection on intron 1. BMC Medical Genetics. 2007 Jul 23; 8:47.

III Juhanson P, **Kepp K,** Org E, Veldre G, Kelgo P, Rosenberg M, Viigimaa M, Laan M. *N-acetyltransferase 8*, a positional candidate for blood pressure and renal regulation: resequencing, association and in silico study. BMC Medical Genetics. 2008 Apr 10; 9:25.

IV **Kepp K**, Org E, Sõber S, Kelgo P, Viigimaa M, Veldre G, Tõnisson N, Juhanson P, Putku M, Kindmark A, Kozich V, Laan M. Hypervariable intronic region in *NCX1* is enriched in short insertion-deletion polymorphisms and showed association with cardiovascular traits. BMC Medical Genetics. 2010 Jan 28; 11:15.

Author's contributions:

Ref. I – performed *in silico* candidate gene selection and contributed to manuscript preparation

Ref. II – participated in experimental design, conducted the experiments, analyzed the data and contributed in manuscript preparation

Ref. III – participated in experimental design, in re-sequencing analysis of promoters of target genes and contributed in manuscript preparation

Ref. IV – participated in experimental design and conducted majority of the experiments, analyzed the data and wrote the first draft of the manuscript

LIST OF ABBREVIATIONS

bp – base pair
BP – blood pressure
CAD – Coronary Artery Disease
CADCZ – Coronary Artery Disease in CZech
CRP – C-Reactive Protein
CVD – Cardiovascular Disease
CD-CV – Common Disease-Common Variant
DBP – Diastolic Blood Pressure
dsDNA – double stranded genomic DNA
EH – Essential Hypertension
GRE – Glucocorticoid Responsive Element
GRU – Glucocorticoid Responsive Unit
GWAS – Genome-Wide Association Study
HDL – High Density Lipoprotein
HR – Heart Rate
HYPEST – HYPertension in ESTonia
indel – insertion/deletion variation
IMT – Intima Media Thickness
kb – kilo base (1000 base pairs)
LD – Linkage Disequilibrium
LDL – Low Density Lipoprotein
MAF – Minor Allele Frequency
MetS – Metabolic Syndrome
MI – Myocardial Infarction
NCX1 – Na^{+}/Ca^{2+} exchanger; sodium/calcium exchanger 1
PNMT – Phenylethanolamine-N-MethylTransferase
QTL – Quantitative Trait Locus
SBP – Systolic Blood Pressure
SNP – Single Nucleotide Polymorphism
TC – Total Cholesterol
TFBS – Transcription Factor Binding Sites
TG – Triglycerides
WNK1 – lysine deficient protein kinase 1 (With No lysine Kinase 1)
WNK4 – lysine deficient protein kinase 4 (With No lysine Kinase 4)

INTRODUCTION

Cardiovascular diseases (CVD) are known as one of the main causes behind the premature deaths and disability in Western societies. CVD is a complex trait influenced by the interplay of multiple genes as well as controllable (stress, salt and alcohol consumption, diet, weight and exercise) and uncontrollable (age, sex, family history, and ethnic background) including environmental determinants. Many genetic loci are known to be involved in the etiology of CVD causing both monogenic and complex forms of the disease. The most commonly known risk factor leading to the development of stroke, myocardial infarction and congestive heart disease, heart failure, peripheral vascular disease and end-stage renal disease is the adverse changes of blood pressure levels (European Society of Hypertension-European Society of Cardiology, 2003). Studies among multiple families and twin cohorts propose that the approximate inheritance of CVD is diverse ranging from 17% up to 66% (Fischer, et al., 2005; Levy, et al., 2007a; Zdravkovic, et al., 2002; Tobin, et al., 2005a; Wienke, et al., 2001).

Despite of the knowledge of multiple disease susceptibility loci leading to CVD many loci are still yet to be described. Recent advances in mapping and genotyping DNA variants together with a better understanding of genetic variations in individuals as well as in populations has broadened the studies of human disease. The success in finding genes underlying human genetic diseases is greatly dependent on the used markers and their physiological location in the genome. Variations within coding regions are known to have a high impact on causing defects directly on gene expression level compared to the rest of the genomic sequence. In addition to protein coding sequence several other genomic regions like promoter (Jin, et al., 2008; Liu, 2010 ; Yan, et al., 2010) and conserved non-coding regions have been shown to have also a remarkable role in gene expression processes (Dermitzakis, et al., 2005; Paul, et al., 2007; Reid, et al., 1990; Rowntree, et al., 2001). Genetic analysis based on linkage and association studies have revealed close to 200 genomic risk loci each with a potential small effect on CVD development (Adeyemo, et al., 2009; Arking and Chakravarti, 2009; Harrap, et al., 2002; Org, et al., 2009; Samani, et al., 2007; Wang, 2005; Wang, et al., 2009).

In this thesis, the literature review gives an overview of the most commonly used genetic markers and approaches in mapping human genetic diseases, description of the importance of selection of genomic regions in genetic studies and the heritability of cardiovascular diseases focusing on hypertension and coronary artery disease.

The experimental part of current study investigates the genetic variation pattern within candidate genes known to have an impact on cardiovascular disease susceptibility. This research included *in silico* CVD candidate gene selection; variation detection within the human CVD candidate gene *PNMT* (phenylethanolamine-N-methyltransferase) and in regulatory regions like

3

promoters and conserved noncoding regions (CNRs) of CVD candidate genes; evaluation of the association of detected variations in *PNMT* and *NCX1* (Na+/Ca2+ exchanger; sodium/calcium exchanger 1) genes with the prevalence of essential hypertension and coronary artery disease as well as with serum lipids and BP in two Eastern-European populations.

I. REVIEW OF LITERATURE

I.1. Human genetic variations and their contribution to human disease

During the past decade, a great knowledge has been achieved by studying human genome and the nature of its genetic variability (Lander, et al., 2001; Venter, et al., 2001). Revealing the complete sequences of four individual diploid genomes (Bentley, et al., 2008; Levy, et al., 2007b; Wang, et al., 2008; Wheeler, et al., 2008) has given even more insight about the number of existing forms of genetic variations and their evolutionary background as well as susceptibility to human diseases. Analyses of human genetic variations in phenotypic differences have become one of the central efforts to understand the function of the genes and genetic variants in predisposition to disease development.

I.1.1. Genetic variation in the human genome

I.1.1.1. Diversity and distribution of human genetic variants

Human genome is estimated to consist of ~3 billion base pairs (bp) and code over 20 000–25 000 distinct protein coding genes (International Human Genome Sequencing Consortium 2005). Comparison of two different human genomes has been shown to exhibit high similarity (99.9%) to one another differing only by 0.1% in DNA level. These differences are mostly represented by the natural genetic variations and are used as markers in dissection of the genetic basis of human disease. Genetic variants in the human genome can be divided into two different nucleotide composition classes: single nucleotide variants and structural variants (Table 1) (Eichler, et al., 2007; Frazer, et al., 2009). First class includes variants where only a single base in DNA sequence (A, T, G, or C) is altered. Second class, structural variants, occur when one or more base pairs vary compared to other genomes resulting with changes in DNA length caused by the insertion, inversion, deletion or duplication events of DNA segment. Structural changes are generally composed of few bases up to 80kb in length (Levy, et al., 2007b). During the last years a new type of variations called copy number variations (CNV) were discovered. CNVs are classically defined as DNA segments greater than 1 kilo base pairs (kb) and have been shown to play an important role in human diseases like autism spectrum disorder (Sebat, et al., 2007) and schizophrenia (Walsh, et al., 2008). CNVs have also been suggested as a major cause that may lead to structural variation formation involving both duplications and deletions in DNA sequence (Kidd, et al., 2008).

First genetic markers used in studying human genetic disease were restriction fragment polymorphisms (RFLPs), representing single nucleotide variance in restriction enzyme recognition site (Chakravarti, 1984). Further, microsatellites, like short tandem repeats (STR), composing of 2–6bp repeat motifs and minisatellites (variable number of tandem repeats – VNTR) were used in linkage analysis to identify disease loci within families (Walsh, et al., 2008). Today the most widely used variations in genetic studies are single nucleotide polymorphisms (SNPs) and are estimated to represent over 90% of all genetic variants in the genome (Wang, et al., 1998). At least 11 million SNPs are proposed to be located in the human genome where approximately 7 million are suggested to be with minor allele frequency (MAF) of 5% (Kruglyak and Nickerson, 2001). The vast majority of SNP are bi-allelic, although tri- and tetra-allelic forms can also be found. Due to the binary property, SNPs are widely used in association studies in automated high-throughput genotyping technologies (Wang, et al., 1998). Furthermore, there are large-scale genotyping platforms available such as Illumina 1M and Affymetrix 6.0 arrays, each representing over one million SNPs mapped in the human genome. Based on the high number of SNPs in the genome, several public recourses (i.e. Haplotype Map (HapMap) Project, etc.) have been built up to provide researchers with the information of their frequency estimates among different populations (2005; Frazer, et al., 2007; Lander, et al., 2001; Venter, et al., 2001).

Table 1. Examples of classes of human genetic variants modified from (Frazer, et al., 2009)

Variation type	**Example**
Single nucleotide variants:	ATTGGCCTTAACC**C**CCGATTATCAGGAT ATTGGCCTTAACC**T**CCGATTATCAGGAT
Structural variants:	
Insertion-deletion variant	ATTGGCCTTAACCC**GAT**CCGATTATCAGGAT ATTGGCCTTAACCC - - - CCGATTATCAGGAT
Block substitution	ATTGGCCTTAAC**CCCC**GATTATCAGGAT ATTGGCCTTAAC**AGTG**GATTATCAGGAT
Inversion variant	ATTGGCCTT**AACCCCCG**ATTATCAGGAT ATTGGCCTT**CGGGGGTT**ATTATCAGGAT
Copy number variant	ATT**GGCCTTAGGCCTTA**ACCCCCGATTATCAGGAT ATT**GGCCTTA** - - - - - - - ACCTCCGATTATCAGGAT

1.1.1.2. Small indels and their role in human complex traits

Insertion and deletion polymorphisms (indels) are considered as structural variations being less than 1kb in size (Scherer, et al., 2007). Small indels are estimated to be the second most frequent variation type in human genome and their occurrence may lead to individual phenotypic differences. Recently, more attention has been paid to the systematic study on mapping indels in the human genome. Comparison of chromosome 22 in chimpanzee and human chromosome 21 has revealed nearly 68 000 indels represented mostly by small variants with the exception of number of larger ones sized up to 54 000bp (Watanabe, et al., 2004). It has been estimated that human genome may harbour approximately 1.5 million indels (Mills, et al., 2006). Re-sequencing of 330 genes with known biological role in lipid metabolism and DNA repair among diverse humans revealed over 2000 small heterozygous indels with length range between 1bp to 543bp (Bhangale, et al., 2006). Mostly because of the detection inaccuracy the diversity of indels between individuals is still questioned. For example, sequencing of three personal genomes discovered different number of indels in the Venter genome (823 396 indels) (Levy, et al., 2007b), in the Watsons genome (22 718 indels) (Wheeler, et al., 2008) and in the Han Chinese genome (135 262 indels) (Wang, et al., 2008) that vary significantly.

There are several classes of indels known to exist in the human genome (Table 2) where majority of indels are single base pair deletions or insertions followed by monomeric or multibase repeat expansions. Remaining variations are described either as transposon insertions or indels containing a random DNA sequence.

Tabel 2. Classification of indels in the human genome. Modified from (Mills, et al., 2006).

Indel class	**Examples**
Single bases	A; T (most common forms); C; G
Repeat expansions:	
Monomeric	(A)n; (T)n; (C)n; (G)n
Dimeric	(AC)n; (GT)n; (TG)n; (CA)n; (TA)n; (AT)n; (CT)n; (AG)n; (GA)n; (TC)n; (GC)n; (CG)n
Trimeric	(AAT)n; (TTA)n; (ATT)n; (TAA)n; (AAG)n; (TTC)n; (TAT)n; (AAC)n; (ATA)n; (TTG)n; (CAA)n; other (NNN)n
Tetrameric to Decameric	
Transposon insertions	
Other	Indels containing random DNA sequence (~99%); mostly <100bp

If an indel occur in the protein-coding region it tends to be more deleterious than SNPs. This is mainly because of their size and chance to alter amino acids and affect the protein production. Therefore, the frequency of small indels in coding areas has been correlated with the genes with lower selection pressure

(Chen, et al., 2007; de la Chaux, et al., 2007). There are several indels known to cause human genetic disease (Table 3), mostly locating in the coding region as well as in regulatory regions of genes (such as promoters). The best described example causing human disease is a 3bp in-frame deletion (ΔF508) in cystic fibrosis transmembrane conductance regulator gene (*CFTR*) leading to the autosomal-recessive condition called cystic fibrosis (Saleheen and Frossard, 2008).

Table 3. Examples of indel variations leading to human genetic disease.

Gene	Location	Indel type	Disease	Reference
FMR1	Xq27.3	(CGG)n repeat expansion	Fragile X mental retardation syndrome	(Penagarikano, et al., 2007)
HTT	4p16.3	CAG repeat	Huntington disease	(Aziz, et al., 2009)
MYBPC3	11p11.2, gene encoding region	25bp deletion	Heritable cardiomyopathies and an increased risk of heart failure in Indian populations	(Dhandapany, et al., 2009)
CCR5	3p21.31	32bp deletion	incomplete HIV-1 resistance, late onset of Schizophrenia	(Rasmussen, et al., 2006; Sheppard, et al., 2002)
PAX8	2q12-q14, exon 7	ACCC deletion (leads to premature stop codon)	Thyroid dysfunction	(de Sanctis, et al., 2004)
GPIBA	17p12, gene encoding region	27bp deletion	Severe bleeding disorder, platelet-type von Willebrand's disease	(Othman, et al., 2005)
MMP1	11q22.3, promoter region	G nucleotide insertion	Cancer (ovarian, lung, colorectal)	(Rutter, et al., 1998)
NFκB	4q24, promoter	ATGG indel	Inflammatory bowel disease ulcerative colitis	(Karban, et al., 2004; Lewander, et al., 2007)
BRCA1	17q21, exon 11	5bp deletion; 4bp insertion, etc.	breast and/or ovarian cancer	(Presneau, et al., 1998)
CFTR	7q31.2	8108bp deletion (exon 2); 182bp insertion (intron 1), etc.	susceptibility to cystic fibrosis	(Faa, et al., 2006)

1.1.2. Approaches in mapping human disease

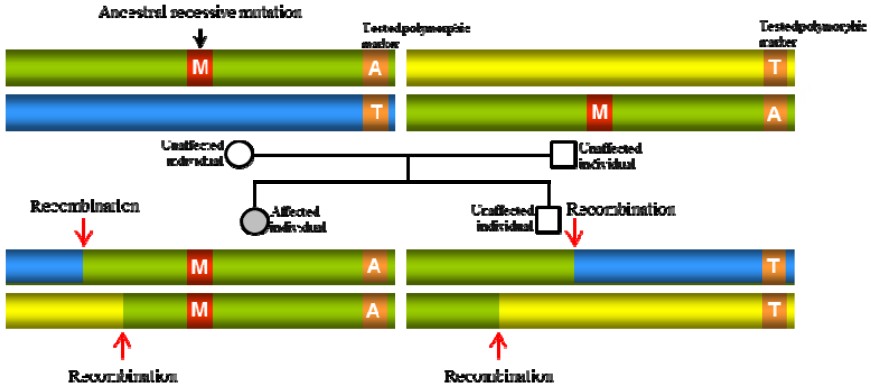
1.1.2.1. Linkage and association mapping

Majority of genetic variations in the genome are acting neutral as they are located mostly in non-coding regions with no major effect on biological traits. Alterations in DNA sequences which affect individual's phenotype usually locate to the protein-coding or regulatory sequences. These modifications may possibly have an impact on a protein's functional performance up to the loss of function.

In order to map genetic variations behind human diseases two main methodological approaches can be used: linkage and association studies (Borecki and Suarez, 2001). In linkage studies the responsible trait loci are assumed to cosegregate from parents to offspring with polymorphic markers at a specific chromosomal region. This is based on the assumption that two loci are physically closely linked (Agarwal, et al., 2005) and because of meiotic recombination a marker that is showing segregation with the trait must be nearby in the genome (Altshuler, et al., 2008) (Figure 1 A). During the evolution cosegregation of two loci separated by longer distances might be broken up by recombination. Linkage mapping is a powerful tool to identify preferentially rare high-risk alleles contributing to the disease susceptibility. To measure the significance of linkage, the logarithm of the odds/lod score is used (Morton, 1955) to describe the recombination fraction between a genetic marker and disease locus in terms of likelihood ratio. This is based on the null hypothesis assuming no linkage between the marker and disease loci.

Association studies are based on a statistical correlation between a specific genetic variation and a trait variation among sample of individuals (Risch and Merikangas, 1996). Because of the effect of the locus variant on the trait variant, this approach enables to measure actual causal risk factor (Borecki and Suarez, 2001). Compared to the linkage analysis an association occurs in short physical distances in the genome. To detect a positive association a large number of common (polymorphic) genetic markers or a combination of markers (haplotype) are required where each contribute with the moderate effects to the disease susceptibility (Figure 1 B**)**. The statistical evidence of association between an allele and a phenotype may arise from the potential variant leading directly to the disease phenotype and is correlated with or is in linkage disequilibrium (LD) with the nearby causal allele. LD is termed as the non-random association between the alleles of different loci (Weiss and Clark, 2002). Usually association studies are based on analysis of unrelated affected (cases) and unrelated unaffected (control) individuals in the population (Cardon and Bell, 2001). For example, if the prevalence of the allele is more frequent in the cases compared to controls it will have high probability of being associated with the diseases susceptibility. In genetic analysis both the linkage as well

A. Linkage analysis



B. Association analysis

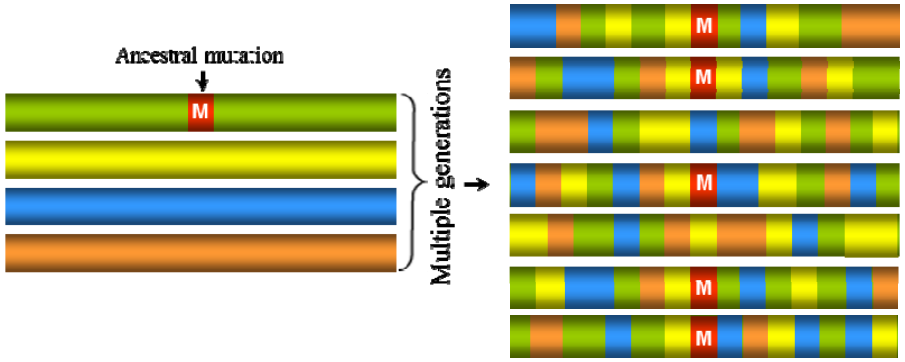


Figure 1. A study of segregation of the mutation (M) (ancestral) and a polymorphic marker (A/T) with the disease from parents to offspring is called linkage analysis **(A).** In this example a recessive mutation is segregating with the polymorphic marker allele A. Affected individual in pedigree is marked in gray. Red arrows refer to the recombination events during the segregation. In association analysis, the causal mutation along with genetic markers is segregating through multiple generations among randomly mating individuals within population **(B)**. Different colors indicate to the chromosomal regions segregating during the generations. Modified from (Cardon and Bell, 2001).

as the association studies may be used as complementary approaches to each other (Hodge, 1993). to study complex genetic traits classically a candidate-gene based association approach is used. This is based on testing the hypothesis that specific gene(s) are associated with the disease risk (Jorgensen, et al., 2009). The availability of high-throughput technologies combining with the cataloging of common human genetic variants (like in dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/) or HapMap (http://www.hapmap.org/cgi-perl/gbrowse/gbrowse)) have promoted a novel hypothesis-free, genome-wide association strategy (GWAS) to identify novel genetic contributors statistically linked to the disease (Frazer, et al., 2007).

1.1.2.2. Critical aspects in design of an association study

Despite of the great effort on technological and theoretical methods the genetic background of human complex diseases is still largely unknown and remains to be challenging. Considering the genetic association studies there are multiple crucial aspects that may arise in studying the disease of interest: (a) sample collection of properly defined phenotype and population of origin; (b) selection of genetic markers and methodologies to describe and to test the presence of an association (c) genetic heterogeneity, where a similar phenotype is caused by different loci or allelic variants; (d) pleiotropy, where one gene may affect many traits simultaneously (Altshuler, et al., 2008; Cardon and Bell, 2001; Hirschhorn, et al., 2002). Compared to family-based studies, population-based association studies of unrelated individuals could provide a more practical and powerful tool for the detection of genetic loci related to complex diseases. However, several problems have arisen from insufficient knowledge of demographic history and population stratification, which can lead to biased or spurious results. Therefore, it is essential to explore the unique genetic and social history of the populations in detail to assure reliable results to be obtained from the population-based association studies (Cardon and Palmer, 2003; Pritchard and Rosenberg, 1999).

During the past years in GWAS a number of novel genetic factors linked to the human diseases and continuous traits of biomedical importance have been revealed (Hindorff, et al., 2009; Johnson and O'Donnell, 2009). GWA approach has a great advantage to detect genetic variations throughout the human genome with no prior knowledge of genes or regions with unknown biological pathways potentially linked to the disease of interest. The limitation of GWA studies is that they are classically based on common-disease common-variant (CD-CV) hypothesis. This theory proposed that common disease-causing alleles are present in all humans leading to the complex disease excluding the role of rare variants. GWA results have explained only a small fraction of the burden human diseases among the entire population (http://genome.gov/gwastudies/). This suggests that common inherited variations are not likely to explain the

majority of common human disease but rather by the remaining genetic variations (for example rare or other forms of genetic (for example epigenetic variations) or by their combined interactions in the genome level (Schork, et al., 2009). A moderate success of GWAS may also be a result of an inadequate coverage of genetic variations available in commercial genotyping arrays mostly built up of SNPs tagged through LD from the HapMap collection and comprise less than a half of the SNPs available in dbSNP database (Ku, et al., 2010). Moreover these platforms do not comprise structural variations like small insertions, deletions and CNVs identified in the human genome.

1.2. Regulatory regions in the human genome

Regulatory regions are DNA sequences where gene regulatory proteins potentially bind and have a control over the gene expression processes. It has been suggested that complex traits may result from noncoding regulatory variants rather than coding variants altering protein structure as it is common for monogenic diseases. Still, non-coding regulatory regions have been often questioned because of their undefined role in the genome function and therefore further investigation is needed (Glazier, et al., 2002).

1.2.1. Promoter regions as potential affectors of human disease susceptibility

The most known regulatory DNA sequences are promoter regions where the transcriptional process is regulated. Promoters are located upstream of a gene and are composed of specific DNA segments termed as core and proximal promoter regions. Core promoter is a minimal DNA segment surrounding the transcription start site (70–80bp) sufficient for initiation of transcription. Proximal promoter (up to –500bp according to ATG site) is composed of sequences needed for binding of transcription regulatory factors to enforce the transcription (Butler and Kadonaga, 2002). The exact length of the true promoter region is gene-specific and can often be defined experimentally, Promoters may locate a few hundred base pairs directly upstream of the site of initiation of transcription or even as far 30–40kb from mRNA start site (Pedersen, et al., 1999).

In addition, regulatory elements like enhancers or silencers can also enhance or repress transcription upon interacting with transcription factors. These regions may range from 60bp up to 900bp, locate anywhere in the genome and stimulate or repress the transcription process. In genetic linkage and association studies several *cis*- and *trans*-acting DNA variants have been identified that potentially influence expression levels of human genes. Different alleles located

of the *cis*- and *trans*-acting variants may have various influences on gene expression profile (Cheung and Spielman, 2009) (Figure 2).

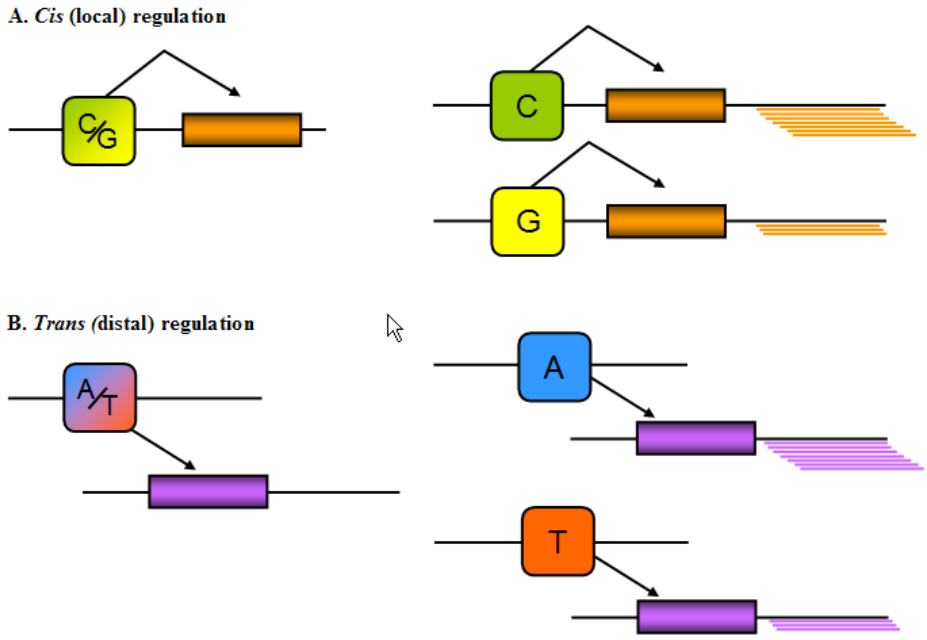


Figure 2. Effects of *cis*- and *trans*-acting DNA variants on different expression levels of genes. Polymorphic forms of regulators that act in *cis* (local) **(A)** or in *trans* (distal) **(B)** to the target gene may lead to the lower or higher expression levels of the gene. Modified from (Cheung and Spielman, 2009).

As indicated in Figure 2A, subjects with the C variant of the *cis* regulator (like promoter region) have a higher expression level of the target gene than individuals possessing the G variant. Similarly, individuals with the A variant of the *trans* regulator (usually a DNA sequence that codes for a gene) have a higher expression level of the target gene compared to the T variant (Figure 2B) (Cheung and Spielman, 2009).

There are multiple examples where promoter variations have been shown to be associated with the increased or decreased risk of human diseases. An association study of human genetic variant (–1535C>T) located in the promoter region of *visfatin* gene have been shown to decrease the risk of CAD up to 40% in CT+TT genotype carriers compared to CC genotype carriers in Chinese population (Yan, et al., 2010). Among HIV-infected and *CCR5Δ32* homozygote individuals, a *CCR5Δ32* 59537-G/A promoter polymorphism possessing 59537-A/A genotype have been found to be strongly associated with the low translational efficiency of the CCR5Δ32 protein and with the loss of CCR5Δ32

protective effects. The results provide an important insight into the mechanism of resistance to HIV-1 infection and the pathogenesis (Jin, et al., 2008). Examples of variations in promoter regions and their allelic combinations have been shown to exist also between hypertension and promoter activity of several human CVD candidate genes. In human *SCNN1A*, G allele of the regulatory SNP (accession no. AF060910) have been shown higher promoter activity compared to the A allele leading to the elevated the risk of hypertension (Iwai, et al., 2002). Similarly, A/G polymorphism at −217 position of the *AGT* gene showed higher promoter activity in the case of A allele resulting to the higher binding affinity of transcriptional complex and increased expression level (Jain, et al., 2005). Studies of the fertility in men have identified a G/T SNP (rs10835638; –211 G/T) located in transcription start site of highly conserved promoter region upstream of human *FSHB* (Follicle Stimulating Hormone (FSH) beta) gene. The T allele of the SNP has been indicated to decrease the mRNA production leading to decreased hormone formation in men and is enriched among male partners of infertile couples (Grigorova, et al., 2008; Grigorova, et al., 2009).

Regulatory sequences that might affect gene expression level may also be found in introns. Multiple intronic regions have been identified possessing regulatory regions leading to the different gene expression profiles. An example of combined activity of the intronic site has been shown to be present in CFTR gene expression levels, where the combination of different transcription factors and cell differentiation and proliferation modifiers has an influence on the fate of cell-specific expression (Paul, et al., 2007). Also, an alternative regulatory element in the intron 1 of *CFTR* gene has shown to increase the intestinal expression level *in vivo* (Rowntree, et al., 2001). Similarily, sequence elements in *HPRT* gene have been identified both in the first and second intron of the gene exhibiting an impact on expression in embryonic stem cells in humans (Reid, et al., 1990).

Approximately 58% of mammalian genes are known to possess alternative promoters and transcription start sites. Majority of genes harbour at least two up to more than 20 alternative promoters (Carninci, et al., 2006). In a single gene, multiple promoter regions may contain different transcription start sites and are potentially related to tissue-specific gene expression. Genome-wide analyses of human microarray data have provided evidence that alternative promoter are positively associated with differential expression and disease susceptibility (Liu, 2010). This kind of high complexity of promoter regions in gene expression profiles makes these regulatory units challenging to study in the etiology of human disease.

1.2.2. Conserved non-coding regions in human disease development

Comparative genomics used in genetic studies has offered a great opportunity to detect and to follow how genomes have been changing throughout the evolution. One of the aims using comparative genomics is to detect evolutionary conserved and functionally relevant sequence elements in the genome. Evolutionarily conserved DNA sequences are classically termed as DNA sequences exhibiting >70% identity over at least 100bp of ungapped alignment of human compared to mouse DNA (Dermitzakis, et al., 2002; DeSilva, et al., 2002). There are also ultra-conserved regions (UCR) existing in the genome defined as DNA segments with 100% sequence identity (spanning longer than 200bp) compared to rodents (Baira, et al., 2008; Bejerano, et al., 2004).

For example comparing human chromosome 21 and mouse syntenic sequences (segments of mouse chromosomes 10, 16 and 17) have revealed numerous highly conserved regions (CNR). Furthermore, additional support for conservation has been observed with other placental mammals like dog (Dermitzakis, et al., 2003). Majority of these CNRs reside in non-coding sequences, outside as well as within introns of the gene regions and comprise approximately 1–2% of the human genome (Dermitzakis, et al., 2005). There are over 327 000 CNR regions estimated to exist in human genome compared to mouse genome, where 65% are proposed as intergenic and 35% are intronic sequences (Giardine, et al., 2003). In recent years it has been shown that 3.5% of noncoding DNA sequence is substantially conserved across diverse mammals where also some regions have conservation with even more distant vertebrates like chicken and fish (Bejerano, et al., 2004; Woolfe, et al., 2005). Evolutionary analyses have also suggested that CNRs might be regulatory and have functional role in the genome. Respect to their function, identified blocks may be exons of unknown genes, non-coding RNAs, *cis*-regulatory regions, or functional sequences of unknown significance (Dermitzakis, et al., 2002), and also include areas that are neutrally evolving with the low substitution rate (Dermitzakis, et al., 2005). Mutations or rearrangements in such conserved non-coding regions can lead to the disruption of regulatory elements and direct phenotypic effects (Rossi, et al., 2008).

1.3. Heritability of cardiovascular disease (CVD)

1.3.1. Challenges in mapping the genetic component of cardiovascular diseases

Cardiovascular disease (CVD) is known to be associated with the alterations in heart, metabolism and blood vasculature and is a heterogeneous trait with the global impact on human morbidity and mortality. CVD development involves complex interplay of environmental and genetic risk factors leading to the difference between patients and population groups. There are several modifiable risk factors, mainly environmental: smoking, diet, exercise; and unmodifiable risk factors, for example age, sex and birth weight that contribute to CVD phenotype (Table 4.) (Goldstein, et al., 2006).

Table 4. Examples of physiological, metabolic and environmental risk factors for CVD.

Risk factors:		
Unmodifiable	**Metabolic**	**Modifiable: environmental or lifestyle**
Gender	Total cholesterol (TC) ↑	Smoking
Age	LDL-cholesterol (>2.8 mmol/l)	Diet
Birth weight	HDL-cholesterol (<1.0mmol/l)	Exercise
Genetic background	Total triglycerides (TG) (>1.7 mmol/l)	Infection
	Obesity	Fetal environment
	Systolic blood pressure levels (SBP) ↑↓	Air pollution
	Diastolic blood pressure levels (DBP) ↑↓	Alcohol consumption
	Lp(a) levels ↑ *	
	Homocysteine levels ↑	
	Fibrinogen levels ↑	
	C-reactive protein (CRP) level ↑	

*Lp(a) – Lipoprotein(a)

Some of the factors are modifiable by adjusting personal dietary habits. Lifestyle changes like exercise, smoking and alcohol consumption may also increase the risk for CVD. To date, there is limited information of the role of genetic risk factors underlying the complexity of cardiovascular diseases. In family studies mutations in single genes have been shown to contribute to severe CVD phenotype, like coronary artery disease and essential hypertension (Table 5 and Table 6). Majority of cardiovascular diseases reflect multiple components of larger number of genes each imparting with small genetic effect to the disease risk (Arking and Chakravarti, 2009; Dominiczak, et al., 2005). Contribution of several pathways to the progression and emerging of CVD development have been indicated. For example, lipids, inflammation, oxidative

stress, renin-angiotensin system, vascular remodeling, sympathetic nervous system, electrolyte and sodium homeostasis pathways (Delles, et al., 2009; Dominiczak, et al., 2005). Genes involved in these pathways have a potential genetic role in CVD pathogenesis. During the recent advances in genotyping technologies, GWA studies, have revealed approximately 160 loci and risk factors associated with cardiovascular diseases (Arking and Chakravarti, 2009). However, there are still many more genes and variations in the human genome remained to be found and characterized in the pathogenesis of CVD.

1.3.2. Cardiovascular diseases relevant to this thesis: essential hypertension and coronary artery disease

1.3.2.1. Genetics of essential hypertension (EH)

Hypertension is defined as the constant presence of high blood pressure exceeding over 140 and/or 90 mmHg – a systolic pressure (SBP) above 140 with a diastolic pressure above (DBP) 90. Essential hypertension (*Hypertensio essentialis (primaria)*-EH), also called primary hypertension, is a complex disease with the prevalence ~27% exhibiting highest rate among people aged 45–69 years worldwide (Delles, et al., 2009; Lawes, et al., 2008). EH comprise approximately 95% of all hypertension forms and has been shown to be associated with an increased risk for cardiovascular disease (Cowley, 2006) like stroke, myocardial infarction, heart failure (Staessen, et al., 2003), and may lead to renal insufficiency (Rosario and Wesson, 2006). From family and twin studies heritability of hypertension has been estimated from 25% up to 60% (Cifkova, et al., 2003; Luft, 2001).

Studies of Mendelian disorders of primary effect on blood pressure regulation have given a great opportunity to understand the molecular etiology of the disease. There are in total 17 genes known to cause Mendelian forms of hypertension and hypotension (Table 5) (Lifton, et al., 2001; Staessen, et al., 2003). Detected mutations have been shown to affect directly the renal tubular electrolyte transport functions, indicating to the physiological importance of kidney in the blood pressure regulation (Lifton, et al., 2001). However, rare alleles identified in Mendelian forms account for less than 1% of human hypertension and have been shown only limited association with common forms of the disease (Kato and Julier, 1999). Efforts on finding association with the rare alleles of genes underling monogenic disorders and between common forms of elevated blood pressure or hypertension have been successful with *WNK1* (lysine deficient protein kinase 1) (Tobin, et al., 2005a), *KCNJ1* (potassium inwardly-rectifying channel, subfamily J, member 1), *SLC12A3* (solute carrier family 12 (sodium/chloride transporters), member 3), and *SLC12A1* (solute carrier family 12 (sodium/potassium/chloride transporters), member 1) (Ji, et al., 2008) genes in general population.

Three main methods have been used to map genetic component underlying the pathogenesis of hypertension: linkage and association based analysis and eQTL (Binder, 2006). Until 2004 linkage studies have revealed 26 hypertension candidate genes belonging to five broad classes, genes involved in: (i) renin-angiotensin-aldosterone system, mostly on the genes involved in signal transduction (like *GNB3* – guanine nucleotide binding protein (G protein), beta polypeptide 3) and salt/water handling (like *ADD1*– adducin 1 (alpha)); (ii) adrenergic pathways; (iii) vascular and (iv) metabolism related genes (like *NOS3*), and (v) genes with potential role in hypertension development (Agarwal, et al., 2005; Delles, et al., 2009). Over 100 hypertension-related QTLs (quantitative trait loci) have been revealed in genome-wide linkage mapping across the genome (particularly chromosomes 1, 2, 3, 17 and 18) each potentially contributing with a small effect to the disease phenotype (Cowley, 2006). Mainly because of the limited knowledge of the biological pathways, gene functions, inconsistent phenotype (disease) determination, and influence of environmental factors, no strong linkage was observed between these genes and elevated blood pressure levels. In BRIGHT (The British Genetics of Hypertension) study, using mostly affected sibling pairs, the analysis of transmission of the disease loci showed several chromosomal loci associated with hypertension like 6q (lod score 3.21; p=0.042), 2q, 5q and 9q (lod score >1.57; p=0.017) (Caulfield, et al., 2003). Still, genes and causative genetic variants in these regions are not yet identified.

To find genetic loci with potential impact on gene expression levels have led to the eQTL (expression quantitative loci) approach using the combination of the QTL mapping together with microarray technology. eQTL combines expression profiling with linkage analysis in segregating populations and further correlates them with phenotypes (Abiola, et al., 2003). An example of this kind of study has identified multiple EH candidate genes (n=73) using inbred strains from spontaneously hypertensive rat (SHR) and the Brown Norway (BN) progenitor strains (Hubler and Scammell, 2004; Hubner, et al., 2005). Experimental models with SHR stroke prone strain have derived several candidate genes like *KCNJ1* (encodes the potassium channel ROMK1) (Tobin, et al., 2008), *WNK1*, *WNK4* (lysine deficient protein kinase 4) (Newhouse, et al., 2005; Tobin, et al., 2005a); *ACE* (angiotensin I converting enzyme) (Sayed-Tabatabaei, et al., 2006) and *ADD1* (adducin 1 (alpha)) (Staessen and Bianchi, 2005) participating in renin-angiotensin-aldosterone system, and Na, K-ATPase activity.

Several candidate gene based association studies have been conducted to reveal polymorphisms in susceptibility to CVD. One of the first studies was focused on screening polymorphisms in coding and/or flanking regions of 36 CVD genes among European populations (Cambien, et al., 1999). To date, over 160 genomic loci are described (http://genecanvas.idf.inserm.fr/infusions/genecanvas/Genes/GenesList.php) with the potential susceptibility to CVD

Table 5. Examples of genes known to be associated with monogenic forms of essential hypertension, modified from (Cowley, 2006).

Gene	Disease	Mutations effect	OMIM
CYP11B1; ***CYP11B2*** (cytochrome P450, subfamily 11B, polypeptide 1 and 2)	Glucocorticoid-remediable hyperaldosteronism	Ectopic expression of aldosterone synthase activity in adrenal fasciculata	#610613 #124080
11BHSD2 (hydroxysteroid 11-β dehydrogenase)*	Apparent mineralocorticoid excess	Loss-of-function mutation resulting in excess stimulation of the mineralocorticoid receptor (MR); hypertension mediated by increased renal cortical collecting tubule epithelial sodium channel (ENaC) activity	#207765
NR3C2 (mineralocorticoid receptor (aldosterone receptor)	Early-onset hypertension with severe exacerbation in pregnancy	S810L missense mutation in the ligand-binding domain converts receptor antagonists (such as progesterone) to agonists	#605115
SCNN1B (sodium channel non-voltage-gated 1β (epithelial))#	Liddle syndrome	*De novo* missense mutation of the β-subunit of ENaC	#600760
SCNN1G (sodium channel, non-voltage-gated 1γ)#	Liddle syndrome	Mutation in the γ-subunit of ENaC that deletes the cytoplasmic C terminus, resulting in excess sodium retention	#600761
WNK1; WNK4 (protein kinase, lysine deficient 1 and 4)#	Pseudohypoaldosteronism type II	WNK serine–threonine kinase defects resulting in hyperkalaemia and hypertension	#145260
PPARG (peroxisome proliferator activated receptor-γ)	Diabetes mellitus and hypertension	Mutations in peroxisome proliferator-activated receptor-γ Loss-of-function mutation resulting in insulin resistance	#601487
	Syndrome of hypertension, hypercholesterolaemia and hypomagnesaemia	Maternal inheritance causes a cytidine substitution in the mitochondrial tRNA	*Not yet Identified*

*autosomal recessive; #autosomal dominant

traits (Arking and Chakravarti, 2009). Though, because of the knowledge of the expression and function of the human genome is incomplete this approach may fail to discover important, novel pathways, which could be detected in a genome-wide scan. Also, multiple phenotypes caused by the pleiotropic effects may decrease the detection of genetic variant underlying the disease pathogenesis of interest.

The availability of HapMap data and advances in high density genotyping platforms have enabled to run hypothesis free genome-wide association approach to dissect the genetic background of CVDs. Several significant associations have been observed with BP traits and hypertension in different populations (Table 6). The two seminal genome wide association studies with hypertension were conducted by Framingham Heart Study (FHS, 100K Project) and WTCCC (Wellcome Trust Case Control Consortium) study to identify genetic variations underling elevated blood pressure in humans (Burton, et al., 2007; Levy, et al., 2007a). Moderate associations were observed between ten loci and BP traits in FHS project and between hypertension (Levy, et al., 2007a) and six genomic regions in WTCCC study ($P>5x10^{-5}$) (Burton, et al., 2007). However, none of these SNPs were common to the two studies. Loci identified by WTCCC have shown no replication in Europeans and in Koreans (Ehret, 2010; Hong, et al., 2009). In the meta-analysis of the GWAS of hypertension and blood pressure traits have identified multiple genomic loci exceeding the genome wide significance level ($P>5x10^{-8}$) (Table 6). In CHARGE Consortium (Cohorts for Heart and Aging Research in Genome Epidemiology) 13 SNPs were significantly associated with SBP, 20 SNPs with DBP and 10 SNPs with the state of hypertension (Levy, et al., 2009). In total eight loci were identified and shown to associated with systolic and diastolic blood pressure levels among individuals with European ancestry from the Global BPgen consortium (Newton-Cheh, et al., 2009). In joint meta-analysis of both CHARGE and Global BPgen data four loci attained genome-wide significance level for SBP, six for DBP and one for hypertension (Levy, et al., 2009; Newton-Cheh, et al., 2009).

Also, independent studies of different discrete populations have revealed several susceptibility loci associated with BP traits. One study has been performed to compare association of blood pressure traits and hypertension risks between SNPs within different ethnic groups like Caucasians and Koreans (Cho, et al., 2009; Hong, et al., 2010b). In total 27 loci exhibiting $P>4x10^{-7}$ significance level in Global BPgen (Caucasians Global Blood Pressure Genetics) and CHARGE studies were analyzed among unrelated individuals from KARE (Korean Association REsource) where only four loci were associated with blood pressure and the risk for hypertension (Hong, et al., 2009; Hong, et al., 2010a; Hong, et al., 2010b) (Table 6). Multiple other GWAS have been conducted to study blood pressure traits and hypertension in Amish, African Americans and European populations

Table 6. Examples of GWAS results with susceptibility to BP traits and hypertension

Study	Population	Genes/loci identified	Reference
Seminal studies			
WTCCC* (500K Affymetrix chip)	British population (n=16179)	six loci showed moderate (P>5x10^{-5}) association with hypertension: 1q43, 8q24, 12p12, 12q23, 13q21, 15q26	(Burton, et al., 2007)
Framingham Heart Study (100K Affymetrix chip)	Framingham Heart Study families, US (n=1327)	10 loci (for example in genes *CAMK4, CIorf118, TMEM144, UGT2A3, OPN5, CDH13*) showed moderate (P>5x10^{-5}) association with SBP and DBP.	(Levy, et al., 2007a)
Meta-analysis			
CHARGE Consortium**	Individuals of European ancestry (n=29136)	13 SNPs associated with SBP; 20 SNPs with DBP and 10 SNPs with hypertension (i.e. in *ATP2B1, CYP17A1, PLEKHA7, SH2B3, ATP2B1, TBX3-TBX5, ULK4* and *ATP2B1*)	(Levy, et al., 2009)
Global BPgen***	17 cohorts of European ancestry (n=34433)	genome wide association (P>5x10^{-8}) at eight loci near: *MTHFR, CYP17A1* and *PLCD3* with SBP; *FGF5, C10orf107, SH2B3, CYP1A2,ZNF652* with DBP	(Newton-Cheh, et al., 2009)
CHARGE Consortium and Global BPgen joint meta-analysis	Individuals of European ancestry (n=63569)	four loci associated with (P>5x10^{-8}) SBP: *ATP2B1, CYP17A1, PLEKHA7, SH2B3*, six, with DBP: *ATP2B1, CACNB2, CSK-ULK3, SH2B3, TBX3-TBX5, ULK4*, and one with hypertension: *ATP2B1*	(Levy, et al., 2009; Newton-Cheh, et al., 2009)
Population specific studies			
Amish study (100K Affymetrix chip)	Subjects of the religious Amish isolate (n= 542 subjects)	association with BP traits and SNPs located within the *STK39* gene	(Wang, et al., 2009)
The KORA GWAS (500K Affymetrix chip)	Southern Germany (n(S3)=1644; n(S4=1830); Estonians (n=1823)	a susceptibility locus (rs11646213) located in *CDH13* gene showed association with SBP, DBP and hypertension in KORA S3# and were replicated in KORA S4 and HYPEST## cohorts	(Org, et al., 2009)
African Americans study on BP traits (600K Affymetrix chip)	African Americans from the Washington DC (n=1017)	Association detected with SBP and five genomic loci near *PMS1, SLC24A4, YWHAZ, IPO7* and *CACANA1H* genes	(Adeyemo, et al., 2009)
KARE§ (500K Affymetrix chip)	Korean cohort (n=8842)	association detected between SBP and variant (rs17249754) located near *ATP2B1* gene	(Cho, et al., 2009)
KARE replication using CHARGE and Global BPgen GWA data (500K Affymetrix chip)	Korean cohort (n= 8512)	association detected with four loci in *ATP2B1, CSK, CYP17A1* and *PLEKHA7* genes	(Hong, et al., 2010b)

*Wellcome Trust Case Control Consortium (European origin)
**Cohorts for Heart and Aging Research in Genome Epidemiology (European origin)
***Global Blood Pressure Genetics Consortium (European origin)
§Korean Association REsource
#Kooperative Gesundheitsforschung in der Region Augsburg
European population-based sample from Estoni

(Adeyemo, et al., 2009; Org, et al., 2009; Wang, et al., 2009) (Table 6). In Amish population a strong association has been detected between blood pressure levels and common genetic variants in *STK39* (a serine/threonine kinase) gene resulting with the increase of 3.3 SBP/1.3DBP mmHg compared to non-Amish (Wang, et al., 2009). In African American population a significant association with genetic markers located near five genomic loci and SBP levels have been described (Adeyemo, et al., 2009.) A novel susceptibility locus, *CDH13*, encoding for the adhesion glycoprotein T-cadherin gene, have been proposed in GWAS for BP traits and hypertension among European populations (Org, et al., 2009). One of the striking features is that to date among the loci identified in GWAS only a small proportion of the disease variability are actually explained. Possible answers could be hidden behind the rare or structural genetic variants that might identify potential susceptibility alleles or their combination related to the development of hypertension and its related traits.

1.3.2.2. Genetics of coronary artery disease

Coronary artery disease (CAD), also known as coronary heart disease (CHD), is a complex inflammatory-metabolic disease influenced by multiple environmental and heritable risk factors. CAD is the most common cause of sudden death worldwide both in men and woman. Coronary arteries in the body have a role to supply the heart with oxygen and other nutrients. The loss of normal barrier function of endothelium of coronary artery can lead to the progressive deposition of lipids and other substances in the blood resulting with the development of plaque (Watkins and Farrall, 2006). The increased level of plaque leads to the condition known as atherosclerosis which reduces the blood flow through the vessels to the heart and may lead to the myocardial infarction (MI) (Arking and Chakravarti, 2009; Watkins and Farrall, 2006) (Figure 3).

MI is a result of rupture of atherosclerotic plaque and the formation of thrombus where heart suffers from severe damage of blood flow and oxygen supply. CAD is known to be highly heritable, ranging around 40% up to 60% but still only a proportion of the cases are described by the rare monogenic effects. There are no certain monogenic forms known underlying the CAD phenotype. Although, there are several intermediate phenotypes of rare Mendelian diseases described to be involved in premature CAD development (Table 7). It includes findings mostly linked to the heritable basis of atherosclerosis and differences in of HDL and LDL cholesterol levels (Watkins and Farrall, 2006).

Twin studies of fatal CAD events have shown differential heritability in men and women. The age range between 36–86 years has higher risk for CAD development; in men, 57%, and in women, 38% (Zdravkovic, et al., 2002). Family history of myocardial infarction has been indicated as an independent and important risk factor for premature CAD progression in humans (Assmann,

et al., 2002; Lloyd-Jones, et al., 2004). There are potentially several other unknown genetic risk factors that may contribute to the higher risk for CAD as well as cholesterol and blood pressure levels. Most forms of CAD are multifactorial in aetiology, extremely complex and poorly understood.

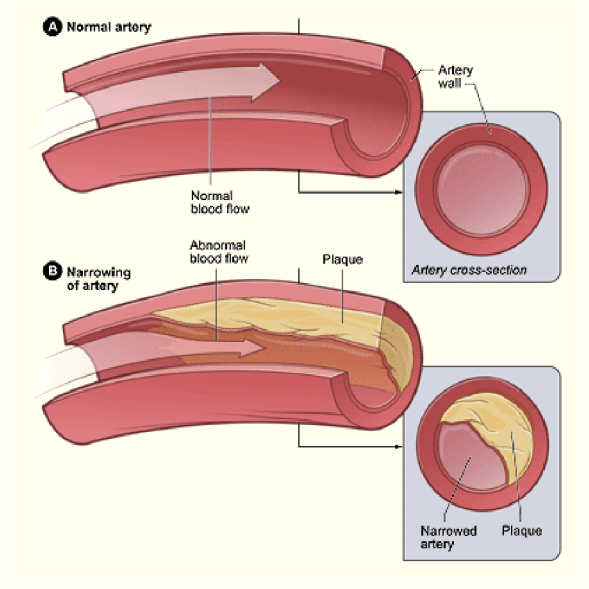


Figure 3. An illustration of plaque genesis in human coronary artery. Adapted from (http://www.nhlbi.nih.gov/) **A.** Normal artery and artery cross-section in the case of normal blood flow. **B.** Plaque formation and lumen narrowing in coronary artery, also called atherosclerosis. Atherosclerotic plaque formation is mainly predisposed by the elevated TG, LDL and lowered HDL levels. Presence of plaque limits blood flow in artery and results in myocardial infarction.

Several linkage based studies have been conducted to find genetic determinants leading to CAD phenotype (Broeckel, et al., 2002; Francke, et al., 2001; Harrap, et al., 2002; Wang, 2005). In candidate gene based linkage studies multiple variations have identified to be associated with HDL and LDL levels leading to the large phenotypic effects (Table 7). For example, both rare and common sequence differences located in *APOB* (apolipoprotein B) and *LDLR* (LDL receptor) genes have been associated with the concentration level regulation of LDL in humans (Breslow, 2000). A linkage study among Icelandic individuals identified *ALOX5AP* (arachidonate 5-lipoxygenase-activating protein) as a susceptibility gene and the involvement of leukotriene pathway for myocardial infarction and stroke. A four-SNP haplotype (HapA, spanning 33kb) located in the gene showed two times higher risk (adjusted P value of 0.005) for myocardial infarction and stroke compared to non-carriers. However, no

8

Table 7. List of genes known to be associated with monogenic forms of premature coronary artery disease, modified from (Watkins and Farrall, 2006).

Gene	Disease	Mutations effect	OMIM
LDLR (LDL receptor)	Familial hypercholesterolaemia	Defective binding of LDL by receptor	#143890
APOB (apolipoprotein B)	Familial defective APOB	Reduced binding affinity of APOB to LDLR	#144010
ABCG5 (ATP-binding cassette, subfamily G, member 5) and ***ABCG8*** (ATP-binding cassette, subfamily G, member 8)	Sitosterolaemia	Increased absorption of plant sterols	#210250
ARH (autosomal recessive hypercholesterolaemia protein)	Autosomal recessive hypercholesterolaemia	Defective endocytosis of LDLR	#603813
APOA1 (apolipoproteinA1)	APOA1 deficiency	Deletion or loss-of-function mutation that leads to very low HDL	#107680
ABCA1 (ATP-binding cassette, subfamily A, member 1)	Tangier disease	Impared cholesterol efflux in macrophages (foam cells)	#205400
CBS (cystathionine β-synthase)	Homocystinuria	Homocysteine increases thrombotic tendency	#236200
MEF2A (Myocyte enchancer factor 2A)	Coronary artery disease (ADCAD)	A mutation in transcription factor results in dominant familial vascular disease.	#608320

association was detected in an independent study between HapA and the risk of myocardial infarction among British population (Helgadottir, et al., 2004). The knowledge of participation of leukotriene pathway in CAD pathogenesis has triggered several candidate-gene association studies to test large numbers of candidate genes and their variants (Watkins and Farrall, 2006). Using whole genome-linkage scan with microsatellites have identified multiple susceptibility loci for CAD/MI on chromosome 14 (LOD= 3.9, $P<0.05$) (Broeckel, et al., 2002); on 16p13 (LOD=3.06; $P=0.00017$) and 10q23 (LOD=2.06; $P=0.00188$) among families of North-Eastern Indian origin (Francke, et al., 2001), and on chromosome 2q36 (LOD=2.63; $P<0.0001$) in Australian families (VFHS – Victorian Family Heart Study) (Harrap, et al., 2002).

The most frequently used method for identifying the susceptibility loci for CAD/MI is candidate gene in combination with case-control based association studies. Multiple QTLs in already known pathophysiological role in atherosclerosis for example studies of apolipoproteins have been shown to be associated with CAD (like lipoprotein A) (Lusis, et al., 2004). Additionally,

multiple genetic variants have been identified in *LTA* (lymphotoxin-α) gene with the susceptibility to myocardial infarction in Japanese cohorts (Iwanaga, et al., 2004; Ozaki, et al., 2002) and in populations with white European ancestry (The PROCARDIS Consortium., 2004).

Genome-wide association studies have revealed a common allele associated with CAD located on human chromosome 9 (Burton, et al., 2007; Helgadottir, et al., 2007; McPherson, et al., 2007; Samani, et al., 2007) (Table 8). The 9p21.3 region has been indicated to be associated with CAD among individuals participated in the Ottawa Heart Study (OHS) (McPherson, et al., 2007) and Icelandic patients with MI and is highly replicated in multiple other cohorts of European descent from three cities from the United States: Philadelphia, Atlanta and Durham (Helgadottir, et al., 2007). To date, using GWAS eight different loci in multiple independent sample collections, have reached for genome-wide significance ($P>5x10^{-8}$) with myocardial infarction (Kathiresan, et al., 2009) (Table 8). Compared to the GWAS of BP traits and hypertension a fewer loci with stronger associations as well as higher replication rate have been observed. This might be explained by the better described as well as by fewer intermediate phenotypes in the CAD pathogenesis than in hypertension or in BP traits.

Animal models describing the pathogenesis of CAD are limited. There are not yet any well described animal models characterizing the CAD phenotype or the plaque formation in coronary arteries (Watkins and Farrall, 2006).

Table 8. Examples of GWAS results in susceptibility to CAD

Study	Population	Genes/loci identified	Reference
WTCCC* (500K Affymetrix chip)	British population (n=16179); 1926 CAD cases and 2938 controls	Locus 9p21 (rs1333049,) associated with susceptibility to CAD ($P>5x10^{-8}$)	(Burton, et al., 2007; Samani, et al., 2007)
GWAS on Myocardial Infarction (Illumina Hap300 chip)	Icelandic population: 4587 cases and 12767 controls	Three SNPs (rs1333040, rs2383207, rs10116277) located on 9p21 had OR 1.22 for the risk allele and $P \sim 1x10^{-6}$	(Helgadottir, et al., 2007)
Ottawa Heart Study (OHS) (500K Affymetrix chip)	Caucasian cohort: 500 CAD cases and 500 controls from Canada	Two SNPs (rs10757274 and rs2383206) located in 9p21 showed significant association with CAD	(McPherson, et al., 2007)
International study of MI (Affymetrix 6.0 chip)	Individuals from US, Sweden, Finland, Spain, and Italy: 2967 cases of early-onset MI and 3075 controls	Six genetic association signals locate near 9p21.3: *CXCL12, SMAD3, MTHFD1L,* and *MIA3,* near *CELSR2 / PSRC1/SORT1*, 2q36, and *PCSK9* genes. Two novel susceptibility loci in *PHACTR1* and *MRPS6/KCNE2* genes.	(Kathiresan, et al., 2009)

* Wellcome Trust Case Control Consortium

1.4. Cardiovascular candidate genes targeted in this study

Based on the genomic location, linkage peaks, knowledge of association with CVD and functional importance in CVD susceptibility, two human cardiovascular candidate genes among 162 genes were selected for the current study: *PNMT* (17q21–22) and *NCX1* (2p22.1) (Figure 4). Both of these genes are known to have a significant role in pathogenesis of CVD in humans.

1.4.1. Human *PNMT* gene as a candidate gene for cardiovascular disease

Several genetic studies have indicated an important role of chromosome 17 in blood pressure regulation in humans, also both in mouse and rat syntenic chromosomes (chromosome 11) and rat (chromosome 10) respectively, leading to hypertension (Hilbert, et al., 1991; Julier, et al., 1997). One of the proposed CVD candidate gene, a human *PNMT*. This gene is located at chromosomal position 17q21–22 (Cui, et al., 2003) and contains only three exons and two introns (Figure 4A). *PNMT* codes for phenylethanolamine-N-methyltransferase which is a key enzyme in the last step of catecholamine biosynthesis, catalyzing the synthesis of epinephrine from norepinephrine. PNMT is present in many tissues throughout the body with higher concentration in the adrenal medulla and the left atrium of the heart (Ziegler, et al., 2002). The adrenomedullar hormone epinephrine has shown to transduce environmental stressors into cardiovascular events like hypertension, adrenergic control of stress, metabolic function, and energy metabolism. Therefore, mutations in this gene region have suggested having an impact on cardiovascular system and leading to early lethality. Still, no distinct clinical or metabolic phenotype with indication on PNMT dysfunction as well as candidate mutation in *PNMT* has been reported (Haavik, et al., 2008).

Although, the significant role of a 5´upstream SNP in *PNMT* gene in hypertension has been detected. Enrichment of the G-allele of *PNMT*-390 promoter variant (Figure 4A) among hypertensives (38.02%) compared to normotensives (27.35%) was identified in African Americans (P=0.019). No significant differences were observed in Greeks and Americans of European decent (Cui, et al., 2003). Furthermore, for neurological diseases such as Alzheimer disease and multiplex sclerosis, a protective effect of the heterozygous status of two SNPs (–390/–184 GA/AG) have been indicated (Mann, et al., 2001; Mann, et al., 2002). Both of the detected promoter polymorphisms, SNP-184; SNP-390, have been previously shown to be associated with Alzheimer disease (Mann, et al., 2001) and multiple sclerosis (Mann, et al., 2002) as well as with the enrichment of G-allele of SNP-390 and hypertension among individuals originated from America, with African decent (Cui, et al., 2003).

1.4.2. Human *NCX1* gene as a candidate gene for cardiovascular disease

Human *NCX1* gene is located on chromosome 2p22.1 and consists of 12 alternatively spliced exons (Figure 4B) (Kraev, et al., 1996) that lead to the formation of different tissue specific isoforms (Quednau, et al., 1997) with different functional properties (Dunn, et al., 2002; Hurtado, et al., 2006). NCX1 is predominantly expressed in the heart, neurons and renal tubules, but also at lower levels in other tissues, including the smooth muscle, skeletal muscle, lung and spleen (Blaustein and Lederer, 1999). Na^{+}/Ca^{+2} exchanger (NCX1) is known as bidirectional calcium transporter which is responsible for calcium homeostasis in cardiac myocytes and in several other cell types, catalyzing the exchange of one Ca^{2+} ion for three Na^{+} ions across plasma membrane. NCX has shown to be responsible for 90% of Ca^{2+} extrusion from the heart where the exchanger is an important regulator of contractility (Noble and Herchuelz, 2007). Na^{+}/Ca^{+2} exchanger participates in regulation of vascular function and therefore, alterations in the exchange process might lead to the CVD development (Blaustein, 1993) like arrhythmias, heart failure (Schillinger, et al., 2003) and salt-sensitive essential hypertension (Kokubo, et al., 2004).

Re-sequencing of the entire coding and promoter regions of *NCX1* gene in Japanese population identified 15 variations where two upstream polymorphisms (−23200T>C and −23181T>C) were significantly associated with the prevalence of hypertension in both men (P=0.04) and women (P=0.03) (Figure 4B). The pathophysiological functional behaviors of these polymorphisms are still remained to be studied (Kokubo, et al., 2004). In Ncx1 –/– deficient mice the complete lack of exchanger activity may lead to the severe alterations in heart development and embryonic lethality (Iwamoto, et al., 2004).

A. Structure of human *PNMT* gene (17q21-22)

B. Structure of human *NCX1* gene (2p23-p22)

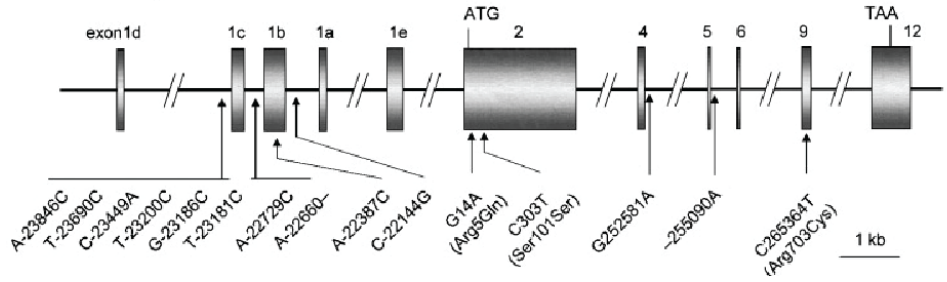


Figure 4. A. Genomic structure of human *PNMT* gene drawn to an approximate scale. The 5' and 3'UTR have been indicated according to NCBI GenBank database (February 28, 2006 release). **B.** Genomic structure of human *NCX1* in arbitrary scale adapted from (Kokubo, et al., 2004).

2. AIMS OF THE PRESENT STUDY

The aims of the present study were:

- to compile a list of cardiovascular disease (CVD) candidate genes to be used in a series of studies exploring the genetic component of CVD in Estonian and other European populations

- to conduct polymorphism screening in promoter and evolutionarily conserved non-coding regions of selected CVD candidate genes in order to identify potential gene regulatory variants

- to perform genetic association analysis between identified genetic variants in selected CVD genes (*PNMT, NCX1*) and cardiovascular traits in two Eastern-European populations

3. RESULTS AND DISCUSSION

3.1. The selection of candidate genes with a potential role in regulation of cardiovascular phenotypes (Ref. I)

Candidate genes which are potentially involved in cardiovascular disease (CVD) genesis were selected based on the information from literature reports and different genome resources such as OMIM, NCBI GeneBank, NCBI Locuslink and Ensembl. Based on the previous evidence of involvement in CVD development, genetic and biological role in cardiovascular systems, a list of 162 candidate genes was gathered for the further analysis for cardiovascular traits (Additional Table 1). This list includes genes currently known to lead to Mendelian forms of hypertension/hypotension (Table 5). Alterations in genes with a primary effect on blood pressure regulation includeing *SCNN1B* and *SCNN1G* genes leading to Liddle syndrome, *CYP11B2* and *CYP11B2* causing glucocorticoid-remediable aldosteronism (GRA), *11BHSD2* resulting in apparent mineralcorticoid excess (AME), *WNK1* and *WNK4* leading to autosomal dominant and highly penetrant disorder as pseudohypoaldosteronism type 2 (PHA2), and mutations in *NR3C2* altering ligand-binding domain and its function that leads to hypertension during the pregnancy. To better understand the nature of the regulation of high blood pressure, genes associated with low blood pressure like *CYP11B2*, *SLC12A3*, *SLC12A1*, *KCNJ1*, *CLCNKA, CLCNKB*, *BSND*, *CYP17A1*, *CYP21A2* were also included in the selection. In addition genes related to complex phenotypes of essential hypertension, myocardial infarction, coronary artery disease as well as genes reported from animal models, in human association studies, and near linkage peaks or quantitative trait loci (QTLs) were added to the list. Among 162 genes 32.1% were previously shown to participate in the development of essential hypertension, 21.6% with coronary artery disease, myocardial infarction and atherosclerosis, and 14.2% with changes in serum lipid levels (Table 9).

Table 9. Generic distribution of 162 CVD candidate genes of the study (Additional Table 1)

Disease/trait involved	Number of genes	Percentage
Essential hypertension/hypotension; Blood pressure: SBP, DBP	52	32.1%
Kidney/renal function, EH	44	27.2%
Coronary artery disease, myocardial infarction, atherosclerosis	35	21.6%
Serum lipids: HDL, LDL, TG	23	14.2%
Metabolic Syndrome, obesity	19	11.7%
Genes with pleiotropic effects	35	21.6%
Other	45	27.7%

Cardiovascular diseases are known for their high complexity with different genetic polymorphisms in many genes leading to multiple phenotypes. Based on their functional role, 21.6% of the listed genes show evidence of pleiotropic effects on multiple cardiovascular phenotypes or traits making the detection of causative variants underlying the disease even more complex to accomplish. For example genetic variants located in *LEP* (leptin precursor), *ADORA2A* (adrenergic, beta-2-, receptor, surface) and *ADD1* (adducin 1 (aplha)) genes have shown association with the susceptibility to coronary artery disease and essential hypertension. Additionally, 45 genes with an indirect implication to CVD development like different receptors and ligands (for example *AVPR1A, AVPR1B* (arginine vasopressin receptor 1A and 1B), and *AVPR2* (Arginine vasopressin receptor 2 (nephrogenic diabetes insipidus)) were included (Additional Table 1).

Currently, there are no certain pathways known to be entirely responsible for cardiovascular phenotype but based on functional role, a great majority of these 162 genes tend to cluster into three major organ systems: renal (salt handling), heart and sympathetic nervous systems (Figure 5).

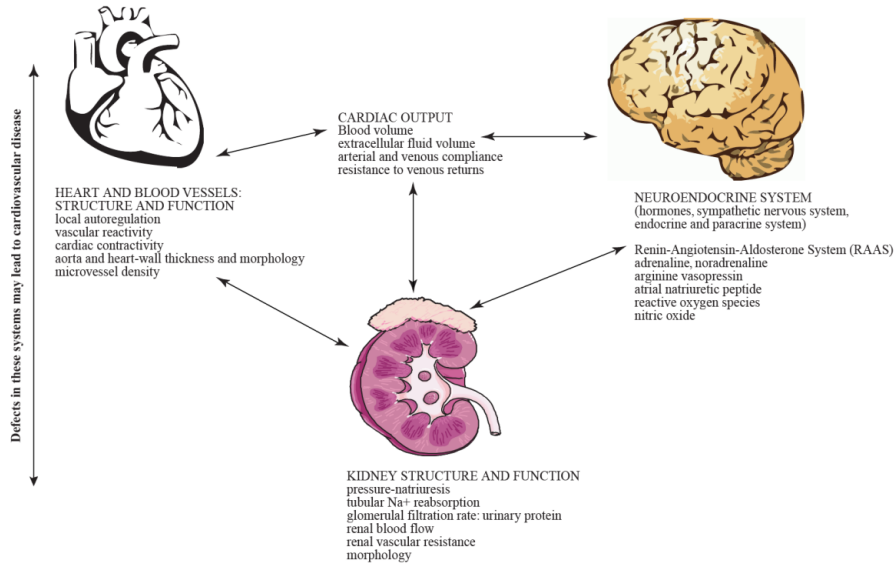


Figure 5. Three organ systems known to be involved in CVD pathogenesis. According to the reported functional roles and association to different cardiovascular pathways, 162 CVD candidate genes (Additional Table 1) can be divided among three major organ systems: renal, heart and sympathetic nervous system. Defects in these systems have been shown to influence CVD development in humans as well as in several model organisms.

A number of reports highlight multiple pathways and defects in these organ systems indicating as important players behind CVD regulation (Delles, et al., 2009; Dominiczak, et al., 2005; Floras, 1993; Helgadottir, et al., 2004). g:Profiler analysis (Reimand, et al., 2007) clustered the majority of the 162 genes into seven distinct KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways involved in cardiovascular regulation: calcium signaling pathway, hypertrophic cardiomyopathy (HCM), aldosterone-regulated sodium reabsorption, neuroactive ligand-receptor interaction, renin-angiotensin system, vascular smooth muscle contraction and type II diabetes mellitus (Figure 6). Although genes and genetic variants underlying CVD have been thoroughly investigated for years, many loci in human genome involved in complex pathogenesis of CVD are still remained unmapped and undescribed.

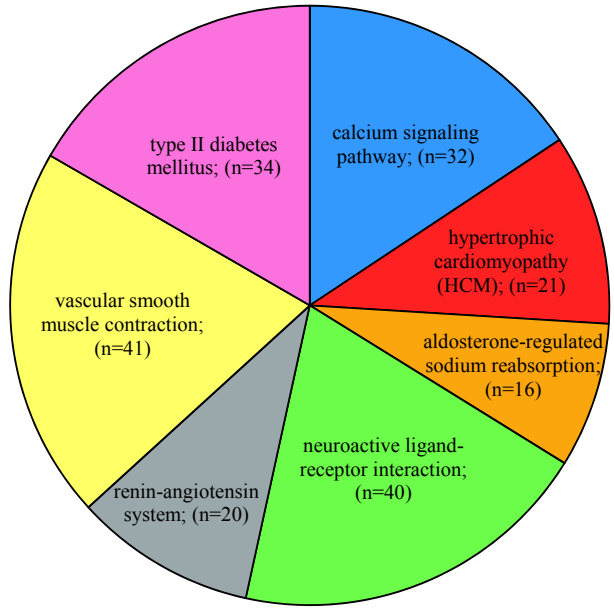


Figure 6. KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway prediction (g:Profiler) for all 162 known CVD susceptibility genes. In total seven distinct pathways were observed: calcium signaling, hypertrophic cardiomyopathy (HCM), aldosterone-regulated sodium reabsorption, neuroactive ligand-receptor interaction, renin-angiotensin system, vascular smooth muscle contraction and type II diabetes mellitus pathways

3.2. CVD phenotype description and subject selection criteria for polymorphism screening and association studies (Ref. II–IV)

Two types of CVD, coronary artery disease (CAD) and essential hypertension (EH) were analyzed in detail among two Eastern-European populations: Czechs and Estonians.

3.2.1. Sample collections used in the study

The collection of HYPEST (HYPertension in ESTonia) and CADCZ (Coronary Artery Disease in CZech) sample sets were approved by the Ethics Committee on Human Research of University of Tartu, Estonia and by the Ethics Committee of Charles University–1st Faculty of Medicine, Prague, Czech Republic. During the recruitment, all participants with Eastern-European ancestry gave their written informed consent and filled out a self-administrated epidemiological questionnaire including questions about their past and present health together with lifestyle factors.

Subjects representing HYPEST case-cohort sample were recruited across Estonia during 2004–2007 with the aim to evaluate risk factors for essential hypertension and related cardiovascular traits (Org, et al., 2009). In this work a subset of HYPEST collection (n=1122) (Table 10) including 470 individuals with essential hypertension (n=470) and 652 healthy individuals as controls were analyzed. Essential hypertension was diagnosed by a cardiologist during the patients' ambulatory visits or hospitalization at the North Estonia Medical Center, Centre of Cardiology, or at the Cardiology Clinic, Tartu University Hospital, Estonia while healthy individuals recruited from the long-term blood donors across Estonia had no personal history of CVD and had never been prescribed any relevant medications.

In the case of CADCZ cohort (n=670) (Table 10), individuals with coronary artery disease (CAD, n=257) were recruited by the Cardiology Department of the 2nd Clinic of Internal Medicine, Faculty Hospital Královské Vinohrady in Prague Czech Republic (Janosikova, et al., 2003). Individuals with no personal history of CAD, essential hypertension, MI, peripheral arterial disease, stroke, and with no recorded prescription of related medications (n=413) were obtained as controls from health clinics across the Czech Republic.

3.2.2. Phenotype definitions used in the study

In total, 727 patients with cardiovascular phenotype and 1065 healthy individuals were defined and analyzed to reveal the genetic component behind the regulation of CVD and its related traits. Patients with essential hypertension (n=470, Table 10) from HYPEST cohort were defined as follows: (a) individuals with SBP>160 mmHg and DBP>100 mmHg readings and (b) individuals diagnosed by blood pressure specialists or subjects with prescribed antihypertensive medications. The matched normotensive control group (n=652, Table 10) consisted of: (a) subjects with SBP<140 mmHg and DBP<90 mmHg readings and (b) with no recorded prescription of antihypertensive medications.

Patients with CAD phenotype (n=257, Table 10) in CADCZ sample set were diagnosed in accordance with WHO criteria (http://www.who.int/en), with one or more large stenosis of a major coronary vessel confirmed by coronarography in all subjects (Janosikova, et al., 2003). The exact diagnosis was prescribed by clinicians in the Cardiology Department of the 2nd Clinic of Internal Medicine, Faculty Hospital Královské Vinohrady in Prague Czech Republic and Czech health clinics between 1998–2000. Healthy individuals (n=413, Table 10) in CADCZ sample set were recruited by health clinics across Czech Republic and represented with no personal history of CAD, essential hypertension, MI, peripheral arterial disease, or stroke, and had never been prescribed any disease related medications.

Subjects of CADCZ sample set were further subgrouped based on having MetS (n=88) as growing evidence of MetS in regulation of CVD, specifically in atherosclerotic diseases leading to CAD, has been reported (Kasai, et al., 2008). MetS in CADCZ was defined using criteria appointed by the International Diabetes Federation (http://www.idf.org): triglycerides>1.7mmol/L; BMI>30kg/m2; SBP>130mmHg; DBP>85mmHg and HDL>1.29mmol/L.

Many end-stage phenotypes like renal insufficiency (Rosario and Wesson, 2006) and heart failure are known to affect CVD. To avoid false positive results in association study, all subjects diagnosed with diabetes and renal diseases were excluded from HYPEST and CADCZ sample sets. Furthermore, to minimize the effect of severe obesity and age related risk factors, individuals only with BMI <35 kg/m^2 and age <65 were analyzed. Reducing the impact of the prescribed medications, two corrections described by Martin Tobin and Jun Wu were applied for individuals who had taken antihypertensive or lipid lowering drugs respectively (Tobin, et al., 2005b; Wu, et al., 2007).

Table 10. Baseline characteristics of patient and control individuals in CADCZ and HYPEST cohorts

Characteristic	HYPEST cohort		CADCZ cohort	
	EH patients	Healthy subjects	CAD patients	Healthy subjects
Number of subjects	470	652	257	413
	Mean (SD)			
Age (years)[1]	43.9 (13.0)	39.0 (4.9)	51.2 (8.1)	49.5 (7.4)
Body Mass Index (BMI) (kg/m^2)	28.7 (3.7)	24.4 (3.3)	27.9 (3.3)	25.3 (3.1)
Systolic blood pressure (SBP) (mmHg)	143.2 (17.6)	130.0 (22.7)	136.6 (19.2)	125.2 (14.0)
Diastolic blood pressure (DBP) (mmHg)	87.4 (10.6)	81.0 (14.6)	85.1 (11.3)	80.3 (9.3)
Total cholesterol (mmol/liter)	5.6 (1.1)	5.0 (1.4)	5.4 (1.0)	5.6 (1.0)
High-density lipoprotein (HDL) (mmol/liter)	1.5 (0.4)	1.7 (0.5)	1.2 (0.3)	1.5 (0.4)
Low-density lipoprotein (LDL) (mmol/liter)	3.8 (1.0)	3.3 (1.1)	3.2 (0.8)	3.4 (0.9)
Tricylglycerols (mmol/liter)	1.8 (1.6)	0.8 (0.3)	2.1 (1.3)	1.5 (1.1)

[1]EH and CAD patients: age at the onset of disease; healthy subjects: age at the recruitment

3.3. Screening for novel genetic markers in CVD candidate genes (Ref. II–IV)

3.3.1. Selection of genomic regions and methods for detection of genetic variants

In order to detect variations underlying the pathogenesis of cardiovascular diseases, potential regulatory regions for the 162 originally selected genes were defined and subjected to polymorphism screening.

Firstly, the entire *PNMT* gene was screened for variations potentially having an impact on CVD pathogenesis. *PNMT* gene, located in chromosome 17, has been identified as a significant gene involved in blood pressure regulation (Knight, et al., 2003; Naber and Siffert, 2004; Rutherford, et al., 2001); accordingly is considered as a candidate gene of high importance in CVD (Cui, et al., 2003; Koike, et al., 1995). The entire *PNMT* gene (Chr.17q21–22, 3187bp) was subjected to polymorphism detection in hypertensive and normotensive individuals among two Eastern-European populations (Figure 7).

Secondly, promoter regions defined among CVD candidate genes were subjected for polymorphism screening. Promoter regions were assessed according to NCBI GenBank sequences spanning from +100bp up to –500bp

I. POLYMORPHISM SCREENING			
Genomic regions studied:	Whole gene re-sequencing	Core promoter regions (+100bp up to –500bp)	105 Conserved Noncoding Regions (CNR)
Studied genes:*	Human *PNMT* gene (chr.17q21–22, 3187bp)	*ACE, ADD1, ADD2, AGT, AGTR1, AGTR2, ATP1A1, BSND, CLCNKA, CLCNKB, CYP11B2, CYP17A1, CYP21A2, HSD11B1, IL1A, KCNJ1, KLK1, NPR1, NPR2, NR3C2, REN, SAH, SCNN1A, SCNN1B, SGK, SLC12A3, SLC14A2, SLC22A2, NCX1*	*ACE, ADD1, ADD2, AGTR1, ATP1A1, GNB3, HSD11B1, HSD11B2, IL1A, KCNJ1, NEDD4L, NPR1, NPR2, NR3C2, WNK1, WNK4, REN, SAH, SCNN1A, SCNN1G, SGK, SLC12A1, SLC14A2, SLC22A2, NCX1*
Technology:	re-sequencing	re-sequencing	DHPLC, re-sequencing
Study sample:	HYPEST and CADCZ	HYPEST	HYPEST and CADCZ
No. of individuals:	n=100	n=14	n=49
Phenotype:	EH and CAD patients and healthy individuals with no CVD	healthy individuals with no CVD	EH and CAD patients
Reference no. :	Ref. II	Ref. III	Ref. IV
II. GENOTYPING AND GENETIC ASSOCIATION ANALYSIS			
Polymorphisms screened:	Two *PNMT* promoter SNPs (–184 (rs876493) and –390 (rs3764351)	In promoter regions of 17 genes no SNPs were detected. In case of 12 genes the identified SNPs were singletons or located in duplicated genes with higher nucleotide diversity. Based on these results no follow-up association analysis with CVD was performed in current thesis.	14bp indel (rs11274804, located in 348bp CNR region in *NCX1* gene intron 2)
Technology:	re-sequencing		DGGE, re-sequencing
Study sample:	HYPEST and CADCZ		HYPEST and CADCZ
Phenotypes:	EH, BP		EH, CAD, MetS, BP, HDL, LDL, TC, TG, HR, IMT
No. of individuals:	n=100		n=1792
Reference no.	Ref. II		Ref. IV

*Detailed description of the analyzed genes are given in Additional Table 2

Figure 7. A schematic overview of the study design applied in this thesis. Three polymorphism detection approaches among 37 CVD genes followed by two genetic association analyses are described. No association analysis was performed between detected promoter variants and CVD.

relative to mRNA transcription initiation sites, harboring both the core and proximal promoter regions. As the accurate determination of the transcription start site and the potential promoter region of the gene is difficult to identify, only 29 genes out of 162 were subjected for further promoter polymorphism screening among Eastern European population (Figure 7).

Thirdly, conserved noncoding regions (CNRs) of the CVD candidate genes were analyzed among two Eastern European populations. Genomic regions exhibiting >70% conservation of 100bp sliding window compared to mouse and rat genomes were addressed. In total 105 CNRs located in 25 genes that fit the above criteria and were further analyzed in the pathogenesis of cardiovascular diseases (Figure 7, Additional Table 2).

Polymorphism screening of all selected genomic regions were assessed by three alternative detection methods: re-sequencing (Sanger method), DHPLC and DGGE assays. Direct sequencing was applied for variation screening of core promoter regions of 29 human genes (given in bold in Additional Table 1) and for human *PNMT* gene. DHPLC and DGGE applications were used to identify potential regulatory variants within 105 CNR regions of 25 CVD genes (are underlined in Additional Table 1). DHPLC is based on denaturation/ renaturation and analysis of formed heteroduplexes of genomic sequences. DGGE is based on different melting properties of dsDNA in an increasing gradient of denaturant (urea and formamide) at fixed elevated temperature.

3.3.2. Diversity and polymorphism pattern within human *PNMT* gene and CVD candidate gene promoter regions (Ref. II–III)

3.3.2.1. Screening of genetic variation of human *PNMT* gene (Ref. II)

The entire human *PNMT* gene was re-sequenced in hypertensive and normotensive individuals with European ancestry with the aim to identify novel hypertension-susceptibility polymorphisms. Re-sequencing of human *PNMT* gene (3187bp; [GeneBank:J03280.1]) was performed among normotensive (n=50) and hypertensive individuals (n=50) from HYPEST and CADCZ sample sets (Figure 7). Samples included equally 25 normotensives and 25 hypertensives from both cohorts. The re-sequenced region spanned the entire 5´and 3´UTR regions (–882bp to +2305bp relative to ATG), three exons (424bp, 208bp, 525bp), and two introns of the gene (Figure 4A). In total seven SNPs along 3187bp region were revealed by re-sequencing, represented by three rare genic polymorphisms and four upstream SNPs indicating to relatively low variation (Ref. II: Figure 1, Table 2). Compared to NCBI reference assembly (build no.131), there are 55 SNPs described to locate in *PNMT* gene from which 15 were in coding regions. Detected variation pattern was similar in both populations. No single SNP was exclusively present in hypertensive

individuals. Only one non-synonymous mutation in exon 3 (SNP+1517; Ala->Thr) was detected, present in both normotensive (MAF=6%) and hypertensive (MAF=4%) individuals. Apart from two upstream SNPs (SNP-184 A/G; SNP –390 A/G) none of the other detected variations have been previously described as common variants in human genome. Detected low variation pattern of the entire human *PNMT* gene, described among 100 Eastern-European individuals, may reflect the significant role of epinephrine in the regulation of cardiovascular and metabolic functions. Also, neurons expressing PNMT have been suggested to play a critical role in regulating the development and maintenance of hypertension (Reis, et al., 1988).

3.3.2.2. Polymorphism screening in promoter regions of 29 CVD candidate genes (Ref. III)

To describe the variation pattern of core promoter regions (+100bp up to –500bp relative to mRNA transcription initiation site) in 29 CVD candidate genes, 14 Estonian normotensive individuals were re-sequenced (Figure 7). In total, 23 SNPs among 12 genes were detected ranging from one up to four SNPs per core promoter region. The most variable core promoter regions were located in *CYP11B2, IL1A* and *SLC14A2* genes, known to have duplicate copies in the genome. No SNPs were detected in 17 (59%) out of 29 genes: *ACE, ADD1, ADD2, AGT, AGTR2, KCNJ1, NPR1, HSD11B1, CLCNKB, KLK1, NR3C2, CYP21A2, SLC22A2, REN, ATP1A1, SCNN1B,* and *SLC8A1*. Observed low variation in these core promoter regions are in concordance with their crucial role in the gene regulation processes. Majority of detected SNPs were singletons or located in duplicated genes and therefore no further studies were performed.

3.3.2.3. Diversity parameters of *PNMT* and 29 CVD gene promoter regions (Ref. II–III)

Population genetics statistics were further used to estimate the evolutionary pressure on human *PNMT* and on promoter regions of 29 CVD genes. The variation pattern observed by re-sequencing of entire *PNMT* gene were the lowest within introns, representing only one rare SNP (SNP +360) located in intron 1 of *PNMT* gene (Table 11, Ref. II: Figure 1). Comparing diversity parameters between intronic, exonic and upstream regions revealed that introns exhibited >15 fold and >7 fold diversity reduction compared to 5' UTR and exons respectively. In contrast, diversity between 5' UTR and exons showed only 2–2.5 fold difference (Table 11). Furthermore, second intron, spanning only 114bp, of *PNMT* gene showed complete lack of diversity.

Table 11. Sequence diversity parameters of human *PNMT* gene and the core promoter regions of 12 cardiovascular disease candidate genes

	Gene (region)	No. SNPs	π[1]	θ[2]	D[3]
***PNMT* gene**					
Sequenced region (3187bp)	All Estonians[4]	6	0.00037	0.00036	0.04368
	Hypertensives		0.00035	0.00035	0.01342
	Normotensives		0.00039	0.00042	–0.16433
	All Czech[4]	7	0.00027	0.00036	–0.56268
	Hypertensives		0.0030	0.00042	–0.72287
	Normotensives		0.00024	0.00035	–0.75219
Exons (1157bp)	All Estonians	2	0.00052	0.00050	0.07805
	Hypertensives		0.00050	0.00039	0.51545
	Normotensives		0.00056	0.00058	–0.07671
	All Czech	2	0.00052	0.00050	0.06616
	Hypertensives		0.00054	0.00058	–0.13715
	Normotensives		0.00050	0.00039	0.50566
Introns (1065bp)	All Estonians	1	0.00002	0.00018	–1.02786
	Hypertensives		NA	NA	NA
	Normotensives		0.00004	0.00021	–1.10280
	All Czech	1	0.00007	0.00018	–0.68607
	Hypertensives		0.00007	0.00021	–0.87191
	Normotensives		0.00007	0.00021	–0.87191
5' UTR region (882bp)	All Estonians	3	0.00116	0.00066	1.38242
	Hypertensives		0.00114	0.00076	1.03138
	Normotensives		0.00120	0.00076	1.19760
	All Czech	4	0.00121	0.00088	0.74751
	Hypertensives		0.00138	0.00101	0.83180
	Normotensives		0.00099	0.00101	–0.03929
Promoter regions for 12 CVD candidate genes[5]					
	AGTR1	1	0.00082	0.00043	1.47
	CYP17A1	2	0.00058	0.00043	0.57
	CLCNKA	2	0.00081	0.00086	–0.11
	CYP11B2	3	0.00168	0.00085	2.02
	NPR2	1	0.00042	0.00043	–0.018
	BSND	2	0.00023	0.00043	–0.74
	SAH	2	0.00071	0.00087	–0.38
	SGK	1	0.00079	0.00043	1.36
	SLC12A3	1	0.00075	0.00043	1.216
	SCNN1A	3	0.00167	0.00128	0.721
	IL1A	2	0.0017	0.00086	2.02
	SLC14A2	4	0.00133	0.00085	1.17

Estimation of nucleotide diversity per site from [1]average pairwise difference among individuals and [2]number of segregating sites (S); [3]Tajima's D statistics (Tajima, 1989); [4]n=25 for normotensives, n=25 for hypertensives. [5]Regions spanning –500bp and +100bp relative to mRNA transcription start point. Upstream regions possessing duplicate copies are indicated in **bold**. Overview about genes and their genomic location is given in Additional Table 1. NA – not applicable.

Further Tajima DT neutrality test was used to test if the pattern of DNA sequence variation in *PNMT* fit the expectations under hypothesis of neutrality. Tajima's D value is indicating to the difference between two estimates: π (per-site heterozygosity derived from the average pairwise sequence differences) and θ (Watterson's θ, per-site heterozygosity based on the number of segregating sites). Under neutral conditions $\pi=\theta$ and Tajima's D=0. No statistically significant difference was observed for the entire *PNMT* between observed (π) and expected (θ) nucleotide diversity parameters.

Interestingly, analyzing nucleotide diversity of 29 re-sequenced promoter regions of CVD candidate genes revealed similar low diversity with the average variation π/bp =0.00040 as *PNMT* upstream region (Table 11). This phenomenon may refer to the selective constraint and/or the maintenance of functioning promoter variants with differential effect on gene expressional level.

3.3.2.4. A potential regulatory unit in intron 1 of human *PNMT* gene (Ref. II)

Based on the low variation detected in human *PNMT* intron 1 and the possibility of purifying selection in the region, *in silico* analysis was implemented to explore the functional regulatory elements potentially important in the gene regulation processes. Previously, a human-specific Glucocorticoid Responsive Element (GRE) inserted by an *Alu*-element mediated transfer has been described in *PNMT* intron 1 (Sasaoka, et al., 1989) (Figure 4A). *In silico* analysis of the entire intron 1 region (951 bp) confirmed the insertion of GRE element *via Alu*-sequence. Furthermore, using MatInspector 2.2 software (http://www.genomatix.de/en/produkte/genomatix-software-suite.html) together with manual inspection of previously reported sequence motifs a human-specific gene regulatory unit, a Glucocorticoid Responsive Unit (GRU) with other *cis*-acting elements were predicted within *PNMT* intron 1 (Ref. II: Figure 2). Enhanced glucocorticoid (GC) response by the binding of other transcription factors to adjacent binding sites to form Glucocorticoid Responsive Units (GRUs) has been observed in a number of genes (Schoneveld, et al., 2004). The regulatory role of GRU unit has been described in rat liver *6-phosphofructo-2-kinase* (*Pfk-2*) gene targeted by both glucocorticoids and insulin (Pierreux, et al., 1999) and also in mouse *c-HA-ras* gene where the regulation occurs jointly by GRE and Estrogen Responsive Element (ERE) (Pethe and Shekhar, 1999). Additionally, sensitivity for GC has been reported to affect *PNMT* promoter activity in rat (Wong, et al., 1998) and bovine (Cahill, et al., 1996), and at least one putative GRE has been identified for every species-specific *PNMT* gene.

Both, the detected low variation between normotensives and hypertensives individuals together with predicted intronic regulatory unit incorporating multiple TF binding sites, suggests that alterations in PNMT expressional profile is not controlled by the polymorphisms located in the genic region but rather by the complex role of expressional regulators of the gene.

3.3.3. Variation patterns in conserved non-coding regions of CVD candidate genes (Ref. IV)

3.3.3.1. Conserved non-coding regions of CVD candidate genes

Conserved non-coding regions in 162 CVD candidate genes were defined and screened among individuals from HYPEST and CADCZ cohorts with the aim to detect regulatory variants affecting susceptibility to CVD and its related traits. Web-based VISTA tools (http://genome.lbl.gov/vista/index.shtml) were used to define conserved non-coding regions of 162 CVD genes presented in humans compared to rat and mouse genomes. Conservation parameters proposed by VISTA website were applied: 100bp sliding window and sequence identity >70% compared to both rodents genome. In total 365 CNR regions were detected between human, mouse and rat sequence comparisons (May 2004, NCBI Build 33) (Tõnisson, unpublished data). Based on the used technological limitations of DHPLC and DGGE assays, sequence length (50–300bp) and location (>200bp from the nearest exon), 105 regions were selected for further polymorphism screening among 25 CVD genes (Kepp, unpublished data) (genes and regions are given underlined in Additional Table 1 and 2). The average length of 105 conserved regions was 140bp (range: 55–291bp) and for PCR fragments 374bp (range: 249–505bp). The average conservation of CNRs compared to rat was 75.9% and to mouse 76%.

3.3.3.2. Polymorphism screening in conserved non-coding regions of CVD candidate genes

To increase the chance of finding genetic variations in susceptibility to human CVD, only patients diagnosed with essential hypertension (n=22, HYPEST) and coronary artery disease (n=24, CADCZ) were subjected for variation detection among 105 selected CNRs from 25 CVD genes (Figure 7). To minimize technological limitations like inappropriate PCR product (too long) and primer design two complemental mutation screening methods were used: DGGE and DHPLC. Information about the 105 regions and their inclusion/exclusion criteria are given in Additional Table 2.

In total 39 genetic variations were detected among 15 CVD genes ranging from 0–14 variations per gene (Table 12 and Additional Table 2, unpublished data). Genes known to have structural isoforms of transcripts like *NCX1, WNK1, WNK4, SLC*-genes and *NEDD4L* exhibited higher number of polymorphisms than genes known to have single transcripts in the genome. Majority of the detected sequence differences were single nucleotide substitutions and were presented as singletons or are available in commercial genotyping platforms already presented in a large number of studies. Additionally, two polymorphic indel variants were detected in CNR regions of human *WNK1* and *NCX1* genes. Sequence analysis revealed a novel polymorphic *AluYb8* element

insertion into *WNK1* intron 10. This *Alu*-insertion was targeted for further evolutionary and population genetic analyses as well as was explored for the association with cardiovascular disease and the effect on the gene expression profile multiple in European populations (Putku, Kepp, et., al manuscript in preparation). The second, common 14bp indel (rs11274804) is located in *NCX1* gene intron 2 (the first intron in the coding region). The localization of the 14bp indel in the intron 2 raised the hypothesis about its potential effect on gene expression and was targeted for further analysis in CVD phenotypes among Eastern-European population.

Table 12. Conserved non-coding regions and variations detected among 25 CVD genes (Kepp, unpublished data).

No.	Gene	Gene ID	Chromosomal location	Analyzed bp per gene	Number of CNRs screened	Genetic variants detected in CNRs
1	*ACE*	1636	17q23	298	1	1 (rs4316818)
2	*ADD1*	118	4p16.3	349	3	0
3	*ADD2*	119	2p14-p13	396	1	2 (novel (A/G), novel (C/T))
4	*AGTR1*	185	3q21-q25	505	1	0
5	*ATP1A1*	476	1p13-11	479	1	0
6	*GNB3*	2784	12p13	282	1	0
7	*HSD11B1*	3290	1q32-q41	300	1	0
8	*HSD11B2*	3291	16q22	387	1	0
9	*IL1A*	3552	2q14	393	4	1 (rs1304037)
10	*KCNJ1*	3758	11q24	351	1	0
11	*NCX1*	6546	2p23-p22	367	16	14*
12	*NEDD4L*	23327	18q21	361	13	3 (two novel (G/A), rs4149609)
13	*NPR1*	4881	1q21-22	357	2	1 (novel (insT))
14	*NPR2*	4882	9p21-p12	354	4	1 (rs13294295)
15	*NR3C2*	4306	4q.31.1	384	12	1 (rs17582031)
16	*REN*	5972	1q32	300	1	0
17	*SAH*	6296	16p13.11	351	3	1 (rs1027457)
18	*SCNN1A*	6337	12p13	366	1	1 (novel (C/T))
19	*SCNN1G*	6340	16p12	332	1	0
20	*SGK*	6446	6q23	357	1	0
21	*SLC12A1*	6557	15q15-q21.1	374	8	2 (novel (A/T), novel (G/C))
22	*SLC14A2*	8170	18q12.1-q21.1	388	3	1 (novel (T/C))
23	*SLC22A2*	6582	6q26	383	2	1 (rs316021)
24	*WNK1*	65125	12p13	393	15	8*
25	*WNK4*	65266	17p21-q22	374	8	1 (novel (A/C))
In total					**105**	**39**

*Information about all the detected variants is in the Additional Table 2.

3.3.4. Characterization of hypervariable region in *NCX1* gene (Ref. IV)

3.3.4.1. Identification of a hypervariable CNR region in *NCX1* intron 2

The further study of the CNR region located in *NCX1* intron 2 in the etiology of CVD phenotypes in larger Eastern-European population sample set (n=1792, Figure 7) revealed its hypervariable nature. Studied sample set included 470 EH patients and 652 healthy subjects from HYPEST cohort and 257 CAD patients and 413 healthy subjects from CADCZ cohort (Table 10). In addition to identified common 14bp indel variant in both sample collections, three alternative deletion polymorphisms: 5bp, 10bp and 43bp, one SNP (C/G) and a duplication of 40bp segment with overlapping breakpoints, were detected by DGGE method in the 348bp analyzed region (Figure 8). Genotyping among 1792 individuals revealed that the common variant, 14bp indel, was represented with the allele frequency of up to 8.51%. Other detected variants were represented as singletons. One of the novel variants, a 43bp deletion, showed enrichment in HYPEST presented in 20 subjects of HYPEST compared to only one in CADCZ cohort. Additionally, one of the HYPEST subject appeared a heterozygote for both 14bp indel/43bp deletion (Ref. IV: Figure 2 and Table 3).

3.3.4.2. *NCX1* 14bp indel as a human-specific deletion compared to ancestral primate sequence

Description of the ancestral status of the *NCX1* 14bp indel (rs11274804) was performed by ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html) using the consensus sequences from two primate species, the common chimpanzee and rhesus macaque. The analysis revealed that the ancestral primate variant is actually the minor human allele containing the sequence of the previously described 14bp indel (rs11274804). The major human variant, represented by the allele frequency of 90.45% among Eastern-European subjects, has evolved through a 14bp human-specific deletion when compared to other primates (Figure 8). Comparison also indicates that all the other detected short indel variants surrounding the detected 14bp indel have possibly occurred on the chromosomal variant carrying the human-specific 14bp heterozygous deletion. Supportingly, all the alternative variants were absent from both chimpanzee and rhesus macaque consensus sequences (Figure 8).

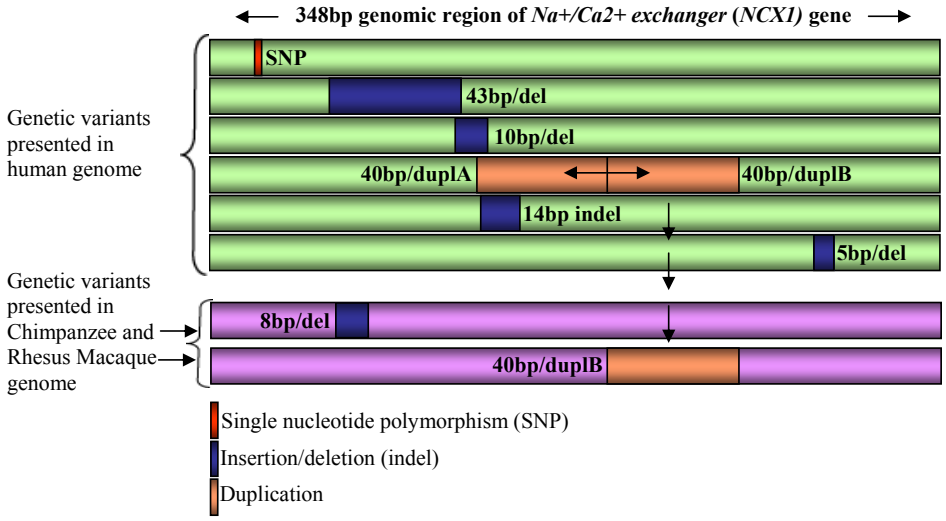


Figure 8. Approximate location of insertion and deletion variants, one SNP and one duplication detected in human CNR of *NCX1* intron 2. Indels confirmed by sequencing in human *NCX1* gene, are depicted on green background and indels presented in in *Ncx1* genes of two primates, common chimpanzee and rhesus macaque are shown on purple background. SNP is indicated in red, deletions in blue and duplications in orange color.

3.3.4.3. Short indels as a source of evolution for the hypervariable regions

Short indels have been previously shown to represent a combination of micro-deletion/micro-insertion events that give rise to the replacement of one or more base pairs by others (Chuzhanova, et al., 2003). Indels tend to form preferably in polymorphic sites of the genome where DNA sequence is more prone to structural changes. Especially the regions undergoing double stranded DNA breaks repair are more susceptible to higher mutation rate (Lercher and Hurst, 2002; Tian, et al., 2008). There is little known about the origin of indels due to their complexity in length and mutational frequency. A potential sequence motif has been reported, acting as an indel hotspot, representing GTAAG sites with the high prevalence within genomic regions prone to indel events (Ball, et al., 2005; Chuzhanova, et al., 2003). Comparing proposed motif with studied 348bp region it appeared to occur on the reverse strand (GTAAG vs. CATTC) within the 14bp indel (CATTCCCTCTCCAT). Furthermore, two additional CATTC motifs were present at the studied hypervariable *NCX1* intronic region (Ref. IV: Figure 2b) indicating to the potentially higher mutational rate along this sequence area.

3.4. Association analysis of detected variants in two candidate genes: *PNMT* and *NCX1* among two Eastern-European populations

In the last part of my thesis, I tested the association of identified polymorphisms located in two CVD candidate genes, *PNMT* and *NCX1,* with cardiovascular traits among two Eastern-European cohorts HYPEST and CADCZ.

3.4.1. Human *PNMT* gene: no common genetic variants associated with blood pressure regulation (Ref. II)

A study was conducted to test for the association of two common upstream SNPs (SNP-184 (rs876493) and SNP-390 (rs3764351)) observed in human *PNMT* gene with the prevalence of high blood pressure levels in two European populations (25 normotensive and 25 hypertensive individuals from both HYPEST and CADCZ cohorts). In allelic and genotypic differentiation tests (Fishers's exact test) among Estonians and Czech no significant differences were observed in either study groups nor between normotensive and hypertensive individuals ($p>0.05$). However, a significant excess of heterozygotes for the two promoter polymorphisms was detected among hypertension patients in a joint case-control analysis (Fisher exact test, $p<0.05$). These results support the role of these polymorphisms in susceptibility to CVD. Previously, the two polymorphisms have shown to be associated with hypertension African American individuals (Cui, et al., 2003). In haplotype analysis (PHASE 2.1) no significant differences were detected between normotensives and hypertensive patients in both sample sets.

3.4.2. Human *NCX1* gene: association of 14bp indel variant with cardiovascular traits (Ref. IV)

3.4.2.1. Association of the 14bp indel with coronary artery disease

To test for the association of detected intronic 14bp indel variant (rs11274804) with the pathogenesis of CVD, two main CVD phenotypes: essential hypertension (HYPEST, cases=470 and controls=652) and coronary artery disease (CADCZ, cases=257 and controls=413) were used (Figure 7). Association was tested using logistic regression under additive and dominant effect models including age, sex, and BMI as covariates. Significant evidence of association was observed with *NCX1* intronic 14bp indel region and subjects with the diagnosis of CAD from CADCZ cohort (P=0.0016, OR=2.02; P=0.0018, OR=2.07; additive and dominant models, respectively) but not with the patients diagnosed with essential hypertension in the Estonian HYPEST

sample collection (P>0.1). Epidemiological studies have pointed out several risk factors for CAD, including high level of low-density lipoprotein cholesterol, low level of high-density lipoprotein cholesterol, high triglyceride levels, and hypertension etc. (Lusis, 2000). Combinations of these risk factors may lead to a condition named metabolic syndrome (MetS). MetS has been implicated to be one of the important factors in CAD developmental processes (Kasai, et al., 2008; Wassink, et al., 2008). Additional analysis of CAD patients with metabolic syndrome (International Diabetes Federation: http://www.idf.org/) revealed even higher association (P=0.0014, OR=2.34; P=0.0016, OR=2.41; additive and dominant models, respectively) compared to CAD cases only (Table 13). After the correction for multiple testing (significance level of Bonferroni correction was used), the risk remained significant (α=0.05/11=0.0045) among CAD subjects with metabolic syndrome.

3.4.2.2. Association of the 14bp indel with quantitative cardiovascular traits

To characterize the role of separate risk factors influencing the susceptibility to cardiovascular disease, association of 14bp indel polymorphism with several quantitative cardiovascular parameters including SBP, DBP, heart rate (HR), intima media thickness (IMT), total cholesterol (TC), HDL, LDL and triglycerides (TG) were tested (linear regression under additive and dominant models). A marginal negative evidence was observed with heart rate (P=0.04, beta= –1.6) and LDL (P=0.04, beta= –0.26) level among healthy European individuals (Table 14, Table 15). Additionally, a supportive evidence among the CAD sample set was detected with serum triglyceride levels (P=0.04, beta=0.25; Table 15). These results support the previous evidence that higher concentration levels of TG are an independent risk factor for CVD, foremost in the CAD pathogenesis (Nordestgaard, et al., 2007). No significant association was detected with other quantitative parameters either in separate or in meta-analyses. In order to confirm the detected association of common 14bp indel (rs11274804) polymorphism with CAD and its related metabolic risk factors additional populations need to be examined. Moreover, it would also be challenging to analyze other observed rare indel variants or their interactions with the common 14bp indel across the intronic CNR region (348bp) in a larger CVD sample sets. Also, functional studies would give an opportunity to reveal whether the alternative *NCX1* intronic indel variants affect the alternative transcript profile of the gene or not. In conclusion, the phenomenon of the enrichment of the novel 14bp deletion in human genome may have been triggered from either natural selection or genetic drift being advantageous in decreasing the risk for CAD and the development of higher triglyceride levels. Similar observation between human and chimpanzee has been pointed by the indels in immunity-associated loci contributing to different responses to viral infections like in the case of human immunodeficiency virus (HIV) (Tian, et al., 2008).

Table 13. Association between cardiovascular disease and the carrier status of the 14bp indel in human *NCX1* intron 2

Disease	Sample	Sample size cases/controls[3]	Minor allele frequency (%)		Association testing using logistic regression[1] Additive model		Dominant model	
			Cases	Controls	P-value	OR [95%CI]	P-value	OR [95%CI]
Hyper-tension	HYPEST	470/652	7.76	9.04	0.14	0.70 [0.44, 1.12]	0.09	0.65 [0.40, 1.06]
Coronary artery disease (CAD)	CADCZ	257/413	12.45	7.02	**0.0016**	2.02 [1.30, 3.13]	**0.0018**	2.07 [1.31, 3.26]
CAD+Metabolic syndrome[2]	CADCZ	88/361	15.07	7.64	**0.0014**	2.34 [1.38, 3.96]	**0.0016**	2.41 [1.39, 4.18]

[1]Logistic regression analysis was performed with the following covariates: sex, age, BMI.
[2]Metabolic syndrome was defined according to IDF (International Diabetes Federation).
[3]Detailed description of cases and controls for essential hypertension, CAD and Metabolic Syndrome is given in Materials and Methods.
Significant differences have been highlighted in bold, **P<0.05**.

Table 14. Association between cardiovascular parameters and the carrier status of the 14bp indel in *NCX1* intron 2

			WT/WT	WT/indel	indel/indel	Association testing using linear regression[1] Additive model		Dominant model	
	Sample	n	Mean (±SD)	Mean (±SD)	Mean (±SD)	P-value	Effect (SE)	P-value	Effect (SE)
SBP (mmHg)[2]	HYPEST	997	140.7 (21.19)	141.4 (22.93)	168.7 (36.14)	0.31	1.50 (1.47)	0.45	1.14 (1.52)
	CADCZ	670	132.4 (20.29)	134.1 (22.00)	135.0 (18.03)	0.44	1.43 (1.85)	0.42	1.54 (1.92)
	Meta-analysis	1667	137.4 (21.23)	138.4 (22.80)	151.8 (31.51)	0.20	1.47 (1.15)	0.28	1.29 (1.19)
	Healthy subjects[3]	1048	128.5 (5.66)	128.0 (10.53)	136.5 (9.19)	0.47	–0.71 (0.98)	0.39	–0.86 (0.99)
DBP (mmHg)[2]	HYPEST	993	86.9 (12.67)	87.2 (13.31)	104.3 (22.28)	0.39	0.82 (0.95)	0.57	0.55 (0.98)
	CADCZ	669	84.2 (11.51)	84.1 (15.45)	86.7 (11.55)	0.71	0.39 (1.07)	0.73	0.38 (1.11)
	Meta-analysis	1662	85.8 (12.29)	85.9 (14.29)	95.5 (18.59)	0.38	0.63 (0.71)	0.51	0.48 (0.74)
	Healthy subjects[3]	1048	81.4 (1.41)	80.5 (7.85)	85.0 (7.07)	0.11	–1.06 (0.66)	0.08	–1.15 (0.67)
Heart rate (bpm)	CADCZ	670	75.1 (5.51)	76.2 (5.83)	69.3 (3.06)	0.15	0.79 (0.54)	0.07	1.03 (0.56)
	Healthy subjects[3]	833	75.0 (8.49)	73.3 (8.06)	70.0 (0.00)	**0.04**	–1.61 (0.79)	**0.05**	–1.60 (0.80)
Intima-media thickness (mm)	CADCZ	670	0.63 (0.20)	0.65 (0.20)	0.87 (0.29)	0.20	0.03 (0.02)	0.33	0.02 (0.02)
	CADCZ controls	413	0.58 (0.17)	0.54 (0.15)	1.20 (0.00)	0.42	–0.02 (0.02)	0.18	–0.03 (0.02)

[1]For association analysis with SBP and DBP regression testing for a linear trend of marker alleles was performed with age, sex and BMI as covariates. Association analysis with heart rate was performed with sex as covariate, and intima-media thickness without covariates.
[2] Correction for antihypertensive treatment was implemented to all treated patients.
[3] Pooled HYPEST and CADCZ control subjects, who had no personal history of cardiovascular disease, including essential hypertension, myocardial infarction, coronary artery disease, stroke, and had never been prescribed cardiovascular medications. Previously, no population differentiation was detected between HYPEST and CADCZ study subjects.
n – number of individuals; Significant differences have been highlighted in bold, **p<0.05**

Table 15. Association between serum lipid biomarkers and the carrier status of the 14bp indel in *NCX1* intron 2

			WT/WT	WT/indel	indel/indel	Association testing using linear regression[1] Additive model		Dominant model	
	Sample	n	Mean (±SD)	Mean (±SD)	Mean (±SD)	P-value	Effect (SE)	P-value	Effect (SE)
Total cholesterol (mmol/L)	HYPEST	459	5.92 (1.15)	5.85 (1.24)	6.40 (2.47)	0.86	–0.02 (0.14)	0.73	–0.05 (0.15)
	CADCZ	670	5.50 (1.04)	5.48 (0.99)	5.55 (0.83)	0.11	0.17 (0.11)	0.14	0.17 (0.11)
	Meta-analysis	1129	5.67 (1.10)	5.63 (1.10)	5.98 (1.71)	0.25	0.10 (0.09)	0.33	0.09 (0.09)
	Healthy subjects[3]	431	5.59 (1.07)	5.36 (0.98)	6.37 (0.00)	0.18	–0.19 (0.14)	0.14	–0.22 (0.15)
HDL (mmol/L)	HYPEST	458	1.52 (0.42)	1.46 (0.42)	1.38 (0.29)	0.23	–0.06 (0.05)	0.24	–0.06 (0.05)
	CADCZ	670	1.41 (0.39)	1.41 (0.42)	1.42 (0.29)	0.81	–0.009 (0.04)	0.82	–0.01 (0.04)
	Meta-analysis	1128	1.45 (0.41)	1.43 (0.42)	1.40 (0.26)	0.35	–0.03 (0.03)	0.37	–0.03 (0.03)
	Healthy subjects[3]	431	1.52 (0.40)	1.61 (0.45)	1.70 (0.00)	0.10	0.09 (0.05)	0.10	0.09 (0.06)
LDL (mmol/L)	HYPEST	459	4.04 (1.03)	4.04 (1.00)	4.66 (2.58)	0.73	0.04 (0.12)	0.88	0.02 (0.13)
	CADCZ	651	3.33 (0.89)	3.22 (0.85)	3.22 (0.63)	0.66	–0.09 (0.20)	0.57	–0.12 (0.21)
	Meta-analysis	1110	3.63 (1.01)	3.54 (1.00)	3.94 (1.85)	0.94	0.01 (0.11)	0.87	–0.02 (0.11)
	Healthy subjects[3]	427	3.41 (0.91)	3.13 (0.85)	3.94 (0.00)	0.06	–0.23 (0.12)	**0.04**	–0.26 (0.13)
Triglycerides (mmol/L)	HYPEST	458	1.77 (1.66)	1.82 (1.02)	1.29 (0.53)	0.99	0.003 (0.19)	0.89	0.03 (0.20)
	CADCZ	670	1.72 (1.19)	1.94 (1.27)	2.00 (1.16)	**0.04**	0.25 (0.12)	**0.04**	0.26 (0.12)
	Meta-analysis	1128	1.74 (1.40)	1.89 (1.18)	1.65 (0.90)	0.08	0.18 (0.10)	0.07	0.19 (0.10)
	Healthy subjects[3]	431	1.52 (1.12)	1.40 (0.81)	1.62 (0.00)	0.44	–0.11 (0.15)	0.42	–0.12 (0.15)

[1] For association analysis with serum lipids regression testing for a linear trend of marker alleles was performed with age, sex and BMI as covariates.
[2] Correction for the treatment with lipid-lowering medication was implemented as described (Janosikova, et al., 2003).
[3] Pooled HYPEST and CADCZ control subjects with available records for serum lipids. The individuals had no personal history of cardiovascular disease, including essential hypertension, myocardial infarction, coronary artery disease, stroke, and had never been prescribed cardiovascular medications. Previously, no population differentiation was detected between HYPEST and CADCZ study subjects
n – number of individuals; Significant differences have been highlighted in bold, **p<0.05**

3.5. Discussion

A classical method to search for the genetic risk factors of the disease of interest is based on the hypothesis of previous knowledge of biological pathways combined with the candidate gene approach. There have been doubts about the reliability of candidate gene approach mainly due to the inadequate knowledge of pathways but recent years have given great insight on this feature. Variations detected within candidate genes as well as in regulatory regions within these pathways may give new insight to the disease susceptibility in the studied populations. Sequencing of one potential CVD candidate gene, the human *PNMT* gene, among two European populations revealed low genetic variation patterns representing only by seven SNPs compared to 55 polymorphisms currently located in NCBI database. This kind of difference may raise a question of the previous reliability of the available information in databases. By contrast, to date the sequencing of 1000 genomes have revealed only 15 SNPs along the entire human *PNMT* gene (http://www.ncbi.nlm.nih.gov/gene). The identified variation pattern in the current study is concordant with the data from 1000 genome and indicates the important role of PNMT-synthesized epinephrine in the regulation of cardiovascular and metabolic functions.

Comparative genetics studies have noted several essential gene regulatory elements that are conserved among species (Drake, et al., 2006). Targeting the evolutionarily conserved non-coding regions in candidate genes for complex disease may pinpoint novel disease susceptibility variants and novel regulatory elements contributing to gene expression profile. Majority of variations, detected among conserved non-coding sequences of 25 CVD candidate genes in the current study were rare, which is concordant with the purifying selection acting on these regions. Similarly, low variation was also observed in promoter regions of 29 genes where the highest number of SNPs was found in duplicate genes.

Currently the most used hypotheses free approach GWAS, based on the CD-CV analyses, is used to map loci in susceptibility to complex disease, where SNPs are used as markers. To date, GWAS have provided hundreds of common variants in susceptibility to complex traits. However, these results have shown to explain only a small fraction of the inherited risk of complex disease. For example, only ~5% of type two diabetes and ~10% of the Crohn's disease, have been explained by the common risk variants (Altshuler, et al., 2008). To uncover the rest of genetic component of the inheritance of complex disease has remained challenging. Currently rare alleles with severe effects have been proposed to cause many medical conditions like breast and ovarian cancer (*BRCA1, BRCA2*), hearing loss (Dror and Avraham, 2009), mental illness (autism, schizophrenia) (Walsh, et al., 2008) and adverse changes in lipid metabolism (Cohen, et al., 2004).

In addition to SNPs, studied extensively in etiology of complex diseases, a whole-genome sequencing approach has unveiled several other unexplored

genetic variants located in the human genome with potential impact on susceptibility of human complex disease. These unexplored variants are mainly structural variations that include large copy number variations, small insertions and deletions. Structural variations in human genome are assumed to have higher significance than single-nucleotide variations altering more than one base pair in DNA sequence and therefore having a higher impact on gene dysfunction as well as on disease development. Also, a number of these mutations have likely no functional role in the genome acting neutral. Polymorphism screening among CNR regions of CVD candidate genes revealed two common indel variants located within intronic regions of *NCX1* and *WNK1*. A 14bp indel located in *NCX1* intron 2 and a novel *Alu*-sequence in *WNK1* intron 10. Genotyping of the CNR harbouring the 14bp indel of *NCX1* gene revealed its hypervariable nature (represented by seven different alleles) and showed association with the increased risk for CAD and elevated triglyceride levels. Furthermore, the comparison of the hypervariable region with chimpanzee and rhesus macaque genomes uncovered that the major human variant was actually human specific deletion. Either natural selection or genetic drift may have triggered the enrichment of 14 bp advantageous deletion variant among humans. *In silico* analysis of studied 348bp region revealed that majority of detected variants had overlapping breakpoints with other variations (Figure 8) and harboured multiple sequence motifs previously shown to lead to higher mutation rates. Consequently, genomic regions with indel heterozygosity are shown to be prone to the double stranded DNA breaks and therefore are targeted to mutational repair which in turn may lead to the higher mutational rate (Lercher and Hurst, 2002; Tian, et al., 2008). Research among different model organisms such as fruit flies and Caenorhabditis elegans, have shown that indels comprise 16% up to 25% of all known genetic variations (Berger, et al., 2001; Wicks, et al., 2001) and therefore, are expected to be abundant in the same level in humans (Dawson, et al., 2001). In comparative analysis of chromosome 22 in chimpanzee and human chromosome 21 have revealed ~68 000 indels, where the majority were small in size. An excess of ~300bp regions, among identified indels in comparative analysis of chimpanzee chromosome 22, were represented by the short transposable *Alu*-elements (Watanabe, et al., 2004). Two *Alu*-sequences were also detected in *PNMT* and *WNK1* genes. An ancient *Alu*-element insertion, in *PNMT* intron 1, revealed potential regulatory role on gene expression profile. Similar, supportive mechanisms have also been shown where *Alu* sequences possessing regulatory elements were inserted and now are acting as central control/enhancement of transcription (Britten, 1996). The potential functional role of the young polymorphic *Alu*-sequence located in *WNK1* intron 10 with cardiovascular traits will further be studied among different European populations (Putku, Kepp, manuscript in preparation).

To date performing a high throughput indel association studies are greatly limited due to the unavailable detection technology. The whole-genome

sequencing data from 1000 Genomes Project as well as other individual whole-genome sequencing studies within and between different populations will reveal considerable locus complexity and provide insight into the different mutational processes that have shaped the human genome. The improved map of human genetic variation will provide an invaluable opportunity to consider the analysis of gene-gene, gene-pathway, gene-environment and genetic polymorphism-polymorphism interactions to understand the complexity of human disease.

CONCLUSIONS

Summary of conclusions of this study:

1. Selection of the genes from literature reports and databases with the potential role in CVD pathogenesis revealed that the high number of genes clustered among three major organ systems: heart, kidney and sympathetic nervous system. The list include loci responsible for monogenic and complex forms of high and low blood pressure, CAD, metabolic syndrome and other related traits. Analysis of the gene selection confirmed their importance in CVD development. Seven distinct KEGG pathways with known role in CVD were defined: calcium signaling pathway, hypertrophic cardiomyopathy, aldosterone-regulated sodium reabsorption, neuroactive ligand-receptor interaction, renin-angiotensin system, vascular smooth muscle contraction and type II diabetes mellitus.

2. The characterization of the variation pattern of the human *PNMT* showed low genetic variation along the entire gene. *In silico* analysis of the intronic regions of human *PNMT* gene identified a major human-specific gene regulatory unit GRU inserted by *Alu*-mediated transfer. The detected low variation pattern together with the *in silico* predicted regulatory complex suggested that the differences in *PNMT* expression between patients and controls may probably be determined not only by the polymorphisms of this gene, but rather by the interplay of gene expression regulators that may vary among individuals.

3. The screening for polymorphisms within promoter and conserved non-coding sequences in selected CVD candidate genes showed relatively low diversity among the two targeted Eastern-European populations. Detailed characterization of CNR regions in human *NCX1* intronic region identified a hypervariable genomic fragment harbouring multiple human-specific polymorphisms enriched by short indels.

4. Studying the role of polymorphisms detected in human *PNMT* and conserved non-coding regions located in *NCX1* genes in CVD development revealed suggestive association with essential hypertension in combined Eastern-European sample. An indel variant located in intronic CNR of *NCX1* gene showed significant association with coronary artery disease and was even higher between the indel variant and among patients additionally diagnosed with metabolic syndrome in Czech population. A suggestive evidence of association with indel variant and serum triglyceride levels was observed in Czechs as well as with indel variant and heart rate and LDL levels in healthy control individuals among both cohorts.

Current research, as well as other recent studies have shown that non-SNP variations are a substantial source of polymorphism in humans and may have larger role in complex disease (like CVD) than previously thought. The studies of multiple cardiovascular traits refer to the pleiotropic role of human *NCX1* gene where the same gene and variant (14bp indel) may affect different cardiovascular traits simultaneously.

REFERENCES

Abiola O, Angel JM, Avner P, Bachmanov AA, Belknap JK, Bennett B, Blankenhorn EP, Blizard DA, Bolivar V, Brockmann GA and others. 2003. The nature and identification of quantitative trait loci: a community's view. Nat Rev Genet 4(11):911–6.

Adeyemo A, Gerry N, Chen G, Herbert A, Doumatey A, Huang H, Zhou J, Lashley K, Chen Y, Christman M and others. 2009. A genome-wide association study of hypertension and blood pressure in African Americans. PLoS Genet 5(7):e1000564.

Agarwal A, Williams GH, Fisher ND. 2005. Genetics of human hypertension. Trends Endocrinol Metab 16(3):127–33.

Altshuler D, Daly MJ, Lander ES. 2008. Genetic mapping in human disease. Science 322(5903):881–8.

Arking DE, Chakravarti A. 2009. Understanding cardiovascular disease through the lens of genome-wide association studies. Trends Genet 25(9):387–94.

Assmann G, Cullen P, Schulte H. 2002. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Munster (PROCAM) study. Circulation 105(3):310–5.

Aziz NA, Jurgens CK, Landwehrmeyer GB, van Roon-Mom WM, van Ommen GJ, Stijnen T, Roos RA. 2009. Normal and mutant HTT interact to affect clinical severity and progression in Huntington disease. Neurology 73(16):1280–5.

Baira E, Greshock J, Coukos G, Zhang L. 2008. Ultraconserved elements: genomics, function and disease. RNA Biol 5(3):132–4.

Ball EV, Stenson PD, Abeysinghe SS, Krawczak M, Cooper DN, Chuzhanova NA. 2005. Microdeletions and microinsertions causing human genetic disease: common mechanisms of mutagenesis and the role of local DNA sequence complexity. Hum Mutat 26(3):205–13.

Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D. 2004. Ultraconserved elements in the human genome. Science 304(5675):1321–5.

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR and others. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456(7218):53–9.

Berger J, Suzuki T, Senti KA, Stubbs J, Schaffner G, Dickson BJ. 2001. Genetic mapping with SNP markers in Drosophila. Nat Genet 29(4):475–81.

Bhangale TR, Stephens M and Nickerson DA. 2006. Automating resequencing-based detection of insertion-deletion polymorphisms. Nat Genet 38(2):1457–62.

Binder A. 2006. Identification of genes for a complex trait: examples from hypertension. Curr Pharm Biotechnol 7(1):1–13.

Blaustein MP. 1993. Physiological effects of endogenous ouabain: control of intracellular Ca2+ stores and cell responsiveness. Am J Physiol 264(6 Pt 1):C1367–87.

Blaustein MP, Lederer WJ. 1999. Sodium/calcium exchange: its physiological implications. Physiol Rev 79(3):763–854.

Borecki IB, Suarez BK. 2001. Linkage and association: basic concepts. Adv Genet 42:45–66.

Breslow JL. 2000. Genetics of lipoprotein abnormalities associated with coronary artery disease susceptibility. Annu Rev Genet 34:233–254.

Britten RJ. 1996. DNA sequence insertion and evolutionary variation in gene regulation. Proc Natl Acad Sci U S A 93(18):9374–7.
Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ, Comuzzie AG, Blangero J, Nurnberg P, Reis A, Riegger GA and others. 2002. A comprehensive linkage analysis for myocardial infarction and its related risk factors. Nat Genet 30(2):210–4.
Burton P, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI OW, Samani NJ, Todd JA, Donnelly P, and others. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447(7145):661–78.
Butler JE, Kadonaga JT. 2002. The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes Dev 16(20):2583–92.
Cahill AL, Eertmoed AL, Mangoura D, Perlman RL. 1996. Differential regulation of phenylethanolamine N-methyltransferase expression in two distinct subpopulations of bovine chromaffin cells. J Neurochem 67(3):1217–24.
Cambien F, Poirier O, Nicaud V, Herrmann SM, Mallet C, Ricard S, Behague I, Hallet V, Blanc H, Loukaci V and others. 1999. Sequence diversity in 36 candidate genes for cardiovascular disorders. Am J Hum Genet 65(1):183–91.
Cardiology ESoH-ESo. 2003. 2003 European Society of Hypertension-European Society of Cardiology guidelines for the management of arterial hypertension. J Hypertens 21(6):1011–53.
Cardon LR, Bell JI. 2001. Association study designs for complex diseases. Nat Rev Genet 2(2):91–9.
Cardon LR, Palmer LJ. 2003. Population stratification and spurious allelic association. Lancet 361(9357):598–604.
Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, Semple CA, Taylor MS, Engstrom PG, Frith MC and others. 2006. Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet 38(6):626–35.
Caulfield M, Munroe P, Pembroke J, Samani N, Dominiczak A, Brown M, Benjamin N, Webster J, Ratcliffe P, O'Shea S and others. 2003. Genome-wide mapping of human loci for essential hypertension. Lancet 361(9375):2118–23.
Chakravarti A. 1984. Aspects of linkage analysis using RFLP's. Prog Clin Biol Res 147:293–5.
Chen FC, Chen CJ, Li WH, Chuang TJ. 2007. Human-specific insertions and deletions inferred from mammalian genome sequences. Genome Res 17(1):16–22.
Cheung VG, Spielman RS. 2009. Genetics of human gene expression: mapping DNA variants that influence gene expression. Nat Rev Genet 10(9):595–604.
Cho YS, Go MJ, Kim YJ, Heo JY, Oh JH, Ban HJ, Yoon D, Lee MH, Kim DJ, Park M and others. 2009. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nat Genet 41(5):527–34.
Chuzhanova NA, Anassis EJ, Ball EV, Krawczak M, Cooper DN. 2003. Meta-analysis of indels causing human genetic disease: mechanisms of mutagenesis and the role of local DNA sequence complexity. Hum Mutat 21(1):28–44.
Cifkova R, Erdine S, Fagard R, Farsang C, Heagerty AM, Kiowski W, Kjeldsen S, Luscher T, Mallion JM, Mancia G and others. 2003. Practice guidelines for primary care physicians: 2003 ESH/ESC hypertension guidelines. J Hypertens 21(10):1779–86.

Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. 2004. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. Science 305(5685):869–72.

Cowley AW, Jr. 2006. The genetic dissection of essential hypertension. Nat Rev Genet 7(11):829–40.

Cui J, Zhou X, Chazaro I, DeStefano AL, Manolis AJ, Baldwin CT, Gavras H. 2003. Association of polymorphisms in the promoter region of the PNMT gene with essential hypertension in African Americans but not in whites. Am J Hypertens 16(10):859–63.

Dawson E, Chen Y, Hunt S, Smink LJ, Hunt A, Rice K, Livingston S, Bumpstead S, Bruskiewich R, Sham P and others. 2001. A SNP resource for human chromosome 22: extracting dense clusters of SNPs from the genomic sequence. Genome Res 11(1):170–8.

de la Chaux N, Messer PW, Arndt PF. 2007. DNA indels in coding regions reveal selective constraints on protein evolution in the human lineage. BMC Evol Biol 7:191.

de Sanctis L, Corrias A, Romagnolo D, Di Palma T, Biava A, Borgarello G, Gianino P, Silvestro L, Zannini M, Dianzani I. 2004. Familial PAX8 small deletion (c.989_992delACCC) associated with extreme phenotype variability. J Clin Endocrinol Metab 89(11):5669–74.

Delles C, McBride MW, Graham D, Padmanabhan S, Dominiczak AF. 2009. Genetics of hypertension: From experimental animals to humans. Biochim Biophys Acta Dec 24.

Dermitzakis ET, Reymond A, Antonarakis SE. 2005. Conserved non-genic sequences – an unexpected feature of mammalian genomes. Nat Rev Genet 6(2):151–7.

Dermitzakis ET, Reymond A, Lyle R, Scamuffa N, Ucla C, Deutsch S, Stevenson BJ, Flegel V, Bucher P, Jongeneel CV and others. 2002. Numerous potentially functional but non-genic conserved sequences on human chromosome 21. Nature 420(6915):578–82.

Dermitzakis ET, Reymond A, Scamuffa N, Ucla C, Kirkness E, Rossier C, Antonarakis SE. 2003. Evolutionary discrimination of mammalian conserved non-genic sequences (CNGs). Science 302(5647):1033–5.

DeSilva U, Elnitski L, Idol JR, Doyle JL, Gan W, Thomas JW, Schwartz S, Dietrich NL, Beckstrom-Sternberg SM, McDowell JC and others. 2002. Generation and comparative analysis of approximately 3.3 Mb of mouse genomic sequence orthologous to the region of human chromosome 7q11.23 implicated in Williams syndrome. Genome Res 12(1):3–15.

Dhandapany PS, Sadayappan S, Xue Y, Powell GT, Rani DS, Nallari P, Rai TS, Khullar M, Soares P, Bahl A and others. 2009. A common MYBPC3 (cardiac myosin binding protein C) variant associated with cardiomyopathies in South Asia. Nat Genet 41(2):187–91.

Dominiczak AF, Graham D, McBride MW, Brain NJ, Lee WK, Charchar FJ, Tomaszewski M, Delles C, Hamilton CA. 2005. Corcoran Lecture. Cardiovascular genomics and oxidative stress. Hypertension 45(4):636–42.

Drake JA, Bird C, Nemesh J, Thomas DJ, Newton-Cheh C, Reymond A, Excoffier L, Attar H, Antonarakis SE, Dermitzakis ET and others. 2006. Conserved noncoding sequences are selectively constrained and not mutation cold spots. Nat Genet 38(2):223–7.

Dror AA, Avraham KB. 2009. Hearing loss: mechanisms revealed by genetics and cell biology. Annu Rev Genet 43:411–37.

Dunn J, Elias CL, Le HD, Omelchenko A, Hryshko LV, Lytton J. 2002. The molecular determinants of ionic regulatory differences between brain and kidney Na+/Ca2+ exchanger (NCX1) isoforms. J Biol Chem 277(37):33957–62.

Ehret GB. 2010. Genome-wide association studies: contribution of genomics to understanding blood pressure and essential hypertension. Curr Hypertens Rep 12(1):17–25.

Eichler EE, Nickerson DA, Altshuler D, Bowcock AM, Brooks LD, Carter NP, Church DM, Felsenfeld A, Guyer M, Lee C and others. 2007. Completing the map of human genetic variation. Nature 447(7141):161–5.

Faa V, Bettoli PP, Demurtas M, Zanda M, Ferri V, Cao A, Rosatelli MC. 2006. A new insertion/deletion of the cystic fibrosis transmembrane conductance regulator gene accounts for 3.4% of cystic fibrosis mutations in Sardinia: implications for population screening. J Mol Diagn 8(4):499–503.

Fischer M, Broeckel U, Holmer S, Baessler A, Hengstenberg C, Mayer B, Erdmann J, Klein G, Riegger G, Jacob HJ and others. 2005. Distinct heritable patterns of angiographic coronary artery disease in families with myocardial infarction. Circulation 111(7):855–62.

Floras JS. 1993. Clinical aspects of sympathetic activation and parasympathetic withdrawal in heart failure. J Am Coll Cardiol 22(4 Suppl A):72A–84A.

Francke S, Manraj M, Lacquemant C, Lecoeur C, Lepretre F, Passa P, Hebe A, Corset L, Yan SL, Lahmidi S and others. 2001. A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27. Hum Mol Genet 10(24):2751–65.

Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM and others. 2007. A second generation human haplotype map of over 3.1 million SNPs. Nature 449(7164):851–61.

Frazer KA, Murray SS, Schork NJ, Topol EJ. 2009. Human genetic variation and its contribution to complex traits. Nat Rev Genet 10(4):241–51.

Giardine B, Elnitski L, Riemer C, Makalowska I, Schwartz S, Miller W, Hardison RC. 2003. GALA, a database for genomic sequence alignments and annotations. Genome Res 13(4):732–41.

Glazier AM, Nadeau JH, Aitman TJ. 2002. Finding genes that underlie complex traits. Science 298(5602):2345–9.

Goldstein LB, Adams R, Alberts MJ, Appel LJ, Brass LM, Bushnell CD, Culebras A, DeGraba TJ, Gorelick PB, Guyton JR and others. 2006. Primary prevention of ischemic stroke: a guideline from the American Heart Association/American Stroke Association Stroke Council: cosponsored by the Atherosclerotic Peripheral Vascular Disease Interdisciplinary Working Group; Cardiovascular Nursing Council; Clinical Cardiology Council; Nutrition, Physical Activity, and Metabolism Council; and the Quality of Care and Outcomes Research Interdisciplinary Working Group. Circulation 113(24):e873–923.

Grigorova M, Punab M, Ausmees K, Laan M. 2008. FSHB promoter polymorphism within evolutionary conserved element is associated with serum FSH level in men. Hum Reprod 23(9):2160–6.

Grigorova M, Punab M, Poolamets O, Kelgo P, Ausmees K, Korrovits P, Vihljajev V, Laan M. 2009. Increased Prevalance of the –211 T allele of follicle stimulating hormone (FSH) beta subunit promoter polymorphism and lower serum FSH in infertile men. J Clin Endocrinol Metab 95(1):100–8.

Haavik J, Blau N, Thony B. 2008. Mutations in human monoamine-related neurotransmitter pathway genes. Hum Mutat 29(7):891–902.
Harrap SB, Zammit KS, Wong ZY, Williams FM, Bahlo M, Tonkin AM, Anderson ST. 2002. Genome-wide linkage analysis of the acute coronary syndrome suggests a locus on chromosome 2. Arterioscler Thromb Vasc Biol 22(5):874–8.
Helgadottir A, Manolescu A, Thorleifsson G, Gretarsdottir S, Jonsdottir H, Thorsteinsdottir U, Samani NJ, Gudmundsson G, Grant SF, Thorgeirsson G and others. 2004. The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke. Nat Genet 36(3):233–9.
Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Jonasdottir A, Sigurdsson A, Baker A, Palsson A and others. 2007. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 316(5830):1491–3.
Hilbert P, Lindpaintner K, Beckmann JS, Serikawa T, Soubrier F, Dubay C, Cartwright P, De Gouyon B, Julier C, Takahasi S and others. 1991. Chromosomal mapping of two genetic loci associated with blood-pressure regulation in hereditary hypertensive rats. Nature 353(6344):521–9.
Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. 2009. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 106(23):9362–7.
Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. 2002. A comprehensive review of genetic association studies. Genet Med 4(2):45–61.
Hodge SE. 1993. Linkage analysis versus association analysis: distinguishing between two models that explain disease-marker associations. Am J Hum Genet 53(2):367–84.
Hong KW, Jin HS, Cho YS, Lee JY, Lee JE, Cho NH, Shin C, Lee SH, Park HK, Oh B. 2009. Replication of the Wellcome Trust genome-wide association study on essential hypertension in a Korean population. Hypertens Res 32(7):570–4.
Hong KW, Jin HS, Lim JE, Cho YS, Go MJ, Jung J, Lee JE, Choi J, Shin C, Hwang SY and others. 2010a. Non-synonymous single-nucleotide polymorphisms associated with blood pressure and hypertension. J Hum Hypertens Feb 11.
Hong KW, Jin HS, Lim JE, Kim S, Go MJ, Oh B. 2010b. Recapitulation of two genomewide association studies on blood pressure and essential hypertension in the Korean population. J Hum Genet 55(6):336–41.
Hubler TR, Scammell JG. 2004. Intronic hormone response elements mediate regulation of FKBP5 by progestins and glucocorticoids. Cell Stress Chaperones 9(3):243–52.
Hubner N, Wallace CA, Zimdahl H, Petretto E, Schulz H, Maciver F, Mueller M, Hummel O, Monti J, Zidek V and others. 2005. Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. Nat Genet 37(3):243–53.
Hurtado C, Prociuk M, Maddaford TG, Dibrov E, Mesaeli N, Hryshko LV, Pierce GN. 2006. Cells expressing unique Na+/Ca2+ exchange (NCX1) splice variants exhibit different susceptibilities to Ca2+ overload. Am J Physiol Heart Circ Physiol 290(5):H2155–62.
International Diabetes Federation. http://www.idf.org/.
Iwai N, Baba S, Mannami T, Ogihara T, Ogata J. 2002. Association of a sodium channel alpha subunit promoter variant with blood pressure. J Am Soc Nephrol 13(1):80–5.

Iwamoto T, Kita S, Uehara A, Imanaga I, Matsuda T, Baba A, Katsuragi T. 2004. Molecular determinants of Na+/Ca2+ exchange (NCX1) inhibition by SEA0400. J Biol Chem 279(9):7544–53.

Iwanaga Y, Ono K, Takagi S, Terashima M, Tsutsumi Y, Mannami T, Yasui N, Goto Y, Nonogi H, Iwai N. 2004. Association analysis between polymorphisms of the lymphotoxin-alpha gene and myocardial infarction in a Japanese population. Atherosclerosis 172(1):197–8.

Jain S, Li Y, Patil S, Kumar A. 2005. A single-nucleotide polymorphism in human angiotensinogen gene is associated with essential hypertension and affects glucocorticoid induced promoter activity. J Mol Med 83(2):121–31.

Janosikova B, Pavlikova M, Kocmanova D, Vitova A, Vesela K, Krupkova L, Kahleova R, Krijt J, Kraml P, Hyanek J and others. 2003. Genetic variants of homocysteine metabolizing enzymes and the risk of coronary artery disease. Mol Genet Metab 79(3):167–75.

Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, Newton-Cheh C, State MW, Levy D, Lifton RP. 2008. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat Genet 40(5):592–9.

Jin Q, Agrawal L, Meyer L, Tubiana R, Theodorou I, Alkhatib G. 2008. CCR5Delta32 59537-G/A promoter polymorphism is associated with low translational efficiency and the loss of CCR5Delta32 protective effects. J Virol 82(5):2418–26.

Johnson AD, O'Donnell CJ. 2009. An open access database of genome-wide association results. BMC Med Genet 10:6.

Jorgensen TJ, Ruczinski I, Kessing B, Smith MW, Shugart YY, Alberg AJ. 2009. Hypothesis-driven candidate gene association studies: practical design and analytical considerations. Am J Epidemiol 170(8):986–93.

Julier C, Delepine M, Keavney B, Terwilliger J, Davis S, Weeks DE, Bui T, Jeunemaitre X, Velho G, Froguel P and others. 1997. Genetic susceptibility for human familial essential hypertension in a region of homology with blood pressure linkage on rat chromosome 10. Hum Mol Genet 6(12):2077–85.

Karban AS, Okazaki T, Panhuysen CI, Gallegos T, Potter JJ, Bailey-Wilson JE, Silverberg MS, Duerr RH, Cho JH, Gregersen PK and others. 2004. Functional annotation of a novel NFKB1 promoter polymorphism that increases risk for ulcerative colitis. Hum Mol Genet 13(1):35–45.

Kasai T, Miyauchi K, Kubota N, Tamura H, Kojima T, Yokoyama K, Kurata T, Daida H. 2008. The relationship between the metabolic syndrome defined by various criteria and the extent of coronary artery disease. Atherosclerosis 197(2):944–50.

Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand S, Engert JC, Samani NJ, Schunkert H and others. 2009. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 41(3):334–41.

Kato N, Julier C. 1999. Linkage mapping for hypertension susceptibility genes. Curr Hypertens Rep 1(1):15–24.

Kidd JM, Cooper GM, Donahue WF, Hayden HS, Sampas N, Graves T, Hansen N, Teague B, Alkan C, Antonacci F and others. 2008. Mapping and sequencing of structural variation from eight human genomes. Nature 453(7191):56–64.

Knight J, Munroe PB, Pembroke JC, Caulfield MJ. 2003. Human chromosome 17 in essential hypertension. Ann Hum Genet 67(Pt 2):193–206.

Koike G, Jacob HJ, Krieger JE, Szpirer C, Hoehe MR, Horiuchi M, Dzau VJ. 1995. Investigation of the phenylethanolamine N-methyltransferase gene as a candidate gene for hypertension. Hypertension 26(4):595–601.

Kokubo Y, Inamoto N, Tomoike H, Kamide K, Takiuchi S, Kawano Y, Tanaka C, Katanosaka Y, Wakabayashi S, Shigekawa M and others. 2004. Association of genetic polymorphisms of sodium-calcium exchanger 1 gene, NCX1, with hypertension in a Japanese general population. Hypertens Res 27(10):697–702.

Kraev A, Chumakov I, Carafoli E. 1996. The organization of the human gene NCX1 encoding the sodium-calcium exchanger. Genomics 37(1):105–12.

Kruglyak L, Nickerson DA. 2001. Variation is the spice of life. Nat Genet 27(3):234–6.

Ku CS, Loy EY, Salim A, Pawitan Y, Chia KS. The discovery of human genetic variations and their use as disease markers: past, present and future. J Hum Genet 55(7):403–15.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W and others. 2001. Initial sequencing and analysis of the human genome. Nature 409(6822):860–921.

Lawes CM, Vander Hoorn S, Rodgers A. 2008. Global burden of blood-pressure-related disease, 2001. Lancet 371(9623):1513–8.

Lercher MJ, Hurst LD. 2002. Human SNP variability and mutation rate are higher in regions of high recombination. Trends Genet 18(7):337–40.

Lewander A, Butchi AK, Gao J, He LJ, Lindblom A, Arbman G, Carstensen J, Zhang ZY, Sun XF. 2007. Polymorphism in the promoter region of the NFKB1 gene increases the risk of sporadic colorectal cancer in Swedish but not in Chinese populations. Scand J Gastroenterol 42(11):1332–8.

Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, Glazer NL, Morrison AC, Johnson AD, Aspelund T and others. 2009. Genome-wide association study of blood pressure and hypertension. Nat Genet May 10.

Levy D, Larson MG, Benjamin EJ, Newton-Cheh C, Wang TJ, Hwang SJ, Vasan RS, Mitchell GF. 2007a. Framingham Heart Study 100K Project: genome-wide associations for blood pressure and arterial stiffness. BMC Med Genet 8 Suppl 1:S3.

Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G and others. 2007b. The diploid genome sequence of an individual human. PLoS Biol 5(10):e254.

Lifton RP, Gharavi AG, Geller DS. 2001. Molecular mechanisms of human hypertension. Cell 104(4):545–56.

Liu S. 2010. Increasing alternative promoter repertories is positively associated with differential expression and disease susceptibility. PLoS One 5(3):e9482.

Lloyd-Jones DM, Nam BH, D'Agostino RB, Sr., Levy D, Murabito JM, Wang TJ, Wilson PW, O'Donnell CJ. 2004. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. JAMA 291(18):2204–11.

Luft FC. 2001. Twins in cardiovascular genetic research. Hypertension 37(2 Part 2):350–6.

Lusis AJ. 2000. Atherosclerosis. Nature 407(6801):233–41.

Lusis AJ, Mar R, Pajukanta P. 2004. Genetics of atherosclerosis. Annu Rev Genomics Hum Genet 5:189–218.

Mann MB, Wu S, Rostamkhani M, Tourtellotte W, MacMurray J, Comings DE. 2001. Phenylethanolamine N-methyltransferase (PNMT) gene and early-onset Alzheimer disease. Am J Med Genet 105(4):312–6.

Mann MB, Wu S, Rostamkhani M, Tourtellotte W, MacMurray JP, Comings DE. 2002. Association between the phenylethanolamine N-methyltransferase gene and multiple sclerosis. J Neuroimmunol 124(1–2):101–5.

McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR and others. 2007. A common allele on chromosome 9 associated with coronary heart disease. Science 316(5830):1488–91.

Mills RE, Luttig CT, Larkins CE, Beauchamp A, Tsui C, Pittard WS, Devine SE. 2006. An initial map of insertion and deletion (INDEL) variation in the human genome. Genome Res 16(9):1182–90.

Morton NE. 1955. Sequential tests for the detection of linkage. Am J Hum Genet 7(3):277–318.

Naber CK, Siffert W. 2004. Genetics of human arterial hypertension. Minerva Med 95(5):347–56.

Newhouse SJ, Wallace C, Dobson R, Mein C, Pembroke J, Farrall M, Clayton D, Brown M, Samani N, Dominiczak A and others. 2005. Haplotypes of the WNK1 gene associate with blood pressure variation in a severely hypertensive population from the British Genetics of Hypertension study. Hum Mol Genet 14(13):1805–14.

Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, Najjar SS, Zhao JH, Heath SC, Eyheramendy S and others. 2009. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet May 10.

Noble D, Herchuelz A. 2007. Role of Na/Ca exchange and the plasma membrane Ca2+–ATPase in cell function. Conference on Na/Ca exchange. EMBO Rep 8(3):228–32.

Nordestgaard BG, Benn M, Schnohr P, Tybjaerg-Hansen A. 2007. Nonfasting triglycerides and risk of myocardial infarction, ischemic heart disease, and death in men and women. JAMA 298(3):299–308.

Org E, Eyheramendy S, Juhanson P, Gieger C, Lichtner P, Klopp N, Veldre G, Doring A, Viigimaa M, Sober S and others. 2009. Genome-wide scan identifies CDH13 as a novel susceptibility locus contributing to blood pressure determination in two European populations. Hum Mol Genet 18(12):2288–96.

Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, Sato H, Sato H, Hori M, Nakamura Y and others. 2002. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. Nat Genet 32(4):650–4.

Othman M, Notley C, Lavender FL, White H, Byrne CD, Lillicrap D, O'Shaughnessy DF. 2005. Identification and functional characterization of a novel 27-bp deletion in the macroglycopeptide-coding region of the GPIBA gene resulting in platelet-type von Willebrand disease. Blood 105(11):4330–6.

Paul T, Li S, Khurana S, Leleiko NS, Walsh MJ. 2007. The epigenetic signature of CFTR expression is co-ordinated via chromatin acetylation through a complex intronic element. Biochem J 408(3):317–26.

Pedersen AG, Baldi P, Chauvin Y, Brunak S. 1999. The biology of eukaryotic promoter prediction – a review. Comput Chem 23(3–4):191–207.

Penagarikano O, Mulle JG, Warren ST. 2007. The pathophysiology of fragile x syndrome. Annu Rev Genomics Hum Genet 8:109–29.

Pethe V, Shekhar PV. 1999. Estrogen inducibility of c-Ha-ras transcription in breast cancer cells. Identification of functional estrogen-responsive transcriptional regulatory elements in exon 1/intron 1 of the c-Ha-ras gene. J Biol Chem 274(43):30969–78.

Pierreux CE, Rousseau GG, Lemaigre FP. 1999. Insulin inhibition of glucocorticoid-stimulated gene transcription: requirement for an insulin response element? Mol Cell Endocrinol 147(1–2):1–5.
Presneau N, Laplace-Marieze V, Sylvain V, Lortholary A, Hardouin A, Bernard-Gallon D, Bignon YJ. 1998. New mechanism of BRCA-1 mutation by deletion/insertion at the same nucleotide position in three unrelated French breast/ovarian cancer families. Hum Genet 103(3):334–9.
Pritchard JK, Rosenberg NA. 1999. Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet 65(1):220–8.
Quednau BD, Nicoll DA, Philipson KD. 1997. Tissue specificity and alternative splicing of the Na+/Ca2+ exchanger isoforms NCX1, NCX2, and NCX3 in rat. Am J Physiol 272(4 Pt 1):C1250-61.
Rasmussen HB, Timm S, Wang AG, Soeby K, Lublin H, Fenger M, Hemmingsen R, Werge T. 2006. Association between the CCR5 32-bp deletion allele and late onset of schizophrenia. Am J Psychiatry 163(3):507–11.
Reid LH, Gregg RG, Smithies O, Koller BH. 1990. Regulatory elements in the introns of the human HPRT gene are necessary for its expression in embryonic stem cells. Proc Natl Acad Sci U S A 87(11):4299–303.
Reimand J, Kull M, Peterson H, Hansen J, Vilo J. 2007. g: Profiler – a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Res 35(Web Server issue):W193–200.
Reis DJ, Morrison S, Ruggiero DA. 1988. The C1 area of the brainstem in tonic and reflex control of blood pressure. State of the art lecture. Hypertension 11(2 Pt 2):I8–13.
Risch N, Merikangas K. 1996. The future of genetic studies of complex human diseases. Science 273(5281):1516–7.
Rosario RF, Wesson DE. 2006. Primary hypertension and nephropathy. Curr Opin Nephrol Hypertens 15(2):130–4.
Rossi S, Sevignani C, Nnadi SC, Siracusa LD, Calin GA. 2008. Cancer-associated genomic regions (CAGRs) and noncoding RNAs: bioinformatics and therapeutic implications. Mamm Genome 19(7–8):526–40.
Rowntree RK, Vassaux G, McDowell TL, Howe S, McGuigan A, Phylactides M, Huxley C, Harris A. 2001. An element in intron 1 of the CFTR gene augments intestinal expression in vivo. Hum Mol Genet 10(14):1455–64.
Rutherford S, Johnson MP, Curtain RP, Griffiths LR. 2001. Chromosome 17 and the inducible nitric oxide synthase gene in human essential hypertension. Hum Genet 109:408–415.
Rutter JL, Mitchell TI, Buttice G, Meyers J, Gusella JF, Ozelius LJ, Brinckerhoff CE. 1998. A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter creates an Ets binding site and augments transcription. Cancer Res 58(23):5321–5.
Saleheen D, Frossard PM. 2008. The cradle of the deltaF508 mutation. J Ayub Med Coll Abbottabad 20(4):157–60.
Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE and others. 2007. Genomewide association analysis of coronary artery disease. N Engl J Med 357(5):443–53.
Sasaoka T, Kaneda N, Kurosava Y, Fujita K, Nagatsu T. 1989. Structure of human phenylethanolamine N-methyltransferase gene: existence of two types of mRNA with different transcription initiation sites. Neurochem Int. Vol. 15(No. 4):pp 555–565.

Sayed-Tabatabaei FA, Oostra BA, Isaacs A, van Duijn CM, Witteman JC. 2006. ACE polymorphisms. Circ Res 98(9):1123–33.
Scherer SW, Lee C, Birney E, Altshuler DM, Eichler EE, Carter NP, Hurles ME, Feuk L. 2007. Challenges and standards in integrating surveys of structural variation. Nat Genet 39(7 Suppl):S7–15.
Schillinger W, Fiolet JW, Schlotthauer K, Hasenfuss G. 2003. Relevance of Na+–Ca2+ exchange in heart failure. Cardiovasc Res 57(4):921–33.
Schoneveld OJ, Gaemers IC, Lamers WH. 2004. Mechanisms of glucocorticoid signalling. Biochim Biophys Acta 1680(2):114–28.
Schork NJ, Murray SS, Frazer KA, Topol EJ. 2009. Common vs. rare allele hypotheses for complex diseases. Curr Opin Genet Dev 19(3):212–9.
Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J and others. 2007. Strong association of de novo copy number mutations with autism. Science 316(5823):445–9.
Sheppard HW, Celum C, Michael NL, O'Brien S, Dean M, Carrington M, Dondero D, Buchbinder SP. 2002. HIV-1 infection in individuals with the CCR5-Delta32/ Delta32 genotype: acquisition of syncytium-inducing virus at seroconversion. J Acquir Immune Defic Syndr 29(3):307–13.
Staessen JA, Bianchi G. 2005. Adducin and hypertension. Pharmacogenomics 6(7):665–9.
Staessen JA, Wang J, Bianchi G, Birkenhager WH. 2003. Essential hypertension. Lancet 361(9369):1629–41.
Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin AI, De Faire U. 2002. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. J Intern Med 252(3):247–54.
Ziegler MG, Bao X, Kennedy BP, Joyner A, Enns R. 2002. Location, development, control, and function of extraadrenal phenylethanolamine N-methyltransferase. Ann N Y Acad Sci 971:76–82.
Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123(3):585–95.
The International HapMap consortium. 2005. A haplotype map of the human genome. Nature 437(7063):1299–320.
The PROCARDIS Consortium TP. 2004. A trio family study showing association of the lymphotoxin-alpha N26 (804A) allele with coronary artery disease. Eur J Hum Genet 12(9):770–4.
Tian D, Wang Q, Zhang P, Araki H, Yang S, Kreitman M, Nagylaki T, Hudson R, Bergelson J, Chen JQ. 2008. Single-nucleotide mutation rate increases close to insertions/deletions in eukaryotes. Nature 455(7209):105–8.
Tobin MD, Raleigh SM, Newhouse S, Braund P, Bodycote C, Ogleby J, Cross D, Gracey J, Hayes S, Smith T and others. 2005a. Association of WNK1 gene polymorphisms and haplotypes with ambulatory blood pressure in the general population. Circulation 112(22):3423–9.
Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. 2005b. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. Stat Med 24(19):2911–35.
Tobin MD, Tomaszewski M, Braund PS, Hajat C, Raleigh SM, Palmer TM, Caulfield M, Burton PR, Samani NJ. 2008. Common variants in genes underlying monogenic hypertension and hypotension and blood pressure in the general population. Hypertension 51(6):1658–64.

Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, Nord AS, Kusenda M, Malhotra D, Bhandari A and others. 2008. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 320(5875):539–43.

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J and others. 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280(5366):1077–82.

Wang J, Wang W, Li R, Li Y, Tian G, Goodman L, Fan W, Zhang J, Li J, Zhang J and others. 2008. The diploid genome sequence of an Asian individual. Nature 456(7218):60–5.

Wang Q. 2005. Molecular genetics of coronary artery disease. Curr Opin Cardiol 20(3):182–8.

Wang Y, O'Connell JR, McArdle PF, Wade JB, Dorff SE, Shah SJ, Shi X, Pan L, Rampersaud E, Shen H and others. 2009. Whole-genome association study identifies STK39 as a hypertension susceptibility gene. Proc Natl Acad Sci U S A 106(1):226–31.

Wassink AM, van der Graaf Y, Olijhoek JK, Visseren FL. 2008. Metabolic syndrome and the risk of new vascular events and all-cause mortality in patients with coronary artery disease, cerebrovascular disease, peripheral arterial disease or abdominal aortic aneurysm. Eur Heart J 29(2):213–23.

Watanabe H, Fujiyama A, Hattori M, Taylor TD, Toyoda A, Kuroki Y, Noguchi H, BenKahla A, Lehrach H, Sudbrak R and others. 2004. DNA sequence and d. Nature 429(6990):382–8.

Watkins H, Farrall M. 2006. Genetic susceptibility to coronary artery disease: from promise to progress. Nat Rev Genet 7(3):163–73.

Weiss KM, Clark AG. 2002. Linkage disequilibrium and the mapping of complex human traits. Trends Genet 18(1):19–24.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA and others. 2001. The sequence of the human genome. Science 291(5507):1304–51.

Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT and others. 2008. The complete genome of an individual by massively parallel DNA sequencing. Nature 452(7189):872–6.

Wicks SR, Yeh RT, Gish WR, Waterston RH, Plasterk RH. 2001. Rapid gene mapping in Caenorhabditis elegans using a high density polymorphism map. Nat Genet 28(2):160–4.

Wienke A, Holm NV, Skytthe A, Yashin AI. 2001. The heritability of mortality due to heart diseases: a correlated frailty model applied to Danish twins. Twin Res 4(4):266–74.

Wong DL, Siddall BJ, Ebert SN, Bell RA, Her S. 1998. Phenylethanolamine N-methyltransferase gene expression: synergistic activation by Egr-1, AP-2 and the glucocorticoid receptor. Brain Res Mol Brain Res 61(1–2):154–61.

Woolfe A, Goodson M, Goode DK, Snell P, McEwen GK, Vavouri T, Smith SF, North P, Callaway H, Kelly K and others. 2005. Highly conserved non-coding sequences are associated with vertebrate development. PLoS Biol 3(1):e7.

Wu J, Province MA, Coon H, Hunt SC, Eckfeldt JH, Arnett DK, Heiss G, Lewis CE, Ellison RC, Rao DC and others. 2007. An investigation of the effects of lipid-

lowering medications: genome-wide linkage analysis of lipids in the HyperGEN study. BMC Genet 8:60.
Yan JJ, Tang NP, Tang JJ, Jia EZ, Wang MW, Wang QM, Zhu J, Yang ZJ, Wang LS, Huang J. 2010. Genetic variant in visfatin gene promoter is associated with decreased risk of coronary artery disease in a Chinese population. Clin Chim Acta 411(1–2):26–30.

SUMMARY IN ESTONIAN

Südameveresoonkonna funktsioonis osalevad geenid: uued pärilikud DNA variandid Eesti ja Tšehhi populatsioonides

Südameveresoonkonna haigusi (SVH) peetakse tänapäeva ühiskonna üheks peamiseks enneaegse surma põhjustajaks, mis on sagedasemad (60–70%) üle 60 aastaste inimeste vanusegrupis. Taolise kõrge riski taga peituvad nii geneetilised faktorid kui ka elustiilist ja keskkonnast tulenevad mõjutused. Olulisteks keskkonnast tingitud riskifaktoriteks peetakse ülekaalulisust, vähest füüsilist aktiivsust, liigset soola- ja alkoholitarbimist ning suitsetamist. Perekonnapõhised ja kaksikute uuringud on näidanud, et SVH pärilik komponent võib varieeruda 17–66%ni. Geneetilised uuringud on tuvastanud, et erinevate SVHte vormide kujunemiseni võivad viia mutatsioonid ühes geenis (üksiku geeni ehk monogeenne vorm) kui mitmetes geenides (multigeenne vorm) korraga. Selleks, et kaardistada inimeste haigustega seotud geene on läbi aegade kasutatud inimese genoomis esinevaid geneetilisi markereid, millest kõige sagedasemad on ühenukleotiidsed polümorfismid (ÜNPd). ÜNP markeritel põhinevate ülegenoomsete aheldusanalüüside ja kandidaatgeenide ning kogu genoomi assotsiatsiooniuuringute tulemusel on kaardistatud mitmeid lookusi, mis on olulised SVHte kujunemisel. Paraku pole aga senistel leidudel põhinevad funktsionaalsed uuringud näidanud ühegi geeni või mitmeid geene hõlmava metaboolse raja selget defineeritud rolli SVHte regulatsioonis. Üheltpoolt võib see tuleneda ebapiisavast markerite või uuritavate genoomsete piirkondade valikust ning teisalt uuritavate indiviidide (või populatsioonide) ja haiguse fenotüübi erinevast määratlemisest. Lisaks ÜNPdele, esineb genoomis ka veel teisi geneetilisi markereid (mitme nukleotiidi muutusi hõlmavad struktuursed variandid nagu insertsioonid, deletsioonid ja duplikatsioonid), mille rolli inimese haiguste kujunemisel on oluliselt vähem uuritud.

Töö kirjanduse osa hõlmab ülevaadet inimese genoomis esinevatest markeritest, geneetiliste haiguste kaardistamise meetoditest ja genoomis esinevate regulatoorsete piirkondade kasutamisest multigeensete haiguste uuringutes. Lisaks antakse ülevaade kahest uuritavast SVHst (essentsiaalne hüpertensioon ja südame koronaartõbi) ning nende käesolevaks hetkeks teadaolevast geneetilisest taustast.

Töö eksperimentaalne osa haarab järgmisi etappe: (i) südame veresoonkonna haigusi põhjustavate kandidaatgeenide valik *in silico* analüüsil; (ii) geneetiliste varieeruvuste tuvastamine SVHte kandidaatgeenides: inimese *PNMT* geenis, evolutsiooniliselt konserveerunud mittekodeerivates DNA regioonides ja promootor piirkondades;

(iii) kahes kandidaatgeenis, *PNMT* ja *NCX1,* tuvastatud geneetiliste variantide assotsiatsiooniuuring SVH fenotüüpidega Eesti ja Tšehhi populatsiooni valimites.

Uurimistöö peamised tulemused võib kokku võtta järgmiselt:

1. Põhinedes kirjanduses ja andmebaasides olevale infole tuvastasin 162 geeni, mida on varasemalt seostatud erinevate SVHte (madal ja kõrge vererõhk, EH, CAD jt.) monogeensete ja multigeensete haigusvormide kujunemisega. Enamik neist geenidest on oma funktsiooni poolest seotud kolme organsüsteemiga: neerud, süda ja sümpaatiline närvisüsteem. Lisaks jaotusid uuritavad geenid oma teadaoleva funktsiooni alusel seitsme metaboolse raja vahel, mille rolli on SVHte kujunemisel oluliseks peetud.

2. SVH kandidaatgeeni *PNMT* re-sekveneerimisel Eesti ja Tšehhi indiviididel ilmnes, et DNA järjestuse varieeruvus terves geenis on suhteliselt madal. Võimalike transkriptsioonifaktorite seondumiskohtade *in silico* ennustamisel ja lisaks inimese ning näriliste (hiir ja rott) geeni piirkondade võrdlemisel ilmnes, et inimese *PNMT* geeni esimesse intronisse paigutub transkriptsiooni reguleeriv üksus. Identifitseeritud geeniekspressiooni reguleeriv piirkond *GRU* on *PNMT* geeni lokaliseerunud inimesespetsiifilise *Alu*-insertsiooni tulemusena. Nii tuvastatud madal geneetiline varieeruvus kui identifitseeritud regulatoorne *GRU* annavad põhjust oletada, et SVH patsientide ja tervete kontrollide vaheline PNMT geeni ekspressioonitasemete erinevus ei ole põhjustatud mitte ainult geenis leiduvate polümorfismide poolt, vaid ilmselt hoopis keerulisemas koostöös individuaalsete ekspressiooni reguleerivate üksustega.

3. SVHte kandidaatgeenide geneetiliste varieeruvuste identifitseerimisel ilmnes, et evolutsiooniliselt konserveerunud mittekodeerivates DNA regioonides ning promotorpiirkondades esines madal DNA järjestuste varieeruvus. Analüüsi käigus tuvastasin ühe konserveerunud indiviiditi väga varieeruva piirkonna inimese *NCX1* geeni teises intronis. Lisaks esmalt tuvastatud 14bp indelile esines seal mitmeid teisi varasemalt kirjeldamata indelite variante. Šimpansi, reesusmakaagi ja inimese vastavate genoomipiirkondade võrdlemisel selgus aga, et tegemist on inimese spetsiifilise deletsiooniga. Šimpansil ja reesusmakaagil esinev variant on inimesel esindatud minoorse alleelina.

4. Assotsiatsioonianalüüs kahes Ida-Euroopa populatsioonis (eestlased, tšehhid) näitas tšehhidel seost *NCX1* geeni teise introni hüpervarieeruvasse alasse paigutuva 14bp indeli kandjastaatuse ja suurenenud CAD haigestumise riski vahel. Risk oli veelgi kõrgem metaboolse sündroomiga CADi patsientidel. Lisaks esines seos ka 14bp indel variandi ja triglütseriidide tasemete vahel tšehhidel ning 14bp indel variandi ja HR ja LDL tasemete vahel mõlema valimi tervetel kontrollidel.

Nii käesolev töö kui ka varasemad tööd on näidanud, et siiani vähe uuritud geneetilised variandid, nagu insertsioonid ja deletsioonid võivad omada palju olulisemat rolli inimeste komplekshaiguste (nagu SVHd) kujunemisel kui seni arvatud. Taoliste indel polümorfismide esinemine inimese genoomi kindlates regulatoorsetes piirkondades, nagu seda on promootorid, enhanserid ja intronid, võivad viia geenieksрessiooni muutuseni ja seeläbi haiguse fenotüübi kujunemiseni. Käesolevas töös tuvastatud *NCX1* geeni indel variandi assotsiatiooniuuring erinevate SVH fenotüüpidega näitas tema potentsiaalset peiotroopset rolli, kus ühes geenis asuv polümorfism võib mõju avaldada mitmele fenotüübilisele tunnusele korraga.

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor professor Maris Laan who has supported me with the ideas in the field of human complex diseases and who has guided me through my fruitful masters and doctoral studies.

The work presented in this thesis has been done in a lab among the young and most joyful people you could ever imagine. I thank you all for the wonderful atmosphere, assistance, and great scientific discussion. Elin, Piret, Liina, Liis, Siim, Jaana, Tiina, Gudrun, Kristiina, Neeme, Tarmo and Rait – thank you for the quality time in the lab. Special thanks are going to Pille and Kärt – persons whom I shared my utopic science views as well as sweets and Coca-Cola with. Peeter Juhanson, Margus Putku and Laura Kasak – it was pleasure to do science with you! You were the best students one could have to supervise in the lab.

I would like to thank the team in the Institute of Molecular and Cell Biology comprising of Tiiu, Lagle and Teele who helped me initiate the Graduate School of Biomedicine and Biotechnology projects in the University of Tartu.
I also wish to thank all my previous colleagues from Asper Biotech days.

The highest gratitude goes to my family, especially to my most loving and caring father and mother with all the invaluable support during my studies in the University of Tartu. I would deeply like to thank my best friend Liisi for the support and for always being there when I needed her. She has been like a sister I never had. Darryl, thank for the valuable linguistic support during the writing of this dissertation.

Finally, my deepest gratitude goes to Toomas and our most wonderful son Karl-Richard who have been supportive and loving during my studies and work.

This study was financially supported by the personal stipends from Kristjan Jaak Stipend by Archimedes Foundation and Estonian Ministry of Education. World Federation of Scientists, Graduate School in Biomedicine and Biotechnology, Estonian Students Fund in the USA, Estonian World Council Inc. stipend and EcoGene fellowship for young women in science (women returning from maternity leave).

APPENDIX

1. The human genetic variation data of *PNMT* gene from Ref. II is deposited in dbSNP database under the accession numbers:

PNMT – SNP-702	262803330
PNMT – SNP-390	262803318
PNMT – SNP-591	262803319
PNMT – SNP-184	262803320
PNMT – SNP+360	262803321
PNMT – SNP+1520	262803322
PNMT – SNP+1587	262803323

ADDITIONAL TABLES

Additional Table 1. - List of and description of 162 cardiovascular trait related genes.

No	Gene Symbol*	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
1	*ABCC8*	ATP-binding cassette, sub-family C, member 8	6833	11p15.1	Putative subunit of the beta-cell ATP-sensitive potassium channel (KATP). Regulator of ATP-sensitive K(+) channels and insulin release.	Hypoglycemia; hyperinsulinemic hypoglycemia type 1 (HHF1), neonatal diabetes mellitus (PNDM); neonatal diabetes mellitus type 2 (TNDM2); non-insulin-dependent diabetes mellitus (NIDDM).
2	***ACE***	angiotensin I converting enzyme (peptidyl-dipeptidase A) 1	1636	17q23	Converts angiotensin I to angiotensin II by release of the terminal HIS-LEU, this results in an increase of the vasoconstrictor activity of angiotensin. Also able to inactivate bradykinin, a potent vasodilatator.	Susceptibility to myocardial infarction.
3	***ADD1***	adducin 1 (aplha)	118	4p16.3	Membrane-cytoskeleton-associated protein that promotes the assembly of the spectrin-actin network. Binds to calmodulin.	Essential salt-sensitive hypertension; hyperlipidemia familial combined, huntington disease, stroke, renal disease, insulin resistance, coronary heart disease
4	***ADD2***	adducin 2 (beta)	119	2p14-p13	Membrane-cytoskeleton-associated protein that promotes the assembly of the spectrin-actin network. Binds to calmodulin. Calmodulin binds preferentailly to the beta subunit.	Essential salt-sensitive hypertension; hereditary elliptocytosis and spherocytosis, hemolysis, anemia.
5	*ADD3*	adducin 3 (gamma)	120	10q24.2-q24.3	Membrane-cytoskeleton-associated protein that promotes the assembly of the spectrin-actin network.	Essential salt-sensitive hypertension.
6	*ADM/AM*	adrenomedullin	133	11p15.4	AM and PAMP are potent hypotensive and vasodilatator agents. Numerous actions have been reported most related to the physiologic control of fluid and electrolyte homeostasis. In the kidney, am is diuretic and natriuretic, and both am and pamp inhibit aldosterone secretion by direct adrenal actions.	Essential hypertension; heart failure; chronic hypotension; hypertension pulmonary.
7	*ADORA1*	adenosine A1 receptor	134	1q32.1	Receptor for adenosine. The activity of this receptor is mediated by G proteins which inhibit adenylyl cyclase.	Ischemia; myocardial ischemia; heart failure congestive; cardiovascular diseases; hypotension.
8	*ADORA2A*	adrenergic, beta-2-, receptor, surface	135	22q11.2	Receptor for adenosine. The activity of this receptor is mediated by G proteins which activate adenylyl cyclase.	Essential hypertension; stroke; parkinson disease; huntington disease; coronary artery disease; coronary artery disease.
9	*ADRA1A*	adrenergic, alpha-1A-, receptor	148	8p21-p11.2	This alpha-adrenergic receptor mediates its action by association with G proteins that activate a phosphatidylinositol- calcium second messenger system. Its effect is mediated by G(q) and G(11) proteins.	Prostatic hyperplasia; attention-deficit hyperactivity disorder.
10	*ADRA2A*	alpha2A adrenergic receptor	150	10q24-q26	Alpha-2 adrenergic receptors mediate the catecholamine- induced inhibition of adenylate cyclase through the action of G proteins.	Tourette syndrome, obesity, disability, schizophrenia, attention-deficit hyperactivity disorder.
11	*ADRB2*	adrenergic, beta-2-receptor, surface	154	5q31-q32	Beta-adrenergic receptors mediate the catecholamine- induced activation of adenylate cyclase through the action of G proteins. The beta-2-adrenergic receptor binds epinephrine with an approximately 30-fold greater affinity than it does norepinephrine	Asthma; essential hypertension; heart failure, obesity, resting heart rate
12	*ADRB3*	Adrenergic, beta-3-, receptor	155	8p12-p11.2	Beta-adrenergic receptors mediate the catecholamine- induced activation of adenylate cyclase through the action of G proteins. Beta-3 is involved in the regulation of lipolysis and thermogenesis	Excess capacity to gain weight in patients with insulin resistance, early onset non insulin dependent diabetes and/or morbid obesity, with fatty liver
13	***AGT***	angitensinogen (serine (or cystenine) proteinase inhibitor, clade A (alpha- 1 antiproteinase antitrypsin), member 8)	183	1q42-q43	In response to lowered blood pressure, the enxyme renin cleaves angiotensin I, from angiotensinogen. ACE (angiotensin converting enzyme) then removes a dipeptide to yield the physiologically active peptide peptide angiotensin II, the most potent pressor substance known, which helps regulate volume and mineral balance of body fluids.	Hypertension, essential, susceptibility to; pregnancy-induced hypertension (PIH) (Preeclampsia)
14	***AGTR1***	Angiotensin-II type 1 receptor	185	3q21-q25	Receptor for angiotensin II. Mediates its action by association with G proteins that activate a phosphatidylinositol - calcium second messenger system.	Hypertension, essential; renal disease, heart failure; cardiovascular diseases; nephropathy obstructive
15	***AGTR2***	Angiotensin receptor 2	186	Xq22-23	Receptor for angiotensin II. May have role in cell morphogenesis and related events in growth and development.	Hypertension, essential; renal disease, heart failure; cardiovascular diseases; nephropathy obstructive
16	*AKR1B1*	aldo-keto reductase family 1, member B1 (aldose reductase)	231	7q35	Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	Atherosclerosis; myocardial infarction; stroke; coronary artery disease; renal disease.

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
17	*APOA1*	Apolipoprotein A-1	335	11q23-q24	APOA-1 participates on the reverse transport of choloesterol from tissues to the liver for excretion by promoting cholesterol efflux from tissues and by acting as a cofactor for the lecithin cholesterol acyltransferase (LCAT).	Atherosclerosis; coronary heart disease; coronary artery disease; hypercholesterolemia
18	*APOA2*	Apolipoprotein A-II	336	1q21-q23	May stabilize HDL (high density lipoprotein) structure by its association with lipids, and affect the HDL metabolism.	Apolipoprotein A-II deficiency; atherosclerosis; coronary heart disease; coronary artery disease; hypercholesterolemia
19	*APOC2*	Apolipoprotein C-II	344	19q13.2	APOC-II is a component of the very low density lipoprotein (VLDL) fraction in plasma, and is and activator of several triacylgycerol lipases. The association of APOC-II with plasma chylomicrons, VLDL, and HDL is reversible, a function of the secretion and catabolism of triglyceride-rich lipoproteins, and changes rapidly.	Hyperlipoproteinemia, type Ib
20	*APOC3*	Apolipoprotein C-III	345	11q23.1-q23.2	APOC-III inhibits lipoprotein lipase and hepatic lipase and decreases the uptake of lymph chylomicrons by hepatic cells. This suggests that APOC-III delays catabolism of triglyceride-rich particles.	Hyperlipoproteinemia
21	*APOC4*	Apolipoprotein C-IV	346	19q13.2	May participate in lipoprotein metabolism.	Coronary artery disease
22	*APOE*	Apolipoprotein E	348	19q13.2	Mediates the binding, internalization, and catabolism of lipoprotein particles. It can serve as a ligand for the LDL (apo B/E) receptor and for the specific apo-E receptor (chylomicron remnant) of hepatic tissues.	Early-onset coronary heart disease; multiple sclerosis; myocardial infarction; hyperlipoproteinemia type III; atherosclerosis.
23	*AQP2*	aquaporin 2	359	12q12-q13	Forms a water-specific channel that provides the plasma membranes of renal collecting duct with high permeability to water, thereby permitting water to move in the direction of an osmotic gradient	Diabetes insipidus nephrogenic; hyponatremia; hypercalciuria; central diabetes insipidus
24	***ATP1A1***	ATPase, Na+K+ transporting, alpha-1 polypeptide	476	1p13-11	This is the catalytic component of the active enzyme , which catalyzes the hydrolysis of ATP coupled with the exange of Na and K ions across theplasma membrane. This action creates the electrochemical gradient of Na and K, providing the energy for active transport of various nutrients.	Hypertension
25	*AVP*	arginine vasopressin (neurophysin II, antidiuretic hormone, diabetes insipidus, neurohypophyseal)	551	20p13	Vasopressin has a direct antidiuretic action on the kidney, it also causes vasoconstriction of the peripheral vessels.	Neurohypophyseal diabetes insipidus; heart failure;
26	*AVPR1A*	arginine vasopressin receptor 1A	552	12q14-q15	Receptor for arginine vasopressin. The activity of this receptor is mediated by G proteins which activate a phosphatidyl- inositol-calcium second messenger system. Has been involved in social behaviors, including affiliation and attachment.	Ischemia; dysmenorrhea
27	*AVPR1B*	arginine vasopressin receptor 1B	553	1q32	Receptor for arginine vasopressin. The activity of this receptor is mediated by G proteins which activate a phosphatidyl- inositol-calcium second messenger system. Has been involved in social behaviors, including affiliation and attachment.	Major depression; tumors
28	*AVPR2*	Arginine vasopressin receptor 2 (nephrogenic diabetes insipidus)	554	Xq28	Receptor for arginine vasopressin. The activity of this receptor is mediated by G proteins which activate adenylate cyclase.	Diabetes insipidus, nephrogenic
29	*BDKRB2*	Bradykinin receptor B2	624	14q32.1-q32.2	Receptor for bradykinin. It is associated with G proteins that activate a phosphatidylinositol-calcium second messenger system.	Inflammation; renal disease, brain tumors, cardiovascular effects
30	*BRS3*	Bombesin-like receptor 3	680	Xq26-q28	Role in sperm cell division, maturation, or function. This receptor mediated its action by association with G proteins that activate phosphatidylinositol-calcium second messenger system.	Obesity; tumors; nsclc

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
31	***BSND***	Barttin	7809	11p32.1	Functions as a beta-subunit for CLCNKA and CLCNKB chloride channels. In the kidney CLCNK/BSND heteromers mediate chloride reabsorption by facilitating its basolateral efflux. In the stria, CLCNK/BSND channels drive potassium secretion by recycling chloride for the basolateral SLC12A2 cotransporter	Bartter syndrome; gitelman syndrome; hypokalemia, hypotension: renal failure
32	*CACNA1C*	calcium channel, voltage-dependent, L type,	775	12p13.3	Voltage-sensitive calcium channels (VSCC) mediate the entry of calcium ions into excitable cells and are also involved in a variety of calcium-dependent processes, including muscle contraction, hormone or neurotransmitter release, gene expression, cell motility, cell division and cell death.	Hypokalemic periodic paralysis
33	*CALCA*	Calcitonin/calcitonin-related polypeptide, alpha	796	11p15.2-p15.1	Calcitonin causes a rapid but short-lived drop in the level of calcium and phosphate in blood by promoting the incorporation of those ions in the bones.	Ostepoprosis
34	***CLCNKA***	Chloride channel, kidney, A	1187	1p36	Voltage-gated chloride channel. Chloride channels have several functions including the regulation of cell volume; membrane potential stabilization, signal transduction and transepithelial transport. May be important in urinary concentrating mechanisms.	Bartter syndrome, type 4, digenic
35	***CLCNKB***	Chloride channel, kidney, B	1188	1p36	Voltage-gated chloride channel. Chloride channels have several functions including the regulation of cell volume; membrane potential stabilization, signal transduction and transepithelial transport. May be important in urinary concentrating mechanisms.	Bartter syndrome, antenatal
36	*CMA1*	chymas 1, mast cell	1215	14q11.2	Major secreted protease of mast cells with suspected roles in vasoactive peptide generation, extracellular matrix degradation, and regulation of gland secretion.	Inflammation; vascular disease; heart failure; dermatitis atopic
37	*CYP11B1*	Cytochrome P450, subfamily XIB, polypeptide 1 (11-beta-hydroxylase; corticosteroid methyl-oxidase II (CMO II))	1584	8q21	Has steroid 11-beta-hydroxylase activity. In addition to this activity, the 18 or 19-hydroxylation of steroids and the aromatization of androstendione to estrone have also been ascribed to cytochrome P450 XIB.	Adrenal hyperplasia, congenital, due to 11-beta-hydroxylasedeficiency (3) Aldosteronism, glucocorticoid-remediable
38	***CYP11B2***	Cytochrome P450, subfamily XIB, polypeptide 2	1585	8q21-q22	Preferentially catalyzes the conversion of 11-deoxycorticosterone to aldosterone via corticosterone and 18-hydroxycorticosterone.	Hypoaldosteronism, congenital, due to CMO II deficiency (3);Hypoaldosteronism, congenital, due to CMO I deficiency, 203400 (3); {Low renin hypertension, susceptibility to} (3); Aldosterone to renin ratioraised (3)
39	***CYP17A1***	Cytochrome P450, family 17, subfamily A, polypeptide 1 (steroid 17-alpha-hydroxylase)	1586	10q24.3	Conversion of pregnenolone and progesterone to their 17-alpha-hydoxylated products and seubsequently to dehydroepiandrosterone (DHEA) and androstenedione. Catalyzes both the 17-alpha-hydroxylation and the 17, 20-lyase reaction. Involved in sexual development during fetal life and at buberty.	Adrenal hyperplasia, congenital, due to 17-alpha-hydroxylasedeficiency (3)
40	**CYP21A2**	cytochrome P450, family 21, subfamily A, polypeptide 2	1589	6p21.3	Specifically catalyzes the 21-hydroxylation of steroids. Required for the adrenal synthesis of mineralocorticoids and glucocorticoids	cardiovascular diseases, heart failure, metabolic disorder
41	*CYP4A11*	cytochrome P450, family 4, subfamily A, polypeptide 11	1579	1p33	Catalyzes the omega- and (omega-1)-hydroxylation of various fatty acids such as laurate, myristate and palmitate. Has little activity toward prostaglandins A1 and E1.	Essential hypertension, myocardial infarction
42	*DBH*	dopamine beta-hydroxylase (dopamine beta-monooxygenase)	1621	9q34	Conversion of dopamine to noradrenaline	Dopamine-beta-hydroxylase activity levels, plasma.
43	*DRD1*	dopamine receptor D1	1812	5q35.1	This is one of the five types (D1 to D5) of receptors for dopamine. The activity of this receptor is mediated by G proteins which activate adenylyl cyclase	schizophrenia, genetic hypertension, essential hypertension

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
44	*DRD2*	dopamine receptor D2	1813	11q23	One of the five types (D1 to D5) of receptors for dopamine. The activity of this receptor is mediated by G proteins which inhibit adenylyl cyclase	schizophrenia, obesity,
45	*ECE1*	endothelin converting enzyme 1	1889	1p36.1	Converts big endothelin-1 to endothelin-1.	Hirschsprung disease, cardiac defects, and autonomic dysfunction; heart failure; cardiovascular diseases
46	*EDN1*	endothelin 1	1906	6p24.1	Endothelins are endothelium-derived vasoconstrictor peptides	Hypertension pulmonary; heart failure; hypertension arterial; cardiovascular diseases; preeclampsia
47	*EDN2*	endothelin 2	1907	1p34	Endothelins are endothelium-derived vasoconstrictor peptides	essential hypertension, cardiovascular diseases, renal failure
48	*EDN3*	endothelin 3	1908	20q13.2-q13.3	Endothelins are endothelium-derived vasoconstrictor peptides	Shah-Waardenburg syndrome; Central hypoventilation syndrome, congenital; Hirschsprung disease
49	*EDNRA*	endothelin-A receptor	1909	4q31.22	Receptor for endothelin-1. Mediates ist action by association with G proteins that activate a phosphatidylinositol - calcium second messenger system.	Heart failure; hypertension arterial; hypertension pulmonary
50	*EDNRB*	endothelin-B receptor	1910	13q22	Non-specific receptor for endothelin1, 2, and 3. Mediates its action by association with G proteins that activate phosphatidylinositol-calcium second messenger system.	Hirschsprung disease-2; ABCD syndrome
51	*F2R*	coagulation factor II (thrombin) receptor	2149	5q13	High affinity receptor for activated thrombin coupled to G proteins that stimulate phosphoinositide hydrolysis. May play a role in platelets activation and in vascular development.	Arterial thrombosis; atherosclerosis; myocardial infarction; renal disease; coronary artery disease; essential hypertension
52	*GAL*	galanin prepropeptide	51083	11q13.2	Contracts smooth muscle of the gastrointestinal and genitourinary tract, regulates growth hormone release, modulates insulin release, and may be involved in the control of adrenal secretion	obesity, hypoglycemia, inflammation, hyperglycemia
53	*GALR1*	galanin receptor 1	2587	18q23	Receptor for the hormone galanin. The activity of this receptor is mediated by G proteins that inhibits adenylate cyclase activity.	Obesity; melanoma
54	*GCG*	glucagon	2641	2q36-q37	1. glucagon promotes hydrolysis of glycogen and lipids, and raises the blood sugar level. 2. GLP stimulates intestinal growth and upregulates villus height in the small intestine, concominant with increased crypt cell proliferation and decreased enterocyte apoptosis.	Hyperglycemia; obesity; coronary artery disease: heart failure; acute myocardial infarction
55	*GCK*	glucokinase (hexokinase 4, maturity onset diabetes of the young 2)	2645	7p15.3-p15.1	Catalyzes the initial step in utilization of glucose by the beta-cell and liver at physiological glucose concentration. Glucokinase has a high Km for glucose, and so it is effective only when glucose is abundant. The role of GCK is to provide G6P for the synthesis of glycogen. Pancreatic glucokinase plays an important role in modulating insulin secretion. Hepatic glucokinase helps to facilitate the uptake and conversion of glucose by acting as an insulin-sensitive determinant of hepatic glucose usage	hyperglycemia, obesity
56	*GFPT1*	glutamine-fructose-6-phosphate transaminase 1	2673	2p13	Controls the flux of glucose into the hexosamine pathway. Most likely involved in regulating the availability of precursors for N- and O-linked glycosylation of proteins	renal disease, obesity
57	*GH1*	growth hormone 1	2688	17q24.2	Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulated both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis on muscle and other tissues.	Isolated growth hormone deficiency, IIIig type with absent GH.
58	*GH2*	growth hormone 2	2689	17q24.2	Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulated both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein in muscle and other tissues.	Tumors; choriocarcinoma
59	*GIPR*	gastric inhibitory polypeptide receptor	2696	19q13.3	This is a receptor for GIP. The activity of this receptor is mediated by G proteins which activate adenylyl cyclase.	Cushing syndrome; hyperplasia
60	*GLP1R*	glucagon-like peptide 1 receptor	2740	6p21	This is a receptor for glucagon-like peptide 1. The activity of this receptor is mediated by G proteins which activate adenylyl cyclase	obesity, metabolic disorder, cardiovascular diseases

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
61	*GNAI1*	guanine nucleotide binding protein (G protein)	2767	19p13.3	Guanine nucleotide-binding proteins (G proteins) are involved as modulators or transducers in various transmembrane signaling systems. Acts as an activator of phospholipase C.	Leukemia, Lymphocytic, Acute, L1
62	*GNB3*	Guanine nucleotide-binding protein, beta polypeptide-3	2784	12p13	Guanine nucleotide-binding proteins (G proteins) are involved as a modulator or transducer in various transmembrane signaling systems. The beta and gamma chains are required for the GTPase activity, for replacement of GDP by GTP, and for G protein-effector interaction.	Hypertension, essential, susceptibility to
63	*GYS1*	glycogen synthase 1 (muscle)	2997	19q13.3	Transfers the glycosyl residue from UDP-Glc to the nonreducing end of alpha-1,4 glucan	hyperinsulinemia, hypoglycemia
64	*GYS2*	glycogen synthase 2 (liver)	2998	12p12.2	Transfers the glycosyl residue from UDP-Glc to the nonreducing end of alpha-1,4 glucan	liver glycogen synthase deficiency
65	*HP*	haptoglobin	3240	16q22.1	Haptoglobin combines with three palsma hemoglobin, preventing loss of iron through the kidneys and protecting the kidneys from damage by hemoglobin, while making the hemoglobin accessible to degradative enzymes.	Anhaptoglobinemia; hypohaptoglobinemia
66	***HSD11B1***	Hydroxysteroid, 11-beta, dehydrogenase 1	3290	1q32-q41	Catalyzes reversibly the conversion of cortisol to the inactive metabolite cortisone.	Cortisone reductase deficiency, 604931 (3)
67	*HSD11B2*	Hydroxysteroid (11-beta) dehydrogenase 2	3291	16q22	Has a role in modulating glucocorticoid activity both at the level of the mineralcorticoid receptor and the glucocorticoid receptor. Uses NADH while 11-DH1 uses NADPH. Catalyzes non reversibly the conversion of cortisol to the inactive metabolite cortisone.	Apparent mineralocorticoid excess, hypertension due to (3);Hypertension, mild low-renin (3)
68	*IAPP*	islet amyloid polypeptide	3375	12p12.3-p12.1	Selectively inhibits insulin-stimulated glucose utilization and glycogen deposition in muscle, while not affecting adipocyte glucose metabolism.	IAPP is the peptide subunit of amyloid found in pancreatic islets of type 2 diabetic patients and in insulinomas.
69	*ICAM1*	intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	3383	19p13.3-p13.2	ICAM proteins are ligands for the leukocyte adhesion LFA-1 protein (integrin alpha-L/beta-2)	Malaria, cerebral
70	*ICAM2*	intercellular adhesion molecule 2	3384	17q23-q25	ICAM proteins are ligands for the leukocyte adhesion LFA-1 protein (integrin alpha-L/beta-2). ICAM2 may play a role in lymphocyte recirculation by blocking LFA-1 dependent cell adhesin. It mediates adhesive interactions important for antigen- specific immune responce, NK-sell mediated clearance, lymphocyte recirculation, and other cellular interactions important for immune response and surveillance.	Inflammation; multiple sclerosis; senile plaques
71	*ICAM3*	intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	3383	19p13.3-p13.2	ICAM proteins are ligands for the leukocyte adhesion protein LFA-1 (integrin alpha-L/beta-2). In case of rhinovirus infection acts as a cellular receptor for the virus. Also acts as receptor for some coxsackieviruses	atherosclerosis, inflammation, coronary artery disease
72	*IGF1*	insulin-like growth factor 1 (somatomedin C)	3479	12q22-q23	The insulin-like growth factors, isolated from plasma, are structurally and functioanlly related to insulin bu have a much higher growth-promoting activity.	Growth retardation with deafness and mental retardation.
73	*IGF2*	insulin-like growth factor 2 (somatomedin A)	3481	11p15.5	The insulin-like growth factors possess growth-promoting activity. In vitro, they are potent mitogens for cultured cells. IGF-II is influenced by placental lactogen and may play a role in fetal development	hypoglycemia, renal failure chronic

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
74	***IL1A***	interleukin 1, alpha	3552	2q14	IL-1 stimulates thymocyte proliferation by inducing IL-2 release, B-cell maturation and proliferation, and fibroblast growth factor activity. IL-1 proteins are involved in the inflammatory response, being identified as endogenous pyrogens, and are reported to stimulate the release of prostaglandin and collagenase from synovial cells.	inflammation,
75	*INS*	insulin	3630	11p15.5	Insulin decreases blood glucose concentration. It increases cell permeability to monosaccharides, amino acids and fatty acids. It accelerates glycolysis, the pentose phosphate cycle, and glycogen synthesis in liver	diabetes mellitus, obesity, atherosclerosis, coronary heart disease, essential hypertension
76	*INSR*	insulin receptor	3643	19p13.3-p13.2	This receptor binds insulin and has a tyrosine-protein kinase activity. Isoform Short has a higher affinity for insulin. Mediates the metabolic functions of insulin. Binding to insulin stimulates association of the receptor with downstream mediators including IRS1 and phosphatidylinositol 3'-kinase (PI3K). Can activate PI3K either directly by binding to the p85 regulatory subunit, or indirectly via IRS1	Diabetes mellitus, Hyperinsulinemic hypoglycemia, essential hypertension, aldosteronism primary, renal disease, coronary artery disease, heart failure, hypertension arterial
77	*IPF1/PDX1*	pancreatic and duodenal homeobox 1	3651	13q12.1	Activates insulin, somatostatin, glucokinase, islet amyloid polypeptide and glucose transporter type 2 gene transcription.	hyperglycemia, myocardial infarction, hypercholesterolemia
78	***KCNJ1***	potassium inwardly-rectifying channel, subfamily J, member 1	3758	11q24	In the kidney, probably palys a major role in K+ homeostasis. Iward rectifier K+ channels are characterized by a greater tendancy to allow potassium to flow into the sell rather than out of it. Their voltage dependance is regulated by yhe concentration of extracellular potassium; as external K+ is rased, the voltage range of the channel opening shifts to more positive voltages. the inward rectification is mainly due to the blockage of outward current by internal magnesium. this channel is activated by internal ATP and can be blocked by external BA2+.	Bartter syndrome type 2, characterized by an hypokalemic,hypochloremic metabolic alkalosis with hyperkaliury,hyperexcretion of prostaglandin E,hyperreninemia hyperaldosteronism with normal blood pressure,insensitivity to AGT2,and hyperplasia of juxtaglomerular apparatus,autosomal recessive. Presenting as an antenatal form with hydramnios prematurity and dehydration at birth and a classic form with failure to thrive.
79	*KCNJ11*	potassium inwardly-rectifying channel, subfamily J, member 11	3767	11p15.1	This receptor is controlled by G proteins. Inward rectifier K+ channels are characterized by a greater tendancy to allow potassium to flow into the cell rather than out of it. Their voltage dependance is regulated by the concentration of extracellular potassium; as external K+ is raised, the voltage range of the channel opening shifts to more positive voltages. The inward rectifications is mainly due to the blockage of outward current by internal magnesium. Can be blocked by extracellular Ba2+ (by similarity).	Persistent hyperinsulinemic hypoglycemia of infancy.
80	*KCNJ6*	potassium inwardly-rectifying channel, subfamily J, member 6	3763	21q22.13-q22.2	This potassium channel may be involved in the regulation of insulin secretion by glucose and/or neurotransmitters acting through G-protein-coupled receptors.	Insulinoma
81	*KCNK2*	potassium channel, subfamily K, member 2	3776	1q41	Outward rectifying potassium channel	acidosis
82	***KLK1***	kallikrein 1, renal/pancreas/salivary	3816	19q13.2-q13.4	Glandural kallikreins cleave Met-Lys and Arg-Ser bonds in kininogen to release Lys-bradykinin	Renal failure chronic; cancer
83	*KNG1*	kininogen 1	3827	3q27	Induction of hypotension; natriuresis and diuresis; decrease in blood glucose level; has a cardioprotective effect	Inflammation; heart failure; essential hypertension; cardiovascular diseases
84	*LDLR*	low density lipoprotein receptor (familial hypercholesterolemia)	6547	19p13.3	Binds LDL, the major cholesterol-carrying lipoprotein of plasma, and transports it into cells by endocytosis.	hypercholesterolemia familial, atherosclerosis
85	*LEP*	leptin precursor	3952	7q31.3	May function as part of a signaling pathway that acts to regulate the size of the body fat depot. An increase in the level of LEP may act directly or indirectly on the CNS to inhibit food intake and/or regulate energy expenditure as part of a homeostatic mechanism to maintain constancy of the adipose mass	Obesity; cardiovascular diseases; renal failure chronic; renal disease; essential hypertension; coronary heart disease

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
86	*LEPR*	leptin receptor	3953	1p31	Receptor for obesity factor (leptin); Involved in the regulation of fat metabolism and, in a hematopoietic pathway, required for normal lymphopoiesis.	Obesity; cardiovascular diseases; atherosclerosis; atherosclerotic plaque; coronary artery disease; essential hypertension
87	*LPL*	lipoprotein lipase	4023	8p22	The primary function of this lipase is the hydrolysis of triglycerides of circulating chylomicrons and very low density lipoproteins (VLDL). The enzyme functions in the presence of apolipoprotein C-2 on the luminal surface of vascular endothelium	endogenous hypertriglyceridemia, atherosclerosis, coronary heart disease, coronary artery disease, coronary atherosclerosis, essential hypertension
88	*LRP8*	low density lipoprotein receptor-related protein 8, apolipoprotein e receptor.	7804	1p34	Cell surface receptor for Reelin (RELN) and apolipoprotein E (apoE)-containing ligands.	Myocardial infarction
89	*MTHFR*	5,10-methylenetetrahydrofolate reductase (NADPH)	4524	1p36.3	Catalyzes the conversion of 5,10- methylenetetrahydrofolate to 5-methyltetrahydrofolate, a co- substrate for homocysteine remethylation to methionine	thrombophilia, venous thrombosis, cardiovascular diseases, atherosclerosis, coronary heart disease
90	***NEDD4L***	neural precursor cell expressed, developmentally down-regulated 4-like	23327	18q21	E3 ubiquitin-protein ligase which accepts ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a thioester and then directly transfers the ubiquitin to targeted substrates. Inhibits TGF-beta signaling by triggering SMAD2 and TGFR1 ubiquitination and proteasome-dependent degradation. Promotes ubiquitination and internalization of various plasma membrane channels	liddles syndrome, essential hypertension, hypotension orthostatic
91	*NOS1*	nitric oxide synthase 1 (neuronal)	4842	12q24.2-q24.31	Produces nitric oxide (NO) which is a messenger molecule with diverse functions throughout the body. In the brain and peripheral nervous system, NO displays many properties of a neurotransmitter.	hypertrophic pyloric stenosis; atherosclerosis; infarct; renal failure; coronary artery disease; coronary heart disease; myocardial infarction; stroke
92	*NOS2A*	Nitric oxide synthase 2A, inducible, hepatocytes	4843	17cen-q11.2, 17q11.2-q12	Produces nitric oxide (NO) which is a messenger molecule with diverse functions throughout the body. In macrophages, NO mediates tumoricial and bactericidal actions.	Hypertension, susceptibility to; Malaria, resistance to
93	*NOS3*	nitric oxide synthase 3 (endothelial cell)	4846	7q36	Produces nitric oxide (NO) which is implicated in vascular smooth muscle relaxation through as SBMP-mediated signal transduction pathway. NO mediates vascular endothelial growth factor (VEGF)-induced angiogenesis in coronary vessels and promotes blood clotting through the activation of platelets.	Preeclampsia, susceptibility to; coronary spasm
94	*NPPA*	natriuretic peptide precursor A	4878	1p36.21	Atrial natriuretic factor (ANF) is a potent vasoactive substance synthesized in mammalian atria and is thought to play a key role in cardiovascular homeostasis. Has a CGMP-stimulating activity.	Heart failure; heart failure congestive; essential hypertension; cardiac hypertrophy
95	*NPPB*	natriuretic peptide precursor B	4879	1p36.2	Acts as a cardiac hormone with a variety of biological actions including natriuresis, diuresis, vasorelaxation, and inhibition of renin and aldosterone secretion. It is thought to play a key role in cardiovascular homeostasis. Helps restore the body´s salt and water balance. improves heart function.	Heart failure; heart failure congestive; heart diseases; cardiac overload
96	*NPPC*	natriuretic peptide precursor C	4880	2q24-qter	Vasorelaxant activity. Has a CGMP-stimulating activity.	Heart failure; heart failure congestive; cardiovascular diseases
97	***NPR1***	natriuretic peptide receptor A/guanylate cyclase A (atrionatriuretic peptide receptor A)	4881	1q21-22	Receptor for atrial natriuretic peptide. Has guanylate cyclase activity on binding of ANF.	Cardiac hypertrophy; heart failure congestive; essential hypertension; heart failure
98	***NPR2***	natriuretic peptide receptor B/guanylate cyclase B (atrionatriuretic peptide receptor B)	4882	9p21-p12	Receptor for atrial natriuretic peptide. Has guanylate cyclase activity on binding of ANF. Seems to be stimulated more effectively by brain natriuretic peptide (BNP) than by ANP.	Acromesomelic dysplasia Maroteaux type (AMDM); heart failure; heart failure congestive; essential hypertension

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
99	*NPR3*	natriuretic peptide receptor C/guanylate cyclase	4883	5p14-p13	Receptor for natriuretic peptide hormones. Has broad specificity and can bind several distinct natriuretic peptides, including atrial natriuretic peptide (NPPA) and brain natriuretic peptide (NPPB). Does not have guanylate cyclase activity	Heart failure; essential hypertension; heart failure congestive; cardiovascular diseases
100	*NPY*	neuropeptide Y	4852	7p15.1	NPY is implicated in the control of feeding and in secretion of gonadotrophin-release hormone.	Essential hypertension; obesity; pheochromocytoma
101	*NPY1R*	neuropeptide Y receptor Y1	4886	4q31.3-q32	Receptor for neuropeptide Y and peptide YY.	Obesity; adenocarcinoma
102	*NR3C1*	Nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)	2908	5q31	The steroid hormones and their receptors are involved in the regulation of eucaryotic gene expression and affect cellular proliferation and differentiation in target tissues. Binds to the GRE target site.	Cortisol resistance
103	***NR3C2***	Nuclear receptor subfamily 3, group C, member 2 (mineralocorticoid receptor; aldosterone receptor)	4306	4q.31.1	Receptor for both mineralcorticoids (MS) such as aldosterone and glucocorticoid (GS) such as cortisone or cortisol. The effect of MC is to increase ion and water transport and thus raise extracellular fluid volume and blood pressure and lower potassium levels.	Pseudohypoaldosteronism type I, autosomal dominant, 177735 (3);Hypertension, early-onset, autosomal dominant, with exacerbation in pregnancy,
104	*PLA2G1B*	phospholipase A2, group IB (pancreas)	5319	12q23-q24.1	PA2 catalyzes the calcium-dependent hydrolysis of the 2-acyl groups in 3-sn-phosphoglycerides.	Inflammation; acute pancreatitis
105	*PNMT*	Phenylethanolamine N-methyltransferase	5409	17q21-q22	Converts noradrenaline to adrenaline	Hypertension, essential,
106	*PPP1CA*	protein phosphatase 1, catalytic subunit, alpha isoform	5499	11q13	Protein phosphatase 1 (PP1) is essential for cell division, and participates in the regulation of glycogen metabolism, muscle contractility and protein synthesis. Involved in regulation of ionic conductances and long-term synaptic plasticity. May play an important role in dephosphorylating substrates such as the postsynaptic density-associated Ca(2+)/calmodulin dependent protein kinase II	Leukemia, Lymphocytic, Acute, L1
107	*PPP1CC*	protein phosphatase 1, catalytic subunit, gamma isoform	5501	12q24.1-q24.2	Protein phosphatase 1 (PP1) is essential for cell division, and participates in the regulation of glycogen metabolism, muscle contractility and protein synthesis. Involved in regulation of ionic conductances and long-term synaptic plasticity. May play an important role in dephosphorylating substrates such as the postsynaptic density-associated Ca(2+)/calmodulin dependent protein kinase II	not known
108	*PRCP*	prolylcarboxypeptidase (angiotensinase C)	5547	11q14	Cleaves C-terminal amino acids linked to proline in peptides such as angiotensin II, III and des-Arg9-bradykinin	chronic hypertension, coronary heart disease
109	*PRKCE*	protein kinase C, epsilon	5581	2p21	This is calcium-independent, phospholipid-dependent, serine- and threonine-specific enzyme	Inflammation; tumors;
110	*PRKCQ*	protein kinase C, theta	5588	10p15	Calcium-independent, phospholipid-dependent, serine- and threonine-specific enzyme. Essential for T-cell receptor (TCR)-mediated T-cell activation, but is dispensable during TCR-dependent thymocyte development. Links the TCR signaling complex to the activation of NF-kappa-B in mature T lymphocytes. Required for interleukin-2 (IL2) production Function: PKC is activated by diacylglycerol which in turn phosphorylates a range of cellular proteins. PKC also serves as the receptor for phorbol esters, a class of tumor promoters	Gastrointestinal stromal tumor; leukemia: inflammation
111	*PRKWNK1 (WNK1; PHA2C)*	Pseudohypoaldosteronism, type IIC	65125	12p13	Controls sodium and chloride ion transport by inhibiting the activity of WNK4, potentially by either phosphorylating the kinase or via an interaction between WNK4 and the autoinhibitory domain of WNK1. WNK4 regulates the activity of the thiazide- sensitive Na-Cl cotransporter, SLC12A3, by phosphorylation. WNK1 may also play a role in actin cytoskeletal reorganization	Pseudohypoaldosteronism, type IIC; essential hypertension

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
112	*PRKWNK4 (WNK4; PHA2B)*	Pseudohypoaldosteronism type II	65266	17p21-q22	Regulates the activity of the thiazide-sensitive Na-Cl cotransporter, SLC12A3, by phosphorylation which appears to prevent membrane trafficking of SLC12A3. Also inhibits the renal K(+) channel, KCNJ1, via a kinase-independent mechanism by which it induces clearance of the protein from the cell surface by clathrin-dependent endocytosis. WNK4 appears to act as a molecular switch that can vary the balance between NaCl reabsorption and K(+) secretion to maintain integrated homeostasis (By similarity)	Pseudohypoaldosteronism type II
113	*PTGER2*	prostaglandin E receptor 2 (subtype EP2)	5732	14q22	Receptor for prostaglandin E2 (PGE2). The activity of this receptor is mediated by G(s) proteins that stimulate adenylate cyclase. The subsequent raise in intracellular cAMP is responsible for the relaxing effect of this receptor on smooth muscle	atherosclerosis, stroke, inflammation
114	*PTGER3*	prostaglandin E receptor 3 (subtype EP3)	5733	1p31.2	Receptor for prostaglandin E2 (PGE2); the EP3 receptor may be involved in inhibition of gastric acid secretion, modulation of neurotransmitter release in central and peripheral neurons, inhibition of sodium and water reabsorption in kidney tubulus and contraction in uterine smooth muscle. the activity of this receptor can couple to both the inhibition of adenylate cyclase mediated by G-I proteins, and to an elevation of intracellular calcium. The various isoforms have identical ligand binding properties but can interact second messenger systems (by similarity).	Atherosclerotic plaque; adenocarcinoma lung; tumors
115	*PTGIR*	prostaglandin I2 (prostacyclin) receptor (IP)	5739	19q13.3	Receptor for prostacyclin (prostaglandin I2 or PGI2). The activity of this receptor is mediated by G(s) proteins which activate adenylate cyclase	atherosclerosis, cardiovascular diseases, arthritis
116	*PTGIS, CYP8A1*	Prostaglandin 12 synthase	5740	20q13.11-q13.13	Catalyzes the isomerization of prostaglandin H2 to prostacyclin (=prostaglandin 12).	Hypertension, essential,
117	*PTGS1*	prostaglandin-endoperoxide synthase 1	5742	9q32-q33.3	May play an important role in regulating or promoting cell proliferation in some normal and neoplastically transformed cells	atherosclerosis, thrombosis, coronary heart disease, renal failure acute, acute myocardial infarction
118	*PTGS2*	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	5743	1q25.2-q25.3	May have a role as a major mediator of inflammation and/or a role for prostanoid signaling in activity-dependent plasticity.	Likely to play a role in inflamatory diseases such as rheumatoid arthritis.
119	*PTHLH*	parathyroid hormone-like hormone	5744	12p12.1-p11.2	Neuroendocrine peptide which is a critical regulator of cellular and organ growth, development, migration, differentiation and survival and of epithelial calcium ion transport. Regulates endochondral bone development and epithelial-mesenchymal interactions during the formation of the mammary glands and teeth	hypercalcemia, renal failure
120	*PTHR1*	parathyroid hormone receptor 1	5745	3p22-p21.1	This is a receptor for parathyroid hormone and for parathyroid hormone-related peptide. The activity of this receptor is mediated by G proteins which activate adenylyl cyclase and also a phosphatidylinositol-calcium second messenger system	pseudohypoparathyroidism, renal failure chronic, renal failure
121	***REN***	renin	5972	1q32	Renin is a highly specific endopeptidase, whose only known function is to generate angiotensin I from angiotensinogen in the plasma, initiating a cascade of reactions that produce an elevation of blood pressure and increased sodium retention by the kidney.	Hyperproreninemia
122	*RENBP*	renin binding protein	5973	Xq28	Catalyzes the interconversion of N-acetylglucosamine to N-acetylmannosamine. Binds to renin forming a protein complex called high molecular weight (HMW) renin and inhibits renin activity	essential hypertension, renal disease

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
123	***SAH***	SA (rat hypertension-associated) homolog	6296	16p13.11	Medium-chain fatty acid:CoA ligase activity with broad substrate specificity (in vitro). Acts on acids from C(4) to C(11) and on the corresponding 3-hydroxy- and 2,3- or 3,4- unsaturated acids (in vitro)	Hypertension, essential
124	*SCN5A*	voltage-gated sodium channel type V alpha	6331	3p21	This protein mediates the voltage-dependent sodium ion permeability of excitable membranes. Assuming opened or closed conformations in response to the voltage difference across the membrane, the protein forms a sodium-selective channel through which Na(+) ions may pass in accordance with their electrochemical gradient. It is a tetrodotoxin-resistant Na(+) channel isoform. This channel is responsible for the initial upstroke of the action potential in the electrocardiogram	Heart block type I (PFHBI); death sudden cardiac; heart diseases; coronary artery disease; heart failure; ischemia
125	***SCNN1A***	Sodium channel, nonvoltage-gated 1, alpha	6337	12p13	Sodium permeable non-voltage-sensitive ion channel inhibited by the diuretic amiloride. Mediate the electodiffusion of the luminal sodium (and water, which follows osmotically) through the apical membrane of epithelial cells. Controls the reabsorption of sodium in kidney, colon, lung and sweat glends. Also plays a role in taste preception.	Pseudohypoaldosteronism, type I,
126	***SCNN1B***	Sodium channel, nonvoltage-gated 1, beta	6338	16p12.2-p12.1	Sodium permeable non-voltage-sensitive ion channel inhibited by the diuretic amiloride. Mediate the electodiffusion of the luminal sodium (and water, which follows osmotically) through the apical membrane of epithelial cells. Controls the reabsorption of sodium in kidney, colon, lung and sweat glands. Also plays role in taste perception.	Liddle syndrome, 177200 (3); Pseudohypoaldosteronism, type I,
127	*SCNN1D*	sodium channel, nonvoltage-gated 1, delta	6339	1p36.3-p36.2	Sodium permeable non-voltage-sensitive ion channel inhibited by the diuretic amiloride. Mediates the electrodiffusion of the luminal sodium (and water, which follows osmotically) through the apical membrane of epithelial cells. Controls the reabsorption of sodium in kidney, colon, lung and sweat glands.	not known
128	*SCNN1G*	Sodium channel, nonvoltage-gated 1, gamma	6340	16p12	Sodium permeable non-voltage-sensitive ion channel inhibited by the diuretic amiloride. Mediate the electodiffusion of the luminal sodium (and water, which follows osmotically) through the apical membrane of epithelial cells. Controls the reabsorption of sodium in kidney, colon, lung and sweat glands. Also plays a role in taste perception.	Liddle Syndrome, pseudohypoaldosteronism type 1
129	*SELE*	selectin E (endothelial adhesion molecule 1)	6401	1q22-125	Expressed on cytokine induced endothelial cells and mediates their binding to leukocytes. The ligand recognized by ELAM-1 is sialyl-lewis X (alpha (1->3) fucosylated derivatives of polylactosamine that are found at the nonreducung termini of glycolipids.	Atherosclerosis, susceptibility to
130	*SEPRINA4*	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 4	5267	14q31-q32.1	Inhibits human amidolytic and kininogenase activities of human tissue kallikrein. Inhibition is achived by formation of an equimolar, heat- and SDS-stable complex between the inhibitor and the enzyme, and generation of a small C-terminal fragment of the inhibitor due to cleavage at the reactive site tissue kallikrein.	Arthritis; ischemia; inflammation
131	***SGK***	serum/glucocorticoid regulated kinase	6446	6q23	Protein kinase that plays an important role in cellular stress response. Activates certain potassium, sodium, and chloride channels, suggesting an involvement in the regulation of processes such as cell survival, neuronal excitability, and renal sodium excretion. Sustained high levels and activity may contribute to conditions such as hypertension and diabetic nephropathy.	essential hypertension, cardiac hypertrophy, renal failure chronic

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
132	***SLC12A1***	Solute carrier family 12 (sodium/potassium/chloride transporters), member 1	6557	15q15-q21.1	Electrically silent transporter system. Mediates sodium and chloride reabsorption. Plays a vital role in the regulation of ionic balance and cell volume	Bartter syndrome, hypokalemic, hypochloremic metabolic alkalosis with hyperkaliury, hyperexcretion of prostaglandin E, hyperreninemia hyperaldosteronism with normal blood pressure, intensitivity to AGT2, and hyperplasia of juxtaglomerular apparatus, autosomal recessive, presenting as an atenatal form with hydramnios prematurity and dehydration at built and a classic form with failure to thrive.
133	***SLC12A3***	Solute carrier family 12 (sodium/potassium/chloride transporters), member 3	6559	16q13	Electrically silent transporter system. Mediates sodium and chloride reabsorption.	Gitelman syndrome characterized by and hypokalemic alkalosis Bartter-like associated with hypocalciuria and hypomagnesemia.
134	*SLC12A4*	solute carrier family 12 (potassium/chloride transporters), member 4	6560	16q22.1	Mediates electroneutral potassium-chloride cotransport when activated by cell swelling. May contribute to cell volume homeostasis in single cells. May be involved in the regulation of basolateral Cl(-) exit in NaCl absorbing epithelia	sickle cell disease
135	***SLC14A2***	solute carrier family 14 (urea transporter), member 2	8170	18q12.1-q21.1	Specialized low-affinity urea transporter. Mediates urea transport in kidney	hypertension
136	***SLC22A2***	Solute carrier family 22 (organic cation transporter), member 2	6582	6q26	solute carrier family 22,member A2,polyspecific transporter oforganic cations,mainly expressed in kidney	Cardiomyopathy; renal failure; hodgkin disease
137	*SLC2A1*	solute carrier family 2	6513	1p35-p31.3	Facilitative glucose transporter. This isoform may be responsible for constitutive or basal glucose uptake. Has a very broad substrate specificity; can transport a wide range of aldoses including both pentoses and hexoses	Autosomal dominant GLUT1 deficiency syndrome; stroke; renal disease; breast cancer
138	*SLC2A2*	solute carrier family 2 (facilitated glucose transporter) member 2	6514	3q26.1-q26.2	Facilitative glucose transporter. This isoform likely mediates the bidirectional transfer of glucose across the plasma membrane of hepatocytes and is responsible for uptake of glucose by the beta cells; may comprise part of the glucose-sensing mechanism of the beta cell. May also participate with the Na+/glucose cotransporter in the transcellular transport of glucose in the small intestine and kidney.	Diabetes mellitus, noninsulin-dependent; Fanconi-Bickel
139	*SLC2A3*	solute carrier family 2 (facilitated glucose transporter), member 3	6515	12p13.3	Facilitative glucose transporter. Probably a neuronal glucose transporter	hyperglycemia, obesity
140	*SLC2A4*	solute carrier family 2 (facilitated glucose transporter) member 4	6517	17p13	Insulin-regulated facilitative glucose transporter.	Diabetes mellitus, noninsulin-dependent.
141	*SLC4A1*	solute carrier family 4, anion exchanger, member 1	6521	17p21-q22	Band 3 is the major integral glycoprotein of the erytrocyte membrane. Band 3 has two functional domains. Its integral domain mediates a 1: exchange of inorganic anions across the membrane, whereas its cytoplasmic domain provides binding sites for cytoskeletal proteins, glycolytic enzymes, and hemoglobin.	Acanthocytosis, one form; Elliptocytosis, Malaysian-Melanesian
142	*SLC4A2*	solute carrier family 4, anion exchanger, member	6522	7q35-q36	Plasma membrane anion exchange protein of wide distribution	biliary cirrhosis primar; liver diseases; epilepsy
143	*SLC4A3*	solute carrier family 4, anion exchanger, member 3	6508	2q36	Plasma membrane anion exchange protein of wide distribution.	neurological disease
144	*SLC4A4*	solute carrier family 4, sodium bicarbonate	8671	4q21	Electrogenic sodium/bicarbonate cotransporter with a Na(+):HCO3(-) stoichiometry varying from 1:2 to 1:3. May regulate bicarbonate influx/efflux at the basolateral membrane of cells and regulate intracellular pH	Renal tubular acidosis, proximal, with ocular abnormalities;

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
145	*SLC5A2*	solute carrier family 5 (sodium/glucose cotransporter), member 2	6524	16p12-p11	Sodium-dependent glucose transporter. Efficient substrate transport in mammalian kidney is provided by the concerted action of a low affinity high capacity and a high affinity low capacity Na(+)/glucose cotransporter arranged in series along kidney proximal tubules	renal glucosuria
146	*SLC6A2*	solute carrier family 6 (neurotransmitter transporter, noradrenalin), member 2	6530	16q12.2	Amine transporter. Terminates the action of noradrenaline by its high affinity sodium-dependent reuptake into presynaptic terminals.	Orthostatic intolerance.
147	***SLC8A1 (NCX1)***	solute carrier family 8 (sodium/calcium exchanger), member 1	6546	2p23-p22	Rapidly transporting Ca2+ during excitation-contraction coupling. Ca(2+) is extruded from the cell during relaxation SO as to prevent overloading of intracellular stores.	Heart failure; arrhythmia; ischemia; cardiomyopathy; essential hypertension; hypertension arterial; cardiovascular diseases
148	*SLC8A2*	solute carrier family 8 (sodium-calcium exchanger), member 2	6543	19q13.3	Rapidly transports Ca(2+) during excitation-contraction coupling. Ca(2+) is extruded from the cell during relaxation so as to prevent overloading of intracellular stores	ischemia
149	*SLC9A1*	solute carrier family 9 (sodium/hydrogen exchanger), member 1	6548	1p36.1-p35	Involved in pH regulation to eliminate acids generated by active metabolism or to counter adverse environmental conditions. Major proton extruding system driven by the inward sodium ion chemical gradient. Plays an important role in signal transduction	essential hypertension, ischemia, heart failure
150	*SLC9A2*	solute carrier family 9 (sodium/hydrogen exchanger), member 2	6549	2q11.2	Involved in pH regulation to eliminate acids generated by active metabolism or to counter adverse environmental conditions. Major proton extruding system driven by the inward sodium ion chemical gradient. Seems to play an important role in colonic sodium absorption	not known
151	*SLC9A5*	solute carrier family 9 (sodium/hydrogen exchanger), member 5	6553	16q22.1	Involved in pH regulation to eliminate acids generated by active metabolism or to counter adverse environmental conditions. Major proton extruding system driven by the inward sodium ion chemical gradient. Plays an important role in signal transduction	not known
152	*TBXA2R*	thromboxane A2 receptor	6915	19p13.3	Receptor for thromoxane A2 (TXA2), a potent stimulatro of patelet aggregation. The activity of this receptor is mediated by a G-protein that activate a phosphatidylinositol-calcium second messenger system. In the kidney, the binding of TXA2 to glomerular TP receptors causes intensive vasocontriction.	Bleeding disorder due to defective thromboxane A2 receptor
153	*TBXAS1*	thromboxane A synthase 1 (patelet, cytochrome P450, family 5, subfamily A)	6916	7q34-q35	Thromboxane synthase,cytochrome p450 superfamily	Thromboxane synthase deficiency; cardiovascular diseases
154	*TGFB1*	Transforming growth factor beta-1 (Camurati-Engelmann disease)	7040	19q13.2	Multifunctional peptide that controls proliferation, differentiation, and other functions in many cell types. Many cells synthesize TGF-beta 1 and essentially all of them have specigic receptors for this peptide. TGF-beta 1 regulated the actions of many other peptide growth factors and determines a positive or negative direction of their effects,	Camurati-Engelmann disease
155	*TH*	tyrosine hydroxylase	7054	11p15.5	Plays an important role in the physiology of adrenergic neurons	Segawa syndrome, neurodegenerative diseases, infarct, cardiovascular diseases, heart failure,
156	*TRH*	thyrotropin-releasing hormone	7200	3q13.3-q21	Functions as a regulator of the biosynthesis of TSH in the anterior pituitary gland and as a neurotransmitter/ neuromodulator in the central and peripheral nervous systems.	Stroke; acromegaly; hypothyroidism; adenoma
157	*TRHR*	thyrotropin-releasing hormone receptor	7201	8q23	Receptor for thyrotropin-releasing hormone. This receptor is mediated by G proteins which activate a phosphatidylinositol-calcium second messenger system.	Isolated central hypothyroidism with TSH and PRL levels unresponsive to TRH (inactivating mutations of TRHR)
158	*TRIP10*	thyroid hormone receptor interactor 10	9322	19p13.3	Required for translocation of GLUT4 to the plasma membrane in response to insulin signaling	wiskott-aldrich syndrome

Additional Table 1. - Continued

No	Gene Symbol	Gene name	Locuslink ID	Chr	Gene function	Disease/Syndrome
159	*UCP3*	uncoupling protein 3	7352	11q13	UCP are mitochondrial transporter proteins that create proton leaks across the inner mitochondrial membrane, thus uncoupling oxidative phosphorylation.	obesity, pulmonary disease chronic obstructive
160	*VEGFA*	vascular endothelial growth factor A	7422	6p12	Growth factor active in angiogenesis, vasculogenesis and endothelial cell growth. Induces endothelial cell proliferation, promotes cell migration, inhibits apoptosis, and induces permeabilization of blood vessels.	chronic myocardial ischemia, atherosclerosis
161	*VEGFB*	vascular endothelial growth factor B	7423	11q13	Growth factor for endothelial cells.	atherosclerosis, cardiovascular diseases, ischemia
162	*VEGFC*	vascular endothelial growth factor C	7424	4q34.1-q34.3	Growth factor active in angiogenesis, and endothelial cell growth, stimulating their proliferation and migration and also has effects on the permeability of blood vessels. May function in angiogenesis of the venous and lymphatic vascular systems during embryogenesis, and also in the maintenance of differentiated lymphatic endothelium in adults.	cardiovascular diseases, myocardial ischemia, carcinoma renal cell

*Genes given underlined (n=25) were used in polymorphism screening among 105 CNR regions in 46 CVD individuals from HYPEST and CADCZ cohorts; genes given in bold were subjected for variation detection in promoter regions of 29 CVD genes in 14 normotensives from HYPEST cohort; genes given in bold and underlined were used in both study stages.

Additional Tabel 2. Characteristics and detected variants of 105 CNR regions

No.	Name of CNR region	Chr. Postion	Start position*	End position*	Sequence type	CNR length (bp)	[2]Rn %	[3]Mm %	DHPLC forward primer 5´-3´	DHPLC reverse primer 5´-3´	DGGE forward primer 5´-3´	DGGE reverse primer 5´-3´	PCR product (bp)	DHPLC/DGGE detected variations; rs
1	ACE_81	17q23	62057853	62057972	noncoding	120	75,8	73,6	CATATCCTTCCCTGCAGATGACT	AGCTTGCAGCCCTCTTTGAC	-	-	298	rs4316818
2	ADD1_11	4p16.3	2877801	2877895	noncoding	95	76,4	77,1	ACCTTTGCCCTCCCTCTCTTT	GGCGCCTCTACTCATAGACACAG	-	-	391	-
3	ADD1_111	4p16.3	2963042	2963134	UTR	93	81,6	86,2	CTGCCTCTCTTCTGTGTCGTTT	CTGCTTCCCAGATTTAGACAAGG	-	-	272	-
4	ADD1_72	4p16.3	2940712	2940885	noncoding	174	80,2	78,3	AGCTCTGTCGTCAGCCTACCTTA	GTTCCTCATCTCCTGGACCTCTT	-	-	384	-
5	ADD2_10	2p14-p13	70863951	70864180	UTR	230	68,9	69,8	ACGGAAGAACATGAGAAGACAGTG	GGCTTTGGGTTGTATAAGGAGGA	-	-	396	novel (A/G), novel (C/T)
6	AGTR1_25	3q21-q25	149747664	149747806	noncoding	143	74,5	78,6	ACAAGCCCTCAGCTTCTCCTAGT	GCTGATGGGATATCATGGACAA	-	-	505	-
7	ATP1A1_50	1p13-11	116266932	116267175	noncoding	244	75,6	76,6	TCAGATCAGATGGGCTCACAGTA	AAATCCCTTTGCTCAGGATGC	-	-	479	-
8	GNB3_5	12p13	6810353	6810465	noncoding	113	75,2	73,6	TCTATCCCTCCCATCTGTCTCTG	AGGCCAGACAGCTAACTCCAAG	-	-	282	-
9	HSD11B1_29	1q32-q41	206945701	206945790	noncoding	90	76,1	75	AGTGTGTGTTAGGAGGAGCAAGG	AGGGTTTGGCAACAGTCTCAA	-	-	300	-
10	HSD11B2_93	16q22	67253843	67253967	noncoding	125	79,2	74,6	CCTTAGGACAAGTCCAACAGCAG	CAGTGCTAGTTCCTGCCTGCTTA	-	-	387	-
11	IL1A_16	2q14	113628112	113628256	noncoding	145	75,8	73,7	TTGGGTATCTCAGGCATCTCCTA	GGGCTTAGAATCACTGTAGTGCTCA	-	-	399	-
12	IL1A_24	2q14	113630323	113630464	noncoding	142	75	73,8	GGACGAAGGATAATGCTGACATC	GCCTAGGCAGGCTCTATGAGAA	-	-	336	-
13	IL1A_42	2q14	113637166	113637403	UTR	238	69	70,2	GCTGTAGCTTTAGAGAAGGCAATG	ACGATTTCTCCCTTCCTCCTAGA	-	-	462	-
14	IL1A_8	2q14	113627420	113627512	UTR	93	77,4	75,3	AACTGAGGTCCCAGAAACTTATTG	GCTGAGCATGTACTGAGCCTTG	-	-	373	rs1304037
15	KCNJ1_17	11q24	128246129	128246222	UTR	94	71,4	70,4	AGCATTCTGTCTTGCAGGAAGGT	GCTCATATAAGGTGGATTGGAGCA	-	-	351	-
16	NCX1_121	2p23-p22	40362540	40362732	noncoding	193	65,3	70,4	TGAACTGTTCTCTCCAGCTGTTAGTT	GAGAAGCAAGTTGGCAAAGCAT	-	-	394	-
17	NCX1_134	2p23-p22	40367504	40367628	noncoding	125	73,8	76	GGGTGCTGAGGAAATGTTAATGG	GTCCTAGAAATGAGGCCATCAGG	-	-	282	novel (T/C)
18	NCX1_238	2p23-p22	40399105	40399193	noncoding	89	75,7	76,1	ACTGAGGGATCCACTTTCTACGC	TACCAGAGAGTGAGGATGCAAGC	-	-	370	rs449383
19	NCX1_273	2p23-p22	40415400	40415566	noncoding	167	74,9	73,3	AGCCCAGCATCTGACAAACAAT	GATTCCTAGCCCTCATCAAATGG	-	-	464	-
20	NCX1_291	2p23-p22	40421956	40422070	noncoding	115	83,5	83,9	CAAGAGAGGAAGAAGCAGGAGGA	AGTGGCAAACAAAGTGCACAGA	-	-	292	rs2540904
21	NCX1_348	2p23-p22	40456707	40456866	noncoding	160	76,2	73,9	TAGGATTGAGCCCTGTCACTGAG	GCAAGTGTGCAGATGGTAAGGAT	GGATTGAGCCCTGTCACTG	CGCCCGCCGCGCCCCCGCCCGG CCCGCCGCCCCCGCCCGGGGGAG AAAAGACACTGTTT	348	rs11274804 and novel C/g; and del43bp; del10bp, del5bp, dup40bp
22	NCX1_364	2p23-p22	40462795	40462906	noncoding	112	76,3	72,1	TGCATGATGGAAGCAATGGAG	TGCAAGCACAAGCTGTTAGTGTG	-	-	250	-
23	NCX1_387	2p23-p22	40470526	40470643	noncoding	118	77,9	74,8	CAGTGAGTGACAGCCCAAGAACT	TTCATAGGAGCATGAGTGCCTGT	-	-	293	novel (A/G)
24	NCX1_418	2p23-p22	40483379	40483491	noncoding	113	74,6	74,8	GCCTTGGCAAGTCTAATGGAGAA	AATCCAGGAGGGCTTTATTCCAC	-	-	334	-
25	NCX1_422	2p23-p22	40483841	40484007	noncoding	167	77,2	73,7	CCCACACAAGAGCGTACCTTTAAT	TGGATGCTCAATGTCTTGGGTA	-	-	395	-
26	NCX1_462	2p23-p22	40500819	40501023	noncoding	205	76,5	76	TTCCAGGCTCCTGATTCTCAAG	GGAGATGAGTTCAGCCACAATTC	-	-	399	-
27	NCX1_504	2p23-p22	40528030	40528158	noncoding	129	78,5	72,9	GTCTTTGCCTGTGCTGCTAAAC	TCCTTTAACTCTTGGCAGTTCG	-	-	361	rs72943138
28	NCX1_533	2p23-p22	40538108	40538300	noncoding	193	73,5	75,5	GGGCCTCTCTTGTCTGCAAT	TTAAGGCATCCTAAAGCGGAAC	-	-	365	-
29	NCX1_650	2p23-p22	40596296	40596570	noncoding	275	79,4	80,7	GGTAGCAAGCAAATGGCAGATT	GATGTTACCTCCTCAGCTCCAGA	-	-	469	rs2192773
30	NCX1_740	2p23-p22	40630039	40630128	noncoding	90	65,6	72,2	GATCTCTGTGTGCAGAGCATATTG	ACTCATGTGAGTGAGAGCATTGG	-	-	367	-
31	NCX1_762	2p23-p22	40635880	40636132	noncoding	253	74,7	75	CCTAGCCCACATCCCACAATA	TAAGAAAGCCAGGCAGCAAATC	-	-	489	rs4952414, rs17026003
32	NEDD4L_170	18q21	54029848	54029931	noncoding	84	77	71,3	TTCTGTATGTTAAGCAGGGAGTGG	ACAGGAAGACCAGTAGACCACCA	-	-	375	-
33	NEDD4L_262	18q21	54075851	54075941	noncoding	91	76,5	70,7	AGGAAGTTGGACCACCATTTACC	GAGAGGACACAGAGGAAAGTCCA	-	-	341	-
34	NEDD4L_265	18q21	54083830	54083925	noncoding	96	71,3	79,9	ATGTCGCCTTAGTATGTGGCAGA	GGGTGAAGCTTGCAGGATAAATG	-	-	396	novel (G/A)
35	NEDD4L_267	18q21	54083968	54084124	noncoding	157	76,8	79,9	GATGCTTTCTGCTGCCTGTGT	CAGAAGGCGCAGAGCAATTT	-	-	325	novel (G/A)
36	NEDD4L_273	18q21	54085218	54085367	noncoding	150	81,7	82,7	TCTATGGGAGACAGTGCTCTTGG	AAATACTGCTCGATCCCATCTCC	-	-	369	-
37	NEDD4L_351	18q21	54150127	54150202	noncoding	76	70,3	72,1	CAGCAAAGAGAACCCTCCTCAAG	TGGGTGGCAGTAACAATTCTCTG	-	-	265	-
38	NEDD4L_361	18q21	54151833	54151922	noncoding	90	68,1	71,3	GGAACCAACGGTAATGATCCACT	ACAGTCTCATGCAATTCCTGTGG	-	-	394	-
39	NEDD4L_369	18q21	54157121	54157238	noncoding	118	74,4	79,8	CAGCAACCCTTTACGTTTGAGAC	GAGTCCAAGTTGTGGTTCGATTG	-	-	388	rs4149609
40	NEDD4L_376	18q21	54158124	54158219	noncoding	96	77,3	72,8	CTCTGTCCTGTGACTCCATTCT	AAGAACGGTGGAAGTCCTGAAA	-	-	249	-
41	NEDD4L_39	18q21	53868179	53868309	noncoding	131	74	77,2	GAAGGGTGGAAGAAGTGTGACAA	CTACAACTTCAGCCACGGTCTTG	-	-	400	-
42	NEDD4L_397	18q21	54167247	54167477	noncoding	231	75	74	GCCTGGCTGGGAAGAAGAAT	ATCCTGCCATAGGACCCTTGA	-	-	480	-
43	NEDD4L_411	18q21	54182003	54182102	noncoding	100	68,7	70,6	GGAGAAGTGCCCACAGTTGTTTA	CGGACATAATTCTCCGATAGGACT	-	-	372	-
44	NEDD4L_444	18q21	54213960	54214100	UTR	141	75,5	70	GCTATTGAGCTGGTAGTGGCAGA	ACAACCTTCCATGGCTTCTCAC	-	-	336	-
45	NPR1_25	1q21-q22	150872905	150872999	noncoding	95	72,8	70,2	CCCTAAGCACAGTCGAGTAGGTG	GCCTCCTCTGCTATTGGTTCTC	-	-	326	-
46	NPR1_5	1q21-q22	150860517	150860595	noncoding	79	100	100	CGCCTACTGTGTCAGGAACTAGC	CGTGCAGTGTTCATAGTTGCCTA	-	-	388	novel (insT)
47	NPR2_115	9p21-p12	35808398	35808539	noncoding	142	74,6	75,2	GCTTGAATGTGAGTGTGATGGAG	CTGACTTCTAGGCCCTGCAGATA	-	-	364	-
48	NPR2_123	9p21-p12	35809187	35809421	noncoding	235	83,8	81	GGCTCTCATACCTGCTCTCCAT	AAAGCATGTCTGCGTGCATAG	-	-	395	-
49	NPR2_70	9p21-p12	35797396	35797528	noncoding	133	66	65,7	CAAGTCTCAGGGCCTCTGCT	CTGGAGAGAAATGTTGGATAGGC	-	-	398	rs13294295
50	NPR2_73	9p21-p12	35797805	35797893	UTR	89	70,9	70,5	GGATGTGGGTCCAGACTGATTTA	GTCTGCCACATGCCAAACAAT	-	-	257	-
51	NR3C2_11	4q.31.1	149577957	149578071	UTR	115	77,6	79,5	TGGAGACACACAAAGAGCCAGTA	CATGCAGGCAACATTACCTTGT	-	-	375	-
52	NR3C2_222	4q.31.1	149743542	149743779	noncoding	238	77,8	75,1	GCCTTGAAGGGAGAGGAAGAGC	CCGGCTCCTCCAAGCACTAA	-	-	400	-
53	NR3C2_234	4q.31.1	149747512	149747593	noncoding	82	76,8	74,5	CTCAGTTGGCCCTTTAGCACAAT	CCCTGCCCATCTCTGAGTTAGAA	-	-	328	-
54	NR3C2_276	4q.31.1	149771035	149771181	noncoding	147	75,5	74,9	GTGGCGGTTTGCAGACAAAT	CCAAGGGTCCTTTCACACACTTA	-	-	323	-
55	NR3C2_37	4q.31.1	149608864	149609084	noncoding	221	75,2	74,4	CAACCCACTCCGATGAAGTCTC	GGTAAAAACTGACCAACCCCAGT	-	-	496	-
56	NR3C2_449	4q.31.1	149875000	149875201	noncoding	202	73,3	72,4	TCAAGGTCAGGAGTAAGGCAAA	TGAGAGTCACTGAGGACAGAGAAA	-	-	380	-
57	NR3C2_45	4q.31.1	149611897	149612101	noncoding	205	78,2	86,2	TAGCATGACTGCGTACAACATCG	GCCCTCAGATTTGCTACAGGTG	-	-	448	-
58	NR3C2_518	4q.31.1	149904354	149904453	noncoding	100	73,8	74,8	CACTTGTAGGGAACGCACAAAGA	TGAGTTGGAGAATAGTGGCAAGG	-	-	387	-
59	NR3C2_531	4q.31.1	149909252	149909393	noncoding	142	72,2	73,9	ATGAGGCAATGGACAGAGAAGTC	ACTCACTGCTTGGAGTGACCTTT	-	-	393	rs17582031
60	NR3C2_54	4q.31.1	149618841	149618935	noncoding	95	71,9	72,6	TCACATGCCCAAAGCTTCATAC	CTGCCAGTCTGGGTTTGATAATG	-	-	391	-
61	NR3C2_665	4q.31.1	149946967	149947055	noncoding	89	76,9	78,9	GGAAAGCCATGGAGTTAAGCTGT	CCCTGTCTTTCTCAACTCTTTCCA	-	-	306	-
62	NR3C2_70	4q.31.1	149626022	149626112	noncoding	91	73,2	75,8	TCCATTGCTGTTTCCATTGC	GAGCTTGCTTCTCTTTCCCAAA	-	-	364	-
63	REN_6	1q32	201294174	201294291	noncoding	118	74,8	74,4	TCCCAGTCCTCTTCCAGTCTTC	CTTGTCTCCTCCACCCTTTAGGA	-	-	300	-
64	SAH_19	16p13.11	20747765	20747858	noncoding	94	71,9	71,2	TGGGCTAGGACCTGATAGTTCTG	ACTTGGCATGACGTAGCATCTTC	-	-	383	rs1027457
65	SAH_5	16p13.11	20743202	20743319	noncoding	118	73,3	75,4	TGGAGTCACCCTGGGAAAGTTAT	GATTTGTTCCCTAGCTCCCATACA	-	-	293	-
66	SAH_78	16p13.11	20778919	20779042	noncoding	124	73,7	76,6	GGATGTTGAGCTGATGAAGACTG	GATTTCAGAGCCTTCTGCTTCTG	-	-	377	-
67	SCNN1A_16	12p13	6326277	6326373	UTR	97	73,7	76,3	AGCTTGGGTGATGACTTGTATCC	ATCCCTTAGAACCCTGCTCAGAC	-	-	366	novel (C/T)
68	SCNN1G_51	16p12	23185578	23185679	noncoding	102	77,1	80,4	ATGTCTGTCTGTCTGAGATGCAAG	GGTTTGAGGCAGGTGGAAAC	-	-	332	-
69	SGK_9	6q23	134471744	134471912	UTR	169	72,3	78	ATATTCGTCATCACAGCCCAGAC	CTCTTTCCTCTGAACCCTGTTAGG	-	-	357	-
70	SLC12A1_131	15q15-q21.1	46267735	46267863	noncoding	129	74,7	74,4	TTTCCTCCCTGCTTATCTTGGA	GGCCATTATGTCTGGTTCTTTCC	-	-	370	-
71	SLC12A1_141	15q15-q21.1	46272810	46272944	noncoding	135	73,6	76,3	AGGCAGTTCAAGCCCTTATAACC	ACAGAGGAATTCACCAAGTGCAG	-	-	384	-

Additional Tabel 2. Continued

No.	Name of CNR region	Chr. Postion	Start position*	End position*	Sequence type	CNR length (bp)	[2]Rn %	[3]Mm %	DHPLC forward primer 5´-3´	DHPLC reverse primer 5´-3´	DGGE forward primer 5´-3´	DGGE reverse primer 5´-3´	PCR product (bp)	DHPLC/DGGE detected variations; rs
72	SLC12A1_162	15q15-q21.1	46276113	46276253	noncoding	141	73,2	78,3	ACAGCTGTAGTTAGGCAGCCAAG	GTCTAGCAGCTGCCCTAACACC	-	-	343	-
73	SLC12A1_197	15q15-q21.1	46291679	46291836	noncoding	158	79,9	75,3	CAGCAGAGCTGTAATTGGTTGTG	GAGCCCATTGCGTTATCTATCTG	-	-	367	novel (A/T)
74	SLC12A1_207	15q15-q21.1	46294125	46294197	noncoding	73	73,5	76,9	TGTGGACTTATGGTATGCAGTCCT	ATCCTGCTGTTTGACTTGGAGAC	-	-	288	-
75	SLC12A1_224	15q15-q21.1	46304615	46304805	noncoding	191	81,7	77,5	GTTGCCGGGTGATACACAAT	CCTTCCTCATCCTTGTTCCTC	-	-	498	novel (G/C)
76	SLC12A1_70	15q15-q21.1	46234683	46234833	noncoding	151	81,4	83	AGGCAAACCAGTTTCTCCTTAGC	AGAACATCGGGCTACTTCCTTGT	-	-	352	-
77	SLC12A1_71	15q15-q21.1	46234834	46234914	noncoding	81	81,4	83	GCAGTGTAGTTTGCAGCAATAGG	CTACAATGTGCCAGAGAGATCCA	-	-	388	-
78	SLC14A2_24	18q12.1-q21.1	41422245	41422462	noncoding	218	72,4	78	CTGAATCCCAGTTTGGCTCTTT	CCAAGACCTGGGAAAGGAGAG	-	-	400	-
79	SLC14A2_34	18q12.1-q21.1	41428105	41428218	noncoding	114	75,4	74,8	CTCTGTTCACAAGGCACCTCTC	ACAGGAAATGCCCATTCTGG	-	-	376	novel (T/C)
80	SLC14A2_75	18q12.1-q21.1	41451491	41451723	noncoding	233	82,4	82	GGTTCTTACAAAGAGGATGGAGGA	CTCCCAAATTGAAGGGAGAGAG	-	-	492	-
81	SLC22A2_30	6q26	160577382	160577436	noncoding	55	70,9	69,9	GGTGAAGGCTTAAGTTGCACAAG	GTGGCTGATCTCCCAGAATAAGA	-	-	375	rs316021
82	SLC22A2_37	6q26	160580530	160580622	noncoding	93	75,3	72,2	AAATGGCTGGCTGAGTAACCTA	AAAGGGACCCAGCTTCAAGA	-	-	391	-
83	WNK1_117	12p13	841698	841796	noncoding	99	73,7	75,8	CATGAACAGGCACATTCTACAGG	GCTTCATGTAGGCAGTTCAGCTA	-	-	296	rs12825084
84	WNK1_139	12p13	846970	847070	noncoding	101	87,6	71,6	GGCCTCTAACATATCACAGGGTAA	ATGCACAACTGGGTAGTTCTGTG	-	-	429	-
85	WNK1_155	12p13	858084	858176	noncoding	93	69,9	80,6	GGGTAACCAACCCTTGAAGTAGG	GGGTACTTCTCAAGTGATTAGGAGC	-	-	353	novel*AluYb8*
86	WNK1_165	12p13	860037	860147	noncoding	111	73,9	73	GTCCTGTCTCTTCTGTTGGAACC	TGTGGGAAAGTACCCAGGTGTAG	-	-	348	-
87	WNK1_168	12p13	860562	860692	noncoding	131	79,5	79,4	CCAGGAGGGAGTTTAGCACAAG	AATGATTGCTGAGAGCACAGGAG	-	-	399	novel (A/C)
88	WNK1_208	12p13	877345	877471	noncoding	127	75	73,3	TGTCTCCAGCTGTATGTGACCTG	CACCGTGCCTAACCAATGACT	-	-	486	-
89	WNK1_214	12p13	880862	881152	noncoding	291	86,8	85,9	CTTGATTGGTAAGGAGGGCAAAC	GCCCAAACTTCTGAGGGTAACA	-	-	487	-
90	WNK1_218	12p13	882302	882503	noncoding	202	75,2	74,1	TCTGATGAGTCAGTGCCTACCTG	GGAGCCAATCCATTTCTAATGC	-	-	373	-
91	WNK1_222	12p13	883173	883364	noncoding	192	72,7	76,1	GGCCCTACATCCTTAGAAGTCACA	GATCCTGTCCCACTTCCTGTTTC	-	-	413	rs10849582
92	WNK1_227	12p13	883843	883984	noncoding	142	78,5	76,6	TGTCTCACCTCCTGCTTTCTCTT	TGTCTGTCACAAGCTAGGCTCAC	-	-	368	-
93	WNK1_246	12p13	888231	888518	noncoding	288	80,5	80,1	GGAACTGCCTCCTTGCAGAAT	AGAGCTCCACCCACTCGAAATTA	-	-	452	-
94	WNK1_250	12p13	889267	889379	noncoding	113	77	75,7	CATGTGGAAGGAAATCGTAGGTC	ACGTCTGTTCAAGCCTAGCTCAA	-	-	339	novel (T/C)
95	WNK1_53	12p13	790344	790469	noncoding	126	92,6	96,2	AAGGGACGGAAGATGTGCAA	CTCTTCACCATTATGGAGCTGGA	-	-	398	novel (G/A), rs36092349
96	WNK1_63	12p13	808723	808848	noncoding	126	72,6	63	GCCTGTTTCTTGGAAATGCCTA	CAGCTACACTCTTGGCATTAACACT	-	-	380	rs72648627
97	WNK1_90	12p13	831113	831231	noncoding	119	75	74,3	GTCTGACTTTGATGTCTGCCTTC	CCTCCTCACACAGAATACACCAT	-	-	381	-
98	WNK4_100	17p21-q22	41322583	41322817	noncoding	235	79,8	85,8	CTTGTCTTTCACCCTTCATCCAC	CTGTGCAGTGTTTGAGCCTACAG	-	-	383	-
99	WNK4_120	17p21-q22	41330143	41330332	noncoding	190	82,8	74,9	CCTAGGTTAGGTGATACAACTCATG	GCAGCATCTGGCCATAGAATAA	-	-	385	-
100	WNK4_122	17p21-q22	41330352	41330468	noncoding	117	74,5	74,9	CCTGAATTGCCAGCATACCTATC	TCTCTTAGGATCCTCCCATCTCC	-	-	378	-
101	WNK4_127	17p21-q22	41331417	41331585	noncoding	169	75,6	74,3	TTGCTTGAACCTGGAAGGTGTA	GGTAGAGGTAGAGAAACCTCCATC	-	-	372	-
102	WNK4_62	17p21-q22	41313222	41313312	noncoding	91	68	70,3	TGACCTATCTCCCTCCTTGTGAA	ATGACTGTCAGAGCAGACACTGG	-	-	323	-
103	WNK4_91	17p21-q22	41321872	41321951	UTR	80	80,5	79,5	GATGTTGGCAGGATGGTGAG	ATGAAGATGGTAGGCAGCAGAAC	-	-	354	-
104	WNK4_92	17p21-q22	41321953	41322091	noncoding	139	72,6	72,7	CACCTTGGAGGTTTCTTCATCAC	GCTGCTAATGTGGTGTGAGCATA	-	-	395	-
105	WNK4_97	17p21-q22	41322339	41322509	noncoding	171	80,7	75,4	CACCACATTAGCAGCAACCAA	TCTGTCCTCTACCCACAAGAACA	-	-	398	novel (A/C)

*Start and end positions of the PCR product

[2]Sequence identity with *Rattus norvegicus*

[3]Sequence identity with *Mus musculus*

CNR - conserved non-coding region

CURRICULUM VITAE

Katrin Kepp

Date of birth:	10 December 1978, Viljandi, Estonia
Address:	Department of Biotechnology Institute of Molecular and Cell Biology, University of Tartu Riia str. 23, 51010, Tartu, Estonia
Phone:	+372 737 5009
E-mail:	katrin.kepp@ut.ee

Education

1994–1997	Carl Robert Jakobson Gymnasium, Viljandi
1997–2001	B.Sc University of Tartu, Institute of Molecular and Cell Biology, molecularbiology
2001–2003	M.Sc University of Tartu, Institute of Molecular and Cell Biology, moleculardiagnostics
2003–…	Ph.D student University of Tartu, Institute of Molecular and Cell Biology, moleculardiagnostics

Professional employment

2000–2001	Asper Biotech Ltd., laboratory assistant
2001–2005	Asper Biotech Ltd., scientist-project manager
2004	University of Tartu, laboratory assistant
2008	University of Tartu, Institute of Molecular and Cell Biology, extraordinary scientist
2006–…	University of Tartu, Institute of Molecular and Cell Biology, Graduate School project manager

Scientific work and activity

Main research interest: pattern of genetic variations in the human genome, detection of genetic variations in susceptibility to the human complex diseases.

List of publications:

1. **Kepp K**, Org E, Sõber S, Kelgo P, Viigimaa M, Veldre G, Tõnisson N, Juhanson P, Putku M, Kindmark A, Kozich V, Laan M. BMC Med Genet. 2010 Jan 28;11:15.PMID: 20109173. Hypervariable intronic region in *NCX1* is enriched in short insertion-deletion polymorphisms and showed association with cardiovascular traits.
2. Annilo T, **Kepp K**, Laan M. BMC Mol Biol. 2009 Aug 11;10:81.PMID: 19671135. Natural antisense transcript of natriuretic peptide precursor A (*NPPA*): structural organization and modulation of NPPA expression.
3. Sõber S, Org E, **Kepp K**, Juhanson P, Eyheramendy S, Gieger C, Lichtner P, Klopp N, Veldre G, Viigimaa M, Döring A; Kooperative Gesundheitsforschung in der Region Augsburg Study, Putku M, Kelgo P; HYPertension in ESTonia Study, Shaw-Hawkins S, Howard P, Onipinla A, Dobson RJ, Newhouse SJ, Brown M, Dominiczak A, Connell J, Samani N, Farrall M; MRC British Genetics of Hypertension Study, Caulfield MJ, Munroe PB, Illig T, Wichmann HE, Meitinger T, Laan M. PLoS One. 2009 Jun 29;4(6):e6034.PMID: 19562039. Targeting 160 candidate genes for blood pressure regulation with a genome-wide genotyping array.
4. Juhanson P, **Kepp K**, Org E, Veldre G, Kelgo P, Rosenberg M, Viigimaa M, Laan M. BMC Med Genet. 2008 Apr 10;9:25.PMID: 18402670. N-acetyltransferase 8, a positional candidate for blood pressure and renal regulation: resequencing, association and in silico study.
5. **Kepp K**, Juhanson P, Kozich V, Ots M, Viigimaa M, Laan M. BMC Med Genet. 2007 Jul 23;8:47.PMID: 17645789. Resequencing *PNMT* in European hypertensive and normotensive individuals: no common susceptibilily variants for hypertension and purifying selection on intron 1.
6. Lahermo P, Liljedahl U, Alnaes G, Axelsson T, Brookes AJ, Ellonen P, Groop PH, Halldén C, Holmberg D, Holmberg K, Keinänen M, **Kepp K**, Kere J, Kiviluoma P, Kristensen V, Lindgren C, Odeberg J, Osterman P, Parkkonen M, Saarela J, Sterner M, Strömqvist L, Talas U, Wessman M, Palotie A, Syvänen AC. Hum Mutat. 2006 Jul;27(7):711-4.PMID: 16786507. A quality assessment survey of SNP genotyping laboratories.

Fellowships:

1. Kristjan-Jaak fellowship, 01.10.04 - 01.03.05, to practice in Uppsala University in Sweden.
2. WSF (World Federation of Scientists) scholarship, 2005/2006 academic year.
3. Estonian Students Fund in USA 2007/2008 academic year
4. Estonian World Council, INC (ÜEKN) for 2008/2009 academic year
5. EcoGene fellowship 2010/2011, for young women in science (women returning from maternity leave)
6. National Human Genetics Society a “National Fellowship” for the European Human Genetics Conference 2010, June 12-15, 2010, Gothenburg, Sweden

Membership in organizations:

2001 - ... Member of the Estonian Society of Human Genetics
2001 - ... Member of the Estonian Biochemical Society

CURRICULUM VITAE

Katrin Kepp

Sünniaeg ja koht:	10 Detsember 1978, Viljandi, Eesti
Aadress:	Tartu Ülikool, Molekulaar- ja Rakubiolooga Instituut, Biotehnoloogia õppetool, Riia 23, 51010, Tartu, Eesti
Phone:	+372 737 5009
E-mail:	katrin.kepp@ut.ee

Haridus

1994–1997	Carl Robert Jakobsoni nimeline Gümnaasium
1997–2001	B.Sc molekulaarbioloogia erialal, Tartu Ülikool, Loodus- ja Tehnoloogiateaduskond, Molekulaar- ja Rakubioloogia Instituut
2001–2003	M.Sc molekulaardiagnostika erialal, Tartu Ülikool, Loodus- ja Tehnoloogiateaduskond, Molekulaar- ja Rakubioloogia Instituut
2003–…	Doktorantuur, molekulaardiagnostika erialal, Tartu Ülikool, Loodus- ja Tehnoloogiateaduskond, Molekulaar- ja Rakubioloogia Instituut

Erialane teenistuskäik

2000–2001	Asper Biotech Ltd., laborant
2001–2005	Asper Biotech Ltd., teadur-projektijuht
2004	Tartu Ülikool, laborant
2008	Tartu Ülikool, Molekulaar- ja Rakubioloogia Instituut, erakorraline teadur
2006– …	Tartu Ülikool, Molekulaar- ja Rakubioloogia Instituut, biomeditsiini ja biotehnoloogia doktorikool (BMBT DK); projektijuht

Teaduslik ja arendustegevus

Peamised uurimisvaldkonnad: inimese genoomis esinevad geneetilised varieeruvused nende detekteerimine ja seos kompleksfenotüüpide kujunemisega.

Publikatsioonide loetelu:

1. **Kepp K**, Org E, Sõber S, Kelgo P, Viigimaa M, Veldre G, Tõnisson N, Juhanson P, Putku M, Kindmark A, Kozich V, Laan M. BMC Med Genet. 2010 Jan 28;11:15.PMID: 20109173. Hypervariable intronic region in *NCX1* is enriched in short insertion-deletion polymorphisms and showed association with cardiovascular traits.
2. Annilo T, **Kepp K,** Laan M. BMC Mol Biol. 2009 Aug 11;10:81.PMID: 19671135. Natural antisense transcript of natriuretic peptide precursor A (*NPPA*): structural organization and modulation of NPPA expression.
3. Sõber S, Org E, **Kepp K**, Juhanson P, Eyheramendy S, Gieger C, Lichtner P, Klopp N, Veldre G, Viigimaa M, Döring A; Kooperative Gesundheitsforschung in der Region Augsburg Study, Putku M, Kelgo P; HYPertension in ESTonia Study, Shaw-Hawkins S, Howard P, Onipinla A, Dobson RJ, Newhouse SJ, Brown M, Dominiczak A, Connell J, Samani N, Farrall M; MRC British Genetics of Hypertension Study, Caulfield MJ, Munroe PB, Illig T, Wichmann HE, Meitinger T, Laan M. PLoS One. 2009 Jun 29;4(6):e6034.PMID: 19562039. Targeting 160 candidate genes for blood pressure regulation with a genome-wide genotyping array.
4. Juhanson P, **Kepp K**, Org E, Veldre G, Kelgo P, Rosenberg M, Viigimaa M, Laan M. BMC Med Genet. 2008 Apr 10;9:25.PMID: 18402670. N-acetyl-transferase 8, a positional candidate for blood pressure and renal regulation: resequencing, association and in silico study.
5. **Kepp K**, Juhanson P, Kozich V, Ots M, Viigimaa M, Laan M. BMC Med Genet. 2007 Jul 23;8:47.PMID: 17645789. Resequencing *PNMT* in European hypertensive and normotensive individuals: no common susceptibilily variants for hypertension and purifying selection on intron 1.
6. Lahermo P, Liljedahl U, Alnaes G, Axelsson T, Brookes AJ, Ellonen P, Groop PH, Halldén C, Holmberg D, Holmberg K, Keinänen M, **Kepp K**, Kere J, Kiviluoma P, Kristensen V, Lindgren C, Odeberg J, Osterman P, Parkkonen M, Saarela J, Sterner M, Strömqvist L, Talas U, Wessman M, Palotie A, Syvänen AC. Hum Mutat. 2006 Jul;27(7):711-4.PMID: 16786507. A quality assessment survey of SNP genotyping laboratories.

Stipendiumid:

1. SA Archimedes Kristjan-Jaagu nimeline stipendium, 01.10.04 - 01.03.05, erialane praktika Uppsala Ülikoolis, Rootsis.
2. WSF (World Federation of Scientists) stipendium 2005/2006 akadeemiliseks õppeaastaks
3. Eesti Üliõpilaste Toetusfondi (USA) stipendium 2007/2008 akadeemiliseks õppeaastaks
4. Ülemaailmne Eesti Kesknõukogu (ÜEKN) stipendium 2008/2009 akadeemiliseks õppeaastaks
5. EcoGene naisteadlaste stipendium 2010/2011 aastaks, lapsepuhkuselt naasvatele naisteadlastele
6. Rahvusvahelise inimesegeneetika ühingu poolt välja antud toetus 2010. aastal aastakonverentsil osalemiseks Gothenburgis Rootsis

Membership in organizations:

2001 - ... Eesti Inimesegeneetika Seltsi liige
2001 - ... Eesti Biokeemia Seltsi liige

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand.** The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro*. Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous crassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.

41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.
42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane.** Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm.** Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg.** Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild.** The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu.** Studies of the TOL plasmid transcription factor XylS. Tartu 2000. 88 p.

61. **Dina Lepik.** Modulation of viral DNA replication by tumor suppressor protein p53. Tartu 2000. 106 p.
62. **Kai Vellak.** Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu 2000. 122 p.
63. **Jonne Kotta.** Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu 2000. 160 p.
64. **Georg Martin.** Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000. 139 p.
65. **Silvia Sepp.** Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira.** On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000. 96 p.
67. **Priit Zingel.** The role of planktonic ciliates in lake ecosystems. Tartu 2001. 111 p.
68. **Tiit Teder.** Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu 2001. 122 p.
69. **Hannes Kollist.** Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu 2001. 80 p.
70. **Reet Marits.** Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu 2001. 112 p.
71. **Vallo Tilgar.** Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002. 126 p.
72. **Rita Hõrak.** Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002. 108 p.
73. **Liina Eek-Piirsoo.** The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002. 74 p.
74. **Krõõt Aasamaa.** Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002. 110 p.
75. **Nele Ingerpuu.** Bryophyte diversity and vascular plants. Tartu, 2002. 112 p.
76. **Neeme Tõnisson.** Mutation detection by primer extension on oligonucleotide microarrays. Tartu, 2002. 124 p.
77. **Margus Pensa.** Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003. 110 p.
78. **Asko Lõhmus.** Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003. 168 p.
79. **Viljar Jaks.** p53 — a switch in cellular circuit. Tartu, 2003. 160 p.
80. **Jaana Männik.** Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003. 140 p.
81. **Marek Sammul.** Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003. 159 p

82. **Ivar Ilves.** Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003. 89 p.
83. **Andres Männik.** Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003. 109 p.
84. **Ivika Ostonen.** Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003. 158 p.
85. **Gudrun Veldre.** Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003. 199 p.
86. **Ülo Väli.** The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004. 159 p.
87. **Aare Abroi.** The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004. 135 p.
88. **Tiina Kahre.** Cystic fibrosis in Estonia. Tartu, 2004. 116 p.
89. **Helen Orav-Kotta.** Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004. 117 p.
90. **Maarja Öpik.** Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004. 175 p.
91. **Kadri Tali.** Species structure of *Neotinea ustulata*. Tartu, 2004. 109 p.
92. **Kristiina Tambets.** Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylogeographic approach. Tartu, 2004. 163 p.
93. **Arvi Jõers.** Regulation of p53-dependent transcription. Tartu, 2004. 103 p.
94. **Lilian Kadaja.** Studies on modulation of the activity of tumor suppressor protein p53. Tartu, 2004. 103 p.
95. **Jaak Truu.** Oil shale industry wastewater: impact on river microbial community and possibilities for bioremediation. Tartu, 2004. 128 p.
96. **Maire Peters.** Natural horizontal transfer of the *pheBA* operon. Tartu, 2004. 105 p.
97. **Ülo Maiväli.** Studies on the structure-function relationship of the bacterial ribosome. Tartu, 2004. 130 p.
98. **Merit Otsus.** Plant community regeneration and species diversity in dry calcareous grasslands. Tartu, 2004. 103 p.
99. **Mikk Heidemaa.** Systematic studies on sawflies of the genera *Dolerus, Empria,* and *Caliroa* (Hymenoptera: Tenthredinidae). Tartu, 2004. 167 p.
100. **Ilmar Tõnno.** The impact of nitrogen and phosphorus concentration and N/P ratio on cyanobacterial dominance and N_2 fixation in some Estonian lakes. Tartu, 2004. 111 p.
101. **Lauri Saks.** Immune function, parasites, and carotenoid-based ornaments in greenfinches. Tartu, 2004. 144 p.

102. **Siiri Rootsi.** Human Y-chromosomal variation in European populations. Tartu, 2004. 142 p.
103. **Eve Vedler.** Structure of the 2,4-dichloro-phenoxyacetic acid-degradative plasmid pEST4011. Tartu, 2005. 106 p.
104. **Andres Tover.** Regulation of transcription of the phenol degradation *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 126 p.
105. **Helen Udras.** Hexose kinases and glucose transport in the yeast *Hansenula polymorpha*. Tartu, 2005. 100 p.
106. **Ave Suija.** Lichens and lichenicolous fungi in Estonia: diversity, distribution patterns, taxonomy. Tartu, 2005. 162 p.
107. **Piret Lõhmus.** Forest lichens and their substrata in Estonia. Tartu, 2005. 162 p.
108. **Inga Lips.** Abiotic factors controlling the cyanobacterial bloom occurrence in the Gulf of Finland. Tartu, 2005. 156 p.
109. **Kaasik, Krista.** Circadian clock genes in mammalian clockwork, metabolism and behaviour. Tartu, 2005. 121 p.
110. **Juhan Javoiš.** The effects of experience on host acceptance in ovipositing moths. Tartu, 2005. 112 p.
111. **Tiina Sedman.** Characterization of the yeast *Saccharomyces cerevisiae* mitochondrial DNA helicase Hmi1. Tartu, 2005. 103 p.
112. **Ruth Aguraiuja.** Hawaiian endemic fern lineage *Diellia* (Aspleniaceae): distribution, population structure and ecology. Tartu, 2005. 112 p.
113. **Riho Teras.** Regulation of transcription from the fusion promoters generated by transposition of Tn*4652* into the upstream region of *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 106 p.
114. **Mait Metspalu.** Through the course of prehistory in india: tracing the mtDNA trail. Tartu, 2005. 138 p.
115. **Elin Lõhmussaar.** The comparative patterns of linkage disequilibrium in European populations and its implication for genetic association studies. Tartu, 2006. 124 p.
116. **Priit Kupper.** Hydraulic and environmental limitations to leaf water relations in trees with respect to canopy position. Tartu, 2006. 126 p.
117. **Heili Ilves.** Stress-induced transposition of Tn*4652* in *Pseudomonas Putida.* Tartu, 2006. 120 p.
118. **Silja Kuusk.** Biochemical properties of Hmi1p, a DNA helicase from *Saccharomyces cerevisiae* mitochondria. Tartu, 2006. 126 p.
119. **Kersti Püssa.** Forest edges on medium resolution landsat thematic mapper satellite images. Tartu, 2006. 90 p.
120. **Lea Tummeleht.** Physiological condition and immune function in great tits (*Parus major* l.): Sources of variation and trade-offs in relation to growth. Tartu, 2006. 94 p.
121. **Toomas Esperk.** Larval instar as a key element of insect growth schedules. Tartu, 2006. 186 p.

122. **Harri Valdmann.** Lynx (*Lynx lynx*) and wolf (*Canis lupus*) in the Baltic region: Diets, helminth parasites and genetic variation. Tartu, 2006. 102 p.
123. **Priit Jõers.** Studies of the mitochondrial helicase Hmi1p in *Candida albicans* and *Saccharomyces cerevisia*. Tartu, 2006. 113 p.
124. **Kersti Lilleväli.** Gata3 and Gata2 in inner ear development. Tartu, 2007. 123 p.
125. **Kai Rünk.** Comparative ecology of three fern species: *Dryopteris carthusiana* (Vill.) H.P. Fuchs, *D. expansa* (C. Presl) Fraser-Jenkins & Jermy and *D. dilatata* (Hoffm.) A. Gray (Dryopteridaceae). Tartu, 2007. 143 p.
126. **Aveliina Helm.** Formation and persistence of dry grassland diversity: role of human history and landscape structure. Tartu, 2007. 89 p.
127. **Leho Tedersoo.** Ectomycorrhizal fungi: diversity and community structure in Estonia, Seychelles and Australia. Tartu, 2007. 233 p.
128. **Marko Mägi.** The habitat-related variation of reproductive performance of great tits in a deciduous-coniferous forest mosaic: looking for causes and consequences. Tartu, 2007. 135 p.
129. **Valeria Lulla.** Replication strategies and applications of Semliki Forest virus. Tartu, 2007. 109 p.
130. **Ülle Reier**. Estonian threatened vascular plant species: causes of rarity and conservation. Tartu, 2007. 79 p.
131. **Inga Jüriado**. Diversity of lichen species in Estonia: influence of regional and local factors. Tartu, 2007. 171 p.
132. **Tatjana Krama.** Mobbing behaviour in birds: costs and reciprocity based cooperation. Tartu, 2007.
133. **Signe Saumaa.** The role of DNA mismatch repair and oxidative DNA damage defense systems in avoidance of stationary phase mutations in *Pseudomonas putida.* Tartu, 2007. 172 p.
134. **Reedik Mägi**. The linkage disequilibrium and the selection of genetic markers for association studies in european populations. Tartu, 2007. 96 p.
135. **Priit Kilgas.** Blood parameters as indicators of physiological condition and skeletal development in great tits (*Parus major*): natural variation and application in the reproductive ecology of birds. Tartu, 2007. 129 p.
136. **Anu Albert**. The role of water salinity in structuring eastern Baltic coastal fish communities. Tartu, 2007. 95 p.
137. **Kärt Padari.** Protein transduction mechanisms of transportans. Tartu, 2008. 128 p.
138. **Siiri-Lii Sandre.** Selective forces on larval colouration in a moth. Tartu, 2008. 125 p.
139. **Ülle Jõgar.** Conservation and restoration of semi-natural floodplain meadows and their rare plant species. Tartu, 2008. 99 p.
140. **Lauri Laanisto.** Macroecological approach in vegetation science: generality of ecological relationships at the global scale. Tartu, 2008. 133 p.
141. **Reidar Andreson**. Methods and software for predicting PCR failure rate in large genomes. Tartu, 2008. 105 p.

142. **Birgot Paavel.** Bio-optical properties of turbid lakes. Tartu, 2008. 175 p.
143. **Kaire Torn.** Distribution and ecology of charophytes in the Baltic Sea. Tartu, 2008, 98 p.
144. **Vladimir Vimberg.** Peptide mediated macrolide resistance. Tartu, 2008, 190 p.
145. **Daima Örd.** Studies on the stress-inducible pseudokinase TRB3, a novel inhibitor of transcription factor ATF4. Tartu, 2008, 108 p.
146. **Lauri Saag.** Taxonomic and ecologic problems in the genus *Lepraria* (*Stereocaulaceae*, lichenised *Ascomycota*). Tartu, 2008, 175 p.
147. **Ulvi Karu.** Antioxidant protection, carotenoids and coccidians in greenfinches – assessment of the costs of immune activation and mechanisms of parasite resistance in a passerine with carotenoid-based ornaments. Tartu, 2008, 124 p.
148. **Jaanus Remm.** Tree-cavities in forests: density, characteristics and occupancy by animals. Tartu, 2008, 128 p.
149. **Epp Moks.** Tapeworm parasites *Echinococcus multilocularis* and *E. granulosus* in Estonia: phylogenetic relationships and occurrence in wild carnivores and ungulates. Tartu, 2008, 82 p.
150. **Eve Eensalu.** Acclimation of stomatal structure and function in tree canopy: effect of light and CO_2 concentration. Tartu, 2008, 108 p.
151. **Janne Pullat**. Design, functionlization and application of an *in situ* synthesized oligonucleotide microarray. Tartu, 2008, 108 p.
152. **Marta Putrinš.** Responses of *Pseudomonas putida* to phenol-induced metabolic and stress signals. Tartu, 2008, 142 p.
153. **Marina Semtšenko.** Plant root behaviour: responses to neighbours and physical obstructions. Tartu, 2008, 106 p.
154. **Marge Starast.** Influence of cultivation techniques on productivity and fruit quality of some *Vaccinium* and *Rubus* taxa. Tartu, 2008, 154 p.
155. **Age Tats.** Sequence motifs influencing the efficiency of translation. Tartu, 2009, 104 p.
156. **Radi Tegova.** The role of specialized DNA polymerases in mutagenesis in *Pseudomonas putida.* Tartu, 2009, 124 p.
157. **Tsipe Aavik.** Plant species richness, composition and functional trait pattern in agricultural landscapes – the role of land use intensity and landscape structure. Tartu, 2008, 112 p.
158. **Kaja Kiiver.** Semliki forest virus based vectors and cell lines for studying the replication and interactions of alphaviruses and hepaciviruses. Tartu, 2009, 104 p.
159. **Meelis Kadaja.** Papillomavirus Replication Machinery Induces Genomic Instability in its Host Cell. Tartu, 2009, 126 p.
160. **Pille Hallast.** Human and chimpanzee Luteinizing hormone/Chorionic Gonadotropin beta (*LHB/CGB*) gene clusters: diversity and divergence of young duplicated genes. Tartu, 2009, 168 p.

161. **Ain Vellak.** Spatial and temporal aspects of plant species conservation. Tartu, 2009, 86 p.
162. **Triinu Remmel.** Body size evolution in insects with different colouration strategies: the role of predation risk. Tartu, 2009, 168 p.
163. **Jaana Salujõe.** Zooplankton as the indicator of ecological quality and fish predation in lake ecosystems. Tartu, 2009, 129 p.
164. **Ele Vahtmäe.** Mapping benthic habitat with remote sensing in optically complex coastal environments. Tartu, 2009, 109 p.
165. **Liisa Metsamaa.** Model-based assessment to improve the use of remote sensing in recognition and quantitative mapping of cyanobacteria. Tartu, 2009, 114 p.
166. **Pille Säälik.** The role of endocytosis in the protein transduction by cell-penetrating peptides. Tartu, 2009, 155 p.
167. **Lauri Peil.** Ribosome assembly factors in *Escherichia coli.* Tartu, 2009, 147 p.
168. **Lea Hallik.** Generality and specificity in light harvesting, carbon gain capacity and shade tolerance among plant functional groups. Tartu, 2009, 99 p.
169. **Mariliis Tark.** Mutagenic potential of DNA damage repair and tolerance mechanisms under starvation stress. Tartu, 2009, 191 p.
170. **Riinu Rannap.** Impacts of habitat loss and restoration on amphibian populations. Tartu, 2009, 117 p.
171. **Maarja Adojaan.** Molecular variation of HIV-1 and the use of this knowledge in vaccine development. Tartu, 2009, 95 p.
172. **Signe Altmäe.** Genomics and transcriptomics of human induced ovarian folliculogenesis. Tartu, 2010, 179 p.
173. **Triin Suvi.** Mycorrhizal fungi of native and introduced trees in the Seychelles Islands. Tartu, 2010, 107 p.
174. **Velda Lauringson.** Role of suspension feeding in a brackish-water coastal sea. Tartu, 2010, 123 p.
175. **Eero Talts.** Photosynthetic cyclic electron transport – measurement and variably proton-coupled mechanism. Tartu, 2010, 121 p.
176. **Mari Nelis.** Genetic structure of the Estonian population and genetic distance from other populations of European descent. Tartu, 2010, 97 p.
177. **Kaarel Krjutškov.** Arrayed Primer Extension-2 as a multiplex PCR-based method for nucleic acid variation analysis: method and applications. Tartu, 2010, 129 p.
178. **Egle Kõster.** Morphological and genetical variation within species complexes: *Anthyllis vulneraria* s. l. and *Alchemilla vulgaris* (coll.). Tartu, 2010, 101 p.
179. **Erki Õunap.** Systematic studies on the subfamily Sterrhinae (Lepidoptera: Geometridae). Tartu, 2010, 111 p.
180. **Merike Jõesaar.** Diversity of key catabolic genes at degradation of phenol and *p*-cresol in pseudomonads. Tartu, 2010, 125 p.

181. **Kristjan Herkül.** Effects of physical disturbance and habitat-modifying species on sediment properties and benthic communities in the northern Baltic Sea. Tartu, 2010, 123 p.
182. **Arto Pulk.** Studies on bacterial ribosomes by chemical modification approaches. Tartu, 2010, 161 p.
183. **Maria Põllupüü.** Ecological relations of cladocerans in a brackish-water ecosystem. Tartu, 2010, 126 p.
184. **Toomas Silla.** Study of the segregation mechanism of the Bovine Papillomavirus Type 1. Tartu, 2010, 188 p.
185. **Gyaneshwer Chaubey.** The demographic history of India: A perspective based on genetic evidence. Tartu, 2010, 184 p.