Genetics of the anticoagulant drug warfarin

Yen-Yu Chen

This dissertation is submitted for the degree of Doctor of Philosophy

St Edmund's College

University of Cambridge

30th November 2007

PREFACE

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text

The dissertation does not exceed the page limit of 300 specified by the Biology Degree Committee

ABSTRACT

The path towards personalised medicine requires understanding how the genetic makeup of each individual patient impacts on drug safety and efficacy. In this thesis I use the most widely prescribed anticoagulant drug, warfarin, as a model to investigate the effect of genetic determinants on drug efficacy and safety. Problematic clinical features of using warfarin include a narrow therapeutic range of PT INR 2-3, inter-individual dose variation of 20 folds and severe bleeding complication in 2% of patients.

Following a literature review of all the genes involved in warfarin pharmacokinetics and pharmacodynamics, 35 candidate genes were selected for investigation. Two independent Swedish cohorts of warfarin-treated patients were analysed. First linkage disequilibrium maps were constructed for each gene. Selected SNPs integrated with putative functional variants were genotyped in 201 patients recruited at the Uppsala University. A panel of 216 haplotype tag SNPs was then derived to analyse an independent cohort of 1496 patients from the prospective Warfarin Genetic study in Sweden (WARG).

The two studies were analysed separately for genetic association to warfarin dose requirement (single marker and haplotypic tests). Common SNPs in the vitamin K epoxide reductase gene (*VKORC1*) are significantly associated with dose in the Uppsala and WARG studies ($p = 1.9x10^{-15}$ and $6.5x10^{-100}$, respectively). Cytochrome P450 2C9 (*CYP2C9*) has been known to affect dose requirement and was confirmed in both Swedish cohorts ($p = 2.3x10^{-5}$ and $4.9x10^{-32}$). The two genes together explain ~40% of warfarin dose variation. SNPs in microsomal epoxide hydrolase (*EPHX1*) and orosomucoid 1 (*ORM1*) genes do not show a broad effect but are associated with dose in both studies. Genes encoding PROC, APOE, CALU, PDIA2 and GGCX showed nominal association with dose in the Uppsala study. Likewise, *PROS1, CYP1A1, CYP3A4, PDIA5, PDIA3* and *F10* showed nominal association to dose in the WARG study. Most of these minor effects, if real, are most likely to be population/treatment specific. A model taking in to account genetic factors (*VKORC1* and *CYP2C9**2 / *3) and non genetic factors (age, gender and drug interaction) together explained more than 50% inter-individual dose variance.

We analysed 64 patients from the Uppsala and WARG studies with recorded severe bleeding episodes using the same 216 common SNPs. Case-control analysis found SNPs in *PDIA4*, *P4HB* and *NR1I3* to be associated ($p \leq 0.01$) with bleeding. Using a recessive model, patients with a gastrointestinal bleeding sub-phenotype in the WARG cohort showed association with common variants in *PDIA6* ($P = 0.0014$, odds ratio = 6.98). We sequenced the exons of 11 of the candidate genes in 36 bleeders and 12 non-bleeders (Uppsala study). However, no high penetrance mutation was discovered.

To my dear parents

Mr Chi-Hwa Chen

Mrs Huei-Wan Liu-Chen

ACKNOWLEDGEMENTS

From the four years of my PhD there are a lot of people I want to mention and thank for giving me help and advice. The first and most important person is Panos Deloukas, my fantastic supervisor who gave me a wonderful thesis project that has let me witness one of the most important pharmacogenomics finding in the history. In addition, Panos and his wife, Christina Hedberg, have also given me guidance and encouragement throughout this study. Without their substantial support, I wouldn't have been able to complete this thesis.

This project is a multi-centre collaboration. I like to thank Mia Wadelius and Niclas Eriksson at the Uppsala University and Anders Rane at the Karolinska Institute in Sweden for the great support. Much appreciation goes to Mia for being a wonderful pen pal during the four years. Big thanks to Rhian Gwilliam for experimental and cultural advice, as well as proofreading my whole thesis. Many thanks to Ralph McGinnis and Fumihiko Takeuchi for statistics tutorials; to Jilur Ghori and Sarah Hunt for bioinformatic assistance; to Suzannah Bumpstead and Kate Downes for genotyping instruction ; to Mark Earthrowl, Richard Norris, Kirsten McLay and Alison Coffey for exon re-sequencing experiments; to Chris Tyler-Smith for being on my thesis committee and Ian Dunham for being my second supervisor. Finally I would like to thank everyone in Team 67 and at the Sanger Institute for all of the help you kindly gave me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A	Adenine
aa	amino acid
ABCB1	P-glycoprotein gene or MDR1 gene
aCGH	Array comparative genomic hybridisation, or array CGH
ADR	Adverse Drug Reaction
ALAT	Alanine aminotransferase
APOC2	Apolipoprotein C-II
APOE	Apolipoprotein E gene
BLAST	Basic Local Alignment Search Tool
bp	base pair(s)
BW	Bodyweight
C	Cytosine
CALU	Calumenin gene
cDNA	complementary DNA
CEPH	Centre d'Etude du Polymorphisme Humain
CEU	Caucasian of European origin
CI	Confident interval
CNV	Copy number variation
CYP	Cytochrome P450
CYP1A1	Cytochrome P450 1A1 gene
CYP1A2	Cytochrome P450 1A2 gene
CYP2C	Cytochrome P450 2C family
CYP2C18	Cytochrome P450 2C18 gene
CYP2C19	Cytochrome P450 2C19 gene
CYP2C8	Cytochrome P450 2C8 gene
CYP2C9	Cytochrome P450 2C9 gene
CYP3A4	Cytochrome P450 3A4 gene
CYP3A5	Cytochrome P450 3A5 gene
dbGaP	Database of Genotype and Phenotype
dbSNP	database of SNPs
DDW	Double distilled water
DMEM	Dulbeco's modified Eagle's medium
DNA	DeoxyriboNucleic Acid
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DTI	Direct thrombin inhibitor
DVT	Deep vein thrombosis

EBI	European Bioinformatics Institute
ECR	Evolutional conserver region
EDTA	EthyleneDiamineTetraAcetic acid
ENCODE	the ENCyclopedia Of DNA Elements
EPHX1	Epoxide hydrolase 1, microsomal gene
ER	Endoplasmic reticulum
EST	Expressed Sequence Tag
EU	European Union
F10	Coagulation factor X gene
F2	Coagulation factor II gene or prothrombin gene
F5	Coagulation factor V gene
F7	Coagulation factor VII gene
F9	Coagulation factor IX gene
FBS	Fetal bovine serum
FII	Coagulation factor II or prothrombin
FIIa	Coagulation factor II activated or thrombin
FIXa	Coagulation factor IX activated
FV	Coagulation factor V
FVII	Coagulation factor VII
FVIIa	Coagulation factor VII activated
FX	Coagulation factor X
FXa	Coagulation factor X activated
G	Guanine
G6PD	Glucose-6-phosphate dehydrogenase
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GAS6	Growth-arrest specific 6
GGCX	Gamma-glutamyl carboxylase gene
GWAS	Genome-wide association study
Hapmap	International Hapmap Project
HDL	High Density Lipoproteins
HGP	Human Genome Project
HMWK	High molecular weight kininogen
HUGO	Human Genome Organization
HWE	Hardy-Weinberg Equilibrium
Kb	Kilo base pairs
LD	Linkage disequilibrium
MAF	Minor Allele Frequency
Mb	Mega base pairs
MDR1	Multidrug resistance protein 1
MHC	Major histocompatibility complex

mRNA	messenger RNA
NAD(P)H	Nicotine adenine dinucleotide phosphate dehydrogenase
NAT	Arylamine N-acetyltransferase
NCBI	National Center for Biotechnology Information
ncSNP	non-coding SNP
NIH	National Institute of Health
NIH	National Institutes of Health
NQO1	NAD(P)H dehydrogenase, quinone 1 gene
NR1I2	Pregnane X receptor gene
NR1I3	Constitutive androstane receptor
NSAID	Non-steroidal anti-inflammatory drugs
nsSNP	Non-synonymous SNP
OMIM	Online Mendelian Inheritance In Man
OR	Odds ratio
ORM1	Orosomucoid 1 gene or Alpha-1-acid glycoprotein 1 gene
ORM2	Orosomucoid 2 gene or Alpha-1-acid glycoprotein 2 gene
P4HB	Prolyl 4- hydroxylase subunit beta
PCR	Polymerase Chain Reaction
PCR	Polymerase Chain Reaction
PDI	Protein disulfide isomerase
PDIA2	Protein disulfide isomerase family A, member 2
PDIA3	Protein disulfide isomerase family A, member 3
PDIA4	Protein disulfide isomerase family A, member 4
PDIA5	Protein disulfide isomerase family A, member 5
PDIA6	Protein disulfide isomerase family A, member 6
PGx	Pharmacogenetics and pharmacogenomics
PIVKA-II	Proteins induced by vitamin K antagonism
PROC	Protein C gene
PROS1	Protein S gene
PROZ	Protein Z gene
PT INR	Prothrombin time international normalised ratio
PXR	Pregnane X receptor
QC	Quality check or quality control
RFLP	Restriction Fragment Length Polymorphism
RNA	RiboNucleic Acid
RNA	Ribonucleic acid
SAEC	Severe Adverse Event Consortium
SAP	Shrimp alkaline phosphatase
SERPINC1	Anti-thrombin III gene
SJS	Stevens-Johnson Syndrome

SNP	Single nucleotide polymorphism
SNP	Single Nucleotide Polymorphism
STR	Short tandem repeats
T	Thymine
TF	Tissue factor
TFPI	Tissue factor pathway inhibitor
TSC	The SNP Consortium
UGT1A1	UDP glucuronosyltransferase 1 family, polypeptide A1
US FDA	United States Food and Drug Administration
UTG	UDP-glucuronosyltransferase
UTR	Untranslated region
VKD	Vitamin K dependent
VKOR	Vitamin K epoxide reductase
VKORC1	Vitamin K epoxide reductase complex subunit 1 gene
VNTR	Variable Number Tandem Repeat
vWF	von Willebrand factor
WARG	Swedish Warfarin Genetics study
WHO	World Health Organisation
WTCCC	Wellcome Trust Case Control Consortium

PUBLICATIONS ARISING FROM THIS WORK

Wadelius M, **Chen LY**, Eriksson N, Bumpstead S, Ghori J, Wadelius C, Bentley D, McGinnis R, Deloukas P. (2007). Association of warfarin dose with genes involved in its action and metabolism. *Hum Genet.* 121(1):23-34.

Chen LY, Eriksson N, Gwilliam R, Bentley D, Deloukas P, Wadelius M. (2005). Gamma-glutamyl carboxylase (GGCX) microsatellite and warfarin dosing. *Blood*. 106(10):3673-4.

Wadelius M, **Chen LY**, Downes K, Ghori J, Hunt S, Eriksson N, Wallerman O, Melhus H, Wadelius C, Bentley D, Deloukas P. (2005). Common VKORC1 and GGCX polymorphisms associated with warfarin dose. *Pharmacogenomics J.* 5(4):262-70.

Chen LY, Wadelius M, Lindh J, Eriksson N, Ghori J, Bumpstead S, Holm L, McGinnis R, Rane A, Deloukas P. The largest prospective warfarin study supports genetic forecasting. (submitted)

CHAPTER I

INTRODUCTION

1.1 PHARMACOGENETICS AND PHAMARCOGENOMICS

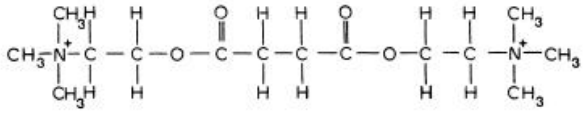
Each human being carries its own unique genetic makeup. In concert with the environmental factors to which he / she is exposed, this may result in differing reactions to external treatments including drugs.

1.1.1 Introduction

The first case of a genetic factor influencing drug response was reported back in the 1950s with the case of metabolising the muscle relaxant, suxamethonium chloride (Figure 1.1A), by serum cholinesterase (Figure 1.1B) (Lehmann and Ryan 1956). In this example, the recessive allele results in a serum cholinesterase peptide which has less efficient enzymatic activity and thus prolongs the drug's effect which in turn causes slower recovery from surgical paralysis. Soon after, the term pharmacogenetics was introduced by Vogel in 1959 (Vogel 1959).

Pharmacogenetics was initially used to refer to the study of inherited differences in genes responsible for drug metabolism i.e. the pharmacokinetics. The term comes from the words pharmacology and genetics. With the decoding of the human genome sequence and the subsequent systematic study of human genome variations, the term pharmacogenomics (pharmacology and genomics) has emerged to refer to a broader study of many different genes which together may determine drug behaviour (covers both pharmacokinetics and pharmacodynamics).

A



B

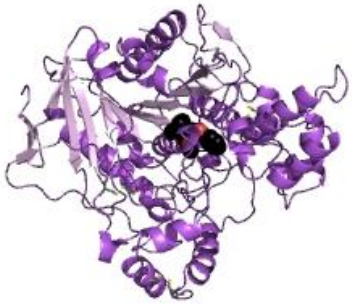


Figure 1.1. Structures for (A) Suxamethonium; (B) Human cholinesterase. Cartoon diagram of figure 1.1B is modified from Protein Data Bank (PDB) 1P0I.

More precisely, pharmacogenetics is the study of differential responses in patients to drug compounds based on their genetic polymorphisms, and it mainly involves the study of patient sample collections (Norbert and Roses 2003). Pharmacogenomics is to identify disease-relevant drug targets at the molecular level and to target drugs to clinical populations with specified genotypes/haplotypes, and it often entails using large scale and high-throughput technologies to identify genes for tractable or screenable targets that are not yet known to be genetically related to diseases (Norbert and Roses 2003). Throughout this dissertation, I will collectively refer to pharmacogenetics and pharmacogenomics as PGx.

1.1.2 Benefit of PGx

It is believed that the way people respond to a drug, i.e. drug efficacy, is a complex trait which is influenced as a whole by many different genes as well as environmental and life

style factors (age, diet, smoking, and state of health). However, understanding the genetic background is a key step in our endeavour to tailor drugs for subgroups of the general population - personalised medicine – with the ultimate goal of offering safe and efficacious therapies. Among the potential benefits coming from PGx research one can highlight the following:

1. *Right drug and right dose.*

   A variety of medicines may be used to treat the same symptoms, given the fact that such drugs may toggle different targets which reside in the same biochemical pathway. However, responses to a given drug may vary due to differences in the genetic makeup of each individual. Knowing the genetic profile of a patient will help clinicians to make the right decision as to which is the most appropriate medicine to administer, replacing current practices of trial-and-error in matching patients with the right drugs and dose. This is particularly useful in cancer therapies (Constable, Johnson et al. 2006; Fujita and Sasaki 2007). Currently, for most drugs, dose is determined empirically by taking into account the patient's age and bodyweight. With a better understanding for PGx of each drug, i.e. how well it acts on the target molecule and how the body metabolises it, a genetic screen combined with environmental factors can provide a quantitative measure for selecting a safe drug and determining the initial dose. This is significantly helpful to shorten the period of adjusting the dose to target the therapeutic range.

2. *Adverse Drug Reaction (ADR) and health care*

   Adverse Drug Reactions *(*ADR) account for 5-7% of hospitalisations in the United States

and in Europe (Ingelman-Sundberg and Rodriguez-Antona 2005). Due to their distinctive genetic makeup, some people will react differently to specific drugs triggering an ADR. If a genetic screen is available before a drug known to cause ADRs is prescribed, patients at risk will be identified and treated accordingly. The elimination of potential ADRs not only benefits the patients but saves tremendous resources for the health care society. PGx information, e.g. dose response and disease susceptibility, will guide the application of the most appropriate therapy and help monitor treatment closely in early stages. It may also allow doctors to recommend to patients lifestyle changes optimal for long term treatments.

3. Powerful medicine and drug discovery

Currently pharmaceutical companies are developing drugs using a "one size fits all" system without much discrimination between patients' ethnic origin and / or disease states (Roses 2004; The Wellcome Trust Case Control Consortium 2007). Large scale genetic studies of common / complex diseases are identifying the molecular basis of such conditions (Roses 2004; The Wellcome Trust Case Control Consortium 2007) which is providing pharmaceutical companies with new targets for developing new drugs. Numerous drugs under development failed in different phases of clinical trials or were removed from the market soon after they are launched due to ADR, e.g. rofecoxib (section 1.3) and Ximelagatran (section 1.7.3). A retrospective pharmacogenomic approach may give a second chance to such failed drugs, for example, identifying utility in subgroups of patients. Meanwhile, genetic screening will help to identify patients who react positively, negatively, or neutrally to this drug in the clinical trial (Roses 2004). This will significantly reduce the time and cost for drug development.

1.1.3 Challenges of PGx practice

Although pharmacogenomics has a great potential to deliver many benefits to our society, it is a very new concept and is still in the early stages. A myriad of challenges with clinical, ethical, social and legal implications have been identified and will need to be resolved before full scale application of PGx in practice (Norbert and Roses 2003). The first challenge is to educate both the health care professionals and the public as to the importance of PGx. Nowadays people are aware of the issues that may arise from third parties obtaining access to their genetic makeup information (work discrimination, eligibility to health insurance). Carelessness in managing personal genetic information, especially disease susceptibility, may cause serious ethical issues. Access to such information needs to be strictly restricted and only be used with the consent of the individual.

Pharmacogenomics may also narrow the use of certain drugs, which means a decrease in market share for this drug. In addition, improvement of drug safety and efficacy by further understanding the genetics of a drug is expensive and time-consuming. Therefore the pharmaceutical industry may be reluctant to invest money and time for maximising the efficacy of a ‘safe’ drug for a small portion of the population.

All these issues can be resolved by education and by law, and the public can be convinced by successful and compelling examples. Currently, drug safety and disease gene screening are pioneering this development.

1.1.4 Personalised medicine

The ultimate goal of drug efficacy and safety studies, as well as disease gene screening is to provide medication tailored to the individual. A tailor-made medication will be determined by a patient's genotype, gene expression, and clinical information to achieve optimal therapy. Initial efforts have been put into understanding of the drug efficacy by collaboration between academic institutions and the pharmaceutical industry.

1.2 DRUG METABOLISM AND ADVERSE DRUG RESPONSE (ADR)

1.2.1 Drug metabolism

Understanding drug metabolism, a complex and sophisticated series of reactions to discharge exotic drugs, is essential to PGx. The substances that result from metabolism, i.e. metabolites, may be inactive, or they may be similar to or different from the original drug in therapeutic activity and / or toxicity. Some drugs, called prodrugs, are administered in an inactive form, which has to be metabolised to give rise to the active form; the resulting metabolites produce the expected therapeutic effects (Stella and Nti-Addae 2007). Metabolites may be processed further instead of being excreted from the body and be excreted thereafter.

Most drugs must pass through the liver, where they are metabolised into more readily excreted polar products. In some situations, hepatic enzymes convert prodrugs to active forms or inactivate drugs (Stella and Nti-Addae 2007). The group of cytochrome P450 (CYP) enzymes is responsible for the liver's primary mechanism for chemically altering drugs, and the reaction rate for drug metabolism is determined by the levels of the CYP enzymes (Schuster and Bernhardt 2007). Because of limitations in their metabolising capacity, the CYPs may become overloaded when a drug is abundant in the blood (Schuster and Bernhardt 2007).

Since these hepatic enzyme systems are only partially developed at birth, infants have decreased drug metabolising abilities with all these enzymes being present but with a decreased activity (Lucier, Lui et al. 1979). This system is fully developed in young adults, but enzymatic activity decreases again with the increase of age. Therefore, age is often taken

into account for evaluating effective dose from potential adverse events (Rawlins, James et al. 1987).

Drug metabolism is a two-step chemical alteration/modification process occurring in the liver. The first step (phase 1) is to increase the polarity of metabolites by oxidation, reduction and hydrolysis. This will help to excrete the metabolites in the urine, for example, aspirin (acetylsalicylate) will be transformed to salicylic acid; the second step (phase 2) is to increase the water solubility of metabolites by conjugating with glucuronic acid, sulphate, and glutathione.

Several liver enzymes cooperate to transform drugs from lipophilic to hydrophilic in order to be eliminated from the body. These enzymes are often located in lipophilic membranes of the endoplasmic reticulum (ER) and are listed in Table 1.1.

Table 1.1 Hepatic enzymes involved in drug metabolism.

Steps	Function	Metabolising enzymes
Phase 1	Oxidation	Cytochrome P450 monooxygenase system
		Flavin-containing monooxygenase system
		Alcohol dehydrogenase and aldehyde dehydrogenase
		Monoamine oxidase
		Co-oxidation by peroxidase
	Reduction	NADPH-cytochrome P450 reductase
		Reduced (ferrous) cytochrome P450
	Hydrolysis	Esterase and amidase
		Epoxide hydrolase
Phase 2	Solubility	Glutathione S-transferase
		UDP-Glucuronyltransferase
		N-Acetyltransferase
		Amino acid N-acyl transferase
		Sulfotransferase

1.2.2 ADR caused by drug metabolising genes

1. N-acetyltransferase

Arylamine N-acetyltransferase (NAT) is a phase II enzyme involved in the detoxification of aromatic and heterocyclic amines and hydroxylamine, arylhydrazines and arylhydrazides in the liver (Westwood, Kawamura et al. 2006). About half of the population of white North Americans carry an autosomal recessive allele for N-acetyltransferase (Evans and White 1964; Drayer and Reidenberg 1977), a liver enzyme that metabolises drugs such as isoniazid, phenelzine, hydralazine and salicylazosulfapyridine. Such people are poor metabolisers, slow acetylators, having inefficient N-acetyltransferase (without wild-type NAT2*4 allele). Drugs such as isoniazid (Hughes, Biehl et al. 1954; Devadatta, Gangadharam et al. 1960), which are metabolised by this enzyme, will remain in the body longer and reach a higher blood level in slow acetylators.

2. UDP-glucuronosyltransferases

UDP-glucuronosyltransferases (UGTs) are also involved in phase II biotransformation and catalyse the transfer of the glucuronyl group from 5'-disphosphoglucuronic acid to endogenous molecules and exogenous substrates; producing less toxic and more easily excreted molecules (Peterson, Bigler et al. 2005). UGT1A1 is the major UGT enzyme responsible for glucuronidation of bilirubin which is an endogenous antioxidant hypothesised to modulate susceptibility to oxidative damage and cancer (Grant and Bell 2000). Approximately 10% of North Americans carry a sequence variant *(UTG1A1*28)* in the *UGT1A1* gene giving rise to a peptide with reduced enzymatic activity to metabolise the drug

irinotecan, which is used to treat colorectal cancer (Nagar and Blanchard 2006). This will also cause high blood levels of the drug and a higher risk of adverse effects. Overdose of irinotecan in poor metabolisers will predominantly cause severe neutropenia (Hasegawa, Ando et al. 2006; Nagar and Blanchard 2006).

3. Glucose-6-phosphate dehydrogenase

Although it is not directly involved in metabolising drugs, glucose-6-phosphate dehydrogenase (G6PD) is an enzyme that protects red blood cells from certain toxic chemicals. About 10% of male Africans and 17.4% of a Chinese subpopulation (Dai) have a deficiency of G6PD (Tishkoff, Varkonyi et al. 2001; Jiang, Yu et al. 2006). Drugs, e.g. primaquine, which is used to treat malaria, destroy red blood cells and cause haemolytic anaemia. However, this is more prevalent in males and less in females because of its location on the X chromosome.

4. Cytochrome P450 2D6

Cytochrome P450 2D6 (CYP2D6) is a hepatic monooxygenase and is the only non-inducible drug metabolising CYP (Ingelman-Sundberg, Sim et al. 2007). CYP2D6 metabolises approximately 25% of all drugs on the market (Evans and Relling 1999; Eichelbaum, Ingelman-Sundberg et al. 2006). Currently, there are 67 functional variants described (http://www.cypalleles.ki.se/cyp2d6.htm) which result in a number of different phenotypes including poor metaboliser (two absence of activity alleles), extensive metaboliser (two normally functioning alleles) and ultrarapid metaboliser (duplicated/multiple normally functioning alleles) (Ingelman-Sundberg, Sim et al. 2007). The ultra-rapid metaboliser

phenotype has been associated with adverse drug reactions, mainly as a result of an increase of 10-to-30-fold in amounts of metabolites (Ingelman-Sundberg, Sim et al. 2007). This can result in lack of response to certain antidepressants (Kawanishi, Lundgren et al. 2004; Rau, Wohlleben et al. 2004). In the case of the prodrug codeine, the ultrarapid metabolisers show decreased levels of several drugs, e.g. morphine (Kirchheiner, Schmidt et al. 2007).

In addition to the three hepatic enzymes mentioned here, a few CYP enzymes have also been reported with decreased enzymatic activities that are associated with common polymorphisms (MAF: 0.01-0.68) in some populations (Pirmohamed and Park 2003).

1.3 INTERNATIONAL EFFORT ON DRUG SAFETY

Most common ADRs are the result of cellular toxicity and increased susceptibility to other diseases. This has been the major reason for the US Food and Drug Administration (FDA) withdrawing drugs from the market. One of the well-known paradigms is the COX-2 inhibitor, rofecoxib, manufactured by Merck which was withdrawn from the market due to this drug being linked to an increased risk of cardiovascular events; including myocardial infarction and stroke (Bresalier, Sandler et al. 2005; Martinez-Gonzalez and Badimon 2007).

Although drug regulators consider the risk / benefit ratio in medicine approvals, ADRs limit the use of many otherwise effective drugs, e.g. rofecoxib, and cause severe burden for the health care system. International initiatives have recently been established in the United States and Europe aiming to unravel the genetic causes of ADRs.

1.3.1 ADR and cost paid by society

In the past ten years, the field of PGx has attracted a lot of interest with its great potential for improving drug safety with tests against ADRs and in general delivering more efficient and personalised medication. In 1994, ADRs accounted for more than 2.2 million serious cases and caused over 100,000 deaths and was one of the major causes of hospitalisation and death in the United States (Lazarou, Pomeranz et al. 1998). A more recent estimate from 2005 indicated that ADRs account for 5-7% of hospitalisation in the United States and in Europe (Ingelman-Sundberg and Rodriguez-Antona 2005). This causes a huge burden to the health care system. In 2000, the US healthcare system spent around $177 billion on treating drug-related mortality and morbidity; which is 10 percent of its total healthcare. In the UK, a

conservative estimate was made in 2004 that admissions related to ADRs cost the NHS up to £466m annually (Pirmohamed, James et al. 2004).

1.3.2 Efforts coordinated in USA

In 2007 with $6 million from pharmaceutical companies, a new alliance, namely Severe Adverse Event Consortium (SAEC), was formed and sets its goal in identifying and qualifying biomarkers for adverse events; including pharmacogenetic markers. Their aim is to identify and validate genetic biomarkers, mainly DNA variants, which might predict ADR. Although there are some pharmaceutical companies still doubting the effectiveness of this consortium, sample collection is well underway around the world. In the initial phase, two projects have been funded by SAEC to understand the drug-related liver toxicity and a rare but very serious drug-related skin problem called Stevens-Johnson Syndrome (SJS). The drug-related liver toxicity project is to study hepatotoxicity as a cause of acute liver failure and involves Eudragene and Diligen (section 1.3.3). Meanwhile, the SJS patient DNAs will be collected from participating companies and academic institutions, and will be compared with control individuals.

Another initiative, namely Drug-Induced Liver Injury Network (DILIN), was funded in 2004 by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) of the U.S. National Institutes of Health, to develop standardized definitions and instruments to identify and fully characterize cases of drug-induced liver injury (http://dilin.dcri.duke.edu/). Two studies are being conducted by DILIN. The retrospective study is to establish a registry of patients who had liver injury due to taking certain prescribed drugs since 1994. The prospective study aims to monitor the patients who were subject to adverse liver events after

taking any drug and/or herbal medicine and investigate the pathogenesis of this liver ADR.

1.3.3 Efforts coordinated in Europe

Eudragene (http://www.eudragene.org/) was a project funded by the European Union (EU) to establish a case-control DNA collection for studying the genetic basis of ADRs. Eudragene is also a member of the SAEC. Six important ADRs which have been identified to cause serious illness were selected for investigation in the SAEC project (Molokhia and McKeigue 2006). The study involves 15 centres in 11 countries and aims to collect at least 500 cases of each ADR alongside with an equal number of healthy volunteers as controls. The co-ordinating centre manages the database and makes samples freely available to academic and industry-based researchers throughout Europe.

Another effort focusing on liver toxicity, namely Diligen, is funded by the UK Department of Health and aims to develop a simple test to identify patients at high risk of developing drug-induced liver disease (http://www.diligen.org/). Blood samples from patients who had liver injury relating to co-amoxiclav and flucloxacillin, and anti-tuberculosis drugs will be analysed by researchers led by Professor Ann Daly in Newcastle, and researchers in Liverpool, Nottingham, and London.

1.3.4 Important issues

Although the ADR attracts big efforts in drug safety and health care system, several issues remain to be fully addressed:

(a) characterisation of ADR cases and controls;

(b) required contact of an optimal genotyping panel;

(c) robust computational methods to perform whole-genome analysis;

(d) publicly available knowledge base to share the information;

(e) intellectual property relating useful markers in predicting ADRs.

1.4 HUMAN GENOME

The completion of a reference sequence of the human genome in 2004 (International Human Genome Sequencing Consortium 2004) provides the foundation for various biomedical applications, e.g. genome dynamics and evolution (Jobling and Tyler-Smith 2003; Feuk, Carson et al. 2006), epigenomics (Jirtle and Skinner 2007; Spivakov and Fisher 2007), and human diseases (Chen, Cooper et al. 2007; Frayling 2007). However the decoding of the human genome has not been the end of the quest to understand human biology. A number of projects have since been proposed and have come to fruition, e.g. the Hapmap project (2003), the ENCyclopedia Of DNA Elements (ENCODE) project (2004). These studies promote and foster the development of other –omics sciences.

1.4.1 The material of inheritance

In a broad sense, the decoding of the human genome could be traced back to the 19th century when Charles Darwin published his work 'On the Origin of Species' in 1859. Soon after, a Czech monk, Gregor Medel, described the basic laws of inheritance in his work 'Experiments in Plant Hybridisation'. More importantly, Mendel's laws initiated the idea that the phenotype (the external appearance of the offspring) is determined by the genotype (some hidden genetic factors). In 1869, Johann Friedrich Miescher extracted a new substance in the nucleus of human pus, which were in fact white blood cells, and he called the substance nuclein and postulated that it is present in all cells and must be concerned with heredity. Miescher's nuclein later came to be called nucleic acid. During 1920s, Phoebus Levene was intensively studying the chemical structure of nucleic acid and proposed the basic components were comprised of a phosphate group, a sugar, and one of the four bases, adenine (A), cytosine (C),

thymine (T), and guanine (G). The terms deoxyribonucleic acid and ribonucleic acid (DNA and RNA) then came into common use. However, at this time there was still no firm evidence that DNA and RNA were, in fact, the molecular basis of heredity.

The chemical structure of the most important genetic element, DNA, was unveiled by the joint effort of James Watson, Francis Crick, Maurice Wilkins, and Rosalind Franklin in 1953. They recognised how the two pairs of complementary bases (adenine-thymine and guanine-cytosine) would have identical shapes when held together by hydrogen bonds. Two long chains of such base pairs would likely form a double helix—roughly, the shape of an enormously long, winding, doubled-railed staircase. The DNA molecule, comprised of long strands of such base pairs in specific and varied sequences, could embed genetic information that, if the strands were separated, could be copied. In 1960s, Marshall Nirenberg cracked the genetic code and discovered the first 'triplet' which is a sequence of three bases of DNA encoding one of the twenty amino acids and elucidated the operation in protein synthesis.

1.4.2 The Human Genome Project

The DNA sequencing technique developed in 1977 by Frederick Sanger paved the way towards sequencing the human genome alongside Leroy Hood's development of an automated sequencer in 1986. At that time, scientists in the United States and elsewhere were debating about the high cost versus usefulness of the information to be derived from sequencing the entire human genome, alongside the construction of genetic and physical maps of the genome sequencing efforts began in US, Japan, France, Italy, the United Kingdom, and Canada. In 1988, Human Genome Organization (HUGO) was privately-funded to coordinate this international effort. The Human Genome Project (HGP) started

soon after in October 1990, and the first physical map of 30,000 human genes was published in 1998 (Deloukas et al, 1998). Initially, the HGP received $3 billion from the US Department of Energy and the National Institutes of Health (NIH). An international consortium emerged with research groups in the United Kingdom, China, France, Germany, and Japan to do the Human Genome sequencing. The first version of draft sequence was announced, on 26th June 2000, jointly by the US president Bill Clinton and British Prime Minister Tony Blair. The team at the Sanger Institute contributed over one third of the finished genome sequence (International Human Genome Sequencing Lander, Linton et al. 2001; Consortium 2004). The sequence of the longest chromosome, chromosome 1, was reported in 2006 (Gregory, Barlow et al. 2006).

1.4.3 Human genetic variation

Analyses of human genome sequences revealed an unexpected scenario, that is, the sequence similarity between any two individual is 99.9%. The 0.1% difference, which accounts for 3 million bases, includes various polymorphisms which are the result of human evolution (International Human Genome Sequencing Consortium 2004).

Although it is a small fraction, however, this 0.1% is important because it is the molecular basis of the various phenotypic differences among individuals including disease risk and variable drug response (Sachidanandam, Weissman et al. 2001; International HapMap Consortium 2005; Frazer, Ballinger et al. 2007). The impact of these genetic variations will be described in the next section.

1.5 HUMAN SEQUENCE VARIATION

David Botstein initiated the use of restriction fragment length polymorphisms (RFLPs) in 1978, to indicate genetic differences among individuals and map genes (Lander and Botstein 1986). Since then various genetic markers have been used for linkage mapping of inheritable diseases. e.g. variable number of tandem repeat (VNTR), short tandem repeats (STRs), microsatellites. However, the above types of genetic markers are unevenly distributed in the genome, and these polymorphisms vary in size (length in base pairs) and mutation rate (Figure 1.2).

Polymorphisms interspersed in the genome have been successfully used to identify the molecular basis of many Mendelian disorders including rare forms of polygenic conditions such as Parkinson's disease (Polymeropoulos, Higgins et al. 1996), and Alzheimer disease (Delabar, Lamour et al. 1986). In the human genome, the most interspersed types of sequence variations are:

(1) variable number of tandem repeat (VNTR, section 1.4.1)

(2) copy number variation (CNV, section 1.4.2)

(3) single nucleotide polymorphism (SNP, section 1.4.3).

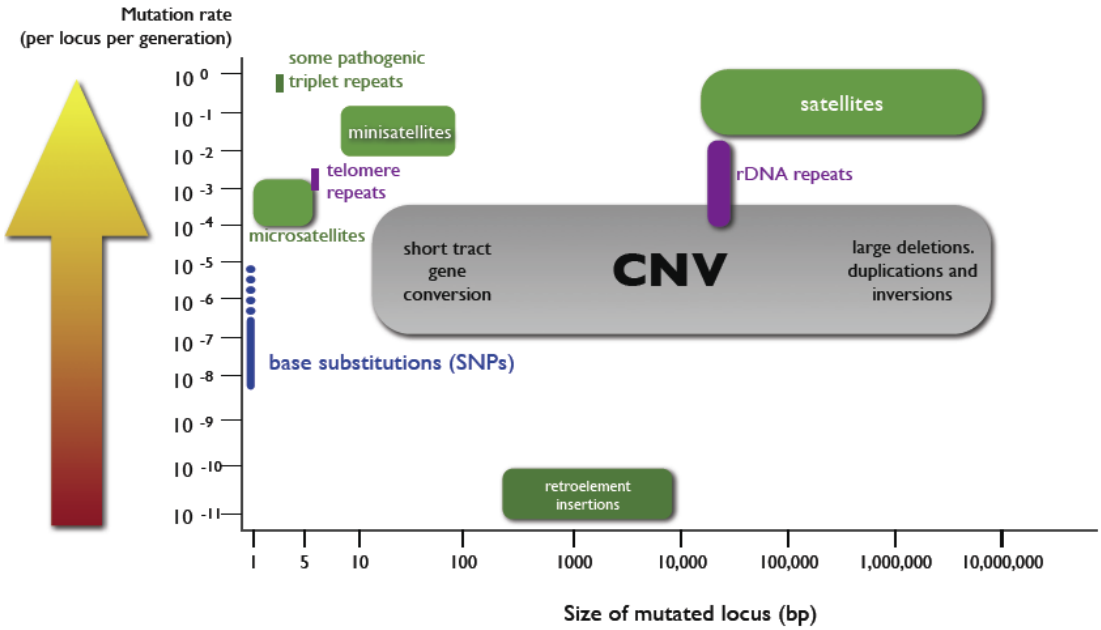


Figure 1.2. The mutated size and rate of genetic polymorphisms. Various sequence variations have been identified during the use of DNA sequencing technology. Smaller satellites (minisatellites and microsatellites) have been extensive applied in genetic disease association and DNA profiling, such as forensic science. Now SNP genotyping replaces satellite typing and is widely used for understanding linkage disequilibrium structure and disease association. CNV variation appears in a diverse length and mutation rate manner. However, a robust technology to look at CNV is imminent. This figure was modified from http://www.sanger.ac.uk/Teams/Team29/.

1.5.1 Variable number of tandem repeat (VNTR)

VNTR polymorphism is one of the non-coding DNA used for early genetic association studies (Permutt and Elbein 1990). VNTR polymorphisms vary in length of repeated units. 'Microsatellites' are the short two to four base pair repeats; those of intermediate length are called 'minisatellites'; or 'midisatellites' and 'macrosatellites' are the larger repeats.

The DNA sequence of minisatellites is hyper-variable being highly polymorphic in size with more than 1000 copies of STRs (Jeffreys, Wilson et al. 1987). In addition, minisatellites DNAs have been also reported to be associated with homologous recombination hotspots in

humans (Wahls, Wallace et al. 1990) and are used for DNA fingerprinting and profiling (Jeffreys, Turner et al. 1991). Compared to minisatellites which are usually detected by RFLP typing, microsatellites based on a much shorter repeat unit of 2, 3, and 4 bases are repeated between 10 and 100 units. They have the advantages that they can be typed by polymerase chain reaction (PCR) and the number of repeated units can be discriminated precisely. There is considerable variation in germline mutation rate at microsatellite loci ranging from undetectable to $8x10^{-3}$ (Mahtani and Willard 1993) and appearing every one thousand base pairs.

1.5.2 Copy number variation (CNV)

Copy number variation (CNV) in DNA sequences appears as a consequence of insertions, deletions, duplications and complex multi-site variants (Fredman, White et al. 2004) and could be functionally significant.

Copy number variation was first discovered in bacteria in 1976 (Lovett, Duvall et al. 1976). The first human CNV was found in 1983 in zeta-globin gene complex by Goodbourn and his colleagues (Goodbourn, Higgs et al. 1983). A first generation of genome-wide CNV map using array comparative genomic hybridisation (array CGH, or aCGH) technology was then published in 2006, conducted mainly by the groups at the Wellcome Trust Sanger Institute, with a total of 1447 CNV regions found on 270 individuals from four populations (the Hapmap collections, section 1.5.1) (Redon, Ishikawa et al. 2006).

With its application in detecting chromosomal rearrangement (Emanuel and Saitta 2007), CNV is now used in clinical diagnostics (Rodriguez-Revenga, Mila et al. 2007), especially

detecting the genomic imbalances in cancer tissues (Michels, De Preter et al. 2007) and cytogenetic diagnosis of constitutional disorders (Lee, Iafrate et al. 2007) and sporadic diseases (Lupski 2007). The association of CNV and X-link mental retardation has demonstrated its value in association studies. However, the resolution is a current issue in fine-mapping the disease susceptibility genes.

1.5.3 Single nucleotide polymorphism (SNP)

SNPs make up around 90% of all human genetic variation with a density of one per 100 to 300 bases (Frazer, Ballinger et al. 2007). It was first reported in 1991 by Ligtenberg and his colleagues as influencing the differential splicing of episialin mRNA (Ligtenberg, Gennissen et al. 1991). This alternative splicing of exon 2 was determined by an A / G SNP which is 8 bases downstream of second splicing acceptor site. It has been estimated that there are circa 10 million common SNPs in the human population (Reich, Gabriel et al. 2003), minor allele frequency ≥ 1%. (11,883,685 RefSNP clusters on dbSNP build 128). In the first instance, a SNP was defined as a bi-allelic polymorphism with a minor allele frequency (MAF) above 1%. However, more and more SNPs are found to be tri-allelic.

The SNP consortium (TSC) (http://snp.cshl.org/) was established in 1999 as a collaboration of several companies and institutions to produce a public resource of SNPs with an initial goal to identify 300 thousand SNPs in two years. In 2001, a map of 1.4 million validated SNPs was published and released to the public domain (Sachidanandam, Weissman et al. 2001).

SNPs have emerged as the most powerful genetic marker in the last decade; due to its

abundance in the human genome (Frazer, Ballinger et al. 2007). The dense distribution of SNP markers further fostered the international collaboration of Hapmap project (2005; Frazer, Ballinger et al. 2007) and genome-wide association studies (GWAS) (Consortium 2007; Plenge, Cotsapas et al. 2007; Thomson, Barton et al. 2007).

Most SNPs are not functional: Few are functional and of those, some are causative of disease. These SNPs usually cause silencing/malfunction in gene expression (Knight, Udalova et al. 1999; Prokunina, Castillejo-Lopez et al. 2002; Tokuhiro, Yamada et al. 2003). SNPs in coding regions can be classified either as synonymous that do not lead to an amino acid change or non-synonymous that result in an amino acid substitution. Non-synonymous changes (nsSNP) can alter the amino acid composition and structural conformation of protein folding, and therefore they are potentially more important.

A few studies (Knight, Udalova et al. 1999; Prokunina, Castillejo-Lopez et al. 2002; Tokuhiro, Yamada et al. 2003) have shown non-coding SNPs (ncSNPs) are also very important, especially those in regulatory regions, such as promoters. These SNPs may change the expression of that particular gene which leads to unbalanced biochemical reactions, for example, drug metabolism (Ingelman-Sundberg, Sim et al. 2007). A genome wide study further suggests that SNPs in the regions that control activity of genes are more likely to be related to common, complex disease, rather than in the regions that specify the protein code, i.e. nsSNP (Stranger, Nica et al. 2007).

1.6 LINKAGE DISEQUILIBRIUM & ASSOCIATION STUDY

The uses of genetic variation in identifying the molecular basis of common complex traits is based on the property of linkage disequilibrium (LD). LD is defined as a non-random association of two or more loci when the genomic recombination occurred at different rate. Therefore, a mutation causing a common (disease) phenotype will be in LD with other nearby common variants.

1.6.1 The HapMap Project

In 2002, an international collaborative effort started to construct a whole genome LD map (or haplotype map) in four human populations with the use of very dense SNP markers. This project gathered scientists and funding from different countries, including Canada, China, Japan, Nigeria, the United Kingdom and the United States, to create a public resource of finding genes in response to diseases and drugs (http://www.hapmap.org).

A total of 270 healthy individuals from four populations of African, European, Chinese and Japanese origin were recruited, referred to as 'the Hapmap collections'. The phase I of the Hapmap project released an LD map with the use of 1 million SNPs in 2005 (2005). A second generation map harbouring more than 3.1 million SNPs was published in October 2007 providing fine-resolution human genome LD information (Frazer, Ballinger et al. 2007).

With such dense SNP maps a few analyses have commenced to globally investigate disease susceptibility. One example is the analysis of these 3 million polymorphisms in HapMap which identified 300 candidate regions to be subject to positive natural selection (Sabeti,

Varilly et al. 2007). Further analyses revealed among these regions genes related to Lassa virus infection, skin pigmentation and development of hair follicles (Sabeti, Varilly et al. 2007).

1.6.2 Haplotype Tag SNPs

One of the early observations that led to the HapMap project was that regions of high LD in humans show a block-like structure with low diversity of common haplotypes (Gabriel, Schaffner et al. 2002). Haplotype blocks, defined typically as regions of high LD with all marker pairs having high D' or r^2, can harbour tens or hundreds of SNPs which can be represented by a small number of SNPs that capture most of the common genetic information in this interval, namely a tag SNP (Patil, Berno et al. 2001; Stram 2004). Therefore, the human haplotype map of 3 million SNPs can be interrogated with a subset of 300,000-1,000,000 markers. African populations require much denser sets of tag SNPs compared to Caucasians because of the more rapid decay of LD in these populations which is the result of their longer evolutionary history.

In a genetic association study, prior to the availability of Hapmap, a small subset of individuals had to be genotyped first with a large number of SNPs; to validate tag SNPs that could then be tested across the study population. The international Hapmap project provides a wealth of information for tag SNPs in four ancestral populations (Frazer, Ballinger et al. 2007) and facilitates the genome-wide association studies.

1.6.3 Association Studies

With the availability of the human genome sequence, comprehensive maps of sequence variation (e.g. HapMap, Perlegen) and recent advances in genotyping and high-throughput sequencing technology, it has become feasible to systematically search disease causing genes genome-wide, without previous understanding of pathogenesis of the diseases (Roses 2004; Peacock and Whiteley 2005). Following the success of understanding the molecular basis of rare, monogenic diseases the challenge has now been to identify genes underlying common diseases (Newman, Hoffjan et al. 2004).

One of the largest projects in common disease genetics is being carried out by the Wellcome Trust Case Control Consortium (WTCCC) which is searching the causative variants for 8 diseases including tuberculosis, coronary artery disease, type 1 diabetes, type 2 diabetes, rheumatoid arthritis, Crohn's disease, bipolar disorder and hypertension. Nine million pounds (18 million US dollars) has been spent to genotype two thousand patients for each disease and a common three thousand control samples (a total of 19 thousand patients) (Consortium 2007).

An international central repository database, called the Database of Genotype and Phenotype (dbGaP, http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap) has been developed by the National Center for Biotechnology Information (NCBI) for the research community.

1.7 COAGULATION AND ANTICOAGULANTS

Anticoagulants are widely used for surgery and various thromboses to stop the amplification of coagulation reaction (haemostasis). Haemostasis consists of two steps: the primary step is the activation of platelets (a.k.a. thrombocytes) which bind to collagen in the injured places (Clemetson and Clemetson 2007), and the secondary step is the coagulation which involves the activation of a number of coagulation factors (Riddel, Aouizerat et al. 2007).

1.7.1 Coagulation

Collagens are bound by circulating platelets by their collagen-specific glycoprotein Ia / IIa receptor on the cell surface. This adhesion is strengthened further by a large multimeric circulating protein von Willebrand factor (vWF), which forms links between the platelet glycoprotein Ib / IX / V and collagen fibrils (Cauwenberghs, Vanhoorelbeke et al. 2000).

The platelets are then activated releasing the content of their granules into the plasma to activate other platelets. The platelets undergo a subsequent change in their shape to expose the phospholipid surface which is required for the coagulation factors (Kornecki, Lenox et al. 1987). Fibrinogen then links the adjacent platelets by adhering surface glycoprotein IIb / IIIa, which is the most abundant adhesion receptor on the platelet cell surface, and is also used to prevent the activation of platelets (Cauwenberghs, Vanhoorelbeke et al. 2000).

The secondary haemostasis is initiated by two pathways; the contact activation pathway (or intrinsic pathway) and the tissue factor pathway (or extrinsic pathway) (Figure 1.3). Both pathways then join to the common pathway to cross-link fibrin clot to prevent further

bleeding (Riddel, Aouizerat et al. 2007).

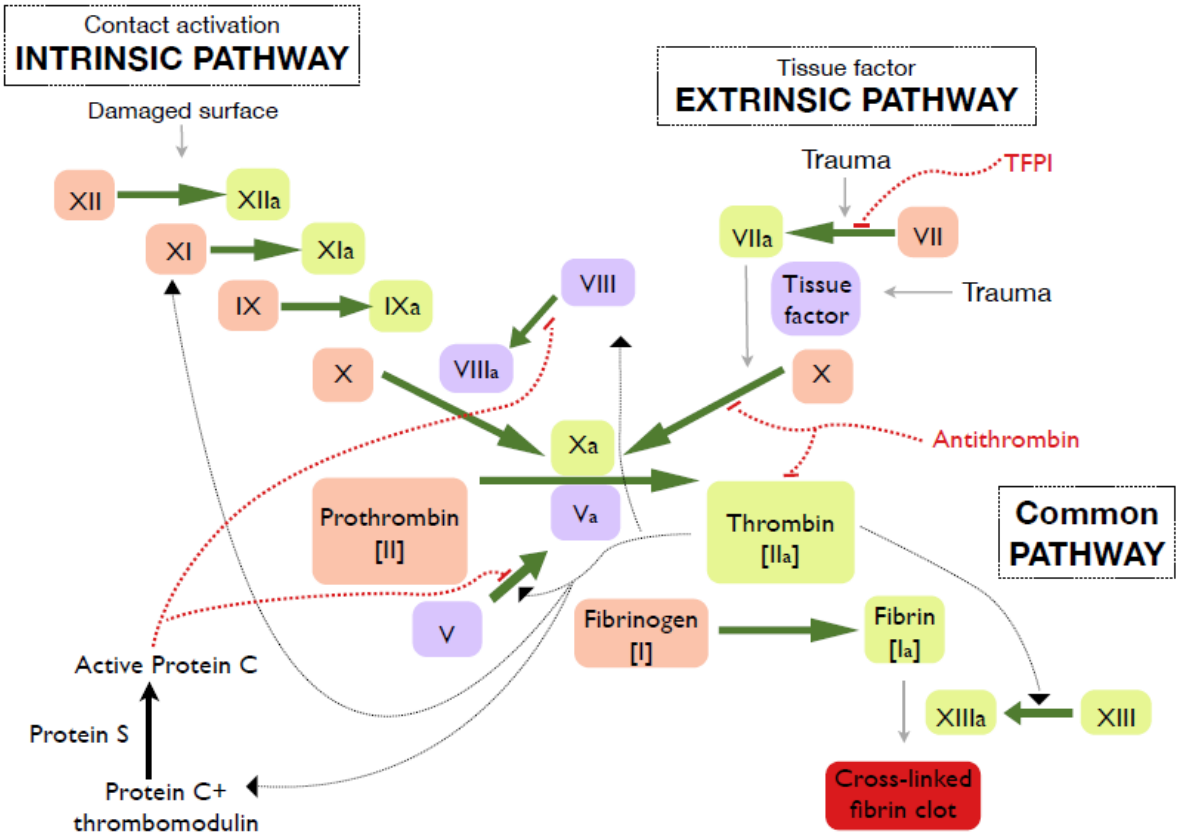


Figure 1.3. The coagulation cascade. The cascade is involved in the phase II coagulation including intrinsic, extrinsic and common pathways. All these coagulation factors exist in inactive form. Protein C (with help from Protein S), Antithrombin, and TFPI act as natural anticoagulant to stop fibrin clot formation (red dotted line). This figure is a modified version from http://en.wikipedia.org/wiki/Coagulation.

The primary (also more important) pathway for the initiation of blood coagulation is the tissue factor (TF) pathway that takes only three steps to activate thrombin. When factor VII is activated by tissue factor released from trauma, it is almost immediately inhibited by the tissue factor pathway inhibitor protein (TFPI). The cascade is triggered when the factor X has been activated by factor VIIa. When the prothombin (factor II) is activated (factor IIa), it feeds back to factor VIII and XI in contact activation pathway. Meanwhile, the thrombin

proteins also feed back to activate more platelets to stop further bleeding.

Collagen is not only involved in platelet activation but in the contact activation pathway by associating with high molecular weight kininogen (HMWK), prekallikrein and factor XII (Hageman factor) to form the primary complex on the damaged surface. However, this pathway has been shown less important than tissue factor pathway because patients with severe deficiencies of FXII, HMWK and prekallikrein do not demonstrate any bleeding disorders (Krijanovski, Proulle et al. 2003; Wynne Jones, Russell et al. 2004).

Factors V, VIII, XI, and XIII are activated by thrombin whilst other factors including prothrombin (II), VII, IX and X are activated by gamma-carboxylation. The gamma-carboxylation is regulated by gamma-glutamyl carboxylase (GGCX) and requires vitamin K as the cofactor (Suttie 1980). The oxidised form of vitamin K is then reduced by vitamin K epoxide reductase (VKOR) in the liver (Fasco and Principe 1980).

Vitamin K antagonists, such as acenocoumarol and warfarin, are used as anticoagulants by preventing the activation of prothrombin (II), VII, IX and X. Meanwhile proteins C and S are also activated through gamma-carboxylation act as natural anticoagulant and prevent over-coagulation (Walker 1981). These coagulation factors which require gamma-carboxylation to become active forms are classified as vitamin K dependent (VKD) proteins.

1.7.2 Anticoagulant Drugs

There are many different anticoagulants currently used for medication by targeting molecules involved in the pathway. Besides vitamin K antagonists, the commonly prescribed drugs

target antithrombin III (encoded by *SERPINC1*), platelet aggregation, plasminogen formation, and inhibition of activated coagulation factors to prevent unexpected coagulation (Table 1.2).

The most commonly used anticoagulants are vitamin K antagonists. These drugs have been used for a long period of time whilst warfarin has a long history back to 1950s. Today more targeted anticoagulants are also available on the market. Except vitamin K antagonists, all other drugs target various proteins in the coagulation cascade (Table 1.2).

Table 1.2. Prescribed anticoagulants and antiplatelet drugs.

Classification	Biological function	Marketed name
Vitamin K antagonists	Antagonising the effects of vitamin K and prevent the activation of vitamin K dependant protein	Acenocoumarol; Clorindione; Coumatetralyl; Dicumarol (Dicoumarol); Diphenadione; Ethyl biscoumacetate; Phenprocoumon; Phenindione; Tioclomarol; Warfarin
Heparin group	Activating antithrombin III which block thrombin from clotting	Antithrombin III; Danaparoid; Heparin; Sulodexide; low molecular weight heparin (Bemiparin, Dalteparin, Enoxaparin, Nadroparin, Parnaparin, Reviparin, Tinzaparin)
Glycoprotein IIb/IIIa inhibitors	Preventing platelet aggregation and thrombus formation	Abciximab; Eptifibatide; Tirofiban
Other platelet aggregation inhibitors	Preventing platelet aggregation and thrombus formation	Acetylsalicylic acid/Aspirin; Aloxiprin; Ditazole; Carbasalate calcium; Cloricromen; Clopidogrel; Dipyridamole; Indobufen; Picotamide; Prasugrel; Ticlopidine; Triflusal; prostaglandin analogue
Plasminogen activators	Activating plasminogen	Alteplase/Reteplase/Tenecteplase, Streptokinase, Urokinase/Saruplase, Anistreplase
Serine endopeptidases	Inhibiting activated coagulation factors	Ancrod, Drotrecogin alfa/Protein C, Fibrinolysin
Direct thrombin inhibitors	Inhibiting thrombin	Argatroban; Bivalirudin; Dabigatran; Desirudin; Hirudin; Lepirudin; Melagatran; Ximelagatran
Other antithrombotics		Defibrotide; Dermatan sulfate; Fondaparinux; Rivaroxaban

1.7.3 Example of drug safety of an anticoagulant

As described in the previous section, cellular toxicity in liver is the primary concern of drug safety. In 2004, Ximelagatran (marketed as Exanta), an oral direct thrombin inhibitor (DTI), was denied approval by the US FDA and was removed from the market in February 2006 after a few reports of severe liver damage and heart attacks. Recently, a retrospective study conducted by the manufacturer, AstraZeneca, demonstrated that the elevated levels of serum alanine aminotransferase (ALAT) is associated with the Major Histocompatibility Complex (MHC) alleles DRB1 *07 and DQA1 *02 (Kindmark, Jawaid et al. 2007).

This provides an explanation of possible immune pathogenesis in long-term treated patients with this oral-administered direct thrombin inhibitor. The result based on DRB1 *07 would have been able to detect patients at risk of the adverse event with sensitivity 47% and specificity of 83% (Kindmark, Jawaid et al. 2007). In addition, this observation suggested that other factors may also contribute to susceptibility of the ADR induced by ximelagatran.

1.8 WARFARIN

The choice of a good drug is important in a study aiming to demonstrate a compelling result as proof of principle for tailor-made personalised medicine. A good / ideal test case must therefore meet the following requirements:

1. A drug whose use could be optimised for each individual.
2. Efficient and safe use of the drug could be determined by differences in the genetic makeup.

Taking into account these requirements, warfarin came out as the first choice from other good candidates because:

1. Narrow therapeutic range (PT INR between 2 and 3) is difficult to target among individuals along with various prescribed dose.
2. Two variants in *CYP2C9* (*CYP2C9**2 and *CYP2C9**3) had been reported to be associated with dose variation (Rettie, Wienkers et al. 1994; Sullivan-Klose, Ghanayem et al. 1996) prior to starting this Thesis.
3. It has known interactions with other drugs (MICROMEDEX, (http://www.micromedex.com/).
4. Serious bleeding episodes have caused large burden to the health care system (Pirmohamed, James et al. 2004).

1.8.1 Introduction

Warfarin (4-hydroxy-3-(3-oxo-1-phenylbutyl)-2H-chromen-2-one, Figure 1.4A) is a vitamin K antagonist that prevents blood coagulation. Before 1940s, cows in northern America had

been bleeding to death from very minor injuries. In 1921, Frank Schofield, a Canadian veterinarian, identified that these animals were eating mouldy sweet clover hay which functioned as a potent anticoagulant. In 1940, Karl Paul Link and his student, Harold Campbell, identified the fungal product effective as a vitamin K antagonist. They named a potent derivative, coumarin derivative 4-hydroxycoumarin, of this antagonist 'warfarin' after the Wisconsin Alumni Research Foundation for 'warf', to which he assigned the patent rights (Sadler 2004), and 'arin' indicating its link with coumarin.

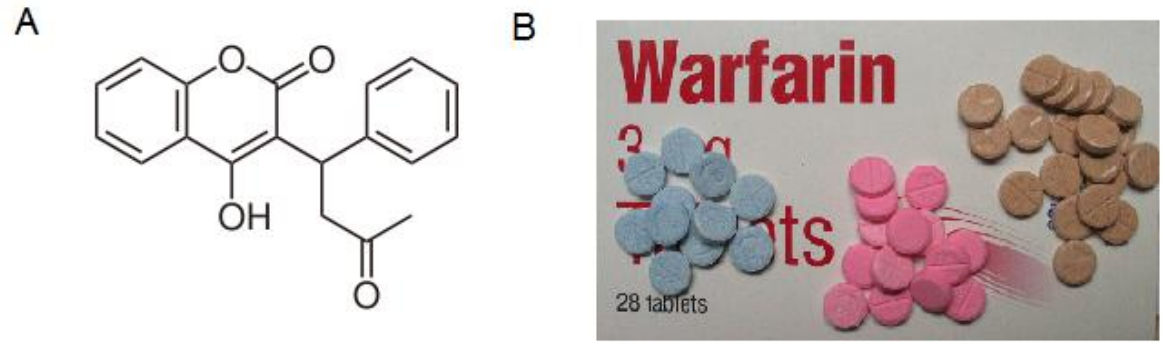


Figure 1.4. Warfarin (A) chemical structure; (B) tablets used in UK (Photo by Gonegonegone). Warfarin ($C_{19}H_{16}O_4$) have a molecular weight of 308.33 g/mol. The prescription warfarin contains mixtures of R- and S-4-hydroxy enantiomers. Different colours are used to distinguish different doses. 5mg (pink), 3mg (blue) and 1mg (brown).

In 1952, Warfarin was registered and used as a rodenticide in agriculture because it causes severe bleeding in large doses. However, after being using for many decades, rats have evolved and developed resistance to warfarin. Today poisons are much more potent and toxic and warfarin is no longer used as rat poison.

The medical application of warfarin started from 1951 when a naval soldier tried but failed to commit suicide by taking warfarin. It was in 1954 that warfarin was firstly approved for clinical use in humans. After 1955, when the US President Dwight Eisenhower was treated

with warfarin following a heart attack, warfarin became the most commonly prescribed anticoagulant.

1.8.2 Warfarin and coumarin derivatives

Nowadays warfarin is the most widely prescribed anticoagulant drug for atrial fibrillation, heart valve prosthesis, recurrent stroke, deep vein thrombosis, and pulmonary embolism (Loewen, Sunderji et al. 1998). It can help prevent the formation of blood clots and help reduce the risk of embolism. Warfarin is normally taken orally and is completely absorbed with more than 99% bound to serum albumin in the plasma. Free warfarin is taken up by the liver where it is metabolised by cytochrome P450 (CYP) enzymes. Different brand names are used in marketing warfarin; including Coumadin, Jantoven, Marevan, and Waran.

Despite the fact that more than 1.5 million patients are prescribed warfarin (Kessler 2006), there are other coumarin derivatives commonly used in some countries: acenocoumarol used in Italy and Spain; Phenprocoumon in Germany, Austria, Belgium, Brazil, Denmark, Switzerland and The Netherlands (Frazer, Ballinger et al. 2007). These anticoagulants share similar pharmacokinetics with warfarin due to structural similarity.

1.8.3 Side effects and adverse reactions

As a rodenticide which causes serious bleeding in rats, warfarin treatment can also cause severe bleeding episode in any organ or tissue. This adverse drug effect happens to less than 5% of patients (severe cases may even be fewer). A UK study (Pirmohamed, James et al. 2004) showed that among different ADRs which cause hospitalisation and the associated

burden on the NHS, warfarin was the second most common drug cause of such admissions (129/1225, 10.5%).

The phenotype of bleeding complications varies accordingly in different locations in the body:

1. bleeding around the brain can cause severe headache and paralysis
2. bleeding in the joints can cause joint pain and swelling
3. bleeding in the stomach or intestines can cause weakness, fainting spells, black tarry stools, vomiting of blood, or coffee ground material
4. bleeding in the kidneys can cause back pain and blood in urine.

Patients who take warfarin may also be subject to the following adverse effects.

- Nausea, vomiting, stomach pain
- Gas and bloating
- Hair loss
- Skin changes or discoloration anywhere in the body
- Purple toes or fingers
- Pain in stomach, back, or sides
- Low fever, loss of appetite, dark urine, jaundice
- Diarrhoea, fever, chills, body aches, flu symptoms
- Feeling weak or light-headed
- Sudden headache, confusion, problems with vision, speech, or balance
- Sudden leg or foot pain
- Sudden numbness or weakness, especially on one side of the body

1.8.4 Drug interaction and effective therapeutic dose

Many drugs, both prescription and non-prescription, can affect the anticoagulant action of warfarin and thereby make dosing complicated. Meanwhile, chemicals presented in considerable amount in daily diets may interact with warfarin and adjusted dose will be necessary (Wittkowsky 2007). These interactions will increase or decrease the therapeutic efficacy, and it has been reported that the required warfarin dose can vary up to 20 fold between patients. Due to high variability of dose, there are several kinds of warfarin tablet available for prescription from 1 mg to 10 mg per tablet (Figure 1.4B).

Interacting drugs which inhibit or potentiate the effect of warfarin cause unsatisfactory therapeutic effect or unexpected bleeding complication. Known examples of such medications include aspirin, paracetamol, alcohol, ibuprofen, cimetidine, oxandrolone, some vitamins, e.g. vitamin K and antibiotics. Some medications can enhance the action of warfarin and cause excessive blood thinning and life-threatening bleeding (Table 1.3).

Table 1.3. Interacting drugs with increased risk of severe bleeding in warfarin treatment.

Category	Therapeutic impact
NSAIDs	Potential for serious gastrointestinal bleeding
Sulfa Drugs	Increased effects of warfarin, with potential for bleeding
Macrolides	Increased effects of warfarin, with potential for bleeding

Non-steroidal anti-inflammatory drugs (NSAIDs) are drugs with analgesic, antipyretic and anti-inflammatory effects - they effectively reduce pain and fever caused by inflammation. NSAIDs increase gastric irritation and erosion of the protective lining of the stomach, resulting in gastrointestinal bleeding. Additionally, NSAIDs decrease the cohesive properties

of platelets and prevent platelets aggregation in blood clot formation (Mustard and Packham 1975).

Sulfa drugs (also called sulphonamides) include several antibiotics and are commonly used to treat pneumocystis jiroveci pneumonia, urinary tract infections, shigellosis, and certain protozoan infections. Sulfa drug has been known for causing various ADRs e.g. Stevens-Johnson syndrome (Mockenhaupt and Schopf 1996) and haemolysis in G6PD deficient (Cohen, Rosenthal et al. 1968). Currently, the mechanism for interaction between warfarin and sulfa drugs is unknown; however, clinicians hypothesize that the warfarin activity is prolonged due to a decreased production of vitamin K by intestinal flora, as a consequence of systemic antibiotic administration.

The macrolides are also a typical group of commonly used antibiotics including azithromycin, clarithromycin, dirithromycin and roxithromycin. Macrolides can interact adversely with commonly used drugs, usually by altering metabolism due to complex formation and inhibition of cytochrome P450 3A4 (CYP3A4) in the liver and in enterocytes (von Rosensteil and Adam 1995). With the concurrent medication of macrolides, the activity of warfarin may also be prolonged due to alterations in the intestinal flora and its production of vitamin K for clotting factor production.

1.9 THIS THESIS

The drug interactions of warfarin and its biochemical pathway have been studied for many decades with molecular biology and biochemistry approaches and in animal models, such as laboratory rat strains which are resistant to warfarin (Berkner and Pudota 1998; Kohn and Pelz 2000; Wallin, Sane et al. 2002). Narrow therapeutic range, dose variation and associated adverse events make warfarin administration difficult.

When this project started in 2003, none of the genome-wide SNP genotyping arrays were available and the only approach to study the genetic effect of warfarin was to comprehensively interrogate the candidate genes. Published studies with regard to warfarin were reviewed to understand the systematic pharmacokinetics and pharmacodynamics (chapter III, see also (Wadelius and Pirmohamed 2007)).

A set of 35 candidate genes was selected on the basis of available functional information at the time to study the genetics of warfarin. Genotyping assays were designed using publicly available SNPs and then used to analyse two collections of warfarin-treated patients from Sweden. The first one comprises 201 patients collected by Dr Mia Wadelius at the Uppsala University hospital (Uppsala study). These samples were used to construct LD maps of each gene. The selection of candidate genes and their LD structure in the Swedish population is discussed in chapter III. Association analysis of the 35 genes with warfarin dose in the Uppsala study is discussed in chapter IV.

Tag SNPs were then selected according to the LD maps built using the samples of the Uppsala study. Tag SNPs and lead SNPs from the association analysis in the Uppsala study

were all tested in an independent sample of 1523 warfarin treated patients enrolled in the Swedish warfarin genetics project (WARG study) (http://www.druggene.org/). The tag SNP selection and association with warfarin dose in the prospective WARG study is discussed in chapter V.

The Uppsala and WARG studies comprise a total of 64 patients with recorded severe bleeding episodes. They were genotyped with the same set of SNPs used for studying dose requirement in the WARG study. A case-control analysis was undertaken for the bleeding phenotype using the 1679 non-bleeders as controls. The bleeding patients from Uppsala and WARG were recruited with different criteria, therefore, each panel was also analysed separately. The association of genotype and bleeding complication is presented in chapter VI.

Since severe bleeding is a complication affecting over 2% of patients in the two studies, exon re-sequencing was undertaken in eleven of the candidate genes including those involved in recycling of vitamin K (*GGCX, VKORC1, NQO1, EPHX1, P4HB, PDIA4*) and other genes of interest (*F5, APOE, CYP2C9, PROC, CALU*). Results on 48 warfarin-treated patients, including 36 bleeders and 12 non-bleeders along 48 CEPH Caucasians are discussed in chapter VI.

CHAPTER II

MATERIALS AND METHODS

2.1 PATIENTS

2.1.1 Uppsala study

In Uppsala University Hospital, 201 patients who were treated with warfarin at the anticoagulation clinic were recruited in 2000. For six consecutive visits in the clinics, five weekly warfarin doses and the corresponding prothrombin time international normalised ratio (PT INR) values were registered. The interval of each visit for each patient ranges between 1 week and 1 month, depending on their medication. The 5 weekly doses were averaged for subsequent analyses, and 166 patients were stable (defined as 3 consistent doses) at the time of recruitment.

The individual warfarin dose varied nearly 20 fold, ranging from 4.5 to 77.25 mg per week. Information such as age, gender, bodyweight (BW), other diseases and indication for treatment was retrieved from the patients' medical records although there are seven patients whose body weight details are unknown. Body weights were obtained from medical records but these values may have varied during the course of warfarin treatment as these were not recorded at each visit.

The patients were mostly Caucasians: 194 were of Swedish origin; four were of European descent and three were from the Middle East. These patients were 28-88 years old (median: 69) when their blood samples were collected, and they had been treated with warfarin for at least 2 months, ranging from 2.4 months to 26 years. These patients were grouped by their treatment indications: patients with heart valve prosthesis which aims to target a higher PT

INR value and patients treated for other indications which aims to a normal 2.0 - 3.0 PT INR value.

The concurrent medications were also recorded and checked according to the MICROMEDEX database (http://www.micromedex.com/). The drugs were categorised as interacting if they had been described to have moderate or major interactions with warfarin. In 201 warfarin patients, there are 107 concurrent medications being prescribed which were known to influence the warfarin therapy. For further analysis in concurrent medication, patients are classified into three groups: individuals requiring high warfarin doses due to interaction with other concurrent prescription; individuals requiring lower warfarin doses because the medications enhance the effect of warfarin and patients without any known interactions.

2.1.2 WARG study

A larger Swedish cohort was used for validating any finding in the Uppsala cohort. The patients were recruited through the Warfarin Genetics (WARG) project: a nation-wide, prospective case-control in warfarin treated patients involving 40 Swedish centres, and coordinated by the Karolinska Institute in Sweden. Patients aged less than 18 years old and those previously treated with warfarin were excluded from this study. Apart from these and established contraindications to warfarin treatment (e.g. pregnancy), this study has no other exclusion criteria. A total number of 1523 patients were recruited and 1496 samples were analyzed. These patients were 18-92 years old (median: 66) when they were first treated with warfarin, and the individual warfarin doses varied also nearly 20 fold, ranging from 6.00 to 113.65 mg per week.

In order to not to interfere with the local management of warfarin treatment, warfarin dose and PT INR measurement intervals are chosen at the discretion of the treating physician. The concurrent medication was recorded and classified according to its potential for pharmacologic interactions with warfarin, by use of the classification presented in the classifications in the Swedish Drug Index (Läkemedelsindustriföreningen 2007), the drug information databases Janusinfo (www.janusinfo.se), and Micromedex (www.thomsonhc.com). Dosages, INR-values, concomitant medication and complications are continuously recorded in the Internet-based medical record system.

2.1.3 Patients subject to severe bleeding

Twelve patients out of the 201 in the Uppsala study were subject to one or more severe bleeding episodes and an additional 24 bleeding patients were recruited from other anticoagulation clinics in Sweden by Dr Mia Wadelius. Severe bleeding is defined as bleeding causing hospitalisation for at least one night but excluding the following: bleeding induced by thrombolysis, post-surgical operation, trauma, and malignancy, and bleeding in patients aged >88 years old. Thus, a total of 36 DNA samples from bleeding patients were sent from Uppsala University. In the WARG study, 28 out of 1523 patients had severe bleeding. In total, 64 DNA samples were collected from warfarin bleeders.

2.1.4 DNA preparation

Patient sample DNA was extracted from whole blood in Uppsala University Hospital (for Uppsala study) and the Karolinska Intitute (for WARG study) in Sweden. The methods and kits for DNA extraction were different in the two institutions (details given below). The

purified DNAs were frozen and shipped to the Sanger Institute on dry ice. When the DNA samples arrived, the DNA was defrosted, centrifuged at 4000 rpm for 5 minutes and then stored at -20^{o}C.

2.1.4.1 Uppsala study

The blood sample was collected in ethylenediaminetetraacetic acid (EDTA)-anticoagulated tube and DNA was extracted using QIAamp DNA Blood Mini Kit (Qiagen). In a reaction, 200 μl EDTA-treated blood, 20 μl QIAgen protease and 200 μl Al buffer (lysis buffer) were mixed in a microcentrifuge tube and vortexed for 15 seconds. The tube was then incubated at 56^{o}C for 10 minutes, followed by a brief centrifugation to remove drops from the inside of the lid. 200 μl of 100% ethanol was then added and the tube vortexed for 15 seconds. The mix was then transferred to the QIAamp Spin Column and centrifuged at 6000g for 1 minute. 500 μl AW1 buffer (wash buffer) was added and centrifuged at 6000g for 1 minute. After centrifugation, 500 μl AW2 buffer (wash buffer) was added and centrifuged at 20000g for 3 minutes. The columns were then placed in Eppendorf tubes and 200 μl AE buffer (elution buffer) was added. The samples were incubated for 1-5 minutes at room temperature before being centrifuged at 6000g for 1 minute.

2.1.4.2 WARG study

DNA was prepared from 1 ml EDTA-treated whole blood using the MagnaPure LC method according to the manufacturer's instructions (MagnaPure DNA Isolation Kit-Large Volume; Roche Diagnostics, Mannheim, Germany).

2.2 GENETIC MARKER SELECTION

2.2.1 Single nucleotide polymorphism (SNP)

SNPs in the candidate genes were selected from Ensembl SNP database (http://www.ensembl.org/) with the aim of at least 5 kb spacing between them. Functional variants on candidate cytochrome P450 genes were tested in the Uppsala study according to the annotation in Allele nomenclature for Cytochrome P450 enzymes website (http://www.cypalleles.ki.se/).

Validated SNPs with minor allele frequency (MAF) greater than 5% were analysed to select tag SNPs using Tagger (section 2.6.2). These tag SNPs, together with SNPs found to be nominally significant in Uppsala study, were then tested in WARG study.

2.2.2 Microsatellite repeat marker

The microsatellite marker in intron 6 of the *GGCX* gene has been reported to be associated with warfarin dose among individuals (Shikata et al. 2004) and was selected for genotyping across the Uppsala cohort. This marker was previously reported to contain CAA repeats with 10, 11 and 13 copies (Shikata *et al*., 2004).

2.3 MASS SPECTROMETRY GENOTYPING

2.3.1 Assay design

Flanking sequences of selected SNPs were downloaded from the Ensembl database (http://www.ensembl.org). The sequences were filtered for repetitive sequences by RepeatMasker (http://repeatmasker.org), and assays were designed with MassARRAY Assay design v3.1 (Sequenom).

2.3.2 PCR amplication of SNP loci

2 μl purified DNA with a concentration of 2 ng/μl for MassEXTEND or 4 ng/μl for iPLEX were dispensed into 384-well microtiter plates using a MULTI-MEK 96 dispenser (Beckman). For genotyping quality control, each 384-well microtiter plate contained two internal duplication samples in 4 wells and water in 16 wells. Before setting up the PCR, the 384-well plate containing the DNA was briefly centrifuged.

2.3.2.1 MassEXTEND

In the MassEXTEND assay, the PCR step is variable depending upon the plexing levels. If each reaction was plexing lower than 8, i.e. detecting less than 8 SNPs in a single reaction, 1 μl PCR mix containing 0.04 μl BD Titanium Taq polymerase (BD BioSciences, Clontech), 0.75 μl PE PCR buffer, 0.2 μl 25m M dNTP and 0.01 μl deionised water was mixed with 2 μl 375 nM primer mix. The total 3 μl mix was then added into the 384-well microtiter plate, which was then centrifuged at 2000 rpm for 1 minute.

When the assay had 8-plex or more, 1 μl PCR mix containing 0.04 μl BD Titanium Taq polymerase, 0.5 μl BD Titanium PCR buffer, 0.2 μl 25mM dNTP and 0.26 μl deionised water was mixed with 2 μl 500 nM primer mix. The total 3 μl mix was then added into the 384-well microtiter plate, which was then centrifuged at 2000 rpm for 1 minute. The thermal cycling conditions were same for both low- and high-plex MassEXTEND assays. Initially the plates were heated at 95℃ for 1 minute, followed by 45 cycles of 95℃ for 20 seconds, 56℃ for 30 seconds, and 72℃ for 1 minute, and a final step at 72℃ for 3 minutes.

2.3.2.2 iPLEX

The composition of the iPLEX PCR reaction is different to that of the MassEXTEND. The PCR setup includes: 8 ng genomic DNA in 2 μl deionised water, 1μl 500 nM primer mix and 2 μl PCR mix (composed of 0.1 μl Qiagen Hotstar Taq polymerase, 0.1 μl 25mM dNTP; 0.325 μl 25mM $MgSO_4$, 0.625 μl Qiagen PCR buffer and 0.85 μl deionised water). The plate was then centrifuged at 2000 rpm for 1 minute. The thermal cycling condition is slightly different to the MassEXTEND assay because of different enzyme used. Initially the 384-well microtiter plates were heated at 94℃ for 15 minute, followed by 45 cycles of 94℃ for 20 seconds, 56℃ for 30 seconds, and 72℃ for 1 minute, and a final step at 72℃ for 3 minutes.

2.3.3 Shrimp alkaline phosphatase treatment

After PCR, shrimp alkaline phosphatase (SAP) treatment is performed to remove the 5'-phosphate group from DNA. The removal of these 5'-phosphate groups is to prevent the PCR products from ligating to each other and the excess dNTPs. 2 μl SAP mix was applied into each well of the 384-well microtiter plate. The plate was then spun at 2000 rpm for 1 minute,

followed by incubation in a thermal cycler at 37°C for 20 minutes, followed by 5 minutes at 85°C to inactivate the enzyme. When the reaction was complete, the plates were re-centrifuged at 2000 rpm for 1 minute. For MassEXTEND the 2 μl SAP mix includes 0.3 μl SAP, 0.2 μl 10x TS buffer and 1.5 μl deionised water whilst for iPLEX it includes 0.3 μl SAP, 0.17 μl 10x SAP buffer and 1.53 μl deionised water.

2.3.4 Oligo extension

Similar to the composition of MassEXTEND and iPLEX PCRs, the oligo extension reaction is different for high and low MassEXTEND, as well as the iPLEX.

2.3.4.1 MassEXTEND

For an oligo extension reaction lower than 8-plex, 2 μl extension mix was added to the 384-well PCR plate including 0.018 μl Thermosequenase, 0.2 μl 10X Thermosequenase buffer, 0.9 μl 0.5mM stop mix and 0.382 μl deionised water was mixed with 0.5 μl 10μM extension primer; for a reaction equal or higher than 8-plex, 1 μl extension mix was added to the 384-well PCR plate including 0.04 μl Thermosequenase, 0.2 μl 10X Thermosequenase buffer, 0.45 μl 1mM stop mix and 0.31 μl deionised water was mixed with 1 μl 9μM extension primer. The plates were then placed on a heated thermal cycler at 94℃ for 2 minutes, followed by 55 cycles of 94℃ for 5 seconds, 52℃ for 5 seconds, and 72℃ for 5 seconds, and a chill down step at 10℃ until the plate was removed from the machine. Importantly, the extension for high-plex assay needs 99 cycles of 94℃ for 5 seconds, 52℃ for 5 seconds, and 72℃ for 5 seconds, compared with 55 cycles for low-plex assay.

2.3.4.2 iPLEX

2 μl extension mix was added to the 384-well microtiter plate which included 0.041 μl iPLEX enzyme, 0.2 μl 10X iPLEX buffer, 0.2 μl iPLEX termination mix, 0.559 μl deionised water and 1 μl primer mix (5.5 μM for low mass extension primers and 1.1 μM for high mass ones). The plates were then placed on a heated thermal cycler at 94℃ for 30 seconds, followed by 45 cycles of 94℃ for 5 seconds, 52℃ for 5 seconds, and 80℃ for 5 seconds, and a final step at 72℃ for 3 minutes. The plate was then chilled down to 10℃ until removed from the machine.

2.3.5 Desalting

After oligo extension, the plate was briefly centrifuged and added with 16 μl HPLC grade water added to each well. Resin was then applied onto a 384-well dimpled plate and it was left at room temperature for 15 minutes to let the resin dry out. The dimpled resin plate was then inverted onto the 384-well microtiter plate and to allow transfer of the resin into each well. The plate was then sealed and left on the rotator for 15 minutes. The plate was subsequently centrifuged at 4000 rpm for 6 minutes before it was ready for spotting.

2.3.6 Sample spotting and analysis

Before spotting, it was important to make sure that the resin was well precipitated in the bottom of each well of the 384-well microtiter plate. The reaction products on the 384-well microtiter plates were transferred onto SpectroCHIP (Sequenom) with MassARRAY Nanodispenser (Sequenom). The SpectroCHIP was then analysed with MassARRAY

Compact Analyzer (Sequenom), and allele calling was performed using TyperAnalyzer v3.4 (Sequenom).

On each SpectroCHIP (i.e. 384 samples), the result from 8 replicated samples and 16 samples including only water were analysed for quality control. Each SNP assay with more than 4 calls from the 16 water samples will be disregarded, whilst data from any chip with more than 2 inconsistent replication results from the 8 replicated samples will be also disregarded.

2.4 OTHER GENOTYPING

2.4.1 Taqman genotyping

Detecting the CYP2C9*2 and CYP2C9*3 SNPs was performed by Taqman SNP Genotyping assay (Applied Biosystems). Pre-Developed Assay Reagents for Allelic Discrimination (TaqMan PDARs for AD) uses the 5' nuclease assay to genotype purified DNA samples and discriminate between the two alleles of a SNP.

For discriminating between the two alleles a 5 μl reaction was setup which contained 2.5 μl 2x Taqman Universal PCR Master Mix, 0.5 μl 10x Allelic Discrimination Mix and 8 ng genomic DNA (in 2 μl deionised water). The PCR reaction was carried out on the MJ thermal cycler including the first initial step at 50^{o}C for 2 minutes to optimise AmpErase UNG enzyme activity and 95^{o}C for 10 minutes to activate AmpliTaq Gold DNA polymerase, and the amplification conditions were 55 cycles for 92^{o}C for 15 seconds and 60^{o}C for 90 seconds.

The reaction plate was read using ABI PRISM 7900HT Sequence Detection System. The genotypes were analysed with SDS software (Applied Biosystem) and the two alleles were called manually in accordance with the allele clusters.

2.4.2 Microsatellite genotyping for GGCX

Genotyping was performed by PCR amplification using HEX-fluorescent labelled primers (Table 2.1). Forward and reverse primers designed to amply the repeated region were generated using Primer 3 (http://frodo.wi.mit.edu/).

Table 2.1. Oligo sequence for amplying microsatellite in intron 6 of GGCX.

Primer	*Sequence*
GGCX-Forward	ggatatgttagaaaacattgaacacc
GGCX-Reverse	gtggctgggtagatgcctaag

PCR reaction was performed in a 25 μl reaction mix containing 2.5 μl 10x NEB buffer, 2.5 μl 2 mM dNTP, 1.6 μl primer mix (5 μM), 0.125 μl Taq polymerase, 10 ng patient DNA and 17.3 μl distilled water. Three different thermal cycling conditions were tested (see Table 2.2). The PCR products (amplicons) were electrophoretically-separated on an ABI PRISM Genetic Analyzer (Applied Biosystems).

Table 2.2. Thermo cycling conditions of GGCX microsatellite PCR.

Condistion	Step 1	Step 2 (10 cycles)			Step 3 (30 cycles)		
1	94^{o}C	93^{o}C	65^{o}C	72^{o}C	93^{o}C	60^{o}C	72^{o}C
2	94^{o}C	93^{o}C	60^{o}C	72^{o}C	93^{o}C	55^{o}C	72^{o}C
3	94^{o}C	93^{o}C	55^{o}C	72^{o}C	93^{o}C	50^{o}C	72^{o}C
Duration	5 min	30 sec	50 sec	50 sec	30 sec	50 sec	50 sec

Allele calling was performed with the Genotyper Software v3.7 (Applied Biosystems). In the 201 Swedish warfarin-treated patients, six different genotypes were detected and the genotype and length of corresponding amplicon are listed in Table 2.3.

Table 2.3. GGCX microsatellite genotype and expected amplicon

Genotype	**Repeat number**	**Amplicon length (bp)**
1	16	436
2	15	433
3	14	430
4	13	427
5	12	424
6	11	421
7	10	418

2.5 EXON RESEQUENCING

2.5.1 Introduction

Genes which have been implicated to play a role in causing bleeding in warfarin patients were chosen for re-sequencing on the bleeders. Twenty-four bleeders from Uppsala cohort and the extra 12 bleeders, together with 12 control Swedish warfarin patients without bleeding episodes, were included in re-sequencing the candidate genes.

2.5.2 Primer design

The genes chosen for resequencing were firstly manually annotated and curated with the help from the Sanger Institute HAVANA group (http://www.sanger.ac.uk/HGP/havana/). Gene sequences, plus 1 kb flanks on both ends, were extracted from the internal sequence database and the primers amplifying each exons for each gene were automatically designed.

Without exception, 125 bp flanks on both ends of each exon were included in the region to be sequenced. Each amplicon (fragment amplified by PCR reaction) is around 500 to 550 bp to ensure perfect sequence read quality. For the exons which are too large to be amplified by 1 amplicon, multiple amplicons were adopted to toggle the exons and a minimal 100 bp overlap was forced. However, in some cases, the automatic amplicon design didn't meet the above criteria and the primers were designed manually.

Primer 3 (http://frodo.wi.mit.edu/) was used to select primers in manual process. Up to 10 amplicons were reported in each design attempt by the software. The first amplicon is

designed as described above whilst the next amplicon targets 100 bases from the 3' end of upstream amplicons. This tolerates 50 'messy' bases at the start of each read. If no amplicon could be designed, the target moves outward 50 bases and the search of primers would be attempted again - this procedure iterates until the whole exon is well included in amplification. Successfully designed primers were then compared to the golden path sequence with IPCRESS developed internally at the Sanger Institute. Any primers with 2 or more matches were removed and the amplicons were recycled.

2.5.3 PCR and sequencing

The PCR was performed using a 384-well PCR plate. 7.5 ng genomic DNA (in 7.5 μl deionised water) was used for each reaction. The 7.5 μl PCR mix (composed of 2.5 μl PCR buffer, 2.5 μl 1 mM dNTP, 0.15 μl HotStar Taq polymerase and 2.35 μl deionised water) was then added into each well.

The thermo cycling conditions included a initial heated step at 95^{o}C for 15 minutes, followed by 39 cycles of 95^{o}C for 30 seconds, 60^{o}C for 30 seconds, and 72^{o}C for 30 seconds and a final cycle of 95^{o}C for 30 seconds, 60^{o}C for 30 seconds and 72^{o}C for 10 minutes. The PCR products were then treated with Shrimp Alkaline Phosphatase at 37^{o}C for 30 minutes followed by 80^{o}C for 15 minutes to inactivate the enzyme.

Sequencing experiments were performed by dedicated sequencing team at the Sanger Institute. The reaction was performed using the same primers as the PCR amplification step in both forward and reverse directions to produce double-stranded sequence. The protocol was modified in-house and each 5 μl reaction contained 1 μl diluted PCR product, 2 μl

diluted primer (15 ng/μl) and 2 μl sequencing mix (0.12 μl BigDye (v3.1), 1.38 μl Sanger BigDye reaction buffer (v2), 0.46 μl double distilled water (DDW) and 0.04 μl dGTP BigDye v3.0). The plates were centrifuged to ensure the content was in the bottom of each well and then placed on the MJ Thermo cycler. The cycling conditions were firstly at 96^{o}C for 30 seconds followed by 45 cycles of 92^{o}C for 8 seconds, 50^{o}C for 8 seconds, and 60^{o}C for 2 minutes, and then chilled down to 10^{o}C until the plate removed. 30 μl of standard sequencing precipitation mix (770 ml ethanol, 16 ml 3 M sodium acetate, 188 ml DDW) was then added to each well and the plate was centrifuged at 4000 rpm at 4^{o}C for 30 minutes. The reaction product was then analysed in ABI 3730 DNA Analyzer.

2.5.4 Sequence analysis

After sequencing experiments, the sequence reads are stored in an internal trace database. 'Capminder' has been adapted to update read length pass information for SNP, and 'Exotrace' was developed in-house by Steven Leonard to compare the sequence reads and to identify SNPs.

2.6 *VKORC1* EXPRESSION IN HUMAN LIVER

2.6.1 RNA extraction

25 human liver biopsies were provided courtesy of Dr Ana Alfirevic (Liverpool) and were stored in liquid nitrogen prior to processing. To extract the RNA, the liver tissues were homogenised in 1ml of Trizol reagent. Homogenised tissue lysate was transferred into an RNase free 1.5 ml tube with 0.22 ml chloroform. The tube was vortexed thoroughly and incubated at room temperature for 3 minutes before being spun at 10000 rpm at 4^{o}C for 10 minutes. The top phase liquid was transferred to a new tube containing 0.6 ml isopropanol followed by thorough vortexing and incubation at room temperature for 10 minutes and another spin at 13000 rpm at 4^{o}C for 10 minutes. The isopropanol was removed and the precipitated pellet was washed with 1 ml of 75% ethanol followed by a centrifugation at 13000 rpm at 4^{o}C for 10 minutes. The ethanol was removed and the tube was dried at room temperature for 10 minutes. The pellet was finally resuspended in 40 μl RNase free water. However, the electrophoresis result indicates that the RNA had been degraded during the preservation of the liver biopsy (also see Figure 4.7).

The extracted RNA was treated with DNase to remove residual DNA contamination. 50 μg RNA was aliquot and treated with 8 units of DNase I at 37^{o}C for 40 minutes (GibcoGRL) followed by extraction with phenol/chloroform (Sigma-Aldrich). After centrifugation, the aqueous layer was transferred to a new 1.5 ml RNase free tube and added with 0.1 volume of 3 M sodium acetate and 3 volumes of 100% ethanol. The tube was incubated at -80^{o}C for an hour followed by a centrifugation at 4^{o}C for 15 minutes. The ethanol was removed and the pellet was washed with 100 μl of 75% ethanol followed by centrifugation at 4^{o}C for 10

minutes. The ethanol was removed and the pellet was resuspended in 20 μl DEPC treated water and stored at -80^{o}C.

2.6.2 Assay of VKORC1 mRNA

The DNase treated RNA samples were converted to cDNA prior to be analysed for *VKORC1* expression. The reverse transcription was carried out with Promega Reverse Transcription System containing 1 μg RNA, 4 μl $MgCl_2$ (25 mM), 2 μl Reverse Transcription 10X Buffer, 2 μl dNTP Mixutre (10 mM), 0.5 μl Recombinant RNasin Ribonuclease Inhibitor 15 units AMV Reverse Transcriptase , 0.5 μg oligo $(dT)_{15}$ primer in 20 μl reaction. The tube was incubated at 42^{o}C for 15 minutes, 95^{o}C for 5 minutes, 0 ^{o}C for 5 minutes and stored at -20 ^{o}C.

1.2 μl of total cDNA was used as template for the quantitative PCR in the presence of SYBR green reporter (Applied Biosystems), as described by Rieder and colleagues (Rieder et al. 2005). Each 25 μl reaction included 12.5 μl 2X SYBR Green PCR Master Mix, 3.0 μl forward and reverse primer (5 μM), 5 μl total cDNA and double distilled water and was performed on an Applied Biosystems 7900HT with standard thermo cycling conditions: 95 ^{o}C for 15 minutes, 40 cycles of 94 ^{o}C for 30 seconds, 60 ^{o}C for 30 seconds and 72 ^{o}C for 30 seconds, and 72 ^{o}C for 5 minutes. The *VKORC1* mRNA level were normalised to *GAPDH* expression levels in each liver sample. The primer sequences on *VKORC1* and *GAPDH* are listed in Table 2.4.

Table 2.4. Primer sequences for *VKORC1* and *GAPDH* gene expression.

Primer	**Sequence**
VKORC1-F	ATCAGCTGTTCGCGCGTC
VKORC1-R	AGAGCACGAAGAACAGGATC
GAPDH-F	ACAGTCAGCCGCATCTTCTT
GAPDH-R	ATGGGTGGAATCATATTGGAAC

2.7 COMPUTATIONAL ANALYSIS

2.7.1 Data processing

SNP genotyping results were stored in an internal database. To assess the data quality control, a PERL script command '~prima/bin/perl/msPlateStatL.pl' was used to produce the genotyping report to check self-priming in negative control (water) and discrepancy calls from duplicate samples. Genotype calls passing quality controls were extracted from the database using a PERL script command '~prima/bin/perl/genoMspec1PlateCallAgg.pl'. Finally, the genotype call file was formatted using a PERL script command 'CallFileFormat.pl', and all the residual discrepant calls were removed. The scripts are listed in Table 2.5.

Table 2.5 Scripts used for data processing.

Script command	Function	Programmed
~prima/bin/perl/msPlateStatL.pl	Data quality check (self-priming and discrepant call)	Jilur Ghori
'~prima/bin/perl/genoMspec1PlateCallAgg.pl	Extract genotyping result from database, including aggressive calls	Jilur Ghori
CallFileFormat.pl	Format the genotyping call file to expected format.	Leslie Chen

2.7.2 Genotype analysis

Genotyped SNPs were analysed and visualised by Haploview (Barrett et al. 2005) which was downloaded from http://www.broad.mit.edu/mpg/haploview/. Analyses including individual patient or SNP exclusion, LD and haplotype block analysis and visualisation, Hardy-Weinberg Equilibrium (HWE) test, and tag SNPs selection were performed in Haploview.

Linkage format files (PED files) recording SNP genotype information were generated with a PERL script described earlier. SNP identifiers and positions were specified in a separate file (INFO file). Individual patients with 50% missing genotype SNPs were excluded for analyses. The haplotype blocks were identified using a predefined definition according to Gabriel and colleagues (Gabriel et al. 2002). The SNP exclusion list was defined with the criteria listed in Table 2.6.

Table 2.6. SNP exclusion criteria in Haploview.

Criteria	**Threshold**
HWE p-value cutoff	0.0010
Minimal genotype call rate	0.7
Maximum number of medel error	1
Minimum MAF	0.05

SNPs giving the same information were excluded in the WARG study, and tag SNPs were selected with the implementation of Tagger (de Bakker et al. 2005). Three different tagging methods can be applied including pairwise tagging with specified r^2 threshold and aggressive tagging using either 2-marker or 2- and 3-marker haplotypes with specified LOD threshold. The tag SNP selection applied in this thesis used a specified r^2 threshold of 0.8.

2.7.3 Multiple sequence alignment

Multiple sequence alignment analysis for both DNA and protein sequence was performed with ClustalX v1.83 (Jeanmougin et al. 1998) downloaded from European Bioinformatics Institute (EBI, ftp.ebi.ac.uk). The DNA and protein sequences were downloaded from

National Center for Biotechnology Information, U.S. National Library of Medicine (http://www.ncbi.nlm.nih.gov/).

2.7.4 Evolutional conserver region (ECR) analysis

To identify potential regulatory regions of *VKORC1*, a comparison for ECR was performed using genomic sequences from human, chimpanzee, mouse, rat, dog and chicken. Analysis was performed using web-based programme, zPicture (Ovcharenko et al. 2004). Human *VKORC1* genomic sequence was used to blast genomes of chimpanzee, mouse, rat, dog, and chicken with Ensembl genome browser (http://www.ensembl.org). The corresponding genomic sequences of the other species were uploaded to the multi-zPicture website (http://zpicture.dcode.org/multiz.php) with human genomic sequencing containing *VKORC1* region.

2.8 STATISTICAL ANALYSIS

2.8.1 Statistics in Uppsala study

Univariate and multivariate regression analysis was assisted by Niclas Eriksson (Uppsala). SAS (SAS Institute Inc) and SPLUS (Insightful Corporation) software were used to perform both univariate and multiple analyses of predictor's impact on the square root of warfarin dose using linear regression models. To address the partial dependence among tests of SNPs in LD, Bonferroni correction for multiple testing based on calculating the effective number of independent tests (Meff) was tested and a Permutation procedure was also applied (Cheverud 2001; Li 2001; Nyholt 2004).

QTPhase component of Unphased software (Dudbridge 2003) was used to estimate haplotype frequencies for *VKORC1* genotypes. Unphased is developed by Frank Dudbridge at MRC Biostatistics Unit, University of Cambridge, United Kingdom and was downloaded from http://www.mrc-bsu.cam.ac.uk/personal/frank/software/unphased/. The means and variances of inter-individual dose were calculated in association with each haplotype. Differences among the haplotype means were also statistically tested with QTPhase.

2.8.2 Statistics in WARG study

Univariate and multivariate regression analysis was assisted by Niclas Eriksson (Uppsala). R version 2.5.1 (http://www.r-project.org/, R foundation for statistical computing,) and SAS version 9.1.3 (SAS Institute Inc) were used for statistical analyses in the WARG study. Univariate and multivariable analyses of predictor impact on the square root of warfarin dose

were calculated using linear regression analyses. To account for partial dependence among tests of SNPs in LD, Bonferroni correction for multiple testing was applied based on the effective number of independent tests (Meff) calculated by a spectral decomposition method (Cheverud 2001; Li 2001; Nyholt 2004).

The coefficient of determination, R^2, was used to measure the proportion of explained variance. Association with over-anticoagulation was evaluated with Log Rank test. Hazard ratios were estimated with Cox regression analyses. The prediction models were based on verified findings and only nominally significant variables ($p<0.05$) were allowed in the final mode, and the R^2 values were calculated.

2.8.3 Statistics in Case/control association

The case/control analysis of warfarin bleeders was performed with PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/) software developed by Shaun Purcell at the Center for Human Genetic Research, Massachusetts General Hospital, and the Broad Institute of Harvard and MIT.

A specified linkage format file (PED file) is needed and contains the individual patient's phenotype and genotype information. SNP marker identifiers and positions were recorded in a separate file (MAP file). Various case-control tests including Cochran-Armitage trend test, Fisher's exact test, two- and three-marker haplotype test using sliding window, recessive gene action test, and gene-gene interaction tests were performed for bleeding complication association. For the cell number smaller than 5 in the recessive gene action test, the p-value, odds ratio and 95% confident interval were calculated with R version 2.6.0.

CHAPTER III

SELECTION OF CANDIDATE GENE AND CONSTRUCTION OF LD MAPS

3.1 INTRODUCTION

Although the number of genes which may affect inter-individual warfarin dose is not known, this chapter will focus on 35 genes which were selected as the best candidates based on pharmacokinetic and pharmacodynamic information and were then tested for association to the two warfarin phenotypes: dose requirement and bleeding complication. Genes interacting with warfarin, function in vitamin K re-cycling or downstream pathways would be the most involved in dose variation or in increased bleeding risk.

Figure 3.1 summarises the genes selected in this study. A first set of 27 genes which does not include members of the protein disulfide isomerase (*PDI*) gene family, vitamin K epoxide reductase (*VKORC1*) and growth arrest-specific gene 6 (*GAS6*), were selected by Drs Mia Wadelius (Uppsala) and Munir Pirmohamed (Liverpool) [see (Wadelius and Pirmohamed 2007) for review]. The *VKORC1*, *GAS6* and *PDI* genes were included later in the project. These genes can be subdivided in five groups based on function:

(i) genes involved in transporting warfarin in the blood (*ORM1* and *ORM2*) and pumping warfarin out of liver cells (*ABCB1*) - may affect drug concentration in the blood and liver, and thereby resulting in dose variations.

(ii) enzymes metabolising warfarin (*CYP1A1*, *CYP1A2*, *CYP2C8*, *CYP2C9*, *CYP2C18*, *CYP2C19*, *CYP3A4* and *CYP3A5*). In addition to the genes encoding these cytochrome P450s, genes that regulate their expression (*NR1I2* and *NR1I3*) may also affect the required dose.

(iii) genes involved in intake and recycling of vitamin K (*APOE*, *VKORC1*, *GGCX*, *CALU*,

EPHX1, *NQO1*, *P4HB*, *PDIA2*, *PDIA3*, *PDIA4*, *PDIA5* and *PDIA6*) constitute another important group to target. The functional variants in these genes may dramatically influence the rate of reducing vitamin K and thereby a lower warfarin dose will be sufficient for effective anticoagulation.

(iv) a number of coagulation factors are activated by gamma-carboxylation. Polymorphisms in the genes encoding the vitamin K dependent (VKD) proteins *F2*, *F7*, *F9*, *F10*, *GAS6*, *PROC*, *PROS1*, and *PROZ* may well influence dose variation and blood coagulation may be easily toggle by low-dose warfarin due to impaired VKD proteins.

(v) some of the above genes are also regulated by *SERPINC1*. Finally, *F5* has been reported to be involved in both thrombosis and haemorrhage and therefore a strong candidate for bleeding complications.

The rational and detailed literature review for each gene is described below along with the results of the first step of this study which was to generate for each gene a comprehensive linkage disequilibrium map in a Swedish sample of 201 warfarin-treated patients enrolled at the Uppsala University by Dr Mia Wadelius and colleagues (see section 2.1.1 and Chapter IV).

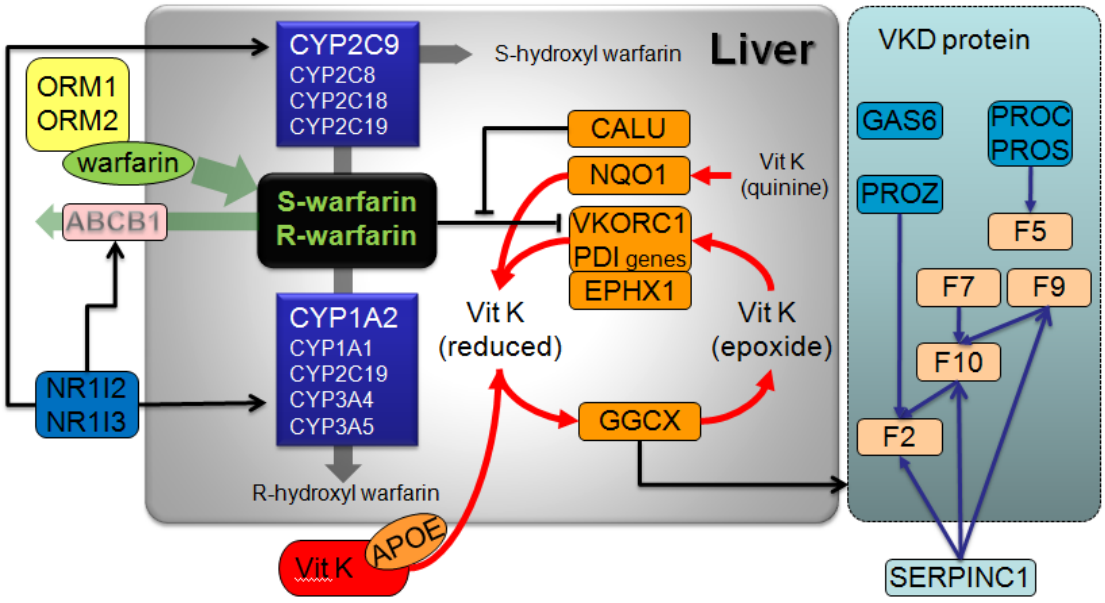


Figure 3.1. Genes selected in this study. (i) genes involved in transporting warfarin in the blood (*ORM1* and *ORM2*) and pumping warfarin out of liver cells (*ABCB1*); (ii) enzymes metabolising warfarin (*CYP1A1*, *CYP1A2*, *CYP2C8*, *CYP2C9*, *CYP2C18*, *CYP2C19*, *CYP3A4* and *CYP3A5*) and upstream regulated (*NR1I2* and *NR1I3*) ; (iii) genes involved in intake and recycling vitamin K (*APOE*, *VKORC1*, *GGCX*, *CALU*, *EPHX1*, *NQO1* and *PDI* genes); (iv) vitamin K dependent (VKD) proteins *F2*, *F7*, *F9*, *F10*, *GAS6*, *PROC*, *PROS1* and *PROZ*, and (v) other coagulation factors (*SERPINC1* and *F5*).

3.2 GENOTYPING STUDY DESIGN

The Uppsala cohort comprises 201 mainly Swedish Caucasian subjects (see section 2.1.1). At the start of this project, there was neither Hapmap data released, nor single nucleotide polymorphisms (SNPs) in public databases such as dbSNP with allele frequency information attached to them (or little, if there was any). Therefore, the first step was to establish a linkage disequilibrium (LD) map for each of the 35 selected genes in this study. We adopted SNPs as the genetic marker of choice to build LD maps because they are both abundant and their bi-allelic nature allows for high throughput genotyping techniques to be applied.

Initially, SNPs were randomly selected from the Ensembl database (http://www.ensembl.org) aiming at 5 kb spacing and minor allele frequency (MAF) equal or greater than 5%. The set of random SNPs was complemented with 73 non-synonymous SNPs and other candidate functional variants giving a total of 728 SNPs which passed assay design for the Sequenom platform (Table 3.1). In the first round of genotyping the 201 Swedish patients, we only obtained 189 common SNPs (MAF $\geq$ 5%) that passed study criteria. At the quality control step we removed SNP assays that failed clustering, those that were out of Hardy-Weinberg equilibrium (HWE, $p < 0.001$) or had a genotype call rate lower than 70% (Whittaker et al. 2005). We also disregarded SNPs with MAF less than 5%. To construct comprehensive LD maps we iteratively selected additional SNPs in gap regions defined physically, over 5 kb, or based on r^2 below 0.8. SNPs within gaps were extracted with a PERL script and followed by manual selection. SNP selection for filling in the outstanding gaps was based on the following criteria:

1. SNP with recorded allele frequency for Caucasians in dbSNP – two hit SNPs
2. SNP which were successfully genotyped in phase I Hapmap project.

3. SNP which had been genotyped internally at the Sanger Institute.

Table 3.1 shows the genotyping summary for each round of assay design. The use of HapMap and dbSNP in subsequent rounds of gap filling has greatly improved success rate from 26% to 43% as can be seen in Table 3.1. Note that in each round gaps are mostly difficult regions, such as highly homologous genes or within regions having repetitive elements, the validated SNP with allele frequency information bring greater opportunity from an initial 26% successful rate to ~43% for the 2nd and 3rd round.

Table 3.1. Summary of iterative genotyping.

Iteration	Attempted SNP	Informative (MAF ≥ 5%)	Disregard	%
1st	728	189	539	25.96%
2nd	202	86	116	42.57%
3rd	169	73	96	43.20%
4th	47	31	16	65.96%
subtotal	1146	379	767	33.07%

In the first three rounds all SNPs were assayed using Sequenom MassEXTEND technology whilst in the fourth round SNPs were assayed using iPLEX which allows higher multiplexing of markers. In the fourth round 6 genes of the protein disulfide isomerase A (PDI) family were examined. With available LD map information from Hapmap, informative SNPs for the six PDI genes were extracted to design the assays and therefore a much higher success assay ratio of 66% was achieved.

A total of 1146 SNPs were tested in order to construct LD maps with an r^2 above 0.8 in the Uppsala sample and over 90% genome coverage for each of the 35 candidate genes. Among

the 73 functional polymorphisms tested (http://www.cypalleles.ki.se/), most of them were excluded in subsequent analysis due to being monomorphic or having MAF lower than 5%. The criterion of 5% MAF was applied as a general threshold for common SNPs as in the HapMap project. Since we looked for common variants which influence warfarin dose, only informative SNPs were further analysed.

We identified 379 SNPs that passed assay quality control and the experimental criteria giving a good coverage for each candidate gene. In 767 disregarded SNP assays, 226 SNPs assays failed to pass the clustering quality control giving a failure rate of total designed SNP assays of 19.7 %. This is mainly because in multiplex experiments the presence of multiple oligo primers often leads to unpredictable behaviour including self priming. The latter causes a marker to show 'water calls' i.e. spurious genotypes in the absence of DNA template, which lower the confidence attached to all genotypes obtained with that assay In the 4^{th} assay design with 47 SNPs, a higher failure ratio of 25.5 % was seen as a result of the higher multiplex design used.

Table 3.2 shows the detailed results for each candidate gene. The number of SNPs which passed the initial clustering quality check (QC) step is counted in the QC column. Each QC SNP was then examined for Hardy-Weinberg equilibrium (HWE), genotype ratio (call rate) and MAF. The SNPs which failed the HWE test were due to high sequence similarity with other loci resulting in a non-specific amplification. Since the Hapmap information was used for the six protein disulfide isomerase genes, no rare or monomorphic SNPs were detected. No allele frequency information for a Swedish population was available at the start of the project but from previously described SNPs, 112 were found to be rare with MAF<5% and 293 could not be detected at all in this population.

Table 3.2. Summary of SNP genotyping for each of the candidate genes.

			MAF ≤ 5%		Failed	
Gene	QC	polyM	rare	monoM	HWE	call rate (70%)
ABCB1	113	38	34	25	3	13
APOE	3	2		1		
CALU	14	9		5		
CYP1A1	15	3	4	7		1
CYP1A2	23	3	7	13		
CYP2C18	22	14	4	3	1	
CYP2C19	35	10	3	16	2	4
CYP2C8	18	12	2			4
CYP2C9	36	19	2	12	1	2
CYP3A4	45	2	6	30	2	5
CYP3A5	24	7	6	8		3
EPHX1	42	25	5	6	2	4
F10	45	15	6	18		6
F2	22	10	3	8		1
F5	74	41	5	19	1	8
F7	32	11	2	15		4
F9	29	11	1	16		1
GAS6	5	4				1
GGCX	16	9	1	4		2
NQO1	18	9	1	7		1
NR1I2	41	20	6	8	2	5
NR1I3	19	9	3	5	1	1
ORM1, & ORM2	29	6		10	8	5
P4HB	6	5				1
PDIA2	4	4				
PDIA3	4	4				
PDIA4	8	7			1	
PDIA5	6	5			1	
PDIA6	7	6			1	
PROC	25	13		5		7
PROS1	50	11	3	20	7	9
PROZ	27	13	3	10		1
SERPINC1	23	9	1	8	1	4
VKORC1LD	39	13	4	13		9
	919	379	112	292	34	102

3.3 CANDIDATE GENES IN WARFARIN TRANSPORTATION

Three genes involved in transporting warfarin in the blood (*ORM1*, *ORM2*) and pumping warfarin out of liver cells (*ABCB1*) were selected (Table 3.3). *ORM1* and *ORM2* are located back to back on chromosome 9q. Both genes have 6 exons, encode peptides comprising 201 amino-acid residues and have very high DNA sequence homology (Figure 3.2). The latter caused extreme difficulty in designing genotyping assays for exonic SNPs. *ABCB1* spans 209 kb on chromosome 7 and belongs to the ATP-binding cassette transporter gene family encoding various cellular pumps in different cells. *ABCB1* is particularly a cellular efflux pump for xenobiotics.

Table 3.3. Candidate genes in warfarin transportation.

Gene	Location	Exons	Transcript	Protein (residues)	Function of protein
ORM1	Chr 9: 114083890 - 114087309 bp	6	802 bps	201	A plasma glycoprotein that functions as a carrier of warfarin in the blood
ORM2	Chr 9: 114171703 - 114175086 bp	6	760 bps	201	A plasma glycoprotein that functions as a carrier of warfarin in the blood
ABCB1	Chr 7: 85668428 - 85877818 bp	29	4643 bps	1279	A cellular efflux pump for xenobiotics.

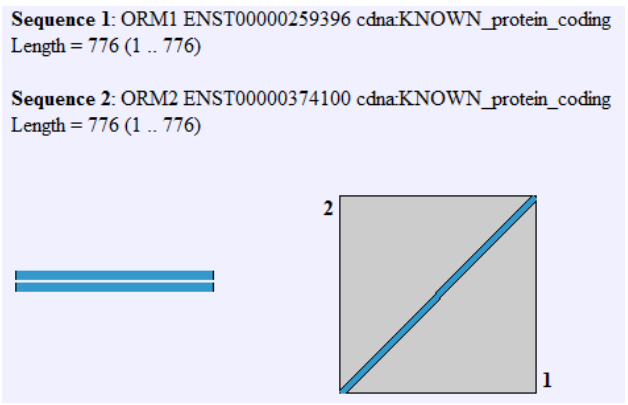


Figure 3.2. cDNA sequence alignment of ORM1 and ORM2. cDNA sequences of both genes contains 776 nucleotides, and only 31 nucleotides are different resulting in a 96% similarity.

3.3.1 ORM1 and ORM2

Warfarin has been intensively studied over decades as an effective anticoagulant. When warfarin is taken, it will be quickly and nearly completely absorbed from the stomach and the upper gastrointestinal tract and then bound to serum albumin and alphal-acid glycoprotein encoded by *ORM1* (orosomucoid 1) and *ORM2* (orosomucoid 2). In plasma, ORM proteins are presented as a mixture of ORM1 and ORM2 at a molar ratio of 3:1 (Yuasa et al. 1997).

A chromatographic study (Nakagawa et al. 2003) has shown that warfarin has different affinity to orosomucoid protein; harbouring amino acid changes due to genetic variation which suggested that orosomucoid protein might influence warfarin dose. Meanwhile, serum albumin has very robust transporting functionality for a broad range of proteins and thus, it is less likely to play a role in inter-individual dose variability. No evidence was reported in the literature for an association between albumin and warfarin dose.

As described earlier, due to high sequence homology (96%), 28 SNPs in *ORM1, ORM2* and flanking regions were tested, but only 6 SNPs located in either intronic or flanking regions gave good results (Table 3.4). None of these 6 SNPs have been reported to be functionally important. LD analysis indicated that the 6 SNPs are in strong LD (Figure 3.3A). Hapmap project phase II data (CEU panel) also show that both *ORM1* and *ORM2* are located at the end of an LD block (Figure 3.3B). In this chapter the phase II result of Hapmap project was compared with Swedish LD structure for each of the candidate genes.

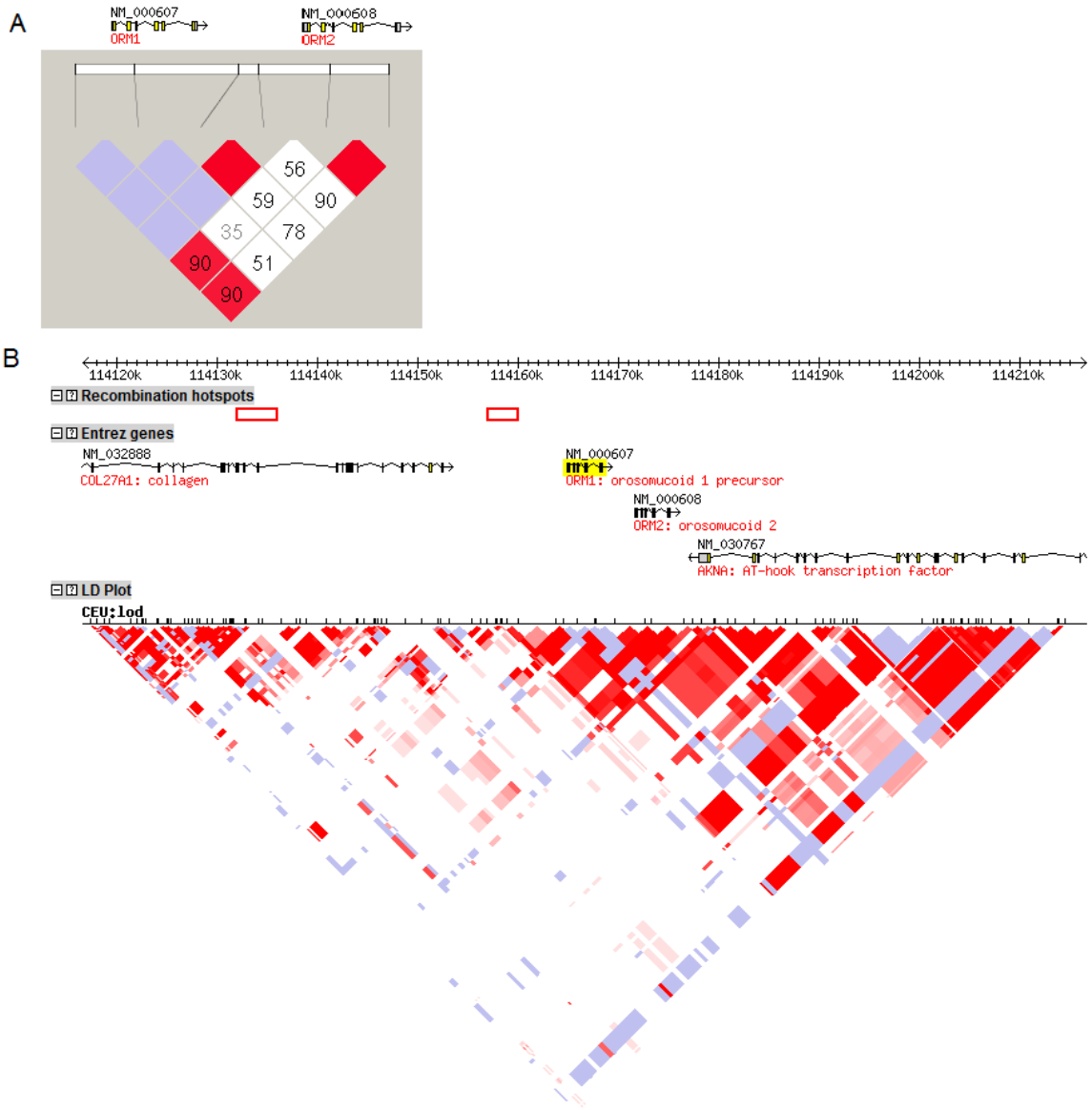


Figure 3.3. (A) Swedish LD structure; (B) LD structure from Hapmap project phase II result. (A) None of exonic SNPs in *ORM1* and *ORM2* were successfully genotyped because of high cDNA sequence similarity (96%). (B) LD structure from Hapmap phase II result indicates the *ORM1* and *ORM2* are at the end of a LD block, and a recombination hot spot is 5 kb away.

Table 3.4. Genotyping summary of *ORM1* and *ORM2*.

Genes	Study SNPs (tested, passed)
ORM1	15 tested, 3 passed study criteria
Coding ns	0
Coding s	0
UTR	0
Intronic	1: rs10982151
Flanking	2: rs2787337, rs1687390
ORM2	13 tested, 3 passed study criteria
Coding ns	0
Coding s	0
UTR	0
Intronic	2: rs17230081, rs1976193
Flanking	1: rs3762055

3.3.2 ABCB1 (MDR1)

Preliminary evidence has demonstrated that P-glycoprotein transports warfarin (as a substrate) through liver plasma membranes using an inhibition assay (Sussman N 2002). This P-glycoprotein is encoded by *ABCB1* (ATP-binding cassette transporter B1) and is also named *MDR1* (multidrug resistance protein 1). Increased expression of *ABCB1* in the small intestine will reduce the absorption of drugs which are substrates for P-glycoprotein, and thus result in a reduced bioavailability and unattained therapeutic plasma concentration of a drug (Cascorbi et al. 2001).

There are a number of reports showing that sequence variants in *ABCB1* influence its transcriptional and translational expression, as well as the pharmacokinetic profiling of various drugs (Ishikawa et al. 2004). A study demonstrated that a haplotype bearing the C3435T variant in exon 26 was over-represented among patients requiring a lower maintenance dose of warfarin (Wadelius et al. 2004). Coincidently, C3435T variation has been shown to reduce drug efflux which is in agreement with the need of a lower dose because of reduced removal of toxic metabolites from the liver cells. There is also another possibility according to which the increased intestinal expression of P-glycoprotein can reduce the absorption of warfarin resulting in a higher maintenance dose (Cascorbi et al. 2001). However, a recent study demonstrated that C3435T affects the timing of co-translational folding and insertion of P-glycoprotein into the membrane (Kimchi-Sarfaty et al. 2007). The observed lower level of membrane insertion of the T variant would be consistent with lower efflux of warfarin from liver cells.

In this study, 113 SNPs in *ABCB1* passed QC and 38 with MAF higher than 5% were

informative. None of these 38 SNPs had been previously reported to be functionally important, however, two nsSNPs cause a change of amino acid composition (N21D and A893S/T) and one nsSNP may even change the translational initiation (Table 3.5).

The genomic structure of *ABCB1* and corresponding LD map is shown in Figure 3.4. Panel A in Figure 3.4 shows the LD map in the Swedish population sample from Uppsala in which the gene seems to be split in the middle into two LD blocks. The same is found in Hapmap (Figure 3.4B) where LD decays as a function of rapidly evolving SNPs in the first intron. Two recombination hot spots are found in Hapmap: one in the centre of *ABCB1* genomic region and the other one at the 3' end of the gene (Figure 3.4).

Table 3.5. Genotyping summary of *ABCB1*.

Genes	Study SNPs (tested, passed)
ABCB1	113 tested, 38 passed study criteria
Coding ns	3: rs2214102 (changes translation initiation), rs9282564 (N21D), rs2032582 (A893S/T)
Coding s	1: rs1045642
UTR	1: rs3842
Intronic	33: rs2188531, rs6465117, rs17328991, rs10267099, rs2157926, rs2214101, rs17149824, rs4728709, rs9282564, rs1858923, rs3789243, rs1202181, rs1202172, rs1989830, rs1202179, rs1202180, rs10260862, rs2235015, rs1202167, rs1202169, rs955000, rs868755, rs1922240, rs2235033, rs2235035, rs2235013, rs2091766, rs2235046, rs1922242, rs4148737, rs2235040, rs6959435, rs4148742, rs2235067

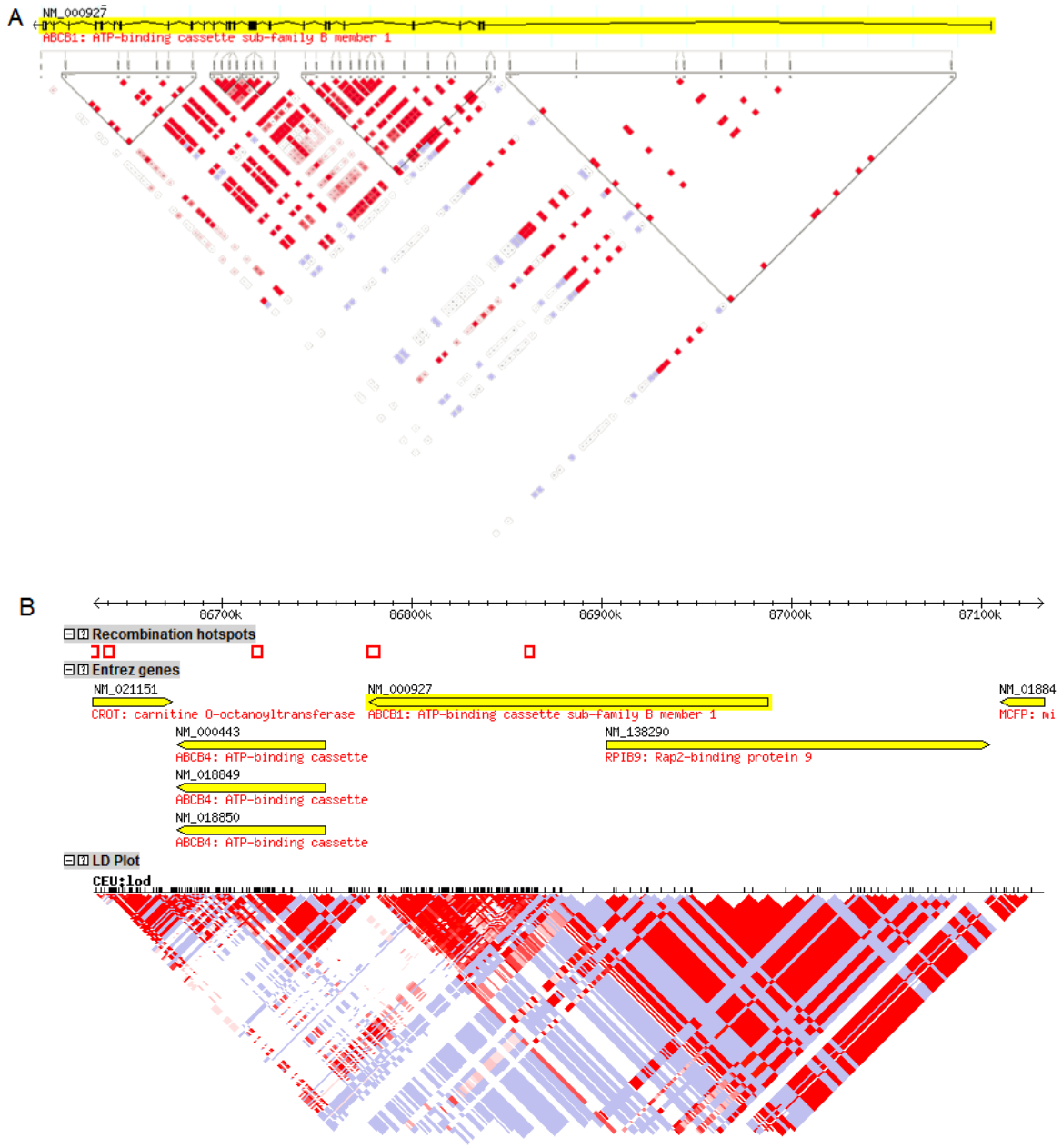


Figure 3.4. LD structure from (A) 201 Swedish and (B) Hapmap CEU. (A) LD in Swedish population suggests a recombination hot spot in the centre of ABCB1. (B) This hot spot is also discovered in HapMap Caucasian population.

3.4 CANDIDATE GENES IN WARFARIN METABOLISM

Warfarin consists of a racemic mixture of two active optical isomers - R and S forms. Each is cleared by different pathways (discussed later this section). S-warfarin has five times the potency of the R-enantiomer with respect to vitamin K antagonism.

Table 3.6 lists the cytochrome P450 metabolising enzymes which convert warfarin to less toxic forms. The *CYP2C9, CYP2C8, CYP2C18, and CYP2C19* genes are clustered in a region of high linkage disequilibrium on chromosome 10 (Ahmadi et al. 2005), whereas *CYP1A1* and *CYP1A2* are located back to back on chromosome 15 and *CYP3A4* and *CYP3A5* on chromosome 7.

Table 3.6. Candidate genes in warfarin metabolism.

Gene	Location	Exons	Transcript	Protein (residues)	Function of protein
CYP2C9	Chr 10: 96688405 - 96739137 bp	9	1847 bps	490	Polymorphic hepatic drug metabolising enzyme. Metabolism of S-warfarin
CYP2C8	Chr 10: 96786520 - 96819244 bp	9	1923 bps	490	Polymorphic hepatic drug metabolising enzyme. Minor pathway for R & S-warfarin
CYP2C18	Chr 10: 96432700 - 96485937 bp	9	2418 bps	490	Found in the liver and lung. Minor pathway for R & S-warfarin
CYP2C19	Chr 10: 96437901 - 96603007 bp	9	1901 bps	490	Polymorphic hepatic drug metabolising enzyme. Minor pathway for R & S-warfarin
CYP1A1	Chr 15: 72798943 - 72804930 bp	7	2601 bps	512	Extrahepatic oxidation, inducible. Metabolism of R-warfarin
CYP1A2	Chr 15: 72828257 - 72834505 bp	7	1618 bps	516	Hepatic oxidation, inducible. Metabolism of R-warfarin
CYP3A4	Chr 7: 97889181 - 97916385 bp	13	2768 bps	503	Hepatic oxidation, inducible. Metabolism of R-warfarin
CYP3A5	Chr 7: 97780394 - 97812183 bp	13	1707 bps	502	Polymorphic hepatic and extrahepatic oxidation. Metabolism of R-warfarin

3.4.1 S-warfarin metabolism (CYP2C8, CYP2C9, CYP2C18, and CYP219)

When warfarin circulates to the liver, the S-enantiomer is mainly converted to 6- and 7-hydroxywarfarin by CYP2C9 whilst CYP2C8 and CYP2C19 metabolise small amounts of S-warfarin to the 4-hydroxyl metabolite (Kaminsky and Zhang 1997).

There are a number of non-synonymous coding variants identified in *CYP2C9* with differential, mostly decreased, enzymatic activities (Schwarz 2003). Some of these variants were reported in particular ethnic populations including the *4 allele which is identified in Japanese, *5 and *6 alleles identified in African-American, and *11 which is relatively rare in European Caucasian and African-American (Schwarz 2003).

Among the 30 reported variants which are functionally important, the *2 and *3 variants have been convincingly shown to be associated with reduced warfarin dose (Aithal et al. 1999; Higashi et al. 2002; Wadelius et al. 2005; Wadelius et al. 2004). Compared to a CYP2C9 metaboliser who is homozygous for the *1 wild type allele, a homozygote for the *2 allele will have a reduced CYP2C9 enzymatic activity of 12%, whereas a homozygote for the *3 allele has only 5% metabolising activity. Many studies have shown that patients with *2 and *3 alleles require lower daily warfarin dose (Aithal et al. 1999; Higashi et al. 2002; Wadelius et al. 2005; Wadelius et al. 2004). A meta-analysis of nine previous studies indicated that the *2 allele would reduce dose requirement by 17% whereas *3 allele comprehensively lowered the required dose by 37% (Sanderson et al. 2005).

Of the genes listed in Table 3.6 only *CYP2C9* has been reported to be associated with warfarin dose. Although many functional variants have been reported in these CYP genes,

only a few of them were found to be polymorphic in the 201 Swedish patients. Table 3.7 lists all informative SNPs including functional variants used for analysis. The LD structure of these four genes is discussed in the chapter 4 (section 4.4) together with the association analysis for warfarin dose requirement. Apart from those alleles described on the CYP website (http://www.cypalleles.ki.se/), SNP rs2281891 (T385M) in *CYP2C18* has been reported to be associated with differential turnover rate for drugs (Goldstein et al. 1994; Kaminsky et al. 1993) and is polymorphic in this Swedish cohort.

Table 3.7. Genotyping summary for candidate genes in S-warfarin metabolism.

Genes	Study SNPs (tested, passed)	SNP aliases
CYP2C9	36 tested, 19 passed study criteria	rs1799853 (formerly rs17110268) = CYP2C9*2, coding R144C, low activity; rs1057910 = CYP2C9*3, coding I359L, very low activity
Coding ns	2 : rs1057910 (I359L), *2 rs1799853 (R144C)	
Coding s	1: rs1057911	
UTR	0	
Intronic	14: rs2298037, rs9332222, rs9332214, rs9332197, rs1934966, rs1934964, rs9325473, rs1856908, rs4917639, rs2153628, rs2475376, rs10509679, rs2860905, rs9332108	
Flanking	2: rs4917636, rs4607998	
CYP2C8	18 tested, 12 passed study criteria	rs17110453 = CYP2C8*1C, 5' upstream, function unknown; rs11572080 = CYP2C8 *3, coding R139K, decreased activity; rs1058930 = CYP2C8 *4, coding I264M, decreased activity
Coding ns	2: rs11572080 (R139K), rs1058930 (I264M)	
Coding s	0	
UTR	1: rs1058932	
Intronic	7: rs2275622, rs3752988, rs1341163, rs947173, rs1891071, rs2275620, rs7898759	
Flanking	2: rs1557044, rs17110453	
CYP2C18	22 tested, 14 passed study criteria	rs2281891= coding T385M, high metabolism of certain substrates, but not warfarin
Coding ns	1: rs2281891 (T385M)	
Coding s	0	
UTR	1: rs2860840	
Intronic	10: rs10509675, rs7919273, rs1926711, rs7898763, rs7099637, rs7896133, rs7478002, rs2901783, rs2860837, rs1926706	
Flanking	2: rs10736086, rs12249418	
CYP2C19	35 tested, 10 passed study criteria	rs17879456 = rs4244285, CYP2C19*2A, splicing defect, no enzyme activity; rs17882687 = CYP2C19*15, coding I19L, effect on enzyme unknown
Coding ns	2: rs17882687 (I19L), rs17879456 (splicing defect)	
Coding s	1: rs3758580	
UTR	0	
Intronic	4: rs1853205, rs4244284, rs4417205, rs17882419	
Flanking	3: rs12248560, rs3814637, rs4250786	

3.4.2 R-warfarin metabolism (CYP1A1, CYP1A2, CYP3A4 and CYP3A5)

The R-enantiomer is mainly metabolised to 6-, 8-, and 10-hydroxylated form. The minor metabolites include 4- and 7-hydroxywarfarin. The 6- and 8-hydroxyl metabolites are formed predominately by CYP1A2 with some contributions from CYP1A1 and CYP2C19 (Kaminsky and Zhang 1997). The 10-hydroxyl form is generated by CYP3A of which activity is derived from CYP3A4 and CYP3A5 with similar substrate specificities (Kaminsky and Zhang 1997; Kuehl et al. 2001). The genotyping results are summarised in Table 3.8.

Table 3.8. Genotyping summary for candidate genes in R-warfarin metabolism.

Genes	Study SNPs (tested, passed)	SNP aliases
CYP1A1	15 tested, 3 passed study criteria	-
Coding ns	0	
Coding s	0	
UTR	0	
Intronic	2: rs4646421, rs2606345	
Flanking	1: rs2470893	
CYP1A2	23 tested, 3 passed study criteria	rs2470890: CYP1A2*1B, coding synonymous, function unknown; rs762551 = CYP1A2*1F, 5'UTR, higher inducibility
Coding ns	0	
Coding s	1: rs2470890	
UTR	1: rs762551	
Intronic	1: rs2472304	
Flanking	0	
CYP3A4	45 tested, 2 passed study criteria	rs11773597 = CYP3A4*1F, 5'upstream, function unknown
Coding ns	0	
Coding s	0	
UTR	0	
Intronic	1: rs11773597	
Flanking	1: rs2242480	
CYP3A5	24 tested, 7 passed study criteria	rs776746 = CYP3A5*3, splicing defect, low activity; g-3844G>A, 5'upstream, function unknown; rs28365067 = intronic g5215C>T, function unknown; rs28365094 = intronic g27050A>G, function unknown
Coding ns	1: rs776746 (splicing defect)	
Coding s	0	
UTR	0	
Intronic	3: rs6976017, rs28365067, rs28365094	
Flanking	3: rs4646457, g-3844G>A (not in dbSNP), rs15524	

The LD map of *CYP1A1* and *CYP1A2* is shown in Figure 3.5A. The poor genotyping coverage in this region is not resulting from high sequence similarity or repetitive regions

(Figure 3.6). In fact, typing of 38 SNPs was attempted across this region and 31 SNPs were discovered which are either rare (11 SNPs, MAF ≤ 5%) or monomorphic (20 SNPs). This is in complete agreement with Hapmap phase II data which report only 7 SNPs in the two genes and 5 kb upstream and downstream region (Figure 3.5B). These results suggest that the CYP1A locus does not tolerate the accumulation of sequence variants which may reflect functional constraints.

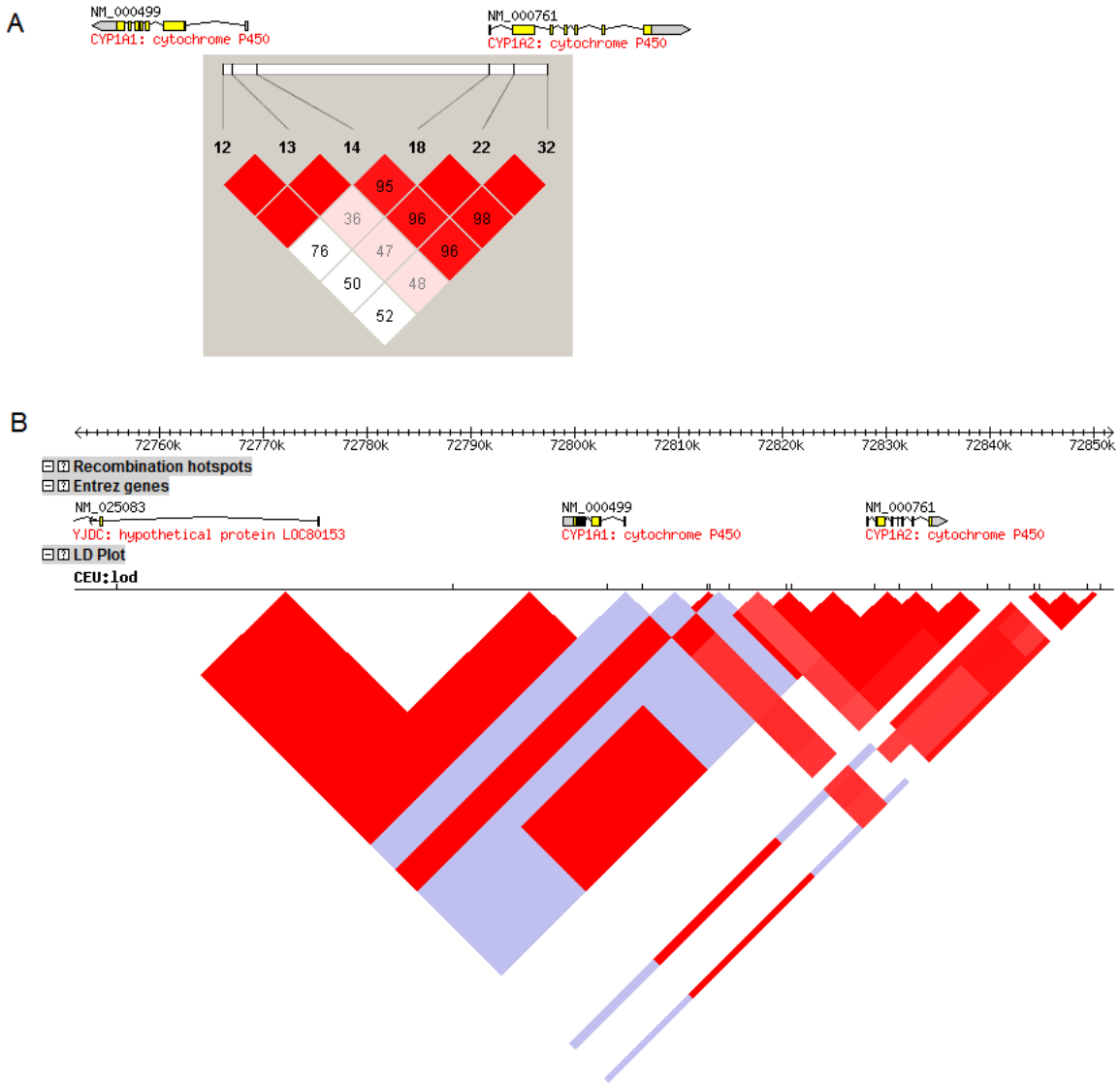


Figure 3.5. LD structure of CYP1A1 and CYP1A2 in (A) Swedish and (B) Hapmap CEU populations.

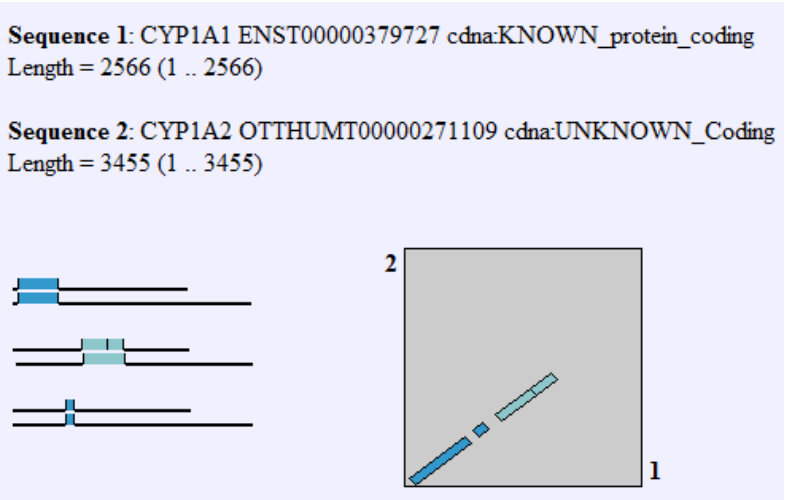


Figure 3.6. Pairwise cDNA sequence alignment of *CYP1A1* and *CYP1A2*. Three segments are high homologous with 88%, 74%, 81%, respectively in the first two third of the genes.

Figure 3.7 shows the LD structures of *CYP3A4* and *CYP3A5* in the Swedish and Hapmap Caucasian sample. After attempting two rounds, only two SNPs were obtained from the last intron and 3' flanking region of *CYP3A4*. A total of 45 SNPs were tested in *CYP3A4* with 6 SNPs found to be rare and 30 monomorphic (Table 3.2). A sparse coverage is also seen in the Hapmap data which show the presence of a recombination hot spot in the centre of *CYP3A4* gene (Figure 3.7B).

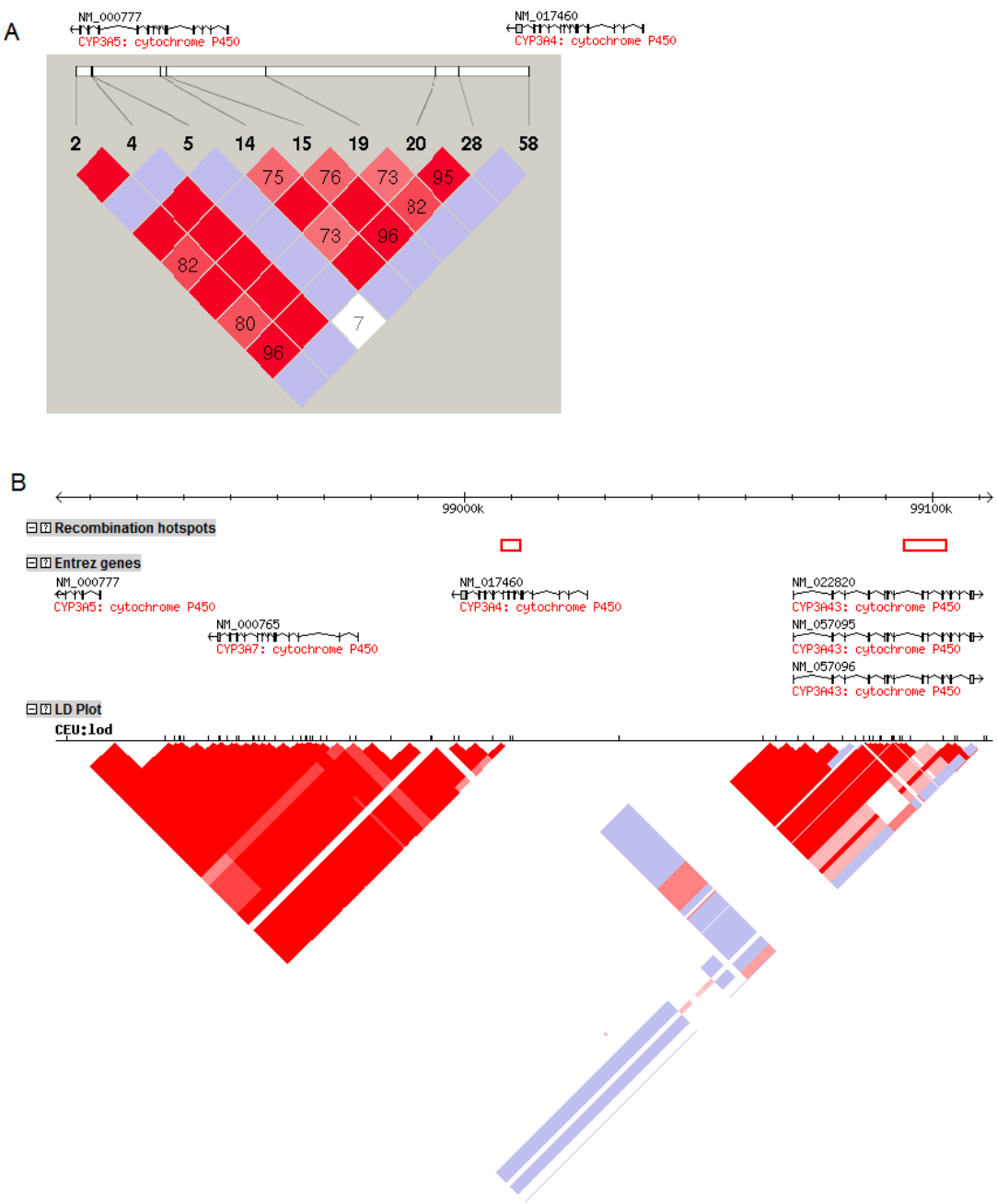


Figure 3.7. LD structure of *CYP3A4* and *CYP3A5* in (A) Swedish and (B) Hapmap CEU populations.

3.4.3 P450 inducibility

Since much of the detoxification of xenobiotic compounds is performed by CYP enzymes, it is also important to look at the genes that regulate CYP expression and activity. Some studies

have shown that these CYP enzymes are inducible especially warfarin is prescribed with other drugs (Lehmann et al. 1998; Moore et al. 2000). In that respect, two genes *NR1I2* and *NR1I3* were identified as regulating the expression of CYP genes (Table 3.9) and therefore selected for analysis as they may harbour variants that influence warfarin dose.

Table 3.9. Candidate genes of regulating warfarin metaboliser.

Gene	Location	Exons	Transcript	Protein (residues)	Function of protein
NR1I2	Chr 3: 120982021 - 121020021 bp	9	2753 bps	473	Mediates induction of CYP2C9, CYP3A4, other CYP enzymes and ABCB1
NR1I3	Chr 1: 158012528 - 158021028 bp	9	1337 bps	348	Transcriptional regulation of a number of genes including CYP2C9 andCYP3A4

The pregnane X receptor (PXR) encoded by *NR1I2* (nuclear receptor subfamily 1, group I, member 2) is activated by a variety of endogenous and exogenous chemicals including St John's Wort; a herbal antidepressant which interacts with warfarin (Table 3.10). PXR is also reported to induce CYP2C9, CYP3A4, and other CYPs (Chen et al. 2004; Lehmann et al. 1998). Interestingly, PXR also regulates the drug efflux by activating expression of ABCB1 (MDR1) (Geick et al. 2001; Synold et al. 2001).

The constitutive androstane receptor (CAR) which is encoded by *NR1I3* (nuclear receptor subfamily 1, group I, member 3), is a nuclear hormone receptor that functions cooperatively with PXR (*NR1I2*) to detoxify xenobiotics (Table 3.10). CAR is also reported to be associated with induced transcription of *CYP2C9* and *CYP3A4* (Assenat et al. 2004).

Table 3.10. Genotyping summary of *NR1I2* and *NR1I3*.

Genes	Study SNPs (tested, passed)
NR1I2	41 tested, 20 passed study criteria
Coding ns	0
Coding s	0
UTR	5: rs3814057, rs1054191, rs3732360, rs3732359, rs1523127
Intronic	13: rs2472682, rs3732357, rs3732356, rs1464602, rs7643645, rs2461818, rs13059232, rs2461823, rs2472677, rs1403527, rs2056530, rs2472672, rs2276706
Flanking	2: rs1523130, rs7643038
NR1I3	19 tested, 9 passed study criteria
Coding ns	0
Coding s	1: rs2307424
UTR	0
Intronic	5: rs2502804, rs6686001, rs3003596, rs2307418, rs4073054
Flanking	3: rs2501870, rs7530560, rs4233368

20 SNPs in *NR1I2* and 9 SNPs in *NR1I3* were analysed (Figure 3.8). The first exon of *NR1I2* is un-translated and two recombination hot spots map in the first and second intron respectively (Figure 3.8A). Hapmap results confirm the presence of the two recombination hot spots in Caucasians which are located 20 kb apart (Figure 3.9). This suggests a higher mutation rate in this gene, and a lot of SNPs have been reported in Ensembl SNP database. In contrast to *NR1I2*, SNPs in *NR1I3* suggest a strong LD across this region which is in agreement with Hapmap (Figure 3.8B).

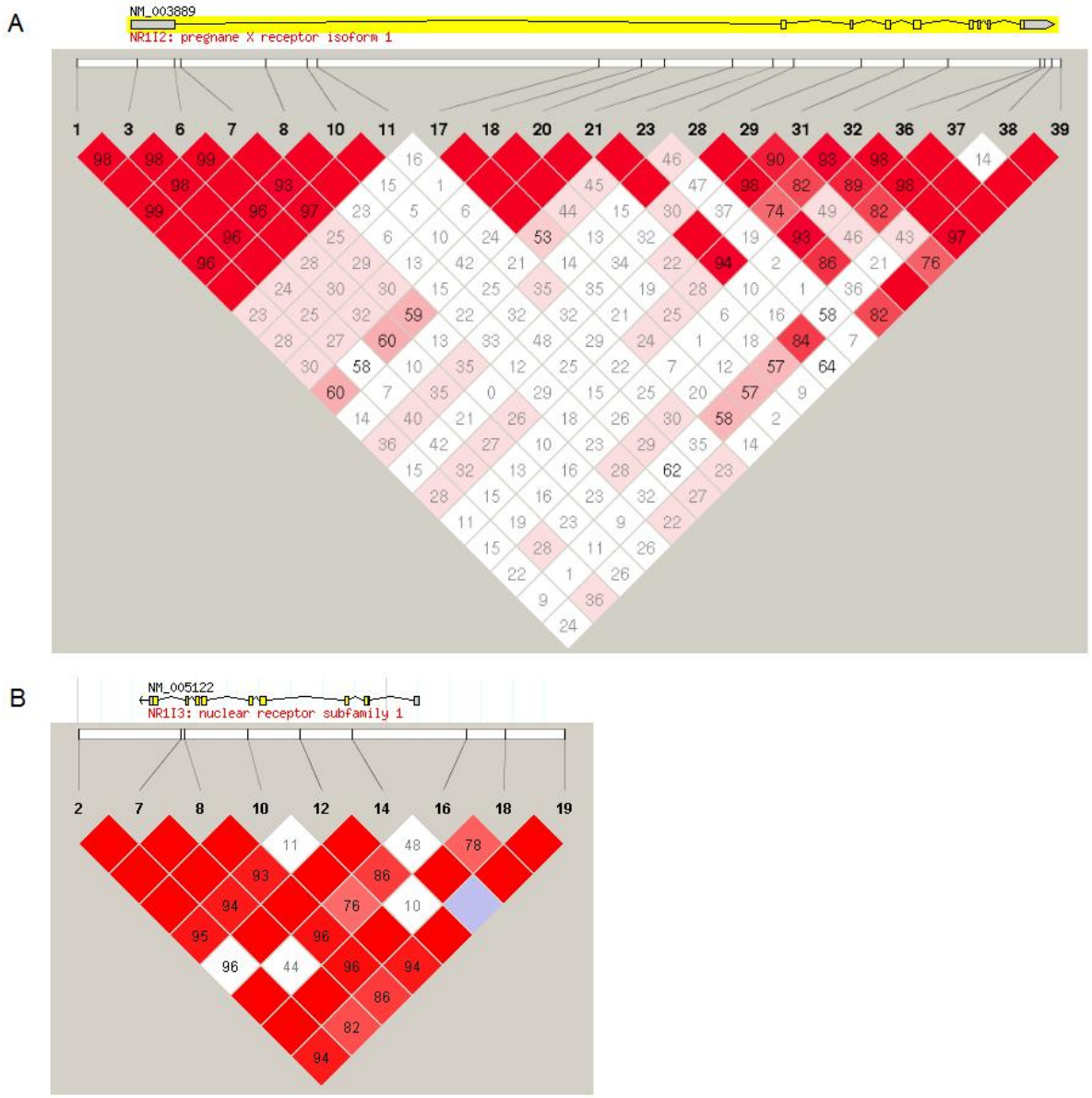


Figure 3.8. LD structure of (A) *NR1I2* and (B) *NR1I3*. (A) The transcription of *NR1I2* starts from the second exon, and the whole exon 1 is untranslated. Two potential recombination hot spots in intron 1 and intron 2 were observed in *NR1I2*. (B) SNPs in and flank on *NR1I3* indicate a strong LD in this region.

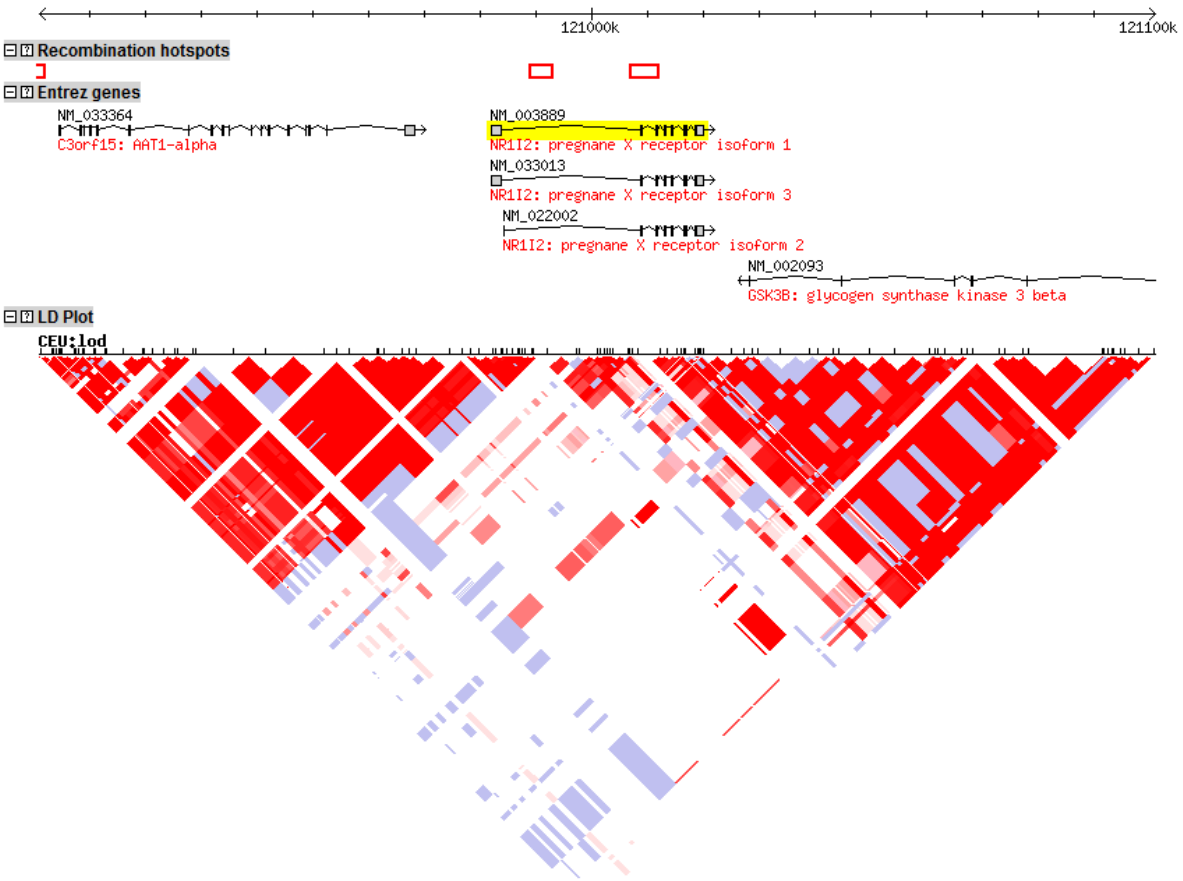


Figure 3.9. Hapmap CEU result for *NR1I2* region. Two recombination hot spots were identified in *NR1I2* in 96 Caucasians. This suggests a potential chromosome instability and higher mutation rate in this gene.

3.5 CANDIDATE GENES IN VITAMIN K INTAKE AND RECYCLING

Vitamin K is an important cofactor to activate certain blood coagulation factors. Warfarin acts on the vitamin K epoxide reductase (VKOR) complex on the endoplasmic reticulum (ER) membrane in liver cells to toggle the recycling of vitamin K and thereby, is interfering with blood coagulation. A diet with high vitamin K update would prevent the action of warfarin and result in a higher maintenance dose. Nowadays, more and more people try to live healthier eating more leafy green vegetables, Brassica vegetables, and fruits which are the major sources of daily vitamin K1 uptake. Table 3.11 lists all genes known to be involved in the recycling of vitamin K1

Table 3.11. Candidate genes in vitamin K intake and recycling.

Gene	Location	Exons	Transcript	Protein (residues)	Function of protein
APOE	Chr 19: 50100879 - 50104489 bp	4	1179 bps	317	Apolipoprotein E serves as a ligand for receptors that mediate the uptake of vitamin K
VKORC1	Chr 16: 31009677 - 31013777 bp	3	997 bps	163	A hepatic epoxide hydrolase that catalyses the reduction of vitamin K. The target of warfarin
EPHX1	Chr 1: 222304587 - 222339995 bp	9	1605 bps	455	A hepatic epoxide hydrolase with the potential to reduce vitamin K
NQO1	Chr 16: 68300807 - 68317893 bp	6	2448 bps	274	A detoxifying enzyme that has the potential to reduce the quinine form of vitamin K
GGCX	Chr 2: 85687865 - 85700237 bp	15	3155 bps	758	Carboxylates vitamin K dependent coagulation factors and proteins in the vitamin K cycle
CALU	Chr 7: 127973368 - 128005478 bp	7	3316 bps	315	Binds to the vitamin K epoxide reductase complex and inhibits the effect of warfarin
P4HB	Chr 17: 77,394,322-77,411,834 bp	11	2580 bps	508	Providing electrons to reduce CXXC centre in VKORC1
PDIA2	Chr16: 273,153-277,216 bp	11	1698 bps	525	Providing electrons to reduce CXXC centre in VKORC1
PDIA3	Chr15: 41,825,882-41,852,769 bp	13	3727 bps	505	Providing electrons to reduce CXXC centre in VKORC1
PDIA4	Chr7: 148,331,087-148,356,666 bp	10	2903 bps	645	Providing electrons to reduce CXXC centre in VKORC1
PDIA5	Chr3: 124,268,599-124,363,641 bp	16	1656 bps	262	Providing electrons to reduce CXXC centre in VKORC1
PDIA6	Chr2: 10,840,968-10,870,421 bp	13	2338 bps	440	Providing electrons to reduce CXXC centre in VKORC1

3.5.1 APOE

Vitamin K1 is a fat-soluble vitamin and is absorbed from the small intestine, together with dietary fat. It is then circulated in the blood by chylomicrons, large lipoprotein molecules secreted by absorptive cells of the small intestine. Vitamin K1 conjugated chylomicron becomes matured by acquiring apolipoprotein C-II (APOC2) and apolipoprotein E (APOE) from High Density Lipoproteins (HDL) and is subsequently cleared by the liver through an APOE receptor specific uptake (Berkner and Runge 2004; Lamon-Fava et al. 1998). It has been reported that different APOE alleles lead to variable uptake efficiency of the chylomicron with evidence showing that the *E4 allele is faster at uptake than *E3 and which is then faster than *E2 (Kohlmeier et al. 1996). Patients carrying the APOE*E2 allele, which is allegedly associated with less efficient uptake of chylomicron and thereby vitamin K1, had an increased risk of warfarin associated intracerebral haemorrhage (Rosand et al. 2000). Table 3.12 shows the two SNPs, rs429358 and rs7412, which passed study criteria and are sufficient to discriminate the APOE haplotypes carrying the E2, E3, and E4 allele.

Table 3.12. Genotyping summary of *APOE*.

Genes	Study SNPs (tested, passed)	SNP aliases
APOE	3 tested, 2 passed study criteria	rs429358 = coding C130R
Coding ns	2: rs429358 (C130R), rs7412 (R176C),	(formerly C112R); rs7412 =
Coding s	0	coding R176C (formerly R158C) ;
UTR	0	These 2 SNPs discriminate
Intronic	0	between the haplotypes E2, E3,
Flanking	0	E4

3.5.2 VKORC1

Warfarin and other vitamin K antagonists preclude the regeneration of vitamin K from its

oxidative forms by targeting VKOR. The gene encoded for VKOR was poorly understood up to 2004, when the coding gene was identified by two studies which demonstrated that the vitamin K epoxide reductase subunit 1 (VKORC1) is the direct target of warfarin by RNAi inhibition and functional assay (Li et al. 2004; Rost et al. 2004a). *VKORC1* is a small gene located on chromosome 16 which comprises 3 exons and encodes a peptide of 163 amino acids. Although this gene was shown to be the target for warfarin, the mechanism by which it functions as a reductase remains unclear.

Twenty-five publicly available SNPs were tested and only four polymorphic SNPs were obtained in our Swedish sample; two intronic SNPs (rs9934438 and rs2359612), one in the 3'-UTR (rs7294), and one 5' upstream (rs9923231) (Table 3.13). According to Hapmap, *VKORC1* is located near one end of a large LD block in Caucasians (CEU panel; Figure 3.10). In later stages of the study a further set of 14 SNPs, lying outside *VKORC1*, were selected to analyse this large LD block. For illustration purposes the genomic structure of *VKORC1* and LD architecture of this locus is discussed alongside the results of the association analysis in chapter 4, section 4.3.1 and 4.3.2).

Table 3.13. Genotyping summary of *VKORC1* and its nearby flanking region.

Genes	Study SNPs (tested, passed)	SNP aliases
VKORC1	39 tested, 13 passed study criteria	rs9923231 = upstream 3673 or -1639 G>A, low expression?; rs9934438 = intronic 6484 or 1173 C>T, function unknown; rs2359612 = intronic 7566 or 2255 C>T, function unknown; rs7294 = 3'UTR 9041 or 3730 G>A, function unknown
Coding ns	0	
Coding s	0	
UTR	1: rs7294	
Intronic	2: rs9934438, rs2359612	
Flanking	10: rs4889537, rs9923231, rs8046978, rs4889599, rs11642603, rs4889630, rs7405035, rs4889490, rs11642466, rs7194347	

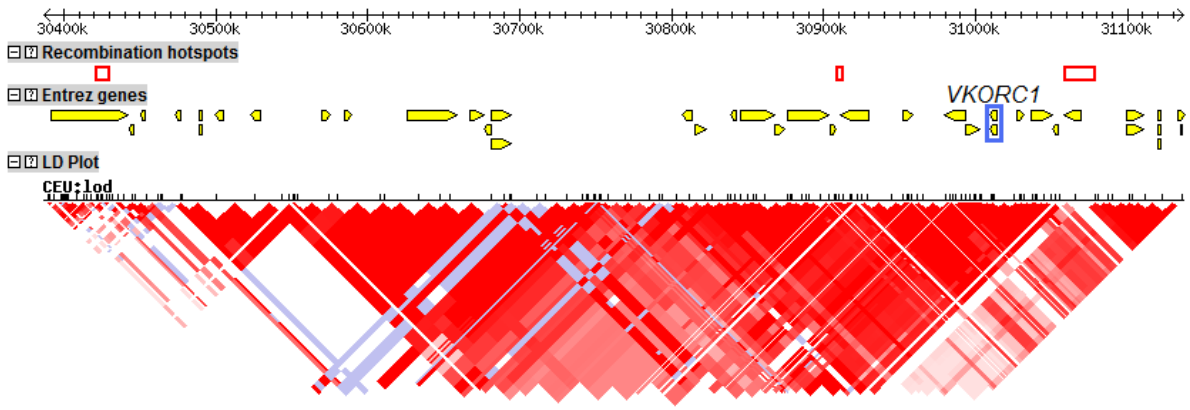


Figure 3.10. LD map of *VKORC1* in Hapmap CEU panel. Genomic architecture in Chr16:30386726 to 31136725 whereas *VKORC1* (Chr16, positions 31,009,676 to 31,013,776) is located in one end of a large LD block.

3.5.3 EPHX1

The microsomal epoxide hydrolase (EPHX1) catalyses the initial epoxide hydration reaction in the vitamin K recycling pathway (Cain et al. 1998) and forms a complex with VKOR. It is encoded by *EPHX1* on chromosome 1 (Table 3.9) which, according to Daly and King (2003), harbours functional polymorphisms that may affect protein stability and could contribute to warfarin response (Daly and King 2003). A recent study in an Israeli population has shown an association between high doses of warfarin and a coding *EPHX1* polymorphism (rs1051740) in *CYP2C9**1 wild type patients (Loebstein et al. 2005).

As shown in Table 3.14, of the 25 SNPs that passed study criteria in *EPHX1*, two are nsSNPs with rs1051740 being common in our Swedish sample. The LD map is drawn with the main contribution from 20 intronic SNPs (Table 3.14 and Figure 3.11). The genetic architecture of this locus indicates two distinct LD blocks and the presence of a recombination hot spot (Figure 3.11A), which is also seen in the Hapmap CEU panel (Figure 3.11B).

Table 3.14. Genotyping summary of *EPHX1*.

Genes	Study SNPs (tested, passed)	SNP aliases
EPHX1	42 tested, 25 passed study criteria	rs1051740 = coding 612 T>C, Y113H, increased warfarin dose requirement?; rs2234922 = coding 691 A>G, H139R, unknown function
Coding ns	2: rs1051740 (Y113H), rs2234922 (H139R)	
Coding s	2 : rs2292566, rs1051741	
UTR	0	
Intronic	18: rs2854461, rs2854447, rs2854450, rs2854451, rs3753658, rs3753659, rs3753660, rs3753661, rs2671272, rs3738047, rs2671270, rs3817268, rs2260863, rs2740170, rs4149223, rs2292567, rs2671266	
Flanking	3: rs4653436, rs3753663, rs2102663, rs6426089	

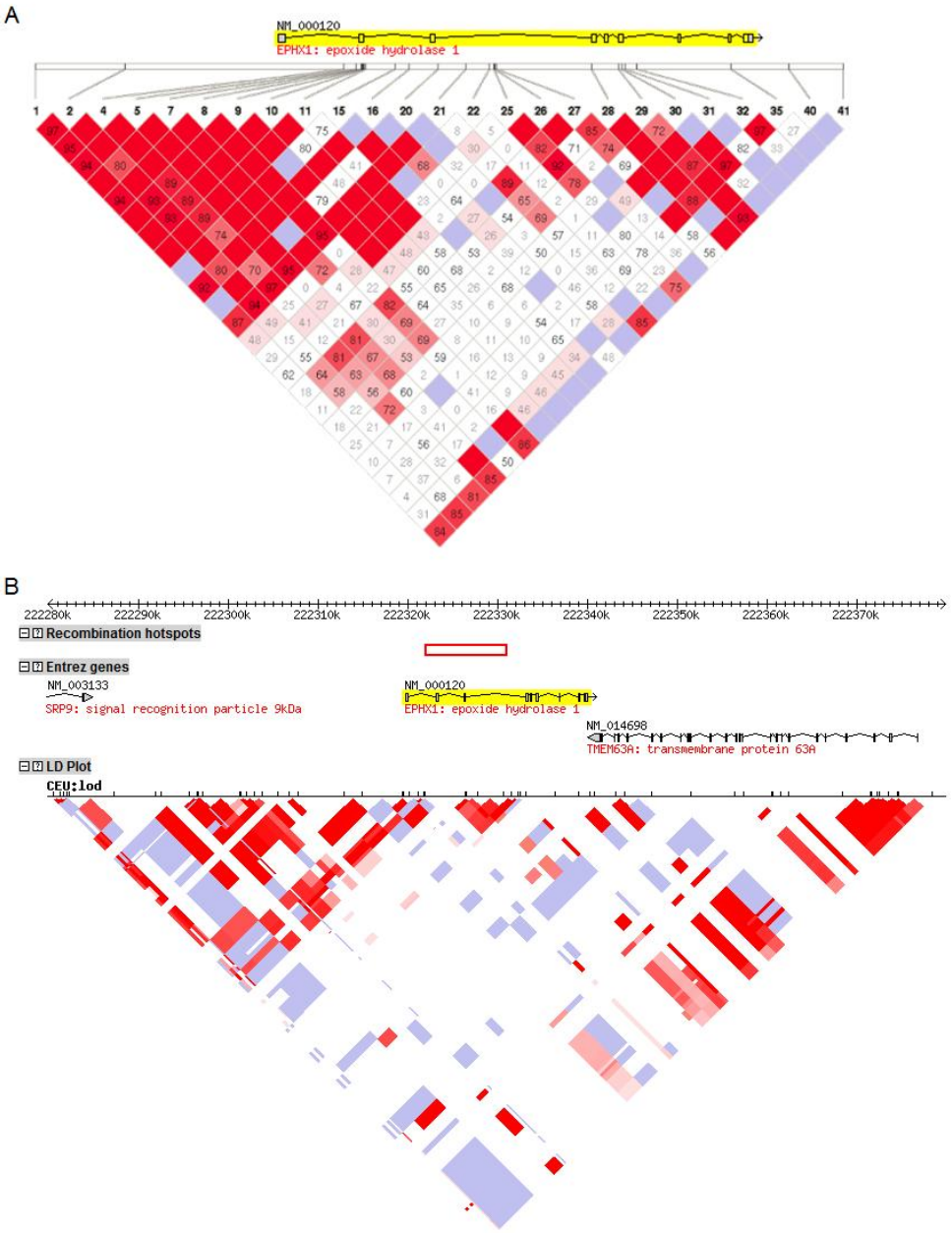


Figure 3.11. LD plots for *EPHX1* region in (A) Uppsala study and (B) Hapmap CEU panel.

3.5.4 NQO1

A few studies have shown that the antioxidant enzyme nicotine adenine dinucleotide phosphate dehydrogenase (NAD(P)H), which is also called flavoprotein DT-diaphorase (NQO1, NMOR1), reduces the quinine form of vitamin K and is of minor importance under physiologic conditions (Wallin et al. 1978). It may play a minor role in the warfarin cycle, but it was included in our study of warfarin dose.

Table 3.15 shows the genotyping summary of *NQO1*. The intronic SNP rs1437135 has been suggested to be associated with protein C level (Peyvandi et al. 2004), However, this evidence is not very strong and requires further validation. SNP rs1800566, a C-to-T substitution at position 609 of NQO1 cDNA, codes for a proline-to-serine change at residue 187 which is associated with absence of activity (Traver et al. 1992; Traver et al. 1997). This SNP is polymorphic in the Swedish population (MAF is 0.172). All SNPs passing study criteria were in LD which is consistent with the CEU Hapmap data (Figure 3.12).

Table 3.15. Genotyping summary of *NQO1*.

Genes	Study SNPs (tested, passed)	SNP aliases
NQO1	18 tested, 9 passed study criteria	rs1437135 = hCV2091258,
Coding ns	1: rs1800566 (P187S)	intronic 2515 C>T, associated
Coding s	1: rs689453	with protein C levels
UTR	0	
Intronic	6: rs2917669, rs2917671, rs1437135, rs689452, rs2965753, rs7186002	
Flanking	1: rs689456	

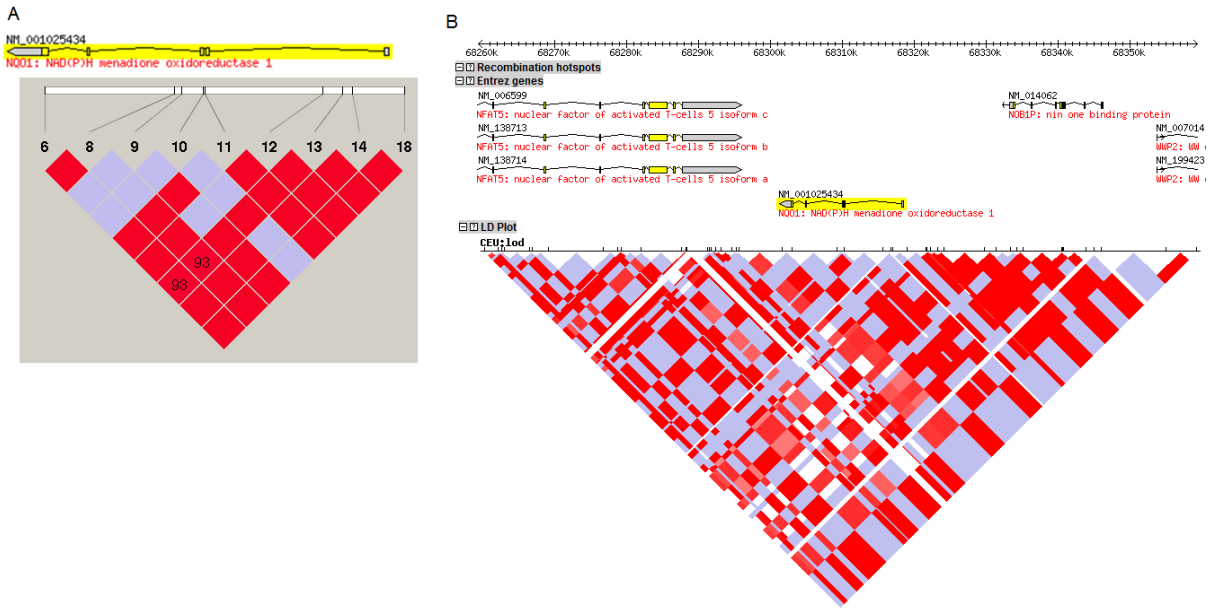


Figure 3.12. Genetic architecture of the *NQO1* locus in (A) Swedish and (B) Hapmap Caucasians. All the SNPs passing the study criteria are in LD (panel A) which is also observed in Hapmap (panel B). *NQO1* is located in the centre of a 400 kb LD block.

3.5.5 GGCX

The dietary vitamin K is an essential cofactor for the activation of coagulation proteins; which are classified as vitamin K-dependent (VKD) proteins and mediated by gamma-glutamyl carboxylase (GGCX) (Rost et al. 2004b; Wu et al. 1997) encoded by *GGCX* on chromosome 2 (Table 3.16). Gamma-glutamyl carboxylase is located in the ER membrane, spatially close to vitamin K epoxide reductase, and forms a complex with VKOR and calumenin (CALU) in order to efficiently recycle the oxidative form of vitamin K (Wajih et al. 2004).

A very rare autosomal recessive bleeding disorder has been reported with combined deficiency of the vitamin K-dependent coagulation factors II, VII, XI, and X, and proteins C, S, and Z resulting from the identification of mutations in *GGCX* (Brenner et al. 1998; Rost et

al. 2004b). The nsSNP rs699664 (C/T) which results in an arginine to glutamine substitution, has an, as yet, unknown functional impact (Table 3.16). Interestingly, the C allele is minor in Africans (MAF = ~0.30) but is major in Caucasians (MAF = 0.58) and Asians (MAF = 0.68) suggesting potential selection. Since warfarin dose requirement differs among different ethnic population, rs699664 may be of interest due to its MAF variation in different population. Table 3.16 shows the SNPs as well as a CAA microsatellite repeats that passed study criteria in *GGCX*.

Table 3.16. Genotyping summary of *GGCX*.

Genes	Study SNPs (tested, passed)	SNP aliases
GGCX	17 tested, 10 passed study criteria	rs699664 (C/T) = coding 8762
Coding ns	1: rs699664 (R325Q)	G>A (formerly 8016 G>A) or 1002
Coding s	1: rs2592551	G>A, R325Q, function unknown
UTR	0	
Intronic	6: rs7568458, rs12714145, rs6738645, rs762684, rs2028898, (CAA) microsatellite	
Flanking	2: rs6547621, rs7605975	

Six alleles of 10, 11, 13, 14, 15 and 16 CAA repeats were identified in *GGCX* intron 6 whereas only 10-, 11- and 13-repeats alleles were reported in Japanese (Shikata et al. 2004). All SNPs in *GGCX* are in LD. For illustration purposes the genomic structure of *GGCX* and LD architecture of this locus in the Swedish population is discussed alongside the results of the association analysis in chapter 4, section 4.3.3. Hapmap results show that *GGCX* is located in the centre of a 70 kb LD block.

Figure 3.13. Genomic architecture of GGCX region in Hapmap CEU panel. GGCX is located in the centre of a 70 kb LD block.

3.5.6 CALU

Calumenin is an ER chaperone protein, encoded by *CALU* on chromosome 7 (Table 3.11), which is associated with gamma-carboxylase. Recent studies have shown that calumenin binds to both VKOR and gamma-carboxylase and acts as an inhibitory protein (Wajih et al. 2004; Wallin et al. 2001). An RNAi experiment showed that the enzymatic activity of gamma-carboxylase increases five times by interfering with CALU mRNA transcription (Wajih et al. 2004). Experiments in rats indicated that over-expression of calumenin prevents

warfarin from interacting with VKOR (Wallin et al. 2001). This protects the rat from rodenticide made of high dose of warfarin. It is suggested that abundant calumenin associated to VKOR would change the conformation and prevent the affinity of warfarin targeting. However, the molecular mechanism of the warfarin resistance is still unclear.

A non-synonymous SNP (rs2290228) in exon 2 has been reported to be associated with warfarin dose (Vecsler et al. 2006), which is also polymorphic in our Swedish patients with an MAF = 0.144 (Table 3.17).

Table 3.17. Genotyping summary of *CALU*.

Genes	Study SNPs (tested, passed)	SNP aliases
CALU	14 tested, 9 passed study criteria	rs2290228 = coding 11 G>A, R4Q;
Coding ns	2: rs2290228 (R4Q), rs2307040 (A82V)	
Coding s	0	
UTR	2: rs11653, rs8597	
Intronic	4: rs2060717, rs339054, rs1006023, rs339098	
Flanking	1: rs339057	

The *CALU* gene is spanned by a single LD block in the Swedish population and Hapmap Caucasians (Figure 3.14). This LD block also contains the gene encoding opsin 1 (*OPN1SW*).

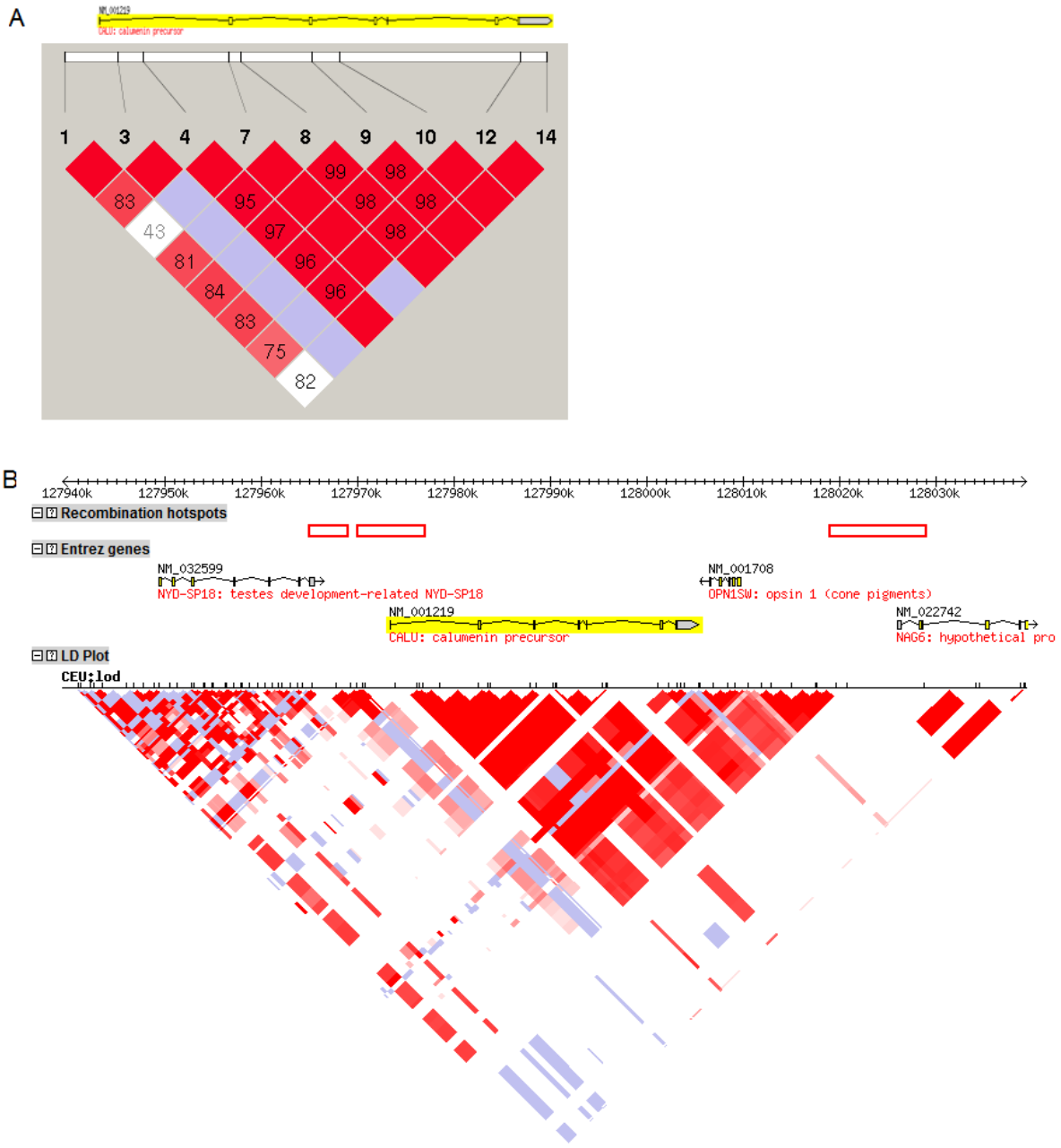


Figure 3.14. Genomic architecture of *CALU* in (A) Swedish and (B) Hapmap CEU. The Swedish LD indicates the SNPs that passed the study criteria are in LD. *CALU* and another gene, *OPN1SW*, are located in a small LD block whereas *CALU* spans across two third of this region.

3.5.7 Protein Disulfide Isomerase (PDI)

Although VKORC1 has been identified as a target for warfarin (Li et al. 2004; Rost et al.

2004a) and the involvement of its thioredoxin-like CXXC centre in the reduction of vitamin K1 2,3-epoxide (Vit.K>O) has also been validated (Wajih et al. 2005), the cellular system providing electrons to the centre was unknown. In 2007, a study using *in vitro* inhibition assay and immunoprecipitation experiment showed that a protein disulfide isomerase (PDI) peptide is tightly associated with VKORC1 and form VKOR complex (Wajih et al. 2007). Wajih and his colleagues further concluded that the energy required for gamma-carboxylation of proteins is provided by dithiol-dependent oxidative protein folding by PDI in the ER. Their proposed model present the PDI as another subunit of the complex in the ER which provides the electrons for the reduction of the thioredoxin-like CXXC centre in VKORC1 (Figure 3.15).

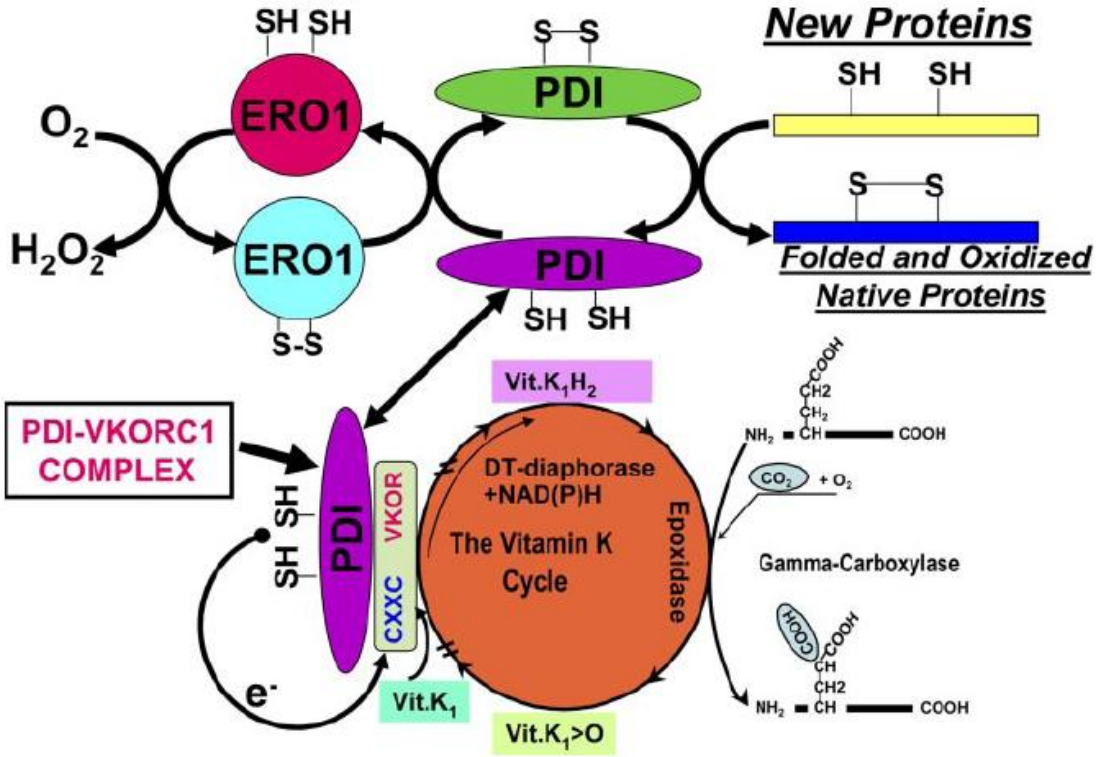


Figure 3.15. Hypothetical model of protein disulfide isomerase functionality. This figure is reproduced from J. Biol. Chem., 2007, 282(4): 2626-2635 (Wajih et al. 2007)

According to the experimental results, it was also proposed that part of the PDI peptide pool in the ER becomes strongly linked to part of the VKORC1 pool and that the complex is the active warfarin-sensitive Vit.K>O reducing enzyme complex of the vitamin K cycle. Since PDI is found in abundance in all eukaryotes and is fairly uniform in its main features, this proposed model also explains the VKOR activity is inhibitory in transfection expression study in various cell lines including insect cells (Li et al. 2004).

In 2006, a study suggested that oxidation/isomerisation by PDI switches tissue factor from a non-coagulant to an active molecule (Ahamed et al. 2006). Later, Versteeg and Ruf reported that tissue factor (TF) mediated coagulation is significantly enhanced and factor X activation is 5-10 times increased with presence of PDI (Versteeg and Ruf 2007). Although Wajih et al (Wajih et al. 2007) described the use of siRNA to inhibit disulfide isomerase activity, no precise DNA or protein sequence was reported. In Homo *sapiens*, there are six genes known as protein disulfide isomerases. In the absence of any further information all known *PDI* genes were tested for association to warfarin-induced bleeding and, in parallel, warfarin dose requirement. Table 3.18 lists a total of 31 SNPs that passed study criteria. The six PDI genes were selected in a later phase of the project. SNPs genotyped in Hapmap were chosen, except for *P4HB* where only one SNP was available with a further two SNPs within 200 kb flanking region. SNPs in *P4HB* were selected from the Ensembl SNP database. Hapmap CEU result indicates that *PDIA5* spans across three LD blocks, i.e. two recombination hot spots were observed in the gene. However, due to selection of tag SNPs, our result in the Swedish population is not reflecting the LD structure properly (Figure 3.17). LD structures of the six PDIs in the Swedish sample are shown in Figure 3.17.

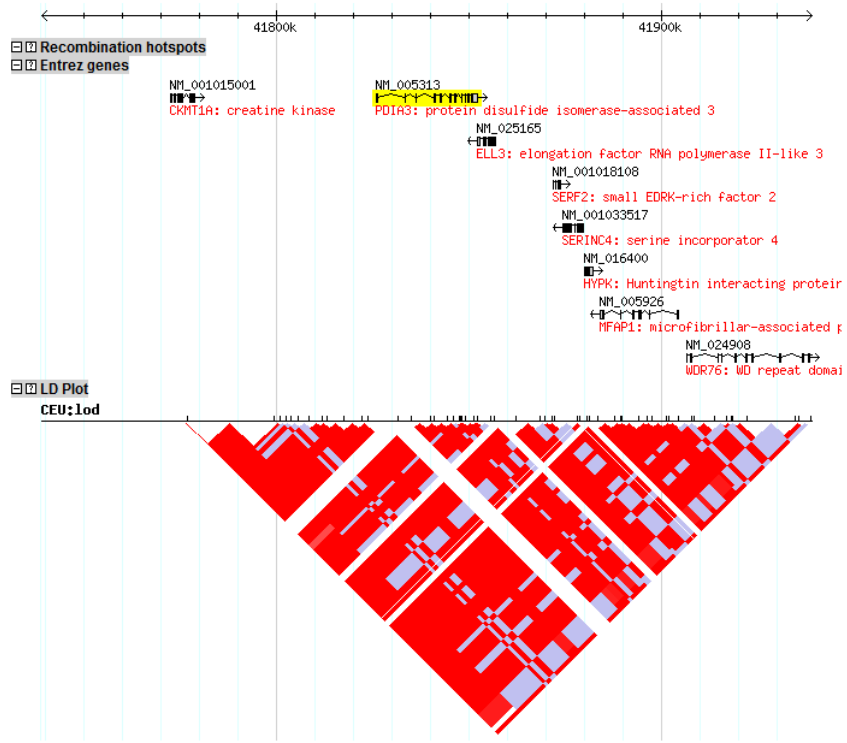
Table 3.18. Genotyping summary of *PDI* family.

Genes	Study SNPs (tested, passed)	SNP aliases
P4HB	6 tested, 5 passed study criteria	-
Coding ns	0	
Coding s	2: rs1130674, rs2070871	
UTR	0	
Intronic	3:rs876017, rs1533756, rs1010954	
Flanking	0	
PDIA2	4 tested, 2 passed study criteria	rs2685127, function unknown; rs400037 (R388Q), function unknown
Coding ns	2: rs2685127 (T286M), rs400037 (R388Q)	
Coding s	0	
UTR	0	
Intronic	0	
Flanking	0	
PDIA3	4 tested, 4 passed study criteria	-
Coding ns	0	
Coding s	0	
UTR	0	
Intronic	4: rs10163054, rs8040336, rs11070411, rs7175032	
Flanking	0	
PDIA4	8 tested, 7 passed study criteria	-
Coding ns	0	
Coding s	0	
UTR	0	
Intronic	7: rs10085877, rs4727005, rs10272564, rs10269104, rs6464929, rs1551927, rs6464930	
Flanking	0	
PDIA5	6 tested, 6 passed study criteria	-
Coding ns	0	
Coding s	0	
UTR	0	
Intronic	6: rs1078982, rs3792366, rs4677875, rs702030, rs836832, rs1107377	
Flanking	0	
PDIA6	7 tested, 7 passed study criteria	-
Coding ns	0	
Coding s	0	
UTR	0	
Intronic	6: rs1198873, rs11904084, rs1686447, rs1734343, rs1734346, rs12471762	
Flanking	1: rs1686482	

A


180k 190k 200k 210k 220k 230k 240k 250k 260k 270k 280k 290k 300k 310k 320k 330k 340k 350k 360k 370k
Recombination hotspots
Entrez genes
NM_201412
LUC7L: LUC7-like
NM_183337
RGS11: regulator of G-protein signalling 11
NM_032039
C16orf9: chromosome 16 open reading frame 9
NM_006428
MRPL28: mitoch
NM_001176
ARHGDIG: Rho GDP dissociation inhibitor (GDI) gamma
NM_006849
PDIA2: protein disulfide isomerase-associated 2
NM_181050
AXIN1: axin 1
LD Plot
CEU:lod

B


41800k
41900k
Recombination hotspots
Entrez genes
NM_001015001
CKMT1A: creatine kinase
NM_005313
PDIA3: protein disulfide isomerase-associated 3
NM_025165
ELL3: elongation factor RNA polymerase II-like 3
NM_001018108
SERF2: small EDRK-rich factor 2
NM_001033517
SERINC4: serine incorporator 4
NM_016400
HYPK: Huntingtin interacting proteir
NM_005926
MFAP1: microfibrillar-associated p
NM_024908
WDR76: WD repeat domai
LD Plot
CEU:lod

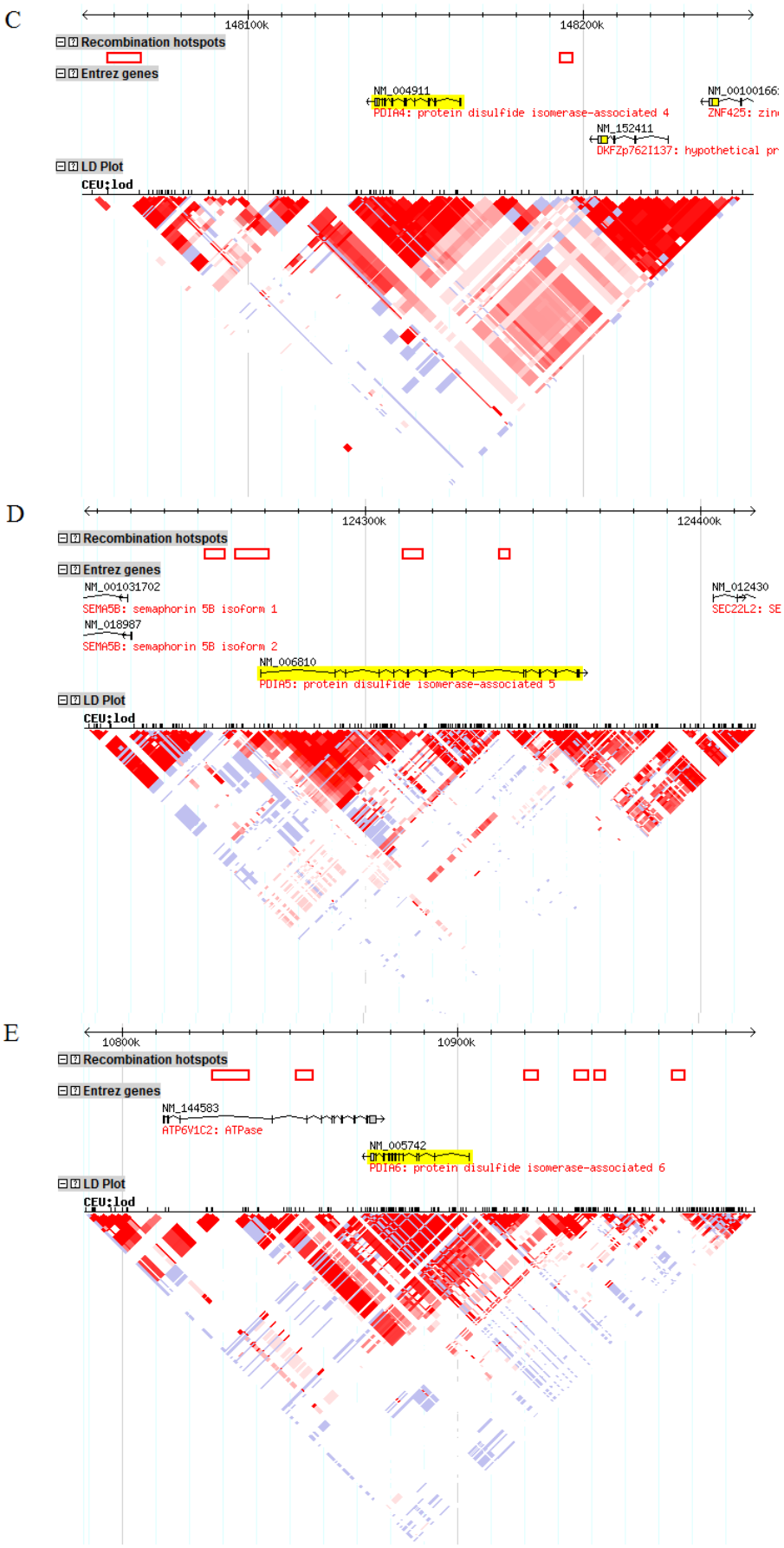


Figure 3.16. Genomic architectures of (A) *PDIA2*, (B)*PDIA3*, (C) *PDIA4*, (D) *PDIA5* and (E) *PDIA6* in Hapmap. Within 200 kb flanking near *P4HB*, only two SNPs were genotyped in Hapmap.

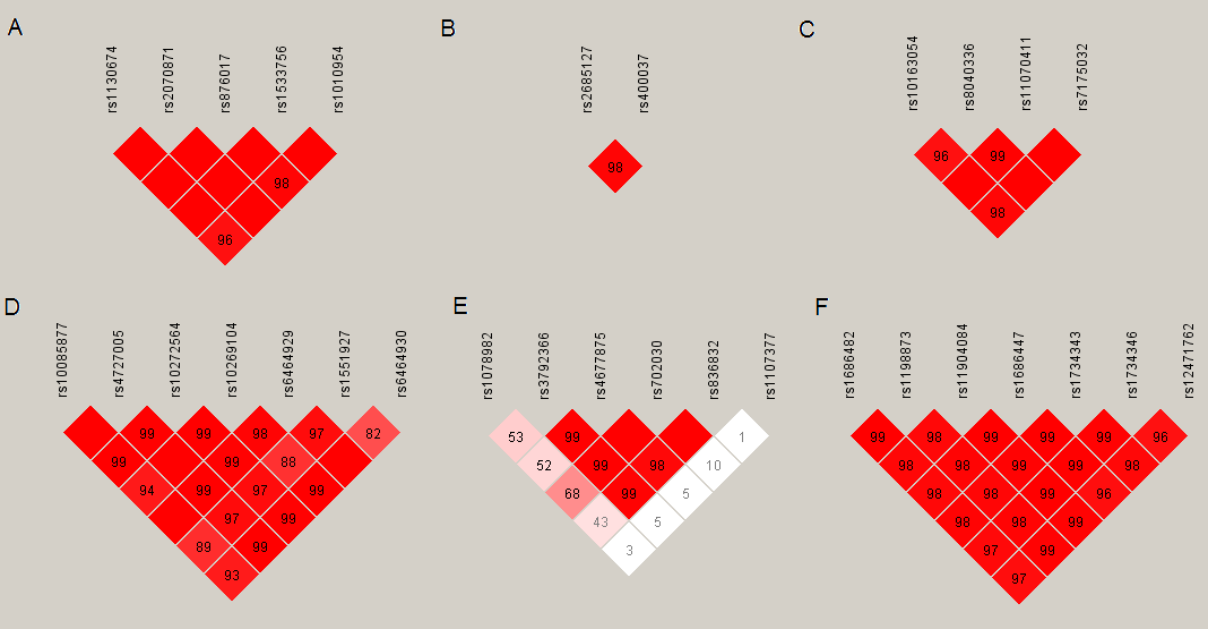


Figure 3.17. LD structures of six genes in PDI family A. (A) *P4HB*, (B) *PDIA2*, (C) *PDIA3*, (D) *PDIA4*, (E) *PDIA5* and (F) *PDIA6*.

3.6 Vitamin K Dependent Proteins

The role of many vitamin K-dependent (VKD) proteins has been investigated in warfarin sensitivity and, at the start of this study, several had been postulated to be implicated (see below). The main VKD proteins are clotting factors II (prothrombin), VII, IX and X, proteins C, S and Z and growth-arrest-specific protein 6, encoded by *F2, F7, F9, F10, PROC, PROS1, PROZ and GAS6* (Berkner 2000; Berkner and Runge 2004). Table 3.19 shows the genomic information for each candidate VKD gene. *F7*, *F10*, and *PROZ* are located back to back in chromosome 13:112808124-112874700 whilst *GAS6* is located 672 kb telomeric to this cluster.

Table 3.19. Candidate genes of vitamin K dependent.

Gene	Location	Exons	Transcript	Protein (residues)	Function of protein
F2	Chr 11: 46697331 - 46717631 bp	14	1997 bps	622	Converts fibrinogen to fibrin, activates FV, FVIII, FXIII, protein C
F7	Chr 13: 112808124 - 112822348 bp	9	2459 bps	466	Is converted to FVIIa and then converts FIX to FIXa and FX to FXa
F9	Chr X: 138340437 - 138373137 bp	8	2780 bps	461	Makes a complex with FVIIIa and then converts FX to its active form
F10	Chr 13: 112825128 - 112851846 bp	8	1524 bps	488	Converts FII to FIIa in the presence of factor Va.
PROC	Chr 2: 127892246 - 127903048 bp	9	1756 bps	461	Activated protein C counteracts coagulation together with protein S by inactivating FVa and VIIIa
PROS1	Chr 3: 95074647-95175395 bp	15	3275 bps	676	Cofactor to protein C that degrades coagulation factors Va and VIIIa
PROZ	Chr 13: 112860971 - 112874700 bp	8	1488 bps	400	Is together with protein Z-dependent protease inhibitor a cofactor for the inactivation of FXa
GAS6	Chr 13: 113546903 - 113590421	15	2499 bps	678	Participates in many processes, e.g. potentiation of agonist-induced platelet aggregation

3.6.1 Gene cluster on chromosome 13 (F7, F10, PROZ)

Factor VII also known as Hageman factor, is a serine protease encoded by the *F7* gene on chromosome 13. In both intrinsic and extrinsic pathways, factor VII is the first to be activated in trauma (extrinsic pathway) and initiate the coagulation cascade (Figure 3.18). A few studies have reported promoter polymorphisms in *F7* to have an effect on warfarin sensitivity (Aquilante et al. 2006; D'Ambrosio et al. 2004; Shikata et al. 2004).

Factor X is also known as Stuart-Prower factor or as thrombokinase and is equipped with serine endopeptidase activity. It is encoded by the *F10* gene located adjacent to *F7*. Studies of promoter variants and a synonymous coding polymorphism in exon 7 of *F10* have reported no effect on warfarin sensitivity (Aquilante et al. 2006; Shikata et al. 2004). However, factor X plays an important role as the first member of the final common pathway (also called thrombin pathway) (Figure 3.18).

Protein Z, encoded by *PROZ*, is a glycoprotein and assists hemostasis by binding thrombin and promoting its association with phospholipid vesicles. It also accelerates 1000-fold of the degradation of factor X which is primarily done by protein Z-related protease inhibitor (ZPI). Mutation in ZPI has been reported to be associated with an increased susceptibility to venous embolisms (Van de Water et al. 2004).

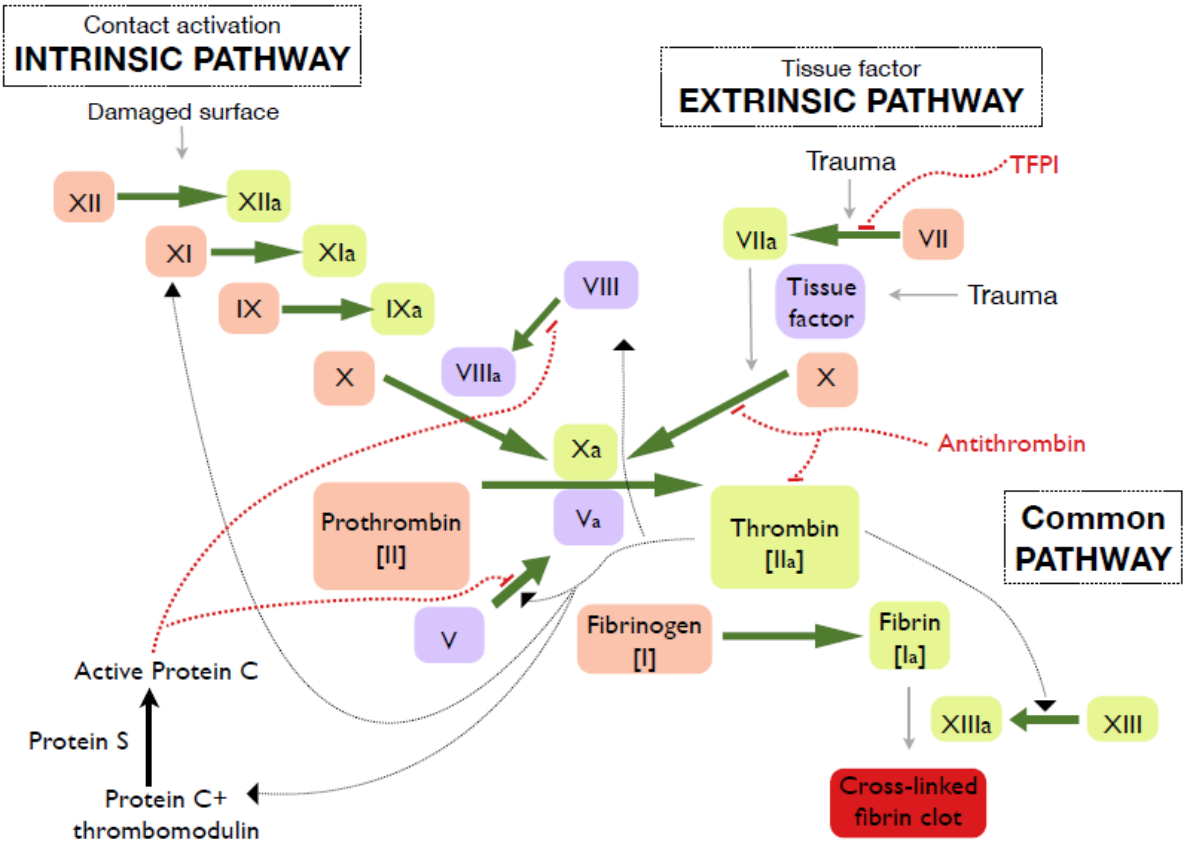


Figure 3.18. The coagulation cascade. (This figure also appears in chapter 1 as Figure 1.3, modified from http://en.wikipedia.org/wiki/Coagulation).

As described previously, *F7*, *F10*, and *PROZ* are sitting back to back on chromosome 13. In this 80 kb region genotype results were obtained for 39 SNPs (Table 3.20). An nsSNP in *F7* (rs6046, R413Q) has been previously suggested to give rise to a peptide with lower activity in blood coagulation. LD analysis shows that *F7* and the first exon of *F10* are located in the same LD block, whereas the last two exons of *F10* sit in the same LD block with *PROZ* (Figure 3.19A). It is evident that there is a recombination hot spot near the intron 6 of *F10* whereas exons 2, 3, 4 and 5 of *F10* appear to form a sub haplotype block suggesting the presence of a weak recombination hot spot.

The result from Hapmap CEU panel replicates our finding in 201 Swedish patients (Figure 3.19B). Two recombination hot spots were observed in *F10*. The Hapmap result further suggests a genomic instability in the region including *F7* and *F10* whereas *PROZ* is located in a distinct LD block (Figure 3.19B).

The Online Mendelian Inheritance in Man (OMIM) database reports 24 different mutations in *F7* of which 22 are associated with deficiency or decrease enzymatic activity and two with a decreased susceptibility to myocardial infarction. In *F10*, 14 mutations are reported to result in decreased activity or deficiency and which cause mild or severe bleeding. In contrast to *F7* and *F10*, *PROZ* has not been reported to harbour any variants associated with disease state.

Table 3.20. Genotyping summary of *F7*, *F10* and *PROZ* cluster.

Genes	Study SNPs (tested, passed)	SNP aliases
F7	32 tested, 11 passed study criteria	rs6046 = coding 1289 G>A, R413Q (formerly 1238 G>A, R353Q), low activity allele
Coding ns	1: rs6046 (R413Q)	
Coding s	0	
UTR	1: rs2476324	
Intronic	6: rs2774030, rs491098, rs493833, rs569557, rs488703, rs6041	
Flanking	3: rs3093229, rs3093230, rs3093233	
F10	45 tested, 15 passed study criteria	rs5960 = synonymous coding T264T, function unknown
Coding ns	0	
Coding s	1: rs5960	
UTR	0	
Intronic	12: rs473598, rs776897, rs3211770, rs2026160, rs2251102, rs3211764, rs2480946, rs693335, rs483949, rs485798, rs776905, rs474810	
Flanking	2: rs563964, rs3093261	
PROZ	27 tested, 13 passed study criteria	rs3024711 = rs17878660, intronic
Coding ns	0	
Coding s	0	
UTR	0	
Intronic	11: rs3024764, rs3024747, rs3024746, rs17881956, rs3024743, rs17886440, rs3024731, rs2480948, rs513479, rs3024718, rs3024711	
Flanking	2: rs2273971, rs7335409	

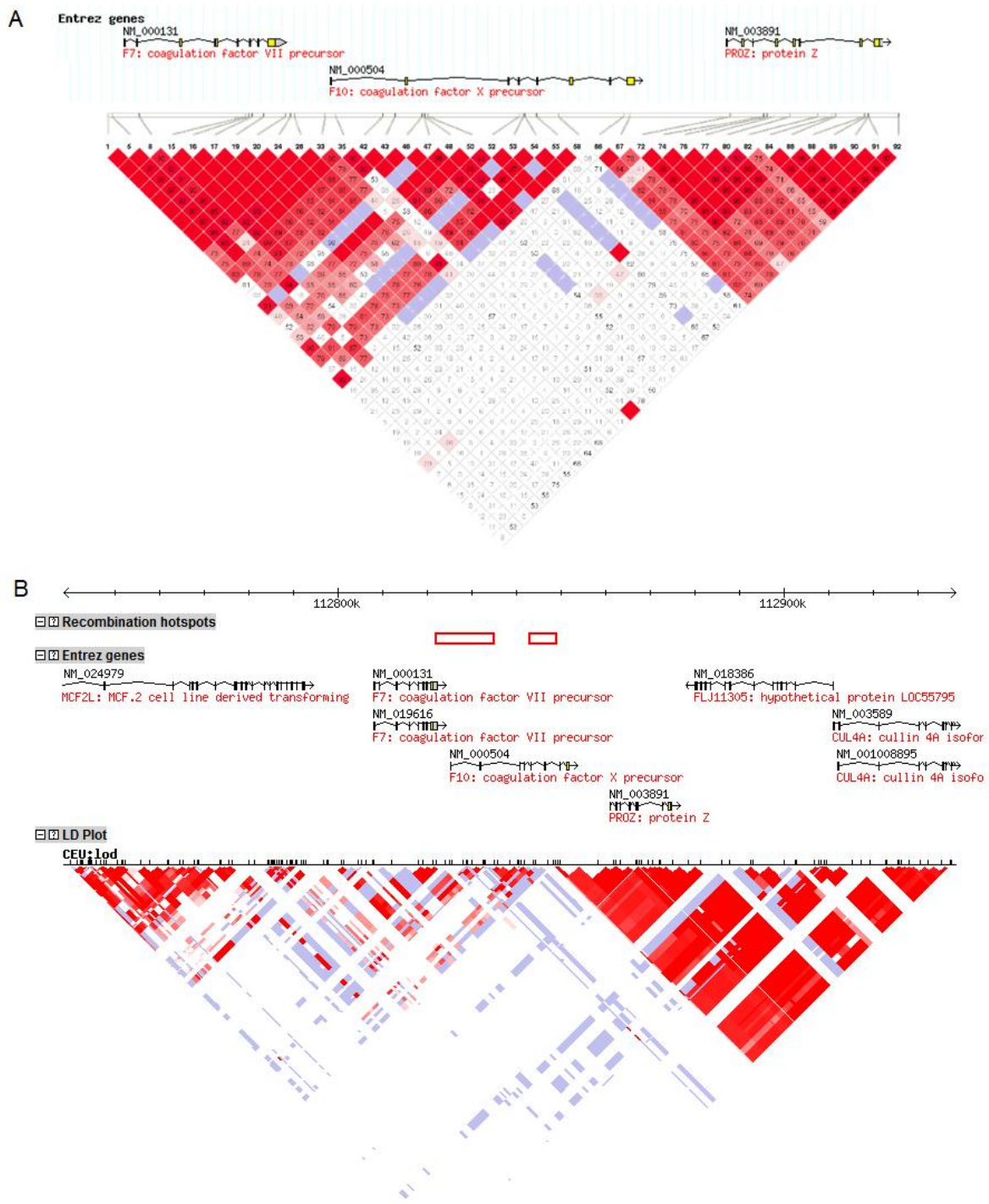


Figure 3.19. Genomic architecture and LD organisation of the chromosome 13 locus harbouring *F7*, *F10* and *PROZ* in (A) Swedish and (B) Hapmap Caucasians.

3.6.2 Other VKD genes (F2, F9, PROC, PROS1 and GAS6)

Beside the chromosome 13 cluster, the following VKD proteins were selected and genotyped as candidate genes: coagulation factor II and IX which are involved in the coagulation cascade; protein C and Z which act as natural anticoagulants by inhibiting factor V and VIII, and finally growth arrest-specific gene 6 which activates platelet formation (Table 3.21).

Table 3.21. Genotyping summary of other VKD candidate genes.

Genes	Study SNPs (tested, passed)	SNP aliases
F2	22 tested, 10 passed study criteria	rs5896 = coding 525 C>T (formerly 494 C>T), T165M, warfarin sensitivity?
Coding ns	1: rs5896 (T165M)	
Coding s	1: rs5898	
UTR	0	
Intronic	8: rs2070850, rs3136435, rs3136447, rs2070851, rs2070852, rs3136460, rs2282687, rs3136516	
Flanking	0	
F9	29 tested, 11 passed study criteria	Chr X: 138344709: coding G>A, A-10T; Chr X: 138344710: coding C>T, A-10V; Both are rare causes of warfarin sensitivity; These mutations were not found in the study
Coding ns	1: rs6048 (T194A)	
Coding s	0	
UTR	1: rs440051	
Intronic	7: rs401597, rs392959, rs398101, rs422187, rs413957, rs110583, rs413536	
Flanking	1: rs434447, rs445691	
PROC	25 tested, 13 passed study criteria	rs1799809 = upstream -1644 A>G (formerly -1641), GG lower protein C activity
Coding ns	0	
Coding s	1: rs5936	
UTR	0	
Intronic	10: rs2069910, rs2069915, rs2069916, rs2069919, rs2069921, rs973760, rs1518759, rs2069924, rs2069928, rs2069931	
Flanking	2: rs2069901, rs1799809	
PROS1	50 tested, 11 passed study criteria	-
Coding ns	0	
Coding s	0	
UTR	0	
Intronic	8: rs8178592, rs5013930, rs8178607, rs8178610, rs4857343, rs8178623, rs4857037, rs8178649	
Flanking	3: rs7650230, rs9713061, rs9683303	
GAS6	5 tested, 4 passed study criteria	-
Coding ns	0	
Coding s	0	
UTR	0	
Intronic	3: rs9577874, rs9604573, rs6602908	
Flanking	1: rs7997328	

1. Coagulation

Prothrombin (factor II) is encoded by *F2*. The activated form (thrombin, factor IIa) converts soluble fibrinogen into insoluble fibrin and activates and further amplifies coagulation cascade by activating factor XI, VIII, and V (Figure 3.18). Over-presence of thrombin will also activate protein C for anticoagulation. Two independent studies have shown that a polymorphism in F2 causing a change from threonine to methionine at residue 165 leads to increased sensitivity to warfarin (D'Ambrosio et al. 2004; Shikata et al. 2004) whereas a third study could not replicate this finding (Aquilante et al. 2006).

Factor IX is also a serine protease and activates factor X in the intrinsic pathway. To date, 110 mutations found in the *F9* gene, which is located on the long arm of the X chromosome, have been registered in the OMIM database. Factor IX deficiency causes Hemophilia B which occurs in the general population at one-sixth the incidence of Hemophilia A, affecting approximately 1 in 30,000 male births. Mutations in the propeptide of factor IX, amino acid substitution change from alanine to valine or threonine at residue 10, lead to a rapid drop of factor IX during warfarin treatment resulting in warfarin sensitivity and thereafter bleeding in rare cases (Kristensen 2002).

The LD result of *F2* and *F9* is shown in Figure 3.20. Our Swedish results indicated both genes are well located in LD blocks and this is confirmed by the Hapmap CEU result. Additionally, Hapmap CEU results indicate that the LD block, where F2 is located, is bigger than 1 Mb whereas F9 is located at the end of a 150 kb block.

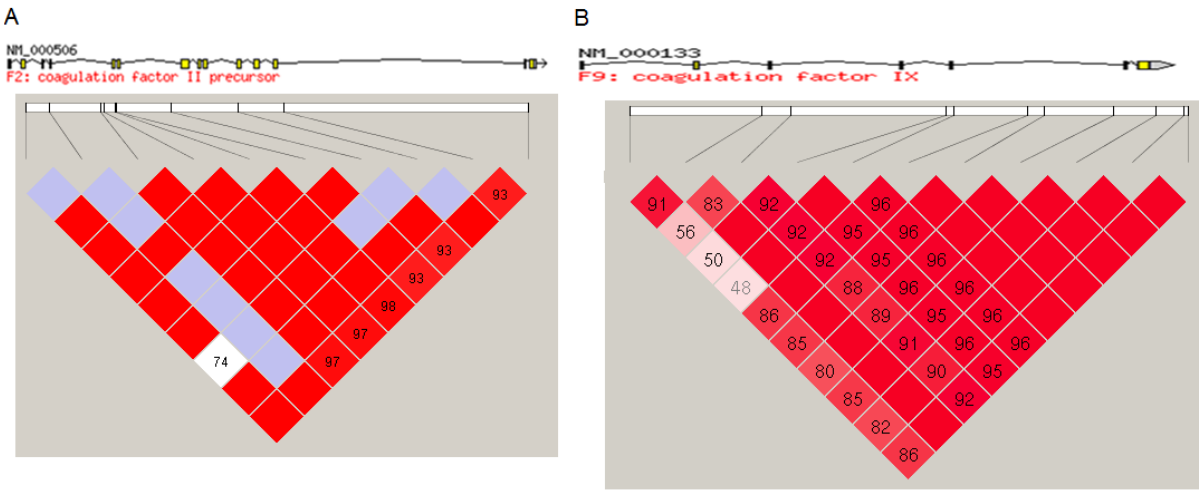


Figure 3.20. LD architecture of (A) F2 and (B) F9.

2. Anticoagulation

Protein C, encoded by *PROC* on chromosome 2, is a VKD serine protease that regulates blood coagulation by inactivating factors Va and VIIIa in the presence of calcium ions and phospholipids (Berkner 2000; Dahlback 2005). Protein S is a VKD plasma glycoprotein synthesized in the liver and is encoded by *PROS1* on chromosome 3. In anticoagulation, it functions as a cofactor to the activated protein C (APC) in the degradation of coagulation factors Va and VIIIa (Figure 3.18) and helps to prevent coagulation and stimulates fibrinolysis (Berkner 2000; Dahlback 2005).

Deficiency of protein C and S has been reported to predispose to thrombophilia (Engesser et al. 1987; Hasstedt et al. 1998; Matsuda et al. 1988). Protein C and S levels decline more rapidly than other VKD proteins resulting in a poor antithrombotic efficacy in the beginning of warfarin treatment (Vigano et al. 1984; Weiss et al. 1987). Warfarin-treated patients with a hereditary deficiency of protein C or S might result in warfarin-induced hypercoagulation (Chan et al. 2000; McGehee et al. 1984) and skin necrosis (Conlan et al. 1988). SNPs that

passed study criteria were evenly distributed alongside both genes (Figure 3.21). Both genes are in the same LD block; which is in agreement with Hapmap.

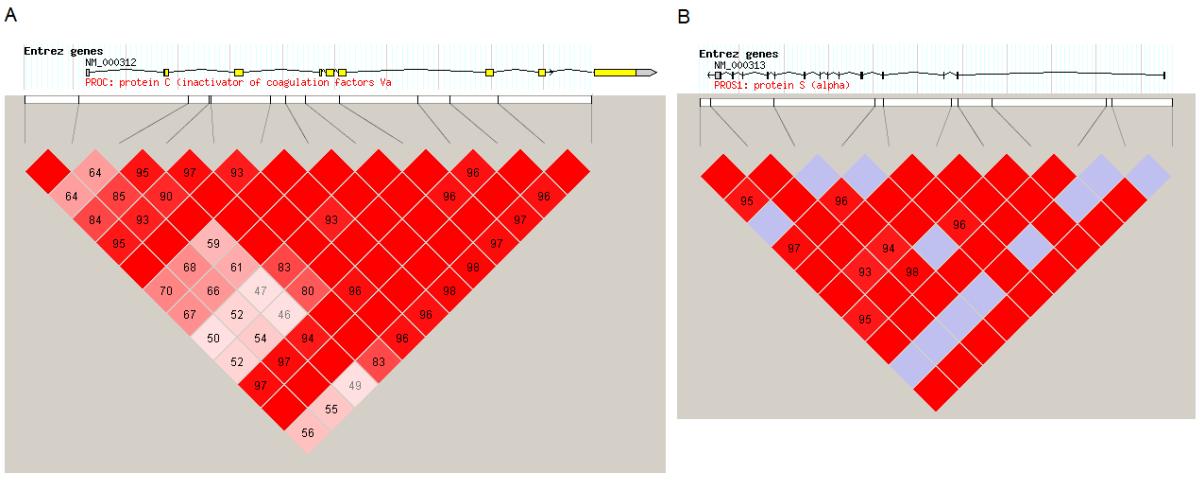


Figure 3.21. Genomic architecture of protein C and protein S in Swedish.

3. Platelet Formation

Growth arrest-specific gene 6 (*GAS6*) was first identified in 1993 with a 44% sequence similarity to coagulation protein S, and to which it is structurally related but without the anticoagulant activity (Manfioletti et al. 1993). Like protein S, GAS6 is also a vitamin K-dependent protein and is secreted from platelet A-granules and, it has been suggested to be involved in platelet activation and thrombus development (Angelillo-Scherrer et al. 2001; Gould et al. 2005).

Five SNPs in *GAS6* were reported to be significantly associated with atherothrombotic disease in a group of 110 healthy controls and 188 patients (Munoz et al. 2004). Together with being a VKD protein, *GAS6* may be associated with warfarin dose.

We obtained 4 SNPs that passed study criteria (Table 3.21) and of those 3 are in strong LD (Figure 3.22A). Hapmap CEU panel also indicate the presence of LD breaks in this 200 kb region. Three recombination hot spots are located in *GAS6* and 30 kb flanking region of both ends.

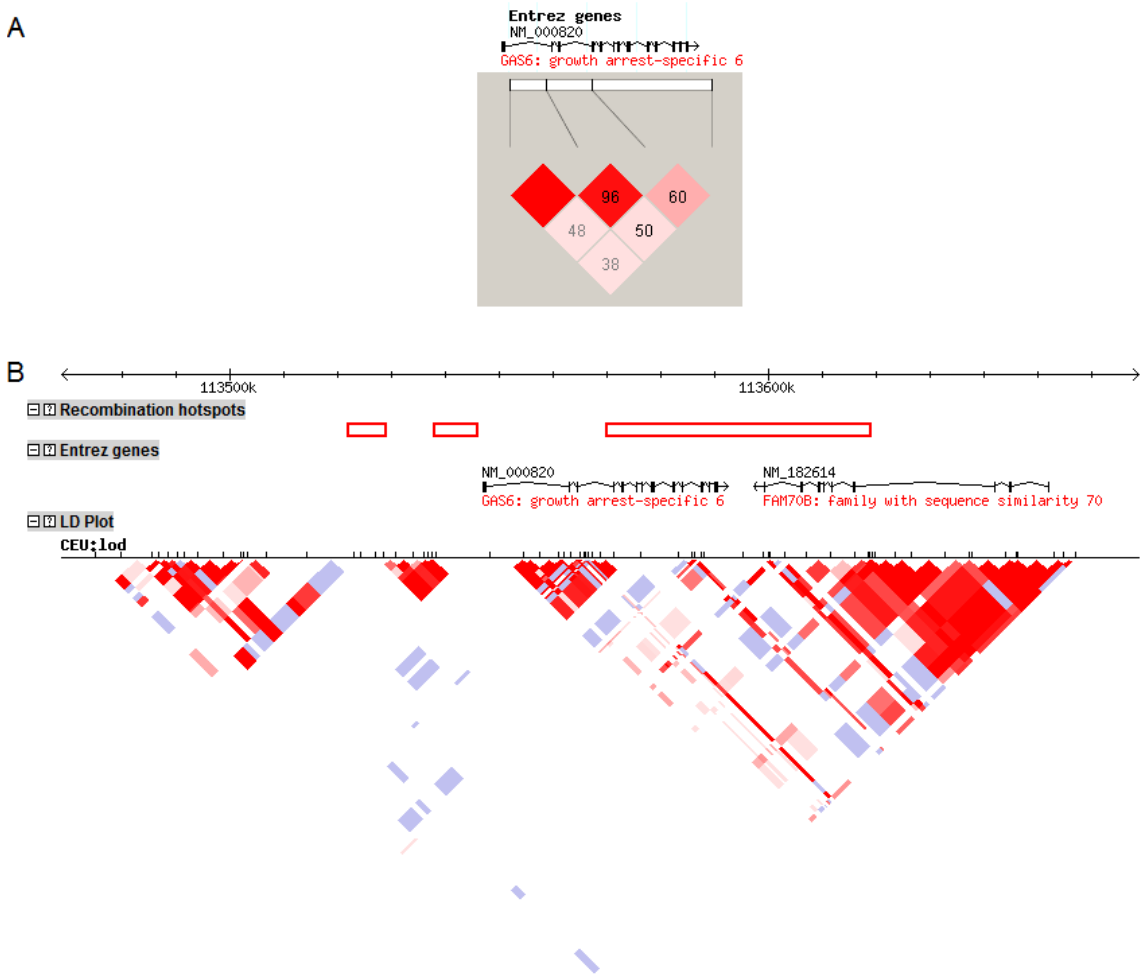


Figure 3.22. Genomic structure of growth arrest-specific gene 6 (GAS6) peptide in (A) Swedish and (B) Hapmap CEU panel.

3.7 OTHER COAGULATION FACTORS

Two other genes, which are important in the coagulation cascade, were selected (Table 3.22). The first one is antithrombin-III precursor encoded by *SERPINC1* on chromosome 1 which inhibits factors II, IX, X, XI and XII. The second one is factor V encoded by *F5* also located on chromosome 1 which activates prothrombin (factor II). A total of 97 SNPs were tested and 50 SNPs passed the study criteria (Table 3.23). Results of seven non-synonymous SNPs in factor V were obtained including FV Leiden mutation which confers increased risk of thrombosis (Larsen et al. 1998).

Table 3.22. Other candidate genes in this study.

Gene	Location	Exons	Transcript	Protein (residues)	Function of protein
SERPINC1	Chr 1: 170604596 - 170618130 bp	7	1559 bps	464	Inhibits FIIa, FIXa, Xa, XIa and XIIa. Anti-thrombin deficiency increases risk of thrombosis
F5	Chr 1: 166215067 - 166287379 bp	25	6914 bps	2224	A cofactor that activates FII together with FXa. A F5 mutation leads to risk of thrombosis

Table 3.23. Genotyping summary of antithrombin III and factor V.

Genes	Study SNPs (tested, passed)	SNP aliases
SERPINC1	23 tested, 9 passed study criteria	-
Coding ns	0	
Coding s	1: rs5878	
UTR	0	
Intronic	6: rs2227590, rs2227593, rs2227594, rs2227607, rs5877, rs2759328	
Flanking	2: rs2227588, rs2146372	
F5	74 tested, 41 passed study criteria	rs6025 = coding 1698G>A, R534Q (formerly R506Q), FV Leiden increased risk of thrombosis
Coding ns	7: rs6033 (M413T), rs6025 (R534Q), rs6018 (N817T), rs4525 (H865R), rs6032 (K925E), rs6030 (M1764V), rs6027 (D2222G)	
Coding s	9: rs6028, rs6029, rs6022, rs6035, rs6015, rs6036, rs6037, rs6024, rs6021	
UTR	0	
Intronic	23 : rs3753305, rs9332504, rs2298905, rs2298908, rs2236870, rs3766121, rs3766120, rs3766119, rs1894702, rs6012, rs3766117, rs1894699, rs6427198, rs721161, rs2298909, rs3766110, rs1557572, rs9332618, rs9332629, rs2213867, rs2213866, rs2227244, rs966751	
Flanking	2 : rs2269648, rs2187952	

3.7.1 Anti-thrombin III

Anti-thrombin III is a serine protease inhibitor in the plasma that regulates the intrinsic coagulation pathway. The inhibition reaction traps the protease including thrombin and activated form of factor IX, X, XI, and XII, by preventing the enzymatic site from specific substrates (Olson and Bjork 1994). Anti-thrombin III deficiency is a rare inheritable disease and patients normally suffer recurrent venous thrombosis and pulmonary embolism. Both the congenital form caused by mutations in *SERPINC1* and the acquired form may result in a hyper-coagulable scenario during warfarin induction (Chan et al. 2000; Dahlback 2005). Analysis of the nine SNPs that passed study criteria shows that *SERPINC1* spans a single haplotype block (Figure 3.23). Hapmap indicates that this block is more than 1 Mb in length.

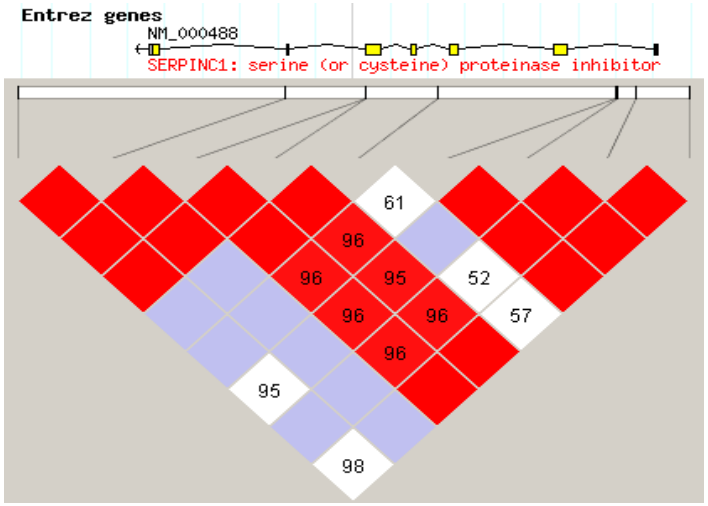


Figure 3.23. Genomic architecture of *SERPINC1*in Swedish.

3.7.2 Coagulation factor V

Factor V is a stimulatory protein without enzymatic activity which circulates in the plasma and binds to activated platelets. Factor V is activated by thrombin (FII) and the activated form of factor V reciprocally activates prothrombin to thrombin (active form) for blood

clotting. Mutations in F5 usually cause either the factor V deficiency or resistance to activated protein C. Deficiency of factor V leads to predisposition for mild or severe hemorrhage in recessive action (Guasch et al. 1998; Lindqvist et al. 1998; van Wijk et al. 2001). Meanwhile, mutations like factor V Leiden (Arg506Gln), commonly cause thromboembolism (Larsen et al. 1998) resulting a stubborn factor V protein which is resistant to cleavage by activated protein C: This keeps it constantly active and increases the rate of thrombin formation. To date, *F5* mutations are not known to be associated with warfarin treatment. *F5* also harbours an intragenic recombination hot spot (Figure 3.24).

Our result is in agreement with Hapmap CEU which also reports that LD breaks in introns 6 and 7 (Figure 3.25A). The first seven exons encode for 372 amino residues, whilst the whole peptide is 2224 residues long. This recombination hot spot is consistently observed in Han-Chinese (Figure 3.25B), Japanese (Figure 3.25C) and African from Yoruba.

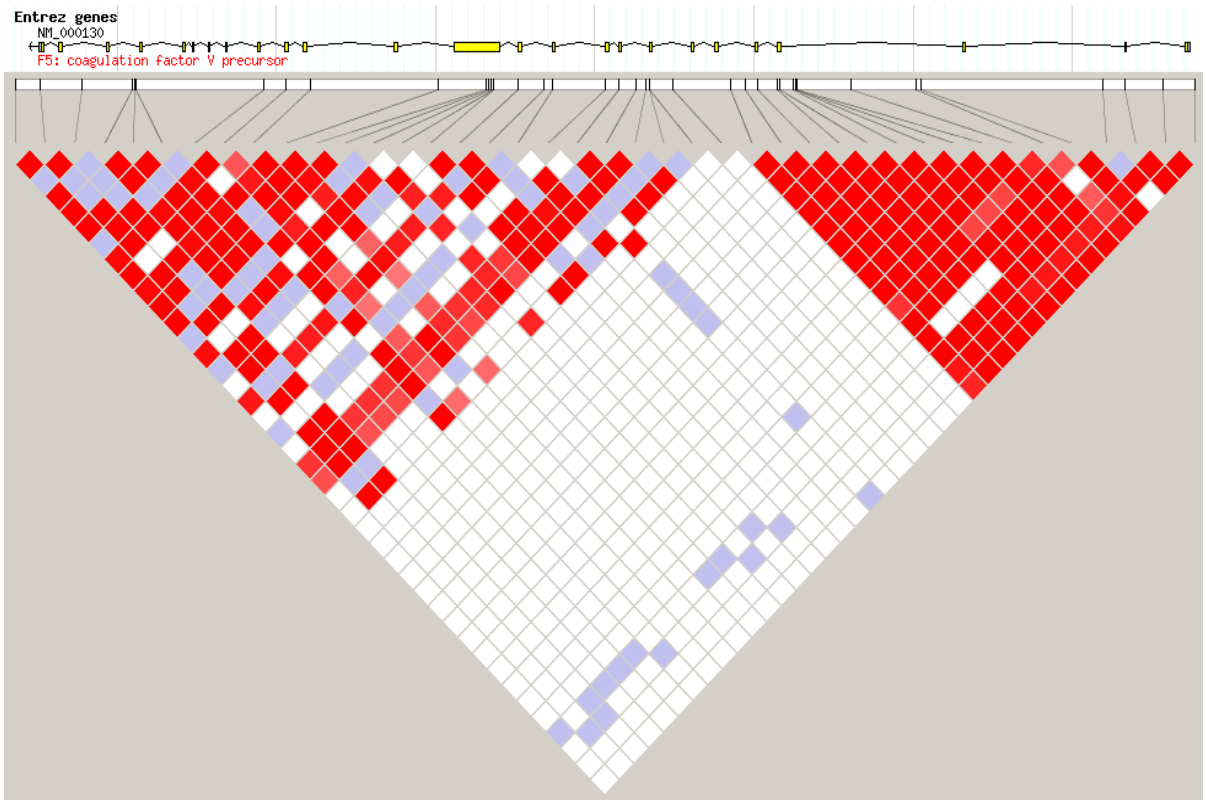


Figure 3.24. Genomic architecture of factor V in Swedish.

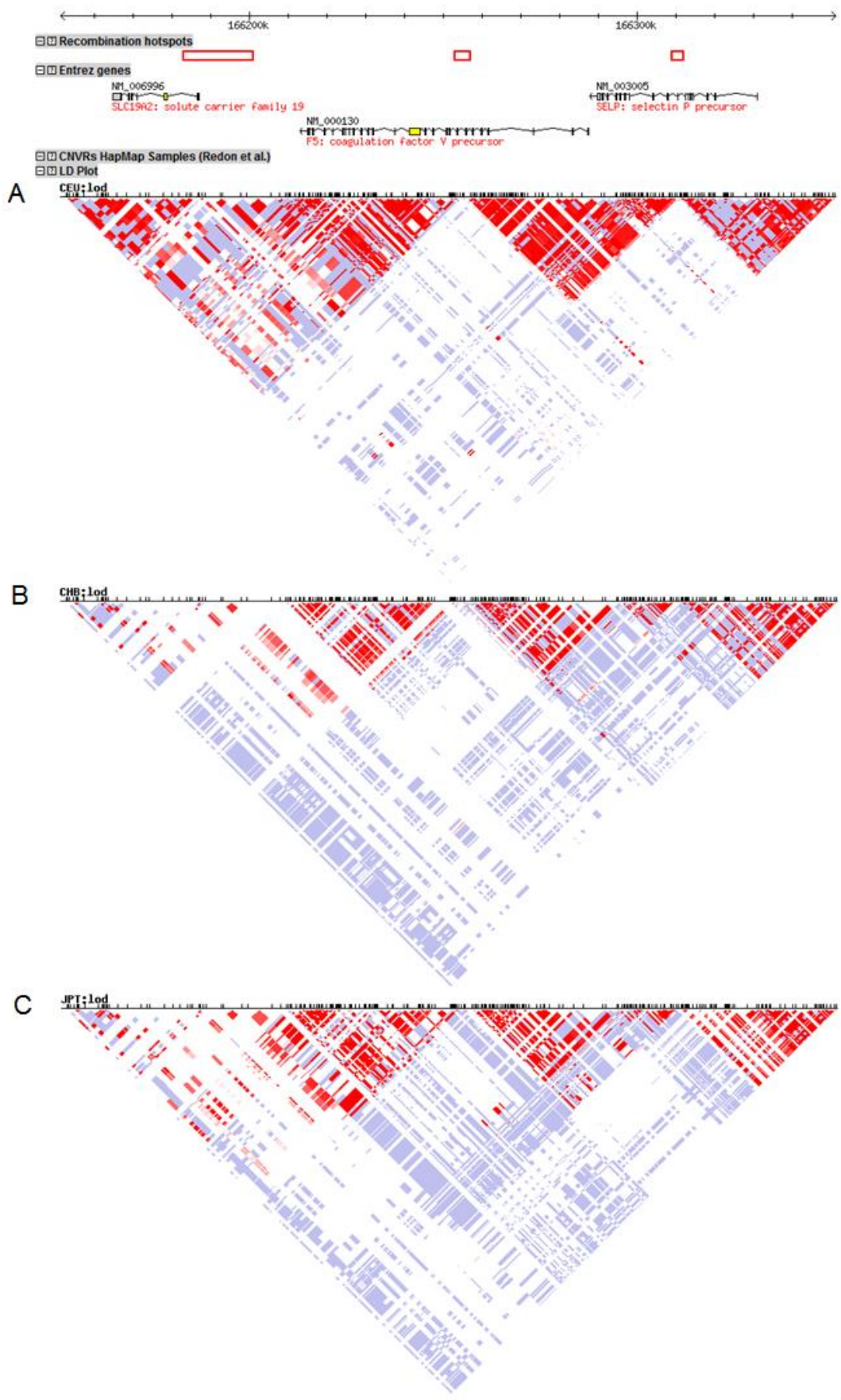


Figure 3.25. Genomic architecture of F5 region in Hapmap (A) Caucasian, (B) Han-Chinese and (C) Japanese.

Most of FV mutations have been reported in Caucasians. By comparing the Hapmap result of Caucasian (CEU), Han-Chinese (CHB) and Japanese (JPT), the LD blocks in Caucasian separate distinctly whereas marginal LD was observed between blocks in Han-Chinese and Japanese. This might suggest more severe genomic instability near the recombination hot spot within F5, which results in a higher mutation rate observed in the CEU panel.

3.8 CONCLUSION

We carefully selected 35 candidate genes that we believe have potential effect in inter-individual variation of warfarin dose and in bleeding complications of warfarin treatment. Except for genes we selected in a later phase (*APOE, GAS6* and *PDI* family) and for which we mainly used tag SNPs from Hapmap, we constructed comprehensive LD maps in our Swedish cohort(s) for each of the candidate genes. Not surprisingly, the genetic structure in the Swedish population is in accordance with that seen in the Hapmap CEU panel which reflects northern and western European ancestry.

In the first phase of this project, a lot of effort was put into validation for the SNPs deposited in public databases. In this regard, the Hapmap project provides a resource of information in genetic variation, and in later phases of the project SNPs validated by Hapmap were preferentially selected to test our patients.

All the SNPs that passed study criteria were analysed for association to warfarin dose in the 201 patients of Uppsala cohort (chapter IV). These SNPs were also used in the selection of tag SNPs to analyse warfarin-treated patients from the WARG study (chapter V), as well as the association with warfarin-induced bleeding complication (chapter VI).

CHAPTER IV

AN INVESTIGATION OF GENETIC DETERMINANTS OF WARFARIN DOSE REQUIREMENT IN 201 SWEDISH PATIENTS (UPPSALA COHORT)

4.1 PATIENT RECRUITMENT

Collection and genetics analysis of subjects in the Uppsala cohort was approved by the local Ethics Committee in Uppsala, Sweden. After informed consent, Dr Mia Wadelius and colleagues (Uppsala University) collected blood samples from 201 patients who were treated with warfarin at the anticoagulation clinic at the Uppsala University Hospital in 2000. These patients were aged from 28 to 88 years old. At the time of recruitment, patients had been treated with warfarin for at least 2 months, and the frequency each patient visited the clinic varied from every week to every month, depending on medications. Their treating period varied from 2.4 months to 26 years and the median duration of treatment was 2 years.

Medical information for each patient was recorded at six consecutive visits in the clinic. For each visit, the administered dose and corresponding PT INR value was recorded in the database. The prescribed warfarin dose for these patients varied from 4.5 to 77.25 mg/week. Meanwhile, information about age, gender, body weight, and any other diseases and indication for concurrent treatment was extracted from the patients' medical records (Table 4.1). 113 (56%) patients in this study had atrial fibrillation. The second largest medical indication was those patients with heart valve prosthesis (49 individuals, 24.4%). These patients were sorted into a separate group because a higher PT INR value is targeted. The average PT INR of majority patients (153 patients) was within 2 and 3 while 13 patients had the average INR below 2 and 35 patients had the average INR above 3. Among the total of 1005 patient-visits, 686 (68.3%) of measured INRs were between 2 and 3.

Most patients had other diseases including hypertension (78 patients), heart failure (51 patients), angina pectoris (35 patients) and type 2 diabetes (18 patients) and received

concurrent medications.

Table 4.1. Medical information statistics of patients in Uppsala cohort.

Characteristics		Patients	(%)	Effects on warfarin
Indication				
	Atrial fibrillation	113	56.2	
	Heart valve prosthesis	49	24.4	
	Deep vein thrombosis/pulmonary embolus	9	4.5	
	Cardiomyopathy	8	4.0	
	Transischemic attack	5	2.5	
Other diseases				
	Hypertension	78	38.8	
	Heart failure	51	25.4	
	Angina pectoris	35	17.4	
	Type 2 diabetes mellitus	18	9.0	
Interacting medication				
	Simvastatin	25	12.4	+
	Aspirin	21	10.4	+
	Paracetamol	18	9.0	+
	Amiodarone	9	4.5	+
	Disopyramide	7	3.5	+
	Dextropropoxyphene	7	3.5	+
	Propafenone	3	1.5	+
	Carbamazepine	3	1.5	-
	Non-steroidal anti-inflammatory drug	2	1.0	+
	Phenytoin	1	0.5	-
	Mianserin	1	0.5	-
Gender				
	Men	135	67.2	
	Women	66	32.8	
Age				
	Mean years (range)	66.9	22-88	

As to the concurrent medications, the drugs were catalogued as interacting if they had moderate or major interactions with warfarin according to the database MICROMEDEXs Healthcare Series (http://www.micromedex.com/). In the 201 Uppsala patients, a total of 107 concurrent medications were known to influence warfarin (Table 4.1). Therefore, patients were divided into three groups for further statistical analysis: four individuals were simultaneously taking drugs that lower the effect of warfarin, 74 patients received

medications that enhanced the effect of warfarin and the 123 patients were not on any medications known to affect the metabolism of warfarin.

Ethnic origin is also an important factor in genetic studies and should be accounted for in order to avoid false positive associations due to confounding effects (Devlin et al. 2001). The patients were mainly Caucasians. Of the 201 patients, 194 were of Swedish origin; four of other European descent and three from the Middle East.

4.2 UNIVARIATE ASSOCIATION ANALYSIS

As described in chapter III 783 SNPs (1146 SNPs attempted) were successfully genotyped in the 35 candidate genes selected to test for association to warfarin dose requirement. To best fit the linear model, the association between SNPs and average weekly dose was analysed using the original value, log of dose and square root of dose. Square root demonstrated better distribution in linearity. Univariate regression analysis was then performed and Table 4.2 shows all the nominally associated SNPs ($P \leq 0.05$) with square root of dose ranked by p-value. Only 379 common SNPs ($MAF \geq 0.05$) were included in the analyses. In Table 4.2, N is the number of successfully genotyped individuals for each SNP, whereas univariate R^2 indicates the ratio of inter-individual dose variation explained by that SNP in a regression model. The LD (r^2) between the SNP with the lowest P-value and any other in the gene or gene cluster (*CYP2C9*, *CYP2C19*, and *CYP2C18*) was also calculated and is shown in the last column (Table 4.2).

The 33 SNPs nominally associated with warfarin dose ($P \leq 0.05$) were found in *VKORC1, CYP2C9-CYP2C19-CYP2C18, PROC, EPHX1*, *CALU*, *PDIA2*, *GGCX*, *ORM1-2* and *APOE*. Notable, the two polymorphisms (rs429358 and rs7412) in *APOE* were used to discriminate haplotypes between APOE*E2, *E3 and *E4 allele, and the association of the dose and *APOE* is significant only when the haplotypes were assessed as E2 + E4 versus E3.

Table 4.2. Univariate analysis in regression model.

Gene	SNP	MAF	N	Univariate R^2	P-value	r^2 with best SNP
VKORC1	rs9923231	0.391	181	0.317	$1.91\times10^{-15**}$	-
	rs2359612	0.389	200	0.29	$2.30\times10^{-15**}$	0.968
	rs9934438	0.383	169	0.292	$3.59\times10^{-13**}$	1
	rs7294	0.384	188	0.208	$4.14\times10^{-10**}$	0.385
	rs4889490	0.446	199	0.16	$3.821\times10^{-8**}$	0.461
	rs4889537	0.372	199	0.142	$3.158\times10^{-7**}$	0.209
	rs4889599	0.366	194	0.124	$3.270\times10^{-6**}$	0.305
	rs8046978	0.214	197	0.047	0.00906	0.173
	rs11642603	0.093	192	0.027	0.02304	0.07
	rs11642466	0.103	195	0.025	0.02623	0.08
	rs7194347	0.343	197	0.032	0.04069	0.153
CYP2C9	rs1057910 (*3)	0.058	201	0.141	$2.784\times10^{-7**}$	-
	rs9332108	0.064	201	0.141	$2.784\times10^{-7**}$	0.89
	rs9325473	0.055	189	0.147	$3.753\times10^{-7**}$	0.908
	rs1057911	0.067	191	0.145	$4.218\times10^{-7**}$	0.89
	rs9332214	0.059	198	0.139	$4.654\times10^{-7**}$	0.878
	rs4917639	0.173	197	0.118	$4.944\times10^{-6**}$	0.276
	rs2860905	0.214	193	0.072	0.0008*	0.224
CYP2C19	rs3814637	0.059	195	0.106	0.00002**	0.838[a]
	rs17882687(*15)	0.08	183	0.044	0.00417*	0.395[a]
CYP2C18	rs7896133	0.056	193	0.074	0.00013	0.869[a]
PROC	rs2069919	0.372	182	0.09	0.00022*	-
	rs1799809	0.433	188	0.078	0.00055*	0.777
	rs2069901	0.441	177	0.072	0.00147*	0.785
	rs2069910	0.387	178	0.046	0.01678	0.414
APOE	rs429358+rs7412[b]	0.251	201	0.051	0.0057*	-
EPHX1	rs4653436	0.266	196	0.048	0.00848	-
CALU	rs11653	0.366	197	0.047	0.00944	-
	rs1006023	0.331	200	0.033	0.03789	0.865
	rs2307040	0.336	200	0.033	0.03811	0.867
	rs339054	0.461	195	0.032	0.04487	0.612
PDIA2	rs400037	0.232	183	0.020	0.02954	-
GGCX	rs12714145	0.408	198	0.034	0.0332	-
ORM1-2	rs1687390	0.062	149	0.026	0.04964	-

** Experiment-wise significance, significant p-value is 1.65E-04.

* Gene-wise significance, significant p-value depends on the effective number of test in each gene or gene cluster.

[a] Linkage disequilibrium with CYP2C9*3 (rs1057910)

[b] The two APOE SNPs are not significant individually, the significance derived from assessed as E2+E4 vs. E3

To minimise false positives it is important to correct p-values for performing multiple tests.

Bonferroni correction is the most frequent approach of correcting for multiple tests, but this

estimate is very conservative as it considers all tests as independent (Balding 2006); for example, tightly linked SNPs in terms of LD constitute one as opposed to multiple tests. The Bonferroni correction based on the effective number of independent traits (Bonferroni Meff) provides a simple way of estimation of significant p-value after multiple testing (Cheverud 2001; Li 2001; Nyholt 2004), which is close to that of the permutation procedure. However, multiple correction using permutation procedure is very computationally demanding. In the dose study, we therefore corrected the p-value using the Bonferroni Meff method. Since there is no evidence showing that the 35 selected genes have co-evolved, we assume that genes are independent of each other, and SNPs tightly linked to the tag SNPs are disregarded. The Meff estimation is obtained using Single Nucleotide Polymorphism Spectral Decomposition (SNPSpD, http://gump.qimr.edu.au/general/daleN/SNPSpD/). The sum of effective tests over all genes is 303 in this study and the experiment-wise significance is $P < 1.65E-04$. Meanwhile, the number of independent effective tests within each gene or gene cluster varied from 2 to 50, resulting in gene-wise corrected significance of $P < 0.025$ (two independent SNPs) to $P < 0.001$ (50 independent SNPs). After experiment-wise correction for multiple testing, only *VKORC1* (7 SNPs), *CYP2C9* (6 SNPs), *CYP2C18* (1 SNP) and *CYP2C19* (1 SNP) were associated with warfarin dose. *PROC* and *APOE*, however, passed the gene-wise correction but failed in the experiment-wise test (Table 4.2).

The two SNPs in *CYP2C19* and one SNP in *CYP2C18* (Table 4.2) are in strong linkage disequilibrium with *CYP2C9**3 allele (rs1057910); which is known to be associate with warfarin dose (Aithal et al. 1999; Higashi et al. 2002; Wadelius et al. 2005; Wadelius et al. 2004). The *VKORC1* and CYP2C cluster loci are discussed in detail in section 4.3 and 4.4.

Association analysis using 2- or 3-marker sliding-window haplotypes was also performed for

each candidate gene. In each sliding-window the haplotype was phased statistically according to alleles of each SNP. We found 13 genes to be nominally associated ($P < 0.05$) with warfarin dose, and Table 4.3 lists only the haplotypes that exhibited the lowest p-value in each candidate gene for association with the square root of dose. Compared to single marker analysis (Table 4.2), six additional genes, *PDIA6*, *F7*, *PDIA3*, *PROZ*, *F9* and *NR1I2*, were nominally associated. *EPHX1*, *CALU*, *PDIA2* and *ORM1-2* did not have a haplotype that reached nominal significance.

Haplotypes in *VKORC1*, *CYP2C9*, *CYP2C18* and *CYP2C19* were the most associated and showed experiment-wise significance as they did in single marker analysis. As with single marker analysis, *PROC* was also significant at the gene-wise level in haplotype analysis. Haplotypes giving experiment-wise significance in *CYP2C9*, *CYP2C18* and *CYP2C19* are associated with either *CYP2C9**2 or *CYP2C9**3 alleles. In *GGCX*, the CAA microsatellite (section 4.3.3) was taken into account with the SNP haplotype. Table 4.3 shows the result of two or three marker haplotype association in dose. The last column in Table 4.3 compares the p-value between single marker and haplotype analysis. Haplotypes in *VKORC1*, *PROC* and *ORM1-2* provide no additional information over single SNPs. The remaining ten genes in Table 4.3 show a slightly smaller p-value in haplotype analysis than in single marker analysis. The experiment-wise significant haplotypes in *CYP2C9*, *CYP2C18* and *CYP2C19* have one or more SNP tightly linked to either *CYP2C9**2, *CYP2C9**3 or both.

Table 4.3. Two or three marker haplotype association in dose variation.

Gene	Haplotype	P-value	smaller p-value
VKORC1	rs9934438-rs9923231	5.76E-15**	**No**
CYP2C9	rs9332214[a] -rs9332222[b]-rs2298037	4.86E-09**	Yes
CYP2C18	rs1926711-rs7919273[a]-rs10509675	3.47E-07**	Yes
CYP2C19	rs2860840-rs3814637[a]	2.08E-06**	Yes
PROC	rs2069919-rs2069921-rs973760	1.36E-03*	**No**
PDIA6	rs11904084-rs1686447	1.00E-02*	Yes
GGCX	Microsatellite-rs762684-rs6738645	1.78E-02	Yes
F7	rs3093229-rs3093233	2.42E-02	Yes
PDIA3	rs10163054-rs8040336	2.64E-02	Yes
PROZ	rs2273971-rs3024711	3.57E-02	Yes
F9	rs401597-rs392959	3.83E-02	Yes
NR1I2	rs2461818-rs7643645	3.93E-02	Yes
ORM1-2	rs1687390-rs3762055	4.93E-02	**No**

** Experiment-wise significance, significant p-value is 1.65E-04.
* Gene-wise significance, significant p-value depends on the effective number of test in each gene or gene cluster.
[a] Linkage disequilibrium with *CYP2C9**3 (rs1057910)
[b] Linkage disequilibrium with *CYP2C9**2 (rs1799853)

4.3 *VKORC1* AND *GGCX*

VKORC1 and GGCX are the two key enzymes in recycling vitamin K in our body. When *VKORC1*, the effect target for warfarin, was identified by Rost et al (Rost et al. 2004) and Li et al (Li et al. 2004), all publicly available SNPs (dbSNP 121) in this gene (including 5 kb up- and down-stream flanking regions) were selected. At that time, 29 *VKORC1* SNPs were examined; including the mutations reported by Rost et al. and 16 SNPs plus a microsatellite marker which was reported by Shikata et al (Shikata et al. 2004) on GGCX.

4.3.1 *VKORC1*

Out of the 29 tested SNPs, 20 passed the assay quality control with only five of them being polymorphic in the Uppsala cohort. This suggested that *VKROC1* does not tolerate much variation, but this may be also due to the lack of validation and allele frequency information in the SNP database. Sequencing efforts have been made by two independent studies to discover novel variants in *VKORC1* and flanking region (Geisen et al. 2005; Rieder et al. 2005). Geisen and colleagues sequenced 200 young healthy subjects (100 female and 100 male) on intragenic, 1.8 kb upstream and 1.5 kb downstream region and discovered four additional common SNPs with MAF ranging between 9.25% and 41.5%. Rieder and colleagues sequenced 186 warfarin-treated patients confirming 7 out of 8 SNPs reported by Geisen and colleagues. Meanwhile, Rieder et al also reported 3 additional SNPs in the 5' flanking region of *VKORC1*.which are in LD with the previously known common SNPs in this region.

Alleles of the five *VKORC1* SNPs co-vary significantly with warfarin dose according to the regression analysis. Four of them are extremely closely associated with dose, $P < 0.0001$. The fifth (rs11150606), which is located downstream of the gene and has a much lower MAF, shows a much less significant association, $P = 0.02$. All five SNPs are good predictors of warfarin maintenance dose, and explain 29 – 30% of inter-individual variability. Outside the above three SNPs (rs2359612 is the best predictor, Figure 4.1), the effect of rs7294 was evaluated in the regression model including rs2359612. The result showed that it explains a further 3% of the variance of warfarin maintenance dose.

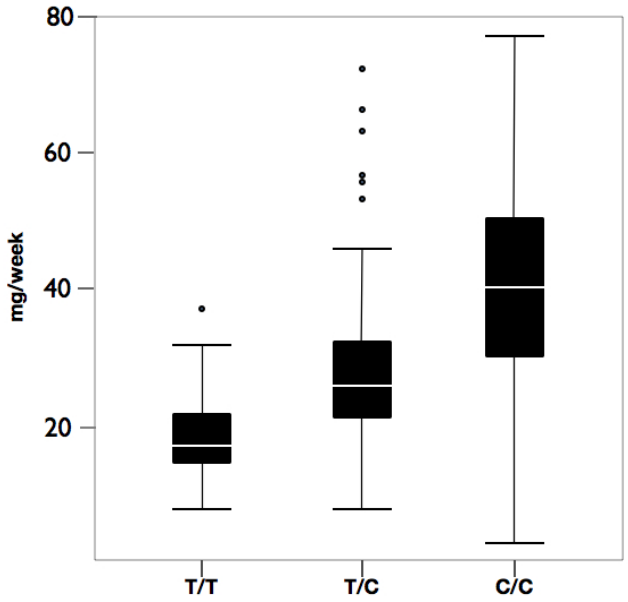


Figure 4.1. Mean weekly dose of rs2359612 genotype in 201 Swedish. The homozygous T/T is associated with low maintenance dose whereas C/C homozygotes need twice dose than T/T homozygotes.

Figure 4.2 shows the genomic structure, LD architecture, and common haplotypes of *VKORC1* and flanking regions. The arrow indicates the transcription from 5' to 3'. Four common SNPs have MAF of circa 40% and are located in the promoter region (rs9923231), the first intron (rs9934438), second intron (rs2359612) and 3' untranslated region (UTR)

(rs7294) of the gene. The fifth SNP (rs11150606) is located downstream of the gene and has a MAF of 4%. Inter SNP distances are 2.8, 1.1, 1.5 and 3.3 kb, respectively (Figure 4.2). We found that the four most common SNPs are in strong LD and give rise to three common haplotypes that are further subdivided into four by the more rare SNP rs11150606 (Figure 4.2). The first three SNPs are tightly linked, pairwise $r^2 = 1$.

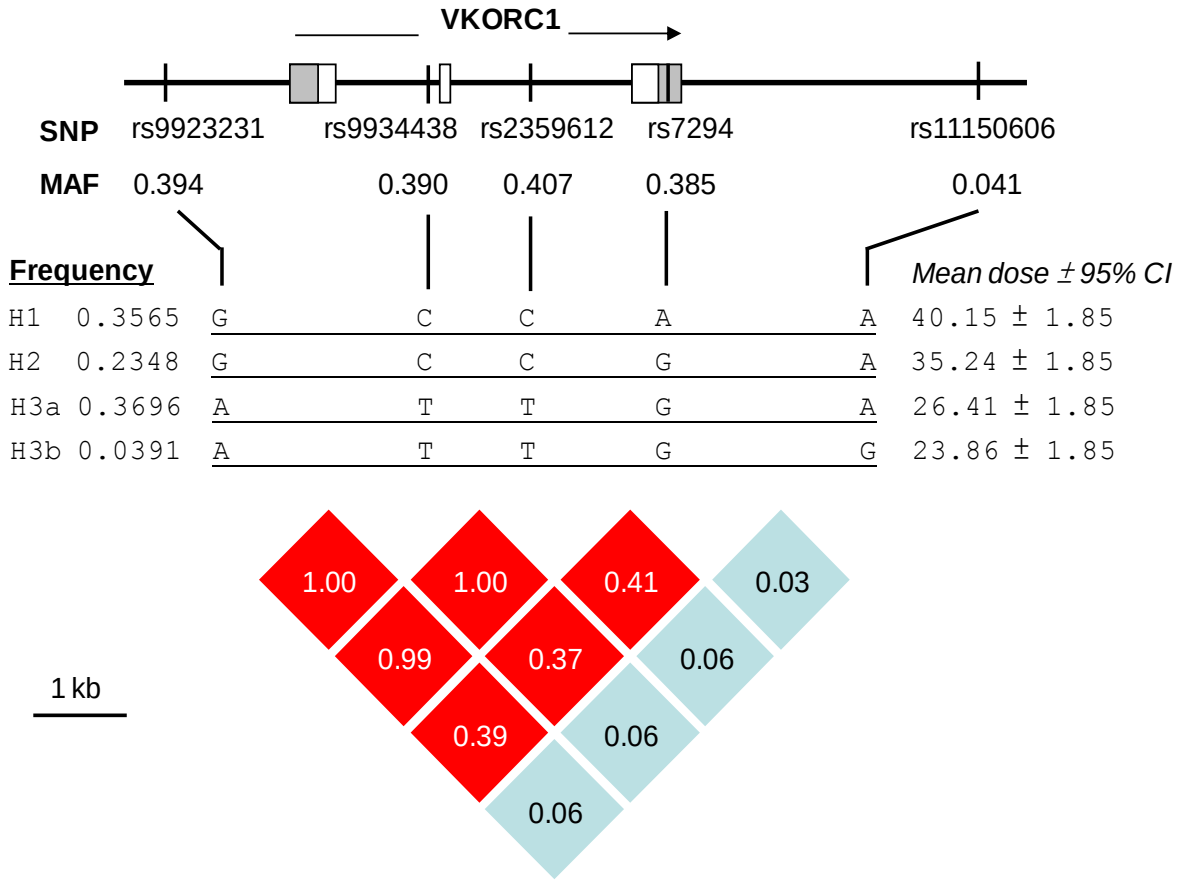


Figure 4.2. Genomic structure, haplotype, and linkage disequilibrium of *VKORC1*. Five SNPs were found to be polymorphic in Uppsala cohort and four of these are common (MAF ≥ 5%). The minor allele frequency (MAF) was noted under each SNP and haplotypes were listed (underline). The frequency of each haplotype and the corresponding mean weekly dose (and 95% confident interval) were calculated. The SNPs in 5' promoter region (rs9923231), intron 1 (rs9934438) and intron 2 (rs2359612) are tightly linked and only two haplotypes (GCC and ATT) were phased. Pairwise LD is represented in red (strong LD) and light blue (weak LD) diamonds with pairwise r^2 shown.

To characterise warfarin dose differences among *VKORC1* haplotypes, the means of each pair

of haplotypes were statistically compared (Figure 4.2 and Table 4.4) – confidence intervals were (CI) were also calculated using QTPhase component of Unphased software (Dudbridge 2003). Haplotypes that share alleles at the first three SNPs (rs9923231, rs9934438 and rs2359612) do not exhibit significant differences in warfarin dose; haplotype H1 G–C–C-A-A versus H2 G–C–C-G-A (P = 0.09), and A–T–T-G-A haplotypes H3a vs. H3b A–T–T-G-G (P = 0.54). However, every pair of haplotypes with different alleles at the first three SNPs (G–C–C vs. A–T–T) shows significant differences (P = $1.2x10^{9}$ to 0.0163). Figure 4.1 and Table 4.4 both illustrate that G–C–C haplotypes (first two rows) require significantly higher mean doses (35.24–40.15 mg) than the A–T–T haplotypes (last two rows, 23.86–26.41 mg). It is possible to subdivide further the high and low haplotypes but this distinction is associated with only 3% additional explanation in dose variation. The haplotypic results and the results from the univariate model discussed earlier indicated that the first three SNPs are the best predictors of warfarin dose, and that rs7294 and rs11150606 provide much less additional predictive information.

Table 4.4. Significance of pairwise haplotype comparison.

Haplotypes	H3b	H3a	H2	H1	Genotype
H1	0.00145198	1.26×10^{-9}	0.0913005	-	G-C-C-A-A
H2	**0.0163264**	**0.000117041**	-		**G-C-C-G-A**
H3a	0.540107	-			A-T-T-G-A
H3b	-				**A-T-T-G-G**

4.3.2 Fine mapping across the 500 kb LD block harbouring VKORC1 in Caucasians

The three SNPs rs2359612, rs9934438 and rs9923231 in *VKORC1* have nearly perfect LD in the Swedish population sample, and it is not possible with the type of analysis described above to dissect which one, if any of them, is causative. None of the reported coding SNPs

that were genotyped is polymorphic in these patients, nor has any other common coding variant been reported by re-sequencing studies totalling 772 chromosomes (Geisen et al. 2005; Rieder et al. 2005). In addition, none of the rare mutations reported by Rost et al (2004) was observed in this study. Hapmap results indicated that *VKORC1* is positioned at the right wing of an extended LD block spanning 285 kb in Caucasians (Figure 4.3). To corroborate the effect of *VKORC1*, all genes located in this large LD block were assessed. *STX4* is located 60 kb to the left of *VKORC1* (Figure 4.3) and encodes syntaxin 4 a peptide thought to be involved in haemostasis and the formation of a platelet plug (Reed et al. 1999). Therefore, it was necessary to investigate whether the observed association signal was due to variants in the *VKORC1* gene or not.

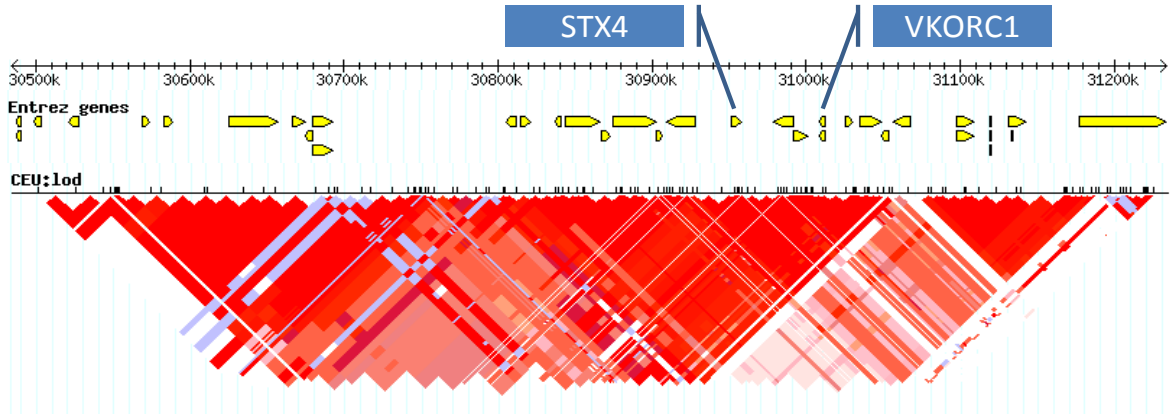


Figure 4.3. Genomic architecture of *VKORC1* region in Hapmap CEU panel (Caucasian).

To fine map the locus and elucidate *VKORC1* role in warfarin dose variation, a set of 13 additional polymorphic SNPs spanning the main and flanking LD blocks were selected from Hapmap (Figure 4.4). SNP genotypes were downloaded from HapMap website (phase I data) whereas LD, haplotype analyses and tag SNP selection were performed with the Haploview software (section 2.7.2). Manual SNP selection was performed if a tag SNP selected by Haploview failed in assay design (Sequenom platform).

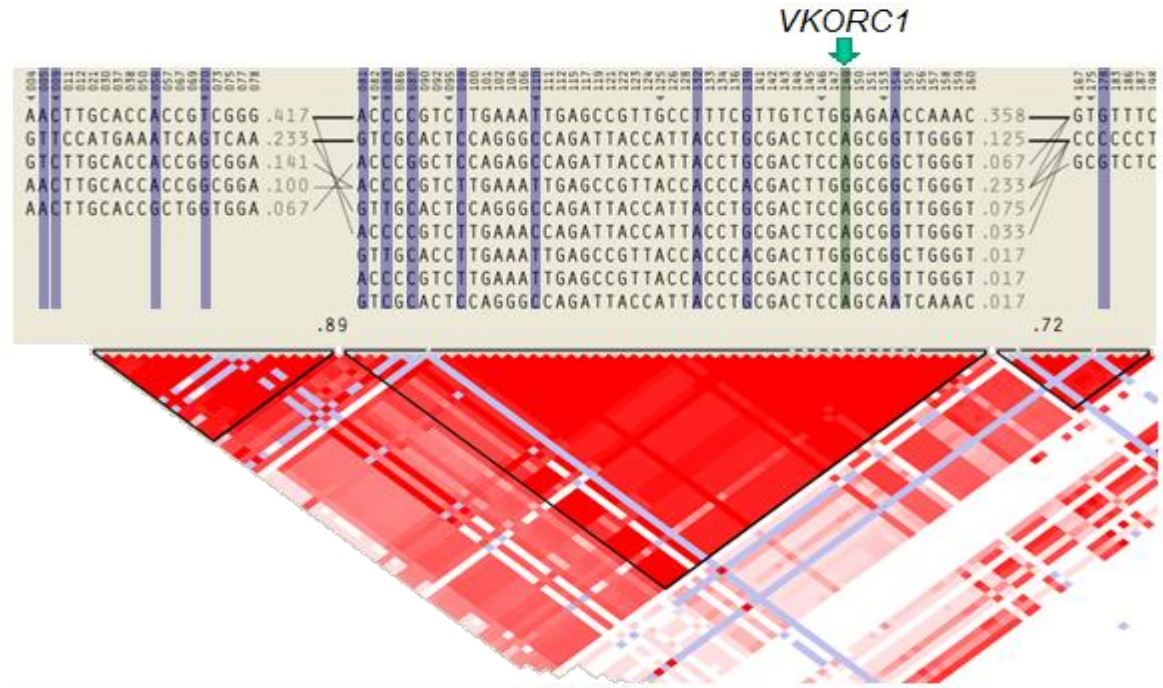


Figure 4.4. SNP selection for fine mapping *VKORC1* locus. SNP genotypes were downloaded from dedicated Hapmap website and haplotypes were phased with Haploview. SNP rs7294 located in *VKORC1* 3' UTR is shaded in green whilst selected tag SNPs are shaded in purple.

Ten SNPs in the main LD block which have either low or moderate r^2 with the four genotyped SNPs in *VKORC1* (Figure 4.5) and one SNP in each of the two flanking LD blocks were selected, giving a total span of ~558 kb. Finally, rs4889490 and rs4889599 which have moderate LD with rs2359612 and rs7294, respectively, were also selected to further refine the signal from VKORC1. Figure 4.5, panel B, shows that none of these SNPs gave higher association signal with warfarin dose compared to the original three linked SNPs which predicted circa 30% of the dose variance in our patients. The two SNPs rs4889490 and rs4889599 gave a signal slightly above the experiment-wise significance threshold due to their moderate LD with the four *VKORC1* SNPs. Further 2- or 3-marker haplotypic analysis indicated that no haplotype yielded association p-values more significant than the three linked SNPs. The fine mapping result confirmed the initial finding that common variants in *VKORC1* are the genetic determinants of warfarin dose.

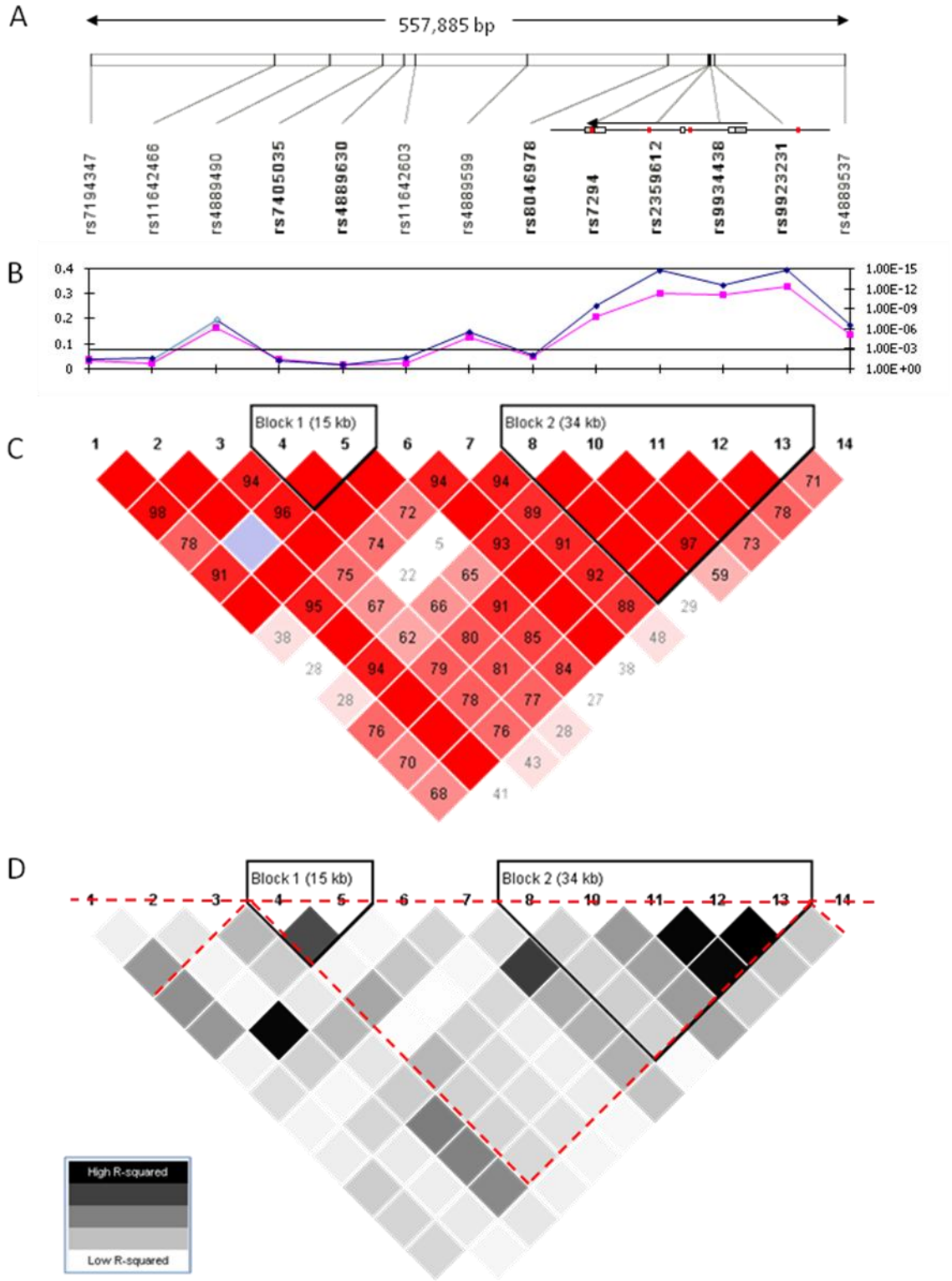


Figure 4.5. Association result of fine mapping *VKORC1* locus. (A) The SNPs flanking on LD blocks, including the four SNPs in *VKORC1*; (B) p-value (blue, right axis) and univariate R^2 (pink, left axis); (C) pairwise D prime and (D) pairwise r2.

4.3.3 The search of causative variants in *VKORC1*

To date there is no common coding sequence variant found in any of the published studies reporting the association of VKORC1 to warfarin dose requirement (D'Andrea et al. 2005; Rieder et al. 2005; Wadelius et al. 2005). Therefore, it is very likely that the causative variant is regulatory rather than one affecting the protein structure of VKORC1. So, does SNP rs9923231 have an effect on the transcribed messenger RNA (m-RNA) expression? This question was initially addressed by studies showing that rs9923231 is correlated with liver levels of mRNA (Rieder et al. 2005) and with activity in a reporter assay done in Hep G2 hepatoma cell lines (Yuan et al. 2005). These two studies both showed that the rs9923231 G allele is associated with higher mRNA expression (Figure 4.6). However, Bodin et al (2005) failed to confirm this reporter assay with the same experimental setup described in Yuan et al.

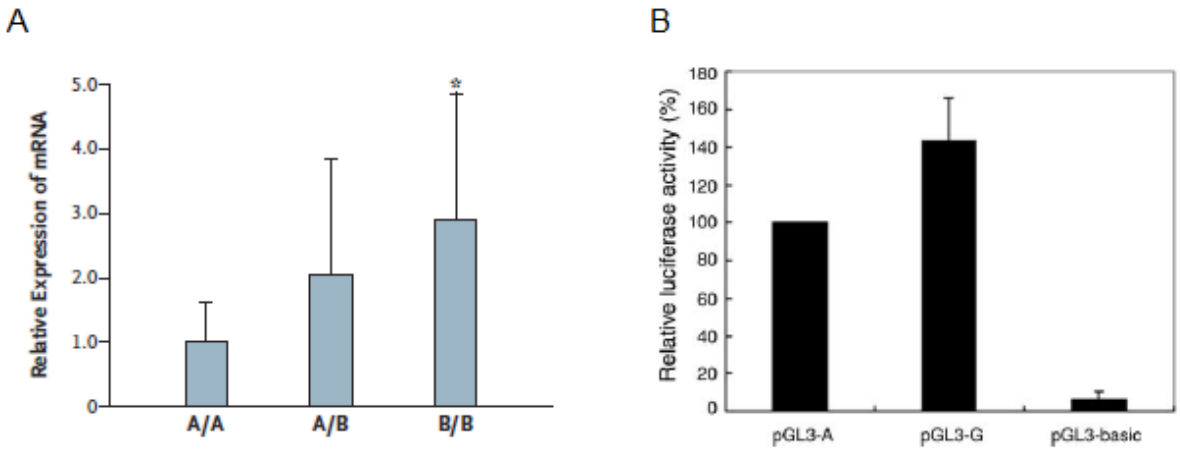


Figure 4.6. *VKORC1* mRNA expression assays. (A) Correlation between *VKORC1* haplotype groups and mRNA expression and (B) Luciferase activity assay with *VKORC1* promoter region. (A) the relative mRNA expression of *VKORC1* in liver was measured accordingly haplotypes A and B. The haplotype A carries minor A allele of rs9923231 whereas haplotype B carries the major G allele. (B) pGL3 luciferase reporter containing either the A (pGL3-A) or the G allele (pGL3-G) of rs9923231 whereas pGL3-basic is negative control without inserted sequence. Figure A is reproduced from N Engl J Med 2005 352:2285-2293 (Rieder et al. 2005) whilst Figure B is reproduced from Hum Mol Gen, 2005, 14(13):1745-1751 (Yuan et al. 2005).

Although the above preliminary results suggested that rs9923231 might be functional in regulating the expression of VKORC1, a further effort to confirm this observation is necessary due to the contradictory results reported by Bodin et al. A collection of liver biopsy samples from 25 individuals were provided courtesy of Dr Ana Alfirevic (Liverpool). The samples were used to test the mRNA expression in association with *VKORC1* genotypes. The RNAs of each liver biopsy were extracted with TRIZOL reagent and thereafter tested with quantitative PCR (Applied Biosystems). However, due to the age of these liver biopsy samples, the extracted RNAs were degraded and it was not possible to obtain reliable results (Figure 4.7). A second attempt was also made by using the Illumina DASL assay which was also unsuccessful (Dr Matthew Forrest).

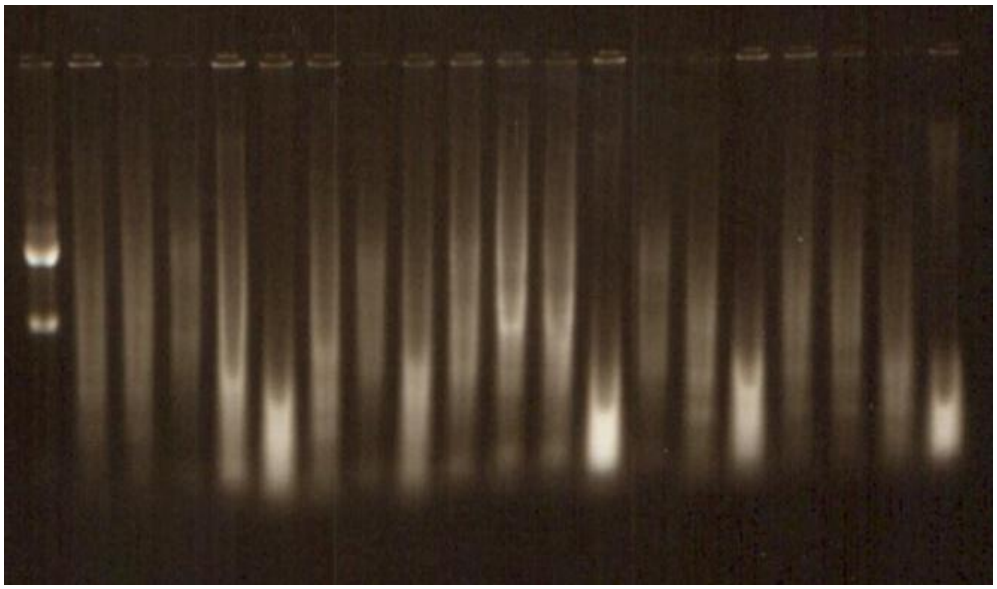


Figure 4.7. Electrophoresis of RNA samples prepared from the liver biopsies. The first lane indicates 18S (lower band) and 28S (upper band) standards.

Comparative genomics has demonstrated its powerfulness in predicting functional important segments, such as coding regions and gene regulatory elements resulting from purifying selection in evolution (Pennacchio and Rubin 2001). Numerous tools have been developed

for comparative sequence analysis such as MultiPIP (Schwartz et al. 2003) and zPicture (Ovcharenko et al. 2004). The corresponding sequences from chimpanzee, mouse, rat, dog, and chicken, corresponding to the human *VKORC1* locus, were downloaded from the Ensembl database and multiple sequence alignment was performed with zPicture (Figure 4.8).

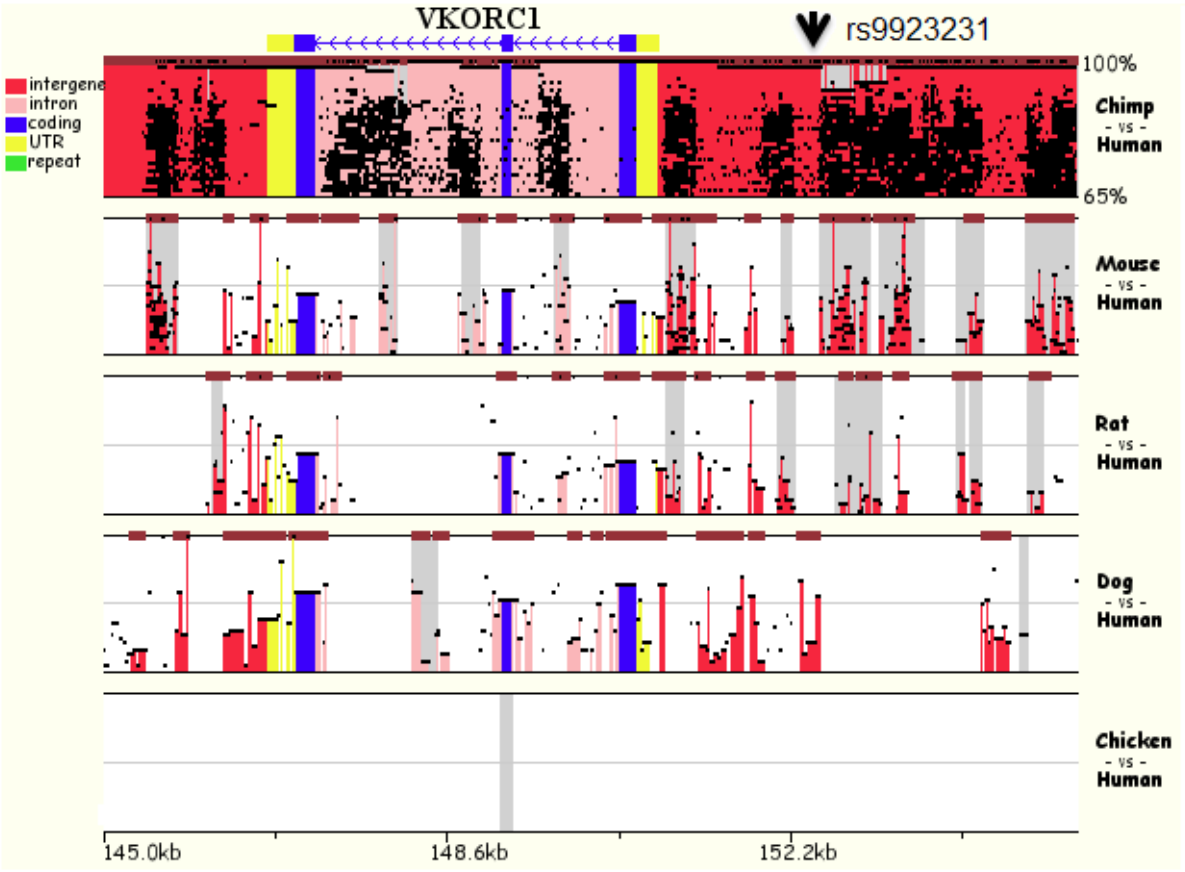


Figure 4.8. Genomic sequence alignment in *VKORC1* locus of human, chimpanzee, mouse, rat, dog and chicken. The black arrow indicates the location of potential functional SNP rs9923231.

Figure 4.8 shows the exon-intron structure of *VKORC1* at the top, whereas the genomic location of the putative promoter SNP rs9923231 is indicated with a black arrow. The three exons are conserved in all species except chicken whereas no significantly conserved region is found outside the genetic region; including the region harbouring rs9923231. This result

suggests that a comprehensive scanning for regulatory elements will be a good alternative approach to elucidate the expression correlation with *VKORC1* genotype.

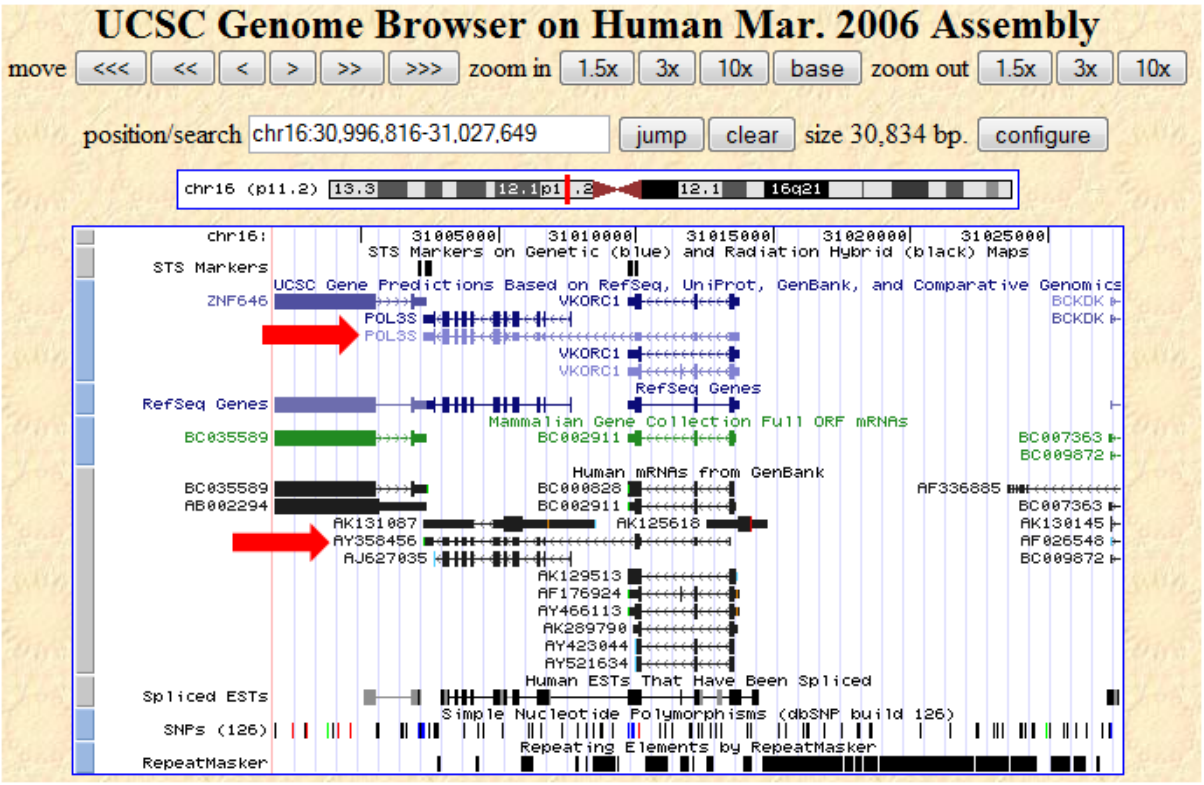


Figure 4.9. Alternative splicing variants of *VKORC1*. The red arrow indicates the evidence of a joint transcript of *VKORC1* and *POL3S* from a full length mRNA AY358456.

A few alternative splice variants have been reported in public databases based on the information of sequencing of full length cDNA libraries and expressed sequence tags (ESTs). Although none of these variants have been confirmed, interestingly, a long transcript which includes all three exons of *VKORC1* and all but the first exon of *POL3S* (polyserase) is supported by a full length mRNA transcript AY358456. The biological relevance and coding potential of such a transcript is unknown. It is worth mentioning that experimental work undertaken in our laboratory confirmed the presence of such transcripts in cDNA libraries from liver and lung tissue.

4.3.4 GGCX

GGCX acts in carboxylating vitamin K dependent proteins including F2, F7, F9, F10, protein C, protein S, and protein Z. For *GGCX*, we obtained results for 9 SNPs that passed study criteria and one microsatellite (Table 3.14). Figure 4.10 shows the LD architecture of the *GGCX* locus based on the nine common polymorphic SNPs, MAF above 30%. They are located in intron one (rs7568458), intron two (rs12714145), intron five (rs6738645), intron six (rs762684), exon eight (rs699664), exon nine (rs2592551), intron 14 (rs2028898) and in the 3' flanking region (rs6547621 and rs7605975). The exon eight SNP rs699664 leads to an arginine to glutamine change in codon 325, whilst rs2592551 in exon nine is a synonymous SNP. Inter-SNP distances are 0.8, 4.2, 1.1, 1.5, 0.4, 2.9, 2.6 and 2.1 kb, respectively. All nine SNPs are within a region of strong LD and define five common haplotypes (Figure 4.10).

In contrast to *VKORC1*, only one of the *GGCX* SNPs that passed study criteria reaches nominal statistical significance, rs12714145 (intron 2), P = 0.0360 (Figure 4.11). GGCX SNPs rs762684 (intron 6) and rs2592551 (exon 9) also show a tendency towards association with warfarin dose (P = 0.0613 and 0.0870). The mean warfarin dose associated with each haplotype was also calculated. A global test for statistical difference among the haplotype means revealed that there is no dose association among these haplotypes (P = 0.757).

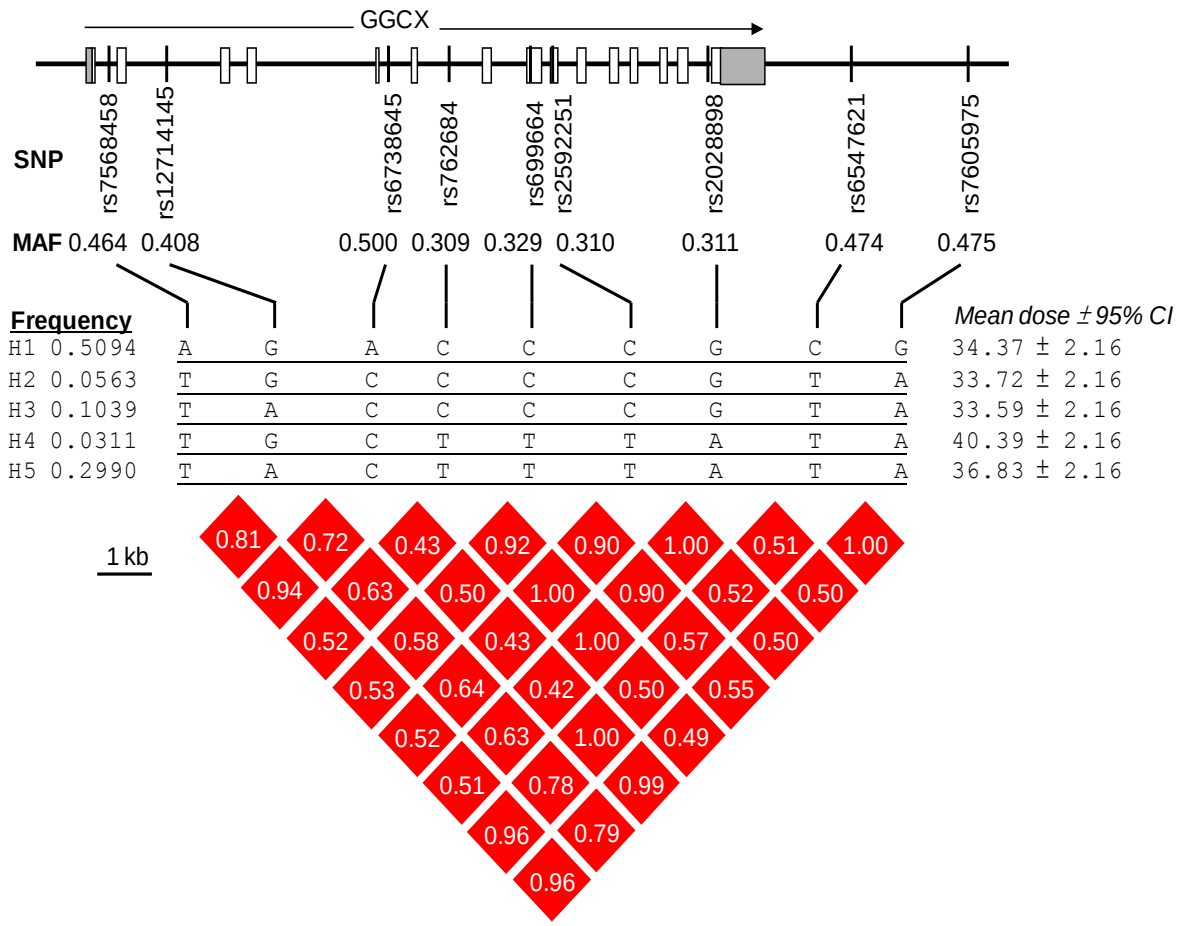


Figure 4.10. Genomic structure, haplotype, and linkage disequilibrium of *GGCX*. Nine SNPs were found polymorphic and common in Uppsala cohort. The MAF was noted under each SNP and ranges from 0.310 to 0.5. Haplotypes were listed with underline with annotation of frequency and mean dose (weekly) and 95% confident interval (CI). All SNPs are in strong linkage disequilibrium. Pairwise LD is represented in red (strong LD) and light blue (weak LD) diamonds with pairwise r^2 on it.

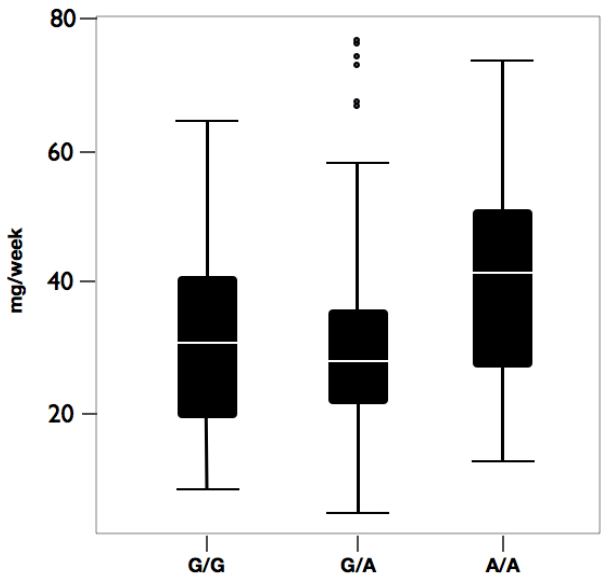


Figure 4.11. Mean weekly dose of rs12714145 genotype in 201 Swedish.

At the time of the above SNP analysis, Shikata et al (Shikata et al. 2004) described a microsatellite marker in intron 6 of the *GGCX* gene that was associated with warfarin dose. Three alleles with 10, 11, and 13 (CAA) repeats were detected in the 45 warfarin-treated Japanese patients. Three individuals which were heterozygous for the 13-repeats allele (10/13 or 11/13) required higher maintenance dose than patients carrying only alleles with 10 and/or 11 repeats. When this marker was typed in the Uppsala cohort, a wider range of alleles than in the Japanese cohort: was detected. Alleles with 10, 11, 13, 14, 15, or 16 repeats were found, with the 10-repeat allele being the most common. In analogy with the Shikata study, patients were divided into groups according to genotype: (1) 10/10 repeats, (2) 10/11 or 11/11 repeats, and (3) 10/13 or 11/13 repeats. In addition, a fourth group of patients has a greater number of (CAA) repeats, that is, homozygous for 13 or heterozygous for 14, 15, or 16 repeats (Figure 4.12).

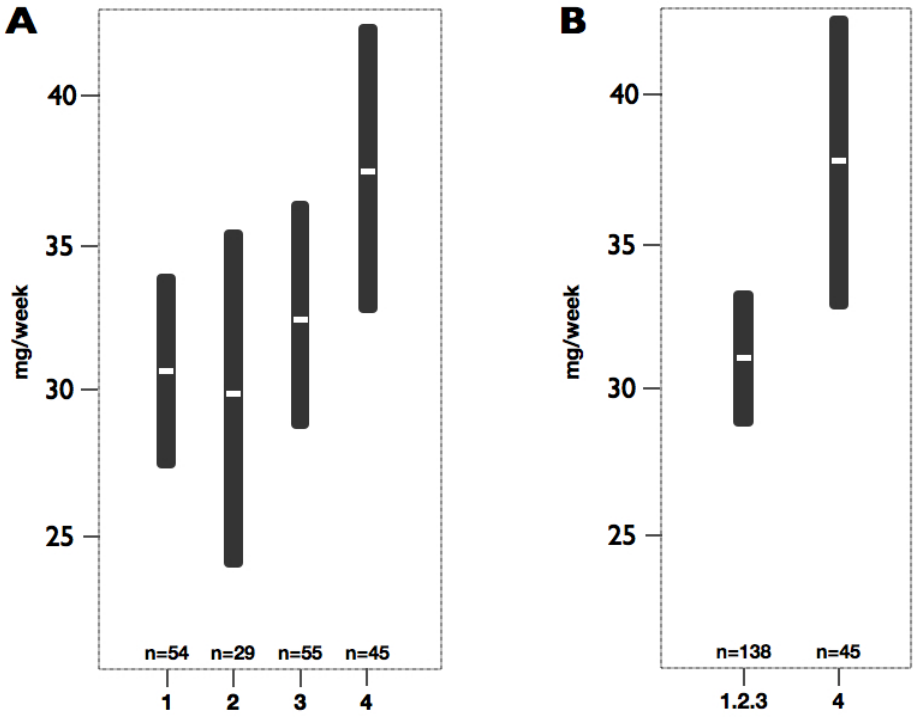


Figure 4.12. Individuals are divided into 4 groups. According to *GGCX* $(CAA)_n$ microsatellite genotype. Group 1: 10/10 repeats; group 2: 10-11/11 repeats; group 3: 10-11/13 repeats; group 4: 13/13 or x/14-16 repeats, x could be any repeat genotypes. (A) Mean weekly dose with 95% CI for four individual group. (B) Combined group 1-3 vs. group 4.

As in Shikata et al, warfarin dose requirement in the Uppsala cohort tends to increase with the number of microsatellite repeats, although the effect is only apparent in patients with higher numbers of repeats (group 4; Figure 4.12A). A combined of group 1, 2 and 3 against the fourth group shows nominally significant association (P = 0.011, Figure 4.12B). A *GGCX* polymorphism, rs12714145, has previously been observed in intron 2 that was associated with an increase in warfarin dose requirement (P = 0 .036). In chapter V, this finding wil be revisited in light of results in the much larger WARG study (see section 5.3.3).

4.4 CYP2C CLUSTER

In the previous chapter the eight P450 genes interrogated as part of this study were discussed (section 3.3). Four of them *CYP2C18, CYP2C19, CYP2C9* and *CYP2C8* constitute the CYP2C cluster on chromosome 10 (Figure 4.13). In the HapMap CEU panel, there is an extended region of strong LD harbouring these four genes. As shown in Figure 4.13 there is some LD granularity in this region with one large LD block harbouring *CYP2C18*, *CYP2C19* and the exonic part of *CYP2C9*, which includes all known functional *CYP2C9* variants, whereas a smaller LD block harbours *CYP2C8* and the flanking region of *CYP2C9* (Figure 4.13).

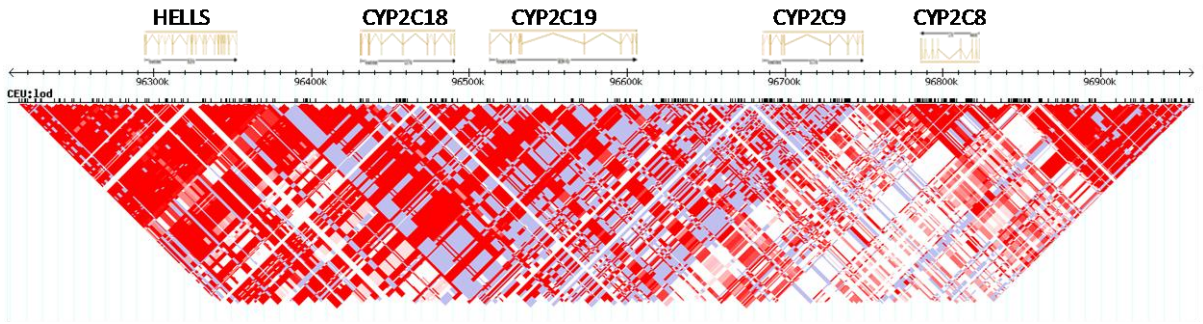


Figure 4.13. Genomic architecture of CYP2C gene cluster on chromosome 10.

Prior to this thesis, the *CYP2C9* gene had already been reported to be associated with warfarin dose (Aithal et al. 1999; Higashi et al. 2002). As expected, several CYP2C SNPs were found to be associated with warfarin dose. These were significant even after correction for multiple testing (Tables 4.2, 4.3).

The association data presented suggested that the CYP2C gene cluster on chromosome 10 was the second most strongly associated region after *VKORC1* (Tables 4.2 and 4.3). *CYP2C8,*

CYP2C9, and *CYP2C19* have been intensively studied and numerous genetic variants which alter amino acid composition have been reported (Human Cytochrome P450 Allele Nomenclature Committee, http://www.cypalleles.ki.se/).

To date *CYP2C9* has 30 reported alleles which may either increase or decrease its enzymatic activity in metabolising drugs. Among these variants, the *2 and *3 alleles have been reported to associate with warfarin dose in different populations (Solus et al. 2004). The functional *CYP2C9**2 polymorphism (rs1799853, R144C) confers a moderate decrease in the metabolism of S-warfarin (Rettie et al. 1994) and was not significant in our univariate analysis in the Uppsala cohort. The *CYP2C9**3 (rs1057910, I359L) severely impairs the efficacy to the hydroxylation of S-warfarin (Sullivan-Klose et al. 1996) and was the most strongly associated SNP in this region in our study (Table 4.2).

In this 400 kb region of high LD, 55 SNPs within, or flanking, the CYP2C cluster passed study criteria and were included for further analysis: 17 in *CYP2C9*, 10 in *CYP2C19,* 14 in *CYP2C18* and 12 in *CYP2C8* (Table 3.7 and Figure 4.14). The genomic architecture in the Swedish sample is similar to that reported by Hapmap, i.e. *CYP2C18, CYP2C19* and *CYP2C9* are in one large block whereas *CYP2C8* is located in separate smaller block. The LD blocks illustrated in Figure 4.14 are according to the block definition described by Gabriel and colleagues (Gabriel et al. 2002). However, an extended LD block can be clearly seen using less stringent thresholds (Ahmadi et al. 2005).

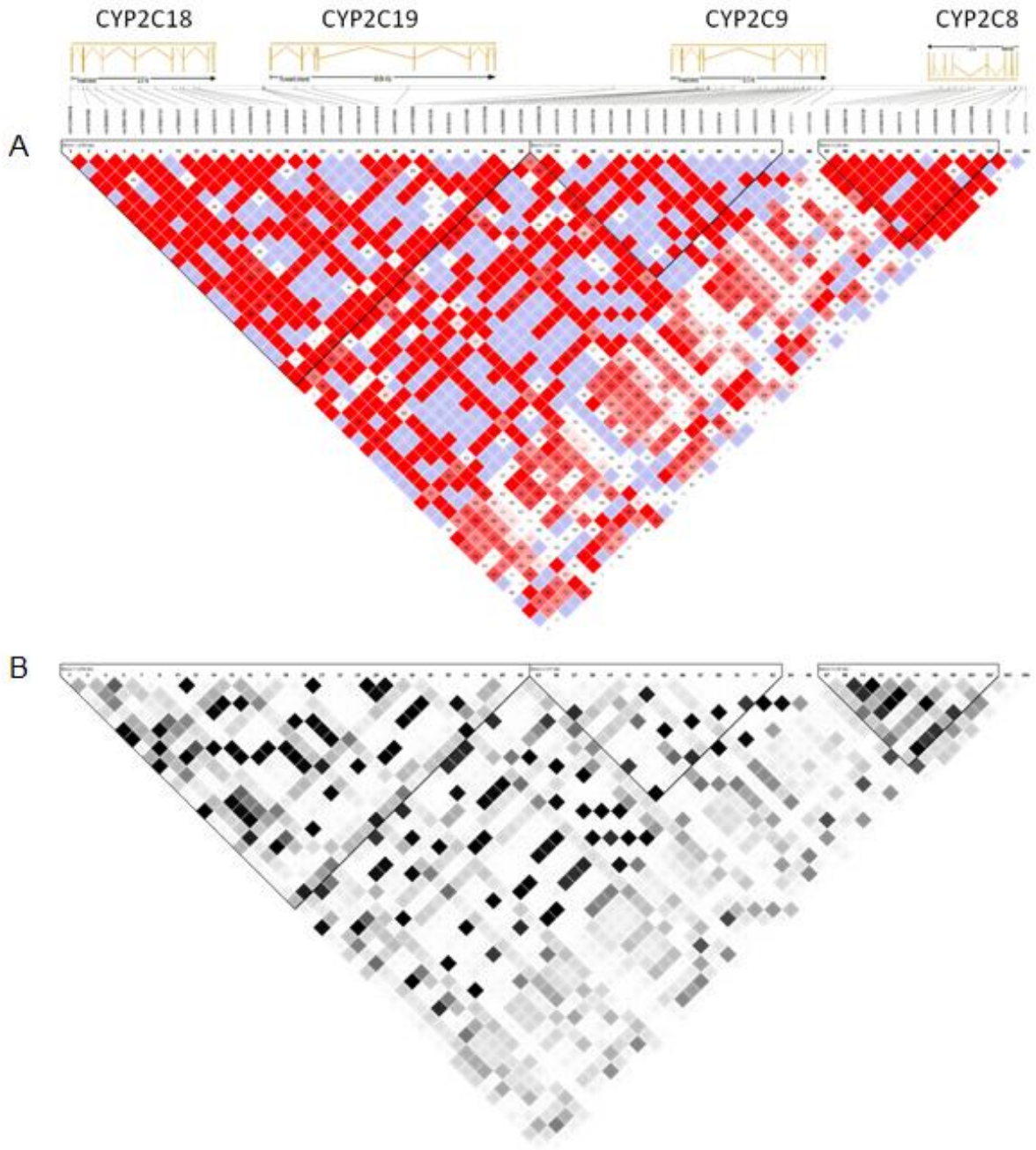


Figure 4.14. Genomic architecture of CYP2C gene cluster in 201 Swedish for (A) Pairwise D prime and (B) Pairwise r^2 calculation. The block is defined with definition described by Gabriel and colleagues (Gabriel et al. 2002).

In univariate analysis, nine SNPs showed significant association. To further explore the dose association of these nine SNPs, two multiple regression models were applied. Both models contained *VKORC1* (rs2359612) and significant non-genetic predictors of warfarin dose (Age, body weight and interacting drugs), whereas one contained *CYP2C9**2 but not *3 and vice

versa.

Each of the 53 SNPs in the CYP2C region was evaluated against the two multiple regression models. The result indicated that all significant results were fully explained by LD with either *CYP2C9**2 or *3, except for rs4917639 in *CYP2C9* (Table 4.5). SNP rs4917639 demonstrated a significant P-value in both the *2 model ($P < 1.33x10^{9}$) and the *3 model ($P < 3.56x10^{3}$). However, haplotype analysis showed that the minor allele of rs4917639 was in perfect LD ($r^2 = 1$) with a composite minor allele formed by aggregating *CYP2C9**2 and *3 into a single allele; although *2 and *3 are rarely carried on the same haplotype. This result suggests that the rs4917639 mutation occurred first, and the *2 and *3 alleles arose independently on the same parent allele. The strong association between rs4917639 and *2/*3 could perhaps be due to positive selection; if rs4917639 lessened the deleterious effect of impaired CYP2C9 metabolism caused by *2 and *3.

Table 4.5. Nominally significant result in multivariate regression model for SNPs in CYP2C gene cluster.

Gene	SNP	R^2 with *2 model	p-value	R^2 with *3 model	p-value
CYP2C8	rs11572080	0.001	0.502	0.017	0.005
CYP2C9	rs9332108	0.109	1.57E-10	0	0.999
CYP2C9	rs1057910 (*3)	0.109	1.57E-10	-	-
CYP2C9	rs1057911	0.114	2.97E-10	0	0.999
CYP2C9	rs9325473	0.112	5.45E-10	0	0.999
CYP2C9	rs4917639	0.1	1.33E-09	0.025	3.56E-03
CYP2C9	rs9332214	0.098	1.50E-09	0	0.999
CYP2C9	rs2860905	0.048	1.86E-04	0.009	0.153
CYP2C9	rs4917636	0.004	0.213	0.026	3.63E-03
CYP2C9	rs4607998	0.004	0.252	0.026	2.85E-03
CYP2C9	rs1799853 (*2)	-	-	0.024	4.00E-03
CYP2C9	rs1934966	0	0.999	0.015	8.72E-03
CYP2C9	rs9332222	0	0.999	0.025	3.86E-03
CYP2C18	rs7896133	0.063	5.17E-07	0	0.999
CYP2C18	rs2901783	0.02	0.029	0.004	0.471
CYP2C19	rs3814637	0.098	3.76E-09	0	0.896
CYP2C19	rs17882687	0.047	5.19E-05	0	0.828

In conclusion, the poor metabolising effect of *CYP2C9**3 alleles in the Swedish sample was confirmed. Compared to *CYP2C9**3, *CYP2C9**2 showed a minor effect in warfarin maintenance dose in the Swedish cohort. In the Uppsala cohort, no independent effect is identified except *2 and *3 alleles and rs4917639 which is linked to both alleles.

4.5 OTHER NOMINALLY ASSOCIATED GENES

Except *CYP2C9* and *VKORC1* which directly interact with warfarin, *PROC* emerged as the most likely factor which may influence the dose variation. In this study, 13 SNPs in *PROC* passed study criteria (Table 3.21.). In univariate analysis, four out of the 13 SNPs tested were significantly associated with dose (Table 4.2). The associated SNPs were located in the 5' regulatory region (rs1799809 and rs2069901), in intron 2 (rs2069910), and in intron 3 (rs2069919). Except for rs2069910, all other SNPs reached gene-wise significance. Haplotype analysis of the 13 SNPs in *PROC* did not provide a stronger signal compared to single marker analysis, although the haplotype derived from rs2069919, rs2069921 and rs973760 reached gene-wise significance (P = 0.00136). In the model with *VKORC1*, *CYP2C9*, body weight and interacting drug, the two SNPs in the promoter region of *PROC* explained 7–9% of the variance in warfarin dose. Previous studies had suggested that homozygotes for the G allele of rs1799809, which is located in the promoter region of *PROC,* have slightly decreased enzymatic activity and reduced levels of PROC concentration (Aiach et al. 1999; Spek et al. 1995). Furthermore, in 2003, Watala et al reported a correlation between the enzymatic activity of PROC and the coagulation rate both in patients treated with oral anticoagulant and in healthy volunteers (Watala et al. 2003). Our results suggested that patients homozygous for the G allele required a lower dose, which is in agreement with the above finding in biological function of protein C. The PROC association reached only gene-wise significance and in a small sample of 201 patients.

Other than *PROC, APOE* was the only other gene that reached gene-wise significance in the univariate analysis for association with warfarin dose. The two *APOE* SNPs (rs429358 and rs7412) that discriminate the widely described *E2, *E3 and *E4 allelic system showed that

patients who carry the common allele *E4 or the rarer *E2 require higher warfarin doses than those with *E3 (Table 4.2). This finding is contradictive to a UK study (Sconce et al. 2006). A Dutch study analysing patients taking the anticoagulant drug phenprocoumon showed that the *E4 is associated with higher maintenance dose, but the *E4 allele carriers required lower maintenance doses of acenocoumarol (Visser et al. 2005). This discrepant result suggests that the association between anticoagulant dose and APOE may result from sampling bias, and a larger cohort treated with phenprocoumon, acenocoumarol and warfarin will help to refine this association (Wadelius et al. 2007).

Besides the genes that have been described above, eight additional genes were nominally associated with warfarin dose but did not pass either the experiment-wise or gene-wise threshold after correction for multiple testing: *EPHX1*, *CALU*, and *ORM1-2* in single marker analysis and *F7*, *PROZ*, *F9* and *NR1I2* in haplotype analysis (Tables 4.2 and 4.3).

In single marker analysis, rs1051740 in *EPHX1* reported in association with high maintenance dose (Loebstein et al. 2005) is not replicated, but, instead, another SNP rs4653436 is nominally significant. A functional variant in *CALU* reported to potentially increase warfarin dose is also not replicated. However, four other SNPs in *CALU* show nominal significance. SNPs in *PDIA2* and *ORM1-2* showed a nominally significant p-value and marginal association with dose. To date, there has not been any report with *ORM1* and *ORM2* in association with warfarin dose.

Haplotypes in *PDIA6* showed gene-wise significance in association with dose whereas in *PDIA3* it is nominally significant (Table 4.3). Although PDI has been demonstrated to be involved in providing electrons to the thioredoxin-like CXXC centre in VKORC1 and in

initiating tissue factor pathway for coagulation, precisely which PDI member is involved in these interaction remains unclear (Ahamed et al. 2006; Versteeg and Ruf 2007; Wajih et al. 2007). The haplotypes in *F7* and *PROZ* showing nominal significance are both comprising SNPs located in the upstream flanking region (Table 4.3), and, interestingly, *F7* and *PROZ* reside back to back on chromosome 13 with *F10*. Our finding of *F7* upstream haplotype is in accordance with previous studies which have reported variations in the *F7* promoter region to be associated with warfarin sensitivity (Aquilante et al. 2006; D'Ambrosio et al. 2004; Shikata et al. 2004). Shikata and colleagues also reported SNPs rs5896 in *F2* and rs5960 in *F10* to be associated with warfarin dose (Shikata et al. 2004), but these observations did not replicate in our Uppsala study. In *F7*, the functional variant R413Q, (rs6046) which was reported to be associated with decreased plasma levels of F7 (Arbini et al. 1994), did not replicate either in our study. A rare variant (A-10V/T) in *F9* which was previously reported to be associated with decreased F9 enzymatic activity, was not polymorphic in the Uppsala sample. However, a haplotype spanning this exon has shown a marginal association. Finally, the haplotype on *NR1I2* is located near the one of the two recombination hot spots and may harbour a novel mutation which influences warfarin dose. To date, no variant in *NR1I2* has been reported in the literature to be associated with warfarin dose and replication in an independent study will validate this finding.

Although some of the variants described above appear to be marginally significant and explain the variability of dose variation in the Uppsala cohort, they should be treated as provisional at this stage pending replication in a larger cohort of the same ethnic origin. This work is described in Chapter V.

4.6 PREDICTIVE MODEL FOR WARFARIN DOSE

Besides the genes showing significant association with warfarin to various degrees, non-genetic factors also contribute to dose variability and age, body weight, and drug interaction (concomitant medication) can better explain the dose variation (Table 4.6). To develop the best multiple regression model, these factors were taken into account with the genetic factors showing at least nominal significance.

Table 4.6. Association test of non-genetic factors.

Variables	P-value	r^2
Age	0.0029	0.092
Bodyweight	0.0075	0.057
Interaction	0.0878	0.036
Gender	0.0314	0.023
Indication	0.0819	0.015
PT INR	0.1272	0.012

Niclas Eriksson (Uppsala) developed a multiple model in which both non-genetic and genetic factors were assessed. For non-genetic factors, age, body weight, interacting drugs, and indication were retained while gender and PT INR value were excluded in the multiple model. In Figure 4.15, the contribution of genetic factors into different models based on the strength of the association signal is shown. Together with non-genetic factors, the two genes showing experiment-wise significance, *VKORC1* and *CYP2C9*, account for 56.0% of the total inter-individual variation in warfarin response. If *PROC* is included in the model with *VKORC1* and *CYP2C9*, 62% dose variation could be explained for the variance in dose (Figure 4.15).

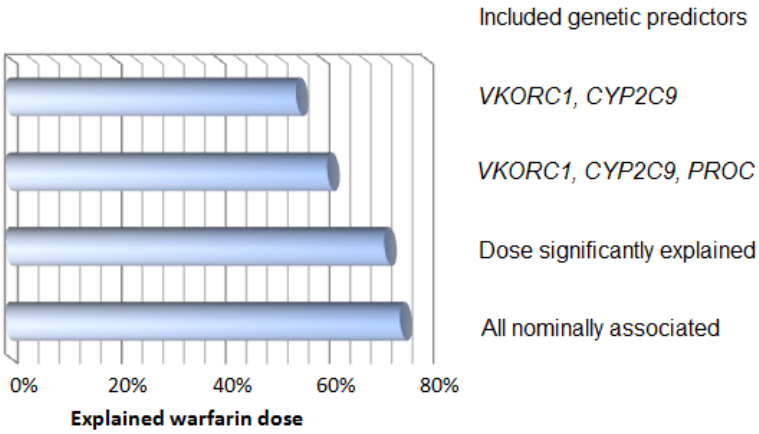


Figure 4.15. Multiple models of dose explained by genetic and non-genetic factors. Different models comprised different combination of predictors were tested to explain the inter-individual dose variation.

Finally, to explore the full potential of all findings from the Uppsala study in describing dose variation, all genes showing nominally significant association including *VKORC1*, *CYP2C9*, *CYP2C19*, *CYP2C18*, *PROC*, *APOE*, *EPHX1*, *CALU*, *GGCX*, *ORM1-2* and non-genetic factors were recruited in the multiple model and a surprising 76% of the inter-individual variation of warfarin dose could be explained (Figure 4.15). However, to refine this model, genetic factors with p-values above 0.2 were removed in a stepwise fashion, together with low-explanatory value (R^2) from the model. A final model which includes *VKORC1*, *CYP2C9**2 and *3, *PROC*, *EPHX1*, *GGCX*, *ORM1-2*, age, body weight, and drug interactions explains 73% of the dose variation (Figure 4.15; Table 4.7).

Table 4.7. Predictors in final multiple model of warfarin dose variation.

Predictor	SNP	p-value	Univariate R^2
VKORC1	rs9923231	<.0001	0.317
CYP2C9	rs1799853 (*2) + rs1057910 (*3)	<.0001	0.159
Age		0.0029	0.092
PROC	rs2069919	0.0416	0.09
Bodyweight		0.0075	0.057
EPHX1	rs4653436	0.1016	0.048
Drug interaction		0.0878	0.036
GGCX	rs12714145	0.026	0.034
ORM1-2	rs1687390	0.0571	0.026

Although p-values of *EPHX1*, *ORM1-2* and drug interaction do not reach even nominal significance, these predictors could explain a substantial variation of dose. Since this final model is evaluated based on a small study sample, validation of these effect would be necessary to confirm the influences.

4.7 CONCLUSIONS

A comprehensive selection of 35 candidate genes was investigated in association with inter-individual dose variation. The result suggests that *VKORC1* and *CYP2C9**2 and *3 are significantly associated with the effect of dose variation. This finding in our Swedish sample was also replicated in different ethnic populations in other studies (D'Andrea et al. 2005; Rieder et al. 2005; Yuan et al. 2005). The outcome of CYP2C9*2 and *3 alleles is well studied and both alleles are known to impair CYP2C9 enzymatic activity. However, the biological explanation of VKORC1 action is still unclear although the two *CYP2C9* alleles explain well the dose variation and preliminary evidence suggests a variation in mRNA expression (Rieder et al. 2005; Yuan et al. 2005). Other possibilities include alternative splicing or remote regulatory elements which require further investigation including functional studies.

A conservative multiple model including age, body weight, drug interaction, genotypes of VKORC1 and CYP2C9*2 and *3 is able to explain 56% whereas additional predictor of PROC contributes another 6% with a total of 62%. This model could be used to develop a dosing algorithm and tested in prospective studies. We found additional, marginal associations which may be population specific. At this point, these findings were tested for replication in a larger cohort; which will be covered in the next chapter.

CHAPTER V

AN INVESTIGATION OF GENETIC DETERMINANTS OF WARFARIN DOSE REQUIREMENT IN 1500 SWEDISH PATIENTS (WARG COHORT)

In the previous chapter, a study of 35 candidate genes in 201 warfarin-treated patients (Uppsala) was described. This study identified SNPs in *VKORC1*, *CYP2C9-CYP2C19*,-*CYP2C18*, *PROC*, *APOE*, *EPHX1*, *CALU*, *PDIA2*, *GGCX* and *ORM1-2* being associated with variability in warfarin dose requirement. The strength of these associations above nominal significance varied and only *VKORC1* and *CYP2C9* reached p-values below 1.65×10^{-4} (experiment-wise).

Our first goal was to replicate the findings from the Uppsala study in a large independent sample. However, the sample size of 201 patients in the discovery study is small and had statistical power to detect only 'strong' effects. We therefore devised a strategy to combine the replication experiment with a *de novo* investigation of all the 35 candidate genes in the much larger sample of the WARG study.

5.1 THE NATIONAL WARFARIN GENETIC (WARG) STUDY

The national warfarin genetic (WARG) study in Sweden (http://www.druggene.org/) is a prospective study to understand the genetic determinants of inter-individual warfarin dosing and warfarin-induced bleeding complication. Patients were recruited in 40 outpatient clinics in Sweden, and the criteria of patient recruitment are described in chapter II, section 2.1.2. Age distribution of WARG patients was comparable to that in the Uppsala study (Figure 5.1). The patients in WARG were 18-92 years old, with the majority of them been aged between 50 and 80. There is no patient younger than 18 years old enrolled in the study; only one recruited patient was aged 18.

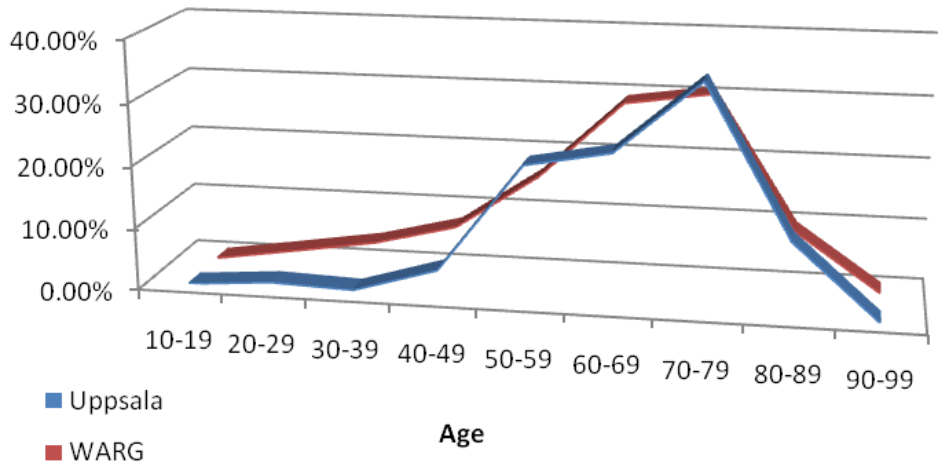


Figure 5.1. Age distribution of patients in the Uppsala and WARG studies. The majority of patients treated with warfarin was aged from 50-80 in both studies.

A total of 1523 patients were eventually enrolled in the WARG study of which 1496 were analysed in this study. The basic characteristics of these 1496 patients are described in Table 5.1. Patients have an average age of 66 years (95% Confidence interval (CI): 57 years and 74 years) and were predominantly male (947 male). The constraints of the approved consent

form and recent practises in Sweden, prevented patient's body weight being recorded. Various medical indications were recorded electronically in each outpatient clinic and thereafter submitted to a central database through the internet (Lindh et al. 2004). More than half of the patients (51%) suffered from atrial fibrillation, which resembles the patient composition of the Uppsala study (56%). The second most diagnosed indication is deep vein thrombosis and pulmonary embolism with 566 patients (38%). This is in contrast to the Uppsala study which comprises only 4.5% of such patients. The Uppsala study in return comprised a bigger portion of patients with heart valve prosthesis (24%) against a merely 4% in the WARG cohort.

Most patients in the WARG study had been monitored for at least three months and relevant medical information was registered into the database. The majority of patients were planned for life-long treatment or for treatment without predefined period. The maintenance dose used to test genetic association was defined as the mean of all doses given to a patient during a minimum series of three consecutive INR measurements between 2 and 3 (therapeutic INR). The stable dose was calculated from all doses that were unchanged over a minimum of three consecutive visits and that lead to a therapeutic INR. Patients outside target INR and the above criteria were removed from further analyses. In the WARG study there is 19-fold variation in warfarin dose requirement (6.0 to 113.75 mg/week) compared to 17-fold in the Uppsala study (4.5 to 77.25 mg/week). The recruitment of the severe bleeding patients is according to the World Health Organisation (WHO) definition, which is: lethal, life-threatening, permanently disabling, or leading to hospital admission (emergency room admissions excluded) or prolongation of hospitalisation (Lindh et al. 2007). During treatment, 146 patients (9.8%) were subject to bleeding complications whereas 28 had severe bleeding episodes (1.9%). The analysis of genetic association with severe bleeding complication is

described in chapter VI.

Warfarin doses and PT INR measurement intervals were chosen at the discretion of the treating physicians (Lindh et al. 2007). The targeted therapeutic range of PT INR is 2.0-3.0 which was measured in each visit to clinics. During the period of a total of 1276 patient-years, 63.8 % had targeted within therapeutic PT INR interval and an average of 24.5 PT INR values were recorded for each patient. The statistics of the number of INR recorded for each patient was shown in Figure 5.2. Information as to patients who did not consent to join the study was not recorded in the study design.

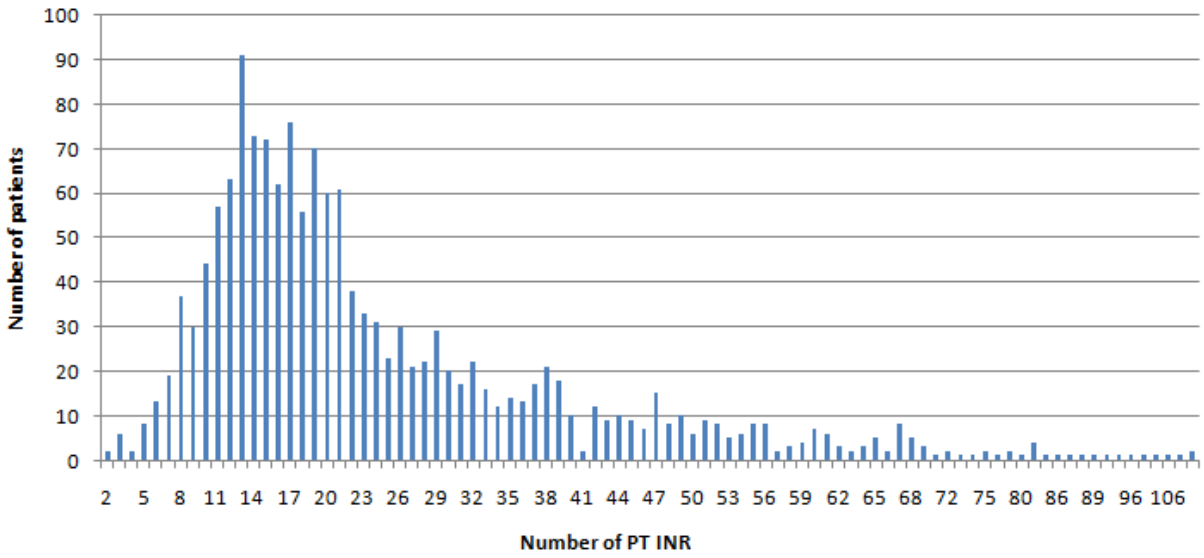


Figure 5.2. The number of PT INR recorded to each patient.

Most of the patients in WARG were receiving other medication with a total of 1528 different drugs being recorded (Table 5.1). Among the prescribed drugs, 781 are known to have no-interaction with warfarin; 56 as having decreasing effect in warfarin administration and 691 drugs potentiating the effect of warfarin which in turn increases anticoagulation and the risk for bleeding complications.

Table 5.1. Medical information statistics of patients in WARG cohort.

Characteristics	
	Patients (%)
Indication for warfarin treatment	
Atrial fibrillation	762 (51 %)
Deep venous thrombosis	378 (25 %)
Pulmonary embolism	188 (13 %)
Heart valve transplant, artificial	34 (2 %)
Cerebral infarction/transient ischemic attack	28 (2 %)
Heart valve transplant, biologic	25 (2 %)
Cardiomyopathy/cardiac failure	23 (2 %)
Other indications†	58 (4 %)
Planned treatment duration	
<3 months	13 (1 %)
3-5 months	214 (14 %)
6-12 months	387 (26 %)
>12 months, specified	14 (1 %)
Infinite or not predefined	868 (58 %)
Patients experiencing bleeding events (%)	
Serious bleeding	28 (1.9%)
All bleeding	146 (9.8%)
INR, international normalized ratio	
Lower limit of therapeutic interval	2.1 (1.5;3.1)
Target value	2.5 (1.8;3.0)
Upper limit of therapeutic interval	3.0 (2.0;3.6)
Concomitant medication, No. of drugs	2 (0; 18)
Time within therapeutic INR interval (%)	63.8 (50.9; 73.5)
Use of drugs interacting with warfarin	
No interacting drug	781 (52 %)
Drugs *decreasing* warfarin effect	56 (4 %)
Drugs *potentiating* warfarin effect	691(46 %)
Gender	
Men	947 (63%)
Women	549 (37%)
Age (95% CI)	66 (57; 74)

A male preponderance was observed in both the Uppsala and the WARG studies with a 2:1 ratio of male/female. Table 5.2 lists the number of male and female patients enrolled in a selection of warfarin genetic studies and only the studies enrolling mainly Caucasian subjects were included. Beside the two study cohorts described in this thesis, the study published by

Bodin et al (2005), which recruited normal subjects, also has male preponderance. However, other studies recruited slightly more male than female patients (5-10%) except the study by Aquilante et al (2006), which recruited a large percentage of patients from two Veterans Administration clinics. The male preponderance in both Swedish studies might be as a result of the earlier onset of heart related diesease in males or a potential bias in recruitment, for exmaple: male is more willing to participate clinical studies.

Table 5.2. Male/female ratio in a selection of warfarin genetic studies.

Year	REF	n	Male	%	Female	%	Ethnicity	Note
2005	Bodin L et al	222	145	65.3	77	34.7	French	Normal subject
2005	*Uppsala study*	*201*	*135*	*67.2*	*66*	*32.8*	*Swedish*	
2005	Sconce EA et al	297	160	53.9	137	46.1	British	
2005	D'Andrea GR et al	180	100	55.6	80	44.4	Italian	
2006	Aquilante CL et al	350	306	87.4	44	12.6	American Caucasian (91%) Afro-American (7%) Hispanic 1% Asian (0.3%)	A large percentage of patients were recruited from 2 Veterans Administration
2006	Carlquist JF et al	213	104	48.8	109	51.2	Predominantly white, of Northern European	
2006	Schalekamp T et al	231	133	57.6	98	42.4	European	
2007	Borgiani P et al	148	78	52.7	70	47.3	Italian	
2008	Schwarz UI et al	297	160	53.9	137	46.1	White (89.2%) Black (9.8%) Hispanic (1.0%)	
2008	*WARG*	*1496*	*947*	*63.3*	*549*	*36.7*	*Swedish*	

5.2 GENOTYPING APPROACH

The aims of this study were to: (i) replicate in WARG the association signals we obtained in the Uppsala study and (ii) investigate the smaller effects in the 35 candidate genes. The strategy of choice involved following a haplotype tagging approach with inclusion of the lead SNPs and important functional variants from the discovery study. Patients recruited in the WARG study were predominantly Swedish; the same as the patients collected in the Uppsala study. In chapter III the construction of detailed LD maps for each of the 35 candidate genes was described. Selection of haplotype tag SNPs was performed using Tagger (de Bakker et al. 2005) built in Haploview (Barrett et al. 2005). This approach reduced the genotyping and labour costs.

5.2.1 SNP selection

All SNPs with $P \leq 0.1$, in either the univariate or multivariate regression analyses of warfarin dose association in the Uppsala study, were selected for inclusion in the WARG study. Tagging was undertaken in all 35 candidate genes with the software Tagger (de Bakker et al. 2005) setting the pairwise r^2 thereshold at 0.8 and MAF $\geq$ 5%. If a tag SNP selected by Tagger was in the same bin (tightly linked SNPs) with any of the pre-selected SNPs, it was then replaced by the pre-selected SNP to represent that bin. A total of 216 SNPs were selected to be tested across the WARG cohort. Any failed genotyping assay was redesigned using different chemistry, i.e. iPLEX to MassEXTEND or vice versa.

5.2.2 Taqman genotyping *CYP2C9**2 and *3

Since the effect of *CYP2C9**2 and *3 alleles on warfarin dose requirement has been confirmed in the Uppsala cohort and other studies (Aithal et al. 1999; Higashi et al. 2002; Wadelius et al. 2005), it was important to obtain accurate genotypes for these two variants in all individuals. However the respective SNP assay designed for the *CYP2C9**2 and *3 allele was shown to suffer from allele drop-out on the Sequenom platform regardless of chemistry used, i.e. both iPLEX and MassEXTEND assays

The *CYP2C9**2 and *3 SNPs were genotyped using Taqman Pre-developed genotyping assays (Applied Biosystems). Experimental setup was as per manufacturer's instructions in 5 μl reaction volume with slight modification. The addition of five amplification cycles increased the genotyping call rate of 15%.

5.2.3 Genotyping summary

Although SNPs which had MAF below 5% were rejected from analysis in the Uppsala study because of the small sample size, a few of them were included with previous evidence of being functionally important in the WARG study. SNPs genotyped in the WARG study, out of Hardy-Weinberg Equilibrium (HWE, $P < 0.001$) and with call rate below 70%, were removed from analyses. 216 SNPs passed the study criteria with an average of genotype call rate of 95% (see Appendix III for call rate of each SNP). The MAF of the 216 SNP is shown in Figure 5.3.

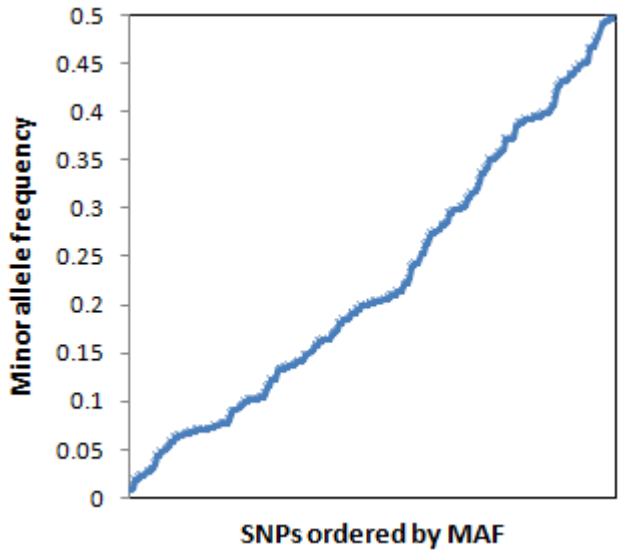


Figure 5.3. Minor allele frequency distribution of 216 SNPs genotyped in WARG study. The SNPs (X-axis) are ordered and plotted from low (1%) to high MAF (50%).

5.3 DOSE VARIATION

The univariate and multivariate linear regression model analyses were performed, with the R statistical package, to examine the association of dose and genetic and environmental factors. To correct for multiple testing, Bonferroni correction based on the effective number of independent tests (Meff) (Cheverud 2001; Li 2001; Nyholt 2004) was applied. The sum of effective tests is 190 in this study and the experiment-wise significance is P< 2.62E-04.

5.3.1 Gender, age, and dose

An unequal number of male and female individuals were recruited in the WARG study. In the Uppsala study, which comprises 201 patients, gender is not significantly associated with the dose when body weight is included in the multiple model. However, it is interchangeable when body weight information is not available. In the WARG study, gender is significantly associated with dose variation (p=6.89E-05). As mentioned earlier body weight information is not available in the WARG patients. Females required ~4 mg (10.7%) less than male patients in mean weekly dose, and the median weekly dose for male and female is 35.51 and 30.83 mg, respectively (Table 5.3; where N is the number of male or female patients).

Table 5.3. Gender effect in warfarin dose association.

Varable	N	Mean	Min	Median	Max
Sex					
M	947	37.42	5.97	35.51	113.75
F	549	33.48	6.45	30.83	103.95

Figure 5.4 is a plot of the weekly dose against age which shows a linear trend for both males

(A) and females (B). The majority of male patients were slightly younger than the female patients; which is in agreement with the earlier onset of heart related diesease in males. The linear trend line in female is more steep than in male suggesting that a smaller amount of warfarin is required for elder, as well as younger, female than male patients. Except for a few outliers, seven (1.27%) women out of 549 female patients were prescribed a dose of >70 mg/week and so were 28 men (2.96%) require a dose of >70 mg/week among 947 males.

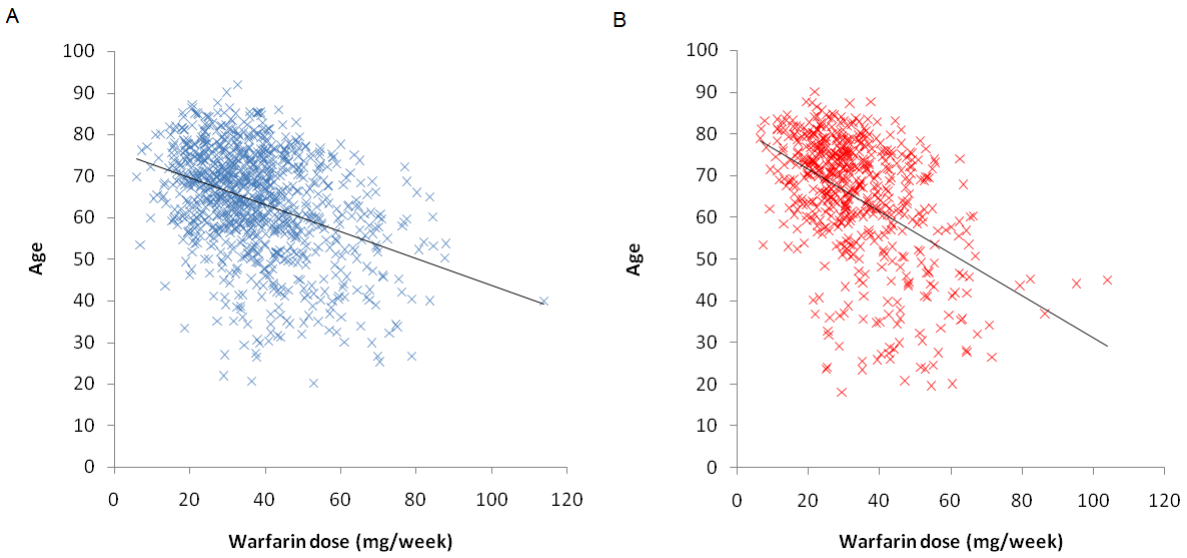


Figure 5.4. Age effect on warfarin dose association in (A) male and (B) female patients Patient's age and maintenance dose is plotted and a linear trend line is shown accordingly.

5.3.2 Univariate regression

Although genotypes of 1496 treated patients were obtained, only those with a stable PT INR between 2 and 3 were included in order to reduce any potential complex and confounding effect (n=1324). Table 5.4 lists the number of patients, if different exclusion criteria were concerned.

Among the 1496 patients, 88% (1324) of them achieved stable therapeutic PT INR of 2-3, while a set of 850 (64.2%) patients achieved stable dose during treatment. Although the inclusion of both stable INR and dose is more stringent, it will lead to 474 patients to be excluded from further analyses reducing the statistical power dramatically. We decided to carry out the analyses with consideration of stable INR criterion.

Table 5.4. Patient in the analyses with different inclusion criteria.

Inclusion criteria	Patients	Average dose	95% C.I.
All pateints	1496	33.8	(25.5, 43.6)
INR*	1324	33.8	(25.6, 44.5)
INR and dose#	850	32.5	(24.6,42.5)

For minimal three consecutive visit in clinic, *a stable INR between 2 and 3, plus #without dose variation

Univariate regression analysis was performed to test the association of inter-individual dose variation and the 216 tag SNPs on the 35 candidate genes. Figure 5.5 shows a Quantile-quantile plot of the observed versus expected p-values. Both expected and observed p-values are plotted in logarithm scale for the SNPs and the black line represents the close adherence of p-value which corresponds to the null hypothesis. The dots deviating from the black line suggest potential significant associations. The two black arrows indicate the result from the two SNPs in *VKORC1*, rs2359612 and rs9923231. The red arrow indicates the result of combined *CYP2C9**2 and *3 in association with dose. A total of 13 SNPs behave differently to the null hypothesis (above the black line) and are significant after correction for multiple tests; listed in Table 5.5.

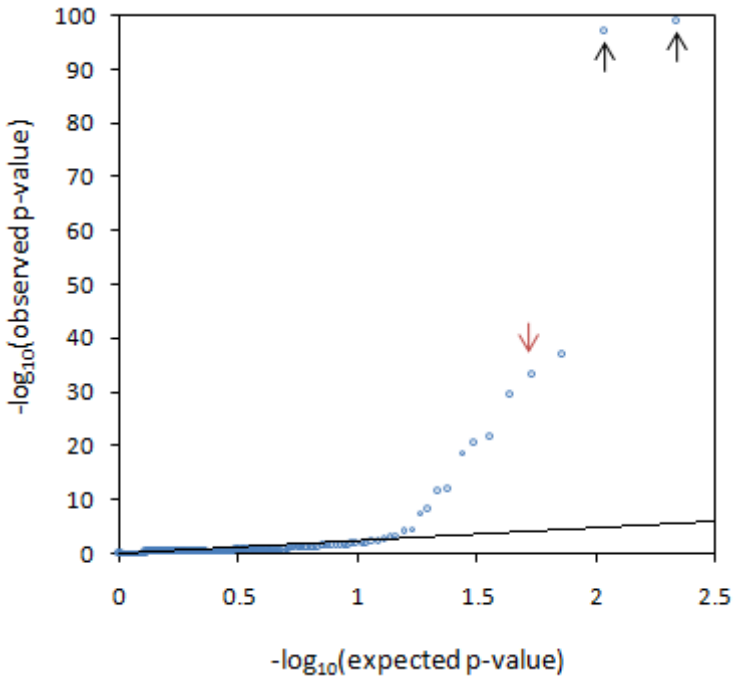


Figure 5.5. Quantile-quantile plot for univariate analysis in dose association. The black line indicates the close adherence of p-value corresponding to the null hypothesis.

A total of thirteen SNPs in *VKORC1* and the *CYP2C9-CYP2C19-CYP2C8* cluster were significantly associated with warfarin after Meff correction for multiple tests ($P \leq 2.62E-04$). The value of R^2 indicates the ratio of dose variation explained in the univariate regression model. The strongest associations with dose were observed for two nearly perfectly concordant SNPs, rs2359612 and rs9923231, in *VKORC1* which explained 29.8% and 29.3% of the inter-individual dose variance, respectively (Table 5.5).

Although rs2359612 gave a slightly better p-value associated with dose, rs9923231 has been demonstrated to potentially influence *VKORC1* mRNA expression (Rieder et al. 2005; Yuan et al. 2005). The slight difference in p-value is down to a small difference in call rate between the two SNPs. Based on the available functional evidence (see Figure 4.6), and the fact that the two SNPs are perfectly correlated, only rs9923231 was included in further analyses to represent the effect of *VKORC1*. In the group of patients with not only stable coagulation but

constant dose, SNPs rs2359612 and rs9923231 explained 33.3% ($P = 5.78x10^{-73}$) and 32.8% ($P = 3.97x10^{-72}$) of dose variation, respectively, which provides strong evidence that *VKORC1* is associated with the stable dose.

Table 5.5. SNPs showing significant association after Meff correction.

Gene	SNP	MAF	Patients	%	R^2	P-value
VKORC1	rs2359612	0.394	1292	97.58%	29.8	9.82E-100
VKORC1	rs9923231	0.393	1293	97.66%	29.3	1.03E-97
VKORC1	rs7294	0.393	1294	97.73%	12.3	1.21E-37
CYP2C9	CYP2C9 (*2 & *3)	n/a	1318	99.55%	11.8	6.63E-34
CYP2C9	rs4917639	0.209	1092	82.48%	11.7	4.61E-30
CYP2C9	rs2860905	0.214	1296	97.89%	7.4	3.05E-22
CYP2C19	rs3814637	0.073	1284	96.98%	7.1	3.14E-21
CYP2C9	rs1057910 (*3)	0.071	1321	99.77%	6.3	1.82E-19
CYP2C9	rs1799853 (*2)	0.114	1321	99.77%	4.1	1.34E-12
CYP2C8	rs11572080	0.1	1256	94.86%	4.1	3.16E-12
CYP2C9	rs1856908	0.433	1291	97.51%	2.9	6.71E-09
CYP2C8	rs10509681	0.093	1147	86.63%	3	3.08E-08
VKORC1	rs11150606	0.027	1309	98.87%	1.3	2.55E-05
CYP2C8	rs2275620	0.417	1204	90.94%	1.6	7.99E-05

Significant p-value after Bonferroni Meff correction is $P \leq 2.62x10^{-4}$.

As expected, the second best associated gene in this study is *CYP2C9*. In the univariate model, *CYP2C9**2 and *CYP2C9**3 individually explained 4.1% ($P = 1.34x10^{-12}$) and 6.3% ($P=1.82x10^{-19}$) of dose variation, respectively. However, when both alleles were taken into account simultaneously, *2 and *3 explained a sum of 11.8% of dose variation ($P = 6.63x10^{-34}$). A further analysis looking at the stable dose subgroup indicated these variants explained 12.3% of the variance ($P = 3.67x10^{-22}$).

Other than *CYP2C9*, SNPs in *CYP2C8* and *CYP2C19* displayed significant associations after correction for multiple testing (Table 5.5). However, LD analysis indicated that these

associations could be fully explained by either the *CYP2C9**2 or *CYP2C9**3 allele. In the multiple regression model (with either *CYP2C9**2 or *CYP2C9**3), none of the additional SNPs remains significant except rs3814637 in *CYP2C19*, and this shall be discussed in section 5.3.3.

The association of dose with *VKORC1* rs9923231 genotypes is illustrated in Figure 5.6. The average dose of patients having GG / AG / and AA genotypes is 42.5 mg (95 % CI: 23.3, 54.3), 32.1 mg (95 % CI: 25.8, 41.2), and 22.2 mg (95 % CI: 16.6, 27.0), respectively. The A allele has been reported to be associated with low mRNA expression (Rieder et al. 2005) which suggests lower warfarin dose could reach the therapeutic effect by blocking vitamin K recycling. In the WARG study, 529 patients are GG homozygotes whilst 714 are AG heterozygotes and 218 AA homozygotes. The patients homozygous for the G allele required a dose almost twice as much as patients homozygous for the A allele (Figure 5.6).

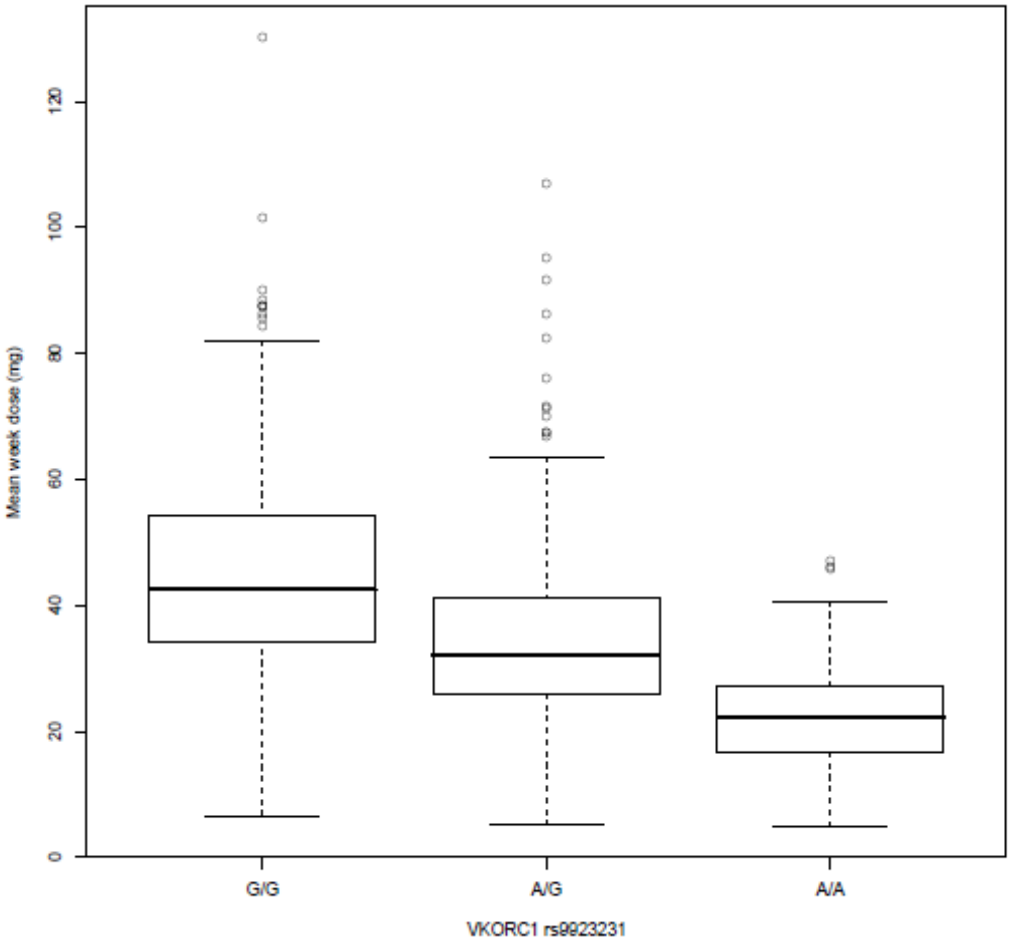


Figure 5.6. Mean weekly dose of rs9923231 genotype in WARG Swedish. The homozygous A / A patients (218) are associated with low maintenance dose whereas G / G homozygotes (529) need twice the dose compared to the A / A homozygotes.

*CYP2C9**2 and *3 alleles are known to impair the enzymatic activity of CYP2C9 in metabolising warfarin (Crespi and Miller 1997; Rettie et al. 1994; Sullivan-Klose et al. 1996). In the WARG study, patients not carrying the *2 and / or *3 alleles require an average warfarin dose of 36.79 mg/week (95% CI: 28.52, 47.56) whereas patients homozygous for the *3 allele (8 patients) required essentially low maintenance dose of 7.87 mg/week (95% CI: 5.48, 11.03) and those homozygous for the *2 allele (21 patients) needed 22.50 mg/week (95% CI: 17.23, 30.52). This result indicates how substantial the effect of *2 and *3 variants is. The patients being *1 / *2 and *1 / *3 heterozygotes require 6.74 and 8.25 mg less dose than *1/*1 homozygotes whereas *2/*3 homozygotes require a weekly dose of 19.86 mg

(Figure 5.7).

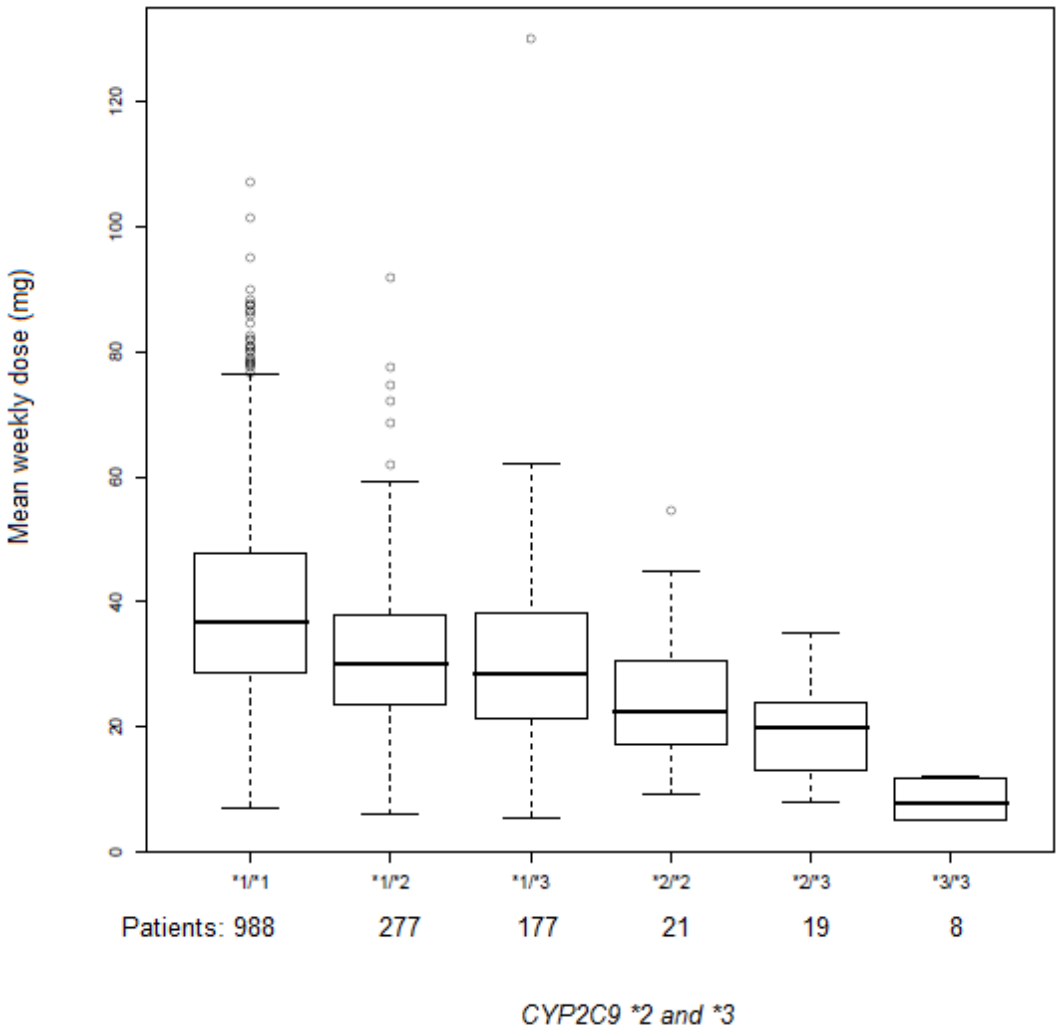


Figure 5.7. Mean weekly dose of rs9923231 genotype in WARG patients. The number of patients carrying different alleles is listed below each genotype.

5.3.3 Comparison of results in the Uppsala and WARG studies

In the Uppsala study (chapter IV), SNPs in *PROC*, *EPHX1*, *GGCX* and *ORM1-2* together explained an additional 17% of dose variation (section 4.6), and these SNPs were tested in the WARG study for replication. Table 5.6 summarises the comparison result in both Uppsala

and WARG studies of the SNPs substantially explaining dose variation in the Uppsala study.

Except *VKORC1* rs9923231 and *CYP2C9**2 and *3, other SNPs significant in the Uppsala study are not replicated in the WARG cohort (Table 5.6; R-square value indicates the explained ratio of inter-individual dose variation for each SNP). *VKORC1* explains an essential 29.3% of dose variation in WARG whilst *CYP2C9**2 and *3 explain 11.8% in univariate analysis. *PROC* SNP rs2069919 reached a gene-wise significant p-value in the Uppsala study explaining a substantial 9% of dose but reports nothing in WARG. SNPs in *EPHX1*, *GGCX* and *ORM1* reach nominal significance in Uppsala study and also explain substantial dose variation but yet do not reach significance in WARG.

The *GGCX* SNP rs4653436 was not directly typed in WARG, and instead SNP rs7568458 which is tightly linked to rs4653436 was tested. The finding in *GGCX* in the Uppsala study is not replicated, a recent Japanese study analysed 828 warfarin-treated patients and found no effect from *GGCX* (Cha et al. 2007). Meanwhile, Rieder and colleagues reported SNP rs11676382 in *GGCX* had a small but significant effect on warfarin maintenance dose (Rieder et al. 2007). Together with our finding in both Uppsala and WARG studies, the effects of GGCX are potentially population/sub-population dependent.

Although the lead SNPs in the Uppsala study for *EPHX1* and *ORM1*, rs4653436 and rs1687390, respectively, did not replicate in WARG (Table 5.6), other SNPs in these two genes, rs3817268 (P = 6.47E-03) and rs6426089 (P = 3.59E-02) in *EPHX1* and rs2787337 in *ORM1* (P = 1.08E-02) did report nominally significant associations (Table 5.7). Despite different SNPs associated with dose in the WARG cohort, however, the recurrence of the two genes suggests a possible minor effect caused by the two genes. Table 5.7 lists SNPs other

than those in the CYP2C cluster that reached nominal significance.

Table 5.6. Comparison between Uppsala and validation studies.

Genes in association	Uppsala study (n=201)		WARG study (n=1496)	
with warfarin dose	R-square	P-value	R-square	P-value
VKORC1 rs9923231	31.70%	$1.91x10^{-15}$	29.30%	$1.03x10^{-97}$
CYP2C9 *2 and *3	15.90%	$2.30x10^{-6}$	11.80%	$6.63x10^{-34}$
PROC rs2069919	9.00%	0.0002	0.20%	0.2073
EPHX1 rs4653436	4.80%	0.0084	0.10%	0.4247
GGCX rs12714145*	3.40%	0.0332	0.10%	0.4966*
ORM1 rs1687390	2.60%	0.0496	0%	0.9145

Table 5.7. Other SNPs showing nominal significant association with dose.

Gene	SNP	MAF	Patients	%	P-value
EPHX1	rs3817268	0.274	1285	97.66%	6.47E-03
PROS1	rs8178633	0.050	1218	97.73%	7.03E-03
ORM1	rs2787337	0.311	1227	99.55%	1.08E-02
CYP1A1	rs2470893	0.342	1271	82.48%	1.49E-02
CYP3A4	rs4986910	0.008	1299	97.89%	1.85E-02
PDIA5	rs1107377	0.496	1241	96.98%	2.46E-02
PROS1	rs9683303	0.353	1223	99.77%	3.26E-02
EPHX1	rs6426089	0.496	1288	99.77%	3.59E-02
PDIA3	rs11070411	0.164	1227	94.86%	4.24E-02
F10	rs2251102	0.192	1167	97.51%	4.82E-02

Apart from *EPHX1* and *ORM1* six more genes *PROS1*, *CYP1A1*, *CYP3A4*, *PDIA5*, *PDIA3* and *F10* also showed marginal association with warfarin dose. These genes are all involved in pharmacokinetics and pharmacodynamics of warfarin and require replication in other populations.

5.3.4 Multivariate

Outside *VKORC1,* SNPs showing significant association are tightly linked to either *CYP2C9**2 or *3 alleles in the WARG study. In the analysis of the Uppsala cohort, rs4917639 explained a substantial 10% and 2.5% of dose in the multivariate model including *CYP2C9**2 or *3 as covariates, respectively (section 4.4). The LD analysis indicated that this SNP is associated with both *2 and *3 and displayed a sum effect of both alleles. This association is also observed in the WARG study. The SNP rs4917639 explained approximately 11.7% ($P = 4.61x10^{-30}$) and 12.1% ($P = 1.99x10^{-20}$) of dose variation in models including either *2 or *3 alleles. It fails to explain any dose in the multiple model including both *2 and *3 allele. The LD between rs4917639 and *2 / *3 is present in a lesser extent in WARG ($r^2 = 0.836$) than in the Uppsala study but could still reflect the outcome of significant p-value and the substantial explanation of dose variation.

Apart from rs4917639, rs3814637 is the only SNP significant and moderately independent to *CYP2C9**2 and *3 in the WARG study ($P = 2.2x10^{-6}$) after Meff correction, although it is tightly linked to *CYP2C9**3 in the Uppsala study. However, in the multiple regression model including *VKORC1* rs9923231 and *CY2C9**2 / *3 alleles, *CYP2C19* SNP rs3814637 explained a merely 0.7% of dose variation and was therefore not considered for inclusion in the final regression model. However, this result does suggest a contribution of *CYP2C19* in warfarin dose requirement.

Except *VKORC1* and *CYP2C9**2 and *3, no SNP explains substantial dose variation in the WARG study. The promoter SNP, rs9923231, in *VKORC1* and *CYP2C9**2 / *3 are together as the only genetic determinants with broad utility for determining warfarin dose.

5.4 REGRESSION MODEL

In the Uppsala study, non-genetic factors including age, gender, body weight, medical indication, drug interaction, and PT INR value were tested and only age, body weight, and drug interaction were included in the final multiple regression model (section 4.6). In the WARG study, information of patient's body weight was not recorded and thus gender was included as a surrogate in the multiple model.

The predictors in the multiple model are listed in Table 5.8 including non-genetic factors age, gender, and drug interaction as well as *VKORC1* rs9923231 and *CYP2C9* *2 and *3 as the genetic determinants. Notably, the drug interaction is significantly associated with dose variation if an overall outcome is potentiating. The dose explained (R^2) and relevant p-value for each predictor is listed on Table 5.8.

Table 5.8 Predictors in multiple regression model.

Predictor		R^2	Patients	P-value
CYP2C9 *2 and *3		11.8	1318	6.63E-34
VKORC1 rs9923231		29.3	1293	1.03E-97
Age		14.5	1324	8.00E-47
Gender		1.2	1324	6.89E-05
Drug interaction	(potentiate)	1.5	1324	8.00E-06
	(decrease)	0	1324	9.03E-01

To achieve the best available coefficient for each predictor, cross validation was applied to test each model. The entire WARG database was thereby divided into two portions, 70% patients as the training dataset whilst 30% were tested for validation for performance. The

parameters for each predictor were determined after 1000 iterations of cross validation and the final model could explain 58.7% of warfarin dose variation (Figure 5.8A).

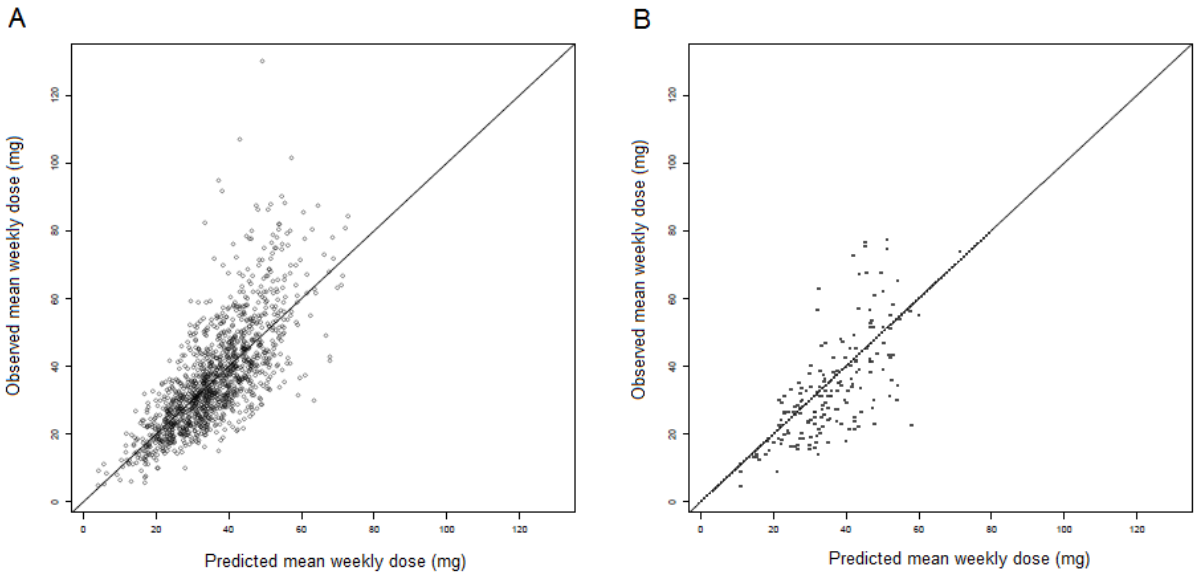


Figure 5.8. Multiple regression model developed accordingly in the WARG study. This model explains (A) 58.7% dose variation in the WARG cohort and (B) 53% in the Uppsala cohort.

The model was then used to estimate the patients in the Uppsala study. Information was available for all predictors in 181 individuals. The developed model predicts slightly higher dose for Uppsala patients, and 53% of dose variation was successfully predicted (Figure 5.8B). Interestingly, results from both WARG and Uppsala studies shows that some outliers required higher dose than the predictive amount, this suggests that some elements remain to be discovered.

5.5 THERAPEUTIC STABILISATION

Warfarin is difficult to administer because of its narrow therapeutic range of PT INR 2-3. When a patient is prescribed with warfarin, a trial dose is given and PT INR is closely monitored before increasing or decreasing the dose. Over anti-coagulation often leads to severe bleeding complication and hospitalisation. Therefore, it is important to investigate the factors which might contribute to the required time of reaching stable INR.

5.5.1 INR stabilisation

The INR values of each patient throughout the treatment period were analysed and plotted with LOESS regression model according to their *VKORC1* rs9923231 genotypes (Figure 5.9). LOESS regression is widely used in analysing localised data subsets and variation in the data point by point without specifying a global function and is, therefore, suitable in analysing PT INR value for each patient over different time points.

Patients with homozygous AA genotypes (green dotted line) had an INR peak at 2.7 before moving towards 2.5 of stable INR in the first week of therapy. Meanwhile, AG heterozygote patients showed slight fluctuation compared to the homozygous GG patients until they were stabilised. This result shows that AA homozygotes for *VKORC1* rs9923231 require lower warfarin dose, but patients were over dosed in the beginning of treatment resulting in over-anticoagulation and unstable PT INR in the first week. Our results suggest that this could be rectified if *VKORC1* genotype information is used. Once the INR is stabilised, *VKORC1* genotype has no effect on further anticoagulation stability.

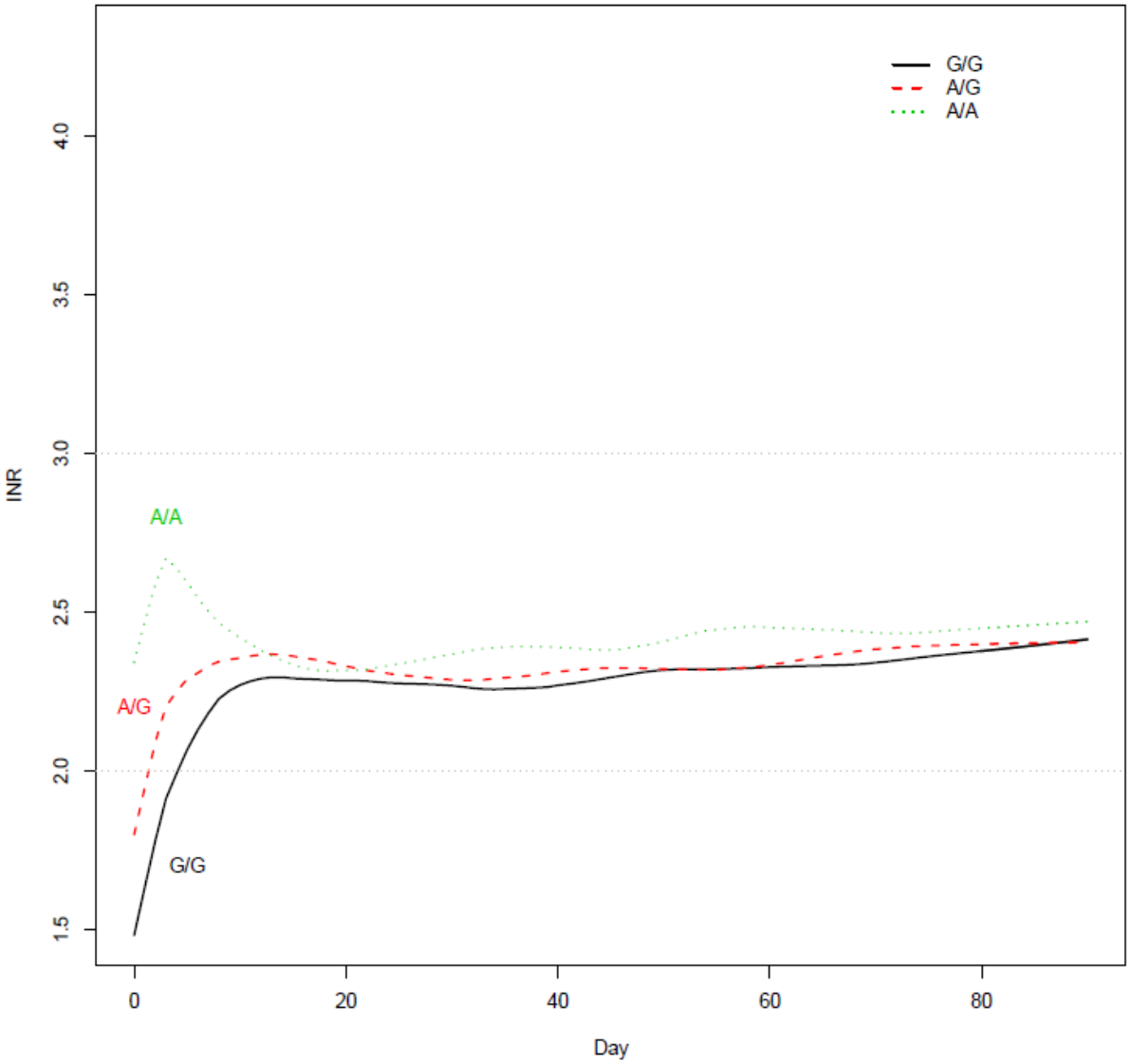


Figure 5.9. Lowess smoothed plot of PT INR values of patients treated with warfarin. Patients with different *VKORC1* rs9923231 genotype were plotted over treatment period.

*CYP2C9**2 and *3 alleles were analysed in LOWESS regression analysis for their effect on PT INR stability. Patients homozygous for *CYP2C9**3 had also extremely unstable INR values, with an average INR peak exceeding 4 during the first two weeks, and subsequent instability for three months (pink line, Figure 5.10). As a consequence of impaired enzymatic activity of CYP2C9, patients homozygous for the *3 allele were very sensitive in response to the increase or decrease in warfarin dose. This PT INR instability is also reflected in patients carrying *2 / *3 (light blue) and *2 / *2 alleles. Warfarin metabolism in heterozygotes

carrying *1 / *2 or *1 / *3 were compensated by the normal allele and they require a comparatively stable dose than those homozygous and heterozygous for *2 and *3 alleles. Only patients homozygous for the *1 allele (solid black line) reached stable anticoagulation of PT INR 2.5 in two weeks.

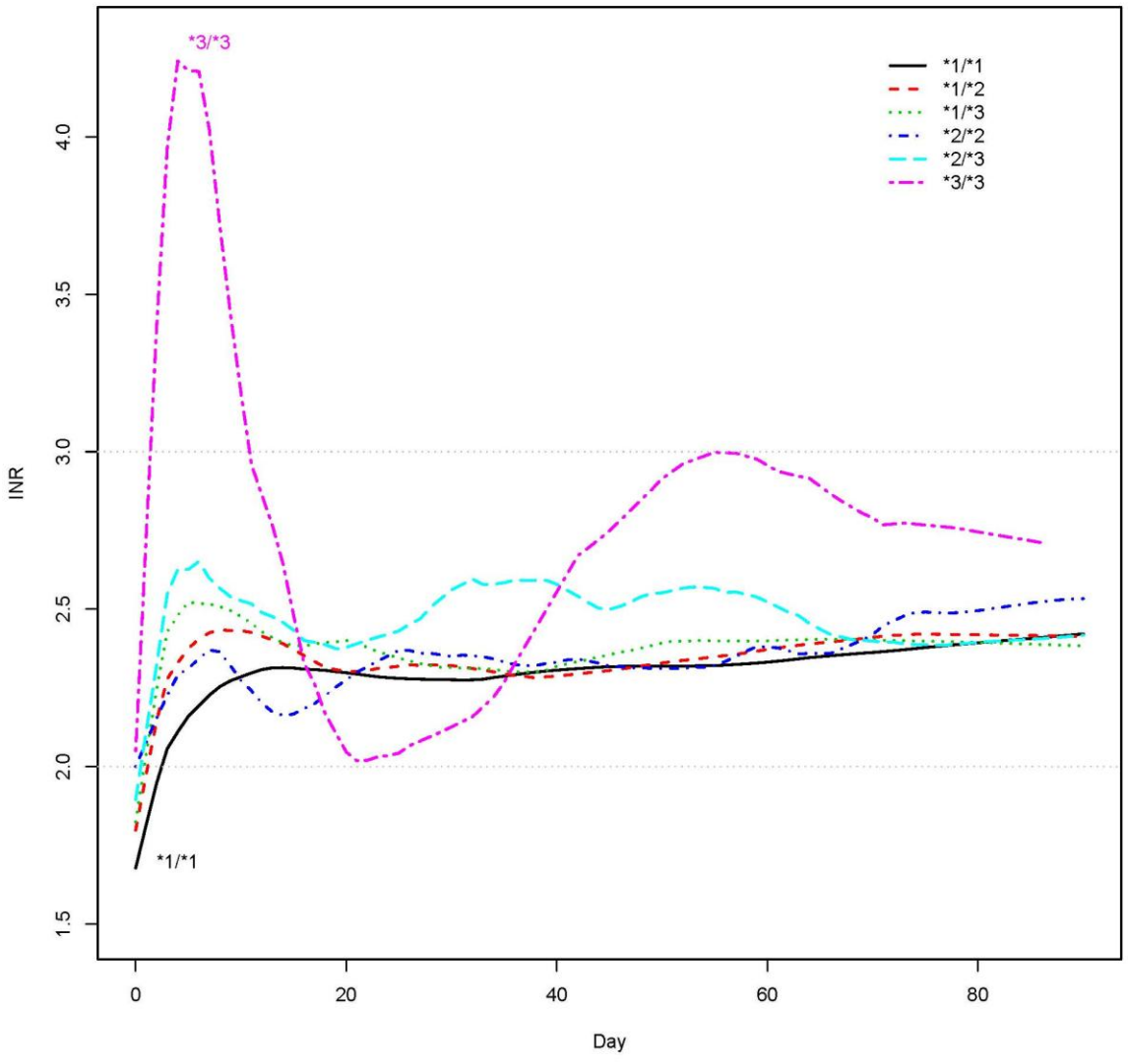


Figure 5.10. Lowess smoothed plot of PT INR values of patients treated with warfarin. Patients carrying normal or *CYP2C9**2 and / or *3 alleles were plotted over treating period.

5.5.2 Over-anticoagulation

To investigate how genetic variants influence over-anticoagulation in the first month and especially the first week treatment, all tag SNPs in the 35 candidate genes were also tested for association, with over-anticoagulation defined as PT INR > 4 within first five weeks of treatment, with log rank test. Log rank test (also called Mantel-Haenszel test or the Mantel-Cox test) is widely used clinically to test the survival distributions of different samples. Each SNP in the 35 candidate genes was tested and the type I error (p-value) was estimated. Table 5.9 only lists the SNPs showing significant p-value after correction for multiple tests.

Table 5.9. SNPs associated with over anti-coagulation in warfarin treatment.

Gene	SNP	MAF	Pateitns	P-value
CYP2C19	rs3814637	0.073	1447	1.11E-16
CYP2C9	rs1057910 (*3)	0.071	1493	1.11E-16
CYP2C9	CYP2C9*2 and *3		1490	1.11E-16
VKORC1	rs2359612	0.394	1455	4.19E-12
VKORC1	rs9923231	0.393	1461	6.56E-12
CYP2C9	rs4917639	0.209	1220	9.78E-06

Five SNPs are significantly associated with over-anticoagulation. The effect of rs3814637 in *CYP2C19* and rs4917639 in *CYP2C9,* are known to be well explained by *CYP2C9**2 and *3 (sections 4.4 and 5.3.4) and were therefore removed from further analysis. Despite the fact that it was not even nominally significant in the test, *CYP2C9**2 was analysed together with the *3 allele. As explained previously, rs9923231 was used for *VKORC1* (section 6.3.2). Therefore, only rs9923231 in *VKORC1* and *CYP2C9**2 / *3 are significantly associated with over-anticoagulation in the first five weeks treatment.

The effect of *VKORC1* rs9923231 and *CYP2C9**2 / *3 in over-anticoagulation was analysed in Cox regression model and illustrated with Kaplan-Meier curve plot. Cox regression model is a survival model for describing the risk changes over time, such as the choice of treatment or the effect of genotype. The advantage of Cox model is that it does not require the calculation of the hazard function. Kaplan-Meier curve is frequently used for illustration of survival function of life-time data. In medical statistics, this is often used to illustrate cumulative probability of clinical events, such as gain of tumour or survival after treatment.

Patients who are AA homozygotes for *VKORC1* SNP rs9923231 appeared to have a significantly increased probability of over-anticoagulation (dotted line, Figure 5.11A). In the first 5 weeks of treatment, 4.56 hazard ratio (95% CI: 2.85, 7.30) was observed in AA homozygous patients with PT INR above 4 comparing to GG homozygous patients (P = 2.4E-10). When the patients is heterozygous, i.e. carrying a G allele, a much lower hazard ratio of 1.74 (95% CI: 1.11, 2.71) was observed (P = 1.5E-02) in comparison with GG homozygotes (Figure 5.11A).

Figure 5.11B shows the effect of *CYP2C9**2 and *3 alleles. The over-anticoagulation is significantly severe in patients carrying homozygous CYP2C9*3 allele (pink dash line, Figure 5.11B). A significant increase of hazard ratio of 21.84 (95% CI: 9.457, 50.42) is observed in patients homozygous for *3 allele. Patients heterozygous for *2 / *3 alleles showed a hazard ratio of 2.98 (95% CI: 1.092, 8.15) comparing to *1 / *1 homozygotes (Figure 5.11B).

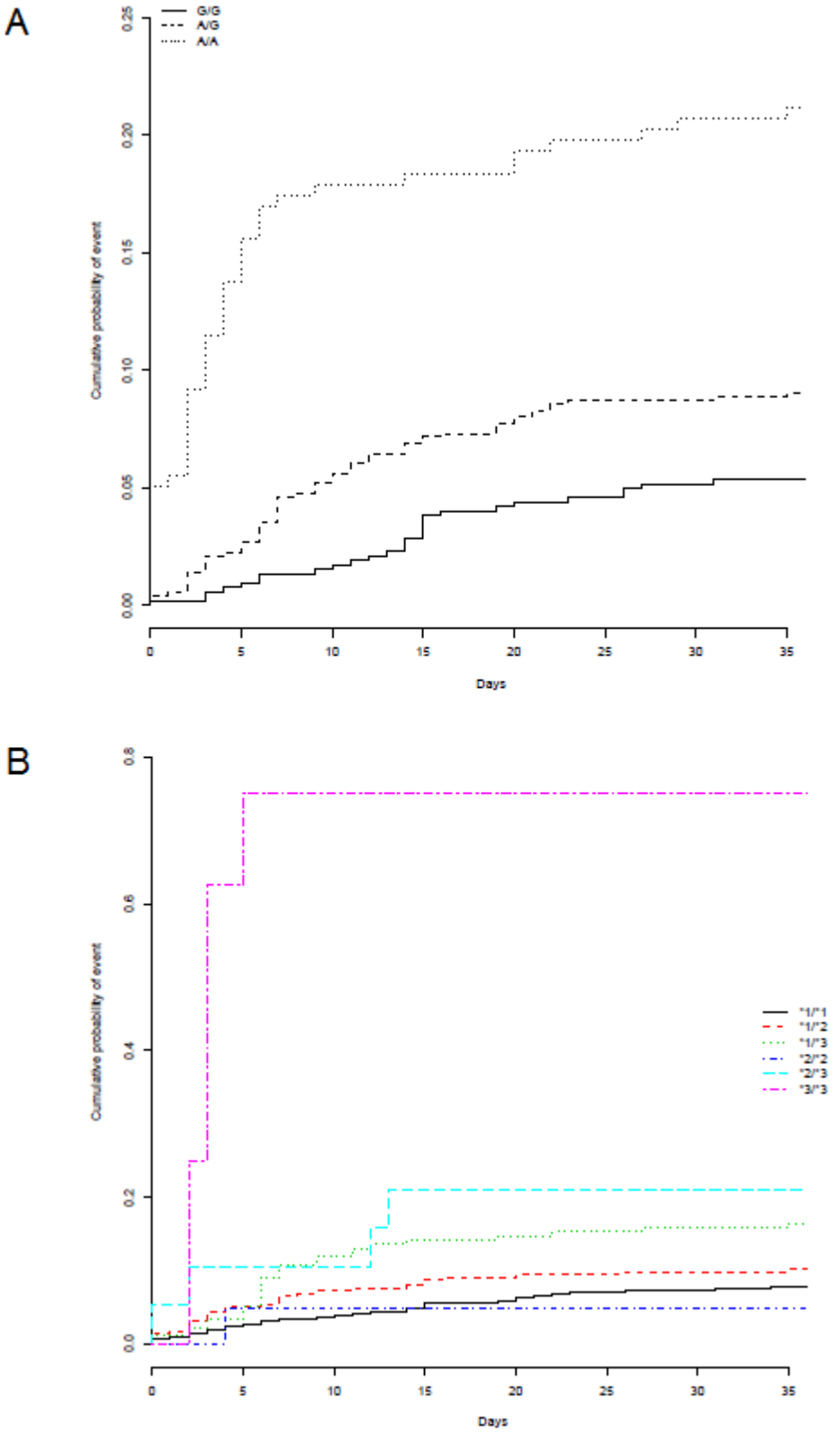


Figure 5.11. Survival (Kaplan-Meier) curve of cumulative probability in patients with PT INR > 4 related to (A) *VKORC1* rs9923231 and (B) *CYP2C9**2 and *3.

5.6 CONCLUSION

The WARG study provided a validation of the initial findings in the Uppsala study and a *de novo* investigation for the impact of the 35 candidate genes on warfarin dose requirement. Only genotypes of *VKORC1* and *CYP2C9* have a demonstrated effect on warfarin dose. Moreover therapeutic stability, that is stable INR and dose, is also influenced by *VKORC1* and *CYP2C9*. The inconsistent findings between the Uppsala and WARG studies suggest minor effects of these genes which may be due to study design/being treatment specific. These small effects were therefore excluded in the development of a global dosing algorithm. Meanwhile, sample size is inevitably important in identifying contributors in association studies.

The CYP2C gene cluster on chromosome 10 has been intensively studied for various drugs. Different methods have been suggested to tag the polymorphisms in this region (Ahmadi et al. 2005; Walton et al. 2005). In the WARG study, this region is comprehensively tagged with 25 SNPs and no effect other than the *CYP2C9**2 and *3 alleles is found.

CHAPTER VI

WARFARIN AND ADVERSE DRUG REACTION

6.1 INTRODUCTION

As described in the introduction, there are various side effects of warfarin therapy. The most and only common adverse drug reaction (ADR) is haemorrhage. Over-anticoagulation and infrequent PT INR monitoring often result in an increased risk in bleeding. The risk of haemorrhage ADR and fatal bleeding has been reported from 0.9 - 2.7% and 0.07 - 0.7%, respectively (Landefeld and Beyth 1993).

To date, a number of risk factors predisposing to bleeding upon warfarin treatment have been suggested. Patients aged older than 75 years, with concomitant atrial fibrillation, have substantially increased risk for intracerebral bleeding (1994; Albers 1994). Interestingly, bleeding that occurs with a PT INR of less than 3 is often associated with an underlying occult gastrointestinal or renal lesion (Levine et al. 1995). Other potential risks of co-morbid diseases include hypertension, cerebrovascular disease, serious heart disease, and renal insufficiency (Hull and Pineo 1995; Routledge et al. 1979).

Cancer patients have an approximately three-fold increase in the rate of recurrent thrombosis and a two-fold increase in the rate of major bleeding during warfarin treatment of deep vein thrombosis (DVT). These complications occur mostly during the first few months of anticoagulation and do not reflect under-anticoagulation or over-anticoagulation but correlate with the extent and severity of the underlying cancer; increased bleeding may be related to bleeding at the primary tumour site (Prandoni et al. 2002). To search for the underlying genetic factors causing severe bleeding ADR, it is important to classify bleeding patients from oncology patients. The confounding effect from oncology patients may dilute the signal from true causations of serious bleeding.

6.2 PATIENTS

In the Uppsala study, 12 out of 201 (5.9%) warfarin treated patients were found having a bleeding episode in their medical record. An additional 24 patients with bleeding complications were recruited from other anticoagulation clinics through Dr Mia Wadelius (Uppsala). In the prospective WARG study, 28 patients (1.9%) have been recorded with severe bleeding complication. Compared to the Uppsala study which is retrospective, the WARG study, with a naturalistic study design, reflects more accurately the real scenario of bleeding complication in warfarin-treated patients. Various definitions of major bleeding have been described (Lindh et al. 2007). In the Uppsala study, including the 24 additional bleeders, severe bleeding is defined as requiring hospital care with exclusion criteria for thrombolysis, surgery or trauma immediately preceding the bleeding. In the WARG study, the World Health Organisation (WHO) definition of severe bleeding, which is: lethal, life-threatening, permanently disabling, or leading to hospital admission (emergency room admissions excluded) or prolongation of hospitalisation, was complied (Lindh et al. 2007). Among a total of 64 bleeders in cohorts, 16 patients and 18 patients in the Uppsala and WARG cohorts, respectively, were within target INR at the time of the bleed.

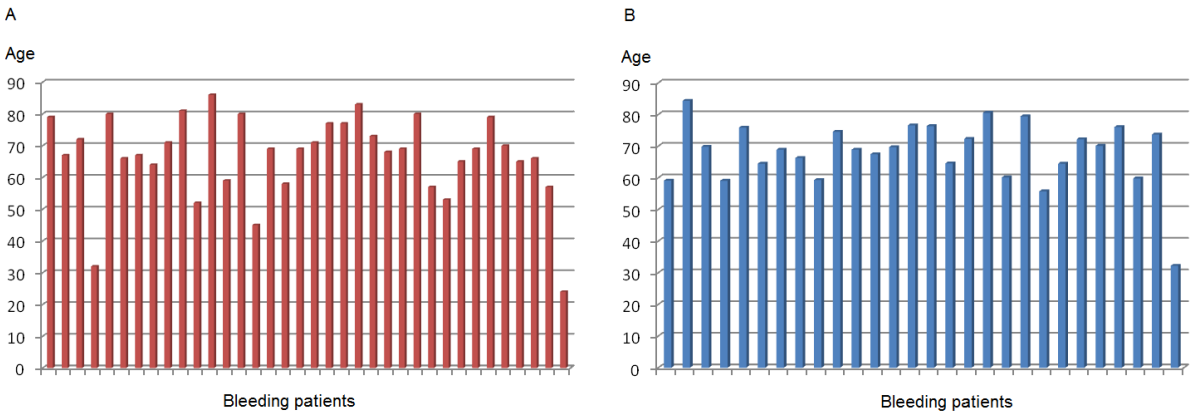


Figure 6.1. Age of each bleeding patient collected in the Uppsala and WARG studies.

Figure 6.1 shows the age of the 64 bleeders collected through the two studies namely Uppsala University (Figure 6.1A) and WARG (Figure 6.1B). It is clear that the haemorrhage recorded upon warfarin treatment happened predominantly among people aged over 50 years in the Uppsala cohort and over 60 years in the WARG cohort. Only 4 individuals (6.25%) aged below 50 years were identified in both cohorts with bleeding ADR (Figure 6.2). This can be explained by the fact that patients receiving warfarin treatment are predominantly older: only 13% of patients were aged below 50 years in a total of 1657 warfarin treated patients (both studies combined).

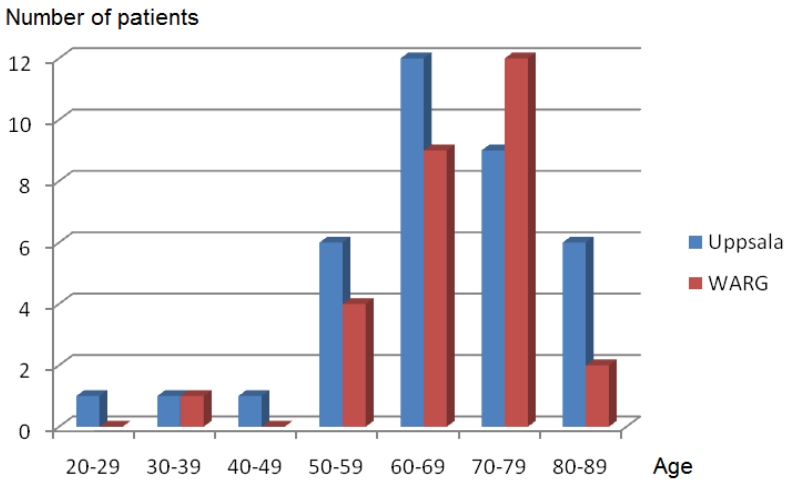


Figure 6.2. Age distribution of bleeders in the Uppsala and WARG studies.

For the purpose of this investigation bleeders will be referred to as ‘cases’ and non-bleeders as ‘controls’, all patients were under warfarin treatment. Table 6.1 shows comparison of age distribution for both case and control patients in the Uppsala and WARG cohorts. Generally, the age distribution of bleeders match the distribution of control patients and the bleeding risk is not particularly associated with age. The majority of bleeding patients were aged between 60 and 80 years (66%, 42 individuals) and this corresponds to the result of the control group. Among the 64 bleeders, the youngest was 24 years and the oldest 86 years old (Table 6.1).

Table 6.1. Age distribution of case/control in the Uppsala and WARG studies.

Age	CASE				CONTROL			
	Uppsala	%	WARG	%	Uppsala	%	WARG	%
10-19	0	0.0%	0	0.0%	0	0.0%	1	0.1%
20-29	1	2.8%	0	0.0%	1	0.5%	31	2.1%
30-39	1	2.8%	1	3.6%	1	0.5%	61	4.2%
40-49	1	2.8%	0	0.0%	9	4.8%	112	7.6%
50-59	6	16.7%	4	14.3%	41	21.7%	243	16.6%
60-69	12	33.3%	9	32.1%	44	23.3%	425	29.0%
70-79	9	25.0%	12	42.9%	70	37.0%	451	30.7%
80-89	6	16.7%	2	7.1%	23	12.2%	141	9.6%
90-99	0	0.0%	0	0.0%	0	0.0%	3	0.2%
Subtotal	36		28		189		1468	

6.3 STATISTICAL POWER

The biggest challenge of most, if not all, ADR studies is the number of available cases, that is, the serious adverse events are typically very rare in a regional area. Furthermore, clinical practice aims at minimising the risk of patients in developing an ADR. Heritability of 'warfarin bleeding' as a trait is unknown but at ~5%, occurrence among patients with mainly heart conditions, could be due to common variants with low frequency lower than 5% such as nsSNPs (Smyth et al. 2006). Whether it is caused by variant(s) in a single or multiple genes it is not possible to say but 'warfarin bleeding' does not appear to have the profile of a monogenic Mendelian disease which is caused by a deleterious rare mutation in the population.

The Uppsala and WARG cohorts have 36 and 28 bleeding patients, respectively. It is clear that this sample size can barely provide the power to detect common variants with strong effects. With this caveat in mind one can at least increase statistical power and maximise the chance in finding the genetic determinants by (i) combining the bleeders from the two studies, 64 in total, and more importantly (ii) using a much larger set of controls; the two studies together have 1657 control patients. However, even a total of 64 cases is still underpowered to detect variants with moderately strong effects. If the causative variant has a moderately strong effect, with odds ratio (OR) of 2, the number of cases required to detect this variant assuming it has a MAF of 1%, 5%, 10%, 25% and 50%, is 895, 201, 116, 72 and 77, respectively; with type 1 error (p-value) set at 0.001 and statistical power at 80% (Figure 6.3). For strong effects, ORs of 3 or 4, our study has some power (Figure 6.3). Under all scenarios, any finding in this study will require replication in a larger cohort.

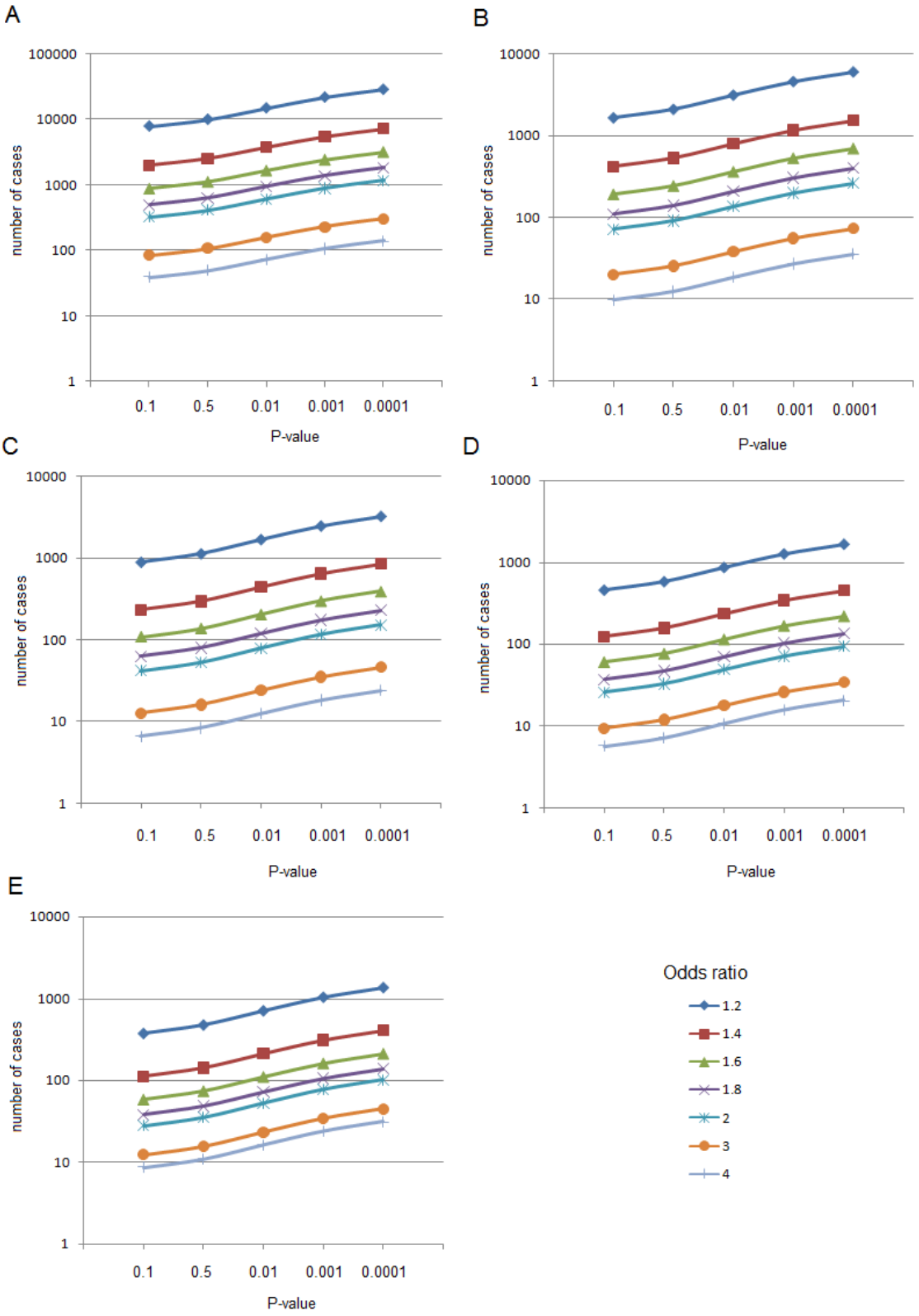


Figure 6.3. Statistical power calculation of cases required to achieve 80% power based on the assumption of 2% bleeding prevalence in population of causative variants of (A) 0.01; (B) 0.05; (C) 0.1; (D) 0.25; (E) 0.5 in MAF.

In the following sections, all bleeders from both Uppsala and WARG cohorts will be initially analysed and discussed with the use of single SNP and 2- and 3-marker haplotype analyses (section 6.4). However, since the bleeding phenotypes and recruitment criteria of the cases in two cohorts are different, same analyses are performed on each cohort independently (section 6.5). Finally, re-sequencing on 11 candidate genes on the Uppsala bleeders were performed and analysed to identify highly penetrant allele(s), if there are any, which is/are associated with severe bleeding ADR (section 6.7).

6.4 CASE-CONTROL ANALYSIS - ALL BLEEDERS

6.4.1 Single marker

The Cochran-Armitage trend test and Fisher's exact test were used to examine the association of genetic determinants and bleeding ADR. If the severe bleeding was a Mendelian trait, the causative SNP(s) would most likely fail in a Hardy-Weinberg equilibrium (HWE) test and be considered as a 'bad' SNP. The Cochran-Armitage test does not assume HWE and instead of the allele, the individual is considered as the unit of association analysis. Compared to other tests, it has the advantage in including SNPs which would be excluded in other tests and it is used to detect the association for the genotype, instead of allele and allele frequency in case/control studies. However, knowing the small number of cases in our study, Fisher's exact test was also performed to validate results.

Table 6.2 lists all the SNPs showing nominal significance ($P < 0.05$) for both tests. Among the 35 genes tested, nine genes gave nominal association with bleeding. The genes *P4HB*, *PDIA4*, and *PDIA6* which are involved in providing electrons to reduce the thioredoxin-like centre in *VKORC1*, the *CYP2C8* gene, the constitutive androstane receptor-beta gene (*NR1I3*) which is known for regulating cytochrome P450, the *EPHX1* gene, the prothrombin gene (*F2*), the protein C gene (*PROC*) which is a natural anticoagulant and down-regulates blood clotting and the *ABCB1* gene encoding the membrane receptor involved in clearance of warfarin. SNPs in *PDIA4* and *P4HB* are significant ($P < 0.01$) in the trend test (Table 6.2) but did not survive after Bonferroni Meff correction for multiple testing ($P < 2.62 \times 10^{-4}$).

Table 6.2. SNPs nominally associated with bleeding complication in the Cochran-Armitage trend test and Fisher's exact test.

TREND			Fisher Exact test					
SNP	P value	Gene	SNP	MAF (Case)	MAF (Control)	P value	OR	Gene
rs1008587	2.72E-03	*PDIA4*	rs2502804	0.10	0.23	3.04E-03	0.3548	*NR1I3*
rs4727005	3.19E-03	*PDIA4*	rs4727005	0.32	0.20	4.93E-03	1.849	*PDIA4*
rs1027256	4.02E-03	*PDIA4*	rs1008587	0.31	0.19	5.92E-03	1.873	*PDIA4*
rs1026910	4.72E-03	*PDIA4*	rs1027256	0.32	0.20	7.68E-03	1.862	*PDIA4*
rs1799919	6.55E-03	*P4HB*	rs1026910	0.30	0.19	8.23E-03	1.828	*PDIA4*
rs6464929	7.79E-03	*PDIA4*	rs1799919	0.34	0.21	1.11E-02	1.859	*P4HB*
rs1536430	1.06E-02	*CYP2C8*	rs6464929	0.31	0.20	1.12E-02	1.787	*PDIA4*
rs2502804	1.12E-02	*NR1I3*	rs1686482	0.37	0.50	1.55E-02	0.5849	*PDIA6*
rs1686482	1.56E-02	*PDIA6*	rs6464930	0.28	0.18	2.49E-02	1.696	*PDIA4*
rs6464930	2.16E-02	*PDIA4*	rs1533756	0.29	0.40	2.61E-02	0.617	*P4HB*
rs3753661	2.18E-02	*EPHX1*	rs2282687	0.05	0.13	2.70E-02	0.377	*F2*
rs1533756	2.36E-02	*P4HB*	rs2069933	0.18	0.28	3.13E-02	0.5498	*PROC*
rs2282687	3.00E-02	*F2*	rs1536430	0.08	0.02	3.24E-02	3.622	*CYP2C8*
rs2069933	3.02E-02	*PROC*	rs1247176	0.08	0.16	3.38E-02	0.4652	*PDIA6*
rs5898	3.55E-02	*F2*	rs3753661	0.13	0.07	3.42E-02	2.051	*EPHX1*
rs1247176	3.96E-02	*PDIA6*	rs1734343	0.39	0.50	4.38E-02	0.6571	*PDIA6*
rs1734343	4.57E-02	*PDIA6*	rs5898	0.16	0.09	4.72E-02	1.837	*F2*
rs2214102	4.67E-02	*ABCB1*						

All tested SNPs were then examined in Fisher's exact test. All genes that showed association in the trend test survived Fisher's exact test, except P-glycoprotein (*ABCB1*, P = 0.06). The odds ratio in Fisher's test ranges between 3.6 (*CYP2C8*) and 1.7 (*PDIA6*). These SNPs are all common variants in the Swedish population with MAF between 5.2% (rs2282687, *F2*) and 37.2% (rs1686482, *PDIA6*). This result indicated that none of the examined candidate genes is showing evidence of warfarin-induced bleeding complication being a Mendelian trait. However, we cannot exclude that a common variant is tagging a haplotype that is carrying a rare variant(s).

6.4.2 Two- and three-maker sliding window haplotype

Beside single marker analysis, it is also important to examine whether haplotypic analysis provides additional information in explaining the disease state. For example, a rare causative

variant which sits on a specific haplotype is most likely going to be missed by single marker analysis but will be detected by haplotype analysis. Therefore, our dataset was used to estimate two- and three-marker haplotypes generated by a sliding window; which were then tested for association with bleeding.

6.4.2.1 Two marker haplotype

The results of the two-marker haplotype association analysis are listed in Table 6.3; all haplotypes reaching nominal significance or below (p-value below 0.05). The most significant association was seen with *PDIA4* (P = 2.84E-03). Compared to the single marker analysis, most genes/SNPs showing nominal significance in the Cochran-Armitage trend test were also reported in haplotype analysis except rs2069933 in protein C. The window with rs2056530 and rs2472677 in *NR1I2* did not pass correction for multiple testing. None of the haplotypes listed in Table 6.3 provides additional information compared to the single marker analysis (Table 6.2).

Table 6.3. P-value of two marker sliding window.

Gene	SNPs in sliding window		Omnibus p-value
PDIA4	rs10085877	rs4727005	2.84E-03
F2	rs2282687	rs3136516	3.88E-03
CYP2C8	rs1536430	rs1058930	1.04E-02
F2	rs2070852	rs5898	1.22E-02
CYP2C8	rs947173	rs1536430	1.26E-02
F2	rs5898	rs2282687	1.56E-02
EPHX1	rs3753661	rs2671272	2.61E-02
ABCB1	rs2214102	rs2214101	2.72E-02
NR1I2	rs2056530	rs2472677	2.94E-02
NR1I3	rs2502804	rs9332618	3.98E-02
PDIA4	rs6464929	rs1551927	1.42E-02
PDIA6	rs1686482	rs1198873	4.08E-02
P4HB	rs1533756	rs1010954	4.40E-02

Although these genes did not have a very significant omnibus p-value (p-value for overall tests with all phased haplotypes in a specified sliding window), further analysis was undertaken by looking at each haplotype separately (Figure 6.4). The second most significant gene is prothrombin (*F2*) and in particular the sliding window containing rs2282687 and rs3136516 (P = 3.88E-03). The two-marker haplotype T-G (rs2282687 and rs3136516) has a frequency of 31.6% in the control group which increases to 46.8% in the bleeder group, giving a p-value of 2.06E-03 (Figure 6.4A). This increase in frequency suggests that a variant in *F2* may lead to an increased susceptibility of bleeding risk.

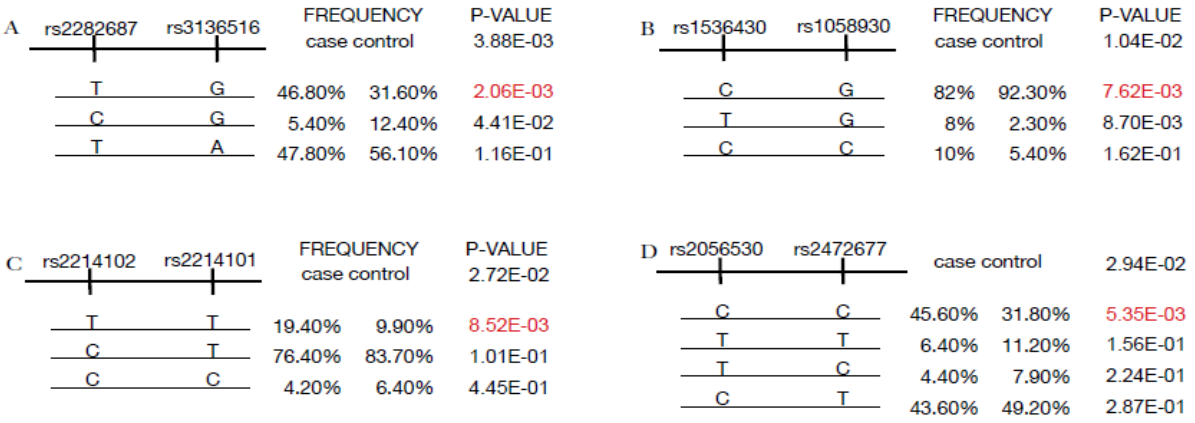


Figure 6.4. Bleeding association of individual haplotype in (A) *F2*, (B) *CYP2C8*, (C) *ABCB1* and (D) *NR1I2*. Omnibus p-value was listed on the top row for each gene whilst p-values for each haplotype in the window were noted next to corresponding haplotype and frequencies in case and control groups.

Figure 6.4B shows the most associated window in *CYP2C8* and here it is the C-G haplotype (rs1536430 and rs1058930), which is associated with a 10% drop in frequency in cases suggesting a protective mechanism in preventing bleeding. In *ABCB1*, the SNP rs2214102 was nominally significant in single marker association (Table 6.2, P = 4.67E-02). The two-marker T-T haplotype (rs2214102- rs2214101) is associated with a 10% increase in cases

(Figure 6.4C) and gives a marginally lower p-value (P = 8.52E-03). Although none of the tested *NR1I2* SNPs gave an association signal, the C-C haplotype (rs2056530-rs2472677) might be of interest (figure 6.4D). The frequency of this haplotype increased from 31.8% in controls to 45.6% in cases. All the haplotypes described so far gave significance p-value above 0.001 but none of them passed Bonferroni Meff correction (P = 2.62E-04).

6.4.2.2 Three marker haplotype

Table 6.4 lists all nominally significant (P < 0.05) sliding windows from the three marker analysis. Only six genes showed association to bleeding which reached nominal significance. Compared to the two marker analysis, protein S (*PROS*) and epoxide hydrolase 1 (*EPHX1*) appear as new candidate genes for association with bleeding. Although the three-marker haplotype rs100855877-rs4727005-rs10272564 in *PDIA4* did not pass the p-value threshold for Bonferroni Meff correction, it provided, in the single marker and two marker haplotype analyses, the strongest signal. This recurrent signal suggests that *PDIA4* may be of interest in association with the risk of warfarin-induced bleeding ADR. None of the 3-marker haplotypes in the other genes listed in Table 6.4 remained significant after correction for multiple tests.

Table 6.4. P-value of three marker sliding window.

Gene	SNPs in sliding window			Omnibus p-value
PDIA4	rs10085877	rs4727005	rs10272564	2.46E-03
F2	rs5898	rs2282687	rs3136516	8.44E-03
F7	rs3093229	rs3093230	rs2774030	1.30E-02
F2	rs2070851	rs2070852	rs5898	1.60E-02
PROS	rs8178633	rs4857037	rs4857343	2.41E-02
F2	rs2070852	rs5898	rs2282687	2.64E-02
NR1I2	rs2056530	rs2472677	rs2461818	3.36E-02
EPHX1	rs3753661	rs2671272	rs2671270	4.24E-02
PDIA4	rs10269104	rs6464929	rs1551927	5.41E-03
PDIA4	rs6464929	rs1551927	rs6464930	1.23E-02
P4HB	rs876017	rs1533756	rs1010954	4.21E-02

6.4.3 Summary of findings for all bleeder groups

Table 6.5 summarises the results for single marker trend test and haplotype analysis with 2- and 3-marker window. For each gene showing an association of nominal significance in a given analysis, only the test result with the lowest p-value is listed. For the trend test result, the p-value, minor allele frequency in both case and control groups, and odds ratio are included, whilst for haplotypic results the omnibus p-value and the frequency in both cases and controls is reported.

Table 6.5. Bleeding association of single marker, 2 SNPs haplotype and 3 SNP haplotype.

Gene	Test	p-value	Omnibus p-value	Allele Freq (Case)	Allele Freq (Control)	OR	Lowest haplotype p-value	Haplotype freq (case)	Haplotype freq (control)
PDIA4	Trend	2.46E-03		31%	19%	1.873			
F2	2SNP		3.88E-03				2.06E-03	47%	32%
P4HB	Trend	6.55E-03		34%	21%	1.859			
CYP2C8	2SNP		0.010				7.62E-03	82%	92%
NR1I3	Trend	0.011		10%	23%	0.3548			
PDIA6	Trend	0.016		37%	50%	0.5849			
EPHX1	Trend	0.022		13%	7%	2.051			
PROS	3SNP		0.024				0.035	13%	7%
ABCB1	2SNP		0.027				8.52E-03	19%	10%
NR1I2	2SNP		0.029				5.35E-03	46%	32%
PROC	Trend	0.030		18%	28%	0.5498			

A total of twelve genes may be of interest to further investigate and replicate in different cohorts: The odds ratios range between 1.7 (*PDIA6*) and 2.8 (*NR1I3*). No genes remained significant in all tests after Bonferroni Meff correction. All the alleles and haplotypes are relatively common in this Swedish sample, apart from the 3-marker haplotype in coagulation

factor 7 which has a frequency of 2% in controls and 6% in cases. As mentioned earlier, although we have a relatively large control group, the small number of cases does not provide enough power. Our results suggest that the bleeding complication might be caused by common variants, in multiple loci but this finding cannot be confirmed without further replication experiments.

6.4.3.1 CYP2C9

Cytochrome 2C9 is the major metaboliser of warfarin. The two variants, *2 and *3, which result in peptides with impaired enzymatic activity and thus poor metabolisers, have been shown to be associated with warfarin dose (Aithal et al. 1999; Higashi et al. 2002; Wadelius et al. 2005; Wadelius et al. 2004), sections 4.2 and 6.3.2). In most instances no genotype test is used prior to prescribing warfarin in clinics, therefore, it is fair to expect that in some prospective studies of warfarin-treated patients an association has to be found between bleeding episode and CYP2C9 genotypes (Limdi et al. 2007; Sanderson et al. 2005).

The *CYP2C9* *2 and *3 alleles were also examined in this study and figure 6.5 shows the result of the 2- and 3-marker haplotype analysis. In the haplotype analysis two further SNPs, rs4244285 and rs4417205, located in *CYP2C19* (Figure 6.5A) were included, since the three CYP2C genes are in good LD. None of the single marker (Appendix IV), nor the specific haplotype, demonstrated nominally significant association for CYP2C9*2 allele (rs1799853, figure 6.5A).

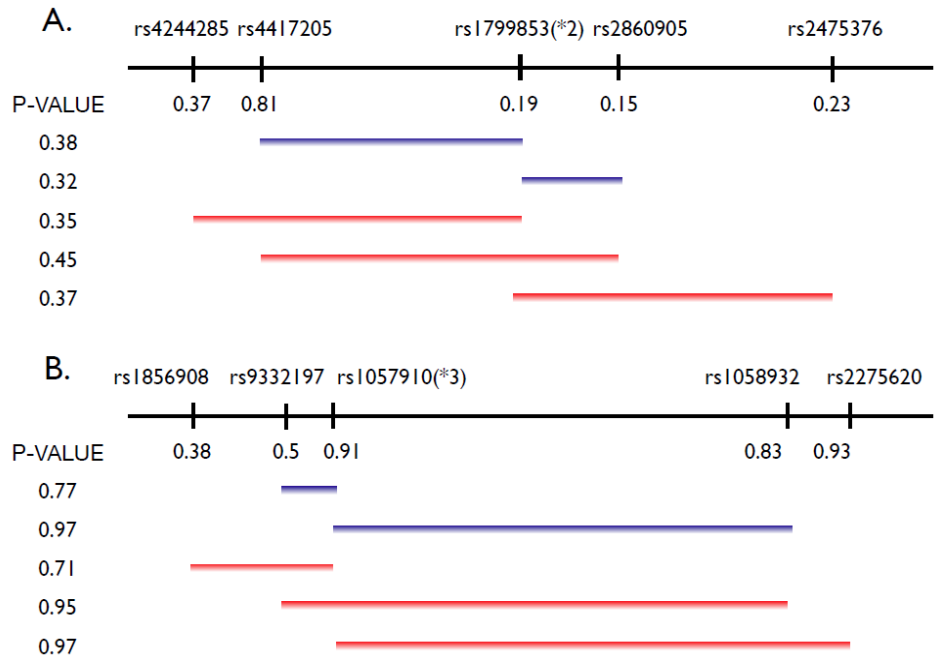


Figure 6.5. Single marker and 2- / 3- marker haplotype analysis of *CYP2C9* (A) *2 and (B) *3 alleles. P-values of single marker trend test are noted below each SNP. The omnibus p-values are noted on the left of each sliding window.

*CYP2C9**3 is the second most important genetic factor affecting warfarin dose that we have found in both Swedish cohorts we have analysed (Chapters IV and V). The two SNPs flanking *CYP2C9**3 were included in 2- and 3-marker haplotype analysis (figure 6.5B). As described earlier, *CYP2C8* is also in weak LD with *CYP2C9* and for the sliding window analysis we included two *CYP2C8* SNPs, rs1058932 and rs2275620. As for the *CYP2C9**2 variant, the p-value for the *3 allele was not significant 0.91 (7.3% in cases and 7.0% in controls).

6.4.3.2 VKORC1

As described in Chapters IV and V, common SNPs in the *VKORC1* locus are significantly associated with warfarin dose (Rieder et al. 2005; Sconce et al. 2005; Wadelius et al. 2005).

Four SNPs, rs9923231 (lead dose SNP), rs2359612, rs7294 (3'-UTR) and rs11150606 (downstream) were genotyped and tested for association with bleeding ADR.

The p-values of the Cochran-Armitage trend test for the four SNPs are shown in Figure 6.6 and are not significant. Evaluation of 2-marker and 3-marker haplotypes gave no significant p-value (Figure 6.6). Our results suggest in a conclusive way that common variants in *VKORC1* are not associated with bleeding, contrary to the strong association we have found between this gene and warfarin dose requirement in both the Uppsala and WARG studies.

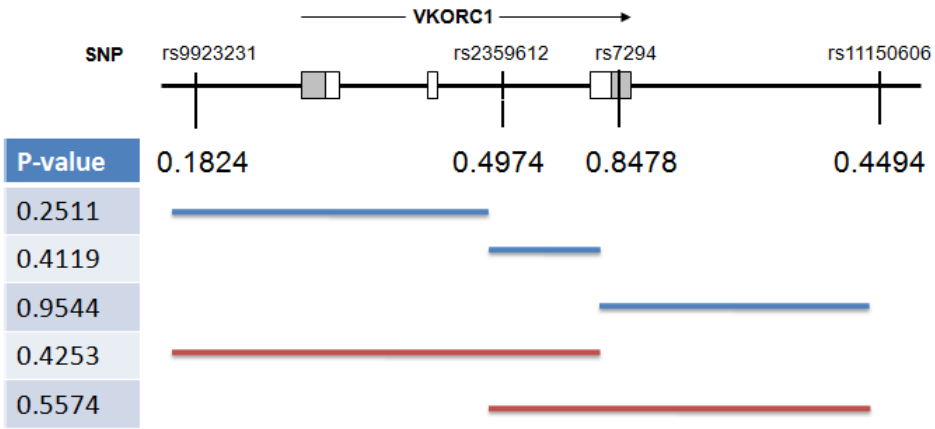


Figure 6.6. Single marker and 2- / 3- marker haplotype analysis of *VKORC1*. P-values from the single marker trend test are noted below each SNP. The omnibus p-values are noted on the left of each sliding window

Reitsma and colleagues (Reitsma et al. 2005) reported the association of bleeding with the C1173T allele in intron 1 of *VKORC1* with patients prescribed phenprocoumon but not acenocoumarol. Though we did not genotype this same SNP, we tested rs9923231 and rs2359612 which are good proxies (r^2 of 1.0 and 1.0 respectively). We failed to replicate the finding by Reitsma et al and this is consistent with the result reported by Limdi and colleagues (Limdi et al. 2007).

6.5 SINGLE STUDY CASE-CONTROL ANALYSIS

In the combined analysis of all bleeders in the Uppsala and WARG studies, we found evidence suggesting that *PDIA4*, *F2* and *P4HB* might be involved in bleeding complications ($P < 0.01$). However, these bleeding patients were recruited in each project with slightly different criteria. In addition, the spectrum of where the bleeding episodes occurred differs between the two cohorts (Table 6.6). Therefore, I assessed the cases of each study independently undertaking the same type of analysis as before.

Table 6.6. Bleeding phenotype of patients in Uppsala and WARG studies.

Type_of_bleeding	Uppsala	WARG
Anaemia		2
Anaemia+gastrointestinal		1
Basal ganglia bleeding	1	
Cerebellar bleeding	2	
Cerebral bleeding	9	3
Gastrointestinal (GI) bleeding	4	8
GI-bleed post-op		2
Haematoma	5	1
Haematuria	1	3
Haemorrhoid bleeding		1
Haemoptysis		1
Intraabdominal bleeding		2
Intraatricular bleeding	2	
Nose bleeding	4	3
Nose bleeding + hematuria	2	
Perirenal bleeding	1	
Subarachnoidal bleeding	1	
Subdural bleeding	4	
Vitreous body bleeding		1
SUM	36	28

Among the 36 bleeding patients in the Uppsala study, 12 were recruited at the Uppsala University hospital as part of the main collection of 201 patients whereas the other 24 patients were collected through different clinics across Sweden: 20 were through the Swedish spontaneous reporting of adverse drug reactions and 4 from an ongoing national study on cerebral bleeding and warfarin. In order to reduce the confounding effect from bias of any subpopulation effect, the comparison was done against 1468 control patients in the WARG project. In other words, both bleeder cohorts were analysed using the same control patient dataset.

6.5.1 Bleeders in the Uppsala study

Table 6.7 lists the strongest association per gene (with $P \leq 0.05$) combining the results of the trend test and sliding window haplotype analyses. The SNP rs1799919 in the protein disulfide isomerise member 1 gene, *P4HB*, is the most associated with P = 8.35E-04 which is lower than the p-value obtained in the combined analysis (P = 6.55E-03) but it still does not pass the significance threshold after correction for multiple tests ($P < 2.62 \times 10^{-4}$).

Table 6.7. Association of bleeders in Uppsala study.

Gene	Analysis	P-value	Marker		
P4HB	Trend	8.35E-04	rs1799919		
PDIA4	3 marker	9.23E-03	rs10269104	rs6464929	rs1551927
NR1I3	Trend	1.03E-02	rs2502804		
F9	2 marker	1.25E-02	rs401597	rs6048	
SERPINC1	Trend	1.59E-02	rs2759328		
EPHX1	Trend	1.98E-02	rs4653436		
F5	Trend	2.11E-02	rs3753305		
ORM1	Trend	2.79E-02	rs1687390		
PDIA5	Trend	3.10E-02	rs1107377		

Nine genes were found to be nominally associated ($P < 0.05$) and *P4HB* SNP rs1799919 has the most significant p-value ($P = 8.35 \times 10^{-4}$) (Table 6.7). Interestingly, none of these nine genes is involved in metabolising warfarin. Instead most of them are involved in pharmacodynamics of warfarin such as recycling vitamin K and in the coagulation cascade; except both *NR1I3* which regulates expression of CYP and ORM1 which transports warfarin to target liver cells.

6.5.2 Bleeders in the WARG study

In the WARG study, which recruited 28 bleeders, the SNP with the lowest p-value in the single marker trend test is rs1686482 in the *PDIA6* gene. Table 6.8 lists all nominally associated SNPs. However, none of these SNPs remained significant after correction for multiple tests. Taking into account the result of 2- and 3-marker sliding window haplotype analyses, a 2-marker window containing rs10272564 and rs10269104 in the *PDIA4* gene gave the most significant p-value of 3.58E-04 which is close to significance after multiple corrections ($P < 2.62 \times 10^{-4}$). As described earlier, candidate genes that fall in to the same LD block were co-analysed for haplotypes. A window with one SNP in coagulation factor X (rs5960) and two SNPs in coagulation protein Z (rs2273971 and rs3024718) was found to be associated with bleeding (P = 8.93E-03). A total of 13 genes indicate marginal association with bleeding episode (Table 6.8).

The observed haplotypes in the most associated 2-marker window of *PDIA4*, rs10272564-rs10269104, are given in Table 6.9. The A-C haplotype is rare (0.93%) in controls but shows a six-fold increase in cases (5.77%). This result suggests the possibility of a causative variant

being associated with this A-C haplotype. However, this finding requires replication in an independent sample.

Table 6.8. Association of bleeders in WARG study.

Gene	Analysis	P-value	Marker		
PDIA4	2 marker	3.58E-04	rs10272564	rs10269104	
CYP2C8	3 marker	4.31E-03	rs947173	rs1536430	rs1058930
PDIA6	Trend	4.99E-03	rs1686482		
F10+PROZ	3 marker	8.93E-03	rs5960	rs2273971	rs3024718
F2	Trend	1.29E-02	rs5898		
PDIA2	Trend	1.75E-02	rs432925		
CYP1A1	Trend	2.23E-02	rs1048943		
ABCB1	Trend	2.50E-02	rs2214102		
SERPINC1	2 marker	2.89E-02	rs5878	rs2227607	
CYP2C9	Trend	2.98E-02	rs1799853		
F10	Trend	3.64E-02	rs3212998		
PDIA3	Trend	3.74E-02	rs11070411		
F7	Trend	4.81E-02	rs6046		

Table 6.9. Association of 2-SNP haplotype (rs10272564-rs10269104) in *PDIA4*.

HAPLOTYPE	Frequency (Case)	Frequency (Control)	CHISQ	P-value
AT	28.85%	18.62%	3.491	6.17E-02
AC	5.77%	0.93%	11.71	6.21E-04
GC	65.38%	80.45%	7.273	7.00E-03
OMNIBUS			15.87	3.58E-04

6.5.3 Summary of findings for separate bleeder groups

By comparing to the same control group, i.e. the non-bleeder patients in the WARG study, the association results obtained for the bleeders in the Uppsala cohort differ from those obtained in the WARG study. The sample number of bleeders in each study is comparable but only three genes appear in common (Table 6.10; the colour in dark blue, blue and light

blue represents p-value of 0.001, 0.01 and 0.05, respectively). Protein disulfide isomerase member 4 and member 2 (*PDIA4* and *PDIA2*), and antithrombin III (*SERPINC1*) indicate possible effects in bleeding complication. *PDIA4* provided the strongest evidence for being a genetic factor influencing serious bleeding.

Table 6.10. Summary of single study analysis.

Study	Uppsala			WARG		
Analysis	Trend	2-marker	3-marker	Trend	2-marker	3-marker
PDIA4	1.9E-02		9.2E-03	9.2E-03	3.6E-04	1.6E-02
SERPINC1	1.6E-02			4.3E-02	2.9E-02	
PDIA2	3.1E-02			1.7E-02	4.0E-02	
CYP2C9				3.0E-02	4.6E-02	4.7E-02
CYP2C8				1.1E-02	1.1E-02	4.3E-03
PDIA6				5.0E-03	1.2E-02	2.1E-02
F10				3.6E-02	3.1E-02	8.9E-03
F2				1.3E-02	1.3E-02	2.0E-02
ABCB1				2.5E-02	3.6E-02	4.5E-02
CYP1A1				2.2E-02		
F7				4.8E-02		
PDIA3				3.7E-02		
P4HB	8.3E-04					
ORM1	2.8E-02					
NR1I3	1.0E-02					
F5	2.1E-02	2.7E-02				
EPHX1	2.0E-02	3.1E-02				
F9		1.2E-02	3.6E-02			

6.5.3.4 Cytochrome P450 2C9

CYP2C9 genotypes have been reported to be associated with warfarin bleeding risk (Limdi et al. 2007; Sanderson et al. 2005). However, this finding is possibly tendentious since patients

were prescribed higher dose of warfarin due to lack of genotypic information of *CYP2C9*, which metabolises the majority of S-warfarin. The combined analysis of association of bleeding risk between *CYP2C9**2 and *3 alleles is not significant; P = 0.189 and P = 0.907, respectively (Table 6.11). Interestingly, in the 28 bleeding patients of the WARG study, the *CYP2C9**2 allele is found rarely (2% in MAF) but six-times more often in the 1468 control warfarin patients, 12% MAF (P=0.030 in trend test; P = 0.026 in Fisher's exact test). This finding is not in agreement with the tendentious association and suggests the possibility that the *2 allele may act protectively against specific bleeding complications in the Swedish subpopulation. This could be explained by negative selection, possibly due to Swedish ancestral exposure to coumarin (or similar compounds). However, this observation is not found in the Uppsala bleeders, which suggests the possibility of sampling bias.

Table 6.11. Result of *CYP2C9* *2 and *3 allele.

SNP	rs1799853 (*2)			rs1057910 (*3)		
	Allele Frequency		P-value	Allele Frequency		P-value
	case	control		case	control	
Combined analysis	7%	12%	0.189	7%	7%	0.907
UPPSALA	14%	12%	0.697	2%	7%	0.221
WARG	2%	12%	0.030	12%	7%	0.219

6.6 THE PROTEIN DISULFIDE ISOMERASE GENE FAMILY

Protein disulfide isomerase has been reported to form a complex with VKORC1, and the thio-redoxin like CXXC centre is where warfarin targets to inhibit blood coagulation (Wajih et al. 2007). Wajih and colleagues reported the evidence of this biochemical interaction, but did not specify which member of PDI multi-gene family is involved (Wajih et al. 2007). Since PDI is the most abundant protein in ER lumen, different members of PDI might work cooperatively to provide electrons to VKOR.

Besides *PDIA4*, SNPs in *P4HB* and *PDIA6* are significantly associated ($P < 0.01$) with bleeding complication in the Uppsala and WARG studies and this result raises a plausible hypothesis that different PDIs might be tissue-specific and impair the initiation of blood coagulation locally. Recent studies demonstrated that PDI works as a switch of the tissue factor pathway and enhances the activation of factor VIIa-dependent substrate factor X by 5-10 fold in the presence of wild-type, oxidised soluble tissue factor (Ahamed et al. 2006; Versteeg and Ruf 2007). Thereby, changes in messenger RNA level or non-synonymous amino acid of PDI may influence the efficiency of coagulation and lead to a bleeding episode. Figure 6.7 shows the *in silico* analysis of mRNA expression using the Unigene database (http:// http://www.ncbi.nlm.nih.gov/UniGene) at NCBI. P4HB, PDIA4 and PDIA6 seem to be abundantly expressed in a variety of tissues. However, their expression profiles are still slight different.

	P4HB		PDIA4		PDIA6	
adrenal gland	1920	64 /33321	60	2 /33321	420	14 /33321
ascites	823	33 /40066	49	2 /40066	499	20 /40066
bladder	1460	44 /30133	0	0 /30133	331	10 /30133
blood	572	71 /124125	96	12 /124125	161	20 /124125
bone	515	37 /71794	97	7 /71794	292	21 /71794
bone marrow	813	40 /49157	203	10 /49157	223	11 /49157
brain	737	814 /1104170	44	49 /1104170	170	188 /1104170
cervix	597	29 /48501	82	4 /48501	288	14 /48501
connective tissue	1497	224 /149629	66	10 /149629	274	41 /149629
ear	61	1 /16341	122	2 /16341	61	1 /16341
embryonic tissue	468	101 /215807	185	40 /215807	379	82 /215807
esophagus	1823	37 /20293	295	6 /20293	640	13 /20293
eye	389	82 /210743	99	21 /210743	204	43 /210743
heart	575	52 /90306	55	5 /90306	188	17 /90306
intestine	947	223 /235317	293	69 /235317	327	77 /235317
kidney	1547	329 /212570	141	30 /212570	211	45 /212570
larynx	163	4 /24439	163	4 /24439	40	1 /24439
liver	945	197 /208304	110	23 /208304	518	108 /208304
lung	680	230 /338054	192	65 /338054	221	75 /338054
lymph	563	25 /44400	45	2 /44400	112	5 /44400
lymph node	97	9 /91861	228	21 /91861	87	8 /91861
mammary gland	1211	187 /154293	246	38 /154293	252	39 /154293
muscle	397	43 /108151	83	9 /108151	83	9 /108151
nerve	697	11 /15765	0	0 /15765	253	4 /15765
ovary	526	54 /102648	224	23 /102648	311	32 /102648
pancreas	720	155 /215273	106	23 /215273	255	55 /215273
parathyroid	48	1 /20634	48	1 /20634	1017	21 /20634
pharynx	192	8 /41489	24	1 /41489	265	11 /41489
pituitary gland	298	5 /16736	239	4 /16736	239	4 /16736
placenta	898	255 /283915	56	16 /283915	380	108 /283915
prostate	1262	241 /190869	193	37 /190869	220	42 /190869
salivary gland	1576	32 /20295	0	0 /20295	98	2 /20295
skin	1768	373 /210896	142	30 /210896	289	61 /210896
soft tissue	456	6 /13146	152	2 /13146	76	1 /13146
spleen	3200	173 /54062	18	1 /54062	92	5 /54062
stomach	1362	132 /96900	319	31 /96900	206	20 /96900
testis	377	125 /331293	84	28 /331293	229	76 /331293
thymus	799	65 /81256	24	2 /81256	61	5 /81256
thyroid	1522	73 /47940	333	16 /47940	417	20 /47940
tongue	1924	127 /65981	75	5 /65981	30	2 /65981
tonsil	411	7 /17031	0	0 /17031	117	2 /17031
trachea	2632	138 /52430	0	0 /52430	152	8 /52430
umbilical cord	1162	16 /13765	72	1 /13765	435	6 /13765
uterus	778	182 /233902	141	33 /233902	299	70 /233902
vascular	3909	203 /51930	154	8 /51930	462	24 /51930

Figure 6.7. *In silico* analysis of mRNA expression using the Unigene database. Left column indicates the tissue of expression. The number on the left of the black spot is the transcripts per million (TPM) reported in the database whilst the intensity of black spot is visually represented based on TPM. The number next to the spot is the number of gene-expressed sequence tag (EST)/Total EST in pool.

6.6.1 Three SNPs in trend and Fisher's exact tests on P4HB, PDIA4 and PDIA6

In the combined analysis of the single marker association, *PDIA4* (P = 2.72E-03) appears to be the most significantly associated gene with bleeding ADR. Meanwhile, in single marker association analyses *P4HB* shows the most significant association in the Uppsala study, whilst the *PDIA6* gave the lowest p-value in WARG study. Table 6.12 lists the most significant associated SNPs in each of the three PDI genes and the result of analysis from both cohorts in combined and as independent cohorts.

Table 6.12. The most significant SNPs in *PDIA4*, *P4HB* and *PDIA6*.

Gene	SNP	Allele	MAF		P-value	Odds ratio
		minor|major	All case	All control		
PDIA4	rs4727005	T|C	0.3208	0.2034	0.005	1.849
P4HB	rs1799919	C|T	0.3372	0.2148	0.011	1.859
PDIA6	rs1686482	C|A	0.3723	0.5035	0.015	0.5849

A total of 27 possible combinations of genotypes of these three SNPs were grouped for case and control patients. The proportion of each combination was calculated for the Uppsala study (blue), WARG study (red), and the combined control patients from the two studies (in purple) (Figure 6.8). The patients from the Uppsala and WARG studies demonstrated very different distributions (correlation coefficient: 0.46), and this can be explained with different bleeding phenotypes in both bleeding cohorts.

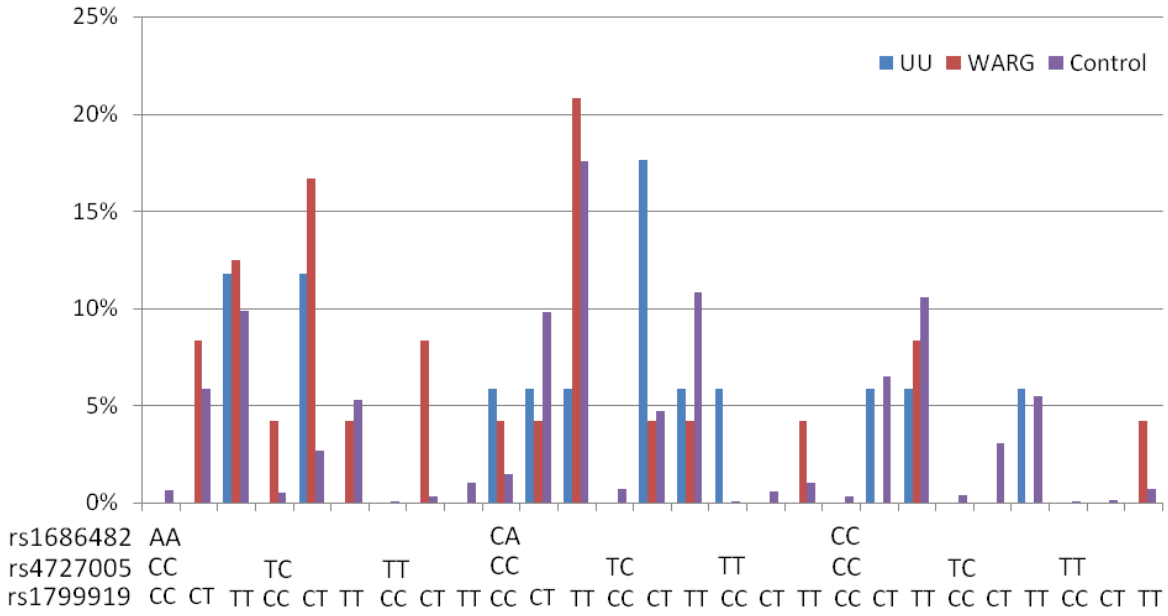


Figure 6.8. All 27 combinations of the top associated SNPs in *P4HB* (rs1799919), *PDIA4* (rs4727005) and *PDIA6* (rs1686482). Colour of the columns represents the proportion of genotype combination in different cohorts of Uppsala bleeder (blue), WARG bleeder (red) and all controls (Uppsala and WARG).

Although the number of cases is small, the frequency of each genotype combination is comparable between all cases and controls, with a correlation coefficient of 0.84 excluding genotype combinations AA-TC-CT, AA-TT-CT, and CA-CC-CC (rs1686482-rs4727005-rs1799919; Figure 6.8). The three excluded combinations have five-time enrichment in cases over controls (Table 6.13). Among 1386 Swedish control patients, only 4 individuals have the AA-TT-CT genotype combination alongside 2 case patients from the WARG study. Inspection of phenotypic information showed that both AA-TT-CT bleeding patients had gastrointestinal bleeding episodes. Overall frequencies for AA-TC-CT, AA-TT-CT, and CA-CC-CC in controls are 2.7%, 0.3%, and 1.4%, respectively (Table 6.13). All 27 genotype combinations were observed in the control patients, but only 14 and 11 of them were found in the WARG and Uppsala cases, respectively. The latter observation may well be due to the small sample size of case patients in the Uppsala cohort. In the control patients, 22 out of the 27 combinations have a frequency lower than 10%. In controls, the AA-TT-CC, CA-TT-CC,

and CC-TT-CC combinations were particularly rare with only one individual bearing each one. Although rs4727005 and rs1799919 are common in the control patients (MAF = 0.2034 and 0.2148, respectively), the TT-CC combination is very rare which implies the possibility of a gene-gene interaction between *PDIA4* and *P4HB*. Furthermore, the CC-TT combination for these two SNPs has a very high frequency (39%), which is also in support of the above hypothesis.

Table 6.13. Population frequency of genotypic combination of the three SNPs in *PDIA6* (rs1686482), *PDIA4* (rs4727005) and *P4HB* (rs1799919).

Genotype			Cases						All	
rs1686482	rs4727005	rs1799919	Uppsala	%	WARG	%	All	%	controls	%
		CC	0	0.0%	0	0.0%	0	0.0%	9	0.6%
	CC	CT	0	0.0%	2	8.3%	2	4.9%	81	5.8%
		TT	2	11.8%	3	12.5%	5	12.2%	137	9.9%
		CC	0	0.0%	1	4.2%	1	2.4%	7	0.5%
AA	TC	CT	2	11.8%	4	16.7%	6	14.6%	37	2.7%
		TT	0	0.0%	1	4.2%	1	2.4%	73	5.3%
		CC	0	0.0%	0	0.0%	0	0.0%	1	0.1%
	TT	CT	0	0.0%	2	8.3%	2	4.9%	4	0.3%
		TT	0	0.0%	0	0.0%	0	0.0%	14	1.0%
		CC	1	5.9%	1	4.2%	2	4.9%	20	1.4%
	CC	CT	1	5.9%	1	4.2%	2	4.9%	136	9.8%
		TT	1	5.9%	5	20.8%	6	14.6%	243	17.5%
		CC	0	0.0%	0	0.0%	0	0.0%	10	0.7%
CA	TC	CT	3	17.6%	1	4.2%	4	9.8%	65	4.7%
		TT	1	5.9%	1	4.2%	2	4.9%	150	10.8%
		CC	1	5.9%	0	0.0%	1	2.4%	1	0.1%
	TT	CT	0	0.0%	0	0.0%	0	0.0%	8	0.6%
		TT	0	0.0%	1	4.2%	1	2.4%	14	1.0%
		CC	0	0.0%	0	0.0%	0	0.0%	4	0.3%
	CC	CT	1	5.9%	0	0.0%	1	2.4%	90	6.5%
		TT	1	5.9%	2	8.3%	3	7.3%	146	10.5%
		CC	0	0.0%	0	0.0%	0	0.0%	5	0.4%
CC	TC	CT	0	0.0%	0	0.0%	0	0.0%	42	3.0%
		TT	1	5.9%	0	0.0%	1	2.4%	76	5.5%
		CC	0	0.0%	0	0.0%	0	0.0%	1	0.1%
	TT	CT	0	0.0%	0	0.0%	0	0.0%	2	0.1%
		TT	0	0.0%	1	4.2%	1	2.4%	10	0.7%

6.6.2 The gastrointestinal bleeding sub-phenotype

As described above, three genotype combinations for rs1686482 (*PDIA6*)-rs4727005 (*PDIA4*)-rs1799919 (*P4HB*) namely AA-TC-CT, AA-TT-CT, and CA-CC-CC, are of particular interest since they appeared to be predominantly present in cases (Table 6.13). To investigate this finding further, I assessed the bleeding phenotype of each patient recruited in the WARG study relative to the genotype combination for the three SNPs. Comparing to the prospective naturalistic clinical design of the WARG study, bleeders recruited in Uppsala study were excluded because various inclusion criteria were applied for the 12 patients recruited in the Uppsala University Hospital, 20 were through the Swedish spontaneous reporting of ADRs and 4 from a study on cerebral bleeding and warfarin.

Among the 28 bleeding patients in the WARG study, the main sub-phenotype is gastrointestinal (GI) bleeding (46.4%, 13 patients, including 2 intra-abdominal bleeders). Two cases (with haematoma and haematuria bleeding) were removed from further analysis due to incomplete genotype information and Table 6.14 summarises the results of the 26 WARG cases which are sorted by genotype combination and bleeding sub-phenotype (colour coded). Three genotype combinations stand out as being enriched for GI bleeders AA-CC-TT (3 cases), AA-TT-CT (2 cases), and CC-CC-TT (2 cases) (Table 6.14). Interestingly, none of the other bleeding sub-phenotypes was found in the three GI-enriched genotypic combinations.

This questions if any SNP is associated with a particular bleeding sub-phenotype. Although the biochemical mechanism of bleeding complication is not yet understood, identification of

risk factors for subdividing serious bleeding could contribute to a bleeding risk model (Shireman et al. 2006) for preventing serious ADRs.

Table 6.14. Bleeding sub-phenotype stratification with genotype combinations in WARG bleeders.

WARG	Genotype			Observed bleeding
Patient ID	rs1686482	rs4727005	rs1799919	
634	A A	C C	C T	Intraabdominal bleeding
1392	A A	C C	C T	Nose bleeding
237	A A	C C	T T	GI-bleeding
362	A A	C C	T T	GI-bleeding
462	A A	C C	T T	GI-bleeding
792	A A	T C	C C	Intraabdominal bleeding
202	A A	T C	C T	GI-bleeding, post-op
219	A A	T C	C T	Hematuria
604	A A	T C	C T	Anaemia (1), GI-bleed (2)
610	A A	T C	C T	Intracerebral bleeding
903	A A	T C	T T	Intracerebral bleeding
339	A A	T T	C T	GI-bleeding
548	A A	T T	C T	GI-bleeding
666	C A	C C	C C	Nose bleeding
108	C A	C C	C T	Anaemia
80	C A	C C	T T	Haemoptysis
110	C A	C C	T T	Hematuria
342	C A	C C	T T	Anaemia
636	C A	C C	T T	Haemorrhoid bleeding
1447	C A	C C	T T	GI-bleeding
214	C A	T C	C T	GI-bleed post-op
407	C A	T C	T T	Nose bleeding
833	C A	T T	T T	Intracerebral bleeding
184	C C	C C	T T	GI-bleeding
253	C C	C C	T T	GI-bleeding
1402	C C	T T	T T	Vitreous body bleeding

6.6.3 Predictors of gastrointestinal bleeding

To examine the gastrointestinal bleeding sub-phenotype stratification, I analysed the patients in the WARG study employing Fisher's exact test. The SNPs showing nominally significant association are listed in Table 6.15 together with their MAF, p-value, and odds ratio. Six SNPs in *PDIA6* top the list, but, not surprisingly, none of these SNPs passed the Bonferroni Meff correction threshold for multiple tests. Besides *PDIA6*, SNPs in *CYP2C8*, *PDIA4* and *CYP2C9* are nominally associated. Notably, the minor allele of *CYP2C8* SNP rs1536430 appears only 2.2% in control patients whereas a substantial increase of 12.5% is found in cases. The MAF in WARG controls is also reflected in the Hapmap CEU panel (1.7%) and the Perlegen EUR panel (2.1%).

Table 6.15. Association of gastrointestinal bleeding in the WARG cohort.

Gene	SNP	Allele	MAF		P-value	OR
		minor|major	Case	Control		
PDIA6	rs1198873	T|C	0.7083	0.3901	2.47E-03	3.797
PDIA6	rs1686482	A|C	0.7917	0.4911	3.56E-03	3.938
PDIA6	rs11904084	T|C	0.6667	0.3856	6.07E-03	3.187
PDIA6	rs1734343	G|T	0.2083	0.4988	6.41E-03	0.2644
PDIA6	rs1734346	G|A	0.6667	0.3914	1.01E-02	3.109
PDIA6	rs1686447	A|G	0.6667	0.393	1.02E-02	3.089
CYP2C8	rs1536430	T|C	0.125	0.02249	1.74E-02	6.208
PDIA4	rs10272564	A|G	0.375	0.1995	4.11E-02	2.407
CYP2C9	rs2860905	A|G	0.04167	0.2175	4.26E-02	0.1565

Genotypes of rs1198873 were extracted from all WARG patients for further analyses. The ratio of the CC / TC / TT genotypes in control patients is 39%, 45%, and 16%, respectively. Interestingly, the TT genotype rises to a ratio of 54% among the gastrointestinal bleeding

patients. The CC and TC genotypes have a similar ratio in patients (1 : 1) and in controls (1 : 1.17) which suggests no sampling bias (Figure 6.9).

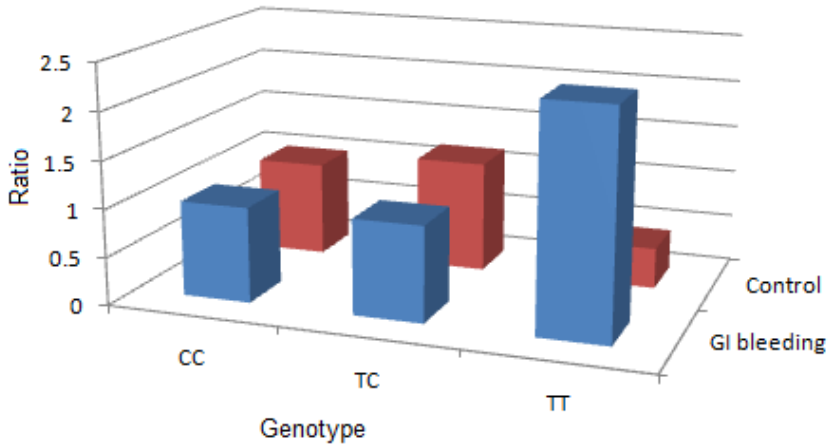


Figure 6.9. Ratio of rs1198873 genotype in the gastrointestinal (GI) bleeding and control patients in the WARG study. The ratio of TC and TT genotypes were presented accordingly to CC genotype in both groups. A significant enrichment of TT genotyped is found in GI bleeders.

A two by two table was generated using rs1198873 as the predictor for GI bleeding (Table 6.16) and recessive mode of inheritance. Seven of the 13 cases could be successfully predicted as positive responders whilst 1191 control could be predicted as negative. The correlative sensitivity and specificity is 53.8% and 83.6%, respectively, with P = 0.0023 and 5.95 in odds ratio (95% confidence interval: 1.70-21.65).

Table 6.16. Prediction of gastrointestinal bleeding using rs1198873 in the WARG study.

		WARG	
		Case	Control
Prediction	Positive	7	233
	Negative	6	1191

Severe bleeding complication was recorded in circa 2% of the patients enrolled in the WARG study. Following the analysis with Fisher's exact test, I also performed an association test using a recessive model of inheritance. All variants showing nominally significant association are listed in Table 6.17. Except the variants in *PDIA6*, SNPs in *PDIA4, CYP2C8*, and *F5* are nominally significant ($P < 0.05$), and the recessive alleles are rare in cases (25%, 20%, and 25%, respectively). Due to the small number of cases, corrected p-values for each SNP were calculated with Fisher's exact test using the R statistical package (Table 6.17).

Table 6.17. Association of gastrointestinal bleeding based on recessive action assumption.

Gene	SNP	TEST	CASE*	CONTROL*	P-value	P (Fisher)	OR	95% CI	
PDIA6	rs1198873	REC	7_5	220_1099	1.3E-04	**1.41E-03**	6.98	1.89	28.13
PDIA6	rs11904084	REC	7_5	220_1082	1.6E-04	**1.52E-03**	6.87	1.86	27.71
PDIA4	rs10272564	REC	3_9	50_1258	2.0E-04	**1.04E-02**	8.35	1.41	34.85
PDIA6	rs1686447	REC	7_5	223_1048	2.5E-04	**1.91E-03**	6.57	1.78	26.49
PDIA6	rs1734346	REC	7_5	226_1059	2.5E-04	**1.94E-03**	6.55	1.77	26.40
CYP2C8	rs1557044	REC	2_10	26_1329	3.3E-04	**2.35E-02**	10.17	1.03	51.35
F5	rs3766110	REC	3_9	71_1264	2.9E-03	**2.45E-02**	5.92	1.01	24.39

*Numbers present the patients homozygous for minor allele and a sum of homozygotes of major allele and heterozygotes.

Taking into account the single marker result of Fisher's exact test and genotypic test in recessive mode, *PDIA6* is likely to play a role in causing, or as a consequence of, gastrointestinal bleeding. The *in silico* expression analysis suggests a higher expression in ascites, esophagus, intestine and stomach (Spot density and TPM in Figure 6.7).

6.6.4 Predictors of non-gastrointestinal bleeding

Although the *PDIA6* SNP rs1198873 is significantly associated with gastrointestinal-related bleeding complication, it is of interest to identify variants explaining the remaining sub-

phenotypes of bleeding complications. Therefore, I examined the 15 non-gastrointestinal bleeders for possible associations to genetic determinants.

Table 6.18 summarises the results of Fisher's exact test for non-gastrointestinal bleeding. SNPs in four genes are nominally significant ($P < 0.05$), and the lead SNP of each gene is common (MAF ≥ 5%), except for rs1048943 in *CYP1A1* (MAF = 2.5%). The Hapmap CEU panel also has a MAF of 4.2% for rs1048943, whereas the WARG bleeders have a MAF of 11.5 %. This SNP is a good candidate for replication studies in other cohorts.

Table 6.18. Fisher's exact test on non-gastrointestinal bleeding in the WARG study.

Gene	SNP	Allele	MAF		P-value	OR	95% CI	
		minor|major	Case	Control				
PDIA3	rs11070411	G|C	32.1%	16.2%	3.60E-02	2.45	1.19	5.23
CYP1A1	rs1048943	C|T	11.5%	2.5%	2.90E-02	5.03	1.42	62.66
ORM1	rs2787337	T|C	11.5%	31.4%	3.21E-02	0.29	0.13	0.67
ABCB1	rs2214102	T|C	21.4%	9.7%	4.98E-02	2.54	1.00	6.03

Recessive genetic association for non-gastrointestinal bleeding was then tested and the p-value was corrected with Fisher's exact test. The SNPs in *PROS1*, *NR1I2*, *F2*, and *PDIA4* were nominally significant (Table 6.19). However, except rs2472677 in *NR1I2*, the number of recessive genotypes in cases is smaller than 5 which might result in an incorrect p-value. Fisher's exact test was applied for corrected p-value, odds ratio, and 95% confidence interval. Five SNPs in *PROS1*, *F2*, and *PDIA4* did not pass the threshold (Table 6.19) and three SNPs in *PROS1*, *NR1I2* and *F2* are nominally significant ($P < 0.05$).

Table 6.19. Association of non-gastrointestinal bleeding based on recessive action assumption.

Gene	SNP	TEST	CASE*	CONTROL*	P-value	P (Fisher)	OR	95% CI
PROS1	rs5013930	REC	2_12	27_1275	2.0E-03	3.62E-02	7.84	0.81 37.91
PROS1	rs4857343	REC	2_10	42_1308	8.2E-03	5.48E-02		
NR1I2	rs2472677	REC	6_8	212_1051	9.9E-03	2.07E-02	3.71	1.05 12.35
F2	rs2070851	REC	3_11	75_1256	1.2E-02	4.30E-02	4.56	0.80 17.75
PDIA4	rs10085877	REC	2_12	43_1276	2.3E-02	7.84E-02		
PDIA4	rs4727005	REC	2_12	49_1272	4.0E-02	9.72E-02		
PDIA4	rs6464929	REC	2_12	50_1264	4.4E-02	1.01E-01		
PDIA4	rs10272564	REC	2_12	50_1258	4.5E-02	1.02E-01		

*Numbers present the patients homozygous for minor allele and a sum of homozygotes of major allele and heterozygotes.

6.7 RE-SEQUENCING

Genotyping of common variants with MAF greater than 5% is a powerful approach in searching for genetic determinants underlying complex traits, but the power to detect moderate effects requires large sample sizes (Wang et al. 2005). Haplotype analysis sometimes gives additional information when the causative variants are carried in one of the haplotypes. However, the above approaches based on tag SNPs, are less well suited for studying variants with frequencies of 1-5% which are often functional e.g. non-synonymous changes that alter amino acid composition. In contrast, re-sequencing exonic sequences in patients and controls allows mining of the full spectrum of common variation (with the exception of larger size structural variants) which can then be further interrogated in larger samples.

6.7.1 Candidate genes for bleeding

Patients who were subject to warfarin induced bleeding complication did not have any bleeding episode prior to warfarin treatment. If a patient had bleeding history, the clinician would have been cautious with the use of anticoagulants and provide more frequent PT INR monitoring. In other words, the patients who had bleeding episodes recorded should have had an apparently normal blood clotting system in the absence of warfarin treatment. Therefore, the genes that directly interact, or are affected by warfarin, may play a role in warfarin-induced bleeding. This hypothesis brought our attention to the genes that are involved in recycling vitamin K including, *CALU, EPHX1, GGCX*, *NOQ1*, *PDIs* (*P4HB* and *PDIA4*), and *VKORC1*. The initial results from analysing the 36 bleeders obtained through Uppsala University and literature searches suggested that *APOE, CYP2C9, F*9, and *PROC* might be of

interest. Among the 35 candidate genes we selected as relevant in studying warfarin genetics, the exons of 11 genes (listed above and shown in Figure 6.10) were sequenced in a panel of 36 bleeders and 12 non-bleeders from the Uppsala study. Note that at the start of this investigation we had no access to the WARG bleeders. Not only the exons from 48 Swedish patients were sequenced but these exons were also sequenced on 48 CEPH Caucasians.

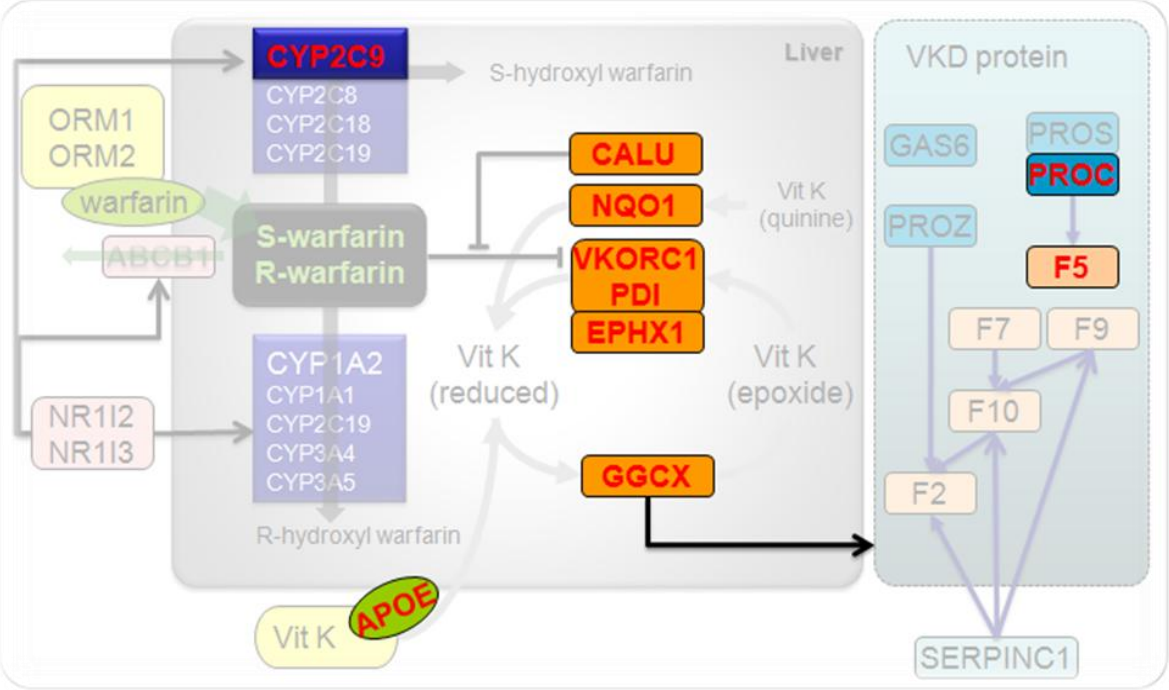


Figure 6.10. Candidate genes for exon sequencing. Among the 35 candidate genes, used for searching for the genetic determinants for associations with warfarin dose and bleeding complications using genotyping technology, a subset of 11 genes were selected to discover common and rare variants using sequencing technology.

6.7.2 Sequencing results

A total of 108 exons in the 11 candidate genes, spanning 28,951 base pairs of sequence were selected. We designed 167 PCR amplicons and PCR products of the 96 individuals were sequenced through the ExoSeq (exon re-sequencing) pipeline at the Sanger Institute (Table

6.20). These exons were sequenced on both the forward and reverse strand and we identified 161 SNPs in 48 Swedish patients and 166 SNPs in 48 CEPH Caucasians for the 11 genes (Table 6.20).

Table 6.20. Sequencing summary of 11 candidate genes.

Genes	Exons	Transcript (bp)	Amplicons	SNPs identified	
				Patients	CEPH
APOE	4	1179	6	3	2
CALU	7	3316	17	13	18
CYP2C9	9	1847	16	18	16
EPHX1	9	1605	11	20	17
F5	25	6914	42	41	40
GGCX	15	3315	19	13	14
NQO1	6	2448	12	11	8
P4HB	11	2580	16	18	24
PDIA4	10	2903	16	12	17
PROC	9	1847	8	10	8
VKORC1	3	997	4	2	2
Subtotal	108	28951	167	161	166

Among the 161 SNPs identified in 48 Swedish patients, 56 SNPs are novel and have not been previously reported in public databases whereas 105 SNPs were already known. Although we targeted the exonic sequence, flanking region of at least 125 bp on both ends of each exon were included. Due to the alleviation in selection pressure, introns accumulate polymorphisms and 86 (53.4%) out of 161 SNPs were intronic (Table 6.21). We identified 26 synonymous coding changes (16.1%) and 20 non-synonymous coding changes which alter the peptide composition (12.4%). This observation agrees with the evolutional nature that polymorphisms altering peptide composition are subject to selection pressure. The study design, permitted identification of a small number of SNPs in putative regulatory regions, upstream and downstream regions, and 5'- or 3' untranslated regions (UTR) (Table 6.21). In

the set of Swedish patients, the exonic SNPs, including UTR SNPs, appear on one of every 432 bp.

Table 6.21. Consequence of SNPs identified in warfarin-treated patients.

Consequence	SNP	
	novel	known
UPSTREAM	1	0
SYNONYMOUS CODING	4	22
REGULATORY REGION	1	0
NON SYNONYMOUS CODING	3	17
INTRONIC	38	48
DOWNSTREAM	1	5
5PRIME UTR	2	3
3PRIME UTR	6	10
Subtotal	56	105

Among the 56 novel SNPs found in patients, 20 SNPs are also found in CEPH Caucasians, suggesting that the remaining 36 novel SNPs may be Swedish specific. The MAF of these 36 novel SNPs is shown with blue columns in Figure 6.11. The red columns in Figure 6.11 indicate the 41 novel SNPs but these SNPs were not observed in patients. Most of these 'population-specific' novel SNPs are rare variants, i.e. MAF = 1%, 22 SNPs are patient specific whilst 29 SNPs are found in CEPH only.

Additionally, 125 SNPs were found in both Swedish patients and CEPH panel, and the frequency difference between the two panels is shown in Figure 6.12. Twenty SNPs have no difference in MAF and 70 SNPs have MAF deviation lower than 5%. Only 5 SNPs show a difference in allele frequency above 10%. The correlation of the allele frequency between the two populations is 0.88. In fact, this is lower than a previous estimation of 0.97 with the

SNPs genotyped in Swedish and Hapmap CEU panels. One possibility is the inclusion of the 36 bleeders from the Uppsala study and the other one that rare SNPs show increased fluctuation in MAF.

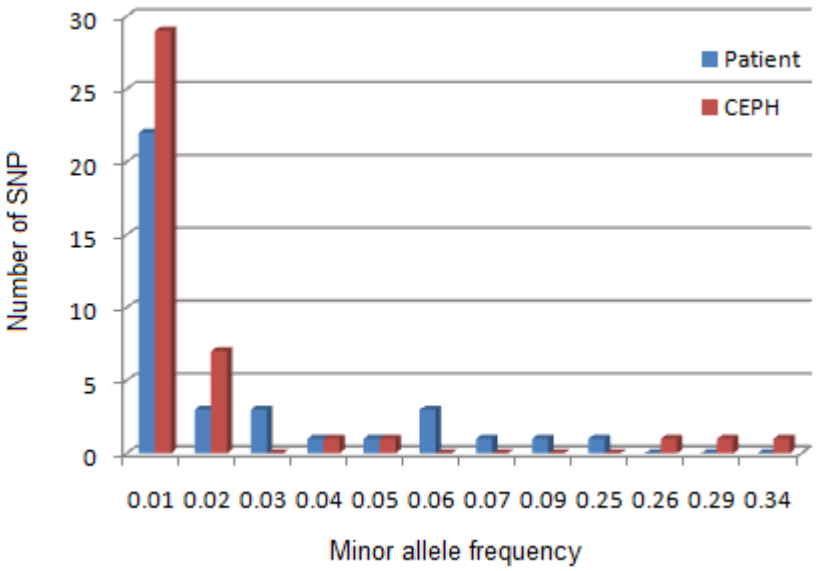


Figure 6.11. Minor allele frequency of novel SNPs found specific in Swedish (blue) and in CEPH (red).

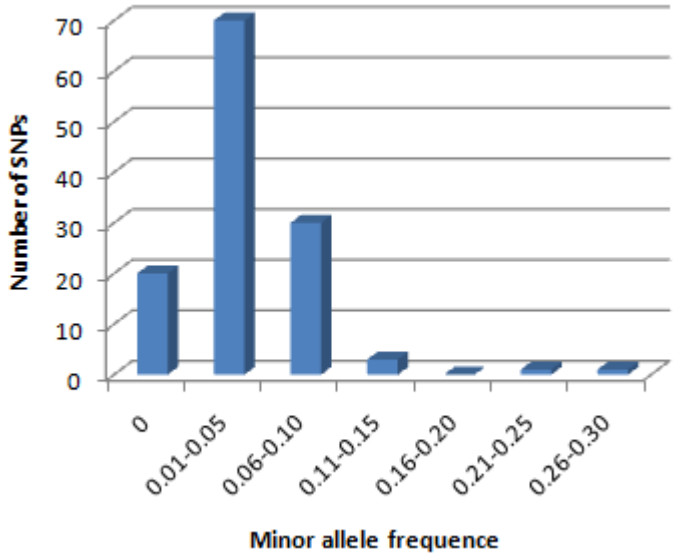


Figure 6.12. Allele frequency differences of SNPs found in both Swedish patients and CEPH Caucasians.

No particular variant was found in the selected bleeders; showing significant enrichment against the CEPH Caucasians. SNPs found to be present only in the Swedish patients will need to be validated in a larger Swedish cohort consisting of both cases (bleeders) and warfarin treated control patients (non-bleeders). Association tests were also performed using CEPH individuals as controls, but no significant association is found due to the small sample size of cases (36 bleeders) and controls (48 CEPH Caucasians and 12 warfarin treated control patients).

6.8 SUMMARY

Currently no genetic determinant is included in contemporary risk assess model for warfarin bleeding (Kakar et al. 2006) although some studies have reported *VKORC1* SNP C1173T (Reitsma et al. 2005) and *CYP2C9**2 and *3 alleles (Limdi et al. 2007; Sanderson et al. 2005). The results of this research are in agreement with Limdi and colleagues; that *VKORC1* genotype has no influence on warfarin related bleeding complication. *CYP2C9**2 and *3 alleles result in poor metabolisers which may cause bleeding due to over-anticoagulation.

In this study I reported associations found in *PDIA4*, *P4HB* and *PDIA6* in the combined and single study analyses. Meanwhile, *PDIA6* is potentially associated with an increased risk of gastrointestinal bleeding. These results support the newly identified function of PDI peptides in initiating the tissue factor pathway of coagulation. A comprehensive identification of sequence variants on the PDI gene family and biological characterisation of different PDI members will further elucidate the mechanism of warfarin induced bleeding complication. With a total of 64 bleeding patients from the Uppsala and WARG studies, the identified associations are with no doubt the result of an under powered study and may well be false positives. Replication in other cohorts or a meta-analysis including different cohorts will help to clarify the finding in the PDI genes which looks cautiously promising.

CHAPTER VII

SUMMARY AND DISCUSSION

7.1 THIS THESIS

When this project commenced, comprehensive whole genome genotyping chips were not yet available, the HapMap project was just scaling up, and genotyping was much more expensive than it is now thus, the candidate gene approach was therefore a realistic route to search for the genetic determinants which influence warfarin dose and the cause of ADR of bleeding. Throughout the project, a total of 35 candidate genes were selected based on pharmacokinetics and pharmacodynamics of warfarin reported in the literature (Chapter III). In the initial phase of SNP selection, only a small number of SNPs in public databases had allele frequency information attached to them. LD maps were constructed for each gene or locus (e.g. CYP2C genes on chromosome 10) using a sample of 201 warfarin treated patients enrolled by the University of Uppsala in Sweden. The maps harbour 379 common SNPs (MAF≥5%) and capture most of the common variation in HapMap ($r^2 \geq 0.8$). SNP genotyping was carried out using the MassExtend and iPLEX assays from Sequenom.

Our group's collaboration with Drs. Mia Wadelius (Uppsala) and Anders Rane (Karolinska) gave us access to two study samples:

- 201 warfarin treated patients (including 12 bleeding patients) plus 24 additional patients with bleeding complication (Uppsala study)

- 1496 warfarin treated patients including 28 bleeders enrolled in the prospective WARG study in Sweden.

In both studies patient information was available on age, gender, indication, concomitant medication, and corresponding dose and PT INR values in each visit to clinics.

We first looked for genetic determinants of warfarin dose requirements (Chapters IV and V). After two rounds of discovery / replication in the Uppsala and WARG sample collections, only two of the 35 candidate genes interrogated showed irrefutable evidence for a strong effect on inter-individual dose variation. Both these genes namely *VKORC1* and *CYP2C9* encode peptides that directly interact with warfarin. Others and we have demonstrated that the *VKORC1* genotype and *CYP2C9* *2 / *3 alleles have predictive value justifying their usefulness as a genotyping test for patients who are prescribed warfarin. Based on the evidence we provided of *VKORC1* rs9923231 and *CYP2C9*2 and *3 association with (1) stable dose, (2) time to reach stable dose, and (3) risk of over-anticoagulation, their use in a clinical setting will narrow the time window for reaching an optimal warfarin dose and reduce hospitalisation. The genetic impact of the two genes is detailed in the respective chapters (chapters IV and V).

In the Uppsala study we also observed associations with warfarin dose for *APOE*, *EPHX1*,

CALU, *PDIA2*, *GGCX*, *ORM1-2* and *PROC*; the latter almost reached experiment-wise significance. However, upon replication in the prospective WARG study of 1496 patients also of Swedish origin, none of these genes reached significance after correction for multiple tests (see Table 5.5). Minor effects from *GGCX*, *EPHX1*, *APOE* and *ORM1-2* have shown a replication trend in both this and other studies (Kohnke et al. 2005; Rieder et al. 2007; Sconce et al. 2006; Wadelius et al. 2007). Our data suggest that these genes do not have a strong effect on dose and are potentially population/treatment specific.

Based on genetic factors determined through the study and in combination with non-genetic factors such age, gender and drug interaction, nearly 60% in dose variation could be explained. However the remaining 40% still needs to be deciphered. It might be due to both genetic factors in other untested genes and environmental factors such as diet for which our studies had not recorded sufficient information. Some relevant non-genetic factors could be recorded using proteomic technologies such as measuring peptide and chemical compositions in the serum and urine.

The ultimate goal of this project is to develop a warfarin-dosing algorithm. It is relevant to mention the work of Niclas Eriksson (Uppsala), which has led to the development of such a tool based on our knowledge. The calculated coefficients of each predictor in his dosing

algorithm are listed in Table 7.1. The number calculated according to this table needs to be squared, which is the predicted weekly maintenance dose.

Table 7.1. Predictors for warfarin dosing algorithm.

Coefficients		Estimate	Standard error	P value
Starting		9.46832	0.11867	$<2x10^{-16}$
CYP2C9	*1/*2	-0.50836	0.05811	$<2x10^{-16}$
	*1/*3	-0.97546	0.07077	$<2x10^{-16}$
	*2/*2	-1.10204	0.19767	$3.0x10^{-8}$
	*2/*3	-1.74761	0.20391	$<2x10^{-16}$
	*3/*3	-3.40061	0.33091	$<2x10^{-16}$
VKORC1 rs9923231	A/G	-0.90112	0.04959	$<2x10^{-16}$
	A/A	-2.01863	0.06799	$<2x10^{-16}$
Age at start		-0.03686	0.00172	$<2x10^{-16}$
Female gender[1]		-0.27698	0.04682	$4.2x10^{-9}$
Interaction[2]		-0.06992	0.01867	0.00019

For example, a 60 year old man with no interacting drugs is genotyped as *CYP2C9* *1/*3 and *VKORC1* rs9923231 A/A alleles. The calculation is 9.46832 + (-0.97546) + (-2.01863) + (-0.03686)*60 + (-0.27698)*0 + (-0.06992)*0 = 4.26263. By squaring the obtained number, the predicted dose for this 60 year old man is 18.17 mg/week. Hopefully, new findings will improve the predictive power of such tools.

In chapter VI, the analysis of the severe bleeding complication phenotype which occurs in

~2% of warfarin treated patients as a result of using this drug is presented. The results suggest that bleeding is likely to be a complex trait, but this needs to be further addressed. Due to the small size of bleeding patients (64 in total) our analyses lacked statistical power, and further samples will be needed to replicate some of the initial findings. Three members of the protein disulfide isomerise gene family, *P4HB*, *PDIA4* and *PDIA6*, may be implicated in warfarin-induced severe bleeding complication and will be of interest to pursue further including some functional characterisation.

For reasons outlined in Chapter VI, 11 of the genes (see Table 6.20) were selected for exon re-sequencing in 48 warfarin treated patients including 36 bleeders and 12 non-bleeders, as well as 48 CEPH Caucasians. To complete our knowledge of common variants possibly enriched in bleeders i.e. no high penetrance variants which cause bleeding complication, investigation of rare functional variants was also performed. Compared to typical Mendelian disorders, the 2% prevalence of bleeding among warfarin treated patients is relatively common and it is more likely to be caused by a common, possibly low frequency (1-5%) variant. Due to the small number of samples sequenced, the common SNPs we identified in the Swedish patients but not in the 48 CEPH individuals may be of interest and need to be further assessed in all patients in the WARG study.

The findings from studying 35 candidate genes were described in this thesis. With the latest developments in genotyping technology, genome-wide association (GWA) studies have become both feasible and cost effective in large sample sizes (Amos 2007). The WARG study has a reasonably large sample size to provide statistical power to detect moderate effects in a genome-wide scan for warfarin dose. Our lab and the collaborating groups in Sweden have come to the view that a GWA approach is the way forward. The first 1000 samples of the WARG study have been scanned with the Illumina Hap370K array. Initial analysis indicated that there are no loci with effects comparable to *VKORC1* and *CYP2C9* (N Soranzo and P Deloukas, personal communication). This is in agreement with a report by Rieder and colleagues presented in the recent Pharmacogenomics meeting (2007) held in Hinxton, United Kingdom, in which an underpowered genome scan of 186 samples with the Illumina 550K chip did not yield any statistically significant findings. The way forward will be to combine many more genome scans and to increase power to detect small effect sizes.

7.2 MOLECULAR MECHANISM

Studies aiming to link genotype to phenotype often bypass many intermediate steps which are biologically relevant. The study of gene expression as intermediate phenotype is well documented in the recent literature (Nevins and Potti 2007; Ozdemir et al. 2006). Expression

data can directly point to new candidate genes based on their response to a stimulus under investigation, e.g. a drug such as warfarin. This information can become even more powerful once combined with genome-wide data on sequence variants (Goring et al. 2007; Stranger et al. 2007).

I have initiated work with two well-established human hepatoma cell lines, Hep G2 and Hep 3B, which I treated with isomeric mixtures of warfarin. The warfarin-treated cell lines could provide systematic information as to which genes are up- or down-regulated in response to warfarin *in vitro*. A previous study has demonstrated the accumulation of des-r-carboxyprothrombin or proteins induced by vitamin K antagonism (PIVKA-II) in Hep G2 cell hepatoma in the presence of warfarin (Lawley et al. 2006). It is planned to look at the effect of warfarin treatment on gene expression in hepatoma cells as previously described by Lawley et al.

This will help understand the molecular interaction when warfarin is administered in patients. Genes showing differential expression response could be influential to, or as an outcome of, inter-individual dose variation and bleeding ADR. We anticipate that such data can be used to provide additional statistical weight to weakly associated SNPs in the genome scan. Although the experiment is done with human hepatoma, this experiment would be best done using

primary hepatocytes as hepatocytes retain a more complete drug metabolising system.

7.3 CONCLUDING REMARKS

In August 2007, the US FDA updated the label of warfarin to include information on pharmacogenetic testing. The encouragement, but not compulsoriness, of genetic testing before initiating warfarin therapy is one of the most debated subjects. This thesis provided results that have contributed to the body of evidence used by the US FDA to reach this decision and in a broader sense to pharmacogenomic research towards personalised medicine.

REFERENCES

Ahamed J, Versteeg HH, Kerver M, Chen VM, Mueller BM, Hogg PJ, Ruf W (2006) Disulfide isomerization switches tissue factor from coagulation to cell signaling. Proc Natl Acad Sci U S A 103: 13932-7

Ahmadi KR, Weale ME, Xue ZY, Soranzo N, Yarnall DP, Briley JD, Maruyama Y, Kobayashi M, Wood NW, Spurr NK, Burns DK, Roses AD, Saunders AM, Goldstein DB (2005) A single-nucleotide polymorphism tagging set for human drug metabolism and transport. Nat Genet 37: 84-9

Aiach M, Nicaud V, Alhenc-Gelas M, Gandrille S, Arnaud E, Amiral J, Guize L, Fiessinger JN, Emmerich J (1999) Complex association of protein C gene promoter polymorphism with circulating protein C levels and thrombotic risk. Arterioscler Thromb Vasc Biol 19: 1573-6

Aithal GP, Day CP, Kesteven PJ, Daly AK (1999) Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. Lancet 353: 717-9

Albers GW (1994) Atrial fibrillation and stroke. Three new studies, three remaining questions. Arch Intern Med 154: 1443-8

Amos CI (2007) Successful design and conduct of genome-wide association studies. Hum Mol Genet 16 Spec No. 2: R220-5

Angelillo-Scherrer A, de Frutos P, Aparicio C, Melis E, Savi P, Lupu F, Arnout J, Dewerchin M, Hoylaerts M, Herbert J, Collen D, Dahlback B, Carmeliet P (2001) Deficiency or inhibition of Gas6 causes platelet dysfunction and protects mice against thrombosis. Nat Med 7: 215-21

Aquilante CL, Langaee TY, Lopez LM, Yarandi HN, Tromberg JS, Mohuczy D, Gaston KL, Waddell CD, Chirico MJ, Johnson JA (2006) Influence of coagulation factor, vitamin K epoxide reductase complex subunit 1, and cytochrome P450 2C9 gene polymorphisms on warfarin dose requirements. Clin Pharmacol Ther 79: 291-302

Arbini AA, Bodkin D, Lopaciuk S, Bauer KA (1994) Molecular analysis of Polish patients with factor VII deficiency. Blood 84: 2214-20

Assenat E, Gerbal-Chaloin S, Larrey D, Saric J, Fabre JM, Maurel P, Vilarem MJ, Pascussi JM (2004) Interleukin 1beta inhibits CAR-induced expression of hepatic genes involved in drug and bilirubin clearance. Hepatology 40: 951-60

Balding DJ (2006) A tutorial on statistical methods for population association studies. Nat Rev Genet 7: 781-91

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263-5

Berkner KL (2000) The vitamin K-dependent carboxylase. J Nutr 130: 1877-80

Berkner KL, Pudota BN (1998) Vitamin K-dependent carboxylation of the carboxylase. Proc Natl Acad Sci U S A 95: 466-71

Berkner KL, Runge KW (2004) The physiology of vitamin K nutriture and vitamin K-dependent protein function in atherosclerosis. J Thromb Haemost 2: 2118-32

Bodin L, Verstuyft C, Tregouet DA, Robert A, Dubert L, Funck-Brentano C, Jaillon P, Beaune P, Laurent-Puig P, Becquemont L, Loriot MA (2005) Cytochrome P450 2C9 (CYP2C9) and vitamin K epoxide reductase (VKORC1) genotypes as determinants of acenocoumarol sensitivity. Blood 106: 135-40

Borgiani P, Ciccacci C, Forte V, Romano S, Federici G, Novelli G (2007) Allelic variants in the CYP2C9 and VKORC1 loci and interindividual variability in the anticoagulant dose effect of warfarin in Italians. Pharmacogenomics 8: 1545-1550

Brenner B, Sanchez-Vega B, Wu SM, Lanir N, Stafford DW, Solera J (1998) A missense mutation in gamma-glutamyl carboxylase gene causes combined deficiency of all vitamin K-dependent blood coagulation factors. Blood 92: 4554-9

Bresalier RS, Sandler RS, Quan H, Bolognese JA, Oxenius B, Horgan K, Lines C, Riddell R, Morton D, Lanas A, Konstam MA, Baron JA (2005) Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med 352: 1092-102

Cain D, Hutson SM, Wallin R (1998) Warfarin resistance is associated with a protein component of the vitamin K 2,3-epoxide reductase enzyme complex in rat liver. Thromb Haemost 80: 128-33

Carlquist JF, Horne BD, Muhlestein JB, Lappe DL, Whiting BM, Kolek MJ, Clarke JL, James BC, Anderson JL (2006) Genotypes of the cytochrome p450 isoform, CYP2C9, and the vitamin K epoxide reductase complex subunit 1 conjointly determine stable warfarin dose: a prospective study. J Thromb Thrombolysis 22: 191-7

Cascorbi I, Gerloff T, Johne A, Meisel C, Hoffmeyer S, Schwab M, Schaeffeler E, Eichelbaum M, Brinkmann U, Roots I (2001) Frequency of single nucleotide polymorphisms in the P-glycoprotein drug transporter MDR1 gene in white subjects. Clin Pharmacol Ther 69: 169-74

Cauwenberghs N, Vanhoorelbeke K, Vauterin S, Deckmyn H (2000) Structural determinants within platelet glycoprotein Ibalpha involved in its binding to von Willebrand factor. Platelets 11: 373-8

Cha PC, Mushiroda T, Takahashi A, Saito S, Shimomura H, Suzuki T, Kamatani N, Nakamura Y (2007) High-resolution SNP and haplotype maps of the human gamma-glutamyl carboxylase gene (GGCX) and association study between polymorphisms in GGCX and the warfarin maintenance dose requirement of the Japanese population. J Hum Genet 52: 856-64

Chan YC, Valenti D, Mansfield AO, Stansby G (2000) Warfarin induced skin necrosis. Br J Surg 87: 266-72

Chen JM, Cooper DN, Chuzhanova N, Ferec C, Patrinos GP (2007) Gene conversion: mechanisms, evolution and human disease. Nat Rev Genet 8: 762-75

Chen Y, Ferguson SS, Negishi M, Goldstein JA (2004) Induction of human CYP2C9 by rifampicin, hyperforin, and phenobarbital is mediated by the pregnane X receptor. J

Pharmacol Exp Ther 308: 495-501

Cheverud JM (2001) A simple correction for multiple comparisons in interval mapping genome scans. Heredity 87: 52-8

Clemetson KJ, Clemetson JM (2007) Collagen receptors as potential targets for novel anti-platelet agents. Curr Pharm Des 13: 2673-83

Cohen SM, Rosenthal DS, Karp PJ (1968) Ulcerative colitis and erythrocyte G6PD deficiency. Salicylazosulfapyridine-provoked hemolysis. JAMA 205: 528-30

Conlan MG, Bridges A, Williams E, Marlar R (1988) Familial type II protein C deficiency associated with warfarin-induced skin necrosis and bilateral adrenal hemorrhage. Am J Hematol 29: 226-9

Constable S, Johnson MR, Pirmohamed M (2006) Pharmacogenetics in clinical practice: considerations for testing. Expert Rev Mol Diagn 6: 193-205

Dahlback B (2005) Blood coagulation and its regulation by anticoagulant pathways: genetic pathogenesis of bleeding and thrombotic diseases. J Intern Med 257: 209-23

Daly AK, King BP (2003) Pharmacogenetics of oral anticoagulants. Pharmacogenetics 13: 247-52

D'Ambrosio D, Lecca P, Constantin G, Priami C, Laudanna C (2004) Concurrency in leukocyte vascular recognition: developing the tools for a predictive computer model. Trends Immunol 25: 411-6

D'Ambrosio RL, D'Andrea G, Cappucci F, Chetta M, Di Perna P, Brancaccio V, Grandone E, Margaglione M (2004) Polymorphisms in factor II and factor VII genes modulate oral anticoagulation with warfarin. Haematologica 89: 1510-6

D'Andrea G, D'Ambrosio RL, Di Perna P, Chetta M, Santacroce R, Brancaccio V, Grandone E, Margaglione M (2005) A polymorphism in the VKORC1 gene is associated with an interindividual variability in the dose-anticoagulant effect of warfarin. Blood 105: 645-9

de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D (2005) Efficiency and power in genetic association studies. Nat Genet 37: 1217-23

Delabar JM, Lamour Y, Gegonne A, Davous P, Roudier M, Nicole A, Ceballos I, Amouyel P, Stehelin D, Sinet PM (1986) Rearrangement of chromosome 21 in Alzheimer's disease. Ann Genet 29: 226-8

Devadatta S, Gangadharam PR, Andrews RH, Fox W, Ramakrishnan CV, Selkon JB, Velu S (1960) Peripheral neuritis due to isoniazid. Bull World Health Organ 23: 587-98

Devlin B, Roeder K, Bacanu SA (2001) Unbiased methods for population-based association studies. Genet Epidemiol 21: 273-84

Drayer DE, Reidenberg MM (1977) Clinical consequences of polymorphic acetylation of basic drugs. Clin Pharmacol Ther 22: 251-8

Dudbridge F (2003) Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol 25: 115-21

Eichelbaum M, Ingelman-Sundberg M, Evans WE (2006) Pharmacogenomics and individualized

drug therapy. Annu Rev Med 57: 119-37

Emanuel BS, Saitta SC (2007) From microscopes to microarrays: dissecting recurrent chromosomal rearrangements. Nat Rev Genet 8: 869-83

ENCODE Project Consortium (2004) The ENCODE (ENCyclopedia Of DNA Elements) Project. Science 306: 636-40

Engesser L, Broekmans AW, Briet E, Brommer EJ, Bertina RM (1987) Hereditary protein S deficiency: clinical manifestations. Ann Intern Med 106: 677-82

Evans DA, White TA (1964) Human Acetylation Polymorphism. J Lab Clin Med 63: 394-403

Evans WE, Relling MV (1999) Pharmacogenomics: translating functional genomics into rational therapeutics. Science 286: 487-91

Fasco MJ, Principe LM (1980) Vitamin K1 hydroquinone formation catalyzed by a microsomal reductase system. Biochem Biophys Res Commun 97: 1487-92

Feuk L, Carson AR, Scherer SW (2006) Structural variation in the human genome. Nat Rev Genet 7: 85-97

Frayling TM (2007) Genome-wide association studies provide new insights into type 2 diabetes aetiology. Nat Rev Genet 8: 657-62

Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Sun W, Wang H, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, et al. (2007) A second generation human haplotype map of over 3.1 million SNPs. Nature 449: 851-61

Fredman D, White SJ, Potter S, Eichler EE, Den Dunnen JT, Brookes AJ (2004) Complex SNP-related sequence variation in segmental genome duplications. Nat Genet 36: 861-6

Fujita K, Sasaki Y (2007) Pharmacogenomics in drug-metabolizing enzymes catalyzing anticancer drugs for personalized cancer chemotherapy. Curr Drug Metab 8: 554-62

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. Science 296: 2225-9

Geick A, Eichelbaum M, Burk O (2001) Nuclear receptor response elements mediate induction of intestinal MDR1 by rifampin. J Biol Chem 276: 14581-7

Geisen C, Watzka M, Sittinger K, Steffens M, Daugela L, Seifried E, Muller CR, Wienker TF, Oldenburg J (2005) VKORC1 haplotypes and their impact on the inter-individual and inter-ethnical variability of oral anticoagulation. Thromb Haemost 94: 773-9

Goldstein JA, Faletto MB, Romkes-Sparks M, Sullivan T, Kitareewan S, Raucy JL, Lasker JM, Ghanayem BI (1994) Evidence that CYP2C19 is the major (S)-mephenytoin 4'-hydroxylase in humans. Biochemistry 33: 1743-52

Goodbourn SE, Higgs DR, Clegg JB, Weatherall DJ (1983) Molecular basis of length polymorphism in the human zeta-globin gene complex. Proc Natl Acad Sci U S A 80: 5022-6

Goring HH, Curran JE, Johnson MP, Dyer TD, Charlesworth J, Cole SA, Jowett JB, Abraham LJ, Rainwater DL, Comuzzie AG, Mahaney MC, Almasy L, Maccluer JW, Kissebah AH, Collier GR, Moses EK, Blangero J (2007) Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. Nat Genet 39: 1208-16

Gould WR, Baxi SM, Schroeder R, Peng YW, Leadley RJ, Peterson JT, Perrin LA (2005) Gas6 receptors Axl, Sky and Mer enhance platelet activation and regulate thrombotic responses. J Thromb Haemost 3: 733-41

Grant DJ, Bell DA (2000) Bilirubin UDP-glucuronosyltransferase 1A1 gene polymorphisms: susceptibility to oxidative damage and cancer? Mol Carcinog 29: 198-204

Gregory SG, Barlow KF, McLay KE, Kaul R, Swarbreck D, Dunham A, Scott CE, Howe KL, Woodfine K, Spencer CC, Jones MC, Gillson C, Searle S, Zhou Y, Kokocinski F, McDonald L, Evans R, Phillips K, Atkinson A, Cooper R, Jones C, Hall RE, Andrews TD, Lloyd C, Ainscough R, Almeida JP, Ambrose KD, Anderson F, Andrew RW, Ashwell RI, Aubin K, Babbage AK, Bagguley CL, Bailey J, Beasley H, Bethel G, Bird CP, Bray-Allen S, Brown JY, Brown AJ, Buckley D, Burton J, Bye J, Carder C, Chapman JC, Clark SY, Clarke G, Clee C, Cobley V, Collier RE, Corby N, Coville GJ, Davies J, Deadman R, Dunn M, Earthrowl M, Ellington AG, Errington H, Frankish A, Frankland J, French L, Garner P, Garnett J, Gay L, Ghori MR, Gibson R, Gilby LM, Gillett W, Glithero RJ, Grafham DV, Griffiths C, Griffiths-Jones S, Grocock R, Hammond S, Harrison ES, Hart E, Haugen E, Heath PD, Holmes S, Holt K, Howden PJ, Hunt AR, Hunt SE, Hunter G, Isherwood J, James R, Johnson C, Johnson D, Joy A, Kay M, Kershaw JK, Kibukawa M, Kimberley AM, King A, Knights AJ, Lad H, Laird G, Lawlor S, Leongamornlert DA, Lloyd DM, et al. (2006) The DNA sequence and biological annotation of human chromosome 1. Nature 441: 315-21

Guasch JF, Cannegieter S, Reitsma PH, van't Veer-Korthof ET, Bertina RM (1998) Severe coagulation factor V deficiency caused by a 4 bp deletion in the factor V gene. Br J Haematol 101: 32-9

Guenthner TM, Cai D, Wallin R (1998) Co-purification of microsomal epoxide hydrolase with the warfarin-sensitive vitamin K1 oxide reductase of the vitamin K cycle. Biochem Pharmacol 55: 169-75

Hasegawa Y, Ando Y, Ando M, Hashimoto N, Imaizumi K, Shimokata K (2006) Pharmacogenetic approach for cancer treatment-tailored medicine in practice. Ann N Y Acad Sci 1086: 223-32

Hasstedt SJ, Bovill EG, Callas PW, Long GL (1998) An unknown genetic defect increases venous thrombosis risk, through interaction with protein C deficiency. Am J Hum Genet 63: 569-76

Higashi MK, Veenstra DL, Kondo LM, Wittkowsky AK, Srinouanprachanh SL, Farin FM, Rettie AE (2002) Association between CYP2C9 genetic variants and anticoagulation-related outcomes during warfarin therapy. JAMA 287: 1690-8

Hughes HB, Biehl JP, Jones AP, Schmidt LH (1954) Metabolism of isoniazid in man as related to the occurrence of peripheral neuritis. Am Rev Tuberc 70: 266-73

Hull RD, Pineo GF (1995) Current concepts of anticoagulation therapy. Clin Chest Med 16: 269-80

Ingelman-Sundberg M, Rodriguez-Antona C (2005) Pharmacogenetics of drug-metabolizing enzymes: implications for a safer and more effective drug therapy. Philos Trans R Soc Lond B Biol Sci 360: 1563-70

Ingelman-Sundberg M, Sim SC, Gomez A, Rodriguez-Antona C (2007) Influence of cytochrome P450 polymorphisms on drug therapies: pharmacogenetic, pharmacoepigenetic and clinical aspects. Pharmacol Ther 116: 496-526

Investigators AF (1994) Risk factors for stroke and efficacy of antithrombotic therapy in atrial fibrillation. Analysis of pooled data from five randomized controlled trials. Arch Intern Med 154: 1449-57

Ishikawa T, Hirano H, Onishi Y, Sakurai A, Tarui S (2004) Functional evaluation of ABCB1 (P-glycoprotein) polymorphisms: high-speed screening and structure-activity relationship analyses. Drug Metab Pharmacokinet 19: 1-14

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with Clustal X. Trends Biochem Sci 23: 403-5

Jeffreys AJ, Turner M, Debenham P (1991) The efficiency of multilocus DNA fingerprint probes for individualization and establishment of family relationships, determined from extensive casework. Am J Hum Genet 48: 824-40

Jeffreys AJ, Wilson V, Wong Z, Royle N, Patel I, Kelly R, Clarkson R (1987) Highly variable minisatellites and DNA fingerprints. Biochem Soc Symp 53: 165-80

Jiang W, Yu G, Liu P, Geng Q, Chen L, Lin Q, Ren X, Ye W, He Y, Guo Y, Duan S, Wen J, Li H, Qi Y, Jiang C, Zheng Y, Liu C, Si E, Zhang Q, Tian Q, Du C (2006) Structure and function of glucose-6-phosphate dehydrogenase-deficient variants in Chinese population. Hum Genet 119: 463-78

Jirtle RL, Skinner MK (2007) Environmental epigenomics and disease susceptibility. Nat Rev Genet 8: 253-62

Jobling MA, Tyler-Smith C (2003) The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 4: 598-612

Kakar P, Lane D, Lip GY (2006) Bleeding risk stratification models in deciding on anticoagulation in patients with atrial fibrillation: a useful complement to stroke risk stratification schema. Chest 130: 1296-9

Kaminsky LS, de Morais SM, Faletto MB, Dunbar DA, Goldstein JA (1993) Correlation of human cytochrome P4502C substrate specificities with primary structure: warfarin as a probe. Mol Pharmacol 43: 234-9

Kaminsky LS, Zhang ZY (1997) Human P450 metabolism of warfarin. Pharmacol Ther 73: 67-74

Kawanishi C, Lundgren S, Agren H, Bertilsson L (2004) Increased incidence of CYP2D6 gene duplication in patients with persistent mood disorders: ultrarapid metabolism of antidepressants as a cause of nonresponse. A pilot study. Eur J Clin Pharmacol 59: 803-7

Kessler CM (2006) Urgent reversal of warfarin with prothrombin complex concentrate: where are the evidence-based data? J Thromb Haemost 4: 963-6

Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM (2007) A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science 315: 525-8

Kindmark A, Jawaid A, Harbron CG, Barratt BJ, Bengtsson OF, Andersson TB, Carlsson S, Cederbrant KE, Gibson NJ, Armstrong M, Lagerstrom-Fermer ME, Dellsen A, Brown EM, Thornton M, Dukes C, Jenkins SC, Firth MA, Harrod GO, Pinel TH, Billing-Clason SM, Cardon LR, March RE (2007) Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. Pharmacogenomics J

Kirchheiner J, Schmidt H, Tzvetkov M, Keulen JT, Lotsch J, Roots I, Brockmoller J (2007) Pharmacokinetics of codeine and its metabolite morphine in ultra-rapid metabolizers due to CYP2D6 duplication. Pharmacogenomics J 7: 257-65

Knight JC, Udalova I, Hill AV, Greenwood BM, Peshu N, Marsh K, Kwiatkowski D (1999) A polymorphism that affects OCT-1 binding to the TNF promoter region is associated with severe malaria. Nat Genet 22: 145-50

Kohlmeier M, Salomon A, Saupe J, Shearer MJ (1996) Transport of vitamin K to bone in humans. J Nutr 126: 1192S-6S

Kohn MH, Pelz HJ (2000) A gene-anchored map position of the rat warfarin-resistance locus, Rw, and its orthologs in mice and humans. Blood 96: 1996-8

Kohnke H, Sorlin K, Granath G, Wadelius M (2005) Warfarin dose related to apolipoprotein E (APOE) genotype. Eur J Clin Pharmacol 61: 381-8

Kornecki E, Lenox RH, Hardwick DH, Bergdahl JA, Ehrlich YH (1987) Interactions of the alkyl-ether-phospholipid, platelet activating factor (PAF) with platelets, neural cells, and the psychotropic drugs triazolobenzodiazepines. Adv Exp Med Biol 221: 477-88

Krijanovski Y, Proulle V, Mahdi F, Dreyfus M, Muller-Esterl W, Schmaier AH (2003) Characterization of molecular defects of Fitzgerald trait and another novel high-molecular-weight kininogen-deficient patient: insights into structural requirements for kininogen expression. Blood 101: 4430-6

Kristensen SR (2002) Warfarin treatment of a patient with coagulation factor IX propeptide mutation causing warfarin hypersensitivity. Blood 100: 2676-7

Kuehl P, Zhang J, Lin Y, Lamba J, Assem M, Schuetz J, Watkins PB, Daly A, Wrighton SA, Hall SD, Maurel P, Relling M, Brimer C, Yasuda K, Venkataramanan R, Strom S, Thummel K, Boguski MS, Schuetz E (2001) Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. Nat Genet 27: 383-91

Läkemedelsindustriföreningen (2007) Akademi-FASS : förteckning över humanläkemedel. Läkemedelsindustriföreningen (LIF), Stockholm

Lamon-Fava S, Sadowski JA, Davidson KW, O'Brien ME, McNamara JR, Schaefer EJ (1998) Plasma lipoproteins as carriers of phylloquinone (vitamin K1) in humans. Am J Clin Nutr 67: 1226-31

Landefeld CS, Beyth RJ (1993) Anticoagulant-related bleeding: clinical epidemiology, prediction, and prevention. Am J Med 95: 315-28

Lander ES, Botstein D (1986) Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. Proc Natl Acad Sci U S A 83: 7353-7

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409: 860-921

Larsen TB, Lassen JF, Dahler-Eriksen BS, Petersen PH, Brandslund I (1998) Effect of anticoagulant therapy on the hypercoagulable state in patients carrying the factor V Arg506Gln mutation. Thromb Res 92: 157-62

Lawley WJ, Charlton AJ, Hughson EJ, Grundy HH, Brown PM, Jones A (2006) Development of a cell culture/ELISA assay to detect anticoagulant rodenticides and its application to analysis of rodenticide treated grain. J Agric Food Chem 54: 1588-93

Lazarou J, Pomeranz BH, Corey PN (1998) Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 279: 1200-5

Lee C, Iafrate AJ, Brothman AR (2007) Copy number variations and clinical cytogenetic diagnosis of constitutional disorders. Nat Genet 39: S48-54

Lehmann H, Ryan E (1956) The familial incidence of low pseudocholinesterase level. Lancet 271: 124

Lehmann JM, McKee DD, Watson MA, Willson TM, Moore JT, Kliewer SA (1998) The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions. J Clin Invest 102: 1016-23

Levine MN, Raskob G, Landefeld S, Hirsh J (1995) Hemorrhagic complications of anticoagulant treatment. Chest 108: 276S-290S

Li H (2001) A permutation procedure for the haplotype method for identification of disease-predisposing variants. Ann Hum Genet 65: 189-96

Li T, Chang CY, Jin DY, Lin PJ, Khvorova A, Stafford DW (2004) Identification of the gene for vitamin K epoxide reductase. Nature 427: 541-4

Ligtenberg MJ, Gennissen AM, Vos HL, Hilkens J (1991) A single nucleotide polymorphism in an exon dictates allele dependent differential splicing of episialin mRNA. Nucleic Acids Res 19: 297-301

Limdi NA, McGwin G, Goldstein JA, Beasley TM, Arnett DK, Adler BK, Baird MF, Acton RT (2007) Influence of CYP2C9 and VKORC1 1173C/T Genotype on the Risk of Hemorrhagic Complications in African-American and European-American Patients on Warfarin. Clin Pharmacol Ther

Lindh JD, Holm L, Dahl ML, Alfredsson L, Rane A (2007) Incidence and predictors of severe bleeding during warfarin treatment. J Thromb Thrombolysis

Lindh JD, Kublickas M, Westgren M, Rane A (2004) Internet based clinical trial protocols -- as applied to a study of warfarin pharmacogenetics. Br J Clin Pharmacol 58: 482-7

Lindqvist PG, Svensson PJ, Dahlback B, Marsal K (1998) Factor V Q506 mutation (activated protein C resistance) associated with reduced intrapartum blood loss--a possible evolutionary selection mechanism. Thromb Haemost 79: 69-73

Loebstein R, Vecsler M, Kurnik D, Austerweil N, Gak E, Halkin H, Almog S (2005) Common genetic variants of microsomal epoxide hydrolase affect warfarin dose requirements beyond the effect of cytochrome P450 2C9. Clin Pharmacol Ther 77: 365-72

Loewen P, Sunderji R, Gin K (1998) The efficacy and safety of combination warfarin and ASA therapy: a systematic review of the literature and update of guidelines. Can J Cardiol 14: 717-26

Lovett PS, Duvall EJ, Keggins KM (1976) Bacillus pumilus plasmid pPL10: properties and insertion into Bacillus subtilis 168 by transformation. J Bacteriol 127: 817-28

Lucier GW, Lui EM, Lamartiniere CA (1979) Metabolic activation/deactivation reactions during perinatal development. Environ Health Perspect 29: 7-16

Lupski JR (2007) Genomic rearrangements and sporadic disease. Nat Genet 39: S43-7

Mahtani MM, Willard HF (1993) A polymorphic X-linked tetranucleotide repeat locus displaying a high rate of new mutation: implications for mechanisms of mutation at short tandem repeat loci. Hum Mol Genet 2: 431-7

Manfioletti G, Brancolini C, Avanzi G, Schneider C (1993) The protein encoded by a growth arrest-specific gene (gas6) is a new member of the vitamin K-dependent proteins related to protein S, a negative coregulator in the blood coagulation cascade. Mol Cell Biol 13:

4976-85

Martinez-Gonzalez J, Badimon L (2007) Mechanisms underlying the cardiovascular effects of COX-inhibition: benefits and risks. Curr Pharm Des 13: 2215-27

Matsuda M, Sugo T, Sakata Y, Murayama H, Mimuro J, Tanabe S, Yoshitake S (1988) A thrombotic state due to an abnormal protein C. N Engl J Med 319: 1265-8

McGehee WG, Klotz TA, Epstein DJ, Rapaport SI (1984) Coumarin necrosis associated with hereditary protein C deficiency. Ann Intern Med 101: 59-60

Michels E, De Preter K, Van Roy N, Speleman F (2007) Detection of DNA copy number alterations in cancer by array comparative genomic hybridization. Genet Med 9: 574-84

Mockenhaupt M, Schopf E (1996) Epidemiology of drug-induced severe skin reactions. Semin Cutan Med Surg 15: 236-43

Molokhia M, McKeigue P (2006) EUDRAGENE: European collaboration to establish a case-control DNA collection for studying the genetic basis of adverse drug reactions. Pharmacogenomics 7: 633-8

Moore LB, Goodwin B, Jones SA, Wisely GB, Serabjit-Singh CJ, Willson TM, Collins JL, Kliewer SA (2000) St. John's wort induces hepatic drug metabolism through activation of the pregnane X receptor. Proc Natl Acad Sci U S A 97: 7500-2

Munoz X, Sumoy L, Ramirez-Lorca R, Villar J, de Frutos PG, Sala N (2004) Human vitamin K-dependent GAS6: gene structure, allelic variation, and association with stroke. Hum Mutat 23: 506-12

Mustard JF, Packham MA (1975) Platelets, thrombosis and drugs. Drugs 9: 19-76

Nagar S, Blanchard RL (2006) Pharmacogenetics of uridine diphosphoglucuronosyltransferase (UGT) 1A family members and its role in patient response to irinotecan. Drug Metab Rev 38: 393-409

Nakagawa T, Kishino S, Itoh S, Sugawara M, Miyazaki K (2003) Differential binding of disopyramide and warfarin enantiomers to human alpha(1)-acid glycoprotein variants. Br J Clin Pharmacol 56: 664-9

Nevins JR, Potti A (2007) Mining gene expression profiles: expression signatures as cancer phenotypes. Nat Rev Genet 8: 601-9

Newman DL, Hoffjan S, Bourgain C, Abney M, Nicolae RI, Profits ET, Grow MA, Walker K, Steiner L, Parry R, Reynolds R, McPeek MS, Cheng S, Ober C (2004) Are common disease susceptibility alleles the same in outbred and founder populations? Eur J Hum Genet 12: 584-90

Norbert PW, Roses AD (2003) Pharmacogenetics and pharmacogenomics: recent developments, their clinical relevance and some ethical, social, and legal implications. J Mol Med 81: 135-40

Nyholt DR (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet 74: 765-9

Olson ST, Bjork I (1994) Regulation of thrombin activity by antithrombin and heparin. Semin Thromb Hemost 20: 373-409

Ovcharenko I, Loots GG, Hardison RC, Miller W, Stubbs L (2004) zPicture: dynamic alignment and visualization tool for analyzing conservation profiles. Genome Res 14: 472-7

Ozdemir V, Williams-Jones B, Glatt SJ, Tsuang MT, Lohr JB, Reist C (2006) Shifting emphasis from pharmacogenomics to theragnostics. Nat Biotechnol 24: 942-6

Patil N, Berno AJ, Hinds DA, Barrett WA, Doshi JM, Hacker CR, Kautzer CR, Lee DH, Marjoribanks C, McDonough DP, Nguyen BT, Norris MC, Sheehan JB, Shen N, Stern D, Stokowski RP, Thomas DJ, Trulson MO, Vyas KR, Frazer KA, Fodor SP, Cox DR (2001) Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. Science 294: 1719-23

Peacock E, Whiteley P (2005) Perlegen sciences, inc. Pharmacogenomics 6: 439-42

Pennacchio LA, Rubin EM (2001) Genomic strategies to identify mammalian regulatory sequences. Nat Rev Genet 2: 100-9

Permutt MA, Elbein SC (1990) Insulin gene in diabetes. Analysis through RFLP. Diabetes Care 13: 364-74

Peterson S, Bigler J, Horner NK, Potter JD, Lampe JW (2005) Cruciferae interact with the UGT1A1*28 polymorphism to determine serum bilirubin levels in humans. J Nutr 135: 1051-5

Peyvandi F, Spreafico M, Siboni SM, Moia M, Mannucci PM (2004) CYP2C9 genotypes and dose requirements during the induction phase of oral anticoagulant therapy. Clin Pharmacol Ther 75: 198-203

Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, Farrar K, Park BK, Breckenridge AM (2004) Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. Bmj 329: 15-9

Pirmohamed M, Park BK (2003) Cytochrome P450 enzyme polymorphisms and adverse drug reactions. Toxicology 192: 23-32

Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, Pe'er I, Burtt NP, Blumenstiel B, Defelice M, Parkin M, Barry R, Winslow W, Healy C, Graham RR, Neale BM, Izmailova E, Roubenoff R, Parker AN, Glass R, Karlson EW, Maher N, Hafler DA, Lee DM, Seldin MF, Remmers EF, Lee AT, Padyukov L, Alfredsson L, Coblyn J, Weinblatt ME, Gabriel SB, Purcell S, Klareskog L, Gregersen PK, Shadick NA, Daly MJ, Altshuler D (2007) Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet

Polymeropoulos MH, Higgins JJ, Golbe LI, Johnson WG, Ide SE, Di Iorio G, Sanges G, Stenroos ES, Pho LT, Schaffer AA, Lazzarini AM, Nussbaum RL, Duvoisin RC (1996) Mapping of a gene for Parkinson's disease to chromosome 4q21-q23. Science 274: 1197-9

Prandoni P, Lensing AW, Piccioli A, Bernardi E, Simioni P, Girolami B, Marchiori A, Sabbion P, Prins MH, Noventa F, Girolami A (2002) Recurrent venous thromboembolism and bleeding complications during anticoagulant treatment in patients with cancer and venous thrombosis. Blood 100: 3484-8

Prokunina L, Castillejo-Lopez C, Oberg F, Gunnarsson I, Berg L, Magnusson V, Brookes AJ,

Tentler D, Kristjansdottir H, Grondal G, Bolstad AI, Svenungsson E, Lundberg I, Sturfelt G, Jonssen A, Truedsson L, Lima G, Alcocer-Varela J, Jonsson R, Gyllensten UB, Harley JB, Alarcon-Segovia D, Steinsson K, Alarcon-Riquelme ME (2002) A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. Nat Genet 32: 666-9

Rau T, Wohlleben G, Wuttke H, Thuerauf N, Lunkenheimer J, Lanczik M, Eschenhagen T (2004) CYP2D6 genotype: impact on adverse effects and nonresponse during treatment with antidepressants-a pilot study. Clin Pharmacol Ther 75: 386-93

Rawlins MD, James OF, Williams FM, Wynne H, Woodhouse KW (1987) Age and the metabolism of drugs. Q J Med 64: 545-7

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME (2006) Global variation in copy number in the human genome. Nature 444: 444-54

Reed GL, Houng AK, Fitzgerald ML (1999) Human platelets contain SNARE proteins and a Sec1p homologue that interacts with syntaxin 4 and is phosphorylated after thrombin activation: implications for platelet secretion. Blood 93: 2617-26

Reich DE, Gabriel SB, Altshuler D (2003) Quality and completeness of SNP databases. Nat Genet 33: 457-8

Reitsma PH, van der Heijden JF, Groot AP, Rosendaal FR, Buller HR (2005) A C1173T dimorphism in the VKORC1 gene determines coumarin sensitivity and bleeding risk. PLoS Med 2: e312

Rettie AE, Wienkers LC, Gonzalez FJ, Trager WF, Korzekwa KR (1994) Impaired (S)-warfarin metabolism catalysed by the R144C allelic variant of CYP2C9. Pharmacogenetics 4: 39-42

Riddel JP, Jr., Aouizerat BE, Miaskowski C, Lillicrap DP (2007) Theories of blood coagulation. J Pediatr Oncol Nurs 24: 123-31

Rieder MJ, Reiner AP, Gage BF, Nickerson DA, Eby CS, McLeod HL, Blough DK, Thummel KE, Veenstra DL, Rettie AE (2005) Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. N Engl J Med 352: 2285-93

Rieder MJ, Reiner AP, Rettie AE (2007) Gamma-glutamyl carboxylase (GGCX) tagSNPs have limited utility for predicting warfarin maintenance dose. J Thromb Haemost 5: 2227-34

Rodriguez-Revenga L, Mila M, Rosenberg C, Lamb A, Lee C (2007) Structural variation in the human genome: the impact of copy number variants on clinical diagnosis. Genet Med 9: 600-6

Rosand J, Hylek EM, O'Donnell HC, Greenberg SM (2000) Warfarin-associated hemorrhage and cerebral amyloid angiopathy: a genetic and pathologic study. Neurology 55: 947-51

Roses AD (2004) Pharmacogenetics and drug development: the path to safer and more effective

drugs. Nat Rev Genet 5: 645-56

Rost S, Fregin A, Ivaskevicius V, Conzelmann E, Hortnagel K, Pelz HJ, Lappegard K, Seifried E, Scharrer I, Tuddenham EG, Muller CR, Strom TM, Oldenburg J (2004) Mutations in VKORC1 cause warfarin resistance and multiple coagulation factor deficiency type 2. Nature 427: 537-41

Rost S, Fregin A, Koch D, Compes M, Muller CR, Oldenburg J (2004) Compound heterozygous mutations in the gamma-glutamyl carboxylase gene cause combined deficiency of all vitamin K-dependent blood coagulation factors. Br J Haematol 126: 546-9

Routledge PA, Chapman PH, Davies DM, Rawlins MD (1979) Pharmacokinetics and pharmacodynamics of warfarin at steady state. Br J Clin Pharmacol 8: 243-7

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Sun W, Wang H, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, et al. (2007) Genome-wide detection and characterization of positive selection in human populations. Nature 449: 913-8

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, Coggill PC, Rice CM, Ning Z, Rogers J, Bentley DR, Kwok PY, Mardis ER, Yeh RT, Schultz B, Cook L, Davenport R, Dante M, Fulton L, Hillier L, Waterston RH, McPherson JD, Gilman B, Schaffner S, Van Etten WJ, Reich D, Higgins J, Daly MJ, Blumenstiel B, Baldwin J, Stange-Thomann N, Zody MC, Linton L, Lander ES, Altshuler D (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409: 928-33

Sadler JE (2004) Medicine: K is for koagulation. Nature 427: 493-4

Sanderson S, Emery J, Higgins J (2005) CYP2C9 gene variants, drug dose, and bleeding risk in warfarin-treated patients: a HuGEnet systematic review and meta-analysis. Genet Med 7: 97-104

Schalekamp T, Brasse BP, Roijers JF, Chahid Y, van Geest-Daalderop JH, de Vries-Goldschmeding H, van Wijk EM, Egberts AC, de Boer A (2006) VKORC1 and CYP2C9 genotypes and acenocoumarol anticoagulation status: interaction between both genotypes affects overanticoagulation. Clin Pharmacol Ther 80: 13-22

Schuster I, Bernhardt R (2007) Inhibition of cytochromes p450: existing and new promising therapeutic targets. Drug Metab Rev 39: 481-99

Schwartz S, Elnitski L, Li M, Weirauch M, Riemer C, Smit A, Green ED, Hardison RC, Miller

W (2003) MultiPipMaker and supporting tools: Alignments and analysis of multiple genomic DNA sequences. Nucleic Acids Res 31: 3518-24

Schwarz UI (2003) Clinical relevance of genetic polymorphisms in the human CYP2C9 gene. Eur J Clin Invest 33 Suppl 2: 23-30

Schwarz UI, Ritchie MD, Bradford Y, Li C, Dudek SM, Frye-Anderson A, Kim RB, Roden DM, Stein CM (2008) Genetic determinants of response to warfarin during initial anticoagulation. N Engl J Med 358: 999-1008

Sconce EA, Daly AK, Khan TI, Wynne HA, Kamali F (2006) APOE genotype makes a small contribution to warfarin dose requirements. Pharmacogenet Genomics 16: 609-11

Sconce EA, Khan TI, Wynne HA, Avery P, Monkhouse L, King BP, Wood P, Kesteven P, Daly AK, Kamali F (2005) The impact of CYP2C9 and VKORC1 genetic polymorphism and patient characteristics upon warfarin dose requirements: proposal for a new dosing regimen. Blood 106: 2329-33

Shikata E, Ieiri I, Ishiguro S, Aono H, Inoue K, Koide T, Ohgi S, Otsubo K (2004) Association of pharmacokinetic (CYP2C9) and pharmacodynamic (factors II, VII, IX, and X; proteins S and C; and gamma-glutamyl carboxylase) gene variants with warfarin sensitivity. Blood 103: 2630-5

Shireman TI, Mahnken JD, Howard PA, Kresowik TF, Hou Q, Ellerbeck EF (2006) Development of a contemporary bleeding risk model for elderly warfarin recipients. Chest 130: 1390-6

Smyth DJ, Cooper JD, Bailey R, Field S, Burren O, Smink LJ, Guja C, Ionescu-Tirgoviste C, Widmer B, Dunger DB, Savage DA, Walker NM, Clayton DG, Todd JA (2006) A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. Nat Genet 38: 617-9

Solus JF, Arietta BJ, Harris JR, Sexton DP, Steward JQ, McMunn C, Ihrie P, Mehall JM, Edwards TL, Dawson EP (2004) Genetic variation in eleven phase I drug metabolism genes in an ethnically diverse population. Pharmacogenomics 5: 895-931

Spek CA, Koster T, Rosendaal FR, Bertina RM, Reitsma PH (1995) Genotypic variation in the promoter region of the protein C gene is associated with plasma protein C levels and thrombotic risk. Arterioscler Thromb Vasc Biol 15: 214-8

Spivakov M, Fisher AG (2007) Epigenetic signatures of stem-cell identity. Nat Rev Genet 8: 263-71

Stella VJ, Nti-Addae KW (2007) Prodrug strategies to overcome poor water solubility. Adv Drug Deliv Rev 59: 677-94

Stram DO (2004) Tag SNP selection for association studies. Genet Epidemiol 27: 365-74

Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, Ingle CE, Dunning M, Flicek P, Koller D, Montgomery S, Tavare S, Deloukas P, Dermitzakis ET (2007) Population genomics of human gene expression. Nat Genet 39: 1217-24

Sullivan-Klose TH, Ghanayem BI, Bell DA, Zhang ZY, Kaminsky LS, Shenfield GM, Miners JO, Birkett DJ, Goldstein JA (1996) The role of the CYP2C9-Leu359 allelic variant in

the tolbutamide polymorphism. Pharmacogenetics 6: 341-9

Sussman N WM, Butler T, Cali J, Riss T, Kelly J (2002) The predictice nature of high throughput toxicity screening using a human hepatocyte cell line. Cell Notes: 7-10

Suttie JW (1980) Mechanism of action of vitamin K: synthesis of gamma-carboxyglutamic acid. CRC Crit Rev Biochem 8: 191-223

Synold TW, Dussault I, Forman BM (2001) The orphan nuclear receptor SXR coordinately regulates drug metabolism and efflux. Nat Med 7: 584-90

The International HapMap Consortium (2003) The International HapMap Project. Nature 426: 789-96

The International HapMap Consortium (2005) A haplotype map of the human genome. Nature 437: 1299-320

The International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. Nature 431: 931-45

The Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661-78

Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J, Donn R, Symmons D, Hider S, Bruce IN, Wilson AG, Marinou I, Morgan A, Emery P, Carter A, Steer S, Hocking L, Reid DM, Wordsworth P, Harrison P, Strachan D, Worthington J (2007) Rheumatoid arthritis association at 6q23. Nat Genet

Tishkoff SA, Varkonyi R, Cahinhinan N, Abbes S, Argyropoulos G, Destro-Bisol G, Drousiotou A, Dangerfield B, Lefranc G, Loiselet J, Piro A, Stoneking M, Tagarelli A, Tagarelli G, Touma EH, Williams SM, Clark AG (2001) Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. Science 293: 455-62

Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, Sawada T, Suzuki M, Nagasaki M, Ohtsuki M, Ono M, Furukawa H, Nagashima M, Yoshino S, Mabuchi A, Sekine A, Saito S, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K (2003) An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. Nat Genet 35: 341-8

Traver RD, Horikoshi T, Danenberg KD, Stadlbauer TH, Danenberg PV, Ross D, Gibson NW (1992) NAD(P)H:quinone oxidoreductase gene expression in human colon carcinoma cells: characterization of a mutation which modulates DT-diaphorase activity and mitomycin sensitivity. Cancer Res 52: 797-802

Traver RD, Siegel D, Beall HD, Phillips RM, Gibson NW, Franklin WA, Ross D (1997) Characterization of a polymorphism in NAD(P)H: quinone oxidoreductase (DT-diaphorase). Br J Cancer 75: 69-75

Van de Water N, Tan T, Ashton F, O'Grady A, Day T, Browett P, Ockelford P, Harper P (2004) Mutations within the protein Z-dependent protease inhibitor gene are associated with venous thromboembolic disease: a new form of thrombophilia. Br J Haematol 127: 190-4

van Wijk R, Nieuwenhuis K, van den Berg M, Huizinga EG, van der Meijden BB,

Kraaijenhagen RJ, van Solinge WW (2001) Five novel mutations in the gene for human blood coagulation factor V associated with type I factor V deficiency. Blood 98: 358-67

Vecsler M, Loebstein R, Almog S, Kurnik D, Goldman B, Halkin H, Gak E (2006) Combined genetic profiles of components and regulators of the vitamin K-dependent gamma-carboxylation system affect individual sensitivity to warfarin. Thromb Haemost 95: 205-11

Versteeg HH, Ruf W (2007) Tissue factor coagulant function is enhanced by protein-disulfide isomerase independent of oxidoreductase activity. J Biol Chem 282: 25416-24

Vigano S, Mannucci PM, Solinas S, Bottasso B, Mariani G (1984) Decrease in protein C antigen and formation of an abnormal protein soon after starting oral anticoagulant therapy. Br J Haematol 57: 213-20

Visser LE, Trienekens PH, De Smet PA, Vulto AG, Hofman A, van Duijn CM, Stricker BH (2005) Patients with an ApoE epsilon4 allele require lower doses of coumarin anticoagulants. Pharmacogenet Genomics 15: 69-74

Vogel F (1959) Moderne probleme der Humangenetik. Ergeb Inn Med Kinderheild 12: 52-125

von Rosensteil NA, Adam D (1995) Macrolide antibacterials. Drug interactions of clinical significance. Drug Saf 13: 105-22

Wadelius M, Chen LY, Downes K, Ghori J, Hunt S, Eriksson N, Wallerman O, Melhus H, Wadelius C, Bentley D, Deloukas P (2005) Common VKORC1 and GGCX polymorphisms associated with warfarin dose. Pharmacogenomics J 5: 262-70

Wadelius M, Chen LY, Eriksson N, Bumpstead S, Ghori J, Wadelius C, Bentley D, McGinnis R, Deloukas P (2007) Association of warfarin dose with genes involved in its action and metabolism. Hum Genet 121: 23-34

Wadelius M, Pirmohamed M (2007) Pharmacogenetics of warfarin: current status and future challenges. Pharmacogenomics J 7: 99-111

Wadelius M, Sorlin K, Wallerman O, Karlsson J, Yue QY, Magnusson PK, Wadelius C, Melhus H (2004) Warfarin sensitivity related to CYP2C9, CYP3A5, ABCB1 (MDR1) and other factors. Pharmacogenomics J 4: 40-8

Wahls WP, Wallace LJ, Moore PD (1990) Hypervariable minisatellite DNA is a hotspot for homologous recombination in human cells. Cell 60: 95-103

Wajih N, Hutson SM, Wallin R (2007) Disulfide-dependent protein folding is linked to operation of the vitamin K cycle in the endoplasmic reticulum. A protein disulfide isomerase-VKORC1 redox enzyme complex appears to be responsible for vitamin K1 2,3-epoxide reduction. J Biol Chem 282: 2626-35

Wajih N, Sane DC, Hutson SM, Wallin R (2004) The inhibitory effect of calumenin on the vitamin K-dependent gamma-carboxylation system. Characterization of the system in normal and warfarin-resistant rats. J Biol Chem 279: 25276-83

Wajih N, Sane DC, Hutson SM, Wallin R (2005) Engineering of a recombinant vitamin K-dependent gamma-carboxylation system with enhanced gamma-carboxyglutamic acid forming capacity: evidence for a functional CXXC redox center in the system. J Biol

Chem 280: 10540-7

Walker FJ (1981) Regulation of activated protein C by protein S. The role of phospholipid in factor Va inactivation. J Biol Chem 256: 11128-31

Wallin R, Gebhardt O, Prydz H (1978) NAD(P)H dehydrogenase and its role in the vitamin K (2-methyl-3-phytyl-1,4-naphthaquinone)-dependent carboxylation reaction. Biochem J 169: 95-101

Wallin R, Hutson SM, Cain D, Sweatt A, Sane DC (2001) A molecular mechanism for genetic warfarin resistance in the rat. FASEB J 15: 2542-4

Wallin R, Sane DC, Hutson SM (2002) Vitamin K 2,3-epoxide reductase and the vitamin K-dependent gamma-carboxylation system. Thromb Res 108: 221-6

Walton R, Kimber M, Rockett K, Trafford C, Kwiatkowski D, Sirugo G (2005) Haplotype block structure of the cytochrome P450 CYP2C gene cluster on chromosome 10. Nat Genet 37: 915-6; author reply 916

Wang WY, Barratt BJ, Clayton DG, Todd JA (2005) Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet 6: 109-18

Watala C, Golanski J, Kardas P (2003) Multivariate relationships between international normalized ratio and vitamin K-dependent coagulation-derived parameters in normal healthy donors and oral anticoagulant therapy patients. Thromb J 1: 7

Weiss P, Soff GA, Halkin H, Seligsohn U (1987) Decline of proteins C and S and factors II, VII, IX and X during the initiation of warfarin therapy. Thromb Res 45: 783-90

Westwood IM, Kawamura A, Fullam E, Russell AJ, Davies SG, Sim E (2006) Structure and mechanism of arylamine N-acetyltransferases. Curr Top Med Chem 6: 1641-54

Whittaker P, Bumpstead S, Downes K, Ghori J, P D (2005) SNP Analysis by MALDI-TOF Mass Spectrometry. In: Celis J CN, Simons K, Small JV, Hunter T, Shotton D (ed) Cell Biology: A Laboratory Handbook., 3 edn. Elsevier, Amsterdam

Wittkowsky AK (2007) Dietary supplements, herbs and oral anticoagulants: the nature of the evidence. J Thromb Thrombolysis

Wu O, Sumii T, Asahi M, Sasamata M, Ostergaard L, Rosen BR, Lo EH, Dijkhuizen RM (2007) Infarct prediction and treatment assessment with MRI-based algorithms in experimental stroke models. J Cereb Blood Flow Metab 27: 196-204

Wu SM, Stafford DW, Frazier LD, Fu YY, High KA, Chu K, Sanchez-Vega B, Solera J (1997) Genomic sequence and transcription start site for the human gamma-glutamyl carboxylase. Blood 89: 4058-62

www.janusinfo.se Janusinfo Drug Interactions. Läkemedelscentrum in Stockholm's county council

www.thomsonhc.com Thomson Micromedex® Healthcare Series. Greenwood Village, Colorado

Wynne Jones D, Russell G, Allford SL, Burdon K, Hawkins GA, Bowden DW, Minaee S, Mumford AD (2004) Severe prekallikrein deficiency associated with homozygosity for an Arg94Stop nonsense mutation. Br J Haematol 127: 220-3

Yuan HY, Chen JJ, Lee MT, Wung JC, Chen YF, Charng MJ, Lu MJ, Hung CR, Wei CY, Chen CH, Wu JY, Chen YT (2005) A novel functional VKORC1 promoter polymorphism is associated with inter-individual and inter-ethnic differences in warfarin sensitivity. Hum Mol Genet 14: 1745-51

APPENDIX I PRIMERS FOR 216 TAG SNPS.

SNP ID	Forward	Reverse	Extension
rs17882687	ACGTTGGATGGCTCTGTCTCTCATGTTTGC	ACGTTGGATGAATCACTGGGAGAGGAGTAG	ttccTGTTTGCTTCTCCTTTCA
rs3753661	ACGTTGGATGGTTTCACAATCCAGAGAGGG	ACGTTGGATGTTCTTAGAACCGACTGCAGC	AATCCAGAGAGGGAGATAGA
rs6976017	ACGTTGGATGTGCCCACAGGGACATAATTG	ACGTTGGATGCTTTACAAACCACAGACTAG	GGGACTGTGGATGGATGTA
rs2275620	ACGTTGGATGCCCCAAGGTAAGCTTGTTTC	ACGTTGGATGTTAGAGGGTTGGAACCAAAC	TCTGTACTTCTGAAATTTCCA
rs1058932	ACGTTGGATGCTGAAGAATGCTAGCCCATC	ACGTTGGATGGATGAGAGGTCAGAGAAGAC	CTAGCCCATCTGGCTGC
rs4986910	ACGTTGGATGTTGGAAGTGGACCCAGAAAC	ACGTTGGATGTGAAGGACTCTGATTAGAGC	CAGAAACTGCATTGGCA
rs17878459	ACGTTGGATGCTTTCAGCCAGTGGGAAATG	ACGTTGGATGATGAAGTGGTGAAGGAAGCC	CCTCTTCCAGAAAACTC
rs4244285	ACGTTGGATGGCAATAATTTTCCCACTATC	ACGTTGGATGCACTTTCCATAAAAGCAAGG	CCCACTATCATTGATTATTTCCC
rs3093230	ACGTTGGATGTGTGGCCCTGTCACTAGGAT	ACGTTGGATGCCTGGGAACAGAGCTGTGAC	GAGGGTCCTTCAGCCCCTAC
rs10509681	ACGTTGGATGACTGACTTCCGTGCTACATG	ACGTTGGATGTATCTAGAAAGTGGCCAGGG	CGTGCTACATGATGACA
coding T26S	ACGTTGGATGCCATCTACTGGTTCATCTCC	ACGTTGGATGCCCACCACCCATCTTCAAGT	CGGGACAAAGAGGAAA
rs1687390	ACGTTGGATGAGGACACGTCCATGGAGACAC	ACGTTGGATGAACTAGGCAAGAGGCGGTTG	GCCCTTCCCTGCCCTGCC
rs11572080	ACGTTGGATGTTTCTCCCTCACAACCTTGC	ACGTTGGATGCAGTGAGCTTCCTCTTGAAC	TTTTGGGATGGGGAAGA
GS30310	ACGTTGGATGAAACAAAGTTTTAGCAAACG	ACGTTGGATGGGCGCATTATCTCTTACATC	TTGTGTCTTCTGTTCTCAAAG
rs17110453	ACGTTGGATGTACACTGATTTCCCTCAAGG	ACGTTGGATGCTGAAGTAAATGATTCTATG	ATTTCCCTCAAGGTCATAAA
rs6046	ACGTTGGATGTACTCGGATGGCAGCAAGGA	ACGTTGGATGTGACGATGCCCGTCAGGTAC	CCCACATGCCACCCACTACC
rs6025	ACGTTGGATGCTGAAAGGTTACTTCAAGGAC	ACGTTGGATGCTCTGGGCTAATAGGACTAC	GGACAAAATACCTGTATTCCT
rs2470890	ACGTTGGATGTCTACGGGCTGACCATGAAG	ACGTTGGATGTGGCCTCAGAATGGTGGTGT	CTGCGCTTCTCCATCAA
rs12721607	ACGTTGGATGCAGAGTCTGTTCCTGGAAAG	ACGTTGGATGCACATACACGGCAGATTTGG	CAACGCAGATGAGGAAGTC
rs2917671	ACGTTGGATGTTTTGCACTGGAGGGACAAC	ACGTTGGATGGAACTAGAGCTTTAAAGTGG	TCACCTCCCCCATCTGT
rs1058930	ACGTTGGATGGCTAATATCTTACCTGCTCC	ACGTTGGATGAGAACACCAAGCATCACTGG	TTTGATCAGGAAGCAATC
rs2234922	ACGTTGGATGACTTCATCCACGTGAAGCCC	ACGTTGGATGAAACTCGTAGAAAGAGCCGG	CCCCCAGCTGCCCGCAGGCC
rs28399504	ACGTTGGATGAGTGCAAGCTCACGGTTGTC	ACGTTGGATGCATGAGAGACAGAGCACAAG	AAGAGGAGAAGGCTTCA
rs1557044	ACGTTGGATGGCTCAGGAGAAGAAACAAGG	ACGTTGGATGTCACTTTCTCTCTTTAGAGC	AACAAGGAGCAGAGCAAGG
rs28365094	ACGTTGGATGGTATGAGTTATTCTCTGGAGC	ACGTTGGATGACTCTCAACTGAGTCCATGC	TCTGGAGCTTCTAATACTTCA
rs3814637	ACGTTGGATGCGACAATACTTACACAAAGCC	ACGTTGGATGAGAGAACTGGAAATAACCTC	GCCATTTTCTTTAAAGATATCATC

SNP ID	Forward	Reverse	Extension
rs4417205	ACGTTGGATGTGGTAAGTATACAATGTGAG	ACGTTGGATGTCCTAGGAATGATTTGATGC	GAGTAATTTTGAATTTACTGTCAT
g-3844G>A	ACGTTGGATGTCAAGTAGGTGTTCACGTGG	ACGTTGGATGCCCTCTGTGTCGACATTTTC	AACAGCCCGGCCTGTGT
rs776746	ACGTTGGATGGTAATGTGGTCCAAACAGGG	ACGTTGGATGACCCAGCTTAACGAATGCTC	GGTCCAAACAGGGAAGAGATA
rs762551	ACGTTGGATGTCTGTGATGCTCAAAGGGTG	ACGTTGGATGCAGCTGGATACCAGAAAGAC	AAGGGTGAGCTCTGTGGGC
rs2282687	ACGTTGGATGCTGCTGTATACCCTAGAACG	ACGTTGGATGGGGAAATGCCATTCATTCAG	ACGTGATAGGCGCTCAATAA
rs5898	ACGTTGGATGATCAGTGACCGCTGGGTCCT	ACGTTGGATGGAAGGTCATTCTCGGTGAAG	CACTGCCTCCTGTACCC
rs3136516	ACGTTGGATGCAAGTTCAAGGTCACATCAG	ACGTTGGATGCCTGGTGAACACATCTTCTG	CAAGGTCACATCAGTATTCC
rs28365083	ACGTTGGATGAGGCTCTGTCCAGTACTTTG	ACGTTGGATGCATTCCCAAAGGGTCAATGG	GGTCATGGTGAAGAGCATAA
rs10260862	ACGTTGGATGCCAAATCAATCTTGGGAAAGC	ACGTTGGATGGGTAGTGACATACTGGCAAC	cATCTTGGGAAAGCTTACATTAT
rs10267099	ACGTTGGATGGGTTCTGCCTGAAGGAATTG	ACGTTGGATGGCATTTATTTCCCTTAAGAG	AAGATGAAGATTGAAAACACAAATG
rs1045642	ACGTTGGATGGCTGAGAACATTGCCTATGG	ACGTTGGATGAAGGCATGTATGTTGGCCTC	GTGTCACAGGAAGAGAT
rs1048943	ACGTTGGATGTATCTTTGGCATGGGCAAGC	ACGTTGGATGGGATAGCCAGGAAGAGAAAG	AAGTGTATCGGTGAGACC
rs1051740	ACGTTGGATGTGGAAGAAGCAGGTGGAGAT	ACGTTGGATGCTGGCGTTTTGCAAACATAC	CAGGTGGAGATTCTCAACAGA
rs1051741	ACGTTGGATGTGGTGCCTGTTGTCCAGTAG	ACGTTGGATGGGCTTTGTGTTCTGCGTTCC	ggGTCCAGTAGAGCATGAC
rs1054191	ACGTTGGATGAGCAGCACAAGGAATTTCCC	ACGTTGGATGGGGAACCTTCACTTGGGTAC	GCTGAGCTGTGATGGC
rs1057910	ACGTTGGATGATGTCACAGGTCACTGCATG	ACGTTGGATGCACATGCCCTACACAGATGC	GTGGGGAGAAGGTCAA
rs10982151	ACGTTGGATGAATCTCCACCAGACTCTTGC	ACGTTGGATGCTAAAGCCAAACAGAAGGCG	accCTTCTCAATAATCTTCCTGTT
rs109829	ACGTTGGATGGAAAAATAAAGTTCAGCCAC	ACGTTGGATGCTCCACTCTTTGCATCAAAC	AGTTCAGCCACATCTTC
rs11150606	ACGTTGGATGCTGCTGTCAGTGTGACTAAC	ACGTTGGATGTGTTGCCCTCCTGAGGCTTG	CCTTTCCACAGCCTGTGGAC
rs11653	ACGTTGGATGCATGTAGACAAACATTAGCTC	ACGTTGGATGCTTATTGATCTCAGGCTTAC	CTCTTTCTCAACCCCTT
rs11773597	ACGTTGGATGTTCTAGTTGCCCACTGTGTG	ACGTTGGATGCTCAAGGGCATAGTCTAGTC	CCACTGTGTGTACAGCAC
rs1202172	ACGTTGGATGCTGTTGCTGCAATAGAACCC	ACGTTGGATGCTGGTCCACAGTAAAACAGC	GCAATAGAACCCTTGTAAGTTTTCTC
rs12248560	ACGTTGGATGCAAATTTGTGTCTTCTGTTC	ACGTTGGATGAGGTCTTCTGATGCCCATCG	aAAATTTGTGTCTTCTGTTCTCAAAG
rs1464602	ACGTTGGATGACCACTGACCCACTGGGTAA	ACGTTGGATGATTACAAGGCCTTTGGGTGG	ctGGCCTCAGCTTGACCT
rs1536430	ACGTTGGATGGGAAGCCCATTTATGATAAGC	ACGTTGGATGCAAAATATCCTACCACAAAC	CATAATCTATGTGCAATATTGATAT
rs1557572	ACGTTGGATGCACAGAACACAAGGGAGTGG	ACGTTGGATGTCCATTCAAGGCTCACAACC	ctttGGTAAAGCAACTCCGA
rs17149866	ACGTTGGATGTGTCTGTGACTAGTTTAGCC	ACGTTGGATGACAAACATGATCCTCAACTC	atTTTAGCCATAAGACTTAAACA
rs17230081	ACGTTGGATGAATCTCCACCAGACTCTTGC	ACGTTGGATGTAAAGTCAACCAGAAGGCGG	ATTGGCCACTTCTCCT

SNP ID	Forward	Reverse	Extension
rs1799809	ACGTTGGATGTAGCTCAGCACGGCTTGTTC	ACGTTGGATGCATCTGTCAAGGGTTTTGCC	tCGAAGCCCACCTCTGCC
rs1799853	ACGTTGGATGCTGCGGAATTTTGGGATGGG	ACGTTGGATGACCCACCCTTGGTTTTTCTC	ttGGGAAGAGGAGCATTGAGGAC
rs1800566	ACGTTGGATGTCCAGGATTTGAATTCGGGC	ACGTTGGATGGCATTTCTGTGGCTTCCAAG	CAATGCTATATGTCAGTTGAG
rs1856908	ACGTTGGATGTTTCTCAGGCAGATCACTAC	ACGTTGGATGAGGTTCGAATGCTGGAGTAG	ccTTTATCCTTCAATAAGGAGAGTTTC
rs1858923	ACGTTGGATGTGCTGATTTTCCTCCAGCTC	ACGTTGGATGCTCTGACTCCTGTGATAAGG	ttccTTTTCTTGTGCTGCCC
rs1882478	ACGTTGGATGATCTTGGGAAAGAGAGCCAC	ACGTTGGATGAAACAAAGCTCAGGAGCCTC	TGAGGAAACTGTCTTCCC
rs1894699	ACGTTGGATGCTCCATCCTTGGCTTTTAGG	ACGTTGGATGCCTTTGTGTGGGTAAGGAAC	AATCACTAGATACTAGATAATGGG
rs1922240	ACGTTGGATGAGAGTGGCTAGGATGTGTTC	ACGTTGGATGAGGGTCAATGTATGAGCAGC	TTCAGTCCTGTGATATACA
rs1922242	ACGTTGGATGTGATAAGGAATAAGGATAGG	ACGTTGGATGGTACAATTCTTACATACGCAC	ATAAGGATAGGATATATTCCTTTAC
rs2026160	ACGTTGGATGGTTCCTGAATTTCCTTTCTGC	ACGTTGGATGACACCCTGAGGGAAAAAGTC	CCTTTCTGCTTTTGTTCT
rs2028898	ACGTTGGATGCCTTCTAAGGGCTTTAGCTG	ACGTTGGATGGTGAGATACAAGTCATCAGG	ACAGGAGTTTGAACTTGGT
rs2056530	ACGTTGGATGAGGCATCACTTGCCTGCATC	ACGTTGGATGGAACTGGCCAGAGGATAAAG	ctTATTCTTATATTACATTCCCCTTAC
rs2060717	ACGTTGGATGCATTTTTGTGAGCTCATGGC	ACGTTGGATGGAGGTTTTACTGTGATTGGG	GCTCATGGCTTTCCTA
rs2069522	ACGTTGGATGAGAAAACTGTGGGATCAACC	ACGTTGGATGCCCATTCATGGCCTTCAAAC	gcatATGGATGGGGAATCCAATAGAG
rs2069525	ACGTTGGATGTGCTGTAGCATGGAAGTGTC	ACGTTGGATGCTGGCCTAGAGGTGTACATT	cagGGAAGTGTCATGATCCCC
rs2069901	ACGTTGGATGACCAATCATAAGGAGGCAGG	ACGTTGGATGAGAAGGAAGATGCAGGTGTG	TGCTCCCTGGGACTCTC
rs2069910	ACGTTGGATGTCCCCTACTCAAATGCACAC	ACGTTGGATGCCAAGAATCATGGCCTCCTC	AATGCACACTGGCCTCA
rs2069919	ACGTTGGATGTACCAAGCTCACCCTACTAC	ACGTTGGATGTTGTCTAAGGGTCAGAGTCC	CTACTACCTAGGGCCA
rs2069928	ACGTTGGATGTCAGGTTGCGTCCATCTTTC	ACGTTGGATGTATGCTCAGGGTGCAGAAAC	CTTCATCATCCCCCAAA
rs2069933	ACGTTGGATGACTCCTGAAAACCAACCAGC	ACGTTGGATGTCAACGGTGACTTTTCCTGC	CCAACCAGCATCCTACC
rs2070851	ACGTTGGATGTGATGTGTACTGAGCACCCG	ACGTTGGATGAGTGGCCTCCAAATATTCGC	tCACCCGACAGTGCCTGTCA
rs2070852	ACGTTGGATGAAGCGTACCTCAAGCCCAAC	ACGTTGGATGAAATTCTCCTGTAGGTCGGC	CTGTTGGGCAATTTCCT
rs2102663	ACGTTGGATGGCACCAGTGGATGATCAAAG	ACGTTGGATGAAAGCTTAGGGTGCAGATAC	ggttGTCTTAGGACCACATGAG
rs2214101	ACGTTGGATGATTCCCCCACATCCTAAAAC	ACGTTGGATGCCTACTGTAAAAAAATATGG	ggCCCACATCCTAAAACATACTTA
rs2214102	ACGTTGGATGTTCTTTGCTCCTCCATTGCG	ACGTTGGATGCATTTGGCTAATGAGCTGCG	ccctCGGTCCCCTTCAAGATCCAT
rs2227590	ACGTTGGATGTGCAGGAAGTCAGCACTCAC	ACGTTGGATGTAGCATTTGCTCTCTTCCCG	GGGGAATCCCCAGGGCCTGC
rs2227607	ACGTTGGATGAGTTAGAACCCCTGCAATAG	ACGTTGGATGGTTTTCTGGGCAGTCACTAC	TGAGAAATCAAAGGTATCCAT
rs2235033	ACGTTGGATGCTGCCTCATGTAAGTTGTCC	ACGTTGGATGTAGTTTCCTAACTTCCTGC	GTCCTTGCCCTTTGCC

SNP ID	Forward	Reverse	Extension
rs2235040	ACGTTGGATGGATGCTGCTCAAGTTAAAGG	ACGTTGGATGTTAGTTTCATGCTGGGGTCC	cacGCCTCCTTTCTACTGGT
rs2235046	ACGTTGGATGTGTGGTTCCCTAGTTTGGTG	ACGTTGGATGCAGGAGGATTCTGGATAACC	TGGTGGGCTAGGGCTAC
rs2242480	ACGTTGGATGCTAAGGTTTCACCTCCTCCC	ACGTTGGATGCTGCAGGAGGAAATTGATGC	cCCTCCCTCCTTCTCCATGTA
rs2251102	ACGTTGGATGGCTTGAGTCACTTAATTATGG	ACGTTGGATGCCAGAATTCATTCTAAAAGG	TCTGTTTTCCCTAATATATTTTTAAAT
rs2260863	ACGTTGGATGGACAGATTTGTTGTGACTGC	ACGTTGGATGGTTGGAGAGATTCAGAACCC	TACAACGTATAAGTACATCTCAGT
rs2273971	ACGTTGGATGTGGGACGCAGCCTGCCATT	ACGTTGGATGTTTGGCTGTGTTTGTAGCCC	CCATTCCCACCCGGAG
rs2276706	ACGTTGGATGTGCTGCTGTCTCCTCATTTC	ACGTTGGATGCAGAGAGCATCAGTAATGGG	gaTGCTGTCTCCTCATTTCTAGGGTG
rs2290228	ACGTTGGATGCACAGGGACAGGCACATAAG	ACGTTGGATGCTGCCTCCTGAATTAACTGC	AGGCACATAAGAAACTGT
rs2292566	ACGTTGGATGTGACATACATCCCTCTCTGG	ACGTTGGATGGAAGCAGGTGGAGATTCTCA	GTTTTGCAAACATACCTTCAAT
rs2292567	ACGTTGGATGGCTGCTAGAGGTTCCATAAC	ACGTTGGATGTCGGGCCAACTCCCTGCTCA	AGGTTCCATAACTGCCCC
rs2298905	ACGTTGGATGCTGCAACTGTAACTAATCAC	ACGTTGGATGGATGCATGCTAGTTACAGAG	CAACTGTAACTAATCACTTAAAA
rs2298909	ACGTTGGATGAGCCTTTGGATCATCCTTTG	ACGTTGGATGGGACAGCAGAGAATTAAAGG	CATCCTTTGTCTGTAGATTA
rs230704	ACGTTGGATGCTGCCTCTGTACAATGTCTG	ACGTTGGATGGGACTGTGAAAAGGGAGTAG	TGCTAGGGCTGCCCTCCTG
rs2307040	ACGTTGGATGCCTCCGTCACAAACCCATCT	ACGTTGGATGCTGGCACCTTGAAACGTAAC	CCGTCACAAACCCATCTTTATCC
rs2307418	ACGTTGGATGATGTTTCACCAACCCCTTCC	ACGTTGGATGAAGCTACATCAAGGGCCAGC	atgtTGTGGCCTCCAAGCCC
rs2307420	ACGTTGGATGGTAGTCAGTGACTTTTCGGG	ACGTTGGATGACCTCCCCTCATCTTTCAGG	TCGGGGTGGATATACAATTTAC
rs2307424	ACGTTGGATGACTGAAGTGTTTGCCTCCTG	ACGTTGGATGTGGTACTGCAAGTCATCAAG	GAGGTCACTCACCGGAAGAC
rs2359612	ACGTTGGATGAGTCTGAACCATGTGTCAGC	ACGTTGGATGTTTGAGTCACCCTTCCCAGC	ACCATGTGTCAGCCAGGACC
rs2461818	ACGTTGGATGCTTAGTTACTGCATCCATCC	ACGTTGGATGTCGGAATAAACAATGACTTC	CCATCCATATTTTAGTTGTATTTTG
rs2470893	ACGTTGGATGTGTTCCCTTCTCTGTCAATC	ACGTTGGATGATATGCGGCCTCGTGCATTG	TGTTCCCTTCTCTGTCAATCGCCAGC
rs2472677	ACGTTGGATGGGAAGACTTATTCTATTCCTG	ACGTTGGATGTTTATCAACTTTTTTGTGCC	cctgTGTGTTTGTTTGTTTTTTAATCA
rs2472682	ACGTTGGATGATAAGCTTAAAGGGGCAGGG	ACGTTGGATGGTCAGTCACTGGATTTCACC	GGGGCAGGGAGAGAAGAATACTTATA
rs2475376	ACGTTGGATGGAAGTACTGATACATGTTGC	ACGTTGGATGAACCTGTTATGTCTGAGTCC	TGCAACTTAGATGAGCCTC
rs2480948	ACGTTGGATGTCACATCCTCGGCCAGAGTC	ACGTTGGATGAGCCAAATCGCGGTTCACAC	aTCCTGAGGTGTAGCCA
rs2502804	ACGTTGGATGACCAACTCATTCTGGTAACC	ACGTTGGATGAAGTGCAGGCTGGAGTCATA	TTAACTCAGTCTCTCTCTCT
rs2606345	ACGTTGGATGCCTTTCCTATCTCATTGACC	ACGTTGGATGTAGGGACTGGTATTTCCAGC	CTGGGAGACAATCAGG
rs2671270	ACGTTGGATGAGCTGACATGAACTTCCTCC	ACGTTGGATGAGAGGGTGCAAGTCTCATTG	aggtGGCTGGCCTTTTAGCCT
rs2671272	ACGTTGGATGTCTGACTCAGGCATAAGGAC	ACGTTGGATGCATCTTCCTCATTCCTAGGG	tTCTGAACAAGAACAGTCT

SNP ID	Forward	Reverse	Extension
rs2740170	ACGTTGGATGGAAAACATGGTGTTGTCTGTC	ACGTTGGATGGGGTGTGAGGTATATGGGAG	ggtcGATTGCTGAAAAACTGCAAAGA
rs2759328	ACGTTGGATGGTTTCAAAAAGCCCCAAAGG	ACGTTGGATGGTGAGATGGGAGAAAGTTGG	CTGAATTCCCATCTGTGG
rs2774030	ACGTTGGATGTCTAGAAACCAGCATCCAGG	ACGTTGGATGACCACCCACTCCTAAAGTTC	GGGAGGCCCTTCTTGGT
rs2787337	ACGTTGGATGTGCTGGGATTACTCAGAAAG	ACGTTGGATGCTACAAACTACACTGAAATATG	GGGATTACTCAGAAAGTATTCTTAC
rs2854461	ACGTTGGATGTAGAGAGTGCCAGGCACCCA	ACGTTGGATGTCGAGCCTTTGCTTTCCCCT	gaTCCCTCGCCAGGGCTGCGGCTTTAT
rs2860840	ACGTTGGATGATTCCTGTCTGAAGAAGGGC	ACGTTGGATGCCCTGATAAGGGAGAATTGC	gcTAGTTTGGCTGCTCCTGTG
rs2860905	ACGTTGGATGGAACATGGGATTAAATGAACC	ACGTTGGATGAGAGCCAAGGGAATTTGCAC	CCTTTTATACCCACACTGTA
rs2901783	ACGTTGGATGATCACGGCAGAAGGCAAAAC	ACGTTGGATGAAGTGTGTAGCACCTCCCAC	gCACTCACATGGCAAAAG
rs3003596	ACGTTGGATGTTGACTTCTGCAAAAGATCC	ACGTTGGATGGGAAGAATGAAAGGAAACTG	GCAAAAGATCCAAGATCA
rs3024718	ACGTTGGATGAGAGAACATGCTTTGGGACC	ACGTTGGATGTCTGCGATTTCCAGGAGGAG	cccaACTGCCGCGGGTCAACTC
rs3024746	ACGTTGGATGCACTTAACACTTCCAGAGCC	ACGTTGGATGGGGATACAAAAGGAGAAAACG	ccACACTTCCAGAGCCAGTACCCGT
rs3093229	ACGTTGGATGCTTGCCTTGAGATGACAACC	ACGTTGGATGTCATTCCCGCCTGCAAGAAG	gCAACCAAAGTTTTCCTGTGTCCTC
rs3093261	ACGTTGGATGCACATTGTGCGTCATTGTGC	ACGTTGGATGTTGGCATCCAGCATCATCAG	gCTCACTGAGCCCCACCC
rs3136435	ACGTTGGATGAAACCCACCCCTGAGCTCTT	ACGTTGGATGATGTCTGGAAGGTGAGCAAC	ccTCTCCCAGGCTCCCCTC
rs3211764	ACGTTGGATGACATTCAGACAAGGGTCACG	ACGTTGGATGTCTAGGCTCTTGATGACCTG	ggggCACGTGCTTGTCAATAGT
rs3212998	ACGTTGGATGTCACTGTCTCTGGTCGATCC	ACGTTGGATGTGTGGCCGCTATCACAGAAC	CTTGCGCCTCCTCTGACCTC
rs3213005	ACGTTGGATGACACCAGGAGACAAGGCTAA	ACGTTGGATGAAGAGCAGATGGAGGCCTTG	ctACAAGGCTAAAGCCAG
rs339042	ACGTTGGATGGTGCATCTGTTGAAATGCTC	ACGTTGGATGGGAAGAAGAGGAACCGGATG	TGTTGAAATGCTCAAGACTT
rs339043	ACGTTGGATGTTATAGGCGTGAGCTATCGG	ACGTTGGATGGTCATCCCTCATAGAGACAA	GAGCTATCGGCACCCAGCC
rs339051	ACGTTGGATGTGAACAACACTCTGGCTGAC	ACGTTGGATGATAGAGCCGCTAAATCTCGG	CTGGCTGACCCTCCAAT
rs339053	ACGTTGGATGAAGTGTTGAAGTGGGTGAGG	ACGTTGGATGGCATCTTTGTCTTTTCAACC	AGGGCCATTAACTTTGT
rs339054	ACGTTGGATGACTACCACACCCATTTCTCC	ACGTTGGATGACGTTAACACTGAGTGATCC	GGGTCTTTGGTCAACCA
rs339056	ACGTTGGATGACTTACCTTTCCTCGTAGCC	ACGTTGGATGTTCCGGTTGGGCGGTGCTTG	TTCCTCGTAGCCCCCACAGC
rs339057	ACGTTGGATGCTAGGAGAATGAGCTAGGTG	ACGTTGGATGTCCGCCATCACATAGACTTC	CATGGAATGAGGGAAGTCAA
rs339096	ACGTTGGATGTGAGTGGAGATCACGCCACTG	ACGTTGGATGGCAGACGTACACAACTTATT	CCTGGGTGACAGAGCAA
rs339097	ACGTTGGATGTGGATTCTGAATCTGGCCAA	ACGTTGGATGCTTTCCCCTTTAGCCTTATA	TTCTGAATCTGGCCAATACTTA
rs339098	ACGTTGGATGTAGCGGATTCATTGAGACAG	ACGTTGGATGGGAGTTTTTGTTACTGTGTG	GATGATTTATACCTGAACTCCAAA
rs3732356	ACGTTGGATGTCTGAGTAAGGACGTGCCGT	ACGTTGGATGTCACATGTGCACGTGTGTTC	ACGTGCCGTGGGTGTG

SNP ID	Forward	Reverse	Extension
rs3732357	ACGTTGGATGTTCAAACCTGAACCTGCAAC	ACGTTGGATGTTCAGGTCACCTCAGATCTC	aaAACCTGCAACAGAATCAC
rs3732359	ACGTTGGATGTCCCTCAGATCCCACTAAAG	ACGTTGGATGAACAAACGTGGGTATGTGGG	ACCAAGCGACCAAGGAT
rs3753305	ACGTTGGATGGACAGTTTGTCTGGGTTGTG	ACGTTGGATGTCCAAAGTTCTGGAGATTCG	CTATGGTTTTGACTCAACAATT
rs3753660	ACGTTGGATGCCAGCACTACCAGGTGTATC	ACGTTGGATGAGAATGCAAGAGATCGTGTT	tAAGAGGAGGGCAAAGTG
rs3753663	ACGTTGGATGTTAGAACGCTGCCCTGGGAC	ACGTTGGATGAGCCTGGGATTGGGAGGAAA	ccCCCCTGGGATGCTCAACATA
rs3766110	ACGTTGGATGGGAAAGGCATGCATGAAGAC	ACGTTGGATGAAATGCCGGAACCCTCATTC	GACATTTCAAGTGATCAGTAAGAT
rs3814057	ACGTTGGATGAGGGCTACATTTCCCAAAAC	ACGTTGGATGTATAGCCACTTGTGAGTAAA	gtGCTACATTTCCCAAAACTAGTTC
rs3817268	ACGTTGGATGGAAAACACAGCAAACCCTGC	ACGTTGGATGTAGCCAGTGATGTGGGAAAC	AGCCACAAGATACTGC
rs3817939	ACGTTGGATGTAGAAATATCTGATGCTGTC	ACGTTGGATGTGCTGGCTGTTAGACTCTTC	ctctAAATATCTGATGCTGTCTTCTTC
rs3842	ACGTTGGATGGGAACAGAGTGAGAGACATC	ACGTTGGATGTAAAATCTACTTTAATTCTG	gggCATCAAGTGGAGAGAAATC
rs392959	ACGTTGGATGCTCTTACTCCTTTGTTTCTC	ACGTTGGATGTGGTAGGCACTTTGAACTAG	TCTTACAGATCAAGCTCC
rs401597	ACGTTGGATGGTGTATGCATCAAAGATGTCC	ACGTTGGATGTAGTACCCTGACACAGTACC	tcCAGACATTACTGAGTTACAACTA
rs4073054	ACGTTGGATGTTACTGTCCTTTCCTTAGGG	ACGTTGGATGCTGAAACGATGTGAGACAGG	gTTCCTTAGGGAATTCAGGTATC
rs413536	ACGTTGGATGAGGAAACACACTGGCTTGAC	ACGTTGGATGTTCATCATGTAGGTACAGGG	gTGACTCCTTGAGATTAATAGTTA
rs4148737	ACGTTGGATGAGTGAATGACCACCACTCTG	ACGTTGGATGTGGACATTTCAAACTGTCCC	CCACATCAGGTTTTCCCCAG
rs4149223	ACGTTGGATGCCTAGATATTGAGTCCTGCC	ACGTTGGATGCCAGTGTCAGCCCTTATTAC	gCTGCCACCCTTAGTGCCCCCGCC
rs4233368	ACGTTGGATGCTCTACCCTCTCTCATTCAG	ACGTTGGATGGGAGGAAGAGGAATTGTGTG	AGAAAAGACACAGAGAATCA
rs4244284	ACGTTGGATGAGGGAGCATGAACCAAATGG	ACGTTGGATGACCCTAGACAAGTCAGTGAG	GTGCTTTTATTTAATTGGACT
rs429358	ACGTTGGATGGCTGGGCGCGGACATGGAG	ACGTTGGATGTCGGTGCTCTGGCCGAGCAT	atCGGACATGGAGGACGTG
rs4646421	ACGTTGGATGTCCTTCATTGATCTGACCAC	ACGTTGGATGAGACTCCTTAGGGACACTTC	ggggCCACTCTTCAAAAGGAGGTA
rs4646425	ACGTTGGATGTCTGGTGTCACGTTGCTTCC	ACGTTGGATGTTCCTCCCAATAACACCAGG	CCCTGTGTTCACACTAA
rs4646453	ACGTTGGATGGTGTGTTGTTCTGCTATGTG	ACGTTGGATGACACTAAGAGGGAGGGCCTTG	caacAATTCTCATCTTCCTGGAATA
rs4646457	ACGTTGGATGTGCATTCCATCTTCACCTCC	ACGTTGGATGATGCTGCCCAATCAACTGAC	TGCAATGAACACTGAATAAAAAAT
rs4653436	ACGTTGGATGGCCATTGAAAGGTTTACGGG	ACGTTGGATGAGAACAACAGCCTTGCTCTG	GGGCAAAAGAATGGGTA
rs4656687	ACGTTGGATGGACCCACGATTATAACTTGC	ACGTTGGATGGAATCAGGGTTCTTTCTGGG	CATTTTTGCTAACCGCTCAC
rs4857037	ACGTTGGATGCAAGTGAGTTTCTTTTGGGTC	ACGTTGGATGCCTTAGGAGTCACAGTGTTC	cGTTTCTTTTGGGTCAGTATA
rs4857343	ACGTTGGATGCTCAAAGTAGAGAGATCAGTG	ACGTTGGATGGGTAAAACAAGAATGGCAGAG	gtTCCCCCACCCCTTTAATAAT
rs4917639	ACGTTGGATGGAAAGCAGCACATCAAAGAG	ACGTTGGATGGGATTACTTTCACCTTTTGAC	tgggGCAGCACATCAAAGAGATATTT

SNP ID	Forward	Reverse	Extension
rs5013930	ACGTTGGATGGGAATAAACCTGAACAACTC	ACGTTGGATGTAAAAATCAGAGGGCTCTAA	AACTCTATTTTATCTATGTGTTTTT
rs563964	ACGTTGGATGCCTGGTGACTGATGAGGATG	ACGTTGGATGCTGGGAACACATCCACTCTG	aTGACTGATGAGGATGAGGTTC
rs5878	ACGTTGGATGATCTCCTCCAATTCATCCAG	ACGTTGGATGTTGCCCAAGCCTGAGAAGAG	TCCAATTCATCCAGCCACTC
rs5936	ACGTTGGATGGCAGAGGTGAGCTTCCTCAA	ACGTTGGATGACCTCCTCTAGGCAGTAATG	AGCTTCCTCAATTGCTC
rs5960	ACGTTGGATGTCTCTTGGCTTGGTAGAGAC	ACGTTGGATGGAAAACGAGGGTTTCTGTGG	aaGATGTAGAACTCGCTCAGAAT
rs6018	ACGTTGGATGGCTCTGCTGTGGAAGAATTG	ACGTTGGATGTCACCAACAAGCCACCACAG	TGTGGAAGAATTGAGAACTGAG
rs6029	ACGTTGGATGATAGGTGTATTCTCGGCCTG	ACGTTGGATGCAGGTGCTTCTTACCTTGAC	CAGCGTCGTCCATCTTCTC
rs6035	ACGTTGGATGCTGGGCTCTGATAATAGGAC	ACGTTGGATGACGAAGATGAGTCCTTCACC	CCCAAAATCCCATCTTC
rs6037	ACGTTGGATGATCCATTGTGACCGTCACAG	ACGTTGGATGCTATGGAAAGAGGCATGAGG	TCCACGCATGGGGAAGAG
rs6041	ACGTTGGATGTGCTCGCCTGGAAGGAAGAA	ACGTTGGATGAGACCTAGAAATGGCCACAG	GGAAGGAAGAAGCCCCC
rs6048	ACGTTGGATGCCAAAATGGTTTCAGCTTCAG	ACGTTGGATGCAAACTTCTAAGCTCACCCG	CCACATCAGGAAAAACAG
rs6426089	ACGTTGGATGTTGTCACAGATGGTGCTGAG	ACGTTGGATGAAAGGAAGGACCTTCCACAG	ttccCCCTGGCTGTCTTGCTGAG
rs6427198	ACGTTGGATGGAATTGTTGACAGAACTGAG	ACGTTGGATGGTGGAAAGCACTGACTGTTG	GTTGACAGAACTGAGATAGGAA
rs6602908	ACGTTGGATGCTCAGGAGGGTGTTGGCAAG	ACGTTGGATGCACCTCTCTCCCACACTTGC	CCACAAAAGGTGTGGC
rs6686001	ACGTTGGATGGGGAATCTGGTGAAATGGAC	ACGTTGGATGGCTTGAGATGTGTTGGATAC	aTGTGCCCAAAGGTCCCCAGG
rs689453	ACGTTGGATGAAGAGCACTGATCGTACTGG	ACGTTGGATGCCTTTCTTCTTCAAAGCCGC	TTCAACTATGCCATGAAGGA
rs693335	ACGTTGGATGTGTCCTTTTGCTGCGCAGGC	ACGTTGGATGCTGCACAGCAGAAGCACTAG	gtCGCAGTGCCGGGTTCGC
rs721161	ACGTTGGATGTTGAAGACAGCCTTACAGGG	ACGTTGGATGCAAGGAGTCTATGTAATTGGG	TGTGGCTTTATTTTCTTTGTCC
rs7294	ACGTTGGATGTTCTAGATTACCCCCTCCTC	ACGTTGGATGAAAAAAGAGCGAGCGTGTGG	CTCCTCCTGCCATACCC
rs7412	ACGTTGGATGTCCTCCGCGATGCCGATGAC	ACGTTGGATGCTCGCGGATGGCGCTGAGG	CGATGACCTGCAGAAG
rs7530560	ACGTTGGATGTTCTCAAAACCCTCTCCTCC	ACGTTGGATGTGCCATAATGTAGGTGAGAG	tTACATCCTTCCACTATCTTAGCTCC
rs753057	ACGTTGGATGACGGATGGACAGAGACAAAC	ACGTTGGATGAATTTTGCAGGGAGTCGCAC	GCGAAATGGGATGAATGCA
rs7568458	ACGTTGGATGTTGAGGCAACCTCATTGAGC	ACGTTGGATGGGCTCCACCTCAAATCAAAG	tTCTGAGCTGTTGGTGC
rs7643645	ACGTTGGATGTGGGCAAGATCACAACATGG	ACGTTGGATGACCATGCTTAGCTACAGCTC	CAACATGGGAAGAAAAATGGCC
rs776905	ACGTTGGATGAACCAGCAGATCAAGACTCA	ACGTTGGATGCAGTACTTAGTACAAAGTTTTC	cACCAGCAGATCAAGACTCATTTGCC
rs7997328	ACGTTGGATGTCCTGAGAGTGCTGTTCTCG	ACGTTGGATGGTAAAATGGCTCAGTGACAG	ggGTGTGAAGTTGTTACATCCTCA
rs8178592	ACGTTGGATGTGGCTGAGCTTAGTTAGCAG	ACGTTGGATGTGTGATCACTACACACACCC	TGAGCTTAGTTAGCAGCTCTTAC
rs8178607	ACGTTGGATGTTGTTTGTTCTTGGTCAGTG	ACGTTGGATGATCCTTGCCAACATCTGAGC	actcGGTCAGTGATCAATGAAGAT

SNP ID	Forward	Reverse	Extension
rs8178610	ACGTTGGATGCTGCCTATGAATTGTATGAC	ACGTTGGATGCAAACCCGCTCTCTGAATTG	GTATGATCGAACTAAGAAAAAAATG
rs8178633	ACGTTGGATGATCCAAGTTCTCTCCTCCAG	ACGTTGGATGGAGTCGAACAAGAAATTCAC	ctcCTTTTTGTTTGCTTGTTTTG
rs8597	ACGTTGGATGCTGTTCGCATTGCAGAATATA	ACGTTGGATGCCCATATTGTGTGTTCTGTG	AATAACCAATATCCAAATTCAAGA
rs9282564	ACGTTGGATGACTCAAATCTCGCAACTATG	ACGTTGGATGGAGGAGCAAAGAAGAAGAAC	agTGAAACAAGCTAGTTACCTTTTAT
rs9332197	ACGTTGGATGCATATACCCCTGAATTGCTAC	ACGTTGGATGCTCTTCCTGGACTTTAGCTG	AATGTGCCATTTTTCTCC
rs9332504	ACGTTGGATGTTCTCTTACCACTCCTGCTC	ACGTTGGATGAGTGAAGCCTTGGTTTTGGG	ATTCTTCCCCAGGGTTC
rs9332618	ACGTTGGATGGTAGTTTTCTCCAGAAATACC	ACGTTGGATGCTCCACTACCTATCACTCTC	gcAATACCAAAAGATATTTGTTCTTAG
rs947173	ACGTTGGATGTTCCCCTGGGTAATATCAGC	ACGTTGGATGGAAAGAGCCACCTACGCTG	ATCAGCCAAGCACAAATCCC
rs955000	ACGTTGGATGAGCCTGGGCAACACATAGAA	ACGTTGGATGTGACTGTGTTGTGTACTACC	ctGCAACACATAGAAACCCCAT
rs9577874	ACGTTGGATGTCTCACCCAGCCACAGAGAG	ACGTTGGATGTCAGCACCCGTTAGAGCTTC	cCAGAGAGGGAGAGGAC
rs9604573	ACGTTGGATGCTGGTCAGGCAGGTCTGATT	ACGTTGGATGATCTCAGCTGCGTTGGGCG	AAGTGGAAAATCATGTTCA
rs9683303	ACGTTGGATGCTTATTTCACTGAAGTCCTC	ACGTTGGATGGATGAAACCCCTTGAAAGCC	gACTGAAGTCCTCAAAACTCTCC
rs9923231	ACGTTGGATGGGATTATAGGCGTGAGCCAC	ACGTTGGATGTCTGGGAAGTCAAGCAAGAG	GCGTGAGCCACCGCACC
rs7175032	ACGTTGGATGTCAAACCAAACCATTCCCCC	ACGTTGGATGCTTCCCGATTCTGAGTTCAC	CCCCCCCTCATCCCTA
rs432925	ACGTTGGATGCCTGCCCCTCACCTGTGTA	ACGTTGGATGAGTACCCTACGCTCAAGTTC	CACCTGTGTACTCCTC
rs836832	ACGTTGGATGTCTTGCTAGCACTTGGCTTC	ACGTTGGATGATGTGAGGCCCCACTGTGAG	GCTTCCCGTTTGTAGC
rs400037	ACGTTGGATGCCTGCTCAAAATTCTTGCCC	ACGTTGGATGTCTCCTGAGCCAGGAGATAC	GAGGGTCTTAACTGGC
rs702030	ACGTTGGATGTTAAGGTGCTCTGGCAAAGG	ACGTTGGATGAAAGTTCAGGACAAGGTGGC	GGCAAAGGTAGTTGTGA
rs11647490	ACGTTGGATGGTACTTGAGAGTAGGGAAGC	ACGTTGGATGCACGAGGACATCATCATTGC	cGCATCCAGCTCGTTGGC
rs2685127	ACGTTGGATGAGACGTCTGCCAAGATCTTC	ACGTTGGATGAAAGCCCGCTAGGAGCTCCC	TGCTGTTTGTCAACCAGA
rs1976715	ACGTTGGATGAAAAGAGGGAGCAGGCAGAG	ACGTTGGATGTTAGATGGGATCCTAGAGGG	tTGGGCACGGGAACATGG
rs4677875	ACGTTGGATGAGTGTCAGTGCACATTCACC	ACGTTGGATGGAAGGTTTTGTGCAGGGAAG	cTAATCTTACACGTCCTCCT
rs6464929	ACGTTGGATGTTGAATCTCATGACCCCTGC	ACGTTGGATGAGGCACTGGACAAAAGTGAG	CTCATGACCCCTGCTTTGCT
rs4727006	ACGTTGGATGAAACACGACTTGGCCTCAAC	ACGTTGGATGCAGGCAATCACCAGTCTCC	AAACCCACAGTGGAACAGGC
rs1686447	ACGTTGGATGAACAGCCTCTGCTCCTTATC	ACGTTGGATGATGTCCAGATAAAGACTCGG	CTCCTTATCTATGTCAGGTAT
rs10163054	ACGTTGGATGTCACAATGATGGTAATACG	ACGTTGGATGAATGGCAACATTCTGGAAGC	TGATGAAAATGTAACCAAATG
rs1625940	ACGTTGGATGCTGGAATGGTTTCCAGTTAC	ACGTTGGATGTTCTGTGAGATACCAGATAG	ATGGTTTCCAGTTACTTGGCTT
rs1010954	ACGTTGGATGTCCAAGCCTCTGGAACCCG	ACGTTGGATGCTGTCTTTAGAGAGGGACTG	AGCTGCCCCTGCCCTTCCAAGTC

SNP ID	Forward	Reverse	Extension
rs12901424	ACGTTGGATGATAGACCAGGTAGGGTTGGG	ACGTTGGATGTGTGAGAAGAGAAGGACCAG	cCAGATAATGGGTCCCCCTAATG
rs6464930	ACGTTGGATGGACAACTCCTTTACAGGGTG	ACGTTGGATGATGTTGAGAGCCCACTTGAG	GTGAATTTTACAAAAAACAAAGA
rs876017	ACGTTGGATGGTTTGGTGATATGTGTCCGC	ACGTTGGATGTGAGCCTTTGCAGGCAAGAG	TGGTGATATGTGTCCGCGAAGGC
rs1734346	ACGTTGGATGATCCCCTTGTTCTCAGTATG	ACGTTGGATGTGTTCAATGGGAAGCTAGTT	CAACAGAACCAGTATTTTCAACTCT
rs1734343	ACGTTGGATGACAGTGATTCCTGTACCCTC	ACGTTGGATGAGCATAGACCTAAGCAAGCG	aCCTCCGAGTCATTCTGAAGATGAA
rs1686482	ACGTTGGATGGATGAAGTAGCCGCTACAAG	ACGTTGGATGGTAAGATGCTGAGCTTACTG	GTATACTGGATGTAGGTATCCAGAA
rs7795577	ACGTTGGATGACTTTGCTCCGCCAAAACTG	ACGTTGGATGAGAGTCTGAGCTGGGACTTG	GAGTTGCTGAAAGGGACCAAGGGCC
rs3792366	ACGTTGGATGAAATCGGGACTGTATTCAGG	ACGTTGGATGACAACTGGTTTGAGGGTTGC	gAATCGGGACTGTATTCAGGAAGTGA
rs1799919	ACGTTGGATGCACCGCTCTCCAGGAATTTC	ACGTTGGATGACTCATCCCTGTCTTCCACA	ATCCAGCGTGCGTTCCCCGTTGTAATC
rs1063495	ACGTTGGATGTGGCCAATGTGTTCAATTCG	ACGTTGGATGATGCTGGAGCTGAATCAGAC	CAATGTGTTCAATTCGATTGTGAAATA
rs1533756	ACGTTGGATGTTCCAGAGGCTTGGAGGGAT	ACGTTGGATGAAACTCCTCACAAGACCCTC	GGATCCCGTTCCTCAG
rs13066716	ACGTTGGATGAGTTCTGGCTGTCAGTGACC	ACGTTGGATGTTTACTCGTGAGGGATCCAG	TTCCTGAACAAGGGACA
rs2070871	ACGTTGGATGATCAAACACAACCAGCTGCC	ACGTTGGATGAACCCACTCCTCCCCATCC	CCAGCTGCCCCTTGTCAT
rs1078982	ACGTTGGATGTTCCCAGCCTGGACCTCACT	ACGTTGGATGCTAAACTGGTTTTGGAGTGG	GGACCTCACTCATTCATA
rs16939823	ACGTTGGATGAGCCCTGGCTTGAGGGAAG	ACGTTGGATGATGCCTGGAAGAGGAGCTG	GGAAGGCCCAGGGCCTGT
rs3844075	ACGTTGGATGAAGGGTTTGTGGAAAGTGGG	ACGTTGGATGAGACCTGGAAAAGGATGGTG	ggGGAGAGGTGCTTAAGAG
rs1551927	ACGTTGGATGCGCTGAAGCCAGAATGTTTC	ACGTTGGATGAGCAGCGTTCATTTACCTCC	TCCGGGGCCAATTATTCTTT
rs10085877	ACGTTGGATGTGGCCACATTTTGTGGGCAG	ACGTTGGATGTAAGCACTGGCACTGAGGG	ggatGGCAGGCCATGCTGTT
rs10272564	ACGTTGGATGCTGAGATTTAATTCGTTACCC	ACGTTGGATGCCAGCCTACAGTGGTCATTT	gcccCATCAGCCAAAGCAAAT
rs4727005	ACGTTGGATGAACAGCTCCCTCTGGTGAAC	ACGTTGGATGACATGGTGCTCCGTTGAGG	gggaAGCTGGAGCTGCCCGAA
rs1198873	ACGTTGGATGCTGTTAGATCTGCAGAAGGG	ACGTTGGATGCTGGATTGTAGCAGGAGTAG	cGAAGGTCAACATCTCATTTTA
rs11904084	ACGTTGGATGTAACATTCCTCACCATCAGC	ACGTTGGATGTTTTCTGATAACGCCCCACC	gCTGGACTACAAGCAGTTGAAG
rs10269104	ACGTTGGATGGCAGCAGTCAGAGTAGAGAG	ACGTTGGATGAGATGGTCTCTCCTGAAGCC	tatcCAGAGTAGAGAGCAGAGC
rs1130674	ACGTTGGATGTCTTAAAACCATCCAGCGTG	ACGTTGGATGCACTCAAGTTCTTTCCTGCC	ccccTGCGTTCCCCGTTGTAATC
rs8040336	ACGTTGGATGGATGAATAATTGCCTTCTC	ACGTTGGATGTGCTTCTTCAAGTGGCTGAC	TGTCATAACTTTTATTTTGACTT
rs11863142	ACGTTGGATGTGAGGGGCAGGCCGGTCAT	ACGTTGGATGCCACTCGGCAATGCCCTCA	ccttGGGACTCCCTGCAGGACCAC
rs11070411	ACGTTGGATGCATTCCTGCTAAATCTTGCTG	ACGTTGGATGCCTCTCAGTGCATTTTACCC	TAGGCATAATAACTGATCTGAAAA
rs1107377	ACGTTGGATGGTCATGACACCTTAGACCTG	ACGTTGGATGTTCAGTGGAAGAAGAGCTCG	gaggCCTGTGTAAATGTTGACTAG

SNP ID	Forward	Reverse	Extension
rs419949	ACGTTGGATGATGAGGGCAGTGACTGTGG	ACGTTGGATGTGGGCCAAGGCCAAGAAGGT	aaattCCCCTCAGGACCTGCAGGAC
rs12471762	ACGTTGGATGGTGGTAAATTTTATTTTATG	ACGTTGGATGCGAGTCTTATTTTCAGTGGG	GGTAAATTTTATTTTATGTTATGTGC

APPENDIX II UNIVARIATE RESULT OF DOSE ASSOCIATION (UPPSALA STUDY).

Gene	Dependent	SNP	DF	Type III SS	R-Square	P-value
ABCB1	sqrt_dose	rs2188531	1	0.1428745	0.000467	0.7632
ABCB1	sqrt_dose	rs6465117	1	0.13951337	0.000463	0.7688
ABCB1	sqrt_dose	rs17328991	1	0.1019175	0.000345	0.7981
ABCB1	sqrt_dose	rs10267099	2	1.46844255	0.004762	0.6324
ABCB1	sqrt_dose	rs2157926	1	0.2286435	0.000737	0.7043
ABCB1	sqrt_dose	rs2214101	1	0.13471973	0.000424	0.7734
ABCB1	sqrt_dose	rs17149824	1	0.14402468	0.000453	0.7649
ABCB1	sqrt_dose	rs4728709	1	0.18263224	0.000586	0.7349
ABCB1	sqrt_dose	rs2214102	2	7.9844733	0.034709	0.0786
ABCB1	sqrt_dose	rs9282564	2	0.21792496	0.000764	0.9335
ABCB1	sqrt_dose	rs1858923	2	9.62596927	0.032526	0.0518
ABCB1	sqrt_dose	rs3789243	2	6.74737669	0.022133	0.1234
ABCB1	sqrt_dose	rs1202181	2	2.53212039	0.008031	0.4593
ABCB1	sqrt_dose	rs1202172	2	3.41923358	0.010961	0.3529
ABCB1	sqrt_dose	rs1989830	2	2.96595808	0.009966	0.3823
ABCB1	sqrt_dose	rs1202179	2	2.36795701	0.007644	0.4787
ABCB1	sqrt_dose	rs1202180	2	2.28844376	0.008014	0.5170
ABCB1	sqrt_dose	rs10260862	2	0.33751716	0.001158	0.9036
ABCB1	sqrt_dose	rs2235015	2	0.24294653	0.000771	0.9290
ABCB1	sqrt_dose	rs1202167	2	1.93349841	0.006185	0.5581
ABCB1	sqrt_dose	rs1202169	2	2.5722116	0.008663	0.4433
ABCB1	sqrt_dose	rs955000	2	3.83224296	0.012704	0.3045
ABCB1	sqrt_dose	rs868755	2	2.26361415	0.007405	0.5010
ABCB1	sqrt_dose	rs1922240	2	1.26254243	0.004321	0.6772
ABCB1	sqrt_dose	rs2235033	2	0.15501537	0.000534	0.9546
ABCB1	sqrt_dose	rs2235035	2	3.22950572	0.017044	0.3002
ABCB1	sqrt_dose	rs2235013	2	0.93393554	0.003161	0.7438
ABCB1	sqrt_dose	rs2091766	2	1.11564984	0.003569	0.7120
ABCB1	sqrt_dose	rs2235046	2	0.56809304	0.001956	0.8376
ABCB1	sqrt_dose	rs1922242	2	0.85172025	0.002795	0.7644
ABCB1	sqrt_dose	rs4148737	2	0.23139447	0.000775	0.9323
ABCB1	sqrt_dose	rs2235040	2	0.892099	0.002813	0.7630
ABCB1	sqrt_dose	rs2032582_3	2	0.21411746	0.000817	0.9383
ABCB1	sqrt_dose	rs6959435	2	0.03883942	0.000138	0.9876
ABCB1	sqrt_dose	rs4148742	2	0.49438749	0.001570	0.8613
ABCB1	sqrt_dose	rs2235067	2	0.38761973	0.001316	0.8859
ABCB1	sqrt_dose	rs1045642	2	0.64644023	0.002156	0.8270
ABCB1	sqrt_dose	rs3842	2	1.06439023	0.003484	0.7266
APOE	sqrt_dose	rs7412	2	2.08206097	0.007666	0.5280
APOE	sqrt_dose	rs429358	2	5.80904527	0.019273	0.1574
CALU	sqrt_dose	rs8597	2	2.01646723	0.007410	0.5495
CALU	sqrt_dose	rs11653	2	12.84034745	0.041134	0.0174

Gene	Dependent	SNP	DF	Type III SS	R-Square	P-value
CALU	sqrt_dose	rs339098	2	7.86815039	0.027129	0.0952
CALU	sqrt_dose	rs2307040	2	8.8565882	0.027851	0.0628
CALU	sqrt_dose	rs1006023	2	8.93422784	0.028095	0.0613
CALU	sqrt_dose	rs2290228	2	1.70872124	0.005560	0.5807
CALU	sqrt_dose	rs339054	2	9.1170075	0.029445	0.0576
CALU	sqrt_dose	rs2060717	2	2.64126295	0.009076	0.4263
CALU	sqrt_dose	rs339057	2	7.04922611	0.022167	0.1099
CYP1A1	sqrt_dose	rs2470893	2	1.08031098	0.003527	0.7289
CYP1A1	sqrt_dose	rs2606345	2	1.69033211	0.005590	0.5888
CYP1A1	sqrt_dose	rs4646421	2	0.85262982	0.002967	0.7619
CYP1A2	sqrt_dose	rs2470890	2	3.26364492	0.010263	0.3620
CYP1A2	sqrt_dose	rs2472304	2	3.48712402	0.011374	0.3531
CYP1A2	sqrt_dose	rs762551	2	4.87851018	0.015424	0.2197
CYP2C18	sqrt_dose	rs10736086	2	0.34984672	0.001109	0.8995
CYP2C18	sqrt_dose	rs2860840	2	2.9847188	0.009675	0.3990
CYP2C18	sqrt_dose	rs2281891_2	2	2.12433905	0.007309	0.5111
CYP2C18	sqrt_dose	rs10509675	2	4.17107664	0.013179	0.2743
CYP2C18	sqrt_dose	rs7919273	2	0.53049901	0.001681	0.8516
CYP2C18	sqrt_dose	rs1926711	2	3.1463342	0.010561	0.3805
CYP2C18	sqrt_dose	rs7898763	2	3.88061458	0.012215	0.3036
CYP2C18	sqrt_dose	rs7099637	2	3.90275221	0.013046	0.2949
CYP2C18	sqrt_dose	rs7896133	1	21.42330419	0.073529	0.0001
CYP2C18	sqrt_dose	rs7478002	2	2.17092589	0.006853	0.5115
CYP2C18	sqrt_dose	rs2901783	2	6.48006569	0.021929	0.1315
CYP2C18	sqrt_dose	rs2860837	2	4.18423274	0.014123	0.2645
CYP2C18	sqrt_dose	rs1926706	2	0.92806719	0.003020	0.7514
CYP2C18	sqrt_dose	rs12249418	2	3.70131673	0.012228	0.3050
CYP2C19	sqrt_dose	GS30424	2	8.79152903	0.031968	0.0685
CYP2C19	sqrt_dose	rs1853205	2	2.66789601	0.008805	0.4335
CYP2C19	sqrt_dose	rs4244284	2	0.92410835	0.003033	0.7608
CYP2C19	sqrt_dose	GS30100	1	16.71871943	0.056501	0.0012
CYP2C19	sqrt_dose	GS30253	2	6.46963001	0.021113	0.1419
CYP2C19	sqrt_dose	rs4417205	2	2.52601287	0.008774	0.4666
CYP2C19	sqrt_dose	rs17879456	2	1.17348711	0.004296	0.6965
CYP2C19	sqrt_dose	rs17882419	2	4.59443162	0.014699	0.2467
CYP2C19	sqrt_dose	rs12248560	2	5.05259273	0.016633	0.2192
CYP2C19	sqrt_dose	rs3814637	2	32.10230095	0.105603	<.0001
CYP2C8	sqrt_dose	rs1557044	2	2.61157221	0.012213	0.4205
CYP2C8	sqrt_dose	rs17110453	2	3.33417342	0.015519	0.3218
CYP2C8	sqrt_dose	rs2275622	2	1.81788969	0.008553	0.5434
CYP2C8	sqrt_dose	rs11572080	1	2.49249684	0.011711	0.1935
CYP2C8	sqrt_dose	rs3752988	2	2.4271239	0.011367	0.4416
CYP2C8	sqrt_dose	rs1058930	1	1.19240743	0.005651	0.3705
CYP2C8	sqrt_dose	rs1341163	2	1.29857521	0.006239	0.6372

Gene	Dependent	SNP	DF	Type III SS	R-Square	P-value
CYP2C8	sqrt_dose	rs947173	2	2.56661918	0.012332	0.4093
CYP2C8	sqrt_dose	rs1891071	2	1.46455774	0.006900	0.6116
CYP2C8	sqrt_dose	rs2275620	2	3.08904361	0.014963	0.3534
CYP2C8	sqrt_dose	rs7898759	2	4.18789688	0.020725	0.2433
CYP2C8	sqrt_dose	rs1058932	2	2.96684837	0.013879	0.3707
CYP2C9	sqrt_dose	rs4607998	2	1.44238256	0.004557	0.6391
CYP2C9	sqrt_dose	rs1057911	2	44.31154644	0.144784	<.0001
CYP2C9	sqrt_dose	rs2298037	2	6.59065715	0.021991	0.1307
CYP2C9	sqrt_dose	rs9332222	2	1.61680924	0.005104	0.6088
CYP2C9	sqrt_dose	rs9332214	2	44.01487444	0.139204	<.0001
CYP2C9	sqrt_dose	rs1057910	2	45.03959448	0.141634	<.0001
CYP2C9	sqrt_dose	rs9332197	2	0.28715136	0.001057	0.9121
CYP2C9	sqrt_dose	rs1934966	1	0.82480581	0.002802	0.4670
CYP2C9	sqrt_dose	rs1934964	2	1.53630853	0.004837	0.6233
CYP2C9	sqrt_dose	rs9325473	2	44.40440128	0.147496	<.0001
CYP2C9	sqrt_dose	rs1856908	2	2.78657292	0.008763	0.4202
CYP2C9	sqrt_dose	rs4917639	2	37.20429074	0.117176	<.0001
CYP2C9	sqrt_dose	rs2153628	2	3.10009247	0.009749	0.3829
CYP2C9	sqrt_dose	rs2475376	2	6.76202224	0.021287	0.1240
CYP2C9	sqrt_dose	rs10509679	2	5.88371769	0.018581	0.1637
CYP2C9	sqrt_dose	rs2860905	2	21.45781125	0.070732	0.0010
CYP2C9	sqrt_dose	rs17110268	2	1.606719	0.005053	0.6072
CYP2C9	sqrt_dose	rs9332108	2	45.03959448	0.141634	<.0001
CYP2C9	sqrt_dose	rs4917636	2	2.8022929	0.008938	0.4378
CYP3A4	sqrt_dose	rs11773597	1	1.04380801	0.003303	0.4201
CYP3A4	sqrt_dose	rs2242480	2	1.75201141	0.005554	0.5842
CYP3A5	sqrt_dose	GS30681	2	8.72136874	0.029442	0.0640
CYP3A5	sqrt_dose	GS30260	2	1.42885469	0.004687	0.6461
CYP3A5	sqrt_dose	GS30600	2	5.12417816	0.016114	0.2019
CYP3A5	sqrt_dose	rs776746	2	1.3267638	0.004292	0.6820
CYP3A5	sqrt_dose	GS30593	2	3.09656456	0.009778	0.3951
CYP3A5	sqrt_dose	rs6976017	2	2.58207852	0.008400	0.4506
CYP3A5	sqrt_dose	rs4646457	2	0.93971197	0.003377	0.7629
EPHX1	sqrt_dose	rs2102663	1	0.51297575	0.001714	0.5676
EPHX1	sqrt_dose	rs3753663	1	1.74888107	0.005543	0.2960
EPHX1	sqrt_dose	rs1051741_2	2	4.67968013	0.014953	0.2302
EPHX1	sqrt_dose	rs2671266	2	4.4678725	0.014155	0.2637
EPHX1	sqrt_dose	rs2292567	2	7.86826941	0.028598	0.0913
EPHX1	sqrt_dose	rs2234922	2	9.23297231	0.032275	0.0605
EPHX1	sqrt_dose	rs4149223	2	2.84598521	0.012301	0.3761
EPHX1	sqrt_dose	rs2740170	2	0.58127545	0.001866	0.8358
EPHX1	sqrt_dose	rs2260863	2	1.34265597	0.005868	0.6546
EPHX1	sqrt_dose	rs2292566	2	2.35747614	0.008596	0.4927
EPHX1	sqrt_dose	rs1051740	2	0.04285224	0.000182	0.9861

Gene	Dependent	SNP	DF	Type III SS	R-Square	P-value
EPHX1	sqrt_dose	rs3817268	2	1.59759588	0.005024	0.6089
EPHX1	sqrt_dose	rs2671270	2	3.29082911	0.010450	0.3686
EPHX1	sqrt_dose	rs3738047	1	0.15745438	0.000519	0.7532
EPHX1	sqrt_dose	rs2671272	2	1.84343982	0.006170	0.5694
EPHX1	sqrt_dose	rs3753661	1	2.11998033	0.006682	0.2535
EPHX1	sqrt_dose	rs3753660	2	1.5995325	0.005206	0.6122
EPHX1	sqrt_dose	rs3753659	2	1.58828987	0.004995	0.6122
EPHX1	sqrt_dose	rs3753658	2	1.51920053	0.004823	0.6302
EPHX1	sqrt_dose	rs2854451	2	3.61602469	0.011386	0.3331
EPHX1	sqrt_dose	rs2854450	2	1.01925999	0.003361	0.7374
EPHX1	sqrt_dose	rs2854447	2	3.52379575	0.011246	0.3513
EPHX1	sqrt_dose	rs2854461	2	1.19396464	0.004152	0.6820
EPHX1	sqrt_dose	rs6426089	2	1.11581491	0.003700	0.7176
EPHX1	sqrt_dose	rs4653436	2	14.39173813	0.045716	0.0112
F10	sqrt_dose	rs5960	2	6.35037606	0.020425	0.1561
F10	sqrt_dose	rs473598	2	5.20104655	0.017318	0.1987
F10	sqrt_dose	rs776897	3	1.00203458	0.003151	0.8918
F10	sqrt_dose	rs3211770	2	5.67837886	0.018390	0.1830
F10	sqrt_dose	rs2026160	2	1.35967122	0.004391	0.6583
F10	sqrt_dose	rs2251102	2	1.38966969	0.005767	0.6259
F10	sqrt_dose	rs3211764	2	1.13640148	0.003953	0.7127
F10	sqrt_dose	rs2480946	2	2.14668455	0.007028	0.5245
F10	sqrt_dose	rs693335	2	3.26463572	0.010845	0.3768
F10	sqrt_dose	rs483949	2	0.05387112	0.000172	0.9834
F10	sqrt_dose	rs485798	2	3.8564801	0.012714	0.3062
F10	sqrt_dose	rs776905	2	6.887572	0.022713	0.1154
F10	sqrt_dose	rs474810	2	1.99234813	0.006498	0.5418
F10	sqrt_dose	rs563964	2	5.25007401	0.018172	0.2182
F10	sqrt_dose	rs3093261	2	0.92468481	0.003094	0.7531
F2	sqrt_dose	rs3136516	2	3.15950427	0.009957	0.3807
F2	sqrt_dose	rs2282687	2	3.09432691	0.009811	0.3805
F2	sqrt_dose	rs5898	2	6.47376991	0.020450	0.1362
F2	sqrt_dose	rs3136460	2	3.23492237	0.010242	0.3703
F2	sqrt_dose	rs5896	2	3.9541219	0.012474	0.2923
F2	sqrt_dose	rs2070852	2	2.77038829	0.008807	0.4240
F2	sqrt_dose	rs2070851	2	7.85489576	0.024969	0.0861
F2	sqrt_dose	rs3136447	2	4.05724411	0.013567	0.2695
F2	sqrt_dose	rs3136435	2	1.09943287	0.003648	0.7197
F2	sqrt_dose	rs2070850	2	1.26385185	0.004935	0.6782
F5	sqrt_dose	rs2269648	2	2.05512126	0.006766	0.5283
F5	sqrt_dose	rs3753305	2	7.97046148	0.027668	0.0823
F5	sqrt_dose	rs6028	2	7.85436773	0.031331	0.0979
F5	sqrt_dose	rs9332504	2	4.17027663	0.013230	0.2729
F5	sqrt_dose	rs2298905	2	4.69192413	0.017056	0.2378

Gene	Dependent	SNP	DF	Type III SS	R-Square	P-value
F5	sqrt_dose	rs2298908	2	2.43214908	0.007898	0.4689
F5	sqrt_dose	rs2236870	2	7.3019912	0.024851	0.1134
F5	sqrt_dose	rs3766121	2	5.3740322	0.019020	0.2031
F5	sqrt_dose	rs3766120	2	7.88491922	0.025221	0.0872
F5	sqrt_dose	rs3766119	2	5.42995615	0.017630	0.1912
F5	sqrt_dose	rs1894702	2	6.35339168	0.020392	0.1487
F5	sqrt_dose	rs6029	2	2.14920561	0.006851	0.5187
F5	sqrt_dose	rs6022	2	2.98118483	0.010302	0.3837
F5	sqrt_dose	rs6012	2	2.50070121	0.008016	0.4674
F5	sqrt_dose	rs3766117	2	2.13737838	0.007440	0.5242
F5	sqrt_dose	rs1894699	2	6.6566744	0.023726	0.1223
F5	sqrt_dose	rs6427198	2	0.48103783	0.001663	0.8768
F5	sqrt_dose	rs6033	2	2.64805427	0.009230	0.4321
F5	sqrt_dose	rs6035	1	3.06308747	0.010709	0.1645
F5	sqrt_dose	rs721161	2	1.47432244	0.004872	0.6506
F5	sqrt_dose	rs2298909	2	2.9580763	0.012419	0.3917
F5	sqrt_dose	rs6015	2	0.74013286	0.002647	0.7867
F5	sqrt_dose	rs6025	2	2.36196859	0.008299	0.4925
F5	sqrt_dose	rs6036	2	1.59299072	0.005260	0.6221
F5	sqrt_dose	rs3766110	2	0.07613453	0.000264	0.9774
F5	sqrt_dose	rs6037	2	1.52447597	0.004794	0.6244
F5	sqrt_dose	rs6024	2	4.39307473	0.014073	0.2602
F5	sqrt_dose	rs6021	2	0.11506032	0.000366	0.9653
F5	sqrt_dose	rs6018	2	3.20136968	0.010456	0.3703
F5	sqrt_dose	rs4525	3	0.48940891	0.001608	0.9621
F5	sqrt_dose	rs6032	2	0.71235853	0.002243	0.8025
F5	sqrt_dose	rs1557572	2	0.56456747	0.001889	0.8371
F5	sqrt_dose	rs9332618	2	2.15058365	0.006988	0.5173
F5	sqrt_dose	rs6030	2	2.33482087	0.007502	0.4909
F5	sqrt_dose	rs9332629	2	3.56616671	0.011254	0.3374
F5	sqrt_dose	rs2213867	2	0.67139079	0.002138	0.8125
F5	sqrt_dose	rs2213866	2	0.67784658	0.002158	0.8101
F5	sqrt_dose	rs2227244	2	0.4006878	0.001332	0.8828
F5	sqrt_dose	rs966751	2	4.3882834	0.013965	0.2629
F5	sqrt_dose	rs6027	2	3.85098202	0.012172	0.3067
F5	sqrt_dose	rs2187952	2	0.75695812	0.002432	0.7964
F7	sqrt_dose	rs2476324	2	2.6995095	0.008489	0.4318
F7	sqrt_dose	rs6046	2	3.39641793	0.011477	0.3559
F7	sqrt_dose	rs6041	2	2.07683276	0.006590	0.5318
F7	sqrt_dose	rs488703	2	1.05060222	0.003829	0.7176
F7	sqrt_dose	rs569557	2	2.63001583	0.008733	0.4601
F7	sqrt_dose	rs493833	2	3.94308895	0.012506	0.2969
F7	sqrt_dose	rs491098	2	3.08529256	0.013818	0.3522
F7	sqrt_dose	rs2774030	2	0.62712618	0.002011	0.8234

Gene	Dependent	SNP	DF	Type III SS	R-Square	P-value
F7	sqrt_dose	rs3093233	2	7.48046198	0.023713	0.0975
F7	sqrt_dose	rs3093230	2	8.7830618	0.030734	0.0612
F7	sqrt_dose	rs3093229	2	6.00590373	0.019704	0.1495
F9	sqrt_dose	rs445691	2	4.33323409	0.013660	0.2670
F9	sqrt_dose	rs434447	2	5.36723351	0.018524	0.1790
F9	sqrt_dose	rs440051	2	4.69599288	0.015367	0.2406
F9	sqrt_dose	rs413536	2	1.5049353	0.005631	0.6101
F9	sqrt_dose	rs110583	2	4.58551641	0.014696	0.2432
F9	sqrt_dose	rs413957	2	4.81223104	0.015499	0.2232
F9	sqrt_dose	rs6048	2	1.70729075	0.005837	0.5887
F9	sqrt_dose	rs422187	2	2.47126951	0.008150	0.4577
F9	sqrt_dose	rs398101	2	8.23167632	0.026263	0.0787
F9	sqrt_dose	rs392959	2	6.5921089	0.020896	0.1303
F9	sqrt_dose	rs401597	2	4.77401881	0.016315	0.2257
GAS6	sqrt_dose	rs9577874	2	0.64898874	0.002122	0.8147
GAS6	sqrt_dose	rs9604573	2	5.97359097	0.020083	0.1595
GAS6	sqrt_dose	rs6602908	2	5.4636253	0.018707	0.1810
GAS6	sqrt_dose	rs7997328	2	1.53257493	0.005157	0.6230
NQO1	sqrt_dose	rs689456	2	0.74585143	0.002371	0.7981
NQO1	sqrt_dose	rs2917669	2	0.27004522	0.000850	0.9204
NQO1	sqrt_dose	rs2917671	2	9.33604926	0.033893	0.0449
NQO1	sqrt_dose	rs1437135	2	0.27144351	0.000991	0.9210
NQO1	sqrt_dose	rs689452	2	0.49152599	0.001554	0.8600
NQO1	sqrt_dose	rs689453	1	0.95199967	0.002994	0.4427
NQO1	sqrt_dose	GS30566	1	0.02355043	0.000077	0.9046
NQO1	sqrt_dose	rs7186002	2	1.87876248	0.006668	0.5458
NQO1	sqrt_dose	rs1800566	2	1.15462818	0.003956	0.7041
NR1I2	sqrt_dose	rs3814057	2	4.79721662	0.015974	0.2219
NR1I2	sqrt_dose	rs1054191	2	0.47055375	0.001612	0.8690
NR1I2	sqrt_dose	rs3732360	2	1.79003247	0.005677	0.5724
NR1I2	sqrt_dose	rs3732359	2	1.13984383	0.004065	0.6889
NR1I2	sqrt_dose	rs2472682	2	1.20981938	0.003853	0.6970
NR1I2	sqrt_dose	rs3732357	2	2.72386327	0.008779	0.4270
NR1I2	sqrt_dose	rs3732356	2	2.49281776	0.008162	0.4705
NR1I2	sqrt_dose	rs1464602	2	7.38256851	0.029139	0.0859
NR1I2	sqrt_dose	rs7643645	2	8.25811216	0.029896	0.0651
NR1I2	sqrt_dose	rs2461818_2	1	2.2592398	0.007593	0.2319
NR1I2	sqrt_dose	rs13059232	2	5.66180334	0.018302	0.1714
NR1I2	sqrt_dose	rs2461823	2	5.46074474	0.017966	0.1819
NR1I2	sqrt_dose	rs2472677	3	11.84315278	0.038942	0.0600
NR1I2	sqrt_dose	rs1403527	2	5.44189137	0.017167	0.1832
NR1I2	sqrt_dose	rs2056530	2	6.35766889	0.022425	0.1455
NR1I2	sqrt_dose	rs2472672	2	6.51952904	0.021041	0.1414
NR1I2	sqrt_dose	rs2276706	2	1.42072339	0.004549	0.6470

Gene	Dependent	SNP	DF	Type III SS	R-Square	P-value
NR1I2	sqrt_dose	rs1523127	2	1.04100685	0.003496	0.7195
NR1I2	sqrt_dose	rs1523130	2	0.01721202	0.000071	0.9946
NR1I2	sqrt_dose	rs7643038	2	1.54672057	0.004874	0.6211
NR1I3	sqrt_dose	rs2501870	2	0.9777107	0.003115	0.7424
NR1I3	sqrt_dose	rs7530560	2	2.60794294	0.008500	0.4406
NR1I3	sqrt_dose	rs2502804	2	0.04854883	0.000197	0.9857
NR1I3	sqrt_dose	rs6686001	2	9.24025622	0.034324	0.0505
NR1I3	sqrt_dose	rs3003596	2	7.27653216	0.025301	0.1049
NR1I3	sqrt_dose	rs2307424	2	7.16041357	0.026913	0.0984
NR1I3	sqrt_dose	rs2307418	2	2.11046244	0.007206	0.5178
NR1I3	sqrt_dose	rs4073054	2	1.99576631	0.008628	0.5500
NR1I3	sqrt_dose	rs4233368	2	0.06198354	0.000202	0.9812
ORM1_2	sqrt_dose	GS30155	1	5.29876961	0.016916	0.0722
ORM1_2	sqrt_dose	GS30283	1	6.88921317	0.027913	0.0424
ORM1_2	sqrt_dose	rs1976193	2	2.27309693	0.007243	0.5013
ORM1_2	sqrt_dose	rs10982151	1	2.92141531	0.009794	0.1720
ORM1_2	sqrt_dose	rs17230081	2	1.00149348	0.003536	0.7244
ORM1_2	sqrt_dose	rs2787337	2	0.11797556	0.000381	0.9641
P4HB	sqrt_dose	rs1799919	2	1.0186033	0.010689	0.3650
P4HB	sqrt_dose	rs1130674	2	2.386251906	0.016979	0.0955
P4HB	sqrt_dose	rs2070871	2	1.363365296	0.009151	0.2588
P4HB	sqrt_dose	rs876017	2	0.850222584	0.005379	0.4292
P4HB	sqrt_dose	rs1533756	2	2.246947202	0.015043	0.1091
P4HB	sqrt_dose	rs1010954	2	0.236199594	0.001383	0.7899
PDIA2	sqrt_dose	rs432925	2	0.417196809	0.002488	0.6596
PDIA2	sqrt_dose	rs2685127	2	1.163145202	0.006771	0.3149
PDIA2	sqrt_dose	rs400037	2	3.594688734	0.020014	0.0295
PDIA3	sqrt_dose	rs10163054	2	2.029765709	0.011574	0.1344
PDIA3	sqrt_dose	rs8040336	2	0.118528655	0.000914	0.8883
PDIA3	sqrt_dose	rs11070411	2	1.326393811	0.009524	0.2687
PDIA3	sqrt_dose	rs7175032	2	0.29073609	0.001724	0.7481
PDIA4	sqrt_dose	rs10085877	2	0.356000694	0.002446	0.7010
PDIA4	sqrt_dose	rs4727005	2	0.377770767	0.002542	0.6860
PDIA4	sqrt_dose	rs10272564	2	0.249943669	0.001884	0.7792
PDIA4	sqrt_dose	rs10269104	2	0.59350858	0.004230	0.5536
PDIA4	sqrt_dose	rs6464929	2	0.17781333	0.001073	0.8373
PDIA4	sqrt_dose	rs1551927	2	1.120921255	0.007818	0.3286
PDIA4	sqrt_dose	rs6464930	2	0.210143256	0.001567	0.8107
PDIA5	sqrt_dose	rs1078982	2	0.675265617	0.005024	0.5106
PDIA5	sqrt_dose	rs3792366	2	1.730843606	0.010673	0.1804
PDIA5	sqrt_dose	rs4677875	2	1.547257543	0.009344	0.2159
PDIA5	sqrt_dose	rs702030	2	0.444883337	0.002561	0.6416
PDIA5	sqrt_dose	rs836832	2	0.026478749	0.000156	0.9739
PDIA5	sqrt_dose	rs1107377	2	0.228777877	0.002173	0.7959

Gene	Dependent	SNP	DF	Type III SS	R-Square	P-value
PDIA6	sqrt_dose	rs1686482	2	0.915588433	0.005517	0.4025
PDIA6	sqrt_dose	rs1198873	2	0.160884135	0.001135	0.8515
PDIA6	sqrt_dose	rs11904084	2	0.196506662	0.001395	0.8218
PDIA6	sqrt_dose	rs1686447	2	0.002565505	0.000016	0.9974
PDIA6	sqrt_dose	rs1734343	2	1.045043314	0.006443	0.3540
PDIA6	sqrt_dose	rs1734346	2	0.003370188	0.000020	0.9966
PDIA6	sqrt_dose	rs12471762	2	0.202306932	0.001554	0.8171
PROC	sqrt_dose	rs2069933	2	6.67530579	0.021880	0.1292
PROC	sqrt_dose	rs2069928	2	1.51334871	0.004988	0.6234
PROC	sqrt_dose	rs2069924	2	9.12724682	0.029936	0.0566
PROC	sqrt_dose	rs1518759	2	5.45356618	0.017300	0.1872
PROC	sqrt_dose	rs5936	2	5.88092628	0.019460	0.1608
PROC	sqrt_dose	rs973760	2	8.63458813	0.034223	0.0606
PROC	sqrt_dose	rs2069921	2	8.49097097	0.027243	0.0715
PROC	sqrt_dose	rs2069919	2	25.08270504	0.087800	0.0003
PROC	sqrt_dose	rs2069916	2	5.51731365	0.017617	0.1832
PROC	sqrt_dose	rs2069915	2	2.77087859	0.011058	0.4018
PROC	sqrt_dose	rs2069910	2	12.46310896	0.043535	0.0208
PROC	sqrt_dose	rs1799809	2	23.30259318	0.076473	0.0007
PROC	sqrt_dose	rs2069901	2	21.05871589	0.072256	0.0015
PROS1	sqrt_dose	rs7650230	2	2.41868926	0.007853	0.4654
PROS1	sqrt_dose	rs8178592	2	2.10882456	0.006667	0.5244
PROS1	sqrt_dose	rs5013930	2	3.11374059	0.009866	0.3879
PROS1	sqrt_dose	rs8178607	2	1.82579137	0.005988	0.5720
PROS1	sqrt_dose	rs8178610	2	0.33182038	0.001483	0.9000
PROS1	sqrt_dose	rs4857343	2	5.67417705	0.018568	0.1800
PROS1	sqrt_dose	rs8178623	2	1.79217551	0.007749	0.5494
PROS1	sqrt_dose	rs4857037	2	4.37091662	0.015886	0.2733
PROS1	sqrt_dose	rs8178649	2	1.55418021	0.004975	0.6180
PROS1	sqrt_dose	rs9713061	2	5.51303856	0.018096	0.1701
PROS1	sqrt_dose	rs9683303	2	0.42690243	0.001536	0.8822
PROZ	sqrt_dose	rs3024764	2	6.17172611	0.021079	0.1502
PROZ	sqrt_dose	rs3024747	2	1.94735584	0.006543	0.5799
PROZ	sqrt_dose	rs3024746	2	3.45952522	0.010906	0.3471
PROZ	sqrt_dose	rs17881956	2	2.51456391	0.007991	0.4611
PROZ	sqrt_dose	rs3024743	2	0.26346762	0.000993	0.9227
PROZ	sqrt_dose	rs17886440	2	1.10691855	0.003494	0.7109
PROZ	sqrt_dose	rs3024731	2	1.2423885	0.003922	0.6817
PROZ	sqrt_dose	rs2480948	2	9.16254677	0.029121	0.0569
PROZ	sqrt_dose	rs513479	2	3.09517632	0.010371	0.3954
PROZ	sqrt_dose	rs3024718	2	5.29253882	0.017112	0.1907
PROZ	sqrt_dose	rs3024711	2	6.80337261	0.022537	0.1173
PROZ	sqrt_dose	rs2273971	2	0.36235782	0.001193	0.8933
PROZ	sqrt_dose	rs7335409	2	5.11718453	0.016475	0.2013

Gene	Dependent	SNP	DF	Type III SS	R-Square	P-value
SERPINC1	sqrt_dose	rs2227588	2	0.92574359	0.002969	0.7483
SERPINC1	sqrt_dose	rs2227590	2	0.90248567	0.002911	0.7636
SERPINC1	sqrt_dose	rs2227593	2	0.22973271	0.000753	0.9306
SERPINC1	sqrt_dose	rs2227594	2	0.71979859	0.002307	0.7993
SERPINC1	sqrt_dose	rs2227607	2	2.2559729	0.007204	0.4995
SERPINC1	sqrt_dose	rs5877	2	0.79834596	0.002685	0.7808
SERPINC1	sqrt_dose	rs5878	2	1.68534442	0.005485	0.5996
SERPINC1	sqrt_dose	rs2759328	2	5.17858637	0.016453	0.2000
SERPINC1	sqrt_dose	rs2146372	2	4.75156343	0.015065	0.2329
VKORC1LD	sqrt_dose	rs4889537	2	42.42121122	0.134033	<.0001
VKORC1LD	sqrt_dose	rs9923231	2	93.54772025	0.327232	<.0001
VKORC1LD	sqrt_dose	rs9934438	2	83.87566494	0.291960	<.0001
VKORC1LD	sqrt_dose	rs2359612	2	94.72991801	0.298437	<.0001
VKORC1LD	sqrt_dose	rs7294	2	64.80046481	0.209333	<.0001
VKORC1LD	sqrt_dose	rs8046978	2	16.12002303	0.050909	0.0065
VKORC1LD	sqrt_dose	rs4889599	2	37.47511211	0.123298	<.0001
VKORC1LD	sqrt_dose	rs11642603	1	6.73545886	0.021825	0.0414
VKORC1LD	sqrt_dose	rs4889630	2	5.18182141	0.016598	0.1989
VKORC1LD	sqrt_dose	rs7405035	2	10.29709659	0.036326	0.0385
VKORC1LD	sqrt_dose	rs4889490	2	50.90208213	0.162005	<.0001
VKORC1LD	sqrt_dose	rs11642466	1	6.29596653	0.020713	0.0453
VKORC1LD	sqrt_dose	rs7194347	2	10.77216988	0.035589	0.0303

APPENDIX III UNIVARIATE RESULT OF DOSE ASSOCIATION (WARG STUDY).

Gene	SNP	MAF	Patients	%	R^2	P-value
ABCB1	rs3842	0.137	1193	97.66%	0.00	9.47E-01
ABCB1	rs2235040	0.108	1284	96.98%	0.20	3.22E-01
ABCB1	rs4148737	0.453	1219	93.50%	0.10	6.99E-01
ABCB1	rs2235046	0.432	1279	95.85%	0.00	8.32E-01
ABCB1	rs2235033	0.493	1260	98.11%	0.00	7.47E-01
ABCB1	rs1922240	0.337	1275	90.48%	0.00	9.08E-01
ABCB1	rs955000	0.102	1238	93.58%	0.10	4.38E-01
ABCB1	rs10260862	0.195	1299	92.07%	0.10	5.97E-01
ABCB1	rs1202172	0.351	1269	95.17%	0.10	5.50E-01
ABCB1	rs9282564	0.132	1239	96.60%	0.10	6.60E-01
ABCB1	rs2214102	0.098	1283	96.30%	0.00	9.14E-01
ABCB1	rs2214101	0.064	1198	96.90%	0.10	6.55E-01
ABCB1	rs10267099	0.239	1293	90.11%	0.30	1.41E-01
APOE	rs429358	0.173	1261	95.24%	0.00	9.68E-01
APOE	rs7412	0.07	1225	92.52%	0.00	9.93E-01
CALU	rs230704	0.285	1294	97.21%	0.00	7.27E-01
CALU	rs339057	0.438	1297	97.96%	0.30	1.27E-01
CALU	rs2060717	0.062	1280	94.56%	0.00	7.62E-01
CALU	rs339054	0.468	1287	97.51%	0.40	9.03E-02
CALU	rs2290228	0.15	1244	84.14%	0.10	7.09E-01
CALU	rs2307040	0.36	1291	97.43%	0.20	2.38E-01
CALU	rs339098	0.356	1290	86.03%	0.20	3.51E-01
CALU	rs339095	0.473	1114	93.96%	0.20	2.62E-01
CALU	rs11653	0.391	1252	97.73%	0.30	1.81E-01
CALU	rs8597	0.07	1139	96.68%	0.10	5.01E-01
CYP1A1	rs1048943	0.026	1216	96.00%	0.00	2.14E-01
CYP1A1	rs4646421	0.081	1302	88.97%	0.10	6.75E-01
CYP1A1	rs2606345	0.317	1178	91.84%	0.30	1.83E-01
CYP1A1	rs2470893	0.342	1271	98.34%	0.70	1.49E-02
CYP1A2	rs2069522	0.032	1290	94.18%	0.00	9.85E-01
CYP1A2	rs762551	0.27	1247	96.30%	0.20	3.41E-01
CYP1A2	rs4646425	0.019	1238	93.50%	0.00	8.55E-01
CYP1A2	rs2470890	0.336	1275	97.43%	0.10	5.98E-01
CYP2C18	rs2901783	0.222	1289	97.36%	1.20	4.09E-04
CYP2C18	rs2860840	0.388	1223	92.37%	0.10	5.53E-01
CYP2C19	rs3814637	0.073	1284	96.98%	7.10	3.14E-21
CYP2C19	rs12248560	0.185	1298	98.04%	0.50	4.78E-02
CYP2C19	rs17878459	0.035	1295	96.83%	0.00	9.24E-01
CYP2C19	rs4244284	0.45	1277	96.45%	0.10	4.94E-01
CYP2C19	rs4244285	0.162	1282	99.17%	0.10	4.30E-01
CYP2C19	rs4417205	0.158	1313	97.81%	0.10	5.15E-01
CYP2C8	rs1058932	0.165	1264	94.86%	0.20	2.51E-01

Gene	SNP	MAF	Patients	%	R^2	P-value
CYP2C8	rs10509681	0.093	1147	86.63%	3.00	3.08E-08
CYP2C8	rs2275620	0.417	1204	90.94%	1.60	7.99E-05
CYP2C8	rs947173	0.295	1312	99.09%	0.90	3.23E-03
CYP2C8	rs1536430	0.022	1303	98.41%	0.40	1.84E-02
CYP2C8	rs1058930	0.059	1288	99.09%	0.00	8.95E-01
CYP2C8	rs11572080	0.1	1256	95.47%	4.10	3.16E-12
CYP2C8	rs17110453	0.147	1021	77.11%	0.10	5.24E-01
CYP2C8	rs1557044	0.139	1312	97.28%	0.40	7.85E-02
CYP2C9	rs1799853	0.114	1321	99.55%	4.10	1.34E-12
CYP2C9	rs2860905	0.214	1296	82.48%	7.40	3.05E-22
CYP2C9	rs2475376	0.152	1258	97.89%	1.00	1.38E-03
CYP2C9	rs4917639	0.209	1092	99.77%	11.70	4.61E-30
CYP2C9	rs1856908	0.433	1291	99.77%	2.90	6.71E-09
CYP2C9	rs9332197	0.073	1280	97.51%	0.20	3.21E-01
CYP2C9	rs1057910	0.071	1321	95.02%	6.30	1.82E-19
CYP2C9	cyp2c9s2s3	n/a	1318	96.68%	11.80	6.63E-34
CYP3A4	GS30681	0.069	1309	98.11%	0.10	4.70E-01
CYP3A4	rs4986910	0.008	1299	93.88%	0.40	1.85E-02
CYP3A4	rs2242480	0.087	1243	92.22%	0.20	3.68E-01
CYP3A4	rs11773597	0.066	1221	98.87%	0.10	4.33E-01
CYP3A5	rs4646457	0.076	1276	98.11%	0.30	1.32E-01
CYP3A5	rs6976017	0.054	1302	97.89%	0.30	1.13E-01
CYP3A5	rs28365083	0.01	1293	98.34%	0.10	2.60E-01
CYP3A5	rs28365094	0.122	1299	96.37%	0.40	5.82E-02
CYP3A5	rs4646453	0.018	1307	97.66%	0.00	4.54E-01
CYP3A5	rs776746	0.073	1296	98.72%	0.40	9.65E-02
EPHX1	rs4653436	0.309	1207	97.05%	0.10	4.25E-01
EPHX1	rs6426089	0.496	1288	97.28%	0.50	3.59E-02
EPHX1	rs2854461	0.328	1284	96.98%	0.30	1.79E-01
EPHX1	rs3753660	0.144	1292	91.77%	0.20	3.36E-01
EPHX1	rs3753661	0.069	1309	95.92%	0.00	8.82E-01
EPHX1	rs2671272	0.21	1272	97.58%	0.10	6.76E-01
EPHX1	rs2671270	0.254	1270	96.83%	0.20	2.94E-01
EPHX1	rs3817268	0.274	1285	91.16%	0.80	6.47E-03
EPHX1	rs1051740	0.294	1215	98.87%	0.20	2.77E-01
EPHX1	rs2292566	0.132	1282	96.60%	0.10	4.20E-01
EPHX1	rs2260863	0.3	1280	95.39%	0.00	8.60E-01
EPHX1	rs2740170	0.201	1291	96.07%	0.00	7.80E-01
EPHX1	rs4149223	0.468	1279	81.95%	0.10	5.77E-01
EPHX1	rs2234922	0.213	1299	97.51%	0.00	9.16E-01
EPHX1	rs2292567	0.101	1085	95.24%	0.10	7.21E-01
EPHX1	rs1051741	0.103	1261	96.68%	0.00	8.51E-01
EPHX1	rs3753663	0.047	1309	98.87%	0.10	4.76E-01
EPHX1	rs2102663	0.157	1263	98.11%	0.10	6.01E-01

Gene	SNP	MAF	Patients	%	R^2	P-value
F10	rs3093261	0.427	1273	88.14%	0.30	1.52E-01
F10	rs3212998	0.064	1266	96.15%	0.00	8.93E-01
F10	rs776905	0.103	1272	94.26%	0.00	7.74E-01
F10	rs693335	0.405	1220	97.43%	0.10	5.86E-01
F10	rs3211764	0.465	1290	90.03%	0.20	2.28E-01
F10	rs2251102	0.192	1167	96.30%	0.50	4.82E-02
F10	rs2026160	0.3	1192	92.15%	0.20	2.31E-01
F10	rs753057	0.052	1297	97.96%	0.10	6.92E-01
F10	rs3213005	0.043	1275	96.07%	0.10	5.61E-01
F10	rs5960	0.135	1248	95.62%	0.20	2.11E-01
F2	rs3136435	0.071	1271	98.56%	0.20	3.40E-01
F2	rs2070851	0.22	1280	96.68%	0.20	3.14E-01
F2	rs2070852	0.301	1297	96.00%	0.00	8.46E-01
F2	rs5898	0.092	1305	98.87%	0.50	5.17E-02
F2	rs2282687	0.131	1309	97.96%	0.00	7.48E-01
F2	rs3136516	0.449	1273	96.15%	0.00	8.63E-01
F5	rs9332618	0.133	1269	99.09%	0.10	4.03E-01
F5	rs4656687	0.321	1293	96.07%	0.00	7.48E-01
F5	rs1557572	0.315	1295	99.09%	0.00	8.13E-01
F5	rs6018	0.058	1133	95.85%	0.00	8.05E-01
F5	rs6037	0.068	1298	91.01%	0.00	8.09E-01
F5	rs3766110	0.229	1289	96.75%	0.10	5.33E-01
F5	rs6025	0.066	1312	91.01%	0.30	1.42E-01
F5	rs2298909	0.303	1205	97.21%	0.10	5.26E-01
F5	rs721161	0.387	1206	97.36%	0.00	7.48E-01
F5	rs6035	0.076	1281	82.55%	0.10	4.87E-01
F5	rs6427198	0.478	1205	97.66%	0.10	4.44E-01
F5	rs1894699	0.49	1279	91.09%	0.00	8.65E-01
F5	rs6029	0.164	1312	85.57%	0.10	4.01E-01
F5	rs2298905	0.301	1093	98.04%	0.10	5.81E-01
F5	rs9332504	0.074	1272	97.81%	0.20	3.50E-01
F5	rs3753305	0.441	1287	96.60%	0.10	5.27E-01
F7	rs3093229	0.24	1205	80.29%	0.00	9.34E-01
F7	rs3093230	0.254	1298	98.34%	0.10	6.97E-01
F7	rs2774030	0.371	1299	98.04%	0.00	7.98E-01
F7	rs6041	0.101	1063	98.11%	0.40	9.56E-02
F7	rs6046	0.103	1302	91.01%	0.10	4.01E-01
F9	rs3817939	0.02	1219	93.43%	0.10	5.85E-01
F9	rs401597	0.283	1273	96.15%	0.30	1.72E-01
F9	rs6048	0.282	1237	95.32%	0.40	7.55E-02
F9	rs413536	0.2	1262	92.07%	0.20	2.19E-01
GAS6	rs7997328	0.287	1256	94.86%	0.30	2.04E-01
GAS6	rs6602908	0.395	1238	93.50%	0.10	6.23E-01
GAS6	rs9577874	0.408	1227	92.67%	0.10	7.14E-01

Gene	SNP	MAF	Patients	%	R^2	P-value
GGCX	rs2028898	0.275	1294	94.56%	0.10	5.00E-01
GGCX	rs7568458	0.451	1252	97.73%	0.10	4.97E-01
NQO1	rs1800566	0.17	1284	96.98%	0.20	2.86E-01
NQO1	rs689453	0.077	1290	97.43%	0.20	3.11E-01
NQO1	rs2917671	0.116	1315	99.32%	0.00	9.06E-01
NR1I2	rs2276706	0.401	1283	96.90%	0.40	9.34E-02
NR1I2	rs2056530	0.191	1266	97.58%	0.10	4.37E-01
NR1I2	rs2472677	0.399	1210	93.35%	0.10	6.77E-01
NR1I2	rs2461818	0.096	1193	96.83%	0.10	7.03E-01
NR1I2	rs7643645	0.352	1292	95.62%	0.30	1.65E-01
NR1I2	rs12721607	0.028	1282	96.22%	0.00	8.03E-01
NR1I2	rs1464602	0.392	1274	97.05%	0.10	4.93E-01
NR1I2	rs3732356	0.067	1236	98.11%	0.30	2.07E-01
NR1I2	rs3732357	0.34	1285	95.69%	0.10	6.30E-01
NR1I2	rs2472682	0.385	1274	91.39%	0.00	8.22E-01
NR1I2	rs3732359	0.261	1282	90.11%	0.10	4.02E-01
NR1I2	rs1054191	0.153	1299	96.83%	0.10	6.45E-01
NR1I2	rs3814057	0.199	1267	96.22%	0.10	6.74E-01
NR1I3	rs4233368	0.263	1245	94.03%	0.30	1.39E-01
NR1I3	rs4073054	0.362	1277	88.97%	0.10	4.60E-01
NR1I3	rs2307418	0.161	1257	96.45%	0.00	9.94E-01
NR1I3	rs2307420	0.023	1236	93.96%	0.00	8.50E-01
NR1I3	rs2307424	0.357	1178	93.35%	0.20	3.14E-01
NR1I3	rs3003596	0.445	1244	93.96%	0.00	8.96E-01
NR1I3	rs6686001	0.135	1244	94.94%	0.00	7.46E-01
ORM1	rs2787337	0.311	1227	92.67%	0.70	1.08E-02
ORM1	rs10982151	0.136	1279	96.60%	0.50	5.18E-02
ORM1	rs1687390	0.044	1310	98.94%	0.00	9.14E-01
ORM2	rs17230081	0.199	1269	95.85%	0.00	8.84E-01
P4HB	rs1799919	0.207399	1235	94.79%	0.00	5.55E-01
P4HB	rs1130674	0.204665	1233	93.13%	0.00	4.96E-01
P4HB	rs2070871	0.200445	1255	93.50%	0.00	4.81E-01
P4HB	rs876017	0.209328	1238	94.64%	0.00	5.01E-01
P4HB	rs1533756	0.399257	1253	93.28%	0.00	5.23E-01
P4HB	rs1010954	0.206231	1252	94.56%	0.00	5.93E-01
PDIA2	rs2685127	0.140483	1216	92.75%	0.00	8.37E-01
PDIA2	rs400037	0.242857	1228	91.84%	0.00	1.72E-01
PDIA3	rs10163054	0.0697227	1196	92.67%	0.00	4.90E-01
PDIA3	rs8040336	0.444694	1248	94.34%	0.00	8.80E-01
PDIA3	rs11070411	0.164399	1227	90.33%	0.00	4.24E-02
PDIA3	rs7175032	0.0754647	1249	94.26%	0.00	3.82E-01
PDIA4	rs10085877	0.195167	1252	94.56%	0.00	4.27E-01
PDIA4	rs4727005	0.203786	1254	94.18%	0.00	6.61E-01
PDIA4	rs10272564	0.202399	1241	93.50%	0.00	6.34E-01

Gene	SNP	MAF	Patients	%	R^2	P-value
PDIA4	rs10269104	0.190656	1231	92.98%	0.00	5.13E-01
PDIA4	rs6464929	0.202612	1238	93.73%	0.00	4.92E-01
PDIA4	rs1551927	0.14158	1247	94.71%	0.00	4.46E-01
PDIA4	rs6464930	0.186003	1232	93.05%	0.00	9.02E-01
PDIA5	rs1078982	0.222802	1249	93.73%	0.00	3.97E-01
PDIA5	rs3792366	0.372175	1166	92.37%	0.00	6.87E-01
PDIA5	rs4677875	0.371201	1211	94.34%	0.00	5.62E-01
PDIA5	rs702030	0.180859	1224	91.47%	0.00	8.44E-01
PDIA5	rs836832	0.185575	1223	88.07%	0.00	9.93E-02
PDIA5	rs1107377	0.495519	1241	92.45%	0.00	2.46E-02
PDIA6	rs1686482	0.495061	1201	91.54%	0.00	3.92E-01
PDIA6	rs1198873	0.39368	1252	90.71%	0.00	5.39E-01
PDIA6	rs11904084	0.388554	1230	89.12%	0.00	6.91E-01
PDIA6	rs1686447	0.395833	1180	94.56%	0.00	5.17E-01
PDIA6	rs1734343	0.495427	1212	92.75%	0.00	3.45E-01
PDIA6	rs1734346	0.394275	1207	92.90%	0.00	7.18E-01
PDIA6	rs12471762	0.168055	1228	91.16%	0.00	5.55E-01
PROC	rs2069901	0.435	1309	97.73%	0.10	6.21E-01
PROC	rs1799809	0.438	1244	98.11%	0.00	9.53E-01
PROC	rs2069910	0.425	1277	97.51%	0.00	7.33E-01
PROC	rs2069919	0.373	1294	98.87%	0.20	2.07E-01
PROC	rs5936	0.276	1291	96.45%	0.10	3.95E-01
PROC	rs2069928	0.18	1313	99.17%	0.00	7.69E-01
PROC	rs2069933	0.279	1299	93.96%	0.20	2.59E-01
PROS1	rs9683303	0.353	1223	91.99%	0.60	3.26E-02
PROS1	rs8178633	0.05	1218	92.37%	0.80	7.03E-03
PROS1	rs4857037	0.076	1256	94.18%	0.00	8.58E-01
PROS1	rs4857343	0.142	1265	95.54%	0.10	6.20E-01
PROS1	rs8178610	0.496	1240	94.86%	0.00	9.18E-01
PROS1	rs8178607	0.244	1222	92.30%	0.00	8.65E-01
PROS1	rs5013930	0.122	1247	93.66%	0.30	1.74E-01
PROZ	rs2273971	0.062	1290	97.43%	0.20	2.91E-01
PROZ	rs3024718	0.175	1300	95.54%	0.10	7.16E-01
PROZ	rs2480948	0.204	1299	98.19%	0.00	8.19E-01
PROZ	rs3024746	0.206	1265	98.11%	0.10	5.68E-01
SERPINC1	rs2759328	0.092	1313	98.41%	0.10	3.79E-01
SERPINC1	rs5878	0.3	1306	99.17%	0.10	3.83E-01
SERPINC1	rs2227607	0.101	1306	98.64%	0.10	4.40E-01
SERPINC1	rs2227590	0.099	1303	98.64%	0.40	6.01E-02
VKORC1	rs11150606	0.027	1309	97.58%	1.30	2.55E-05
VKORC1	rs7294	0.393	1294	97.66%	12.30	1.21E-37
VKORC1	rs2359612	0.394	1292	97.73%	29.80	9.82E-100
VKORC1	rs9923231	0.393	1293	98.87%	29.30	1.03E-97

APPENDIX IV TREND TEST RESULT IN BLEEDING (ALL BLEEDERS).

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
1	rs2502804	TREND	8_74	638_2094	6.437	1	1.1E-02
1	rs3753661	TREND	12_84	192_2756	5.266	1	2.2E-02
1	rs4653436	TREND	18_68	836_1902	3.574	1	5.9E-02
1	rs2307418	TREND	9_83	474_2434	2.816	1	9.3E-02
1	rs2227590	TREND	5_89	299_2631	2.445	1	1.2E-01
1	rs2854461	TREND	37_57	945_1967	2.107	1	1.5E-01
1	rs2260863	TREND	22_72	880_2026	2.002	1	1.6E-01
1	rs3753305	TREND	48_46	1279_1615	1.787	1	1.8E-01
1	rs2307424	TREND	26_64	955_1723	1.778	1	1.8E-01
1	rs1894699	TREND	40_54	1422_1466	1.689	1	1.9E-01
1	rs4656687	TREND	24_68	908_1890	1.632	1	2.0E-01
1	rs6025	TREND	3_89	182_2700	1.438	1	2.3E-01
1	rs3817268	TREND	19_69	791_2147	1.262	1	2.6E-01
1	rs2307420	TREND	4_90	70_2800	1.231	1	2.7E-01
1	rs6037	TREND	4_92	205_2719	1.182	1	2.8E-01
1	rs3003596	TREND	45_47	1237_1597	0.9863	1	3.2E-01
1	rs2671272	TREND	24_70	623_2289	0.9364	1	3.3E-01
1	rs2102663	TREND	11_77	468_2444	0.8253	1	3.6E-01
1	rs4149223	TREND	36_50	1359_1541	0.8171	1	3.7E-01
1	rs5878	TREND	25_71	884_2058	0.7048	1	4.0E-01
1	rs3753660	TREND	11_85	424_2514	0.6808	1	4.1E-01
1	rs6426089	TREND	48_42	1440_1488	0.6203	1	4.3E-01
1	rs4233368	TREND	27_65	745_2139	0.5841	1	4.4E-01
1	rs2740170	TREND	17_77	605_2331	0.3463	1	5.6E-01
1	rs1557572	TREND	28_68	939_2007	0.3139	1	5.8E-01
1	rs2671270	TREND	22_72	748_2164	0.2624	1	6.1E-01
1	rs2292566	TREND	10_78	384_2534	0.2403	1	6.2E-01
1	rs3766110	TREND	23_69	658_2226	0.2392	1	6.2E-01
1	rs6427198	TREND	40_48	1295_1411	0.1884	1	6.6E-01
1	rs9332618	TREND	14_80	398_2536	0.1399	1	7.1E-01
1	rs2227607	TREND	10_84	280_2656	0.1257	1	7.2E-01
1	rs3753663	TREND	4_92	145_2817	0.1045	1	7.5E-01
1	rs6035	TREND	6_86	213_2675	0.09738	1	7.6E-01
1	rs2759328	TREND	10_86	282_2674	0.08451	1	7.7E-01
1	rs2234922	TREND	18_72	606_2294	0.04526	1	8.3E-01
1	rs4073054	TREND	33_55	1047_1825	0.03997	1	8.4E-01
1	rs6029	TREND	16_76	492_2462	0.03568	1	8.5E-01
1	rs1051740	TREND	27_63	804_1948	0.02582	1	8.7E-01
1	rs2298909	TREND	26_62	807_1865	0.01684	1	9.0E-01
1	rs9332504	TREND	7_87	205_2669	0.01403	1	9.1E-01
1	rs6686001	TREND	11_73	385_2485	0.007153	1	9.3E-01
1	rs721161	TREND	37_59	1055_1699	0.002096	1	9.6E-01

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
1	rs1051741	TREND	10_80	327_2585	0.001093	1	9.7E-01
2	rs2069933	TREND	16_74	827_2103	4.7	1	3.0E-02
2	rs5936	TREND	19_75	810_2102	2.587	1	1.1E-01
2	rs7568458	TREND	48_44	1307_1601	1.934	1	1.6E-01
2	rs2069928	TREND	13_81	533_2413	1.16	1	2.8E-01
2	rs2069910	TREND	42_50	1209_1661	0.4354	1	5.1E-01
2	rs2028898	TREND	28_64	827_2123	0.2614	1	6.1E-01
2	rs2069919	TREND	37_57	1098_1834	0.1382	1	7.1E-01
2	rs2069901	TREND	43_53	1280_1662	0.06353	1	8.0E-01
2	rs1799809	TREND	41_51	1271_1623	0.01503	1	9.0E-01
2	rs1686482	TREND	35_59	1422_1402	5.843	1	1.6E-02
2	rs12471762	TREND	8_88	459_2349	4.234	1	4.0E-02
2	rs1734343	TREND	40_62	1404_1430	3.994	1	4.6E-02
2	rs1198873	TREND	49_55	1137_1767	2.51	1	1.1E-01
2	rs11904084	TREND	49_57	1115_1755	2.162	1	1.4E-01
2	rs1686447	TREND	48_56	1109_1695	1.686	1	1.9E-01
2	rs1734346	TREND	47_55	1116_1716	1.681	1	1.9E-01
3	rs2472677	TREND	46_46	1102_1654	3.559	1	5.9E-02
3	rs2056530	TREND	11_85	555_2347	3.527	1	6.0E-02
3	rs8178633	TREND	1_83	141_2629	2.683	1	1.0E-01
3	rs4857037	TREND	3_85	219_2619	2.246	1	1.3E-01
3	rs2276706	TREND	31_63	1166_1766	1.831	1	1.8E-01
3	rs12721607	TREND	1_93	77_2799	0.8593	1	3.5E-01
3	rs4857343	TREND	16_74	424_2512	0.7104	1	4.0E-01
3	rs8178610	TREND	32_40	1392_1418	0.6898	1	4.1E-01
3	rs9683303	TREND	26_58	985_1823	0.601	1	4.4E-01
3	rs3732357	TREND	29_65	994_1930	0.3962	1	5.3E-01
3	rs1054191	TREND	16_78	446_2494	0.2458	1	6.2E-01
3	rs2472682	TREND	39_57	1110_1792	0.2245	1	6.4E-01
3	rs7643645	TREND	34_56	1039_1891	0.2004	1	6.5E-01
3	rs2461818	TREND	8_86	265_2471	0.1407	1	7.1E-01
3	rs5013930	TREND	13_83	354_2488	0.09422	1	7.6E-01
3	rs1464602	TREND	34_56	1132_1754	0.07756	1	7.8E-01
3	rs3814057	TREND	19_73	568_2314	0.04673	1	8.3E-01
3	rs3732359	TREND	20_60	762_2194	0.02358	1	8.8E-01
3	rs3732356	TREND	7_89	201_2619	0.003564	1	9.5E-01
3	rs8178607	TREND	22_68	680_2134	0.003569	1	9.5E-01
3	rs1107377	TREND	55_39	1401_1447	2.985	1	8.4E-02
3	rs702030	TREND	13_93	520_2342	2.396	1	1.2E-01
3	rs3792366	TREND	30_68	1027_1743	1.58	1	2.1E-01
3	rs4677875	TREND	32_72	1043_1789	1.49	1	2.2E-01
3	rs1078982	TREND	20_84	643_2251	0.509	1	4.8E-01
3	rs836832	TREND	20_86	523_2353	0.03056	1	8.6E-01
7	rs2214102	TREND	14_70	287_2613	3.955	1	4.7E-02

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
7	rs2235040	TREND	5_91	307_2617	2.894	1	8.9E-02
7	rs10260862	TREND	11_79	563_2361	2.809	1	9.4E-02
7	rs2060717	TREND	2_92	188_2766	2.734	1	9.8E-02
7	rs28365094	TREND	16_80	353_2571	1.749	1	1.9E-01
7	rs2242480	TREND	5_89	260_2626	1.455	1	2.3E-01
7	rs2290228	TREND	10_84	428_2416	1.393	1	2.4E-01
7	rs1922240	TREND	36_56	989_1949	1.159	1	2.8E-01
7	rs9282564	TREND	8_74	389_2477	0.9251	1	3.4E-01
7	rs11653	TREND	31_59	1115_1719	0.8928	1	3.4E-01
7	rs11773597	TREND	4_88	190_2594	0.8738	1	3.5E-01
7	rs4148737	TREND	48_48	1255_1525	0.8321	1	3.6E-01
7	rs28365083	TREND	0_60	28_2692	0.6303	1	4.3E-01
7	rs10267099	TREND	19_73	693_2239	0.4282	1	5.1E-01
7	GS30681	TREND	5_91	202_2740	0.3985	1	5.3E-01
7	rs4646453	TREND	2_76	52_2900	0.2848	1	5.9E-01
7	rs339098	TREND	31_63	1030_1868	0.2661	1	6.1E-01
7	rs3842	TREND	15_81	377_2357	0.2561	1	6.1E-01
7	rs6976017	TREND	4_90	160_2774	0.2533	1	6.1E-01
7	rs776746	TREND	6_88	221_2703	0.1835	1	6.7E-01
7	rs1202172	TREND	32_64	1026_1912	0.09854	1	7.5E-01
7	rs339057	TREND	43_51	1297_1639	0.09834	1	7.5E-01
7	rs955000	TREND	9_85	301_2575	0.07531	1	7.8E-01
7	rs339054	TREND	44_52	1367_1541	0.05268	1	8.2E-01
7	rs4646457	TREND	8_86	230_2682	0.04788	1	8.3E-01
7	rs2235046	TREND	42_52	1267_1633	0.03467	1	8.5E-01
7	rs2214101	TREND	5_79	175_2571	0.02564	1	8.7E-01
7	rs2307040	TREND	34_62	1055_1887	0.008183	1	9.3E-01
7	rs2235033	TREND	45_47	1416_1494	0.002128	1	9.6E-01
7	rs10085877	TREND	33_73	565_2341	8.983	1	2.7E-03
7	rs4727005	TREND	34_72	592_2318	8.697	1	3.2E-03
7	rs10272564	TREND	32_68	581_2299	8.274	1	4.0E-03
7	rs10269104	TREND	32_74	549_2321	7.983	1	4.7E-03
7	rs6464929	TREND	31_69	579_2303	7.081	1	7.8E-03
7	rs6464930	TREND	27_71	526_2346	5.276	1	2.2E-02
7	rs1551927	TREND	20_86	415_2485	1.721	1	1.9E-01
9	rs1687390	TREND	7_77	130_2768	2.896	1	8.9E-02
9	rs17230081	TREND	13_79	581_2339	1.832	1	1.8E-01
9	rs2787337	TREND	23_69	871_1921	1.639	1	2.0E-01
9	rs10982151	TREND	13_81	387_2521	0.02192	1	8.8E-01
10	rs1536430	TREND	4_48	61_2651	6.534	1	1.1E-02
10	rs1557044	TREND	20_76	409_2543	3.768	1	5.2E-02
10	rs947173	TREND	21_75	872_2076	2.679	1	1.0E-01
10	rs11572080	TREND	5_89	282_2560	2.187	1	1.4E-01
10	rs2860905	TREND	15_81	641_2299	2.127	1	1.4E-01

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
10	rs1799853	TREND	7_89	346_2620	1.726	1	1.9E-01
10	rs2475376	TREND	19_77	448_2476	1.347	1	2.5E-01
10	rs2860840	TREND	32_60	1096_1692	0.787	1	3.8E-01
10	rs4244285	TREND	11_79	449_2411	0.7655	1	3.8E-01
10	rs1856908	TREND	46_50	1277_1663	0.7458	1	3.9E-01
10	rs9332197	TREND	5_83	223_2703	0.4334	1	5.1E-01
10	rs17878459	TREND	2_92	92_2812	0.3354	1	5.6E-01
10	rs2901783	TREND	19_73	649_2277	0.1159	1	7.3E-01
10	rs4417205	TREND	14_82	454_2478	0.05555	1	8.1E-01
10	rs1058932	TREND	15_81	472_2394	0.04548	1	8.3E-01
10	rs1058930	TREND	5_89	166_2722	0.03018	1	8.6E-01
10	rs12248560	TREND	17_79	538_2394	0.02572	1	8.7E-01
10	rs1057910	TREND	7_89	207_2759	0.01363	1	9.1E-01
10	rs3814637	TREND	7_87	206_2678	0.01225	1	9.1E-01
10	rs2275620	TREND	37_51	1156_1564	0.006972	1	9.3E-01
10	rs4244284	TREND	41_51	1273_1603	0.003275	1	9.5E-01
11	rs2282687	TREND	5_91	375_2573	4.709	1	3.0E-02
11	rs5898	TREND	15_81	269_2669	4.424	1	3.5E-02
11	rs3136516	TREND	48_44	1267_1623	2.471	1	1.2E-01
11	rs2070852	TREND	27_67	880_2080	0.04374	1	8.3E-01
11	rs2070851	TREND	21_73	625_2279	0.03426	1	8.5E-01
11	rs3136435	TREND	7_87	210_2728	0.01223	1	9.1E-01
13	rs6046	TREND	5_91	310_2610	2.835	1	9.2E-02
13	rs5960	TREND	17_75	385_2513	2.055	1	1.5E-01
13	rs3212998	TREND	9_85	180_2744	1.875	1	1.7E-01
13	rs3093230	TREND	14_68	683_2235	1.814	1	1.8E-01
13	rs2480948	TREND	24_72	585_2361	1.542	1	2.1E-01
13	rs776905	TREND	6_82	314_2626	1.338	1	2.5E-01
13	rs2774030	TREND	31_65	1101_1835	1.107	1	2.9E-01
13	rs3213005	TREND	6_88	124_2782	0.95	1	3.3E-01
13	rs3024746	TREND	15_77	595_2341	0.8507	1	3.6E-01
13	rs3093261	TREND	44_48	1246_1650	0.8435	1	3.6E-01
13	rs7997328	TREND	30_64	813_2051	0.5543	1	4.6E-01
13	rs3211764	TREND	36_48	1351_1555	0.4364	1	5.1E-01
13	rs2026160	TREND	25_67	833_1905	0.4158	1	5.2E-01
13	rs753057	TREND	4_92	158_2786	0.2631	1	6.1E-01
13	rs693335	TREND	34_46	1108_1658	0.1989	1	6.6E-01
13	rs9577874	TREND	35_57	1127_1681	0.1759	1	6.8E-01
13	rs3093229	TREND	23_69	663_2139	0.08732	1	7.7E-01
13	rs6602908	TREND	35_53	1101_1723	0.02283	1	8.8E-01
13	rs2273971	TREND	6_88	179_2759	0.0132	1	9.1E-01
13	rs3024718	TREND	16_80	502_2442	0.009844	1	9.2E-01
14	rs230704	TREND	26_70	847_2083	0.1488	1	7.0E-01
15	rs1048943	TREND	4_86	70_2700	1.172	1	2.8E-01

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
15	rs2470890	TREND	28_64	969_1923	0.3787	1	5.4E-01
15	rs4646425	TREND	1_93	53_2775	0.3356	1	5.6E-01
15	rs762551	TREND	22_68	764_2058	0.308	1	5.8E-01
15	rs2470893	TREND	34_60	986_1918	0.1934	1	6.6E-01
15	rs2606345	TREND	27_63	868_1834	0.175	1	6.8E-01
15	rs4646421	TREND	8_88	240_2708	0.004477	1	9.5E-01
15	rs2069522	TREND	3_93	89_2847	0.002861	1	9.6E-01
15	rs11070411	TREND	20_80	439_2407	1.559	1	2.1E-01
15	rs7175032	TREND	5_99	218_2676	1.086	1	3.0E-01
15	rs10163054	TREND	5_95	203_2613	0.6927	1	4.1E-01
15	rs8040336	TREND	46_58	1275_1607	3.58E-06	1	1.0E+00
16	rs9923231	TREND	28_58	1153_1755	1.778	1	1.8E-01
16	rs2917671	TREND	26_60	1004_1726	1.531	1	2.2E-01
16	rs11150606	TREND	4_92	84_2858	0.5722	1	4.5E-01
16	rs2359612	TREND	34_60	1149_1749	0.4605	1	5.0E-01
16	rs1800566	TREND	17_75	485_2371	0.1418	1	7.1E-01
16	rs689453	TREND	7_89	232_2676	0.06061	1	8.1E-01
16	rs7294	TREND	37_55	1146_1776	0.03684	1	8.5E-01
16	rs2685127	TREND	17_89	405_2455	0.2967	1	5.9E-01
16	rs400037	TREND	27_79	696_2174	0.08015	1	7.8E-01
17	rs1799919	TREND	29_57	599_2189	7.393	1	6.5E-03
17	rs1533756	TREND	31_75	1165_1739	5.128	1	2.4E-02
17	rs876017	TREND	30_74	605_2289	3.781	1	5.2E-02
17	rs1010954	TREND	30_76	599_2311	3.714	1	5.4E-02
17	rs1130674	TREND	29_77	588_2282	2.936	1	8.7E-02
17	rs2070871	TREND	28_78	583_2327	2.576	1	1.1E-01
19	rs7412	TREND	8_86	193_2609	0.3449	1	5.6E-01
19	rs429358	TREND	17_77	509_2373	0.01136	1	9.2E-01
X	rs413536	TREND	13_79	586_2282	1.367	1	2.4E-01
X	rs3817939	TREND	1_95	53_2739	0.2287	1	6.3E-01
X	rs401597	TREND	24_66	843_2073	0.1316	1	7.2E-01
X	rs6048	TREND	27_65	813_2025	0.01298	1	9.1E-01

APPENDIX V TREND TEST RESULT IN BLEEDING (UPPSALA BLEEDERS).

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
1	rs2502804	TREND	1_35	598_1932	6.59	1	1.03E-02
1	rs2759328	TREND	9_35	262_2452	5.811	1	1.59E-02
1	rs4653436	TREND	5_33	773_1725	5.434	1	1.98E-02
1	rs3753305	TREND	26_16	1185_1485	5.318	1	2.11E-02
1	rs2307420	TREND	3_41	61_2567	3.716	1	5.39E-02
1	rs2854461	TREND	20_24	867_1809	3.555	1	5.94E-02
1	rs1894699	TREND	15_27	1298_1362	2.92	1	8.75E-02
1	rs3817268	TREND	7_37	737_1959	2.89	1	8.91E-02
1	rs2292566	TREND	2_34	359_2353	1.811	1	1.78E-01
1	rs3753663	TREND	4_40	128_2592	1.779	1	1.82E-01
1	rs6686001	TREND	2_32	363_2295	1.751	1	1.86E-01
1	rs2260863	TREND	9_33	826_1898	1.522	1	2.17E-01
1	rs3753661	TREND	5_39	180_2526	1.518	1	2.18E-01
1	rs2234922	TREND	5_33	564_2120	1.474	1	2.25E-01
1	rs2307424	TREND	12_32	883_1575	1.418	1	2.34E-01
1	rs2307418	TREND	4_38	436_2238	1.398	1	2.37E-01
1	rs4656687	TREND	10_32	863_1797	1.391	1	2.38E-01
1	rs6426089	TREND	23_17	1331_1373	1.09	1	2.97E-01
1	rs1557572	TREND	11_33	857_1847	0.895	1	3.44E-01
1	rs5878	TREND	16_28	822_1884	0.7291	1	3.93E-01
1	rs6029	TREND	5_37	449_2265	0.6587	1	4.17E-01
1	rs3003596	TREND	21_21	1152_1454	0.5587	1	4.55E-01
1	rs4233368	TREND	13_29	696_1952	0.4686	1	4.94E-01
1	rs9332504	TREND	2_42	187_2447	0.4545	1	5.00E-01
1	rs2671272	TREND	11_33	565_2117	0.4089	1	5.23E-01
1	rs4149223	TREND	15_21	1260_1450	0.3291	1	5.66E-01
1	rs6037	TREND	2_42	179_2503	0.3203	1	5.72E-01
1	rs2298909	TREND	10_28	756_1732	0.284	1	5.94E-01
1	rs9332618	TREND	7_37	363_2337	0.2297	1	6.32E-01
1	rs6025	TREND	2_40	173_2487	0.2114	1	6.46E-01
1	rs3766110	TREND	8_32	615_2055	0.2037	1	6.52E-01
1	rs2102663	TREND	5_31	434_2244	0.1411	1	7.07E-01
1	rs2671270	TREND	12_32	680_1994	0.08093	1	7.76E-01
1	rs721161	TREND	16_28	970_1552	0.07925	1	7.78E-01
1	rs4073054	TREND	13_25	977_1709	0.07482	1	7.84E-01
1	rs6427198	TREND	21_21	1207_1291	0.04531	1	8.31E-01
1	rs2740170	TREND	9_33	545_2157	0.04008	1	8.41E-01
1	rs2227590	TREND	4_38	275_2421	0.02088	1	8.85E-01
1	rs3753660	TREND	6_38	387_2313	0.01726	1	8.96E-01
1	rs1051740	TREND	11_27	746_1796	0.002906	1	9.57E-01
1	rs2227607	TREND	4_38	262_2434	0.00175	1	9.67E-01
1	rs1051741	TREND	4_36	273_2411	0.00131	1	9.71E-01

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
1	rs6035	TREND	3_37	197_2461	0.000456	1	9.83E-01
2	rs2069933	TREND	8_32	760_1930	1.328	1	2.49E-01
2	rs2069928	TREND	5_37	489_2221	1.11	1	2.92E-01
2	rs7568458	TREND	23_21	1194_1472	0.9993	1	3.18E-01
2	rs5936	TREND	9_33	743_1927	0.8396	1	3.60E-01
2	rs2069910	TREND	21_23	1133_1521	0.4276	1	5.13E-01
2	rs2069901	TREND	21_23	1176_1536	0.3417	1	5.59E-01
2	rs1799809	TREND	20_22	1163_1495	0.2465	1	6.20E-01
2	rs2028898	TREND	11_31	745_1965	0.03586	1	8.50E-01
2	rs2069919	TREND	16_26	1004_1700	0.01583	1	9.00E-01
2	rs12471762	TREND	5_39	441_2149	0.9696	1	3.25E-01
2	rs1686482	TREND	24_20	1268_1314	0.4751	1	4.91E-01
2	rs1734343	TREND	23_27	1283_1289	0.2791	1	5.97E-01
2	rs1198873	TREND	19_33	1029_1609	0.123	1	7.26E-01
2	rs1734346	TREND	20_32	1006_1564	0.009112	1	9.24E-01
2	rs1686447	TREND	21_33	999_1543	0.003446	1	9.53E-01
2	rs11904084	TREND	21_33	1004_1600	0.002282	1	9.62E-01
3	rs4857037	TREND	0_38	201_2429	3.126	1	7.70E-02
3	rs2472677	TREND	21_19	1007_1519	2.511	1	1.13E-01
3	rs8178633	TREND	0_34	135_2457	1.911	1	1.67E-01
3	rs1054191	TREND	4_38	410_2304	1.028	1	3.11E-01
3	rs2056530	TREND	6_38	517_2177	0.8682	1	3.51E-01
3	rs3814057	TREND	11_33	523_2127	0.705	1	4.01E-01
3	rs7643645	TREND	11_27	955_1743	0.6727	1	4.12E-01
3	rs9683303	TREND	10_24	919_1669	0.5358	1	4.64E-01
3	rs5013930	TREND	4_40	321_2283	0.401	1	5.27E-01
3	rs2461818	TREND	3_41	241_2259	0.3822	1	5.36E-01
3	rs8178607	TREND	8_32	624_1964	0.3458	1	5.57E-01
3	rs2276706	TREND	15_27	1072_1620	0.3044	1	5.81E-01
3	rs1464602	TREND	14_26	1049_1633	0.2823	1	5.95E-01
3	rs2472682	TREND	18_26	1012_1656	0.1654	1	6.84E-01
3	rs8178610	TREND	10_12	1308_1324	0.149	1	7.00E-01
3	rs3732357	TREND	14_30	911_1773	0.08567	1	7.70E-01
3	rs12721607	TREND	1_41	75_2585	0.02696	1	8.70E-01
3	rs3732359	TREND	7_21	702_2020	0.008655	1	9.26E-01
3	rs3732356	TREND	3_41	175_2415	0.000247	1	9.88E-01
3	rs4857343	TREND	6_36	387_2313	7.02E-05	1	9.93E-01
3	rs1107377	TREND	28_14	1300_1326	4.654	1	3.10E-02
3	rs4677875	TREND	13_39	958_1622	3.045	1	8.10E-02
3	rs3792366	TREND	12_36	937_1579	2.817	1	9.33E-02
3	rs702030	TREND	5_49	472_2130	2.787	1	9.51E-02
3	rs1078982	TREND	11_43	589_2045	0.1204	1	7.29E-01
3	rs836832	TREND	9_45	483_2127	0.1134	1	7.36E-01
7	rs2290228	TREND	2_40	389_2215	3.377	1	6.61E-02

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
7	rs10260862	TREND	4_34	535_2171	2.047	1	1.53E-01
7	rs28365094	TREND	8_36	324_2360	1.445	1	2.29E-01
7	rs2060717	TREND	1_43	174_2548	1.204	1	2.73E-01
7	rs339054	TREND	24_20	1261_1411	0.9556	1	3.28E-01
7	rs10267099	TREND	8_36	645_2053	0.7599	1	3.83E-01
7	rs3842	TREND	8_36	340_2162	0.747	1	3.87E-01
7	rs955000	TREND	3_41	272_2362	0.5649	1	4.52E-01
7	rs4646453	TREND	1_25	51_2671	0.5504	1	4.58E-01
7	rs11773597	TREND	4_38	170_2372	0.5283	1	4.67E-01
7	rs9282564	TREND	3_29	358_2282	0.4412	1	5.07E-01
7	rs2214102	TREND	4_28	264_2462	0.2789	1	5.97E-01
7	rs2242480	TREND	3_41	239_2407	0.2456	1	6.20E-01
7	rs2214101	TREND	3_33	161_2343	0.2256	1	6.35E-01
7	rs1202172	TREND	14_30	947_1749	0.1988	1	6.56E-01
7	rs11653	TREND	16_28	1021_1575	0.1615	1	6.88E-01
7	rs4646457	TREND	4_38	214_2502	0.1578	1	6.91E-01
7	rs2235040	TREND	4_40	293_2389	0.1561	1	6.93E-01
7	rs2307040	TREND	17_27	970_1730	0.1419	1	7.06E-01
7	rs28365083	TREND	0_10	27_2627	0.1038	1	7.47E-01
7	rs2235046	TREND	20_24	1162_1512	0.06756	1	7.95E-01
7	rs6976017	TREND	2_40	149_2549	0.04542	1	8.31E-01
7	rs1922240	TREND	14_26	910_1804	0.03713	1	8.47E-01
7	rs2235033	TREND	19_21	1313_1369	0.03105	1	8.60E-01
7	rs339057	TREND	19_25	1188_1506	0.01597	1	8.99E-01
7	rs776746	TREND	3_39	205_2485	0.0137	1	9.07E-01
7	rs339098	TREND	16_28	954_1726	0.01134	1	9.15E-01
7	rs4148737	TREND	20_24	1151_1407	0.003484	1	9.53E-01
7	GS30681	TREND	3_41	189_2519	0.001707	1	9.67E-01
7	rs10269104	TREND	17_37	491_2111	5.482	1	1.92E-02
7	rs4727005	TREND	17_37	532_2110	4.317	1	3.77E-02
7	rs10085877	TREND	16_38	508_2130	3.756	1	5.26E-02
7	rs10272564	TREND	14_34	522_2094	2.524	1	1.12E-01
7	rs6464929	TREND	14_34	526_2102	2.491	1	1.15E-01
7	rs6464930	TREND	12_34	482_2138	1.769	1	1.84E-01
7	rs1551927	TREND	8_46	368_2264	0.03095	1	8.60E-01
9	rs1687390	TREND	4_28	122_2584	4.834	1	2.79E-02
9	rs17230081	TREND	5_35	541_2155	1.372	1	2.42E-01
9	rs2787337	TREND	12_30	802_1754	0.1565	1	6.92E-01
9	rs10982151	TREND	6_38	361_2313	0.000698	1	9.79E-01
10	rs1557044	TREND	10_34	382_2328	2.656	1	1.03E-01
10	rs1058930	TREND	0_44	144_2502	2.488	1	1.15E-01
10	rs2475376	TREND	10_34	413_2271	1.684	1	1.94E-01
10	rs3814637	TREND	1_43	190_2452	1.564	1	2.11E-01
10	rs1057910	TREND	1_43	192_2532	1.501	1	2.21E-01

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
10	rs17878459	TREND	0_44	83_2579	1.462	1	2.27E-01
10	rs9332197	TREND	1_35	203_2509	1.073	1	3.00E-01
10	rs2860905	TREND	7_37	588_2116	0.8823	1	3.48E-01
10	rs12248560	TREND	6_38	500_2206	0.6843	1	4.08E-01
10	rs4244284	TREND	16_26	1174_1472	0.6544	1	4.19E-01
10	rs947173	TREND	11_33	801_1911	0.4356	1	5.09E-01
10	rs1856908	TREND	21_23	1173_1525	0.3126	1	5.76E-01
10	rs4244285	TREND	5_33	423_2225	0.2148	1	6.43E-01
10	rs1058932	TREND	6_38	428_2200	0.2142	1	6.44E-01
10	rs4417205	TREND	8_36	424_2286	0.2048	1	6.51E-01
10	rs1799853	TREND	6_38	319_2405	0.1519	1	6.97E-01
10	rs2275620	TREND	17_25	1051_1439	0.04944	1	8.24E-01
10	rs11572080	TREND	4_40	263_2339	0.04908	1	8.25E-01
10	rs2901783	TREND	9_33	595_2107	0.008186	1	9.28E-01
10	rs2860840	TREND	17_27	1002_1558	0.004778	1	9.45E-01
10	rs1536430	TREND	0_0	61_2651	NA	NA	NA
11	rs2282687	TREND	2_42	354_2354	2.742	1	9.77E-02
11	rs2070852	TREND	10_32	818_1900	0.7735	1	3.79E-01
11	rs2070851	TREND	7_35	579_2083	0.597	1	4.40E-01
11	rs3136516	TREND	22_22	1178_1474	0.5315	1	4.66E-01
11	rs5898	TREND	5_39	243_2455	0.2829	1	5.95E-01
11	rs3136435	TREND	3_39	196_2506	0.000752	1	9.78E-01
13	rs3093230	TREND	0_30	683_2013	10.49	1	1.20E-03
13	rs3093261	TREND	25_19	1150_1520	3.336	1	6.78E-02
13	rs2273971	TREND	0_42	167_2533	2.7	1	1.00E-01
13	rs3213005	TREND	4_40	116_2554	2.217	1	1.37E-01
13	rs3024718	TREND	4_40	465_2241	2.025	1	1.55E-01
13	rs3024746	TREND	6_38	552_2146	1.22	1	2.69E-01
13	rs7997328	TREND	16_28	761_1871	1.161	1	2.81E-01
13	rs776905	TREND	2_34	285_2415	0.9463	1	3.31E-01
13	rs9577874	TREND	15_29	1047_1521	0.8483	1	3.57E-01
13	rs2774030	TREND	15_29	1003_1693	0.1837	1	6.68E-01
13	rs3212998	TREND	2_40	171_2521	0.182	1	6.70E-01
13	rs693335	TREND	12_20	1023_1521	0.1006	1	7.51E-01
13	rs2480948	TREND	8_36	542_2162	0.09427	1	7.59E-01
13	rs6046	TREND	4_40	279_2407	0.07688	1	7.82E-01
13	rs3211764	TREND	16_20	1258_1444	0.06444	1	8.00E-01
13	rs753057	TREND	2_42	146_2556	0.06339	1	8.01E-01
13	rs5960	TREND	6_36	353_2307	0.03694	1	8.48E-01
13	rs2026160	TREND	13_31	757_1745	0.00964	1	9.22E-01
13	rs6602908	TREND	16_24	1029_1561	0.001235	1	9.72E-01
13	rs3093229	TREND	10_32	606_1958	0.000688	1	9.79E-01
14	rs230704	TREND	17_27	762_1926	2.201	1	1.38E-01
15	rs2470893	TREND	19_25	914_1760	1.507	1	2.20E-01

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
15	rs2069522	TREND	0_44	82_2612	1.425	1	2.33E-01
15	rs2470890	TREND	11_33	884_1766	1.348	1	2.46E-01
15	rs1048943	TREND	0_40	64_2468	0.9685	1	3.25E-01
15	rs762551	TREND	9_35	696_1884	0.9356	1	3.33E-01
15	rs4646425	TREND	0_44	47_2539	0.8293	1	3.63E-01
15	rs4646421	TREND	2_42	222_2490	0.7474	1	3.87E-01
15	rs2606345	TREND	13_31	782_1678	0.09722	1	7.55E-01
15	rs11070411	TREND	6_42	421_2173	0.4963	1	4.81E-01
15	rs7175032	TREND	3_49	201_2437	0.251	1	6.16E-01
15	rs8040336	TREND	22_30	1166_1458	0.09405	1	7.59E-01
15	rs10163054	TREND	4_50	180_2370	0.009616	1	9.22E-01
16	rs11150606	TREND	3_41	75_2633	2.584	1	1.08E-01
16	rs2917671	TREND	10_30	910_1590	2.204	1	1.38E-01
16	rs9923231	TREND	11_25	1060_1622	1.211	1	2.71E-01
16	rs2359612	TREND	16_28	1049_1609	0.1744	1	6.76E-01
16	rs689453	TREND	3_41	214_2452	0.08646	1	7.69E-01
16	rs7294	TREND	15_25	1059_1625	0.06251	1	8.03E-01
16	rs1800566	TREND	7_35	449_2187	0.003973	1	9.50E-01
16	rs2685127	TREND	10_44	365_2231	0.8631	1	3.53E-01
16	rs400037	TREND	13_41	632_1976	0.000701	1	9.79E-01
17	rs1799919	TREND	15_19	541_2083	11.16	1	8.35E-04
17	rs876017	TREND	15_37	546_2082	1.986	1	1.59E-01
17	rs1130674	TREND	15_39	530_2076	1.788	1	1.81E-01
17	rs1010954	TREND	15_39	541_2103	1.726	1	1.89E-01
17	rs1533756	TREND	17_37	1060_1578	1.687	1	1.94E-01
17	rs2070871	TREND	14_40	526_2116	1.185	1	2.76E-01
19	rs429358	TREND	10_34	464_2176	0.8016	1	3.71E-01
19	rs7412	TREND	4_38	180_2396	0.378	1	5.39E-01
X	rs413536	TREND	3_37	531_2111	2.455	1	1.17E-01
X	rs401597	TREND	10_32	757_1927	0.2442	1	6.21E-01
X	rs6048	TREND	14_30	738_1870	0.161	1	6.88E-01
X	rs3817939	TREND	1_43	51_2499	0.01019	1	9.20E-01

APPENDIX VI TREND TEST RESULT IN BLEEDING (WARG BLEEDERS).

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
1	rs5878	TREND	9_43	822_1884	4.114	1	0.04253
1	rs2227590	TREND	1_51	275_2421	3.896	1	0.04841
1	rs3753661	TREND	7_45	180_2526	3.715	1	0.05393
1	rs2759328	TREND	1_51	262_2452	3.589	1	0.05816
1	rs3753663	TREND	0_52	128_2592	2.482	1	0.1151
1	rs6025	TREND	1_49	173_2487	1.69	1	0.1936
1	rs2307418	TREND	5_45	436_2238	1.436	1	0.2308
1	rs2502804	TREND	7_39	598_1932	1.364	1	0.2429
1	rs6029	TREND	11_39	449_2265	1.072	1	0.3005
1	rs3766110	TREND	15_37	615_2055	0.9644	1	0.3261
1	rs3753660	TREND	5_47	387_2313	0.9404	1	0.3322
1	rs6427198	TREND	19_27	1207_1291	0.8588	1	0.3541
1	rs2102663	TREND	6_46	434_2244	0.8241	1	0.364
1	rs2671270	TREND	10_40	680_1994	0.8005	1	0.3709
1	rs6686001	TREND	9_41	363_2295	0.7888	1	0.3745
1	rs2671272	TREND	13_37	565_2117	0.7295	1	0.393
1	rs2740170	TREND	8_44	545_2157	0.7168	1	0.3972
1	rs2260863	TREND	13_39	826_1898	0.6691	1	0.4134
1	rs6037	TREND	2_50	179_2503	0.6677	1	0.4138
1	rs9332504	TREND	5_45	187_2447	0.6497	1	0.4202
1	rs2307424	TREND	14_32	883_1575	0.5952	1	0.4404
1	rs2234922	TREND	13_39	564_2120	0.5163	1	0.4724
1	rs4656687	TREND	14_36	863_1797	0.4382	1	0.508
1	rs4149223	TREND	21_29	1260_1450	0.3956	1	0.5294
1	rs4653436	TREND	13_35	773_1725	0.3212	1	0.5709
1	rs3003596	TREND	24_26	1152_1454	0.2831	1	0.5947
1	rs4073054	TREND	20_30	977_1709	0.2749	1	0.6001
1	rs6035	TREND	3_49	197_2461	0.2043	1	0.6513
1	rs2292566	TREND	8_44	359_2353	0.2022	1	0.653
1	rs2227607	TREND	6_46	262_2434	0.1891	1	0.6637
1	rs1051741	TREND	6_44	273_2411	0.1885	1	0.6642
1	rs3753305	TREND	22_30	1185_1485	0.09183	1	0.7619
1	rs721161	TREND	21_31	970_1552	0.07798	1	0.78
1	rs4233368	TREND	14_36	696_1952	0.07551	1	0.7835
1	rs2854461	TREND	17_33	867_1809	0.06094	1	0.805
1	rs2298909	TREND	16_34	756_1732	0.05868	1	0.8086
1	rs1051740	TREND	16_36	746_1796	0.05042	1	0.8223
1	rs1557572	TREND	17_35	857_1847	0.02346	1	0.8783
1	rs2307420	TREND	1_49	61_2567	0.02217	1	0.8816
1	rs9332618	TREND	7_43	363_2337	0.01334	1	0.9081
1	rs6426089	TREND	25_25	1331_1373	0.01194	1	0.913
1	rs1894699	TREND	25_27	1298_1362	0.01083	1	0.9171

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
1	rs3817268	TREND	12_32	737_1959	9.09E-05	1	0.9924
2	rs2069933	TREND	8_42	760_1930	3.675	1	0.05524
2	rs5936	TREND	10_42	743_1927	1.878	1	0.1706
2	rs2028898	TREND	17_33	745_1965	1.065	1	0.3022
2	rs7568458	TREND	25_23	1194_1472	1.045	1	0.3066
2	rs2069928	TREND	8_44	489_2221	0.2568	1	0.6123
2	rs2069919	TREND	21_31	1004_1700	0.2221	1	0.6375
2	rs1799809	TREND	21_29	1163_1495	0.06012	1	0.8063
2	rs2069901	TREND	22_30	1176_1536	0.02338	1	0.8785
2	rs2069910	TREND	21_27	1133_1521	0.02067	1	0.8857
2	rs1686482	TREND	35_15	1268_1314	7.881	1	4.99E-03
2	rs1198873	TREND	30_22	1029_1609	7.01	1	8.10E-03
2	rs1734343	TREND	17_35	1283_1289	5.647	1	1.75E-02
2	rs11904084	TREND	28_24	1004_1600	4.601	1	3.20E-02
2	rs12471762	TREND	3_49	441_2149	4.555	1	3.28E-02
2	rs1734346	TREND	27_23	1006_1564	4.12	1	4.24E-02
2	rs1686447	TREND	27_23	999_1543	4.053	1	4.41E-02
3	rs2056530	TREND	5_47	517_2177	3.033	1	0.08161
3	rs1054191	TREND	12_40	410_2304	2.552	1	0.1101
3	rs2276706	TREND	16_36	1072_1620	1.829	1	0.1763
3	rs7643645	TREND	23_29	955_1743	1.715	1	0.1903
3	rs4857343	TREND	10_38	387_2313	1.479	1	0.224
3	rs12721607	TREND	0_52	75_2585	1.395	1	0.2376
3	rs2472677	TREND	25_27	1007_1519	1.372	1	0.2415
3	rs5013930	TREND	9_43	321_2283	1.098	1	0.2948
3	rs8178633	TREND	1_49	135_2457	1.059	1	0.3035
3	rs8178610	TREND	22_28	1308_1324	0.6144	1	0.4331
3	rs8178607	TREND	14_36	624_1964	0.3857	1	0.5346
3	rs3732357	TREND	15_35	911_1773	0.337	1	0.5616
3	rs3814057	TREND	8_40	523_2127	0.2651	1	0.6066
3	rs9683303	TREND	16_34	919_1669	0.2604	1	0.6098
3	rs4857037	TREND	3_47	201_2429	0.1878	1	0.6648
3	rs2472682	TREND	21_31	1012_1656	0.1338	1	0.7146
3	rs3732356	TREND	4_48	175_2415	0.06669	1	0.7962
3	rs1464602	TREND	20_30	1049_1633	0.01654	1	0.8977
3	rs3732359	TREND	13_39	702_2020	0.0161	1	0.899
3	rs2461818	TREND	5_45	241_2259	0.007027	1	0.9332
3	rs1078982	TREND	9_41	589_2045	0.5358	1	4.64E-01
3	rs702030	TREND	8_44	472_2130	0.2591	1	6.11E-01
3	rs836832	TREND	11_41	483_2127	0.2252	1	6.35E-01
3	rs1107377	TREND	27_25	1300_1326	0.1134	1	7.36E-01
3	rs3792366	TREND	18_32	937_1579	0.03011	1	8.62E-01
3	rs4677875	TREND	19_33	958_1622	0.00727	1	9.32E-01
7	rs2214102	TREND	10_42	264_2462	5.047	1	0.02467

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
7	rs2235040	TREND	1_51	293_2389	4.471	1	0.03448
7	rs11773597	TREND	0_50	170_2372	3.559	1	0.05921
7	rs1922240	TREND	22_30	910_1804	1.705	1	0.1917
7	rs339054	TREND	20_32	1261_1411	1.581	1	0.2087
7	rs2060717	TREND	1_49	174_2548	1.555	1	0.2125
7	rs4148737	TREND	28_24	1151_1407	1.519	1	0.2178
7	rs2242480	TREND	2_48	239_2407	1.445	1	0.2294
7	rs10260862	TREND	7_45	535_2171	1.3	1	0.2543
7	rs11653	TREND	15_31	1021_1575	0.8596	1	0.3538
7	GS30681	TREND	2_50	189_2519	0.7656	1	0.3816
7	rs339098	TREND	15_35	954_1726	0.6802	1	0.4095
7	rs28365094	TREND	8_44	324_2360	0.5016	1	0.4788
7	rs9282564	TREND	5_45	358_2282	0.4965	1	0.481
7	rs2214101	TREND	2_46	161_2343	0.4269	1	0.5135
7	rs339057	TREND	24_26	1188_1506	0.3259	1	0.5681
7	rs6976017	TREND	2_50	149_2549	0.273	1	0.6013
7	rs776746	TREND	3_49	205_2485	0.2545	1	0.6139
7	rs2307040	TREND	17_35	970_1730	0.2372	1	0.6263
7	rs955000	TREND	6_44	272_2362	0.1438	1	0.7045
7	rs10267099	TREND	11_37	645_2053	0.02476	1	0.875
7	rs2235033	TREND	26_26	1313_1369	0.02067	1	0.8857
7	rs2290228	TREND	8_44	389_2215	0.008011	1	0.9287
7	rs2235046	TREND	22_28	1162_1512	0.005657	1	0.94
7	rs1202172	TREND	18_34	947_1749	0.005598	1	0.9404
7	rs4646457	TREND	4_48	214_2502	0.002522	1	0.9599
7	rs4646453	TREND	1_51	51_2671	0.000692	1	0.979
7	rs3842	TREND	7_45	340_2162	0.000686	1	0.9791
7	rs10272564	TREND	18_34	522_2094	6.778	1	9.23E-03
7	rs10085877	TREND	17_35	508_2130	5.98	1	1.45E-02
7	rs6464929	TREND	17_35	526_2102	5.096	1	2.40E-02
7	rs4727005	TREND	17_35	532_2110	5.023	1	2.50E-02
7	rs6464930	TREND	15_37	482_2138	3.663	1	5.56E-02
7	rs1551927	TREND	12_40	368_2264	3.49	1	6.17E-02
7	rs10269104	TREND	15_37	491_2111	3.292	1	6.96E-02
9	rs2787337	TREND	11_39	802_1754	2.059	1	0.1513
9	rs17230081	TREND	8_44	541_2155	0.6775	1	0.4105
9	rs1687390	TREND	3_49	122_2584	0.1968	1	0.6574
9	rs10982151	TREND	7_43	361_2313	0.01061	1	0.918
10	rs1536430	TREND	4_48	61_2651	6.534	1	0.01058
10	rs1799853	TREND	1_51	319_2405	4.708	1	0.03002
10	rs11572080	TREND	1_49	263_2339	3.572	1	0.05877
10	rs947173	TREND	10_42	801_1911	2.641	1	0.1042
10	rs1058930	TREND	5_45	144_2502	1.928	1	0.1649
10	rs3814637	TREND	6_44	190_2452	1.647	1	0.1994

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
10	rs1057910	TREND	6_46	192_2532	1.521	1	0.2175
10	rs2860840	TREND	15_33	1002_1558	1.28	1	0.2579
10	rs2860905	TREND	8_44	588_2116	1.236	1	0.2662
10	rs1557044	TREND	10_42	382_2328	1.103	1	0.2935
10	rs4244285	TREND	6_46	423_2225	0.7309	1	0.3926
10	rs4417205	TREND	6_46	424_2286	0.6389	1	0.4241
10	rs4244284	TREND	25_25	1174_1472	0.6238	1	0.4296
10	rs1856908	TREND	25_27	1173_1525	0.4314	1	0.5113
10	rs12248560	TREND	11_41	500_2206	0.2437	1	0.6216
10	rs2475376	TREND	9_43	413_2271	0.1362	1	0.7121
10	rs17878459	TREND	2_48	83_2579	0.13	1	0.7185
10	rs2901783	TREND	10_40	595_2107	0.1133	1	0.7364
10	rs1058932	TREND	9_43	428_2200	0.03728	1	0.8469
10	rs2275620	TREND	20_26	1051_1439	0.02891	1	0.865
10	rs9332197	TREND	4_48	203_2509	0.002964	1	0.9566
11	rs5898	TREND	10_42	243_2455	6.202	1	0.01276
11	rs2282687	TREND	3_49	354_2354	2.372	1	0.1235
11	rs3136516	TREND	26_22	1178_1474	1.757	1	0.185
11	rs2070851	TREND	14_38	579_2083	0.756	1	0.3846
11	rs2070852	TREND	17_35	818_1900	0.1616	1	0.6877
11	rs3136435	TREND	4_48	196_2506	0.01445	1	0.9043
13	rs3212998	TREND	7_45	171_2521	4.397	1	0.036
13	rs6046	TREND	1_51	279_2407	3.898	1	0.04836
13	rs2480948	TREND	16_36	542_2162	3.675	1	0.05522
13	rs5960	TREND	11_39	353_2307	3.205	1	0.07341
13	rs2273971	TREND	6_46	167_2533	2.435	1	0.1187
13	rs3024718	TREND	12_40	465_2241	1.253	1	0.263
13	rs2774030	TREND	16_36	1003_1693	0.9281	1	0.3353
13	rs693335	TREND	22_26	1023_1521	0.6448	1	0.422
13	rs2026160	TREND	12_36	757_1745	0.5728	1	0.4492
13	rs3211764	TREND	20_28	1258_1444	0.4576	1	0.4988
13	rs776905	TREND	4_48	285_2415	0.4456	1	0.5044
13	rs753057	TREND	2_50	146_2556	0.2467	1	0.6194
13	rs3093261	TREND	19_29	1150_1520	0.2357	1	0.6273
13	rs3093229	TREND	13_37	606_1958	0.1507	1	0.6979
13	rs3024746	TREND	9_39	552_2146	0.08378	1	0.7722
13	rs3093230	TREND	14_38	683_2013	0.0709	1	0.79
13	rs7997328	TREND	14_36	761_1871	0.01999	1	0.8876
13	rs9577874	TREND	20_28	1047_1521	0.0168	1	0.8969
13	rs3213005	TREND	2_48	116_2554	0.01346	1	0.9076
13	rs6602908	TREND	19_29	1029_1561	0.000437	1	0.9833
14	rs230704	TREND	9_43	762_1926	3.004	1	0.08304
15	rs1048943	TREND	4_46	64_2468	5.235	1	0.02213
15	rs2069522	TREND	3_49	82_2612	1.305	1	0.2533

CHR	SNP	TEST	AFF	UNAFF	CHISQ	DF	P-value
15	rs4646421	TREND	6_46	222_2490	0.7394	1	0.3899
15	rs2470893	TREND	15_35	914_1760	0.3702	1	0.5429
15	rs2470890	TREND	17_31	884_1766	0.08903	1	0.7654
15	rs762551	TREND	13_33	696_1884	0.03777	1	0.8459
15	rs2606345	TREND	14_32	782_1678	0.03688	1	0.8477
15	rs4646425	TREND	1_49	47_2539	0.009313	1	0.9231
15	rs11070411	TREND	14_38	421_2173	4.333	1	3.74E-02
15	rs10163054	TREND	1_45	180_2370	1.63	1	2.02E-01
15	rs7175032	TREND	2_50	201_2437	1.048	1	3.06E-01
15	rs8040336	TREND	24_28	1166_1458	0.06104	1	8.05E-01
16	rs9923231	TREND	17_33	1060_1622	0.6353	1	0.4254
16	rs1800566	TREND	10_40	449_2187	0.3068	1	0.5797
16	rs2359612	TREND	18_32	1049_1609	0.247	1	0.6192
16	rs7294	TREND	22_30	1059_1625	0.1707	1	0.6795
16	rs11150606	TREND	1_51	75_2633	0.1367	1	0.7116
16	rs2917671	TREND	16_30	910_1590	0.05061	1	0.822
16	rs689453	TREND	4_48	214_2452	0.007803	1	0.9296
16	rs400037	TREND	14_38	632_1976	0.1931	1	6.60E-01
16	rs2685127	TREND	7_45	365_2231	0.01514	1	9.02E-01
17	rs1533756	TREND	14_38	1060_1578	3.777	1	5.20E-02
17	rs1010954	TREND	15_37	541_2103	2.193	1	1.39E-01
17	rs876017	TREND	15_37	546_2082	1.995	1	1.58E-01
17	rs2070871	TREND	14_38	526_2116	1.559	1	2.12E-01
17	rs1130674	TREND	14_38	530_2076	1.356	1	2.44E-01
17	rs1799919	TREND	14_38	541_2083	1.231	1	2.67E-01
19	rs429358	TREND	7_43	464_2176	0.4394	1	0.5074
19	rs7412	TREND	4_48	180_2396	0.03611	1	0.8493
X	rs3817939	TREND	0_52	51_2499	0.6548	1	0.4184
X	rs6048	TREND	13_35	738_1870	0.02082	1	0.8853
X	rs413536	TREND	10_42	531_2111	0.01495	1	0.9027
X	rs401597	TREND	14_34	757_1927	0.01329	1	0.9082

APPENDIX VII SEQUENCE VARIANTS IDENTIFIED WITH EXON RE-SEQUENCING.

SNP	Gene	allele	CHR	Position	MAF (Patient)	MAF(CEPH)	Consequence
rs769452	APOE	C|T	19	50102950		0.01	NON_SYNONYMOUS_CODING
rs429358	APOE	C|T	19	50103781	0.16	0.12	NON_SYNONYMOUS_CODING
NT_011109.15_17677924	APOE	C|T	19	50101546	0.01		INTRONIC
NT_011109.15_17680755	APOE	T|C	19	50104377	0.01		REGULATORY_REGION
NT_007933.14_53593466	CALU	A|G	7	128197126	0.01	0.01	3PRIME_UTR
NT_007933.14_53594540	CALU	G|A	7	128198200		0.01	3PRIME_UTR
NT_007933.14_53591356	CALU	C|T	7	128195016	0.01	0.01	INTRONIC
NT_007933.14_53563005	CALU	G|T	7	128166665		0.01	5PRIME_UTR
NT_007933.14_53563186	CALU	A|G	7	128166846		0.02	INTRONIC
NT_007933.13_53564648	CALU	T|A	7	128175735		0.02	INTRONIC
NT_007933.13_53576255	CALU	T|G	7	128187342		0.02	INTRONIC
NT_007933.14_53593458	CALU	T|C	7	128197118		0.02	3PRIME_UTR
NT_007933.14_53571635	CALU	G|A	7	128175295	0.01	0.04	INTRONIC
NT_007933.14_53578384	CALU	T|C	7	128182044	0.01	0.05	INTRONIC
rs2290228	CALU	A|G	7	128175884	0.05	0.09	NON_SYNONYMOUS_CODING
rs8597	CALU	T|C	7	128198658	0.16	0.11	3PRIME_UTR
rs1043595	CALU	A|G	7	128197248	0.26	0.20	3PRIME_UTR
rs12538139	CALU	A|G	7	128175777	0.36	0.35	INTRONIC
rs2307040	CALU	T|C	7	128181842	0.37	0.36	NON_SYNONYMOUS_CODING
rs1043550	CALU	G|A	7	128196461	0.34	0.37	3PRIME_UTR
rs11653	CALU	A|T	7	128196816	0.36	0.39	3PRIME_UTR
NT_007933.13_53575380	CALU	C|T	7	128186467	0.38	0.44	INTRONIC
NT_007933.14_53593768	CALU	C|T	7	128197428	0.01		3PRIME_UTR
rs9332132	CYP2C9	A|G	10	96699216	0.03	0.02	INTRONIC
rs1057911	CYP2C9	T|A	10	96738727	0.07	0.03	SYNONYMOUS_CODING
NT_030059.11_15450889	CYP2C9	C|T	10	96692353	0.06	0.03	INTRONIC

SNP	Gene	allele	CHR	Position	MAF (Patient)	MAF(CEPH)	Consequence
rs1057910	CYP2C9	C|A	10	96731043	0.04	0.03	NON_SYNONYMOUS_CODING
rs9332230	CYP2C9	T|A	10	96735974	0.04	0.04	INTRONIC
rs9332245	CYP2C9	A|T	10	96739171		0.04	DOWNSTREAM
NT_030059.11_15450863	CYP2C9	T|C	10	96692327	0.17	0.07	INTRONIC
rs9332197	CYP2C9	C|T	10	96730898	0.03	0.07	INTRONIC
rs9332174	CYP2C9	G|A	10	96722087	0.15	0.09	INTRONIC
rs9332242	CYP2C9	G|C	10	96738883	0.11	0.13	3PRIME_UTR
NT_030059.12_15451082	CYP2C9	C|T	10	96692546	0.10	0.15	INTRONIC
NT_030059.12_15450998	CYP2C9	G|A	10	96692462	0.16	0.15	INTRONIC
rs9332104	CYP2C9	C|T	10	96688680	0.16	0.16	INTRONIC
rs9332172	CYP2C9	G|A	10	96721778	0.16	0.16	INTRONIC
rs2860905	CYP2C9	A|G	10	96692285	0.15	0.21	INTRONIC
rs1934969	CYP2C9	T|A	10	96738485		0.29	INTRONIC
NT_030059.12_15451550	CYP2C9	A|T	10	96693014	0.01		INTRONIC
rs28371685	CYP2C9	T|C	10	96730971	0.03		NON_SYNONYMOUS_CODING
rs28371678	CYP2C9	C|T	10	96692542	0.06		INTRONIC
rs9332119	CYP2C9	C|G	10	96691591	0.09		INTRONIC
rs2292568	EPHX1	T|C	1	224094282	0.08	0.01	SYNONYMOUS_CODING
NT_004559.11_2203769	EPHX1	A|C	1	224094198		0.01	NON_SYNONYMOUS_CODING
NT_004559.11_2192461	EPHX1	G|A	1	224082890	0.10	0.02	INTRONIC
rs2234698	EPHX1	C|T	1	224086123	0.04	0.03	SYNONYMOUS_CODING
rs3738047	EPHX1	A|G	1	224083256	0.08	0.04	INTRONIC
NT_004559.11_2203742	EPHX1	T|C	1	224094171	0.03	0.06	SYNONYMOUS_CODING
rs3738040	EPHX1	A|G	1	224079664	0.08	0.06	INTRONIC
rs2292567	EPHX1	A|G	1	224093239	0.10	0.08	DOWNSTREAM
rs1051741	EPHX1	T|C	1	224098852	0.07	0.13	SYNONYMOUS_CODING
rs2292566	EPHX1	A|G	1	224086276	0.11	0.14	SYNONYMOUS_CODING
rs2671266	EPHX1	T|C	1	224093819	0.10	0.16	DOWNSTREAM

SNP	Gene	allele	CHR	Position	MAF (Patient)	MAF(CEPH)	Consequence
rs2234922	EPHX1	G|A	1	224093029	0.12	0.18	NON_SYNONYMOUS_CODING
rs1051740	EPHX1	C|T	1	224086256	0.27	0.26	NON_SYNONYMOUS_CODING
rs2260863	EPHX1	G|C	1	224086397	0.28	0.28	INTRONIC
rs4149225	EPHX1	G|A	1	224093824	0.21	0.36	DOWNSTREAM
NT_004559.11_2192575	EPHX1	A|G	1	224083004	0.37	0.42	INTRONIC
NT_004559.11_2192591	EPHX1	A|G	1	224083020	0.37	0.42	INTRONIC
NT_004559.12_2192885	EPHX1	T|G	1	224083314	0.01		INTRONIC
NT_004559.12_2193017	EPHX1	T|C	1	224083446	0.01		INTRONIC
rs2234699	EPHX1	C|T	1	224086222	0.01		SYNONYMOUS_CODING
NT_004559.12_2208718	EPHX1	A|G	1	224099147	0.02		DOWNSTREAM
NT_004487.17_19920127	F5	A|G	1	167778027	0.01	0.01	SYNONYMOUS_CODING
NT_004487.17_19964604	F5	G|A	1	167822288	0.01	0.01	5PRIME_UTR
NT_004487.17_19907940	F5	C|T	1	167765624		0.01	NON_SYNONYMOUS_CODING
NT_004487.17_19937493	F5	A|G	1	167795177		0.01	INTRONIC
NT_004487.17_19960900	F5	T|C	1	167818584		0.01	INTRONIC
NT_004668.16_8060933	F5	T|A	1	167822085		0.01	INTRONIC
rs6024	F5	C|T	1	167778663	0.05	0.02	SYNONYMOUS_CODING
rs6007	F5	G|T	1	167776363		0.02	NON_SYNONYMOUS_CODING
rs6023	F5	T|C	1	167795008	0.08	0.04	INTRONIC
rs6009	F5	T|C	1	167765458	0.13	0.05	INTRONIC
rs2239854	F5	A|G	1	167792432	0.07	0.06	INTRONIC
rs9332609	F5	G|T	1	167776107	0.06	0.06	INTRONIC
rs6019	F5	G|C	1	167808137	0.02	0.07	NON_SYNONYMOUS_CODING
rs7523043	F5	T|G	1	167796701	0.03	0.07	INTRONIC
rs6015	F5	A|G	1	167786518	0.09	0.08	SYNONYMOUS_CODING
rs6427201	F5	T|C	1	167795346	0.10	0.19	INTRONIC
rs6022	F5	A|C	1	167796450	0.11	0.19	SYNONYMOUS_CODING
rs6029	F5	T|C	1	167796597	0.11	0.19	SYNONYMOUS_CODING

SNP	Gene	allele	CHR	Position	MAF (Patient)	MAF(CEPH)	Consequence
rs7534848	F5	C|T	1	167796717	0.11	0.19	INTRONIC
rs7545236	F5	C|A	1	167796694	0.11	0.19	INTRONIC
rs6012	F5	T|C	1	167795204	0.10	0.20	INTRONIC
GS30742	F5	T|C	1	167822674	0.12	0.21	UPSTREAM
rs6021	F5	C|T	1	167778651	0.21	0.21	SYNONYMOUS_CODING
rs9332643	F5	T|C	1	167759300	0.22	0.22	INTRONIC
rs6028	F5	C|T	1	167818306	0.34	0.24	SYNONYMOUS_CODING
rs6016	F5	A|G	1	167778744	0.20	0.24	SYNONYMOUS_CODING
rs6662696	F5	A|G	1	167779275	0.21	0.24	INTRONIC
rs2239851	F5	A|C	1	167779121	0.19	0.25	INTRONIC
rs6662593	F5	A|G	1	167779218	0.19	0.25	INTRONIC
rs6017	F5	G|A	1	167778717	0.20	0.25	SYNONYMOUS_CODING
rs6675244	F5	C|T	1	167779186	0.20	0.25	INTRONIC
rs6032	F5	C|T	1	167778179	0.21	0.25	NON_SYNONYMOUS_CODING
rs4524	F5	C|T	1	167778379	0.20	0.26	NON_SYNONYMOUS_CODING
rs9332635	F5	C|T	1	167762057	0.19	0.27	INTRONIC
rs4525	F5	C|T	1	167778358	0.22	0.27	NON_SYNONYMOUS_CODING
rs2301515	F5	T|C	1	167760820	0.23	0.28	INTRONIC
rs2239852	F5	T|C	1	167779148	0.26	0.31	INTRONIC
rs6030	F5	C|T	1	167765599	0.27	0.31	NON_SYNONYMOUS_CODING
rs6686805	F5	C|A	1	167779267	0.28	0.32	INTRONIC
rs12131397	F5	A|C	1	167760577	0.50	0.43	INTRONIC
NT_004487.17_19920291	F5	T|C	1	167777863	0.01		NON_SYNONYMOUS_CODING
NT_004487.17_19921046	F5	C|T	1	167778730	0.01		NON_SYNONYMOUS_CODING
NT_004487.17_19928806	F5	C|A	1	167786490	0.01		INTRONIC
rs6034	F5	C|G	1	167765644	0.01		NON_SYNONYMOUS_CODING
rs9332658	F5	G|A	1	167756570	0.01		INTRONIC
NT_004487.17_19892172	F5	C|T	1	167750185	0.06		NON_SYNONYMOUS_CODING

SNP	Gene	allele	CHR	Position	MAF (Patient)	MAF(CEPH)	Consequence
NT_022184.13_64593799	GGCX	G|A	2	85631372		0.01	NON_SYNONYMOUS_CODING
NT_022184.13_64598759	GGCX	A|G	2	85636332		0.01	INTRONIC
NT_022184.13_64603901	GGCX	A|G	2	85641474		0.01	SYNONYMOUS_CODING
NT_022184.14_64592035	GGCX	G|C	2	85629606		0.01	DOWNSTREAM
NT_022184.14_64594133	GGCX	T|C	2	85631704		0.01	SYNONYMOUS_CODING
NT_022184.14_64595619	GGCX	G|C	2	85633190		0.01	NON_SYNONYMOUS_CODING
rs13406935	GGCX	T|A	2	85629519	0.11	0.08	DOWNSTREAM
rs11676382	GGCX	G|C	2	85631144	0.12	0.09	INTRONIC
rs10179904	GGCX	A|G	2	85633618	0.10	0.09	SYNONYMOUS_CODING
rs1254898	GGCX	T|C	2	85641651	0.31	0.27	INTRONIC
rs2592551	GGCX	A|G	2	85633642	0.25	0.35	SYNONYMOUS_CODING
rs2028898	GGCX	A|G	2	85630781	0.26	0.35	INTRONIC
rs699664	GGCX	T|C	2	85634047	0.14	0.39	NON_SYNONYMOUS_CODING
rs7568458	GGCX	A|T	2	85641686	0.42	0.42	INTRONIC
NT_022184.14_64595668	GGCX	G|A	2	85633239	0.01		INTRONIC
NT_022184.14_64592216	GGCX	G|T	2	85629787	0.04		3PRIME_UTR
NT_022184.14_64594203	GGCX	G|C	2	85631774	0.05		INTRONIC
rs6751560	GGCX	T|C	2	85639585	0.06		SYNONYMOUS_CODING
rs1254896	GGCX	T|G	2	85634829	0.07		SYNONYMOUS_CODING
NT_010498.14_18473923	NQO1	A|G	16	68309507	0.03	0.01	INTRONIC
NT_010498.14_18473472	NQO1	T|C	16	68309958	0.01	0.01	INTRONIC
NT_010498.15_23374789	NQO1	G|C	16	68318091		0.01	5PRIME_UTR
rs689453	NQO1	A|G	16	68309874	0.05	0.02	SYNONYMOUS_CODING
rs689452	NQO1	G|C	16	68309965	0.11	0.09	INTRONIC
rs10517	NQO1	T|C	16	68301261	0.12	0.12	3PRIME_UTR
rs1800566	NQO1	T|C	16	68302646	0.15	0.21	NON_SYNONYMOUS_CODING
NT_010498.15_23362858	NQO1	C|T	16	68306160	0.33	0.36	INTRONIC
NT_010498.15_23374635	NQO1	A|G	16	68317937	0.01		5PRIME_UTR

SNP	Gene	allele	CHR	Position	MAF (Patient)	MAF(CEPH)	Consequence
rs34590494	NQO1	A|T	16	68301743	0.01		3PRIME_UTR
NT_010498.15_23358698	NQO1	A|G	16	68302000	0.03		3PRIME_UTR
rs4986998	NQO1	T|C	16	68306370	0.03		NON_SYNONYMOUS_CODING
NT_010663.14_15934	P4HB	A|G	17	77395116	0.01	0.01	3PRIME_UTR
NT_010663.14_19476	P4HB	G|A	17	77398658	0.01	0.01	INTRONIC
NT_010663.14_15329	P4HB	G|A	17	77394511		0.01	3PRIME_UTR
NT_010663.14_16084	P4HB	T|C	17	77395266	0.01	0.01	INTRONIC
NT_010663.14_18339	P4HB	T|C	17	77397521		0.01	INTRONIC
NT_010663.14_30968	P4HB	A|G	17	77410150		0.01	INTRONIC
rs2277706	P4HB	G|T	17	77394960	0.01	0.01	3PRIME_UTR
NT_010663.14_17640	P4HB	A|G	17	77396822		0.01	SYNONYMOUS_CODING
NT_010663.14_27400	P4HB	T|C	17	77406582		0.01	INTRONIC
NT_010663.14_16992	P4HB	A|G	17	77396174		0.01	INTRONIC
NT_010663.14_27335	P4HB	T|C	17	77406517	0.04	0.02	INTRONIC
NT_010663.14_16981	P4HB	A|C	17	77396163		0.02	INTRONIC
NT_010663.14_27411	P4HB	G|A	17	77406593	0.06	0.03	INTRONIC
rs8069408	P4HB	G|T	17	77411341	0.21	0.20	INTRONIC
NT_010663.14_32266	P4HB	C|T	17	77411448	0.18	0.22	INTRONIC
NT_010663.14_31757	P4HB	G|A	17	77410939	0.24	0.23	INTRONIC
NT_010663.14_17540	P4HB	A|G	17	77396722	0.22	0.25	INTRONIC
rs2070871	P4HB	A|G	17	77398423	0.22	0.25	SYNONYMOUS_CODING
rs1010954	P4HB	A|G	17	77410294	0.24	0.26	INTRONIC
rs1130664	P4HB	G|T	17	77411649	0.26	0.26	5PRIME_UTR
rs8324	P4HB	A|C	17	77394577	0.24	0.26	3PRIME_UTR
rs11558886	P4HB	T|G	17	77411549		0.26	SYNONYMOUS_CODING
rs1799919	P4HB	G|A	17	77396390	0.25	0.27	SYNONYMOUS_CODING
rs1533756	P4HB	T|C	17	77410215	0.40	0.31	INTRONIC
NT_010663.14_27007	P4HB	G|A	17	77406189	0.01		INTRONIC

SNP	Gene	allele	CHR	Position	MAF (Patient)	MAF(CEPH)	Consequence
NT_010663.14_19110	P4HB	G|A	17	77398292	0.25		INTRONIC
NT_007914.13_9278378	PDIA4	A|G	7	148333295	0.04	0.01	SYNONYMOUS_CODING
NT_007914.13_9287949	PDIA4	G|T	7	148342866		0.01	INTRONIC
NT_007914.14_9281679	PDIA4	C|T	7	148336596		0.01	INTRONIC
NT_007914.13_9287958	PDIA4	A|G	7	148342875		0.01	INTRONIC
NT_007914.14_9277252	PDIA4	T|C	7	148332169		0.01	NON_SYNONYMOUS_CODING
NT_007914.14_9301480	PDIA4	G|A	7	148356397		0.01	SYNONYMOUS_CODING
NT_007914.14_9294341	PDIA4	A|G	7	148349258		0.01	INTRONIC
rs6971353	PDIA4	T|C	7	148332101		0.05	SYNONYMOUS_CODING
rs11546289	PDIA4	A|G	7	148356534	0.34	0.07	5PRIME_UTR
rs6952916	PDIA4	C|T	7	148355878	0.24	0.23	INTRONIC
rs1052549	PDIA4	C|A	7	148331475	0.31	0.26	3PRIME_UTR
NT_007914.13_9278173	PDIA4	T|C	7	148333090	0.25	0.29	INTRONIC
rs7777113	PDIA4	T|C	7	148333429	0.25	0.29	INTRONIC
rs7795577	PDIA4	G|A	7	148333419	0.25	0.30	INTRONIC
NT_007914.14_9292336	PDIA4	T|C	7	148347253	0.26	0.31	INTRONIC
rs10272564	PDIA4	A|G	7	148336423	0.28	0.31	INTRONIC
rs9065	PDIA4	T|C	7	148331782		0.34	3PRIME_UTR
NT_007914.14_9292170	PDIA4	C|T	7	148347087	0.01		SYNONYMOUS_CODING
rs2290971	PDIA4	A|G	7	148343025	0.02		NON_SYNONYMOUS_CODING
NT_007914.14_9294383	PDIA4	C|G	7	148349300	0.02		INTRONIC
NT_005079.12_1238944	PROC	T|C	2	127893963		0.01	INTRONIC
NT_022135.15_16892518	PROC	A|G	2	127901066		0.01	INTRONIC
NT_022135.15_16884145	PROC	C|G	2	127892693		0.02	INTRONIC
rs2069928	PROC	T|G	2	127900384	0.16	0.19	INTRONIC
rs5937	PROC	C|T	2	127901240	0.35	0.23	SYNONYMOUS_CODING
NT_022135.14_16883972	PROC	T|C	2	127892700	0.32	0.28	INTRONIC
rs1799810	PROC	T|A	2	127892510	0.46	0.32	5PRIME_UTR

SNP	Gene	allele	CHR	Position	MAF (Patient)	MAF(CEPH)	Consequence
rs1158867	PROC	C|T	2	127893847	0.44	0.39	INTRONIC
NT_022135.15_16884044	PROC	A|G	2	127892592	0.01		INTRONIC
NT_022135.15_16885346	PROC	T|C	2	127893894	0.01		INTRONIC
NT_022135.15_16885383	PROC	A|G	2	127893931	0.01		INTRONIC
NT_022135.15_16894449	PROC	T|C	2	127902997	0.01		3PRIME_UTR
NT_022135.15_16885587	PROC	T|C	2	127894135	0.01		INTRONIC
rs2884737	VKORC1	C|A	16	31013055	0.19	0.29	INTRONIC
rs7294	VKORC1	T|C	16	31009822	0.37	0.30	DOWNSTREAM