# PHARMACOGENETICS OF WARFARIN

Thesis submitted in accordance with the requirements of

the University of Liverpool for the degree of

Doctor in Philosophy

by

**Jieying Eunice Zhang**

February 2012

# DECLARATION

This thesis is the result of my own work. The material contained within this thesis has not been presented, nor is currently being presented, either wholly or in part for any other degree or qualification.

Jieying Eunice Zhang

This research was carried out in the Department of Molecular and Clinical Pharmacology, in the Institute of Translational Medicine, at The University of Liverpool.

# CONTENTS

**ABSTRACT** ..... i

**ACKNOWLEDGEMENTS** ..... ii

**PUBLICATIONS AND COMMUNICATIONS** ..... iii

**ABBREVIATIONS** ..... v

**CHAPTER 1:** General Introduction ..... 1

**CHAPTER 2:** Effects of *CYP4F2* Genetic Polymorphisms and Haplotypes on Warfarin Clinical Outcomes ..... 40

**CHAPTER 3:** Genetic variability in the *CYP4F2*, *CYP4F11* and *CYP4F12* genes affects liver mRNA levels and plays a role in warfarin response ..... 70

**CHAPTER 4:** Novel *VKORC1* mutations identified in warfarin resistant patients 113

**CHAPTER 5:** Investigation of the functional consequences of novel *VKORC1* mutations identified ..... 146

**CHAPTER 6:** Clotting factors and warfarin treatment – genetic, biochemical and clinical variability ..... 183

**CHAPTER 7:** Final discussion ..... 226

**BIBLIOGRAPHY** ..... 240

**APPENDIX 1** ..... 260

**APPENDIX 2** ..... 270

# ABSTRACT

Warfarin is one of the most commonly used oral anticoagulants worldwide and is highly efficacious for the treatment and prevention of thromboembolic disorders. However, due to its narrow therapeutic index and large interindividual variability, it remains a challenging drug to prescribe. Genetic factors (*CYP2C9* and *VKORC1*), together with clinical factors (age and body weight), account for up to 60% of warfarin dose variance but the remaining ~40% variability remains unexplained.

A polymorphism rs2108622 in *CYP4F2*, a vitamin K oxidase, has previously been associated with increased warfarin stable dose requirements, accounting for 1-7% dose variability. In our cohort of prospectively recruited patients (n = 311), we were unable to confirm these results. Interestingly, after fine mapping of the *CYP4F2* gene region, we found a SNP rs2189784, which is in LD with rs2108622, to be associated with time to therapeutic INR ($P_c$ = 0.03). Further fine mapping of the *CYP4F* gene cluster together with the utilisation a bank of well characterized Caucasian surgical liver samples (n = 149) and data from a genome-wide association study (n = 714), showed that *CYP4F2* rs2108622 and rs2189784 SNPs were found to be associated with increasing *CYP4F2* and decreasing *CYP4F11* or *CYP4F12* mRNA expression, respectively. Interestingly, a *CYP4F11* variant rs1060467 (in LD with rs2108622) was associated with reduced *CYP4F2* mRNA expression. Furthermore, rs1060467 contributes to 2.5% of warfarin dose variability and was associated with reduced warfarin dose requirement (~1 mg/day, $P_c$ = 0.003), an effect in the opposite direction previously reported with *CYP4F2* rs2108622 by Caldwell *et al.* (2008) and other studies.

Warfarin-resistant patients have been reported to harbour *VKORC1* missense mutations. Extended regions of *VKORC1* were sequenced in our patients (n = 65) with resistance to warfarin, defined by clinical and pharmacodynamic criteria. Seven novel heterozygous mutations were identified and *in silico* analyses predicted the promoter c.-160G>C mutation creates a putative Sp1 transcription factor binding site and that the missense mutation c.79C>G to be deleterious. To confirm these predictions, *in vitro* functional studies were carried out using EMSA, transient transfection assays, and DNA methylation. c.-160G>C was found to create a weak binding site for Sp1 transcription factor, and caused an increase in promoter activity by ~20% ($P$ = 0.003). The c.79C>G mutation reduced levels of PIVKA-II by ~10%. Associations of *VKORC1* genotypes with DNA methylation did not remain significant after correction for multiple testing.

The effect of warfarin on the rate of decline of vitamin K-dependent clotting factors, and the role of SNPs in the clotting factor genes, is not known. Using a large prospective cohort of patients (n = 619), SNPs in *F7* and *F10* genes showed association with variability in factor VII levels. The rate at which the plasma levels of factors II, X and protein C decline affect how patients respond to warfarin, in particular the achievement of warfarin stable dose and time to therapeutic INR. Furthermore, the change in clotting factor X level accounted for 1.4% of warfarin dose variability.

In conclusion, the results presented in this thesis demonstrate that multiple genetic, clinical and biochemical factors account for the variability in warfarin response. Further understanding of such complex interactions, along with the advent of genomics technologies and development of new computational and conceptual tools, will yield insights to the accurate prediction of drug efficacy and toxicity, which will hopefully translate into improved outcomes for patients.

# ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Munir Pirmohamed, for giving me this fantastic opportunity to do a PhD. His kindness, support, encouragement, guidance, unsurpassed knowledge and inspiring ideas have made my study immensely enjoyable and rewarding. I would not have come this far without his help. Thanks go out to Professor Kevin Park for his supervision and helpful suggestions. I would also like to acknowledge the financial and academic support of the Department of Health, UK and the University of Liverpool.

A big thank you to Dr. Ana Alfirevic for being such a fabulous mentor, guardian and great friend. Her invaluable guidance, advice, encouragement and support throughout my PhD study (as well as my work-life) have taught me many lessons in life that I will never forget. Thanks for all the inspiring and motivating discussions and help with proofreading of my thesis.

There have been many research collaborations with different research groups throughout my PhD study that have made this thesis possible. Thanks to Professor Uli Zanger and Dr. Kathrin Klein for providing the human liver samples for the CYP4F work. Thanks to Stephane Bourgeois for the GWAS data and I am grateful to Jack Waterson and Jon Cresswell for their help with PLINK and the computer cluster. Thanks to Mrs Kym Jones for sequencing the DNA samples for the VKORC1 work. Thanks to Dr. Chris Goldring who kindly provided the HepG2 cells and the EMSA apparatus. I am grateful to Dr. Ian Copple for his helpful tips on site-directed mutagenesis and to Craig Benson and Pete Metcalfe for their help with the EMSA technique. The plasma clotting factor levels were measured using the MDA-180 in Professor Cheng Hock Toh's research lab in Haematology at the Royal Liverpool Hospital, with helpful guidance from Alan Smith, Colin Downey and Keith Wright – thank you. Thanks also go out to Dr. Andrea Jorgensen and Dr. Steven Lane who provided superb statistical expertise with all analyses undertaken in this thesis. Many thanks to the research nurses Anita Hanson, Diane VanEker, Karen Hawkins, and Lisa Stevens for all their hard work in recruiting the patients, without whom, the study would never have been possible.

I have been blessed with lovely colleagues at the Department of Molecular and Clinical Pharmacology. To Adam, Alison, Alvin, Andrea Davies, Angela, Andrew, Athina, Azizah, Cerys, Claire, Clare, Cris, David, Elena, Eleanor, Fabio, George, Gill, Han, Hayley, Helen C, Jan, John F, John T, Kanvel, Kate, Kath, Laith, Laurence, Lucy, Luke, Maike, Manal, Mas, Monday, Natalie, Philippe, Phil Roberts, Prathibha, Rana, Ray, Sabah, Sarah, Sameh, Sammy, Steffen, Teresa, and others whom have already been named - thank you for your help in one way or another, friendship, as well as the crazy and fun times we had in and outside the lab. You have all made my stay in Liverpool unforgettable. I would also like to thank Dr. Dan Carr and Dr. Vittorio Simeon for being such great friends and their help with the VKORC1 work. My very good friends Haiyi, Silviya (my squash buddy), and Yugo, and my pals (especially Alicia, Amy and Zann) and relatives back home in Singapore – thank you for your care, concern, support and encouragement.

A special thank you to Greepol for always being there for me even though we are thousands of miles apart. "Kob kun mak mak ka ti rak" for all the love, understanding, and support throughout the years.

Finally, words of thanks will never be enough for my amazing parents who have been there for me all my life. Thank you for your constant support, encouragement and endless love throughout the duration of my studies.

# PUBLICATIONS AND COMMUNICATIONS

## Papers published

**Zhang JE**, Jorgensen AL, Alfirevic A, Williamson PR, Toh CH, Park BK, Pirmohamed M. (2009). Effects of *CYP4F2* genetic polymorphisms and haplotypes on clinical outcomes in patients initiated on warfarin therapy. *Pharmacogenetics and Genomics*. Oct;19(10):781-9.

Jorgensen AL, Al-Zubiedi S, **Zhang JE**, Keniry A, Hanson A, Hughes DA, Eker D, Stevens L, Hawkins K, Toh CH, Kamali F, Daly AK, Fitzmaurice D, Coffey A, Williamson PR, Park BK, Deloukas P, Pirmohamed M. (2009). Genetic and environmental factors determining clinical outcomes and cost of warfarin therapy: a prospective study. *Pharmacogenetics and Genomics*. Oct;19(10):800-12.

## Manuscripts in preparation

**Zhang JE**, Klein K, Jorgensen AL, Alfirevic A, Bourgeois S, Deloukas P, Zanger UM, Pirmohamed M. (2012). Genetic variability in the *CYP4F2*, *CYP4F11* and *CYP4F12* genes affects liver mRNA levels and plays a role in warfarin response

**Zhang JE**, Jones K, Carr DF, Simeon V, Alfirevic A, Kamali F, Mountford R, Pirmohamed M. (2012). Warfarin resistance and novel variants identified in the *VKORC1* gene.

**Zhang JE**, Lane S, Jorgensen AL, Downey C, Wright K, Alfirevic A, Bourgeois S, Deloukas P, Toh CH, Pirmohamed M. (2012). Clotting factors and warfarin treatment – genetic, biochemical and clinical variability.

## Oral presentation

"Novel VKORC1 Mutations Identified in Warfarin Resistant Patients". Cold Spring Harbor Laboratory Meeting on Pharmacogenomics and Personalized therapy, New York, 18th November 2010.

## Poster presentations

**Jieying Eunice Zhang**, Kathrin Klein, Andrea Jorgensen, Ana Alfirevic, Ulrich Zanger and Munir Pirmohamed: Genetic variability in the *CYP4F2*, *CYP4F11* and *CYP4F12* genes affects liver mRNA levels and plays a role in warfarin response. Joint Cold Spring Harbor Laboratory/Wellcome Trust Conference on PHARMACOGENOMICS & PERSONALIZED MEDICINE, Hinxton, September 2011.

**Jieying Eunice Zhang**, Ana Alfirevic, Kimberley Spencer, Roger Mountford and Munir Pirmohamed: *VKORC1* AND WARFARIN RESISTANCE. British Pharmacological Society's Winter Meeting, London, December 2009.

**Jieying Eunice Zhang**, Andrea Jorgensen, Kathrin Klein, Ana Alfirevic, Ulrich Zanger and Munir Pirmohamed: *CYP4F2* GENOTYPE-PHENOTYPE CORRELATION IN HUMAN LIVER. British Pharmacological Society's Winter Meeting, London, December 2009.

**Jieying Eunice Zhang**, Ana Alfirevic, Kimberley Spencer, Roger Mountford and Munir Pirmohamed: *VKORC1* and warfarin resistance. Joint Cold Spring Harbor Laboratory/Wellcome Trust Conference on PHARMACOGENOMICS & PERSONALIZED MEDICINE, Hinxton, September 2009.

**Jieying Eunice Zhang**, Andrea Jorgensen, Kathrin Klein, Ana Alfirevic, Ulrich Zanger and Munir Pirmohamed: *CYP4F2* genotype-phenotype correlation in human liver. Joint Cold Spring Harbor Laboratory/Wellcome Trust Conference on PHARMACOGENOMICS & PERSONALIZED MEDICINE, Hinxton, September 2009.

**Jieying Eunice Zhang**, Andrea Jorgenson, Ana Alfirevic, Paula Williamson, Cheng Hock Toh, Munir Pirmohamed: Effect of *CYP4F2* genetic polymorphisms on clinical outcomes in patients initiated on warfarin therapy: a prospective cohort study. European Science Foundation (ESF-UB) Conference, Pharmacogenetics and Pharmacogenomics: Adverse Drug Reactions, Spain, 27th June - 2nd July 2008.

# ABBREVIATIONS

| | |
|---|---|
| ΔΔG value | free energy change of protein stability |
| ABCB1 | ATP binding cassette, subfamily B, member 1 |
| ADRs | adverse drug reactions |
| AHR | aryl hydrocarbon receptor |
| Align-GVGD | Align-Grantham Variation and Grantham Deviation |
| ANOVA | analysis of variance |
| AP-1 | activator protein 1 |
| APC | activated protein C |
| APOE | apolipoprotein E |
| aPTT | activated partial thromboplastin time |
| ATP | adenosine triphosphate |
| BMI | body mass index |
| BRD4 | bromodomain containing protein 4 |
| BSA | body surface area |
| CALU | calumenin |
| CEPH | Centre d'Etude du Polymorphisme Humain |
| CEU | Utah residents with ancestry from northern and western Europe |
| ChIP | chromatin immunoprecipitation |
| CIVC | cysteine132-isoleucine-valine-cysteine135 |
| CL | cytoplasmic loop |
| conc. | concentration |
| CT | cytoplasmic tail |
| CYP | cytochrome P450 |
| CYP1A1 | cytochrome P450, family 1, subfamily A, polypeptide 1 |
| CYP1A2 | cytochrome P450, family 1, subfamily A, polypeptide 2 |
| CYP2C8 | cytochrome P450, family 2, subfamily C, polypeptide 8 |
| CYP2C9 | cytochrome P450, family 2, subfamily C, polypeptide 9 |
| CYP2C18 | cytochrome P450, family 2, subfamily C, polypeptide 18 |
| CYP2C19 | cytochrome P450, family 2, subfamily C, polypeptide 19 |
| CYP3A4 | cytochrome P450, family 3, subfamily A, polypeptide 4 |
| CYP3A5 | cytochrome P450, family 3, subfamily A, polypeptide 5 |
| CYP4F2 | cytochrome P450, family 4, subfamily F, polypeptide 2 |
| CYP4F3 | cytochrome P450, family 4, subfamily F, polypeptide 3 |
| CYP4F8 | cytochrome P450, family 4, subfamily F, polypeptide 8 |
| CYP4F11 | cytochrome P450, family 4, subfamily F, polypeptide 11 |

| | |
|---|---|
| CYP4F12 | cytochrome P450, family 4, subfamily F, polypeptide 12 |
| DMEM | Dulbecco's modified Eagle medium |
| DMET | drug-metabolizing enzymes and transporters |
| DMSO | dimethyl sulphoxide |
| dNTP | deoxyribonucleotide triphosphate |
| EDTA | ethylenediaminetetraacetic acid |
| EPHX1 | epoxide hydrolase 1 |
| ER | endoplasmic reticulum |
| F | forward |
| F10 | coagulation factor 10 |
| F2 | coagulation factor 2 |
| F5 | coagulation factor 5 |
| F7 | coagulation factor 7 |
| F9 | coagulation factor 9 |
| FDA | Food and Drug Administration |
| FDR | false discovery rate |
| G6PD | glucose-6-phosphate dehydrogenase |
| GAS6 | growth arrest-specific 6 |
| GGCX | gamma-glutamyl carboxylase |
| Gla | γ-carboxyglutamic acid |
| Glu | glutamic acid |
| GS | Gene Splicer |
| GWAS | genome-wide association studies |
| HepG2 | human hepatoma cell line |
| HETEs | hydroxyeicosatetraenoic acids |
| HMWK | high molecular weight kininogen |
| HNF4 | hepatocyte nuclear factor 4 |
| HPLC | high-performance liquid chromatography |
| HRP | horseradish peroxidise-conjugated |
| HSF | Human Splicing Finder |
| HWE | Hardy-Weinberg Equilibrium |
| ILVBL | acetolactate synthase-like protein |
| INR | international normalised ratio |
| IPTG | isopropyl-1-thio-β-D-galactopyranoside |

| | |
|---|---|
| ISI | international sensitivity index |
| LAR II | luciferase assay reagent II |
| LD | linkage disequilibrium |
| LRT | likelihood ratio test |
| LTB4 | leukotriene $B_4$ |
| MAF | minor allele frequency |
| MALDI-TOF MS | matrix-assisted laser desorption/ionization time-of-flight mass spectrometry |
| MDA | multi-channel discrete analyzer |
| MES | MaxEntScan |
| MGB | minor groove binding |
| MgCl2 | magnesium chloride |
| M-PVA | polyvinyl alcohol particles |
| MTT | 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide |
| NCBI | National Centre for Biotechnology Information |
| NHS | National Health Service |
| NNSplice | splice site prediction by Neural Network |
| NQO1 | NAD(P)H dehydrogenase, quinone 1 |
| NR1I2 | nuclear receptor subfamily 1, group I, member 2 |
| NR1I3 | nuclear receptor subfamily 1, group I, member 3 |
| ORM1 | orosomucoid 1 |
| ORM2 | orosomucoid 2 |
| PBS | phosphate buffered saline |
| PCR | polymerase chain reaction |
| PGs | prostaglandins |
| PK | prekallikrein |
| PolyPhen | Polymorphism Phenotyping |
| PPAR | peroxisome proliferator-activated receptor |
| PROC | protein C |
| PROCR | protein C receptor |
| PROS1 | protein S |
| PROZ | protein Z |
| PT | prothrombin time |
| PXR | pregnane X receptor |
| QC | quality control |

| | |
|---|---|
| R | reverse |
| RAR | retinoic acid receptor |
| RXR | retinoid X receptor |
| | |
| SAP | shrimp alkaline phosphatase |
| SBE | single base extension |
| SDS-PAGE | sodium dodecyl sulphate-polyacrylamide gel electrophoresis |
| SERPINC1 | serpin peptidase inhibitor, clade C (antithrombin), member 1 |
| SIFT | Sorting Intolerant From Tolerant |
| SLC35E1 | solute carrier family 35, member E1 |
| SNPs | single nucleotide polymorphisms |
| SREBP-1 | sterol regulatory element-binding protein 1 |
| SSF | Splice Site Finder |
| Std Dev | standard deviation |
| SVM | support vector machine |
| | |
| TBS | tris-buffered saline |
| TESS | Transcription Element Search System |
| TF | tissue factor |
| TFPI | tissue factor pathway inhibitor |
| TM1 | transmembrane domain 1 |
| TM3 | transmembrane domain 3 |
| TYA | Thr-Tyr-Ala |
| | |
| VKOR | vitamin K epoxide reductase |
| VKORC1 | vitamin K epoxide reductase complex, subunit 1 |
| VLDL | very low-density lipoprotein |
| VNTRs | variable numbers of tandem repeats |
| vs | versus |
| vWF | von Willebrand factor |
| | |
| ZPI | protein Z-dependent protease inhibitor |

# CHAPTER 1

# General Introduction

# CONTENTS

**1.1 VARIABILITY IN DRUG RESPONSE** ..... **3**

**1.2 PHARMACOGENETICS AND PHARMACOGENOMICS** ..... **3**

**1.3 ADVANCES IN PHARMACOGENETICS** ..... **6**

**1.4 BASIC CONCEPTS ON HUMAN GENETIC VARIATIONS** ..... **7**

**1.5 WARFARIN** ..... **8**

**1.6 HISTORY OF WARFARIN** ..... **9**

**1.7 CLINICAL USE OF WARFARIN** ..... **9**

**1.8 THE BLOOD COAGULATION SYSTEM** ..... **11**

1.8.1 Extrinsic pathway ..... 12

1.8.2 Intrinsic pathway ..... 13

1.8.3 Common pathway ..... 14

1.8.4 Coagulation pathway regulators ..... 16

**1.9 WARFARIN PHARMACOKINETICS** ..... **16**

**1.10 WARFARIN PHARMACODYNAMICS** ..... **17**

**1.11 PROPOSED WARFARIN BINDING SITE ON VKOR** ..... **20**

**1.12 CLINICAL AND ENVIRONMENTAL FACTORS AFFECTING WARFARIN RESPONSE** ..... **22**

**1.13 GENETIC FACTORS AFFECTING WARFARIN DOSE REQUIREMENTS** ..... **24**

1.13.1 *CYP2C9* ..... 29

1.13.2 *VKORC1* ..... 31

1.13.3 *CYP4F2* and other genetic factors ..... 33

**1.14 WARFARIN DOSING ALGORITHMS** ..... **35**

**1.15 THESIS AIMS** ..... **37**

## 1.1 Variability in drug response

It has long been recognized that variability in drug response is observed across all areas of medicine. The father of modern medicine, Sir William Osler (1849 – 1919), conceded that "variability is the law of life, and as no two faces are the same, no two bodies are alike, and no two individuals react alike and behave alike under the abnormal conditions we know as disease" (Osler 2003; Ginsburg and Willard 2009).

Patients have varied responses to drugs, both desirable and undesirable. Adverse drug reactions (ADRs) represent a major public health issue world-wide (Wiffen 2002), contributing significantly to patient morbidity and in serious cases, fatalities; and impose a considerable financial burden on the healthcare system. In the UK, approximately 6.5% of hospital admissions are related to ADRs, with an associated mortality of 0.15%, costing the National Health Service (NHS) £466 million annually (Pirmohamed *et al.,* 2004). In the USA, adverse drug reactions are one of the leading causes of death in the population (Lazarou *et al.,* 1998; Moore *et al.,* 2007). The lack of response to drug therapy, although not uncommon, leads to inefficient use of health care resources and delay in patients receiving appropriate alternative therapies (Limdi and Veenstra 2010). Clearly, the 'one dose fits all' regime is not ideal for patients and is not cost-effective for the health service.

## 1.2 Pharmacogenetics and pharmacogenomics

Heterogeneity in drug response can be explained by clinical and environmental factors such as co-morbidities, concomitant medications, severity of disease, nutritional status, and environmental exposures. In addition, genetics is

thought to be a contributor to variation in drug response. The term "pharmacogenetics" was first coined in 1959 by Friedrich Vogel, a German Pharmacologist (Vogel 1959) and can be defined as the study of how genetic differences affect inter-individual variation in response to medication. Interestingly, the first pharmacogenetic observation dates back to 510 B.C (Table 1.1). in the village of Croton in southern Italy, where Pythagoras first recognized the "dangers of some, but not other, individuals who ate the fava beans" (Rowan 1859; Meletis and Konstantopoulos 2004). The adverse reaction, haemolytic anaemia (favism), is now known to be caused by glucose-6-phosphate dehydrogenase (G6PD) deficiency, a prevalent X-linked trait seen in 1 every 3 males in southern Italy and Sardinia (Schiliro *et al.,* 1979; Meloni *et al.,* 1983; Calabro *et al.,* 1989). G6PD deficiency is now known to be the commonest human enzyme deficiency in the world, affecting approximately 600 million people worldwide, with over 140 variants identified (Cappellini and Fiorelli 2008; Pirmohamed 2011).

There is considerable overlap between the term "pharmacogenetics" and the much newer term "pharmacogenomics", which was introduced in 1997 by Marshall (Marshall 1997) and describes the knowledge and technology used to evaluate the effects of multiple genes, at genome-wide level, on drug response; while pharmacogenetics traditionally focuses on a single gene (Pirmohamed 2001). These two terms, pharmacogenetics and pharmacogenomics, are often used interchangeably and in this PhD thesis, the term pharmacogenetics is used throughout.

**Table 1.1. Historical overview of important advances in pharmacogenetics.**

| Year | Individual(s) | Landmark | References |
|---|---|---|---|
| 510 bc | Pythagoras | Recognition of the dangers of ingesting fava beans, later characterised to be because of deficiency of G6PD | (Nebert 1999) |
| 1866 | Mendel | Establishment of the rules of heredity | (Mendel 1866) |
| 1906 | Garrod | Publication of 'Inborn Errors of Metabolism' | (Garrod 1909) |
| 1932 | Snyder | Characterisation of the 'phenylthiourea nontaster' as an autosomal recessive trait | (Snyder 1932) |
| 1956 | Alving *et al.* | Discovery of glucose-6-phosphate dehydrogenase deficiency | (Alving *et al.*, 1985) |
| 1957 | Motulsky | Further refined the concept that inherited defects of metabolism could explain individual differences in drug response | (Motulsky 1957) |
| 1957 | Kalow and Genest | Characterisation of serum cholinesterase deficiency | (Kalow and Genest 1957) |
| 1957 | Vogel | Coined the term pharmacogenetics | (Vogel 1959) |
| 1960 | Price Evans | Characterisation of acetylator polymorphism | (Evans *et al.*, 1960) |
| 1962 | Kalow | Publication of 'Pharmacogenetics - Heredity and the Response to Drugs' | (Kalow 1962) |
| 1977/79 | Mahgoub *et al.* and Eichelbaum *et al.* | Discovery of the polymorphism in debrisoquine hydroxylase | (Mahgoub *et al.*, 1977; Eichelbaum *et al.*, 1979) |
| 1988 | Gonzalez *et al.* | Characterisation of the genetic defect in debrisoquine hydroxylase, later termed CYP2D6 | (Gonzalez *et al.*, 1988) |
| 1988-2000 | Various | Identification of specific polymorphisms in various phase I and phase II drug metabolising enzymes, and latterly in drug transporters | |
| 2001-2003 | Public-private partnership | Completion of the initial draft and complete sequence of the human genome | (Lander *et al.*, 2001; Venter *et al.*, 2001) |
| 2003 | The International HapMap Project | Completion of map of human genome sequence variation | (The International HapMap Consortium 2003) |
| 2006 | Reddon *et al.* | Global map of copy number variation | (Redon *et al.*, 2006) |
| 2007 | Wellcome Trust Case-Control Consortium | Genome-wide association in 14,000 cases in seven diseases | (Burton *et al.*, 2007) |
| 2011 | 1000 genomes project | A map of human genome variation based on population-scale genome sequencing | (Altshuler *et al.*, 2010) |

Adapted from Pirmohamed 2011.

## 1.3 Advances in pharmacogenetics

Phenotype-driven assessment of variation in drug metabolising enzyme genes was the hallmark of research undertaken from the end of the 1950s to the end of the 1980s (Meyer 2004; Pirmohamed 2011). This usually requires the administration of a probe drug and the measurement of the ratio between the probe drug and its metabolite, the ratio being used to depict whether the individual had an absolute or partial deficiency of an enzyme. There is an advantage to understanding the phenotype of a particular gene because it enables the identification of many polymorphisms, even those that have not been discovered, and determination of phenocopy (where there is no functional polymorphism in the gene, but the function is decreased because of the co-administration of a drug that inhibits that enzyme). However, disadvantages include the labour intensive nature of the techniques, the associated cost, the low throughput and the fact that in some cases, the probe substance might not be specific for the one enzyme (Pirmohamed 2011).

The advent of molecular biological techniques enabled pharmacogenetics to enter a new era in the 1990s, where the phenotypic assessments could be directly related to nucleotide substitutions (and other variants) in the causative genes (Pirmohamed 2011). Although the wide availability of polymerase chain reaction (PCR)-based techniques enabled molecular assessment of many genes, predominantly the drug metabolising enzyme genes, most studies were still largely limited to single genes, and often single variants within that gene (Pirmohamed 2011).

Advances in pharmacogenetics truly began in the last decade following the completion of the Human Genome Project (Venter *et al.*, 2001) and the International HapMap Project (The International HapMap Consortium 2003), along with the rapid

development of genotyping and sequencing technologies, which have enabled the assessment of the whole genome and have greatly affected pharmacogenetic discoveries (Meyer 2004; Pirmohamed 2011). Table 1.1 highlights some of the major advances that have occurred this century. Pharmacogenetics offers the opportunity to personalise medicines where inter-individual differences in genetic information can be used to improve drug efficacy and reduce drug toxicity, optimising patient care and health outcomes.

## 1.4 Basic concepts on human genetic variations

The human genome consists of approximately 3.3 billion base pairs and comprises over 27, 000 genes. Although the human DNA sequence is more than 99% identical across different people (Lander *et al.*, 2001), the 1% difference between 2 individuals' genomes includes more than 12 million potential variations (Sachidanandam *et al.*, 2001; The International HapMap Consortium 2005; Frazer *et al.*, 2007). Genetic variants that occur infrequently at less than 1% in the population are called "mutations", whereas variants that occur more frequently at 1% or greater are called "polymorphisms". Polymorphisms include insertions or deletions, copy-number variations, variable numbers of tandem repeats (VNTRs) and single nucleotide polymorphisms (SNPs). The different forms or variants of a particular polymorphism are called "alleles".

SNPs occur every 100 to 300 bases along the human genome and are the most common polymorphisms, accounting for approximately 90% of all human genetic variations. SNPs can occur in coding (protein) and non-coding regions of the genome. SNPs occurring in the coding region include (i) nonsynonymous SNPs which lead to a change in an amino acid sequence of the resultant protein, and (ii)

synonymous SNPs which do not result in amino acid change. Other SNPs that do not directly code for protein may still influence cell function through other means, such as controlling the amount of protein the cell builds, thereby influencing drug response.

SNPs are also evolutionarily stable, not changing much from generation to generation. This natural process of non-random association of 2 or more alleles at two or more physically proximate loci is known as "linkage disequilibrium" (LD). Alleles that tend to occur together on the same chromosome and which tend to be inherited together are collectively known as a "haplotype". The regions of DNA that are in LD remain unchanged during recombination and thus "travel together" in transmitting the genomic material from parent to offspring. This efficiently permits the investigation of only 1 representative SNP that can serve as a "tag" or "marker" for nearby SNPs and haplotype blocks by decreasing the number of total SNPs that need to be tested directly, which is utilised in the design of genome-wide association studies (GWAS).

## 1.5 Warfarin

Warfarin (3-α-acetylbenzyl-4-hydroxycoumarin; also known as Coumadin) is the most commonly used oral anticoagulant worldwide. Almost 20 million prescriptions are written in the US each year (Marketos 2004). In the UK, it has been estimated that at least 1% of the whole population and 8% of those aged over 80 years are taking warfarin (Pirmohamed 2006). Owing to its narrow therapeutic window and large inter-individual variability (up to 20-fold difference between different individuals), warfarin has become an interesting and important case study for pharmacogenetics.

## 1.6 History of warfarin

In the 1920s livestock farmers in North Dakota, United States, observed an outbreak of severe, unexplained bleeding in cattle. This was later linked to their diet of mouldy sweet clover hay (*Melilotus alba* and *M. officinalis*) (Schofield 1924; Roderick 1929) which was found to contain a haemorrhagic factor that caused a plasma prothrombin defect (Roderick 1931). However, it was not until 10 years later that the coumarin dicoumarol, 3,3'-methyllenbis(4-hydroxycoumarin), was identified as the active agent in sweet clover responsible for the bleeding disorder (Campbell and Link 1941). This led to the commercialisation of dicoumarol in 1941, and efforts to develop an effective rodenticide resulted in the synthesis of warfarin in 1948 (Link 1959). The name "warfarin" was derived from the acronym *WARF* for Wisconsin Alumni Research Foundation, with the suffix *–arin* from coumarin (Link 1959). The survival of a navy recruit after an unsuccessful attempted suicide by the use of a large amount of warfarin-based rodenticide (estimated to have ingested 567 mg of warfarin) led to clinical trials of warfarin as an anticoagulant in humans, and in 1954, warfarin was approved for medical use (Shapiro 1953; Clatanoff *et al.,* 1954; Link 1959).

## 1.7 Clinical use of warfarin

Warfarin is a highly effective drug for the prevention of thromboembolic events in patients with atrial fibrillation, deep vein thrombosis, pulmonary embolism and mechanical heart valves. Therapy is usually started empirically on a fixed dosed (such as 5 or 10 mg/day) during the first two to three days of warfarin initiation. However, to achieve safe and effective anticoagulation, warfarin treatment needs to

be monitored closely via the international normalised ratio (INR), a laboratory test universally used to measure the clotting ability of blood; and therapeutic doses subsequently titrated according to the INR response for each patient.

INR can be determined from the prothrombin time (PT) test result by a standard formula: (PT of the patient/geometric mean of the PT reference range) $^{\text{ISI of the reagent}}$, where ISI is the international sensitivity index. The universally adopted INR standards for the clinical management of warfarin range between 2.0 to 3.0 for most indications. The exceptions are some types of mechanical prosthetic heart valves and in certain patients with thrombosis and the antiphospholipid syndrome (see Table 1.2).

Studies in patients with atrial fibrillation taking warfarin showed that patients were outside the INR target range 30-50% of the time (Jones *et al.,* 2005; Boulanger *et al.,* 2006). Under dosing can lead to low INR, causing thrombosis while excessive anticoagulation can lead to high INR, resulting in serious adverse reactions such as bleeding (Palareti *et al.,* 1996; Stroke Prevention in Atrial Fibrillation Investigators 1996). Indeed, warfarin is often among the top 3 drugs that lead to hospitalisation from ADRs (Pirmohamed *et al.,* 2004; Budnitz *et al.,* 2007; Wysowski *et al.,* 2007). Bleeding is the most common ADR of warfarin and occurs in up to 41% of patients treated with warfarin, with major bleeding frequencies as high as 10-16% (Gulløv *et al.,* 1999; Petty *et al.,* 1999; Wysowski *et al.,* 2007). The risk of adverse events is highest during the dose-titration period within the first few weeks to months of warfarin therapy (Fihn *et al.,* 1993; Landefeld and Beyth 1993; White *et al.,* 1999; Beyth *et al.,* 2000; Hylek *et al.,* 2007). In addition to mortality and morbidity, substantial hospitalisation costs are also associated with warfarin-related bleeds. An analysis by Fanikos and colleagues showed that the average cost per patient of a

bleeding episode was nearly $16, 000 (range $2, 707–$64, 446) with a mean length of stay of 6 days (Fanikos *et al.*, 2005).

**Table 1.2. Optimal INR therapeutic range for warfarin therapy*.**

| Indication | INR |
|---|---|
| Prophylaxis of venous thrombosis (high-risk surgery)<br>Treatment of venous thrombosis<br>Treatment of pulmonary embolism<br>Prevention of systemic embolism in<br>Atrial fibrillation<br>Valvular heart disease<br>Acute myocardial infarction (to prevent systemic embolism)**<br>Tissue heart valves | 2.0 - 3.0 |
| Mechanical prosthetic heart valves (high risk) | 2.5 - 3.5 |
| Bileaflet mechanical heart valve in aortic position | 2.0 - 3.0 |

*Recommended by the American College of Chest Physicians
** An INR of 2.5 - 3.5 is recommended for recurrent myocardial infarction, consistent with US Food and Drug Administration recommendations.
Table adapted from Hirsh *et al.*, 2001

Therefore, strategies to individualise the initial warfarin dose have been sought. Numerous clinical and genetic factors influence warfarin dose response in individuals. These will be discussed in more detail in subsequent sections.

## 1.8 The blood coagulation system

The coagulation system is designed to prevent bleeding at the site of vessel injury through the formation of a blood clot. Clot formation proceeds in two phases: primary and secondary haemostasis, and involves the interactions among three components: the blood vessel wall, cellular components within the blood (predominantly platelets in the arterial circulation), and plasma clotting factors. Primary haemostasis occurs immediately after blood vessel injury. Vasoconstriction

occurs and glycoproteins in the circulating platelets adhere to the exposed von Willebrand factor (vWF) on the damaged vessel wall. On contact with collagen fibres on the vessel surface, the platelets become activated and platelet aggregation occurs where more platelets and circulating fibrinogen are attracted to the damaged vessel wall, forming a soft plug. This is followed by secondary haemostasis in which the initial platelet plug is consolidated in a meshwork of fibrin via the coagulation pathway.

The blood coagulation pathway is a proteolytic cascade. All of the coagulation factors involved are serine proteases (enzymes) except factor V and factor VIII which are glycoproteins (cofactors), and factor XIII which is a transglutaminase. Each coagulation factor is present in the plasma as a zymogen (inactive form), which on activation undergoes proteolytic cleavage to release the active factor from the precursor molecule. The coagulation pathway functions as a series of positive and negative feedback loops which controls the activation process. The ultimate goal of the pathway is to produce thrombin, which converts soluble fibrinogen into insoluble fibrin, forming a clot. The reactions leading to fibrin formation can be divided into the (i) extrinsic, (ii) intrinsic and (iii) common pathways (Davie and Ratnoff 1964; Macfarlane 1964), as illustrated in Figure 1.1.

### 1.8.1 Extrinsic pathway

The extrinsic pathway (Figure 1.1), also known as the tissue factor pathway, is the primary pathway for the initiation of blood coagulation. Upon vascular injury, the extrinsic pathway is initiated in response to the exposed tissue factor (TF), a subendothelial cell-surface glycoprotein that binds calcium ions and phospholipids.

When the circulating factor VII comes into contact with TF, a TF-VIIa complex is formed. This complex rapidly converts factor X to its enzyme form, factor Xa.

### 1.8.2 Intrinsic pathway

The intrinsic pathway (Figure 1.1) requires the coagulation factors VIII, IX, X, XI, and XII. Other requirements are the proteins prekallikrein (PK) and high molecular weight kininogen (HMWK), as well as calcium ions and phospholipids secreted from platelets. Initiation of the intrinsic pathway occurs when PK, HMWK, and factor XII forms a complex with the negatively charged phospholipids on the damaged vessel surface. This is termed the "contact phase". As such the intrinsic pathway is also known as the contact activation pathway. The assembly of the contact phase components results in the conversion of prekallikrein to kallikrein, which in turn activates factor XII to factor XIIa. Factor XIIa then cleaves the factor XI zymogen to its active form, factor XIa, which then converts factor IX to factor IXa. Subsequently, factor IXa cleaves factor X, activating it to factor Xa. In the presence of minute quantities of thrombin, factor VIII is activated to factor VIIIa (cofactor). Factor VIIIa acts as a receptor for factors IXa and X, forming the tenase complex. This complex in turn activates factor X, and factor Xa catalyses thrombin formation from prothrombin. As the concentration of thrombin increases, factor VIIIa is ultimately cleaved by thrombin and inactivated. This dual action of thrombin, upon factor VIII, acts to limit the extent of tenase complex formation and thus the extent of the coagulation cascade.

### 1.8.3 Common pathway

The common point (Figure 1.1) in both the extrinsic and intrinsic pathways is the activation of factor X to factor Xa. Factor Xa activates prothrombin (factor II) to thrombin (factor IIa). Further activation of thrombin occurs on the surface of activated platelets via the formation of the prothrombinase complex. The glycoprotein factor V is activated to factor Va by means of small amounts of thrombin. Factor Va acts as a cofactor in the formation of the prothrombinase complex by binding to specific receptors on the surfaces of activated platelets and forms a complex with prothrombin and factor Xa. Akin to factor VIII activation in the formation of the tenase complex, factor Va is inactivated by increased levels of thrombin, limiting the extent of the coagulation cascade.

Finally, thrombin converts fibrinogen (factor I) to fibrin (factor Ia). Thrombin also activates factor XIII to factor XIIIa, a highly specific transglutaminase that introduces cross-links between the fibrin monomers, solidifying the fibrin clot.

In addition to its role in the activation of fibrin clot formation, thrombin plays an important regulatory role in coagulation to prevent excessive clotting. Thrombin combines with thrombomodulin, an endothelial cell surface protein, forming a complex that converts the major physiological anticoagulant, protein C, to activated protein C (APC). The cofactor protein S and APC degrade factors Va and VIIIa (Esmon *et al.,* 1982; Comp *et al.,* 1984), thereby limiting their activities in the coagulation cascade.

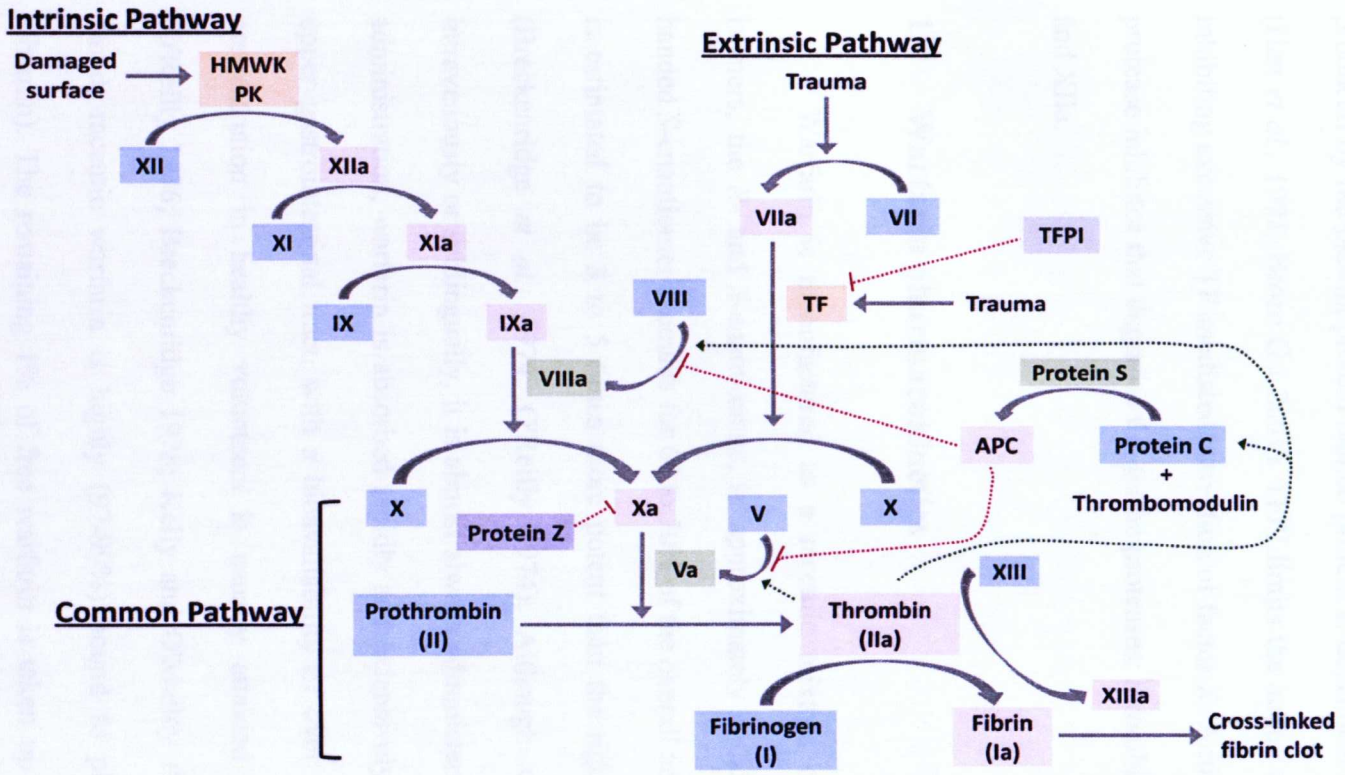



**Figure 1.1. The coagulation cascade.** The extrinsic pathway is initiated upon vascular injury which leads to exposure of the tissue factor (TF) that binds phospholipids. The intrinsic cascade is initiated when contact is made between blood and exposed negatively charged surfaces. The two pathways converge at the activation of factor X to Xa. Activated factor Xa activates prothrombin to thrombin. Thrombin then activates factors V and VIII (as indicated by green arrows), furthering the cascade. Ultimately, thrombin converts fibrinogen to fibrin and activates factor XIII to XIIIa, which cross-links fibrin polymers to solidify the clot. Thrombin also regulates the clotting process by converting the anticoagulant protein C (depicted by green arrow) to activated protein C (APC) which inactivates Va and VIIIa (as indicated by red lines). Protein Z inhibits factor Xa activity (as indicated by the red line). Tissue factor pathway inhibitor (TFPI) inhibits the action of TF. HMWK: high molecular weight kininogen; PK: Prekallikrein.

### 1.8.4 Coagulation pathway regulators

Other regulators of the coagulation pathway include protein Z, tissue factor pathway inhibitor (TFPI) and anti-thrombin III. Protein Z is a cofactor which forms a calcium ion-dependent complex with factor Xa at phospholipid surfaces (Han *et al.*, 1998) and this protein Z–factor Xa interaction enhances the inhibition of factor Xa produced by the plasma protein named protein Z-dependent protease inhibitor (ZPI) (Han *et al.*, 1998; Broze G.J 2001). TFPI limits the action of tissue factor, thereby inhibiting excessive TF-mediated activation of factor X. Anti-thrombin III is a serine protease inhibitor that degrades the serine proteases: thrombin, factors IXa, Xa, XIa, and XIIa.

## 1.9 Warfarin pharmacokinetics

Warfarin is manufactured as a racemic mixture of two optically active isomers, the *R*- and *S*-enantiomers, in approximately equal proportions. The left-handed *S*-enantiomer accounts for 60 to 70% of the overall anticoagulant activity and is estimated to be 3 to 5 times more potent than the right-handed *R*-enantiomer (Breckenridge *et al.*, 1974; O'Reilly 1974). Although warfarin can be given intravenously or sublingually, it is almost always administered by mouth. After oral administration, warfarin is absorbed rapidly and extensively from the stomach and upper gastrointestinal tract, with a bioavailability of over 90%, and peak plasma concentration in healthy volunteers is usually attained within 60-90 minutes (O'Reilly 1976; Breckenridge 1978; Kelly and O'Malley 1979). In the circulating blood, racemic warfarin is highly (97-99%) bound to plasma proteins (mainly albumin). The remaining 1% of free warfarin is taken up by the liver, primarily microsomes, where it exerts its pharmacological actions (Wilting *et al.*, 1980; Porter

and Sawyer 1992). Peak anticoagulant effect occurs 36-72h after drug administration when clotting factors, especially prothrombin, are cleared from the circulation (O'Reilly and Aggeler 1968).

The termination of warfarin pharmacological effect and the elimination of warfarin are dependent on hepatic metabolism, which is catalysed by the cytochrome P450 (CYP) complex (see Figure 1.2). *S*-warfarin is metabolised almost exclusively by CYP2C9 to its major metabolite 7-hydroxywarfarin and to a much lesser extent, 6-hydroxywarfarin (Rettie *et al.,* 1992). Both metabolites are inactive and are excreted in the bile. *R*-warfarin is metabolised by CYP1A2 (Zhang *et al.,* 1995), CYP2C19 (Kaminsky *et al.,* 1993) and CYP3A4 (Brian 1990; Rettie *et al.,* 1992) to 6-, 7-, and 8-hydroxywarfarin that are excreted in the urine. The metabolic elimination of *S*-warfarin is 3 times faster than that of *R*-warfarin. (Breckenridge *et al.,* 1974; O'Reilly 1974).

## 1.10 Warfarin pharmacodynamics

The procoagulant factors II, VII, IX, X, and anticoagulant proteins C, S and Z are known to be vitamin K-dependent proteins which require γ-carboxylation of their glutamic acid residues to become fully functional. The conversion of glutamic acid (Glu) residues on the N-terminal regions of vitamin-K-dependent proteins to γ-carboxyglutamic acid (Gla) residues causes a conformational change which promotes their binding to cofactors on phospholipid surfaces in the presence of calcium ions (Stenflo *et al.,* 1974; Nelsestuen 1976). This post-translational modification reaction is accomplished by the enzyme γ-glutamyl carboxylase (GGCX) and requires reduced vitamin $K_1$ (vitamin $K_1$ hydroquinone or vitamin $K_1H_2$) as a cofactor (Nelsestuen *et al.,* 1974). Vitamin $K_1H_2$ is formed by the reduction of vitamin $K_1$

quinone (vitamin $K_1$), which is catalysed by the enzyme vitamin K 2,3-epoxide reductase (VKOR). Concomitant with the γ-carboxylation reaction, vitamin $K_1H_2$ is oxidised to vitamin $K_1$ 2, 3-epoxide. Due to the limited availability of vitamin K in tissues *in vivo*, the epoxide must be rapidly recycled to vitamin $K_1H_2$. This is again catalysed by VKOR. This cyclic inter-conversion is known as the vitamin K cycle (Suttie 1978) and is depicted in Figure 1.2.

Warfarin exerts its anticoagulant effect by inhibiting VKOR (Figure 1.2), thereby preventing vitamin K recycling. The decrease in vitamin $K_1H_2$ availability leads to diminished γ-carboxylation of vitamin K-dependent clotting factors II, VII, IX and X, thus inhibiting coagulation (Bell *et al.,* 1972; Whitlon *et al.,* 1978; Choonara *et al.,* 1988). In addition to its anticoagulant effect, warfarin inhibits carboxylation of the regulatory anticoagulant proteins C and S and therefore has the potential to exert a procoagulant effect.

When the liver concentration of vitamin $K_1$ is elevated, vitamin $K_1$ can be reduced to vitamin $K_1H_2$ by an alternative enzyme, flavoprotein DT-diaphorase, which is a nicotinamide adenine dinucleotide phosphate (NAD(P)H) dehydrogenase (Hochstein 1983; Wallin and Martin 1987). Warfarin does not inhibit DT-diaphorase (Wallin and Martin 1987). Therefore, warfarin intoxication can be counteracted by administering high doses of vitamin $K_1$ which drives cofactor vitamin $K_1H_2$ production for GGCX, restoring normal functioning of the blood coagulation system (Figure 1.2).

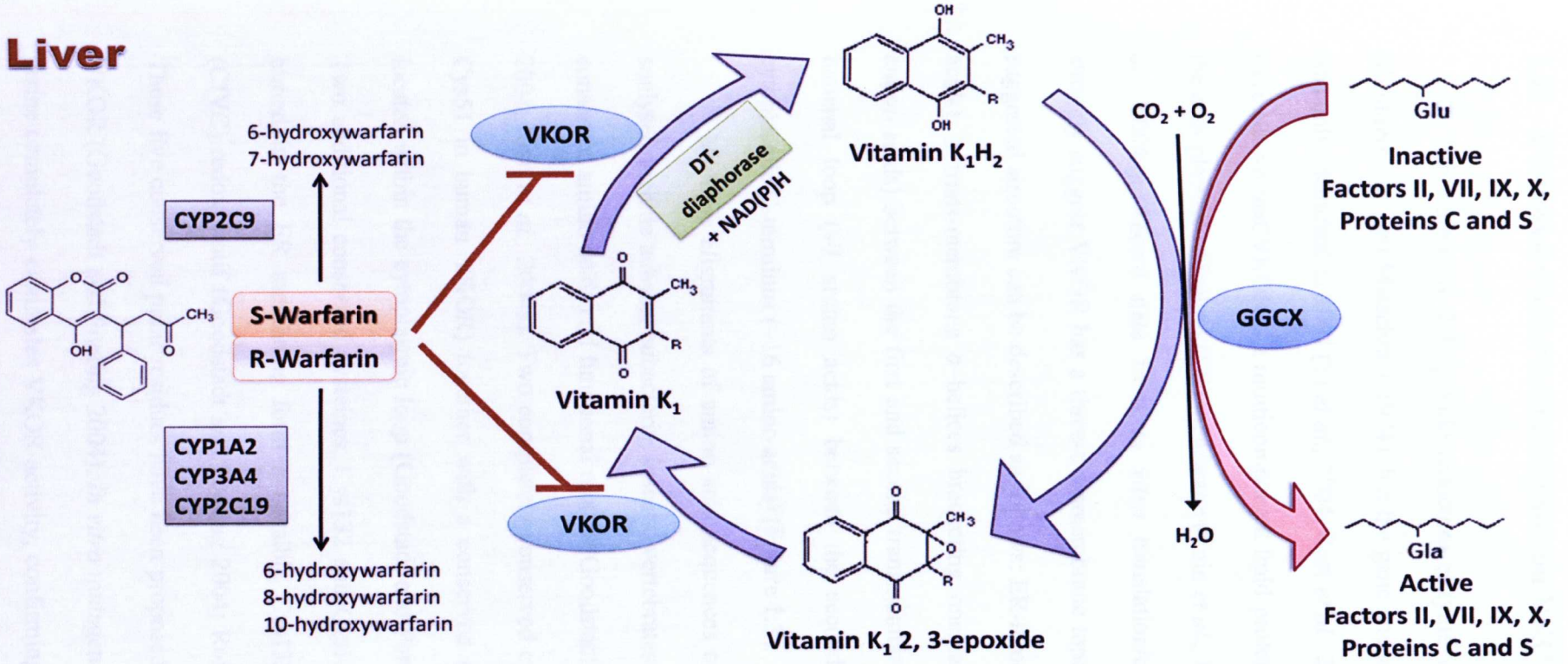



**Figure 1.2. Pharmacology of warfarin.** Warfarin, an enantiomeric mixture of equal concentrations of *R*- and *S*-forms, is absorbed from the gastrointestinal tract and metabolised by hepatic enzymes into inactive hydroxylated metabolites. *S*-warfarin is ~5x more potent than *R*-warfarin. GGCX, with vitamin $K_1H_2$ as a cofactor, converts vitamin K-dependent proteins to γ-carboxyglutamic acid (Gla) containing proteins by adding $CO_2$ to glutamic acid (Glu) residues in newly synthesized proteins. Concomitant with γ-carboxylation, vitamin $K_1H_2$ is converted to vitamin $K_1$ 2,3-epoxide. The epoxide is reduced by VKOR to the vitamin $K_1H_2$ cofactor, creating a cyclic inter-conversion known as the vitamin K cycle. Warfarin inhibits VKOR, decreasing vitamin $K_1H_2$ availability, diminishing active clotting factors, thereby inhibiting coagulation. Vitamin $K_1$ can alternatively be reduced to vitamin $K_1H_2$ by DT-diaphorase, an enzyme which is not inhibited by warfarin. VKOR: vitamin K epoxide reductase; GGCX: γ-glutamyl-carboxylase; NAD(P)H: nicotinamide adenine dinucleotide phosphate; $CO_2$: carbon dioxide; $O_2$: oxygen; $H_2O$: water.

## 1.11 Proposed warfarin binding site on VKOR

The vitamin K 2,3-epoxide reductase enzyme was first identified in 1974 (Zimmermann and Matschiner 1974), but the gene encoding this enzyme, *VKORC1*, was only identified in 2004 (Li *et al.*, 2004; Rost *et al.*, 2004a). Biochemical studies have shown that VKOR is a multicomponent lipid-protein enzyme system located in the endoplasmic reticulum (ER) membrane (Cain *et al.*, 1997; Li *et al.*, 2004; Rost *et al.*, 2004a). Recent data from *in vitro* translation/cotranslocation experiments strongly suggest VKOR has a three-transmembrane topology (Tie *et al.*, 2005) The suggested structure can be described as follows: ER-lumenal N-terminus (~10 amino acids), 3 trans-membrane α–helices bracketing one large cytoplasmic loop (~69 amino acids) between the first and second transmembrane helices and a small ER-lumenal loop (~7 amino acids) between the second and third helices, and a cytoplasmic C-terminus (~16 amino acids) (Figure 1.3).

Multiple alignments of amino acid sequences of VKOR orthologs (species analysed include achaea, eubacteria, insects, vertebrates and plants) indicate several conserved amino acids and functional motifs (Goodstadt and Ponting 2004; Li *et al.*, 2004; Rost *et al.*, 2004a). Two completely conserved cysteine residues (Cys43 and Cys51 in human VKOR) together with a conserved serine/threonine (Ser57) are located within the cytoplasmic loop (Goodstadt and Ponting 2004; Tie *et al.*, 2005). Two additional conserved cysteines, Cys132 and Cys135, predicted to be partially buried in the ER membrane form a possible Cys132-Isoleucine-Valine Cys135 (CIVC) redox motif (Goodstadt and Ponting 2004; Rost *et al.*, 2005) (Figure 1.3). These five conserved polar residues have been proposed to form the active centre of VKOR (Goodstadt and Ponting 2004). *In vitro* mutagenesis of Cys132 or Cys135 to serine completely eliminates VKOR activity, confirming that the CIVC redox motif

plays a crucial role in vitamin K epoxide reduction (Wajih *et al.,* 2005). *In vitro* mutagenesis of tyrosine 139 to phenylalanine (a substitution found in warfarin-resistant rats) provided indirect evidence for Tyr139 being part of the warfarin binding site. Although tyrosine and phenylalanine differ by only one hydroxyl group, substitution led to nearly complete warfarin resistance. Tyr139 is flanked by threonine and alanine; this hydrophobic sequence motif Thr-Tyr-Ala (TYA) has been proposed to be the warfarin binding site (Ma *et al.,* 1992).

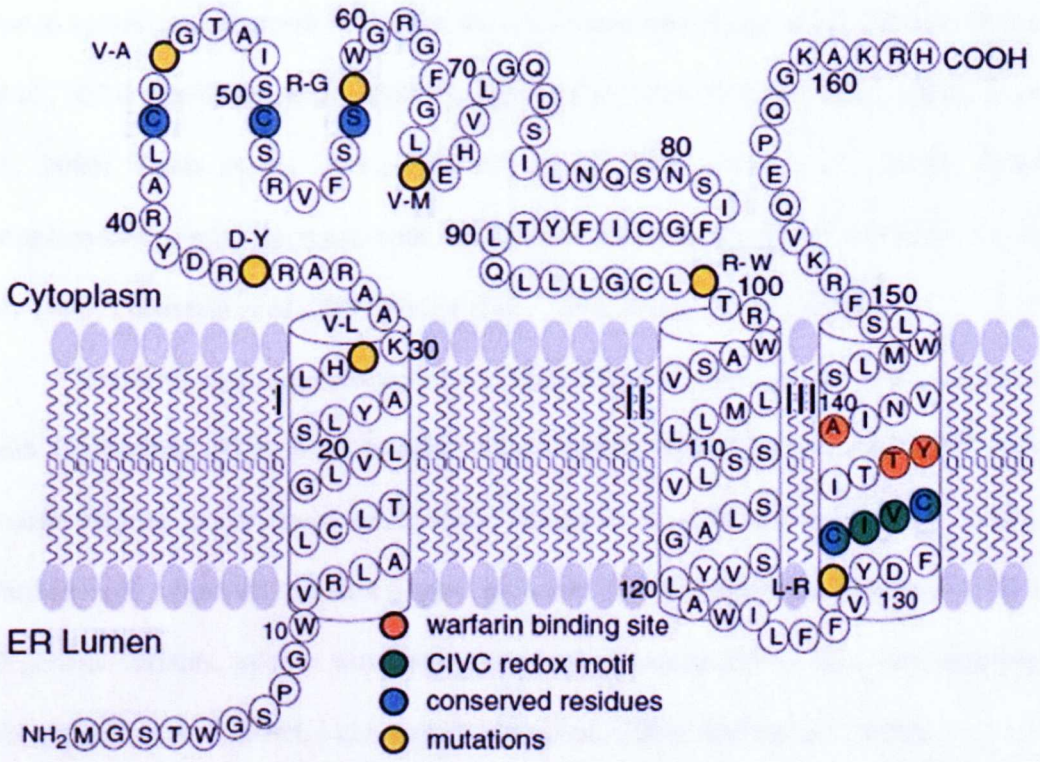



**Figure 1.3. Proposed membrane topology of VKOR.** The model predicts three transmembrane α-helices. The amino-terminal part (aa 1–10) of the enzyme is located within the ER-lumen, followed by α-helix 1 (aa 11–30), a large cytoplasmic loop of 69 aa, α-helix 2 (aa 101–120), and a small ER lumenal loop of 7 aa. The third α-helix (aa 128–147) leads over to the cytoplasmic carboxy terminal part of 16 aa. Functional motifs comprising TYA warfarin binding site (aa 138–140, orange circles), CIVC redox motif (aa 132–135, green circles), ER retention signal (aa 159–163), and five amino acids conserved throughout all species (aa Cys43, Cys51, Ser/Thr57, Cys132, Cys135, blue circles) are highlighted. aa: amino acid; ER: endoplasmic reticulum; CIVC: Cysteine132-Isoleucine-Valine-Cysteine135; TYA: Threonine-Tyrosine-Alanine. Figure adapted from Tie *et al.*, 2005.

## 1.12 Clinical and environmental factors affecting warfarin response

Many clinical and environmental factors influence warfarin dose requirement and response. They include age, ethnicity, weight, height, medications, diet, illness, smoking and adherence.

Increasing patient age has consistently been associated with a higher sensitivity to warfarin, which may be caused by the significant negative correlation between age and warfarin clearance, and by the fall in total hepatic content of VKOR due to age-related decrease in hepatic mass requirements (Gage *et al.,* 2004; Hillman *et al.,* 2004; Aquilante *et al.,* 2006; Carlquist *et al.,* 2006; Herman *et al.,* 2006; Li *et al.,* 2006; Tham *et al.,* 2006; Caldwell *et al.,* 2007; Wu *et al.,* 2008). Dose requirements usually decrease with age by ~8–10% per decade of life (Gurwitz *et al.,* 1992; Loebstein *et al.,* 2001; Gage *et al.,* 2004; Sconce *et al.,* 2005).

Warfarin dose requirements have been found to vary by race. As compared with Caucasians, African-Americans require higher doses (Gage *et al.,* 2004) and Asians require lower doses on average (Dang *et al.,* 2005; Voora *et al.,* 2005). Studies have suggested that this is most probably due to differences in the prevalence of genetic variants, as race was found to contribute minimally to dose requirements after adjusting for *VKORC1* genotype (Gage *et al.,* 2008; Klein *et al.,* 2009).

Medications can affect the pharmacokinetics of warfarin by reducing its absorption from the intestine, by altering its clearance, or by competitive protein binding. Drugs can also influence the pharmacodynamics of warfarin by mechanisms such as inhibition of the synthesis of vitamin K-dependent coagulation factors or increased clearance of these factors. A list of major medications that interact with warfarin is listed in Table 1.3. Studies have reported that patients on amiodarone

require nearly 30% lower doses of warfarin for stable anticoagulation (Sanoski and Bauman 2002; Gage *et al.*, 2008; Klein *et al.*, 2009).

**Table 1.3. Some major drug interactions with warfarin.**

| INCREASED Effect of Warfarin | | DECREASED Effect of Warfarin |
|---|---|---|
| **Anti-Platelet Agents** Abciximab (ReoPro), Aspirin, Dipyridamole, NSAIDs, Clopidogrel, Tirofiban | **Analgesics** Paracetamol (large doses i.e. 4 to 7g per week), Tramadol | Ascorbic Acid (large doses) Vitamin K |
| **COX-2 Inhibitors** Celecoxib, Rofecoxib | **Anticonvulsants** Phenytoin | **Anticonvulsants** Carbamazepine, Phenytoin |
| **Antibiotics** Cephalosporins, Macrolides, Metronidazole, Sulphonamides, Quinolones, Vancomycin | **Selective Serotonin Reuptake Inhibitors** Fluoxetine | **Antibiotics** Rifampicin, Rifabutin |
| **Antifungals** Itraconazole, Fluconazole, Ketoconazole | **Tricyclic Antidepressants** | **Sedatives** Barbiturates |
| **Antiarrythmics** Amiodarone, Mexiletine, Verapamil, Quinidine | **Selective Estrogen Receptor Modulator** Raloxifene, Tamoxifen | |
| **Herbal Medicines** Dong Quai, Garlic, Papaya, St. John's Wort, Ginkgo, Ginger and Garlic (large amounts), Guarana | | **Herbal Medicines** Ginseng, Slippery Elm Bark, Green Tea, Co-Enzyme Q10 |

Warfarin dose requirements may also be affected by dietary factors such as alcohol consumption or vitamin K intake. Alcohol may affect warfarin metabolism and high dietary intake of vitamin K (found in green vegetables) can diminish the action of warfarin. However, there is conflicting evidence on the association between warfarin maintenance doses and vitamin K intake (Loebstein *et al.*, 2001; Absher *et al.*, 2002; Gage *et al.*, 2004; Sconce *et al.*, 2005; Sconce *et al.*, 2007).

Various illnesses such as liver disease, malnutrition, decompensated heart failure, hypermetabolic states (e.g. febrile illnesses, hyperthyroidism) can affect warfarin dose requirements (Ansell *et al.,* 2001; Gage and Eby 2003; D'Andrea *et al.,* 2008; Gage *et al.,* 2008).

Cigarette smoking can induce CYP1A2 activity, the major enzyme responsible for *R*-warfarin metabolism. With increased smoking, the CYP1A2 increases its metabolism of *R*-warfarin, which translates into a need for higher dosages. During longer non-smoking periods, the enzyme activity slows down and thus less warfarin is needed. Therefore, a change in smoking habit may affect warfarin coagulation response, for example after smoking cessation, patients need to be carefully monitored and warfarin doses reduced accordingly (Faber and Fuhr 2004).

Adherence to warfarin also clearly affects the degree of anticoagulant control. Patients with missed doses would have an increased risk of under-anticoagulation while patients who had extra doses would be at increased risk of over-anticoagulation (Kimmel *et al.,* 2007).

## 1.13 Genetic factors affecting warfarin dose requirements

Over 30 genes have been postulated to alter warfarin response. Figure 1.4 illustrates an overview of their interactive pathways with warfarin and Table 1.4 details their protein functions (see Figure 1.4). However, two genes, *CYP2C9* and *VKORC1*, have consistently been reported to play a key role in warfarin response.

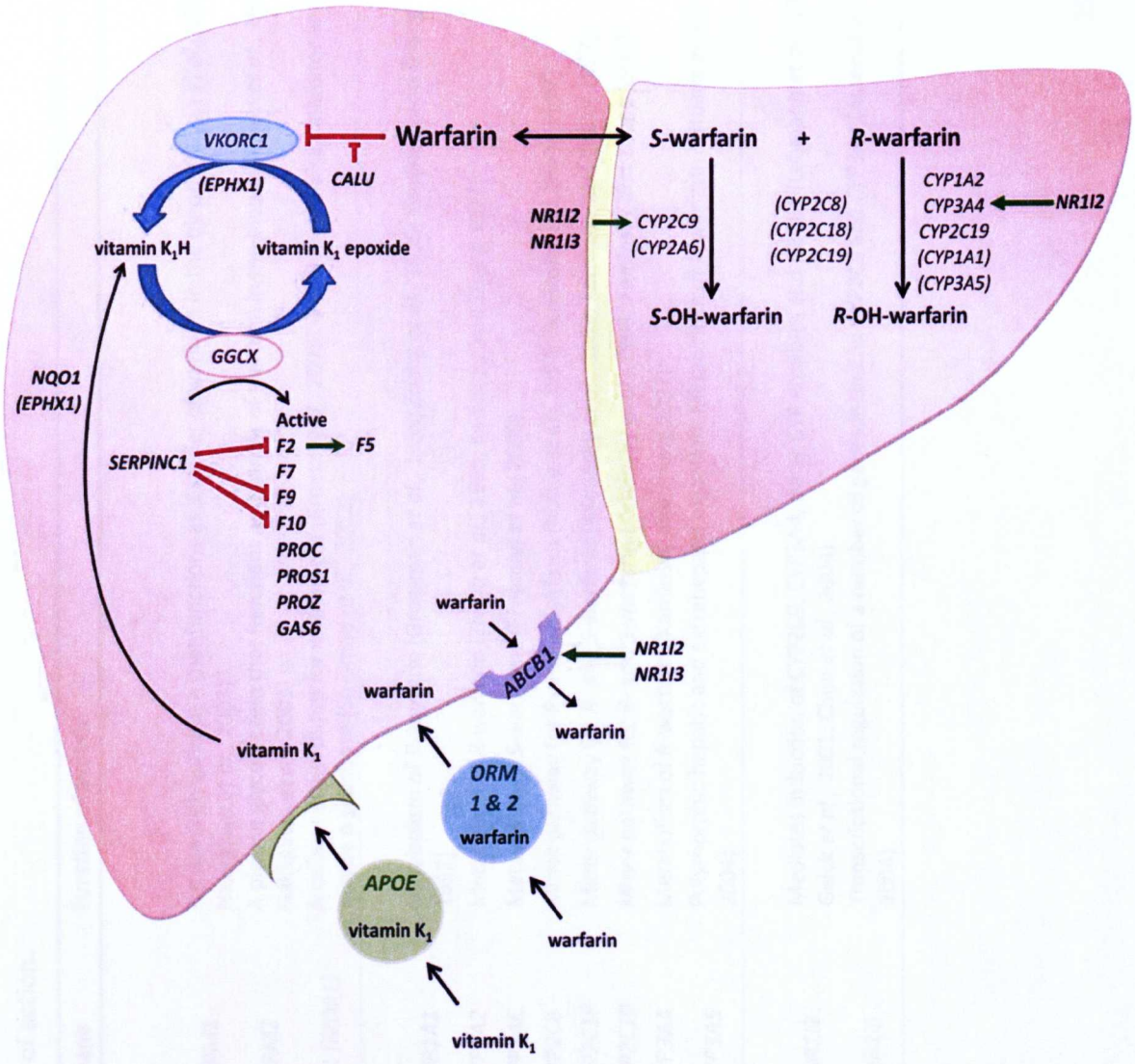



**Figure 1.4. An overview of warfarin interactive pathways.** This figure illustrates the genes thought to be involved in the action and biotransformation of warfarin and vitamin K. Adapted from Wadelius *et al.*, 2006.

**Table 1.4. Other genes involved in warfarin mechanism of action.**

| Protein name | Gene | Function of protein |
|---|---|---|
| **Biotransformation of warfarin** | | |
| **Transport** | | |
| Alpha-1-acid glycoprotein 1, Orosomucoid 1 | *ORM1* | A plasma glycoprotein that functions as a carrier of warfarin in the blood (Otagiri *et al.*, 1987; Nakagawa *et al.*, 2003) |
| Alpha-1-acid glycoprotein 2, Orosomucoid 2 | *ORM2* | A plasma glycoprotein that functions as a carrier of warfarin in the blood (Otagiri *et al.*, 1987; Nakagawa *et al.*, 2003) |
| P-glycoprotein, Multidrug resistance protein 1 | *ABCB1 (MDR1)* | A cellular efflux pump for xenobiotics (Kroetz *et al.*, 2003). Warfarin is a weak inhibitor and maybe a substrate (Sussman *et al.*, 2002) |
| **Metabolism** | | |
| Cytochrome P450 1A1 | *CYP1A1* | Metabolism of *R*-warfarin (Grossman *et al.*, 1993; Zhang *et al.*, 1995; Kaminsky and Zhang 1997) |
| Cytochrome P450 1A2 | *CYP1A2* | Metabolism of *R*-warfarin (Zhang *et al.*, 1995; Kaminsky and Zhang 1997) |
| Cytochrome P450 2A6 | *CYP2A6* | Metabolism of *S*-warfarin? (Freeman *et al.*, 2000) |
| Cytochrome P450 2C8 | *CYP2C8* | Minor pathway for *R*- and *S*-warfarin (Rettie *et al.*, 1992; Kaminsky and Zhang 1997) |
| Cytochrome P450 2C18 | *CYP2C18* | Minor pathway for *R*- and *S*-warfarin (Kaminsky *et al.*, 1993; Kaminsky and Zhang 1997) |
| Cytochrome P450 2C19 | *CYP2C19* | Minor pathway for *R*- and *S*-warfarin (Kaminsky *et al.*, 1993; Kaminsky and Zhang 1997) |
| Cytochrome P450 3A4 | *CYP3A4* | Metabolism of *R*-warfarin (Kaminsky and Zhang 1997) |
| Cytochrome P450 3A5 | *CYP3A5* | Polymorphic hepatic and extrahepatic oxidation. Metabolism of *R*-warfarin? (Huang *et al.*, 2004) |
| **Cytochrome P450 inducibility** | | |
| Pregnane X receptor (PXR) | *NR1I2* | Mediates induction of CYP2C9, CYP3A4, other CYP enzymes and ABCB1 (Lehmann *et al.*, 1998; Geick *et al.*, 2001; Chen *et al.*, 2004) |
| Constitutive androstane receptor (CAR) | *NR1I3* | Transcriptional regulation of a number of genes including *CYP2C9* and *CYP3A4* (Assenat *et al.*, 2004) |

**Table 1.4. Other genes involved in warfarin mechanism of action continued.**

| Protein name | Gene | Function of protein |
|---|---|---|
| **Biotransformation of vitamin K** | | |
| **Transport** | | |
| Apolipoprotein E | *APOE* | Apolipoprotein E serves as a ligand for receptors that mediate the uptake of vitamin K (Saupe *et al.*, 1993; Kohlmeier *et al.*, 1996; Lamon-Fava *et al.*, 1998; Berkner and Runge 2004) |
| **Vitamin K cycle** | | |
| Epoxide hydrolase 1, microsomal | *EPHX1* | A hepatic epoxide hydrolase in the endoplasmic reticulum that may be complexed with VKOR (Cain *et al.*, 1997; Loebstein *et al.*, 2005; Morisseau and Hammock 2005) |
| NAD(P)H dehydrogenase, quinone 1 | *NQO1* | A detoxifying enzyme that has the potential to reduce the quinine form of vitamin K (Wallin and Hutson 1982; Berkner and Runge 2004; Ross and Siegel 2004) |
| Calumenin | *CALU* | Binds to the vitamin K epoxide reductase complex and inhibits the effect of warfarin (Wallin *et al.*, 2001; Wajih *et al.*, 2004) |
| Gamma-glutamyl carboxylase | *GGCX* | Carboxylates vitamin K-dependent coagulation factors and proteins in the vitamin K cycle (Wu *et al.*, 1997; Berkner 2000; Rost *et al.*, 2004b) |
| **Vitamin K-dependent proteins** | | |
| Coagulation factor II, prothrombin | *F2* | Converts fibrinogen to fibrin, activates FV, FVIII, FXI, FXIII, protein C (Berkner 2000; Dahlbäck 2005) |
| Coagulation factor VII | *F7* | Is converted to FVIIa and then converts FIX to FIXa and FX to FXa (Berkner 2000; Dahlbäck 2005) |
| Coagulation factor IX | *F9* | Makes a complex with FVIIIa and then converts FX to its active form (Berkner 2000; Dahlbäck 2005) |
| Coagulation factor X | *F10* | Converts FII to FIIa in the presence of factor Va (Berkner 2000; Dahlbäck 2005) |

**Table 1.4. Other genes involved in warfarin mechanism of action continued.**

| Protein name | Gene | Function of protein |
|---|---|---|
| Protein C | *PROC* | Activated protein C counteracts coagulation together with protein S by inactivating FVa and VIIIa (Berkner 2000; Dahlbäck 2005) |
| Protein S | *PROS1* | Cofactor to protein C that degrades coagulation factors Va and VIIIa (Berkner 2000; Dahlbäck 2005) |
| Protein Z | *PROZ* | Is together with protein Z-dependent protease inhibitor, a cofactor for the inactivation of FXa (Berkner 2000; Broze G.J 2001) |
| Growth-arrest-specific protein 6 | *GAS6* | Participates in many processes, for example, potentiation of agonist-induced platelet aggregation (Berkner and Runge 2004) |
| **Other coagulation proteins** | | |
| Anti-thrombin III | *SERPINC1* | Inhibits FIIa, FIXa, Xa, XIa and XIIa. Anti-thrombin deficiency increases risk of thrombosis (Dahlbäck 2005) |
| Coagulation factor V | *F5* | A cofactor that activates FII together with FXa. An *F5* mutation leads to risk of thrombosis (Dahlbäck 2005) |

Table adapted from Wadelius and Pirmohamed., 2007

### 1.13.1 *CYP2C9*

The cytochrome P450 superfamily are the major enzymes involved in drug metabolism, accounting for approximately 75% of the total metabolism (Ingelman-Sundberg 2004). *CYP2C9* was the first gene documented to affect warfarin dose requirement (Furuya *et al.*, 1995). The *CYP2C9* gene is located on chromosome 10q24.2, spans approximately 55 kb, contains 9 exons, and encodes a 60 kDa microsomal protein (Goldstein and de Morais 1994).

The frequency of the wild-type allele, designated *CYP2C9*1*, varies from 81-96% in different ethnic groups (Table 1.5) (Gage *et al.*, 2004; Sanderson *et al.*, 2005; Voora *et al.*, 2005; D'Andrea *et al.*, 2008). The most common *CYP2C9* functional variant alleles are *CYP2C9*2* (rs1799853) and *CYP2C9*3* (rs1057910). *CYP2C9*2* induces an Arg144Cys amino substitution in exon 3 while *CYP2C9*3* encodes for an Ile359Leu amino acid change in exon 7. The allelic frequencies of *CYP2C9*2* and *CYP2C9*3* diverge considerably among different ethnic groups. *CYP2C9*2* occurs at a frequency of 6-14% in Caucasians but are less frequent in Asian and African-American populations (Table 1.5). Indeed, *CYP2C9*2* has not been reported in Asians and only 2-4% of African Americans carry this allele. *CYP2C9*3* is present in 6-10% of Caucasians, 1-4% of Chinese, Korean and Japanese populations, and 1-2% of African-Americans (Stubbins *et al.*, 1996; Takahashi *et al.*, 1998; Higashi *et al.*, 2002; Lee *et al.*, 2002; Xie *et al.*, 2002; Gage *et al.*, 2004; Aquilante *et al.*, 2006; Yin and Miyata 2007; Wu *et al.*, 2008).

*In vitro* studies have shown that the *CYP2C9*2* and **3* polymorphisms are associated with reduced metabolic efficiency and warfarin clearance (Rettie *et al.*, 1994; Haining *et al.*, 1996; Sullivan-Klose *et al.*, 1996). Possession of either of these

genes, therefore, is associated with lower warfarin dose requirements, compared to individuals carrying the wild-type *CYP2C9*1* allele.

When compared with *CYP2C9* homozygous *1*1 wild-type carriers, individuals heterozygous for the *2 and *3 allele require 15-20% and 30-40% lower daily maintenance doses, respectively. Individuals homozygous for *2 and *3 require ~35% and ~75% lower daily maintenance doses, respectively, than do homozygous wild-type individuals (Higashi *et al.,* 2002; Xie *et al.,* 2002; Takahashi and Echizen 2003; Gage *et al.,* 2004; Ingelman-Sundberg 2004; Hillman *et al.,* 2005; Sanderson *et al.,* 2005; Voora *et al.,* 2005; Aquilante *et al.,* 2006; Lee *et al.,* 2006; Marsh *et al.,* 2006; Wu *et al.,* 2008; Lindh *et al.,* 2009).

Using multiple linear regression models, several observational studies have shown that *CYP2C9* polymorphisms account for ~10-15% of the variance in warfarin maintenance dosage (Gage *et al.,* 2004; Sconce *et al.,* 2005; Aquilante *et al.,* 2006; Carlquist *et al.,* 2006; Anderson *et al.,* 2007; Caldwell *et al.,* 2007; Wu *et al.,* 2008; Wadelius *et al.,* 2009).

In addition to association with dose requirements, patients with the *CYP2C9*2*, **3*, or a combination of both **2*3* variant alleles have been associated with longer times to INR and dose stabilization, supratherapeutic anticoagulation and a two- to three-fold higher risk of a serious or life-threatening bleeding event, particularly at the beginning of therapy (Aithal *et al.,* 1999; Margaglione *et al.,* 2000; Taube *et al.,* 2000; Higashi *et al.,* 2002; Bodin *et al.,* 2005a; Schwarz *et al.,* 2008; Wadelius *et al.,* 2009).

Numerous other *CYP2C9* polymorphisms with respect to their functional effects have also been reported (http://www.imm.ki.se/CYPalleles/cyp2c9.htm). In particular, *CYP2C9* alleles **4* (found in the Japanese), **5* and **6* (identified in

African-Americans) and **11* (rare in both Caucasians and African-Americans) were all found to lead to a reduction in warfarin dose requirement (Schwarz 2003; Tai *et al.*, 2005).

**Table 1.5. Frequencies of genetic variants in Caucasians, Asians and African-Americans.**

| Variant | Caucasian (%) | Asian (%) | African–American (%) |
|---|---|---|---|
| *CYP2C9 *1* | 80 | 96 | 94.7 |
| *CYP2C9 *2* | 6-14 | 0 | 2-4 |
| *CYP2C9 *3* | 6-10 | 1-4 | 1-2 |
| *VKORC1* (group A) | 35–37 | 83–89 | 10–23 |
| *VKORC1* (group B) | 58–64 | 10–13 | 49–80 |

Adapted from Stubbins *et al.*, 1996; Takahashi *et al.*, 1998; Higashi *et al.*, 2002; Lee *et al.*, 2002; Xie *et al.*, 2002; Gage *et al.*, 2004; Sanderson *et al.*, 2005; Voora *et al.*, 2005; Veenstra *et al.*, 2005; Aquilante *et al.*, 2006; Limdi *et al.*, 2007; Yin and Miyata, 2007; D'Andrea *et al.*, 2008; Limdi *et al.*, 2008; Wu *et al.*, 2008.

### 1.13.2 *VKORC1*

The *VKORC1* gene is located on chromosome 16p11.2, and is approximately 4 kb long (Li *et al.*, 2004; Rost *et al.*, 2004a). Since its cloning in 2004, a series of different rare *VKORC1* mutations have been identified in patients resistant to warfarin, who required warfarin doses up to 20-fold higher than average to achieve an anticoagulant effect. Common genetic polymorphisms in the *VKORC1* have also been identified in different studies and all showed an association with warfarin dose variability.

The *VKORC1* intron 1 polymorphism c.1173C>T (rs9934438) was first identified by D'Andrea and colleagues in 2005, and was associated with ~3 mg/day lower warfarin dosage requirements (D'Andrea *et al.*, 2005). Soon after, this intronic c.1173C>T polymorphism was shown to be in LD with a new polymorphism c.-1639G>A (rs9923231) occurring at the second nucleotide of an E-box (CANNTG) in the 5' untranslated (UTR) region of the gene (Yuan *et al.*, 2005). This

polymorphism is associated with low mRNA levels in liver specimens (Rieder *et al.,* 2005). Some studies have also reported that this SNP reduces *VKORC1* gene promoter activity (Yuan *et al.,* 2005) but this has not been confirmed by others (Bodin *et al.,* 2005a).

Rieder and colleagues (2005) resequenced the entire *VKORC1* gene in 186 European-Americans on long-term warfarin maintenance therapy. Using 10 common non-coding SNPs, five common haplotypes (H1, H2, H7, H8 and H9) with > 5% frequency were inferred and these haplotypes were segregated into two groups according to their association with warfarin dose requirement. Group A, comprising H1 and H2, was associated with low warfarin dosage, and group B, comprising H7, H8 and H9, was associated with high dosage; and significant differences in warfarin maintenance dose were observed between the three combinations of haplotype groups: A/A (2.7 mg/day), A/B (4.9 mg/day), and B/B (6.2 mg/day), explaining 25% of dose variance (Rieder *et al.,* 2005). Different frequencies of *VKORC1* allele distribution have been observed among ethnic groups. The prevalence of the AA haplotype group is only 10-23% in African-Americans, but slightly higher in Caucasians (35–37%), and much higher, approximately 83–89%, in Asian subjects. On the other hand, the haplotype group B has a higher prevalence in the Caucasian population, approximately 58–64%, but is only found in 10–13% of Asian subjects (Table 1.5) (Rieder *et al.,* 2005; Veenstra *et al.,* 2005; Limdi *et al.,* 2008a).

Among the 10 SNPs that constitute the haplotypes, 5 polymorphisms are strongly correlated and include the intronic c.1173C>T and promoter c.-1639G>A SNPs. Association between warfarin dose requirements and these two SNPs (c.1173C>T, rs9934438; c.-1639G>A, rs9923231) have been confirmed in other studies in populations from around the world (Italy, Sweden, Hong Kong, Japan,

Singapore, Taiwan, UK, Brazil, US) (Sconce *et al.,* 2005; Wadelius *et al.,* 2005; Yuan *et al.,* 2005; Carlquist *et al.,* 2006; Lee *et al.,* 2006; Li *et al.,* 2006; Mushiroda *et al.,* 2006; Takahashi *et al.,* 2006; Perini *et al.,* 2008; Klein *et al.,* 2009).

Using multiple linear regression models, *VKORC1* polymorphisms have been shown to be strongly predictive of ~20-35% of warfarin dose variability (Sconce *et al.,* 2005; Aquilante *et al.,* 2006; Carlquist *et al.,* 2006; Herman *et al.,* 2006; Tham *et al.,* 2006; Perini *et al.,* 2008; Wadelius *et al.,* 2009). In addition to their effects on dose requirements, variant *VKORC1* alleles have been associated with shorter period of time required to achieve therapeutic INR, as well as first out-of-range INR (INR > 4) (Schwarz *et al.,* 2008). Association with an increased risk of bleeding has also been reported in some studies (Reitsma *et al.,* 2005; Schwarz *et al.,* 2008; Wadelius *et al.,* 2009) but not in others (Crawford *et al.,* 2007; Limdi *et al.,* 2008b).

### 1.13.3 *CYP4F2* and other genetic factors

In addition to *VKORC1* and *CYP2C9*, an exon 2 genetic variation (rs2108622) in the cytochrome P450 4F2 (*CYP4F2*) gene was recently found to be associated with increased warfarin dose requirements, accounting for 1-7% of dose variance (Caldwell *et al.,* 2008; Borgiani *et al.,* 2009; Takeuchi *et al.,* 2009).

Other genes that might have influential effects on warfarin response include those involved in the biotransformation of vitamin K, warfarin and the vitamin K-dependent clotting factors (see Figure 1.4 and Table 1.4). Few studies have investigated the contribution to warfarin dose variability from genes coding for γ-glutamyl carboxylase (*GGCX*), clotting factors such as factors II, VII, IX and X (*F2*, *F7*, *F9* and *F10*), as well as apolipoprotein E (*ApoE*), calumenin (*CALU*), microsomal epoxide hydrolase (*EPHX1*) and P-glycoprotein (*ABCB1*).

Two studies have shown that a polymorphism (rs762684) and a microsatellite (CAA repeat) both located within intron 6 of *GGCX* have a modest effect on warfarin dose requirements (Shikata *et al.,* 2004; Wadelius *et al.,* 2005). In addition, one study reported an association of a nonsynonymous *GGCX* Arg325Glu polymorphism (rs699664) in exon 8 with warfarin dose variability (Kimura *et al.,* 2007) but two other studies did not show this (Loebstein *et al.,* 2005; Wadelius *et al.,* 2005).

A nonsynonymous polymorphism in *F2* (rs5896) was found to lead to increased warfarin sensitivity in two independent studies (D'Ambrosio *et al.,* 2004; Shikata *et al.,* 2004), whereas a third study did not show this (Aquilante *et al.,* 2006). It has also been reported that promoter polymorphisms in *F7* have an effect on warfarin sensitivity (D'Ambrosio *et al.,* 2004; Shikata *et al.,* 2004; Aquilante *et al.,* 2006). Mutations in the propeptide of *F9* which cause an amino acid change from alanine to valine or threonine at residue -10, lead to a rapid drop in factor IX during warfarin treatment and have been associated with rare cases of bleeding (Kristensen 2002; van der Heijden *et al.,* 2004).

Apolipoprotein E is a very low-density lipoprotein (VLDL) which facilitates cellular uptake of chylomicrons, the main vehicle of vitamin K transport to liver (Lamon-Fava *et al.,* 1998). The *ApoE *E4* allele has been found to be associated with increased warfarin dose requirements, suggesting patients carrying the **E4* allele have enhanced vitamin K uptake into the liver (Kohnke *et al.,* 2005; Sconce *et al.,* 2006; Kimmel *et al.,* 2008).

The endoplasmic reticulum chaperone protein, calumenin, can bind to the vitamin K cycle and inhibit its activity (Wallin *et al.,* 2001; Wajih *et al.,* 2004). Thus

far, only one study has shown that a nonsynonymous polymorphism in *CALU* (rs2290228) is related to warfarin dose requirements (Vecsler *et al.,* 2006).

Human microsomal epoxide hydrolase (encoded by *EPHX1*) is a xenobiotic-metabolizing enzyme that detoxifies reactive epoxides to more water-soluble compounds. It resides in the endoplasmic reticulum, and earlier findings have shown that it harbours a vitamin K 2,3-epoxide binding site (Cain *et al.,* 1997), suggesting it may complex with VKOR to produce a multiprotein complex that is responsible for vitamin K epoxide reduction (Morisseau and Hammock 2005). Interestingly, a coding *EPHX1* polymorphism (rs1051740) has been associated with high warfarin doses in CYP2C9 extensive metabolisers (Loebstein *et al.,* 2005).

Warfarin has been shown to be a substrate for P-glycoprotein, an efflux membrane transporter, suggesting P-glycoprotein contributes to warfarin disposition (Sussman *et al.,* 2002). P-glycoprotein is encoded by the adenosine triphosphate (ATP)-binding cassette transporter B1 (*ABCB1*) gene. A synonymous variant within exon 26 of *ABCB1*, rs1045642, was found to modulate warfarin sensitivity (Wadelius *et al.,* 2004).

## 1.14 Warfarin dosing algorithms

Due to the multifactorial nature of warfarin response, the concept of dosing algorithms utilising clinical variables, such as age, body weight, gender, concurrent medication and indication of warfarin regimen, to improve anticoagulation management, reduce complications, and enhance efficacy has existed for decades (Theofanous and Barile 1973; Williams and Karl 1979; Ovesen *et al.,* 1989). Computer programs that incorporate clinical variables have been developed and studies have demonstrated that their usage helps to maintain a more stable INR and

reduce complication rates in both the induction and maintenance phase (Ageno *et al.,* 2000; Manotti *et al.,* 2001).

Since the discovery of *CYP2C9* and *VKORC1* genes, numerous studies have examined their association with warfarin dose requirements. Patients carrying the *CYP2C9*2*, *CYP2C9*3* or/and *VKORC1* -1639 genetic variations have consistently been found to be "sensitive" to warfarin and require lower warfarin dose (Gage *et al.,* 2004; Bodin *et al.,* 2005a; D'Andrea *et al.,* 2005; Sconce *et al.,* 2005; Carlquist *et al.,* 2006; Takahashi *et al.,* 2006). Noting the importance of these observations, the Food and Drug Administration (FDA) updated the warfarin prescription label in 2007 to include the effect of *CYP2C9* and *VKORC1* polymorphisms on warfarin initial dosing.

Many studies have also developed and tested dosing algorithms that incorporate both genotype and clinical characteristics with warfarin dosing requirement. Results from selected studies are summarised in Table 1.6. On average, SNPs in *CYP2C9* and *VKORC1* contribute approximately 30-40% warfarin dose variability, and together with clinical variables, they account for nearly up to 60% of dose variance (Bodin *et al.,* 2005a; Aquilante *et al.,* 2006; Carlquist *et al.,* 2006; Klein *et al.,* 2009; Wadelius *et al.,* 2009). The largest study to date was published in 2009 by the International Warfarin Pharmacogenetics Consortium (IWPC). They derived both a clinical and a pharmacogenetic dosing model from a cohort of over 4043 patients from 21 research groups across 9 countries and subsequently validated the dosing models in a separate group of 1008 patients (Klein *et al.,* 2009). Their results demonstrated that the addition of genetic information (*CYP2C9* and *VKORC1* SNPs) provided a dosage prediction that was significantly closer to the actual dosage required than estimates derived from a clinical algorithm or the fixed-dose approach.

Based on these findings, the FDA made a second revision to the drug label for warfarin in January 2010 to include dose ranges based on pharmacogenetic information on *CYP2C9* and *VKORC1*.

**Table 1.6. Contribution of genetic variants and clinical variables† to warfarin dosage.**

| Study | *VKORC1* (%) | *CYP2C9* (%) | Gene + clinical† (%) |
|---|---|---|---|
| Bodin *et al.*, 2005 | 37 | 14 | 54 |
| Sconce *et al.*, 2005 | 15 | 17 | 54.2 |
| Aquilante *et al.*, 2006 | 28.8 | 11.4 | 51.4 |
| Carlquist *et al.*, 2006 | 15.1 | 18.3 | 44.6 |
| IWPC, 2009‡ | 27.7 | 5.5 | 31.4 |
| Wadelius *et al.*, 2009‡ | 29.3 | 11.8 | 58.7 |

†Clinical variables include age, gender, drug interaction, ethnicity, body mass index (BMI).

‡Percentage of genetic contribution in univariate analysis.

IWPC: International Warfarin Pharmacogenetics Consortium.

## 1.15 Thesis aims

Although warfarin is highly effective, it is a challenging drug to prescribe due to its narrow therapeutic range and wide inter-individual variability. A multitude of studies have now clearly established that clinical factors such as age, gender, body weight, height and concomitant medications contribute to ~15-20% warfarin dose variability; and genetic variations in *CYP2C9* and *VKORC1* account for an even higher variation in dose requirements, ~30-40%. Despite incorporating both clinical and genetic factors, the best pharmacogenetic algorithm developed utilising a prospective cohort of patients explained only ~60% of the variance in dosing (Wadelius *et al.*, 2009). Up to 40% of warfarin dose variance still remains poorly understood, suggesting that other factors might be involved in warfarin's pharmacological effect.

A nonsynonymous polymorphism in *CYP4F2* (rs2108622) has recently been reported to explain some of the variance in warfarin stable dose requirements (2-7%), but studies published were of retrospective nature and other warfarin-related clinical outcomes were not investigated (Caldwell *et al.,* 2008; Borgiani *et al.,* 2009; Takeuchi *et al.,* 2009). Furthermore, the implications of other SNPs in *CYP4F2* in relation to warfarin response have not been investigated.

Although the IWPC pharmacogenetic algorithm was more accurate than a clinical algorithm, it accurately identified < 50% of patients needing low doses and < 25% of patients needing high doses (Figure 1.5) (Klein *et al.,* 2009). This suggests that there remains substantial room for improvement in the understanding of factors predicting warfarin dose, especially for patients requiring high doses. A handful of studies have investigated the genetic factors affecting warfarin resistance and have reported that patients resistant to warfarin carry rare mutations in the *VKORC1* but their functional roles have not been explored.

Warfarin acts through vitamin K antagonism and thereby inhibits the formation of the vitamin K dependent clotting factors (II, VII, IX, X) and anticoagulants (Protein C, S and Z). However, the rate at which these decline, whether this is variable between different individuals, and whether such variability is related to common polymorphisms in the clotting factor genes is not known.

To further improve the clinical management of warfarin, a better understanding of the factors described above is required. The aims of this PhD study were therefore (i) to fine map *CYP4F2* and assess the SNP-effects on warfarin dose variability and additional warfarin clinical outcomes in our cohort of prospective patients; (ii) to elucidate the functionality of *CYP4F2* as well as other *CYP4F* genes in the *CYP4F* cluster in human liver, and to characterise their genotype-phenotype

relationships; (iii) to identify mutations in the *VKORC1* gene in warfarin resistant patients via sequencing, and to explore the functions for any novel mutations identified; and (iv) to evaluate the levels of clotting factors and the rate they decline in our cohort of prospective patients, and relate this to SNPs in the clotting factor genes and warfarin clinical outcomes.

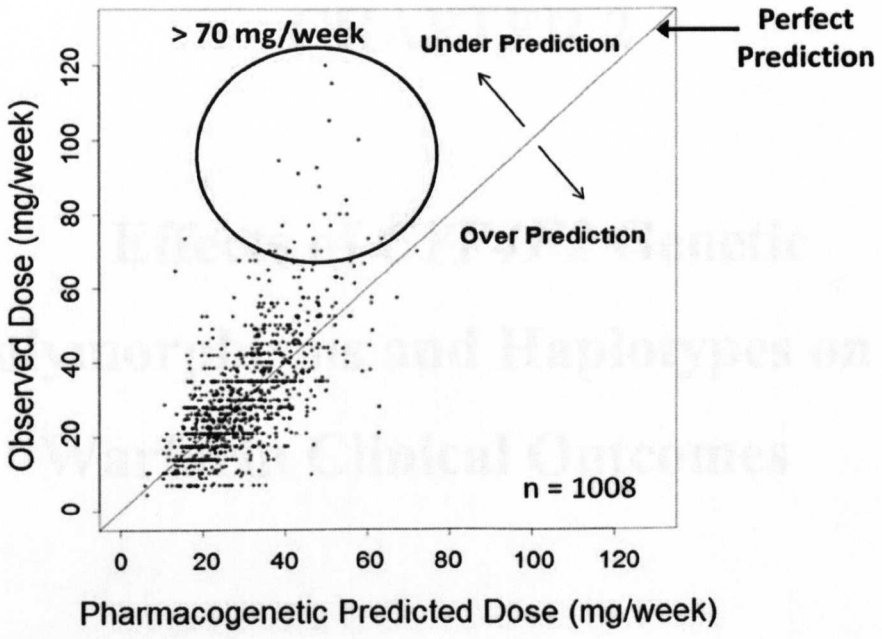



**Figure 1.5. Scatter plot showing the accuracy of the pharmacogenetic algorithm developed by Klein *et al.*, 2009.** Predicted versus observed stable therapeutic warfarin dose for 1008 patients in the validation cohort from the study by Klein *et al.*, 2009 are shown. Each dot represents one patient. Diagonal solid line indicates perfect prediction. Dots below and above the solid line indicate patients were over and under predicted by the pharmacogenetic algorithm, respectively. Patients requiring warfarin doses of > 70 mg/week are circled in red and the pharmacogenetic algorithm did not accurately identify these patients. Adapted from Klein *et al.*, 2009.

# CHAPTER 2

# Effects of *CYP4F2* Genetic Polymorphisms and Haplotypes on Warfarin Clinical Outcomes

# CONTENTS

**2.1 INTRODUCTION** ..... **42**

**2.2 PATIENTS, MATERIALS AND METHODS** ..... **44**

2.2.1 Patients and data collection ..... 44

2.2.2 Clinical outcome measures ..... 45

2.2.3 Interim analysis ..... 46

2.2.4 SNP selection and fine mapping ..... 46

2.2.5 *CYP4F2* primer design ..... 47

2.2.6 *CYP4F2* genotyping ..... 47

2.2.7 *CYP2C9* and *VKORC1* genotyping ..... 51

2.2.8 Statistical analysis ..... 51

*2.2.8.1 Univariate analysis* ..... *52*

*2.2.8.2 Multiple regression analysis* ..... *52*

*2.2.8.3 Haplotype analysis* ..... *54*

*2.2.8.4 False discovery rate* ..... *55*

**2.3 RESULTS** ..... **56**

2.3.1 Patient characteristics ..... 56

2.3.2 Univariate analysis of association between SNPs and outcomes ..... 58

2.3.3 Multiple regression models ..... 60

2.3.4 Association of rs2108622 with stable dose ..... 61

2.3.5 Haplotype Associations ..... 65

**2.4 DISCUSSION** ..... **67**

## 2.1 INTRODUCTION

Caldwell and colleagues (2008) recently utilised the Affymetrix drug-metabolizing enzymes and transporters (DMET) genotyping panel in 429 Caucasian patients stabilized on warfarin to identify other potential predictors of dose requirements. Out of the 1228 SNPs genotyped, a non-synonymous SNP in *CYP4F2*, rs2108622, was associated with stable warfarin dose. This association was validated in additional patient cohorts, and accounted for approximately 2% of the variability in therapeutic warfarin dose. The rank order from the lowest to highest doses was CC>CT>TT amongst the genotype groups (Caldwell *et al.,* 2008). In a retrospective Italian patient population, Borgiani and colleagues have also shown that TT patients require a warfarin dose that is approximately 2.5 mg/day higher than patients with the CC genotype, with rs2108622 explaining about 7% of warfarin dosing variability (Borgiani *et al.,* 2009). More recently, a genome wide association study has also shown an association with *CYP4F2*, but only after adjusting for *CYP2C9* and *VKORC1* (Takeuchi *et al.,* 2009).

CYP4F2 is expressed in the kidney, liver, lung and white blood cells, and is involved in the ω-hydroxylation of arachidonic acid and leukotriene $B_4$ ($LTB_4$) (Powell *et al.,* 1998; Lasker *et al.,* 2000; Sontag and Parker 2002). A previous study has shown that variation at the SNP rs2108622 was associated with reduced ω-hydroxylation of arachidonic acid, but did not affect $LTB_4$ metabolism (Stec *et al.,* 2007). Caldwell and colleagues (2008) postulated that *CYP4F2* may be involved in vitamin K metabolism. Indeed, a recent study has shown that *CYP4F2* is capable of metabolising vitamin $K_1$, with patients who have the rs2108622 variant having lower vitamin $K_1$ oxidase activity, which would result in higher hepatic vitamin $K_1$ levels, and therefore higher dose requirements for warfarin (McDonald *et al.,* 2009).

Given the association between *CYP4F2* and warfarin dose requirements, and the emerging evidence that this P450 isoform is involved in the metabolism of vitamin $K_1$, a comprehensive analysis of *CYP4F2* SNPs and haplotypes in a prospectively recruited cohort of patients from two UK clinics was undertaken in this chapter. The specific aims were (i) to replicate the association of warfarin stable dose with *CYP4F2* rs2108622 (Caldwell *et al.,* 2008) in our cohort of patients initiated on warfarin; (ii) to fine map the whole *CYP4F2* region and define the haplotype structure of *CYP4F2*; and (iii) to perform a comprehensive analysis, including SNP and haplotype-based analyses, to assess for association between genetic variation within *CYP4F2* and several outcomes of warfarin response, including stable dose.

## 2.2 PATIENTS, MATERIALS AND METHODS

### 2.2.1 Patients and data collection

Between November 2004 and May 2007, patients (n = 1000) starting warfarin therapy (either while in-patients or from the anticoagulant clinic) were recruited prospectively from two hospitals in Liverpool, the Royal Liverpool and Broadgreen University Hospitals Trust and University Hospital Aintree. The main indications for warfarin therapy in our patients were treatment of venous thromboembolism and prophylaxis against systemic emboli in patients with atrial fibrillation. All patients recruited were ≥ 18 years of age. The study was approved by the Birmingham South research ethics committee and written informed consent was obtained from all patients.

The study design was observational. Patients received usual clinical care where the warfarin loading dose and subsequent maintenance doses were determined according to in-hospital guidelines. Each patient had four scheduled visits: the first was at the time of warfarin therapy commencement (index visit), then three subsequent follow-up visits at one week, eight weeks and twenty-six weeks of warfarin therapy. Patients also attended the anticoagulant clinic between these four fixed visits as per usual clinical practice, the frequency being determined clinically by the stability of their anticoagulation control. Throughout the study, any changes in warfarin doses and INRs were collected.

Data on ethnicity, age, gender, height, weight, indication for warfarin treatment, concomitant medication, co-morbidities were gathered from all patients, as were details of smoking history, current medications and alcohol intake. Some

patients, particularly those with venous thromboembolism, were initially treated with heparin, in which case dose and duration of treatment were recorded.

All adverse events which occurred during the study period were reported and assessed for causality. A bleeding event was defined as major or minor according to the classification by Fihn and colleagues (1996). A bleeding event was classed as serious if it was lethal, life-threatening, permanently disabling or led to hospital admission or prolongation of hospital stay (Fihn *et al.,* 1996). Only haemorrhagic events considered to be possibly, probably or definitely associated with warfarin were included in the analyses.

### 2.2.2 Clinical outcome measures

In order to capture both efficacy and toxicity of warfarin, two primary and several secondary outcome measures were chosen. The primary outcome measures were:

- incidence of INR > 4 in the first week on warfarin, and
- warfarin sensitivity (a dose of $\leq$ 1.5 mg/day on three successive clinic visits).

The secondary outcome measures were:

- warfarin resistance (a dose of $\geq$ 10 mg/day on three successive clinic visits),
- stable warfarin dose (an unchanged daily dose at three or more consecutive clinic visits where INR measurements were within the individual's target range). As the distribution of stable dose was skewed, the outcome was transformed by taking its square-root for the purpose of the analyses of association.

- time to therapeutic INR (defined as the time to achieving first INR measurement within the individual's target range, providing INR remained within the target range at the subsequent clinic visit),
- time to achieving warfarin stable dose from initiation, and
- bleeding complications associated with warfarin.

### 2.2.3 Interim analysis

An interim analysis on the first 311 patients recruited between November 2004 and March 2006 has previously been undertaken to evaluate the impact of genetic (29 genes but specifically *CYP2C9* and *VKORC1*) and clinical factors on determining the inter-individual variability in warfarin response (Jorgensen *et al.*, 2009). In this chapter, analyses were conducted using this cohort of 311 patients.

### 2.2.4 SNP selection and fine mapping

To replicate the association of rs2108622 with warfarin dose requirements and to search for other candidate SNPs in *CYP4F2*, a total of 80 SNPs spanning 52,563 bp across the chromosomal 19p13.11 region were selected. SNPs were a combination of tagging SNPs and functional variants. Selection was based on the following criteria: (i) known function from the literature and the National Centre for Biotechnology Information (NCBI) SNP database (build 126), (ii) SNP coverage in the CEU population (Utah residents with ancestry from northern and western Europe) available on HapMap data release 23, NCBI build 36 assembly, (iii) block-tagging ability using HaploView version 4.1 ($r^2 > 0.8$), and (iv) a minor allele frequency (MAF) 1% or greater.

### 2.2.5 *CYP4F2* primer design

Amplification and extension primers were designed using the eXTEND suite (www.realsnp.com/default.asp) in conjunction with the Sequenom Assay Designer software (version 3.1). The eXTEND suite consists of three functionalities that increases assay success rate and reduces genotyping errors: (1) ProxSNP – SNP sequences are mapped to the human genome for proximal SNPs and any SNPs in the vicinity of the denoted assay SNP will be demarked. (2) PreXTEND – Pairs of PCR primers are scanned against the human genome to ensure unique binding and amplification of target regions before SNP assays are designed. (3) PleXTEND – After assay design, the combination of all the multiplexed primers are screened for cross-binding possibilities that could lead to false positives. Three different multiplex PCR assays were developed which included 80 SNPs. All PCR primers had a tag sequence (ACGTTGGATG) to permit more efficient amplification. Amplification and extension primer sequences are listed in Appendix 1.1.

### 2.2.6 *CYP4F2* genotyping

Genomic DNA from ethylenediaminetetraacetic acid (EDTA) whole blood samples (n = 311) was extracted at the Sanger Institute using the standard technique incorporating proteinase K lysis, phenol-chloroform and ethanol precipitation methods.

Genotyping was performed on the Sequenom MassARRAY iPLEX platform (Sequenom, Hamburg, Germany), using Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight Mass Spectrometry (MALDI-TOF MS) technology (van den Boom and Ehrich 2007), in accordance with the manufacturer's instructions. Genotype determination is based on the mass separation of single base extension (SBE)

products with the incorporation of mass-modified dideoxynucleotide terminators. Each multiplex assay consisted of four reaction steps:

(i) pre-amplification by PCR

Multiplex PCRs (performed in a 384 well microtiter plate, final volume 5 µl) contained 10 ng of DNA, 0.1 µM of primers (Metabion GmbH, Germany), 500 µM of deoxyribonucleotide triphosphate (dNTP) mix, 1.625 mM of magnesium chloride ($MgCl_2$) 1.25x PCR buffer and 0.5 U of HotStar Taq® polymerase (PCR reagent set, Sequenom). PCR conditions were: denaturing step at 94°C for 15 min followed by 45 cycles at 94°C for 20 s, 56°C for 30 s and 72°C for 1 min, and a final extension at 72°C for 3 min.

(ii) shrimp alkaline phosphatase (SAP) treatment

Any excess or unincorporated dNTPs was dephosphorylated in a final volume of 7 µl with 0.03 U/µl of SAP in 0.17x SAP buffer (Sequenom), at 37°C for 40 min followed by 5 min at 85°C.

(iii) iPLEX™ primer extension

Following addition of 0.2 µl of iPLEX buffer, 0.2 µl of iPLEX terminator mix, and 0.04 µl of iPLEX enzyme (iPLEX® Gold reaction kit, Sequenom), extension primers (Metabion) were sorted into four groups based on mass and added to final concentration of 7, 9.33, 11.66 or 14 µM (final volume of 9 µl), with the highest mass group diluted to 14 µM and the lowest mass group to 7 µM. This adjustment in extension primer concentration is to ensure that the extension primers are as equal in intensity as possible, in order to account for variable signal-to-noise ratios due to the inverse relationship between peak intensity and analyte mass. Extension reaction conditions were: 94°C for

30 s; 40 cycles consisting of a 94°C step for 5 s and five subcycles of 52°C for 5 s and 80°C for 5 s; and a final extension at 72°C for 3 min.

(iv) a clean resin step

A total of 6 mg of Clean Resin (Sequenom) and 16 µl of water were added to the reactions followed by incubation for 20 min to remove the salts and a 20 min centrifugation at 3000 rpm. Samples need to be properly desalted to prevent sodium and potassium adducts from complicating accurate heterozygote allele discrimination of SNPs.

For MALDI-TOF MS analysis, samples were dispensed onto a 384 SpectroCHIP® Array (Sequenom) using a nanodispenser, and introduced into a MassARRAY® Compact mass spectrometer (Sequenom). Automated spectra acquisition was performed using SpectroAcquire. Data analysis was performed with MassARRAY® Typer software version 3.4 and some examples of genotype data output are illustrated in Figure 2.1. To ensure data quality, 10% duplicate DNAs and 8 negative controls (water) were included per 384-well plate and genotypes for each subject were also checked manually.

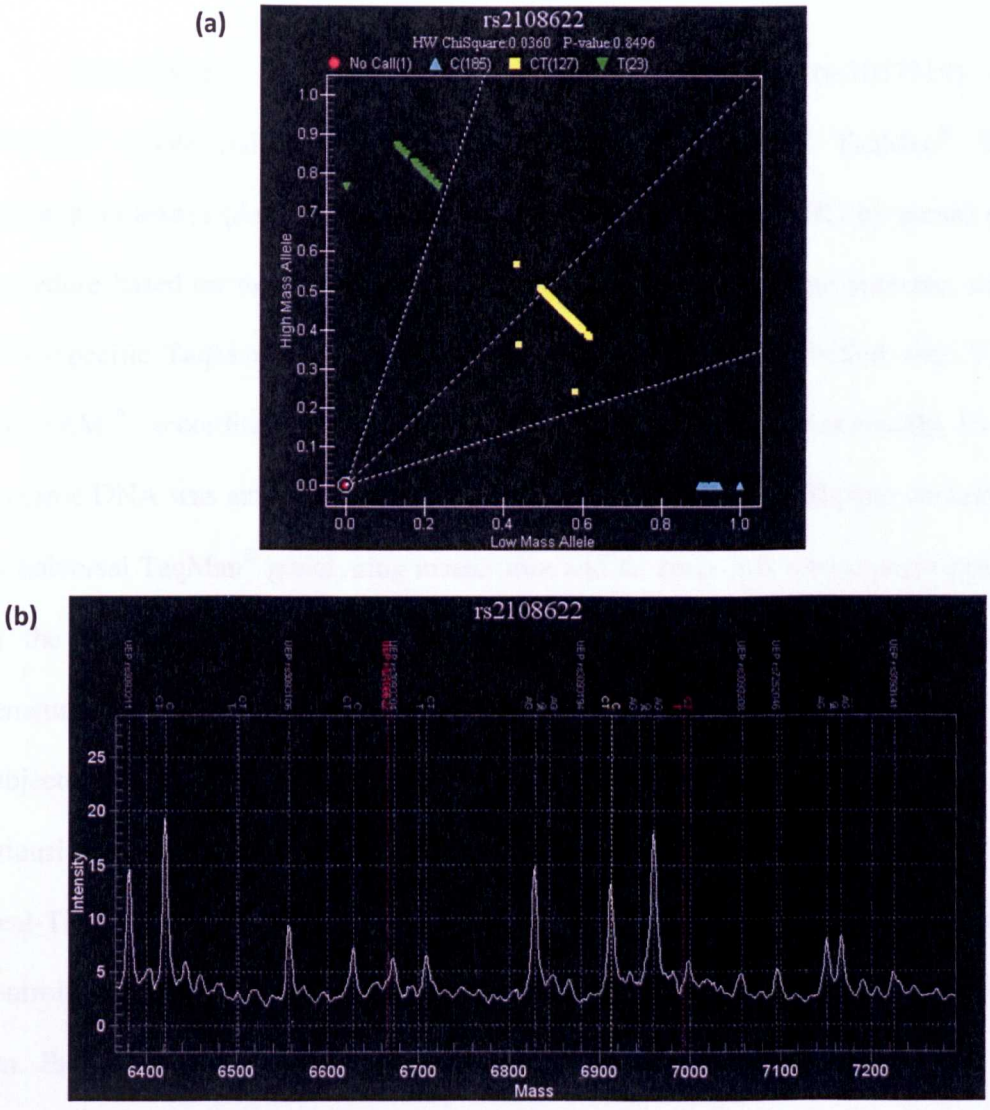



**Figure 2.1. Genotype data output from Sequenom MassARRAY® Typer.** (a) Cluster plot for SNP rs2108622, with the green triangle representing homozygous wild-type TT genotype, the yellow square representing heterozygous CT genotype and the blue triangle representing homozygous mutant CC genotype. (b) Spectrum of a sample homozygous for the rs2108622 minor C allele, where the raw data peak of the C allele is indicated by the dotted yellow line while the absence of the major T allele is indicated by the red dotted line.

### 2.2.7 *CYP2C9* and *VKORC1* genotyping

Genotyping of *CYP2C9*2* (rs1799853), *CYP2C9*3* (rs1057910) and *VKORC1* -1639 (rs9923231) were performed with Custom TaqMan® SNP Genotyping assays (Applied Biosystems, Warrington, Cheshire, UK) by means of a procedure based on the 5'–3' exonuclease activity of Taq DNA polymerase, using allele-specific TaqMan® minor groove binding (MGB) probes labelled with VIC® and FAM™, according to the manufacturer's instructions. Approximately 10 ng genomic DNA was amplified in 5 µl reaction mixture in a 384-well plate containing 1x universal TaqMan® genotyping master mix and 1x assay mix containing a premix of the respective primers and fluorescent-labelled MGB probes. After initial denaturation and enzyme activation at 95°C for 10 min, the reaction mixture was subjected to 40 cycles of denaturation at 95°C for 15 s and combined annealing and extension at 60°C for 1 min. The reactions were performed on an ABI 7900HT Fast Real-Time PCR System (Applied Biosystems). As part of quality control, negative controls containing water instead of DNA and 10% duplicates were included in every run. End-point fluorescence and allelic discrimination were determined using the SDS version 2.2 software (Applied Biosystems).

### 2.2.8 Statistical analysis

All statistical analyses were performed in the genetics package of R (http://cran.r-project.org/web/packages/genetics/index.html) or in SPSS version 16. Deviation from Hardy-Weinberg Equilibrium (HWE) was tested for each SNP using the Chi-Square test. A *P*-value < 0.001 was assumed to indicate deviation from HWE.

### *2.2.8.1 Univariate analysis*

To evaluate the individual effect of each SNP on each outcome, two univariate tests of association were conducted and the maximum test statistic referred to in each case. The first made no assumptions regarding the underlying mode of inheritance whilst the second assumed an additive mode. For the binary outcomes (INR > 4, warfarin sensitivity, warfarin resistance and bleeding events), the Pearson's chi-square test or Fisher's exact test and Cochrane-Armitage test for trend were used. For the time to event outcomes, the log-rank and log-rank test for trend were used. For the continuous outcome of stable warfarin dose, both analysis of variance (ANOVA) and univariate linear regression analyses were undertaken.

### *2.2.8.2 Multiple regression analysis*

The purpose of this analysis was to investigate whether there was an association between the *CYP4F2* genetic region as a whole and response to warfarin. For each outcome, two regression models were fitted and compared using the likelihood ratio test (LRT). The first model, the 'baseline model', included clinical factors found significant ($P < 0.05$) in univariate analyses, as reported previously in the interim analysis (Jorgensen *et al.*, 2009). A list of clinical covariates included for each outcome is shown in Table 2.1. The second model, the 'genetic model', was the same but also included covariates to represent all genotyped SNPs, with the exception of those excluded due to deviation from HWE (n = 10), due to being non-polymorphic (n = 3) and due to having a call rate below 90% (n = 8). To minimise the risk of co-linearity, only SNPs with the least missing genotype data from any group with correlation coefficient $r^2 > 0.9$ were included. For each outcome, the genetic model was fitted twice: first making no assumptions regarding the underlying

mode of inheritance and second assuming an additive mode of inheritance. The proportion of variability explained by the clinical and genetic covariates combined was calculated using Nagelkerke's $R^2$ statistic.

**Table 2.1. List of clinical covariates represented in the multiple regression models.**

| Outcome | Clinical factors represented in multiple regression models[a] | $R^2$ (%) |
|---|---|---|
| INR > 4 | Gender; BMI; clotting factor VII; warfarin loading dose | 17.1 |
| Warfarin Sensitivity | None | N/A |
| Warfarin Resistance | Age; warfarin loading dose | 32.6 |
| Time to stable dose | Clotting factor IX and indication for warfarin | 4 |
| Time to therapeutic INR | BMI; clotting factors VII; warfarin loading dose and ethnic origin | 5.9 |
| Stable warfarin dose | Age, BMI, clotting factors II, warfarin loading dose, and suffering from a neurological condition | 15.1 |
| All haemorrhagic events | None | N/A |
| Major haemorrhagic events | Clotting factor IX, warfarin loading dose and concurrent use of omeprazole | 13.5 |

BMI: Body Mass Index.
[a] Included on the basis that they were found significant (*P*-value < 0.05) in previous univariate analyses (Jorgensen *et al.*, 2009).

### *2.2.8.3 Haplotype analysis*

The pattern of LD between SNPs included in the multiple regression models were visualized using the HaploView software version 4.1 (Barrett *et al.*, 2005). Haplotype blocks were assigned using the internally developed Solid Spine of LD method, in which the two end markers are in strong LD with intervening markers but intervening markers are not necessarily in LD with each other (Barrett *et al.*, 2005). The most probable haplotype pair at each block was allocated to each patient using the PHASE software, version 2.1 (Stephens *et al.*, 2001; Stephens and Donnelly 2003). Any patients where the most likely haplotype-pair allocation had a probability of less than 90% for at least one haplotype block were excluded from the haplotype analysis (n = 8). For each outcome measure, 4 different regression models were built. Model 1, or the 'baseline model', included only the clinical covariates found significant ($P < 0.05$) in the previous univariate analyses. Model 2 included clinical covariates and covariates representing the different haplotypes observed for block 1. Model 3 included clinical covariates and covariates representing the different haplotypes observed for block 2. Model 4 included clinical covariates and covariates representing the different haplotypes observed both for haplotype block 1 and 2. To assess for association with the haplotype blocks, each of models 2-4 were compared to the baseline model using the LRT. In each model, to represent the different haplotypes within a haplotype block, the most common haplotype was assumed to be the baseline category and covariates were added to the model to represent each of the other observed haplotypes.

### *2.2.8.4 False discovery rate*

To account for multiple testing, the false discovery rate (FDR) (Benjamini *et al.,* 2001) was calculated in the genetics package of R, version 2.6.2 (http://cran.r-project.org/web/packages/genetics/index.html). In calculating the FDR, all tests for association undertaken on the dataset, including those referred to in the interim analysis (Jorgensen *et al.,* 2009), were taken into account. FDR-corrected *P*-values are denoted as $P_c$-values.

## 2.3 RESULTS

### 2.3.1 Patient characteristics

Demographics of the 311 patients including underlying co-morbidities are summarised in Table 2.2. The majority of patients were Caucasians with atrial fibrillation being the most common indication for warfarin therapy.

Of the 311 patients recruited, 57 (18%) experienced an INR greater than 4 during their first week on warfarin. 204 (66%) patients achieved stable dose during the course of follow-up, while 274 (88%) achieved therapeutic INR during the follow-up period. 68 (22%) patients experienced a haemorrhagic complication, of which 16 (6%) experienced major bleeding. Complete dosage information was unavailable for 38 patients. Hence, of the remaining 273 patients, 33 (12%) were sensitive to warfarin while 10 (4%) were warfarin resistant.

**Table 2.2. Clinical profile of 311 warfarin patients.**

| Characteristic | N (%) |
|---|---|
| **Gender - Male** | 184 (59) |
| **Age in years, mean (range)** | 66 (19-95) |
| **BMI[a], mean (range)** | 28 (15-61) |
| **Ethnicity** | |
| White | 305 (98) |
| Black African | 1 (0.3) |
| Black Caribbean | 1 (0.3) |
| Black other | 4 (1.3) |
| **Indication for warfarin** | |
| Atrial Fibrillation | 165 (53) |
| Pulmonary Embolism | 75 (24) |
| Deep Vein Thrombosis | 43 (14) |
| Cerebrovascular accident and Transient ischaemic attacks | 10 (3) |
| Myocardial infarction | 1 (0.3) |
| Heart Valve Replacement | 1 (0.3) |
| Other[b] | 16 (5) |
| **Co-morbidity** | |
| Cardiovascular disease | 208 (67) |
| Musculoskeletal problems | 142 (46) |
| Respiratory disease | 102 (33) |
| Gastrointestinal disease | 84 (27) |
| Neurological disease | 73 (23) |
| Urological condition | 45 (14) |
| History of falls | 27 (9) |
| Renal disease | 23 (7) |
| Hepatic disease | 9 (3) |

BMI: Body Mass Index.

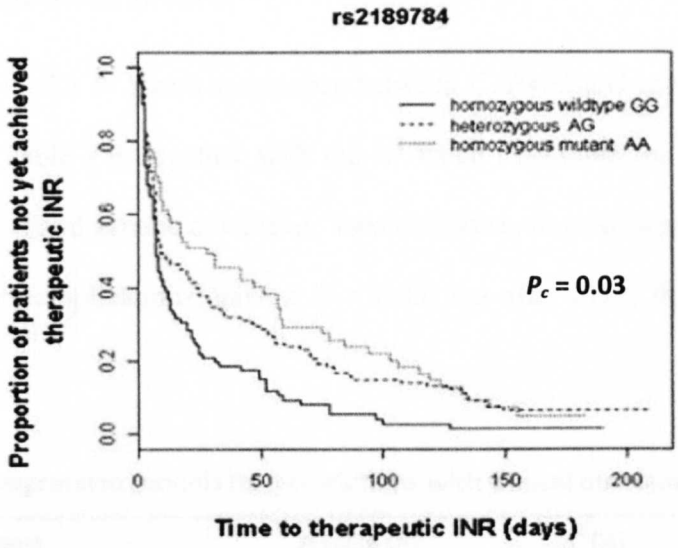
[a] BMI missing for 5 patients.

[b] Other indications include: prevention of clotting in arm for dialysis; axillary vein thrombosis; short saphenous vein thrombosis; valvular heart disease; sagittal sinus thrombosis; dilated left ventricle; occluded graft in leg; aortic and mitral regurgitation; poor liver function and pseudoaneurysm; ischaemic leg; brachial artery thrombosis; mitral stenosis; and post-surgery.

### 2.3.2 Univariate analysis of association between SNPs and outcomes

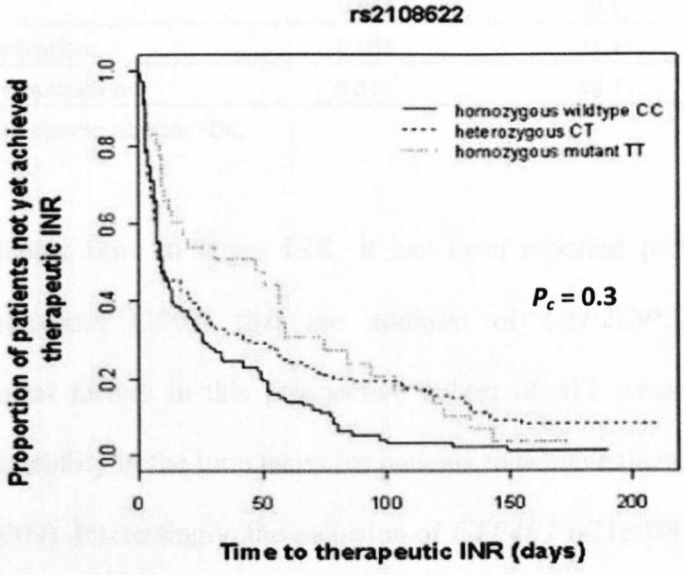
Of the 80 SNPs genotyped, 3 were monomorphic, 8 failed to give a call rate > 90%, and 10 deviated from HWE. Among the remaining 59 SNPs, 17 had been typed by the International HapMap project (data release 23, March 2008) and gave similar MAF. The allele frequencies of the SNPs investigated are listed in Appendix 1.2. Appendix 1.3 summarizes the results of the univariate analyses for each SNP-outcome combination where a $P < 0.05$ was obtained. Several SNP-outcome combinations gave a $P$-value below 0.05 including INR > 4 in the first week on warfarin, warfarin sensitivity, warfarin resistance, stable warfarin dose, time to therapeutic INR, time to achieving warfarin stable dose, and major bleeding complications. However, after correction for multiple testing using FDR which accounted for all tests performed in this study and in our analysis of 29 genes (Jorgensen *et al.,* 2009), only the association of rs2189784 with time to therapeutic INR remained significant ($P_c = 0.03$). Figure 2.2a shows the Kaplan-Meier curves of time to therapeutic INR, stratified by rs2189784 genotype. As the curves illustrate, patients with homozygous wild-type genotype (GG) required less time to achieve therapeutic INR than heterozygous patients, with patients with the homozygous mutant genotype (AA) taking the longest to achieve therapeutic INR. rs2189784 is in linkage disequilibrium with rs2108622 (D' = 0.98, $r^2 = 0.59$) and consistent with this, rs2108622 also showed an association with time to therapeutic INR (Figure 2.2b), but this did not remain significant after FDR.

(a)




| Time to therapeutic INR (days) | 0 | 50 | 100 | 150 | 200 |
|---|---|---|---|---|---|
| No. of patients at risk | | | | | |
| GG | 96 | 17 | 7 | 6 | 6 |
| AG | 149 | 48 | 31 | 22 | 22 |
| AA | 61 | 27 | 17 | 9 | 8 |

(b)




| Time to therapeutic INR (days) | 0 | 50 | 100 | 150 | 200 |
|---|---|---|---|---|---|
| No. of patients at risk | | | | | |
| CC | 140 | 33 | 16 | 14 | 14 |
| CT | 131 | 41 | 29 | 18 | 17 |
| TT | 35 | 18 | 10 | 5 | 5 |

**Figure 2.2. Kaplan-Meier curve showing time to therapeutic INR (days) for (a) rs2189784 and (b) rs2108622 genotypes.** INR indicates international normalized ratio. The tables beneath the graphs show the number of patients who had not yet achieved therapeutic INR at respective time points.

### 2.3.3 Multiple regression models

Results of the LRT to assess association between *CYP4F2* and each outcome are summarized in Table 2.3, together with the $R^2$ value calculated for the model including both clinical and genetic covariates. Two outcomes, time to target INR and major haemorrhagic complications, gave a $P < 0.05$, but after FDR, this was not significant.

**Table 2.3. Multiple regression models for associations with clinical outcomes.**

| Outcome | *P*-value LRT | $R^2$ (%) |
|---|---|---|
| INR > 4 | 0.19 | 35.7 |
| Warfarin sensitive | 0.657 | 9.7 |
| Warfarin resistant | 0.796 | 47.3 |
| Time to stable dose | 0.062 | 17.8 |
| Time to therapeutic INR | 0.006[a] | 18.2 |
| Stable warfarin dose | 0.995 | 19.6 |
| All haemorrhagic complications | 0.098 | 21.2 |
| Major haemorrhagic complications | 0.010[a] | 46.7 |

[a] *P*-value did not remain significant after FDR.

For the outcome time to target INR, it has been reported previously by Jorgensen and colleagues (2009) that the addition of *CYP2C9**2 genotype (rs1799853) to clinical factors in this prospective cohort of 311 patients did not contribute to any variability in the time taken for patients to achieve therapeutic INR (Jorgensen *et al.*, 2009). Interestingly, the inclusion of *CYP4F2* rs2189784 genotype increased the $R^2$ by approximately 5%. However, this contribution was not significant after FDR. The variability explained for these models is shown in Table 2.4.

**Table 2.4. Contribution of variables to time to therapeutic INR.**

| Covariate | $R^2$ (%) | *P*-value |
|---|---|---|
| Clinical only (BMI, clotting factor VII, warfarin loading dose, ethnic origin) | 6 | 0.000001[b] |
| Clinical plus *CYP2C9**2 | 6 | 0.137 |
| Clinical plus *CYP2C9**2 and *CYP4F2* rs2189784 | 11 | 0.001[a] |

[a] *P*-value did not remain significant after FDR.

[b] *P*-value remained significant after FDR.

### 2.3.4 Association of rs2108622 with stable dose

An important goal of our study was to try to replicate the association between SNP rs2108622 and warfarin stable dose (Caldwell *et al.*, 2008; Borgiani *et al.*, 2009; Takeuchi *et al.*, 2009). However, rs2108622 did not show a significant association with stable warfarin dose univariately in our study ($P = 0.4$). The mean weekly warfarin stable doses per rs2108622 genotype groups are presented in Figure 2.3 for the purpose of comparison with those presented in the paper by Caldwell *et al.* (2008). For all genotype groups the median weekly dose was 30 mg/week. In an attempt to reduce the heterogeneity between the patient populations, non-Caucasian patients (n = 7) were excluded but the association still remained non-significant ($P = 0.34$) with no change in warfarin dose across the genotype groups. Tests for association were also repeated in a subgroup of patients without congestive heart failure, renal disease, hepatic disease and cancer (n = 53), which all constitute pre-existing conditions specified as exclusion criteria by Caldwell *et al.* (2008). However, the result was again not significant ($P = 0.16$).

In a further attempt to replicate the analysis undertaken by Caldwell *et al.* (2008), the multiple regression model they presented, including covariates to represent gender, age, body mass index (BMI), target INR, *VKORC1* and *CYP2C9*

SNPs and the *CYP4F2* SNP rs2108622, was fitted to our dataset. BMI was used instead of body surface area (BSA) as the latter was not measured in the current study. Models using the LRT both including and excluding a covariate to represent rs2108622 were compared; however, no change to the $R^2$ value was observed suggesting that the SNP had no effect over and above that of the *CYP2C9* and *VKORC1* SNPs. The $R^2$ value also remained unchanged after excluding the 7 non-Caucasian patients. Restricting the analysis to the subgroup of patients without congestive heart failure, renal disease, hepatic disease and cancer also showed no change in the variability. A summary of the $R^2$ results is presented in Table 2.5.

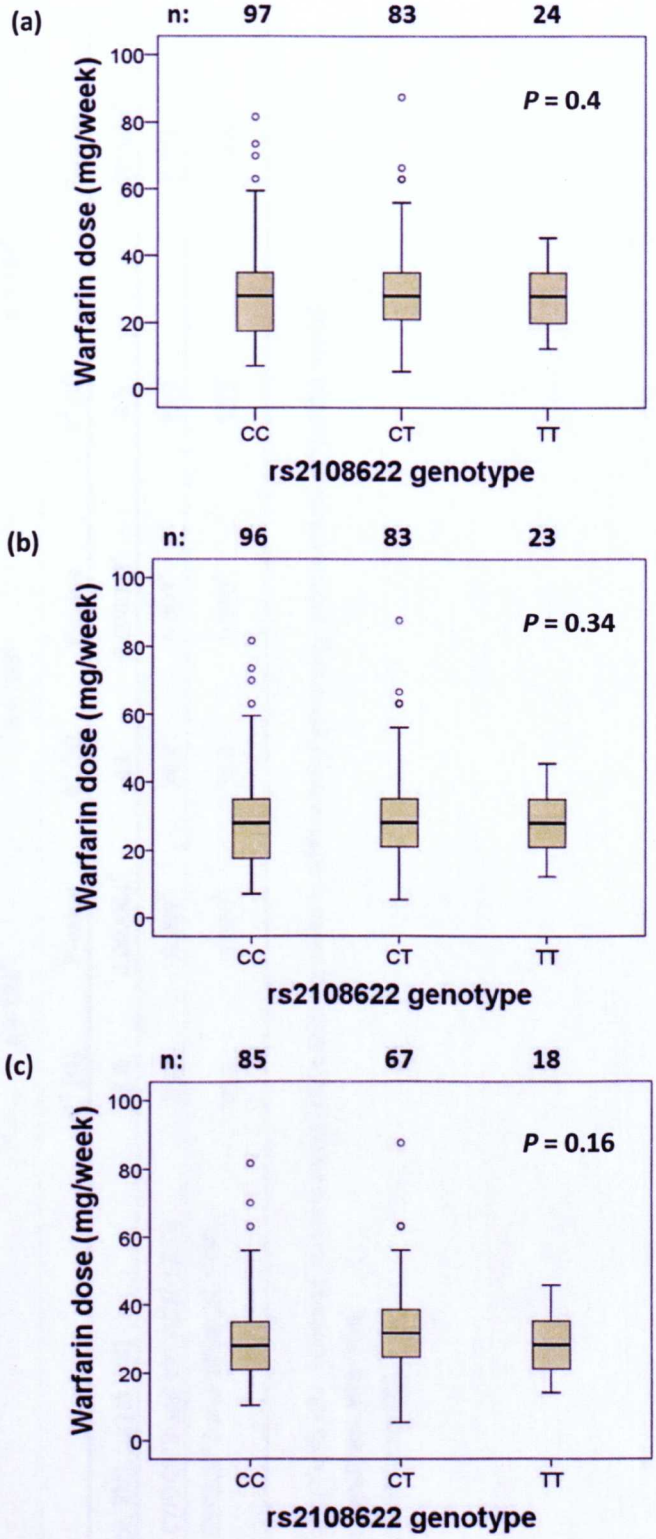



**Figure 2.3. Box and whisker plots showing the distribution of warfarin weekly doses based on *CYP4F2* rs2108622 genotype groups in (a) all prospective patients, (b) Caucasian patients only and, (c) Caucasian patients without congestive heart failure, renal disease, hepatic disease and cancer.** Boxes represent 25th-75th percentiles of warfarin doses, whiskers represent 5th-95th percentiles, solid lines represent median dose in each group, and open dots represent outliers.

**Table 2.5. Contribution of variables to warfarin stable dose requirement.**

| Covariate | All patients n = 191[a] | | Caucasians only n = 189[a] | | Caucasians without congestive heart failure, renal disease, hepatic disease and cancer n = 157[a] | |
|---|---|---|---|---|---|---|
| | $R^2$ (%) | *P*-value | $R^2$ (%) | *P*-value | $R^2$ (%) | *P*-value |
| Clinical only (gender, age, BMI, target INR) | 12.6 | 0.000001[c] | 12 | 0.000001[c] | 8.5 | 0.000086[c] |
| Clinical plus *CYP2C9**2, *CYP2C9**3 and *VKORC1*(-1639) | 39.8 | 0.009[b] | 39.6 | 0.009[b] | 33.3 | 0.056 |
| Clinical plus *CYP2C9**2, *CYP2C9**3 and *VKORC1*(-1639) and *CYP4F2* rs2108622 | 39.8 | 0.020[b] | 39.6 | 0.020[b] | 33.3 | 0.105 |

BMI: Body Mass Index.

[a] n depicts the number of patients who achieved stable warfarin dose and any patients with missing genotypes were excluded from the model.

[b] *P*-value did not remain significant after FDR.

[c] *P*-value remained significant after FDR.

### 2.3.5 Haplotype Associations

Two distinct haplotype blocks were identified using the 'Solid spine of LD' method in HaploView as shown in Figure 2.4. Haplotype block 1 consists of 9 SNPs (rs2189784, rs2079288, rs7252046, rs12610189, rs17756654, rs1272, rs3093204, rs3093198, rs2108622). Haplotype block 2 consists of 11 markers (rs3093195, rs12984060, rs2886296, rs3093168, rs3093150, rs3093145, rs3093144, rs3093135, rs2016503, rs984692, rs3093097).

*P*-values from the LRTs are reported in Table 2.6. A *P*-value < 0.05 was obtained for the association between haplotype block 2 and the outcomes of time to therapeutic INR and major haemorrhagic complications, but these associations did not remain significant after FDR.

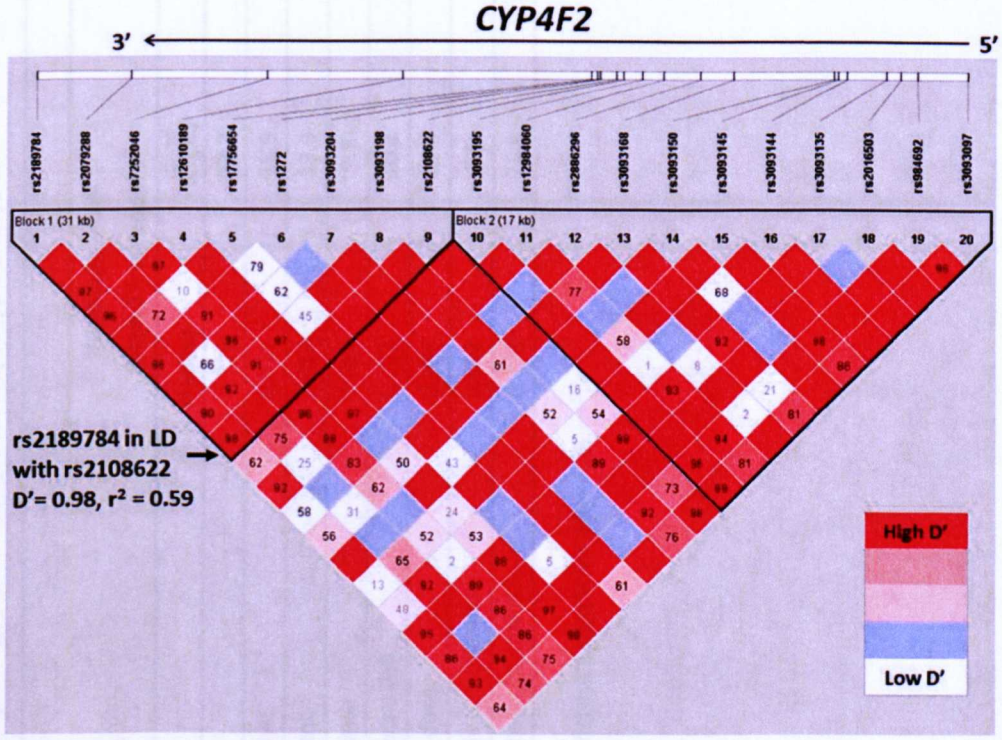



**Figure 2.4. Linkage disequilibrium (D') across chromosomal region 19p13.11.** Tagging SNPs ($r^2 > 0.9$) encompassing the *CYP4F2* gene are represented by 2 distinct haplotype blocks. Haplotype blocks were defined by the 'Solid spine of LD' option in HaploView version 4.1.

**Table 2.6. Multiple regression models for haplotype analysis.**

| Outcome | Assessing association with: | | | | | |
|---|---|---|---|---|---|---|
| | Haplotype block 1 (Model 2 vs model 1) | | Haplotype block 2 (Model 3 vs model 1) | | The gene as a whole (Model 4 vs model 1) | |
| | *P*-value | $R^2$ (%) | *P*-value | $R^2$ (%) | *P*-value | $R^2$ (%) |
| INR > 4 | 0.728 | 23.2 | 0.868 | 22.2 | 0.838 | 27.7 |
| Warfarin sensitive | 0.311 | 11.9 | 0.416 | 10.8 | 0.321 | 13.3 |
| Warfarin resistant | 0.754 | 39.9 | 0.51 | 42.8 | 0.737 | 48.5 |
| Time to stable dose | 0.283 | 9.5 | 0.063 | 12.1 | 0.29 | 13.7 |
| Time to therapeutic INR | 0.222 | 10.2 | 0.027[a] | 13.1 | 0.098 | 15.1 |
| Stable warfarin dose | 0.905 | 22.6 | 0.984 | 16.1 | 0.994 | 22.9 |
| All haemorrhagic complications | 0.612 | 6.5 | 0.154 | 10.2 | 0.232 | 16.2 |
| Major haemorrhagic complications | 0.179 | 25.9 | 0.053 | 29.8 | 0.043[a] | 40.7 |

vs: versus.

[a] *P*-value did not remain significant after FDR.

## 2.4 DISCUSSION

Recent studies have shown that in patients stabilised on warfarin therapy, carriers of the variant homozygous TT alleles for the *CYP4F2* DNA variant, rs2108622, require approximately 1-2.5 mg more warfarin per day than those carrying the wild-type homozygous CC alleles (Caldwell *et al.,* 2008; Borgiani *et al.,* 2009; Takeuchi *et al.,* 2009). However, in our study, rs2108622 did not show association with warfarin stable dose. To reduce heterogeneity in patient characteristics, our analysis of association between rs2108622 and stable dose was repeated in a subgroup of patients which excluded 7 non-Caucasian patients and 53 patients who had congestive heart failure, renal disease, hepatic disease or cancer, in an attempt to mirror the study population of Caldwell *et al.* (2008). However, the results still remained non-significant, although the subgroup included only 235 patients (only 157 of whom achieved stable warfarin dose during course of study) and thus there may not have been adequate power to detect an effect. To investigate how an increased sample size may have influenced our conclusions, confidence intervals for the differences in means between genotype groups were constructed. As the distribution of stable dose was skewed, a bootstrapped approach was undertaken. The 95% confidence interval for the difference in means between the wild-type CC homozygotes and CT heterozygotes was calculated as -0.9965 to 0.1532, whilst the difference in means between CT heterozygotes and mutant TT homozygotes was -0.6591 to 0.7700. Although increasing the sample size would make these confidence intervals narrower (even if the lower limits increased to above zero), it is unlikely that the range of differences spanned by them would be clinically significant. It is also important to note that among the three cohorts tested by Caldwell and colleagues

(2008), they did not find an association between rs2108622 and stable dose in the cohort from University of Washington in St Louis (n = 269, $P = 0.382$).

In addition to looking at rs2108622, this study has undertaken the most comprehensive analysis of the SNPs and haplotypes across the *CYP4F2* gene, and related this to several clinically relevant outcomes. At least 80 SNPs were identified, of which 59 were finally included in the analyses. Haplotype analysis showed that there were 2 blocks, with a high degree of linkage disequilibrium across the gene. While rs2108622 has been shown to be functionally relevant (Stec *et al.,* 2007), the function of most of the other SNPs has not been investigated. Although associations with many of the SNPs and the various outcomes were found in our cohort, none apart from one SNP rs2189784 near the 3' end of the *CYP4F2* gene, remained significant after FDR. rs2189784 showed an association with time to achieve therapeutic INR – patients with two AA variant alleles required a longer period of time to achieve therapeutic INR compared with patients with two GG wild-type alleles. The association with time to therapeutic INR with rs2108622 did not withstand FDR correction despite the fact that rs2108622 and rs2189784 are in strong LD. Whether rs2189784 is functionally important is not known, and given the high LD across the gene, it is difficult to be sure whether either of these SNPs are the causal variants.

The recent demonstration that CYP4F2 metabolises vitamin K (McDonald *et al.,* 2009) provides biological plausibility for the association between rs2108622 and stable warfarin dose (Caldwell *et al.,* 2008; Borgiani *et al.,* 2009; Takeuchi *et al.,* 2009), and indeed with our finding between rs2189784 and time to therapeutic INR. The reason for the different association in our study is likely to be complex and related to the patient populations studied. For instance, in comparison with the

retrospectively recruited patients studied by Caldwell *et al.* (2008) and Borgiani *et al.* (2009), our patient population was recruited prospectively and followed up for 6 months. The determinants of body stores of vitamin K are complex (Custodio das Dores *et al.,* 2007; Booth and Al Rajabi 2008; Shea *et al.,* 2008; Shea *et al.,* 2009), and may well be different in patients studied soon after starting on warfarin when they have had an acute clinical event, compared with patients who have been stable on warfarin for a period of time when they will have been given clinical advice to have a stable diet (Custodio das Dores *et al.,* 2007). Therefore, in a population studied soon after the start of warfarin, time to therapeutic INR may be a more sensitive marker of warfarin response than stable dose in relation to parameters which can modulate vitamin K levels.

# CHAPTER 3

# Genetic variability in the *CYP4F2*, *CYP4F11* and *CYP4F12* genes affects liver mRNA levels and plays a role in warfarin response

# CONTENTS

**3.1 INTRODUCTION** ..... **72**

**3.2 PATIENTS, MATERIALS AND METHODS** ..... **75**

3.2.1 Patient DNA and liver samples ..... 75

3.2.2 SNP selection ..... 75

3.2.3 Genotyping ..... 76

3.2.4 Haplotype determination ..... 76

3.2.5 Determination of *CYP4F2, CYP4F3, CYP4F8, CYP4F11* and *CYP4F12* mRNA expression levels in human liver ..... 77

3.2.6 Liver microsome preparation ..... 78

3.2.7 Bradford assay ..... 79

3.2.8 Analysis of CYP4F2 protein expression ..... 79

3.2.9 Genome-wide genotyping and SNPs imputation ..... 80

3.2.10 Statistical analysis ..... 81

*3.2.10.1 Liver mRNA and protein analysis* ..... *81*

*3.2.10.2 Warfarin outcome analysis* ..... *82*

3.2.11 Identification of SNPs associated with mRNA expression *in silico* ..... 82

3.2.12 Network building method ..... 82

**3.3 RESULTS** ..... **83**

3.3.1 Effects of *CYP4F2* variants on hepatic mRNA expression of the *CYP4F* gene cluster ..... 83

3.3.2 Effects of *CYP4F11* and *CYP4F12* SNPs on hepatic mRNA expression of the *CYP4F* gene cluster ..... 86

3.3.3 Effect of haplotypes in the *CYP4F12-CYP4F2-CYP4F11* region on hepatic mRNA expression of the *CYP4F* gene cluster ..... 90

3.3.4 Association of microsomal CYP4F2 protein with SNPs and haplotypes in the *CYP4F2* and *CYP4F11* region ..... 94

3.3.5 Effect of rs1060467 in *CYP4F11* on warfarin stable dose and time to therapeutic INR. 100

3.3.6 *In silico* genotype-phenotype analysis ..... 100

3.3.7 Effect of rs7248867 and rs2074568 on warfarin stable dose and time to therapeutic INR ..... 101

3.3.8 Interaction between *CYP4F2, CYP4F11* and *CYP4F12* ..... 105

**3.4 DISCUSSION** ..... **107**

## 3.1 INTRODUCTION

The *CYP4F* gene subfamily comprises six members, namely *CYP4F2* (Kikuta *et al.,* 1993), *CYP4F3* (*CYP4F3A* and *CYP4F3B*) (Kikuta *et al.,* 1998), *CYP4F8* (Bylund *et al.,* 2000), *CYP4F11* (Cui *et al.,* 2000), *CYP4F12* (Bylund *et al.,* 2001; Hashizume *et al.,* 2001) and *CYP4F22* (Lefevre *et al.,* 2006). Structurally, these six *CYP4F* genes are largely similar, with more than 65% amino acid sequence homology.

To date, studies have focused on *CYP4F2, CYP4F3, CYP4F8, CYP4F11* and *CYP4F12* and little is known about the expression and function of *CYP4F22.* The splice sites of *CYP4F2, CYP4F3, CYP4F8, CYP4F11* and *CYP4F12* are almost identical, suggesting that this cluster of five genes may have evolved by gene duplication (Bylund *et al.,* 1999; Kikuta *et al.,* 1999; Cui *et al.,* 2000; Bylund *et al.,* 2001). *CYP4F2, CYP4F3, CYP4F8, CYP4F11* and *CYP4F12* genes reside together on chromosome 19p13.1-2, spanning over 320 kb (Figure 3.1). They are all expressed in the liver and are known to metabolise xenobiotics, arachidonic acid and its oxygenated derivatives (eicosanoids) such as leukotrienes, prostaglandins (PGs), lipoxins, and hydroxyeicosatetraenoic acids (HETEs) (Jin *et al.,* 1998; Powell *et al.,* 1998; Kikuta *et al.,* 1999; Bylund *et al.,* 2000; Lasker *et al.,* 2000; Bylund *et al.,* 2001; Christmas *et al.,* 2001; Hashizume *et al.,* 2001; Hashizume *et al.,* 2002; Kalsotra *et al.,* 2004; Stark *et al.,* 2005a; Stark *et al.,* 2005b). *CYP4F2* has also been implicated in the ω-hydroxylation of tocopherol phytyl side chain in the first step of vitamin E inactivation (Sontag and Parker 2002). As discussed in chapter 2, a polymorphism in *CYP4F2*, namely rs2108622 (V433M, +1297 C>T), has been reported to be associated with the dose requirement of warfarin (Caldwell *et al.,*

2008; Borgiani *et al.,* 2009; Perez-Andreu *et al.,* 2009; Takeuchi *et al.,* 2009; Cha *et al.,* 2010; Pautas *et al.,* 2010)

In chapter 2, fine mapping of the *CYP4F2* region was performed to determine the influence of *CYP4F2* SNPs and haplotypes on various warfarin response outcomes (Zhang *et al.,* 2009). An association between rs2189784, a SNP in LD with rs2108622, with time to therapeutic INR, but not with stable dose was found. Given these differences in association found between *CYP4F2* SNPs and different clinical outcomes that have been utilised to assess the responses with warfarin (Zhang *et al.,* 2009), and the high degree of homology and LD across the *CYP4F* gene cluster (as shown in Figure 3.1), the aims of this chapter were (i) to undertake a genotype-phenotype assessment of the *CYP4F* genes utilising a well characterised liver bank and a prospective patient cohort who were followed up for 6 months from the time of intake of warfarin, and (ii) to perform *in silico* analysis using publicly available data repositories to investigate additional SNP-gene associations and the interaction between the *CYP4F* genes.

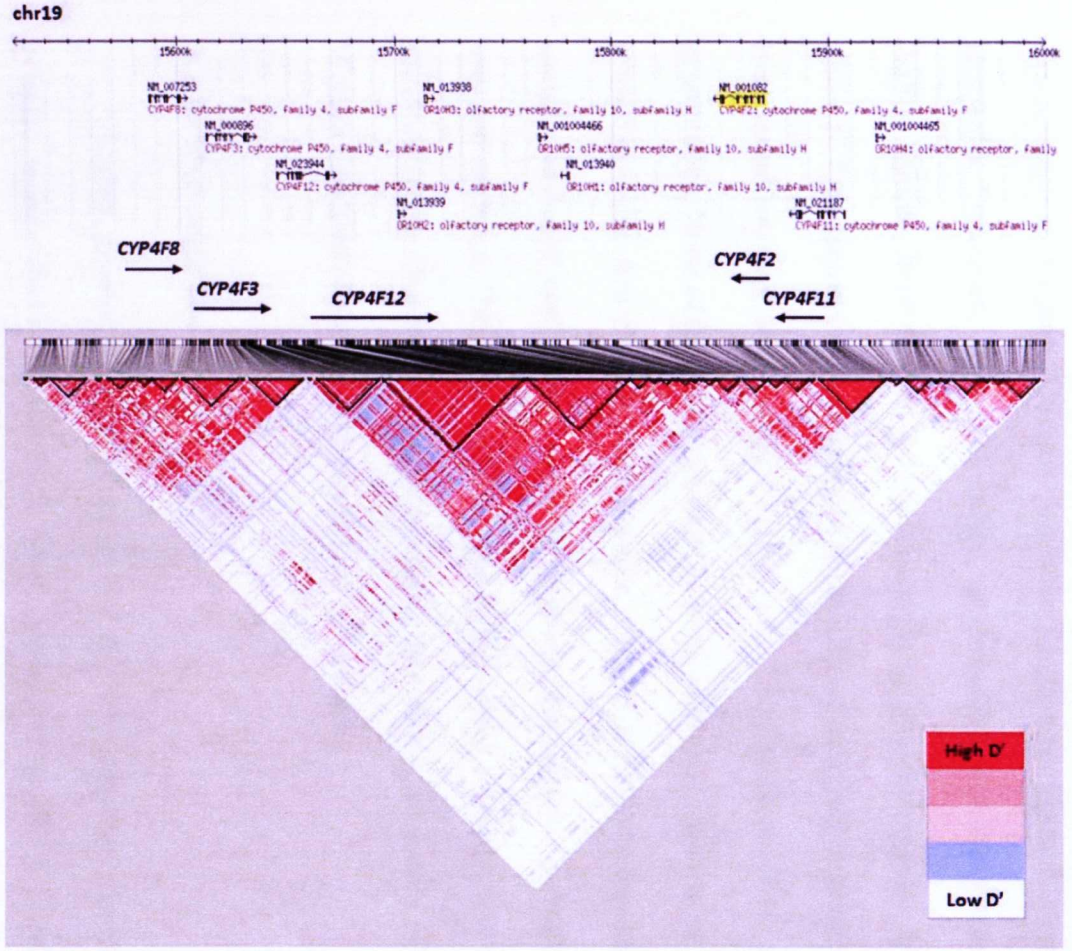



**Figure 3.1. Genomic structure of the *CYP4F* gene cluster encompassing *CYP4F2, CYP4F3, CYP4F8, CYP4F11* and *CYP4F12*.** LD pattern and haplotype blocks were generated using HaploView version 4.2 based on the CEU population genotype data on chromosome 19p13 region 15529000 – 16040000, obtained from HapMap data release 27, NCBI build 36 assembly.

## 3.2 PATIENTS, MATERIALS AND METHODS

### 3.2.1 Patient DNA and liver samples

Blood and liver tissue samples were previously collected from 149 Caucasian patients undergoing liver surgery at the Department of General, Visceral, and Transplantation Surgery, Campus Virchow, University Medical Centre Charité, Humboldt University, Berlin, Germany, as described previously (Gomes *et al.*, 2009). Normal liver tissues were obtained from adjacent regions of surgically removed liver tumours or metastases or hepatic tissue resected for other reasons. All liver tissue samples were certified to be free of malignant cells by pathological examination. None of these samples were from patients with hepatitis, or cirrhosis, or from those who had chronic alcohol abuse. Clinical patient documentation for all samples included age, gender, medical diagnosis, presurgical medication, alcohol use, and smoking. The study was approved by the ethics committees of the Medical Faculties of the Charité, Humboldt University, and of the University of Tuebingen. Written informed consent was obtained from all patients.

### 3.2.2 SNP selection

Eighty genetic polymorphisms in the *CYP4F2* gene have been previously selected as described in section 2.2.4. SNPs encompassing *CYP4F11* and *CYP4F12* across the chromosomal 19p13.11 region were selected using the same criteria described in section 2.2.4 except that data from HapMap data release 27, NCBI build 36 assembly was used. A total of 130 SNPs in the *CYP4F11* and *CYP4F12* region were successfully designed and subdivided into 6 multiplex assays using Sequenom's online Human GenoTyping Tools (www.mysequenom.com/Tools) with

the same principles and procedures described in section 2.2.5. Primer sequences are listed in Appendix 2.1.

### 3.2.3 Genotyping

Genomic DNA of the liver surgery patients (n = 149) was extracted from whole blood using the QIAamp DNA Mini Kit (QIAGEN GmbH, Hilden, Germany) according to the manufacturer's instructions.

All 210 SNPs across *CYP4F2*, *CYP4F11* and *CYP4F12* were genotyped using the Sequenom MassARRAY iPLEX™ platform (Sequenom) in accordance with the manufacturer's instructions as described in section 2.2.6. As summarised in Appendix 2.2, markers which deviated from HWE (n = 10), those with less than 90% call rate (n = 27), and those which were monomorphic (n = 26), were excluded from downstream association analysis. Among the remaining 147 SNPs, 88 had been typed by the International HapMap project (data release 27, February 2009) and gave similar MAF.

### 3.2.4 Haplotype determination

The pattern of pairwise linkage disequilibrium (LD) between the remaining SNPs was visualised using the program HaploView version 4.2 (Barrett *et al.*, 2005). Haplotype blocks were defined using the default algorithm by Gabriel *et al.* (2002) in HaploView, where 95% confidence intervals for D' between pairs of SNPs were calculated. Only groups of SNPs uninterrupted by recombination (i.e. $\geq$ 95% of informative comparisons are in strong LD) were considered haplotype blocks. The MAF cut-off was decreased from 0.05 to 0.01 to allow less common SNPs to be included. The default algorithm sorts the list of all possible blocks and starts with the

largest and keeps adding blocks as long as they do not overlap with an already declared block.

The most probable combinations of haplotype-pairs at each block were inferred using the program PHASE version 2.1.1 (Stephens *et al.*, 2001; Stephens and Scheet 2005), which utilizes a Bayesian statistical method for reconstructing haplotypes from population genotype data. Default iteration settings were used with the command line: PHASE -n -x10 filenamephase.txt filenamephase.out 10000 1 10000. Any individuals with a haplotype-pair probability of < 90% (n = 11) for at least one haplotype block were excluded from tests of association. Within a haplotype block, haplotypes with frequencies < 1% were grouped together as a single covariate for analysis. The presence or absence of haplotypes was coded as absent (= 0), present once (= 1) or twice (= 2).

### 3.2.5 Determination of *CYP4F2, CYP4F3, CYP4F8, CYP4F11* and *CYP4F12 m*RNA expression levels in human liver

Methods described in sections 3.2.5 to 3.2.8 were carried out by colleagues in Professor Ulrich Zanger's lab, at Dr Margarete Fischer-Bosch Institute of Clinical Pharmacology, Stuttgart, Germany.

RNA was extracted from the human liver tissue (n = 149) using TRIzol® reagent (Invitrogen, Paisley, UK) with subsequent RNA clean-up using QIAGEN RNeasy-Mini Kit with on-column DNase treatment. All RNA preparations were of high quality with RNA integrity number (RIN) > 7, as measured on the Agilent Bioanalyzer (Nano-Lab Chip Kit, Agilent Technologies, Waldbronn, Germany). 200 ng of total RNA was amplified and labelled using the Illumina TotalPrep RNA Amplification kit (Ambion Applied Biosystems, Darmstadt, Germany). cRNA

quality was assessed by capillary electrophoresis on Agilent 2100 Bioanalyzer (Agilent Technologies). Expression levels of over 48, 000 mRNA transcripts were assessed by the HumanWG-6 version 2.0 Expression BeadChips (Illumina, Eindhoven, Netherlands) as described in Schröder *et al.,* 2011 (Schroder *et al.,* 2011). Hybridization was carried out according to the manufacturer's instructions.

Illumina BeadStudio version 3.0 (Illumina, San Diego, CA) was used for all low-level pre-processing steps of the Illumina HumanWG-6 version 2.0 Expression BeadChips, including background estimation and correction, probe set summary, and normalization. After these low-level pre-processing steps, genes with detection *P*-value > 0.1 or more than 10% missing values were filtered out and removed from the dataset. Missing signal intensities were estimated using the "k nearest neighbour" algorithm (KNN) implemented in R BioConductor (Troyanskaya *et al.,* 2001; Gentleman *et al.,* 2004). Finally, after all pre-processing steps, the raw data of 48, 701 probe signal intensities were mapped and reduced to signal intensities corresponding to 15, 439 unique genes.

Gene expression levels of *CYP4F2*, *CYP4F3*, *CYP4F8*, *CYP4F11* and *CYP4F12* were selected to test our hypotheses in this chapter.

### 3.2.6 Liver microsome preparation

Human liver microsomes were prepared by differential ultracentrifugation as described previously (Lang *et al.,* 2001). Briefly, approximately 1 g of tissue was homogenized in 1 mm EDTA, 1 mm DTT, 10 mm HEPES pH 7.4, 0.2 mm Pefabloc (Roth, Karlsruhe, Germany) and 0.15 mm KCl and differentially centrifuged at 15, 000 g and 105, 000 g. The microsomal pellet was washed once with 0.1 M

sodium pyrophosphate buffer (pH 7.5) and the final pellet was resuspended in 0.1 M sodium phosphate buffer (pH 7.4) and immediately frozen at –80 °C in aliquots.

### 3.2.7 Bradford assay

Total microsomal protein content was determined using the Bradford based Bio-Rad Protein Assay and bovine serum albumin as a reference (Bio-Rad Laboratories, Munich, Germany). This assay is based on the formation of a complex between proteins and the dye, Brilliant Blue G-250, in an acidic solution (Bradford 1976). The hydrophobic and ionic interactions between the proteins and the dye stabilize the anionic form of the dye, causing a visible colour change from red to blue, and a consequent shift in absorbance maximum of the dye from 465 to 595 nm. The amount of absorption is proportional to the protein present.

### 3.2.8 Analysis of CYP4F2 protein expression

Of the 149 human liver samples, 27 had matching microsomal protein samples and their expression levels of CYP4F2 protein were determined by western blotting. In brief, 50 µg of microsomal protein was separated on a 10% Tris-glycine SDS-polyacrylamide gel and transferred to nitrocellulose membrane, blocked for 1 h with 5% skimmed milk in TBS-T. The blot was incubated with goat IgG anti-human CYP4F2 primary antibody (clone N-19; Santa Cruz Biotechnology, Inc., 1:250 dilution) overnight at 4°C. Bound CYP4F2 antibody was detected by secondary IR Dye 800 anti-goat IgG antibody (1:10, 000 dilution) using the infrared imaging system Odyssey (LI-COR Biosciences, NE, USA). For quantification a serial dilution of supersomes with recombinantly expressed human CYP4F2 (BD Biosciences, Heidelberg, Germany) was included in each run. One of the 27 microsomal protein

samples analysed had undetectable CYP4F2 protein level and was excluded from the association analysis.

### 3.2.9 Genome-wide genotyping and SNPs imputation

For the 1000 patients initiated onto warfarin therapy as described in section 2.2.1, DNA was extracted from patients' EDTA blood samples using the standard phenol-chloroform method (as briefly described in section 2.2.6). Of the 1000 DNA samples, genome-wide genotyping was carried out on 752 DNA samples with the use of Illumina Human 610k chip, at the Wellcome Trust Sanger Institute. All quality controls and imputations were performed by Stephane Bourgeois at the Sanger Institute.

Quality control measures were undertaken using the open-source genome-wide analysis toolset PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/) (Purcell *et al.,* 2007). All SNPs with a genotyping success rate < 90%, HWE threshold of $P < 0.001$ and those with MAF < 2% were excluded from the dataset. Subjects with genotyping success rate < 95% were also removed. Principal component analysis was performed to assess genetic markers for ethnicity. Only individuals with genetically matching ethnicity were included into the association analysis (n = 714).

Imputation of additional SNPs throughout the whole genome in the 714 subjects on warfarin therapy was based on the reference genotype data from the 1000 Genomes project and was performed using the freely available program IMPUTE version 2 (http://www.stats.ox.ac.uk/~marchini/software/gwas/gwas.html) (Howie *et al.,* 2009). Using a command line utility program, QCTOOL, imputed variants with information value (measure of amount of statistical information the genotypes

provide on allele frequency) below 0.5 and variants with a minor allele frequency < 1% were excluded.

For the purpose of our analysis, genotype data of specific SNPs (rs1060467, rs7248867 and rs2074568) from the genome-wide scan and imputation were extracted using PLINK.

### 3.2.10 Statistical analysis

Statistical analyses were conducted with the statistics software packages SPSS, version 18. A *P*-value < 0.05 was regarded as statistically significant. HWE for each SNP was determined using a computationally efficient exact $P_{HWE}$ test statistic implemented in HaploView (Wigginton *et al.*, 2005). A *P*-value <0.001 was assumed to indicate deviation from HWE. The proportion of variability explained by the genetic covariates was calculated using Nagelkerke's $R^2$ statistic.

#### *3.2.10.1 Liver mRNA and protein analysis*

Relationships between each of the phenotypic parameters evaluated were examined by Spearman correlation analysis. All except *CYP4F2* mRNA expression was skewed. To ensure normal distribution, the expression data were either square-root transformed (*CYP4F3* mRNA, *CYP4F12* mRNA and *CYP4F2* protein), natural log transformed (*CYP4F8* mRNA), or log transformed (*CYP4F11* mRNA). To evaluate the association of each SNP or haplotype with mRNA and protein expression levels, two univariate tests of association were performed: (i) one-way analysis of variance (ANOVA) which makes no assumption on the mode of inheritance and (ii) univariate linear regression which assumes an additive mode of inheritance; the minimum *P*-value was referred to in each analysis. *P*-values from all

genotype-phenotype association tests undertaken in the functional study were adjusted for multiple testing using the FDR (Benjamini *et al.,* 2001). FDR-corrected *P*-values are denoted as $P_c$-values.

#### *3.2.10.2 Warfarin outcome analysis*

Univariate linear regression and log-rank test for trend (both assuming additive mode of inheritance) were employed to test for the association of SNPs with warfarin stable dose and time to therapeutic INR, respectively.

### 3.2.11 Identification of SNPs associated with mRNA expression *in silico*

Putative expression quantitative trait loci (eQTLs) in the *CYP4F* gene cluster were identified using the eQTL browser (http://eqtl.uchicago.edu/cgi-bin/gbrowser/eqtl/), a database that summarises results from large-scale studies which identified eQTLs in the liver (Schadt *et al.,* 2008), brain (Myers *et al.,* 2007), fibroblasts (Dimas *et al.,* 2009), T-cells (Dimas *et al.,* 2009), monocytes (Zeller *et al.,* 2010), and lymphoblastoid cell lines (Stranger *et al.,* 2007; Veyrieras *et al.,* 2008; Dimas *et al.,* 2009; Montgomery *et al.,* 2010; Pickrell *et al.,* 2010).

### 3.2.12 Network building method

The interactions between CYP4F2, CYP4F11 and CYP4F12 were visualized in MetaCore™ (GeneGo Inc., St. Joseph, MI, USA), an interactive database derived from manually curated literature publications on proteins and small molecules of biological relevance in humans. MetaCore generates an interaction network around the proteins and finds the clusters of objects directly connected.

## 3.3 RESULTS

### 3.3.1 Effects of *CYP4F2* variants on hepatic mRNA expression of the *CYP4F* gene cluster

The significant effects of *CYP4F2* variants on hepatic mRNA expression of the *CYP4F* gene cluster are summarised in Table 3.1. Contrary to a recent report by McDonald *et al.*, (2009), a significant association between rs2108622 and liver *CYP4F2* mRNA expression was found (Figure 3.2a), with subjects homozygous for the rs2108622 minor T allele showing greater *CYP4F2* expression compared to subjects homozygous for the major C allele (TT = 1.47 ± 0.29, CC = 0.97 ± 0.31, $P_c$ = 5.75 x $10^{-4}$, $R^2$ = 12.6%). Moreover, several other *CYP4F2* variants were also associated with significant up-regulation of *CYP4F2* expression including rs2189784 ($P_c$ = 0.014, $R^2$ = 7.9%, Figure 3.2b), a SNP located 29 kb downstream of the *CYP4F2* gene. Interestingly, in addition to being associated with *CYP4F2* mRNA expression, rs2108622 demonstrated significant association with *CYP4F11* mRNA down-regulation ($P_c$ = 2.56 x $10^{-4}$, $R^2$ = 13.7%, Figure 3.2c) while rs2189784 was significantly associated with lower levels of *CYP4F12* mRNA expression ($P_c$ = 0.014, $R^2$ = 8.3%, Figure 3.2d). No associations were found between *CYP4F2* variants and *CYP4F3* or *CYP4F8* mRNA expression (data not shown).

**Table 3.1. Effects of *CYP4F2* SNPs on hepatic mRNA expression levels of *CYP4F2*, *CYP4F11* and *CYP4F12*.**

| Gene | SNP | Localization | mRNA Expression | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | *CYP4F2* | | | *CYP4F11* | | | *CYP4F12* | | |
| | | | $P_c$-value | ↑ or ↓ | $R^2$ (%) | $P_c$-value | ↑ or ↓ | $R^2$ (%) | $P_c$-value | ↑ or ↓ | $R^2$ (%) |
| - | rs2189784 | Downstream of *CYP4F2* | **0.014** | ↑ | **7.9** | 0.179 | - | - | **0.014** | ↓ | **8.3** |
| *CYP4F2* | rs3093209 * | 3` near gene | **$1.39 \times 10^{-4}$** | ↑ | **14.6** | **0.055** | ↓ | **5.3** | 0.157 | - | - |
| *CYP4F2* | rs2108622 | Exon 11, Missense, Met433Val | **$5.75 \times 10^{-4}$** | ↑ | **12.6** | **$2.56 \times 10^{-4}$** | ↓ | **13.7** | 0.296 | - | - |
| *CYP4F2* | rs3093173 * | Intron 9 | 0.283 | - | - | **0.016** | ↓ | **7.2** | 0.279 | - | - |
| *CYP4F2* | rs3093169 * | Intron 9 | 0.381 | - | - | **0.016** | ↓ | **5.8** | 0.156 | - | - |

Effects of statistically significant SNPs with *P*-values < 0.05 after FDR are shown in bold.

$R^2$ refers to proportion of variability.

* Tagging SNP ($r^2 \geq 0.9$). Tagged SNPs are detailed in Appendix 2.3.

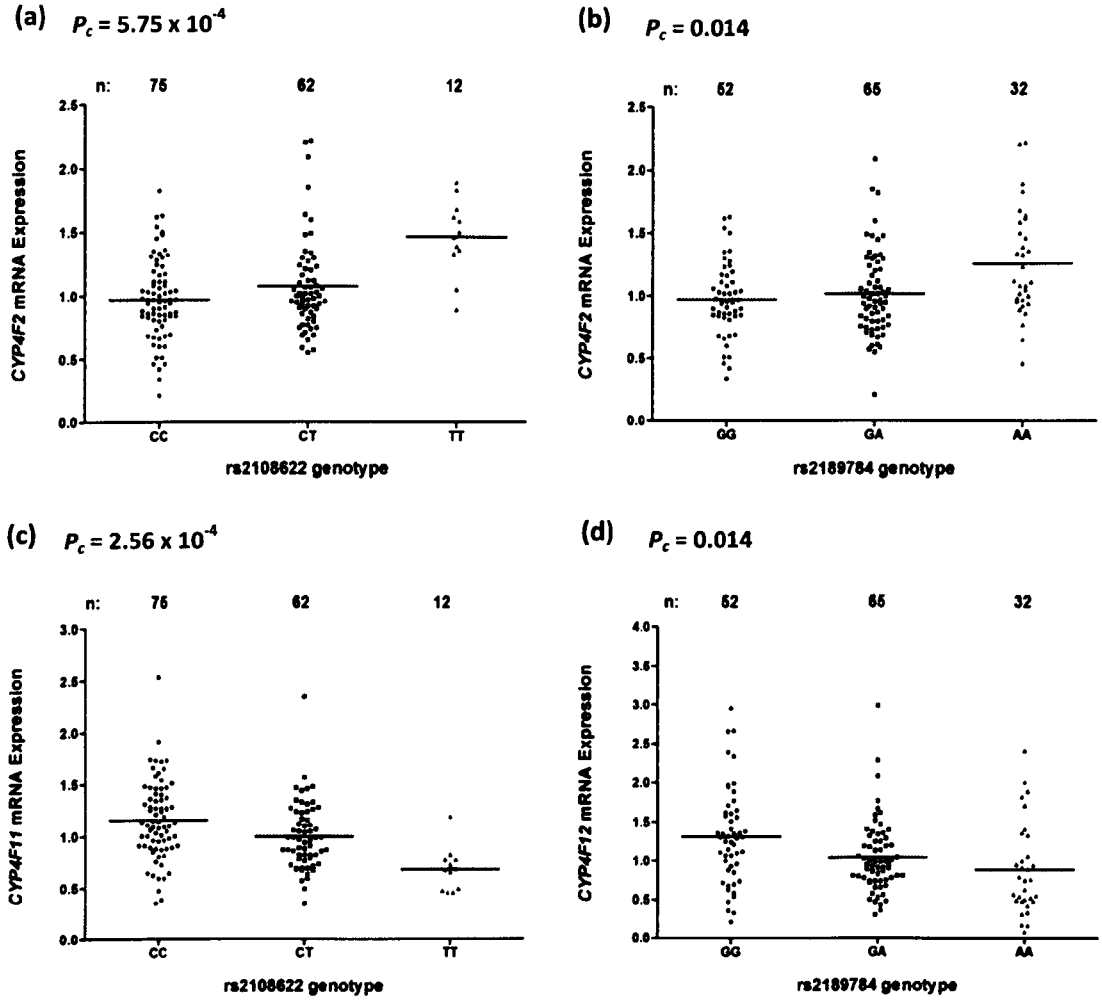



**Figure 3.2. Levels of *CYP4F2*, *CYP4F11* and *CYP4F12* mRNA in normal liver tissue donated from 149 patients in relation to *CYP4F2* SNPs.** *CYP4F2* mRNA expression stratified by (a) rs2108622 and (b) rs2189784 genotypes. *CYP4F11* and *CYP4F12* mRNA expression stratified by (c) rs2108622 and (d) rs2189784 genotypes, respectively. *P*-values after FDR are shown in the upper left corner. Each dot represents an individual and the solid lines represent the mean values.

### 3.3.2 Effects of *CYP4F11* and *CYP4F12* SNPs on hepatic mRNA expression of the *CYP4F* gene cluster

It is noteworthy that *CYP4F11* and *CYP4F12* mRNA were both correlated with *CYP4F2* mRNA expression ($r_s = 0.250$, $P = 0.002$; $r_s = 0.384$, $P = 1.35 \times 10^{-6}$, respectively, Figures 3.3a & b). These correlations are reflected in the associations of *CYP4F2* variants with *CYP4F11* and *CYP4F12* mRNA expression.

Looking at the region encompassing the *CYP4F* gene cluster on HapMap database (Figure 3.1), high LD is seen in the *CYP4F12-CYP4F2-CYP4F11* locus, suggesting that SNPs across the *CYP4F11* and *CYP4F12* regions could influence the mRNA expression of *CYP4F2* and possibly other *CYP4F* gene cluster members. To examine the genetic contribution of variants in *CYP4F11* and *CYP4F12* on the hepatic mRNA expression of the *CYP4F* gene cluster, fine mapping of the *CYP4F11* and *CYP4F12* gene regions was conducted and significant associations are summarised in Table 3.2.

rs1060467, a genetic variant located in the 3' untranslated region (UTR) of *CYP4F11* demonstrated a significant association with decreasing *CYP4F2* mRNA expression ($P_c = 0.016$, $R^2 = 7.2\%$, Figure 3.4a); whilst an opposite trend for increasing *CYP4F11* mRNA expression was observed which was not statistically significant ($P_c = 0.179$, Figure 3.4b).

Eight SNPs spanning *CYP4F12* were significantly associated with *CYP4F12* mRNA expression. No significant association with *CYP4F3* or *CYP4F8* mRNA expression was observed with any SNPs in the *CYP4F11* or *CYP4F12* region.

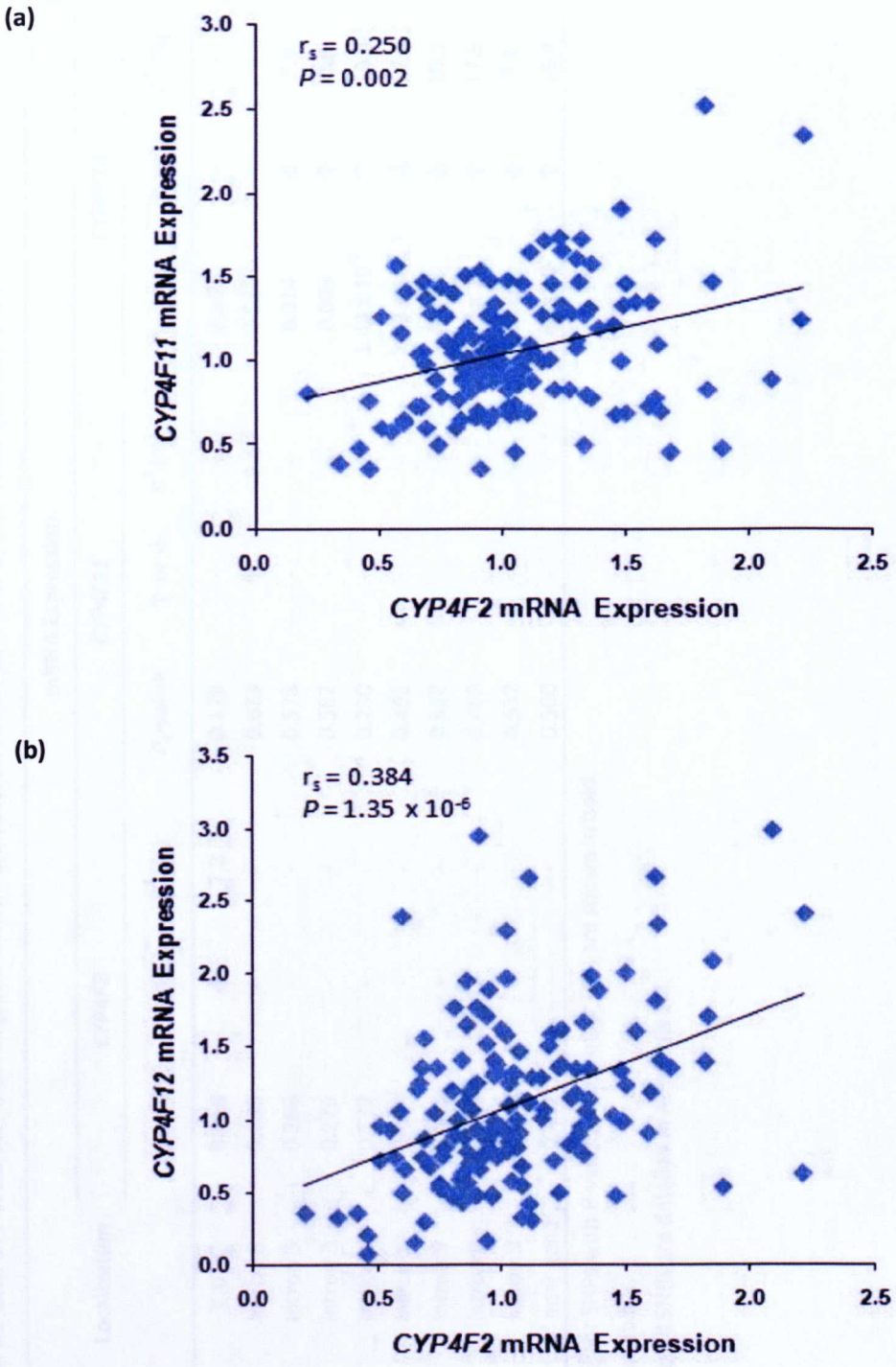



**Figure 3.3. Phenotypic correlations in Caucasian human liver tissues. (a)** Correlation between *CYP4F11* mRNA and *CYP4F2* mRNA levels (n = 149); **(b)** correlation between *CYP4F12* mRNA and *CYP4F2* mRNA levels (n = 149). The Spearman's rho correlation coefficient ($r_s$) and *P*-value for each comparison are given.

**Table 3.2. Effects of *CYP4F11* and *CYP4F12* SNPs on hepatic mRNA expression levels of *CYP4F2*, *CYP4F11* and *CYP4F12*.**

| Gene | SNP | Localization | mRNA Expression | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | *CYP4F2* | | | *CYP4F11* | | | *CYP4F12* | | |
| | | | $P_c$-value | ↑ or ↓ | $R^2$ (%) | $P_c$-value | ↑ or ↓ | $R^2$ (%) | $P_c$-value | ↑ or ↓ | $R^2$ (%) |
| *CYP4F11* | rs1060467 | 3' UTR | **0.016** | ↓ | **7.2** | 0.179 | - | - | 0.369 | - | - |
| *CYP4F11* | rs12977516 * | Intron 8 | 0.085 | - | - | **0.023** | ↑ | **6.4** | 0.510 | - | - |
| *CYP4F12* | rs17682485 | Intron 3 | 0.266 | - | - | 0.578 | - | - | **0.014** | ↓ | **7.9** |
| *CYP4F12* | rs12460703 | Intron 3 | 0.279 | - | - | 0.582 | - | - | **0.009** | ↑ | **8.9** |
| *CYP4F12* | rs2074568 | Intron 4 | 0.537 | - | - | 0.250 | - | - | **$1.49 \times 10^{-5}$** | ↑ | **19.5** |
| *CYP4F12* | rs10409750 | Intron 5 | 0.404 | - | - | 0.493 | - | - | **$1.39 \times 10^{-4}$** | ↓ | **14.8** |
| *CYP4F12* | rs10410357 | Intron 9 | 0.443 | - | - | 0.532 | - | - | **0.003** | ↓ | **10.3** |
| *CYP4F12* | rs11879787 | Intron 9 | 0.545 | - | - | 0.469 | - | - | **$1.65 \times 10^{-5}$** | ↑ | **17.9** |
| *CYP4F12* | rs627971 * | Intron 9 | 0.290 | - | - | 0.532 | - | - | **0.016** | ↓ | **7.1** |
| *CYP4F12* | rs2886476 * | 3' near gene | 0.497 | - | - | 0.500 | - | - | **$1.49 \times 10^{-5}$** | ↑ | **18.4** |

Effects of statistically significant SNPs with *P*-values < 0.05 after FDR are shown in bold.

$R^2$ refers to proportion of variability.

* Tagging SNP ($r^2 \geq 0.9$). Tagged SNPs are detailed in Appendix 2.3.

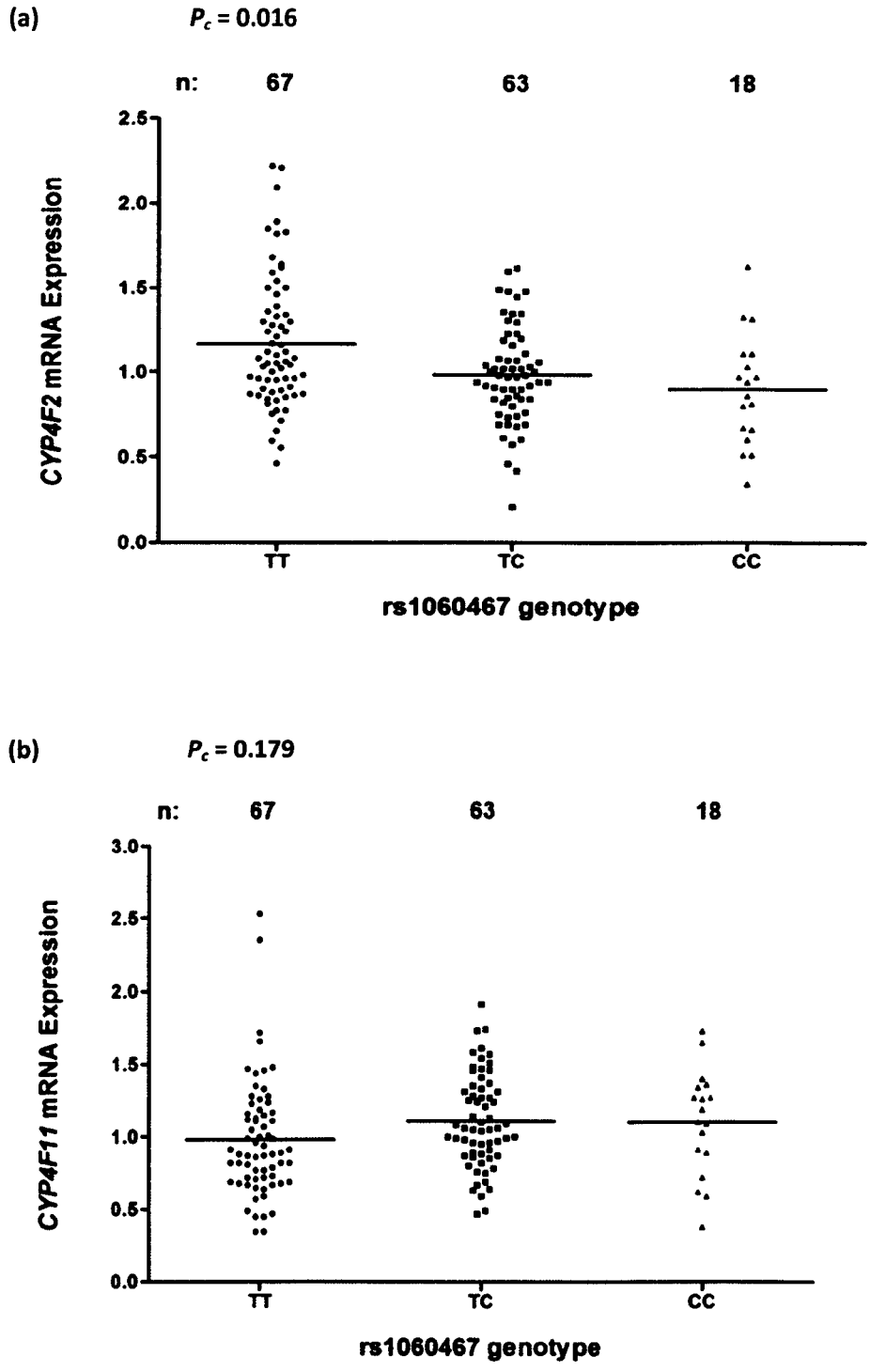
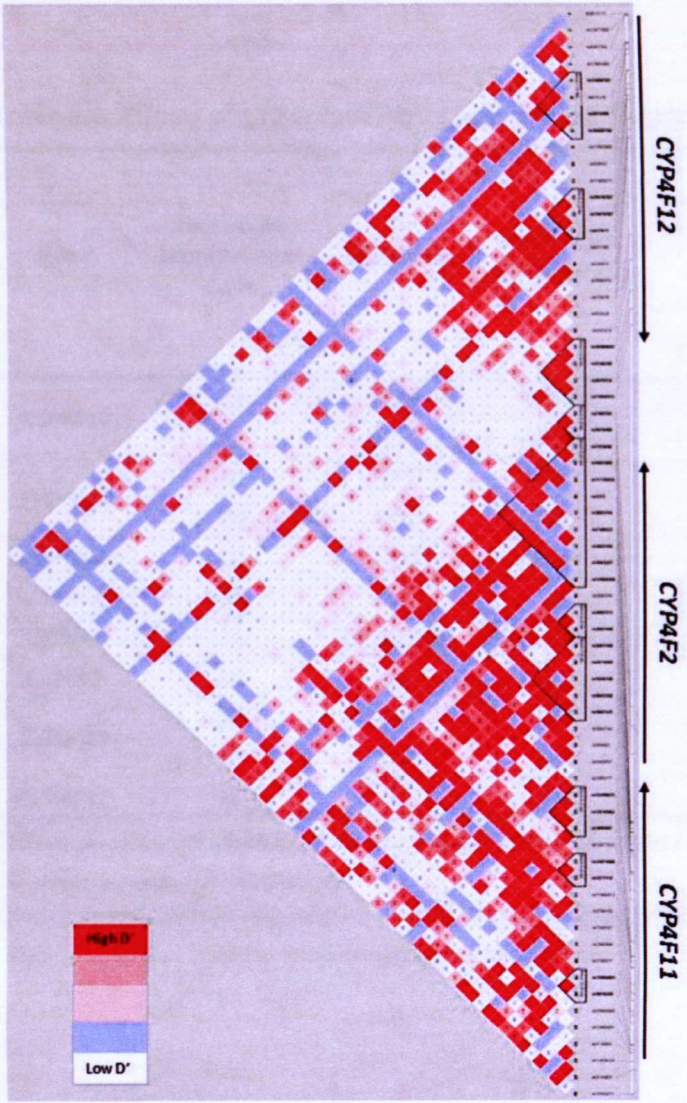



**Figure 3.4. Levels of *CYP4F2* and *CYP4F11* mRNA in normal liver tissue donated from 149 patients in relation to *CYP4F11* SNP rs1060467 genotype.** (a) *CYP4F2* mRNA expression; (b) *CYP4F11* mRNA expression. *P*-values after FDR are shown in the upper left corner. Each dot represents an individual and the solid lines represent the mean values.

### 3.3.3 Effect of haplotypes in the *CYP4F12-CYP4F2-CYP4F11* region on hepatic mRNA expression of the *CYP4F* gene cluster

To explore the complex genetic architecture of *CYP4F* locus containing *CYP4F2*, *CYP4F11* and *CYP4F12*, haplotypes across these three genes were constructed based on the genotype data. 10 haplotype blocks were identified as shown in Figure 3.5, with details of haplotypes inferred and their estimated frequencies. Effects of *CYP4F2*, *CYP4F11* and *CYP4F12* haplotypes on hepatic mRNA expression of the *CYP4F* gene cluster were evaluated and significant associations are reported in Table 3.3.

Haplotype 4A harbouring the sequence 'AT' with a frequency of 43.3% was associated with significant increase in hepatic *CYP4F2* ($P_c$ = 0.014, $R^2$ = 7.9%, Figure 3.6a) and reduced *CYP4F12* ($P_c$ = 0.014, $R^2$ = 8.3%, Figure 3.6b) mRNA expression, mirroring the effect of rs2189784. Corresponding to the effect of rs2108622, haplotype 5A 'TGCGGTGGG' (frequency = 28.3%) was significantly associated with increased *CYP4F2* ($P_c$ = 0.001, $R^2$ = 12.6%, Figure 3.6c) and decreased *CYP4F11* ($P_c = 2.74 \times 10^{-4}$, $R^2$ = 13.8%, Figure 3.6d) mRNA expression. Resembling the effect of rs1060467, haplotype 8B (sequence 'TGC', frequency = 33.2%) was associated with down-regulation of *CYP4F2* ($P_c$ = 0.016, $R^2$ = 7.0%, Figure 3.6e) and showed a non-significant up-regulating effect on *CYP4F11* ($P_c$ = 0.203, Figure 3.6f) mRNA expression.




| Gene | SNPs | Haplotype Block | Haplotype Identification Code | Haplotype Sequence | Haplotype Frequency (N = 138) | |
|---|---|---|---|---|---|---|
| | | | | | *n* | (%) |
| *CYP4F12* | rs12460703 | 1 | 1A | AGTG | 129 | 43.3 |
| | rs675326 | | 1B | GGTA | 80 | 27.2 |
| | rs2074568 | | 1C | GGAA | 65 | 21.6 |
| | rs10409750 | | 1D | AGTA | 23 | 7.7 |
| | | | 1E | GATA | 1 | 0.2 |
| *CYP4F12* | rs10410357 | 2 | 2A | AAC | 99 | 33.2 |
| | rs11879787 | | 2B | AAT | 80 | 26.8 |
| | rs627971 | | 2C | AGT | 72 | 24.1 |
| | | | 2D | GAC | 47 | 15.7 |
| Intergenic | rs16980800 | 3 | 3A | GGGT | 85 | 28.5 |
| | rs2240228 | | 3B | AGTC | 80 | 26.8 |
| | rs2079234 | | 3C | AAGT | 74 | 24.8 |
| | rs11666521 | | 3D | AGTT | 59 | 19.8 |
| *CYP4F2* | rs2189784 | 4 | 4A | AT | 129 | 43.3 |
| | rs2079288 | | 4B | GT | 108 | 36.3 |
| | | | 4C | GC | 61 | 20.5 |
| *CYP4F2* | rs7252046 | 5 | 5A | TGCGGTGGG | 84 | 28.3 |
| | rs3093209 | | 5B | CCCCGCGGA | 78 | 26.2 |
| | rs17756654 | | 5C | TCCGGCAGG | 78 | 26.1 |
| | rs1272 | | 5D | TGCGGCGGG | 22 | 7.4 |
| | rs3093204 | | 5E | CCCGACGGA | 17 | 5.7 |
| | rs2108622 | | 5F | TCAGGCGGA | 13 | 4.3 |
| | rs3093195 | | 5G | CCCCGCGTA | 3 | 1.0 |
| | rs3093227 | | 5G | CGCGGTGGG | 2 | 0.6 |
| | rs12984060 | | 5G | TCCCGCGGA | 1 | 0.3 |
| *CYP4F2* | rs3093173 | 6 | 6A | GG | 248 | 83.2 |
| | rs3093169 | | 6B | AA | 46 | 15.4 |
| | | | 6C | AG | 4 | 1.3 |
| *CYP4F2* | rs3093168 | 7 | 7A | TGGCC | 94 | 31.6 |
| | rs2074900 | | 7B | CAACA | 67 | 22.4 |
| | rs1558139 | | 7C | CGACA | 64 | 21.5 |
| | rs3093150 | | 7D | CGGCC | 28 | 9.4 |
| | rs3093145 | | 7E | CGACC | 22 | 7.4 |
| | | | 7F | TGGTC | 22 | 7.3 |
| | | | 7G | TGGCA | 1 | 0.4 |
| *CYP4F11* | rs12459933 | 8 | 8A | CAT | 109 | 36.3 |
| | rs12610962 | | 8B | TGC | 99 | 33.2 |
| | rs1060467 | | 8C | TGT | 49 | 16.4 |
| | | | 8D | TAT | 40 | 13.7 |
| | | | 8E | CGT | 1 | 0.4 |
| *CYP4F11* | rs12971888 | 9 | 9A | GT | 167 | 56.0 |
| | rs1471112 | | 9B | AG | 79 | 26.5 |
| | | | 9C | AT | 52 | 17.5 |
| *CYP4F11* | rs12985091 | 10 | 10A | AC | 118 | 39.5 |
| | rs3826950 | | 10B | GC | 102 | 34.4 |
| | | | 10C | GT | 78 | 26.1 |

**Figure 3.5. Pairwise LD among polymorphisms in the region of *CYP4F12-CYP4F2-CYP4F11* genes in 149 Caucasian samples ($r^2 \geq 0.9$).** The left panel shows 10 distinct haplotype blocks defined by the confidence interval algorithm in HaploView 4.2. The right panel details the haplotypes sequences and their frequencies inferred by PHASE 2.1.

**Table 3.3. Effects of *CYP4F2*, *CYP4F11* and *CYP4F12* haplotypes on hepatic *CYP4F2*, *CYP4F11* and *CYP4F12* mRNA expressions.**

| Gene | Haplotype Identification Code | SNP Code | Haplotype Sequence | mRNA Expression | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *CYP4F2* | | | *CYP4F11* | | | *CYP4F12* | | |
| | | | | $P_c$-value | ↑ or ↓ | $R^2$ (%) | $P_c$-value | ↑ or ↓ | $R^2$ (%) | $P_c$-value | ↑ or ↓ | $R^2$ (%) |
| ***CYP4F12*** | 1A | 1112 | AGTG | 0.347 | - | - | 0.500 | - | - | **$1.81 \times 10^{-4}$** | ↓ | **14.4** |
| | 1C | 2121 | GGAA | 0.522 | - | - | 0.273 | - | - | **$1.49 \times 10^{-5}$** | ↑ | **19.2** |
| ***CYP4F12*** | 2C | 121 | AGT | 0.545 | - | - | 0.469 | - | - | **$1.64 \times 10^{-5}$** | ↑ | **17.9** |
| | 2D | 212 | GAC | 0.443 | - | - | 0.532 | - | - | **0.003** | ↓ | **10.3** |
| **Intergenic** | 3D | 1121 | AGTT | 0.192 | - | - | 0.283 | - | - | **$1.39 \times 10^{-4}$** | ↑ | **15.2** |
| ***CYP4F2*** | 4A | 21 | AT | **0.014** | ↑ | **7.9** | 0.179 | - | - | **0.014** | ↓ | **8.3** |
| ***CYP4F2*** | 5A | 121112111 | TGCGGTGGG | **0.001** | ↑ | **12.6** | **$2.74 \times 10^{-4}$** | ↓ | **13.8** | 0.347 | - | - |
| ***CYP4F2*** | 6B | 22 | AA | 0.381 | - | - | **0.016** | ↓ | **5.8** | 0.156 | - | - |
| ***CYP4F11*** | 8B | 112 | TGC | **0.016** | ↓ | **7.0** | 0.203 | - | - | 0.367 | - | - |
| | 8D | 121 | TAT | **0.010** | ↑ | **8.8** | **0.031** | ↓ | **6.0** | 0.582 | - | - |
| ***CYP4F11*** | 9C | 21 | AT | **0.006** | ↑ | **9.4** | 0.126 | - | - | 0.225 | - | - |

Effects of statistically significant SNPs with *P*-values < 0.05 after FDR are shown in bold.

$R^2$ refers to proportion of variability.

Haplotype identification code refers to the SNP constitution given in Figure 3.5.

SNPs are coded as 1 or 2 (1 = frequent allele, 2 = minor allele).

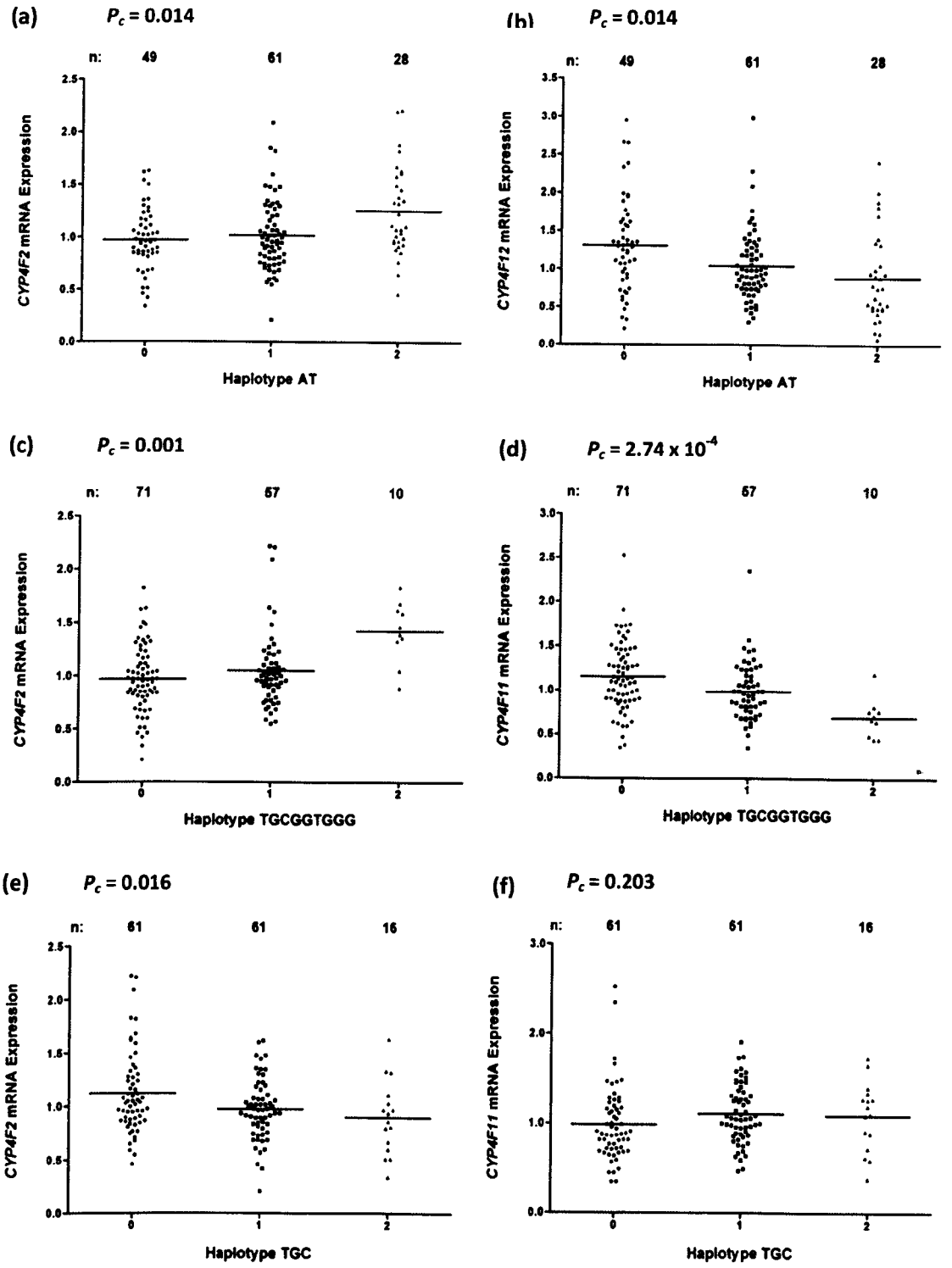



**Figure 3.6. Levels of *CYP4F2*, *CYP4F11* and *CYP4F12* mRNA in normal liver tissue donated from 149 patients in relation to corresponding haplotypes across the *CYP4F2-CYP4F11* locus.** (a)-(b) Haplotype 4A; (c)-(d) Haplotype 5A; (e)-(f) Haplotype 8B. *P*-values after FDR are shown in the upper left corner. Each dot represents an individual and the solid lines represent the mean values.

### 3.3.4 Association of microsomal CYP4F2 protein with SNPs and haplotypes in the *CYP4F2* and *CYP4F11* region

The association of rs2108622 with decreasing human hepatic CYP4F2 protein concentration reported by McDonald and colleagues (McDonald *et al.*, 2009) was confirmed in our cohort of 26 microsomal samples ($P_c$ = 0.05, $R^2$ = 25.6%, Figure 3.7a). As summarised in Table 3.4, several other SNPs across the *CYP4F2* region demonstrated an overall down-regulation of CYP4F2 protein expression including rs2189784 ($P_c$ = 0.103, Figure 3.7b) while *CYP4F11* SNPs were associated with increasing CYP4F2 protein expression. However, none of these associations attained statistical significance after correction for multiple testing using FDR.

It is important to mention that *CYP4F11* mRNA was significantly correlated with CYP4F2 protein levels ($r_s$ = 0.724, $P$ = 2.93 x $10^{-5}$, Figure 3.8a) but neither *CYP4F2* nor *CYP4F12* mRNA was significantly correlated with CYP4F2 protein levels ($r_s$ = -0.214, $P$ = 0.293; $r_s$ = 0.351, $P$ = 0.079, Figures 3.9b & c).

Table 3.5 summarises the *CYP4F2* and *CYP4F11* haplotype associations with CYP4F2 protein expression. Similar to the SNP analyses, haplotypes in the *CYP4F2* region showed an overall down-regulating effect on CYP4F2 protein levels including haplotype 4A (Figure 3.9a) and haplotype 5A (Figure 3.9b) while haplotype 9B in *CYP4F11* exhibited an up-regulating effect on CYP4F2 protein expression. Unfortunately, none of these associations remained statistically significant after FDR except haplotype 5A containing rs2108622 ($P_c$ = 0.05).

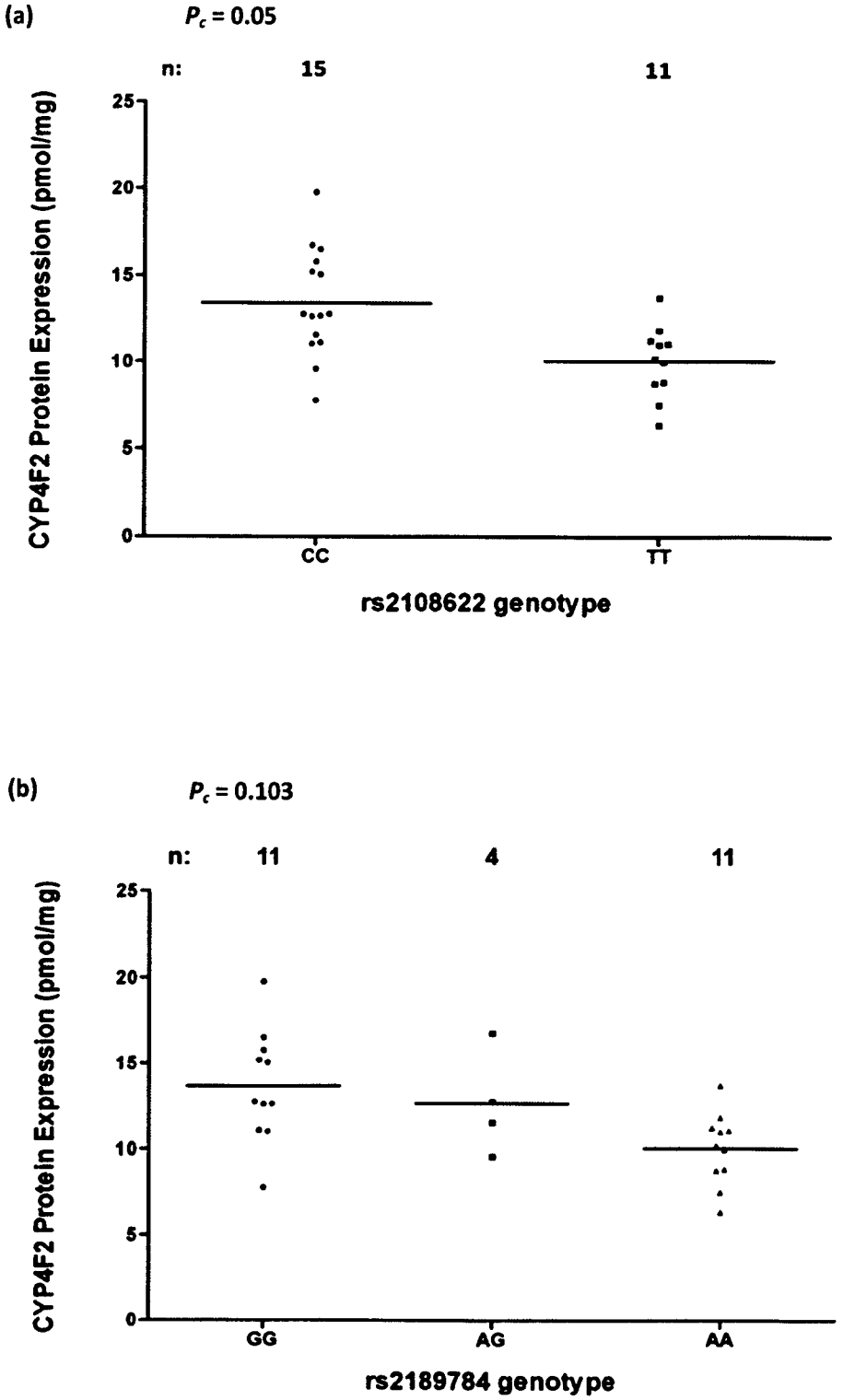



**Figure 3.7. Levels of CYP4F2 microsomal protein in normal liver tissue donated from 149 patients in relation to *CYP4F2* SNPs.** (a) rs2108622; (b) rs2189784. *P*-values after FDR are shown in the upper left corner. Each dot represents an individual and the solid lines represent the mean values.

**Table 3.4. Effects of *CYP4F2* and *CYP4F11* SNPs on hepatic CYP4F2 protein levels.**

| Gene | SNP | Localization | Protein Expression | | | |
|---|---|---|---|---|---|---|
| | | | CYP4F2 | | | |
| | | | *P*-value | $P_c$-value | ↑ or ↓ | $R^2$ (%) |
| **SNPs in *CYP4F2* region** | | | | | | |
| - | rs2189784 | Downstream of *CYP4F2* | 0.018 | 0.103 | ↓ | 23.4 |
| *CYP4F2* | rs3093209 * | 3` near gene | 0.007 | 0.064 | ↓ | 29.7 |
| *CYP4F2* | rs2108622 | Exon 11, Missense, Met433Val | 0.005 | **0.050** | ↓ | **25.6** |
| *CYP4F2* | rs3093173 * | Intron 9 | 0.018 | 0.103 | ↓ | 23.2 |
| *CYP4F2* | rs3093169 * | Intron 9 | 0.021 | 0.117 | ↓ | 16.9 |
| *CYP4F2* | rs984692 | Intron 3 | 0.009 | 0.073 | ↓ | 27.6 |
| **SNPs in *CYP4F11* and *CYP4F12*** | | | | | | |
| *CYP4F11* | rs1471112 | Intron 8 | 0.050 | 0.187 | ↑ | 16.3 |
| *CYP4F11* | rs11086012 | Intron 8 | 0.050 | 0.187 | ↑ | 16.3 |
| *CYP4F11* | rs7249167 | Intron 4 | 0.050 | 0.187 | ↑ | 11.6 |

Effects of statistically significant SNPs with *P*-values < 0.05 after FDR are shown in bold.

$R^2$ refers to proportion of variability.

* Tagging SNP ($r^2 \geq 0.9$). Tagged SNPs are detailed in Appendix 2.3.

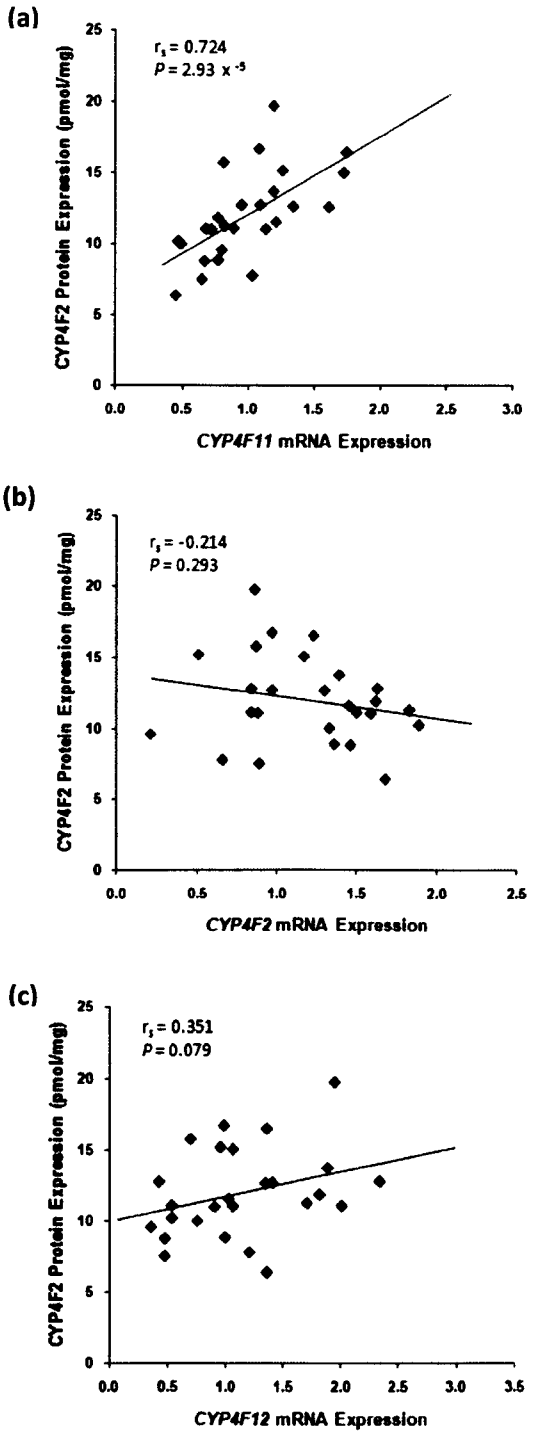



**Figure 3.8. Phenotypic correlations in Caucasian human liver tissues.** **(a)** Correlation between CYP4F2 protein levels and *CYP4F11* mRNA levels (n = 26); **(b)** correlation between CYP4F2 protein levels and *CYP4F2* mRNA levels (n = 26); **(c)** correlation between CYP4F2 protein levels and CYP4F12 mRNA levels (n = 26). The Spearman's rho correlation coefficient ($r_s$) and *P*-value for each comparison are given.

**Table 3.5. Effects of *CYP4F2* and *CYP4F11* haplotypes on hepatic CYP4F2 protein levels.**

| Gene | Haplotype Identification Code | SNP Code | Haplotype Sequence | Protein Expression | | | |
|---|---|---|---|---|---|---|---|
| | | | | CYP4F2 | | | |
| | | | | *P*-value | $P_c$-value | ↑ or ↓ | $R^2$ (%) |
| ***CYP4F2*** | 4A | 21 | AT | 0.018 | 0.103 | ↓ | 23.4 |
| ***CYP4F2*** | 5A | 121112111 | TGCGGTGGG | 0.005 | **0.050** | ↓ | **25.6** |
| ***CYP4F2*** | 6B | 22 | AA | 0.021 | 0.117 | ↓ | 16.9 |
| | 6C | 21 | AG | 0.006 | 0.059 | ↓ | 24.0 |
| ***CYP4F2*** | 7D | 11211 | CGGCC | 0.006 | 0.059 | ↓ | 24.0 |
| ***CYP4F11*** | 9B | 22 | AG | 0.050 | 0.187 | ↑ | 16.3 |

Effects of statistically significant SNPs with *P*-values < 0.05 after FDR are shown in bold.

$R^2$ refers to proportion of variability.

Haplotype identification code refers to the SNP constitution given in Figure 3.5.

SNPs are coded as 1 or 2 (1 = frequent allele, 2 = minor allele).

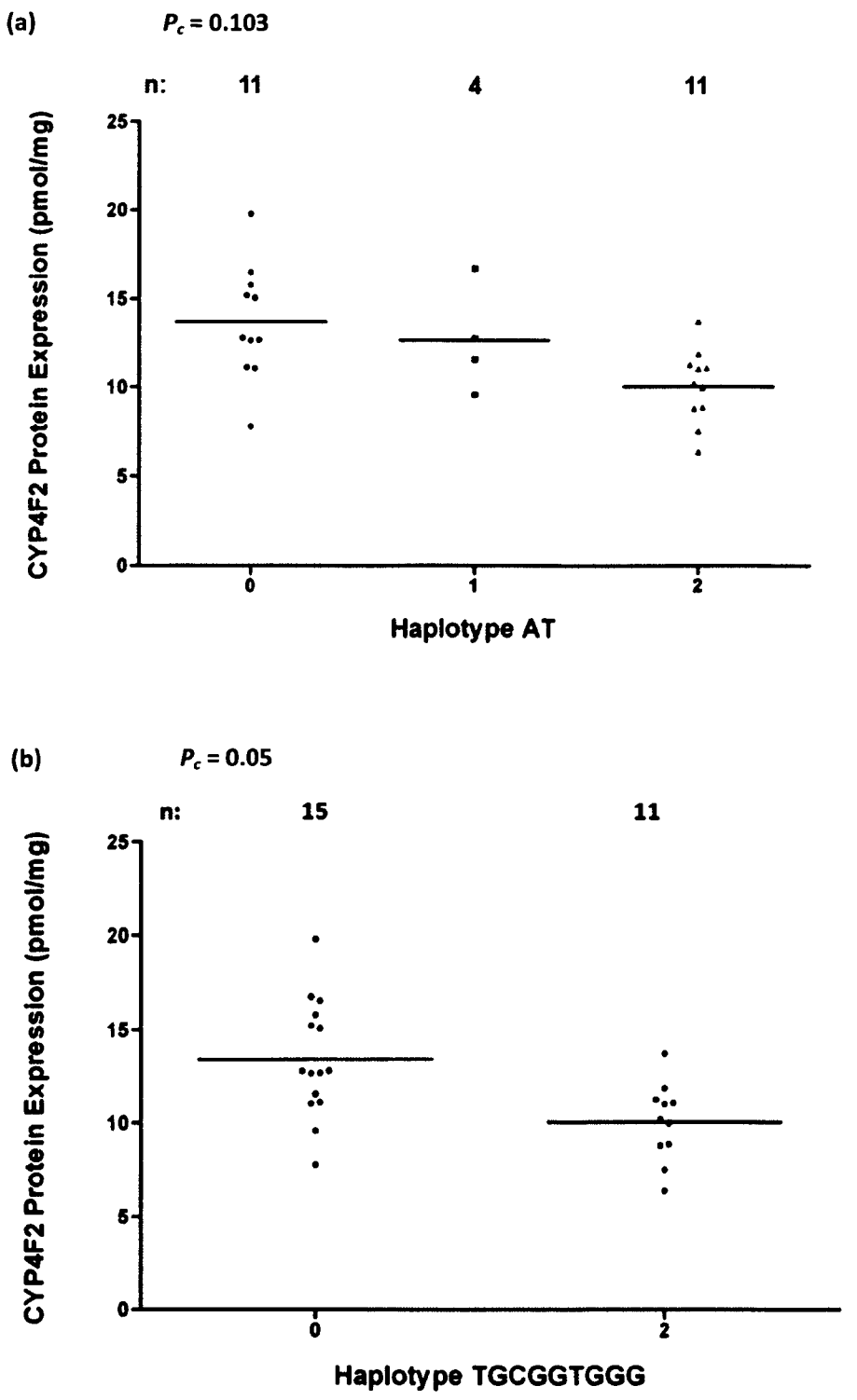



**Figure 3.9. Levels of CYP4F2 microsomal protein in normal liver tissue donated from 149 patients in relation to *CYP4F2* haplotypes.** (a) Haplotype 4A; (b) Haplotype 5A. *P*-values after FDR are shown in the upper left corner. Each dot represents an individual and the solid lines represent the mean values.

### 3.3.5 Effect of rs1060467 in *CYP4F11* on warfarin stable dose and time to therapeutic INR

LD analysis of genotypes in our cohort of 149 liver tissues revealed that rs1060467 in *CYP4F11* is in LD with rs2108622 (D' = 1.0, $r^2$ = 0.206) but in low LD with rs2189784 (D' = 0.557, $r^2$ = 0.121). To assess the role of rs1060467 in warfarin response, data from a previously conducted genome-wide association study (GWAS) looking at variability in warfarin response (Stephane *et al.*, unpublished) was used to test the association of rs1060467 with two clinical outcomes: (i) warfarin stable dose and (ii) time to therapeutic INR. Demographic details of the 714 patients are summarised in Table 3.6. Of the 714 patients, 352 (49%) achieved warfarin stable dose, and 592 (83%) achieved therapeutic INR during the follow-up period. Figure 3.10a illustrates warfarin stable dose established in patients, stratified by rs1060467 genotype. Patients with a C allele exhibited reduced stable dose requirements (mg/week: TT = 32.4 ± 1.5, TC = 28.0 ± 1.1, CC = 26.2 ± 1.7; $P$ = 0.003). The proportion of warfarin dose variability explained by rs1060467 was 2.5%. No association between rs1060467 and time to therapeutic INR was observed ($P$ = 0.820).

### 3.3.6 *In silico* genotype-phenotype analysis

In addition to our genotype-phenotype correlation analyses in human liver tissues, eQTLs in the region encompassing *CYP4F2*, *CYP4F11* and *CYP4F12* genes (chr19: 15635000-16050000) was also assessed using the eQTL database (Imai *et al.*, 2000)hosted by the Pritchard laboratories at the University of Chicago. Table 3.7 outlines the significant SNP-gene associations available on the eQTL database. Of particular interest is the positive association of rs7248867, a SNP located between

*CYP4F12* and *CYP4F2*, with *CYP4F11* transcript levels in livers from individuals of European descent. Using genotype data available on HapMap, LD analysis revealed that this intergenic SNP is in LD with both rs2189784 (D' = 1.0, $r^2$ = 0.103) and rs2108622 (D' = 1.0, $r^2$ = 0.046). rs7248867 also tags several SNPs (using $r^2 > 0.8$) including a *CYP4F12* intronic SNP, rs2074568 (D' = 1.0, $r^2$ = 0.846) which was analysed in our cohort of 149 individuals who had donated liver samples. rs2074568 showed significant association with hepatic mRNA expression of *CYP4F12* ($P_c$ = 0.003) but not *CYP4F11* ($P_c$ = 0.251) nor *CYP4F2* ($P_c$ = 0.779).

### 3.3.7 Effect of rs7248867 and rs2074568 on warfarin stable dose and time to therapeutic INR

Genotypes from the 1000 genomes project were imputed to evaluate the effect of rs7248867 and rs2074568 on warfarin stable dose and time to therapeutic INR. As depicted in Figure 3.10b, patients carrying the minor rs7248867 T-allele required lower warfarin doses compared to patients carrying the major C-allele (mg/week: CC = 30.5 ± 1.0, CT = 26.5 ± 1.3, TT = 22.2 ± 2.5; $P$ = 0.007). Similarly, rs2074568 (Figure 3.10c) showed a recessive effect on warfarin dose requirements with the minor A-allele (mg/week: TT = 30.1 ± 1.0, TA = 27.8 ± 1.4, AA = 21.7 ± 2.4; $P$ = 0.041). No association with time to therapeutic INR was found with neither SNPs.

**Table 3.6. Clinical profile of 714 warfarin patients.**

| Characteristic | N (%) |
|---|---|
| **Gender - Male** | 397 (56) |
| **Age in years, mean (range)** | 69 (19-95) |
| **BMI[a], mean (range)** | 28 (13-55) |
| **Ethnicity** | |
| White | 713 (99.9) |
| Black | 1 (0.1) |
| **Indication for warfarin** | |
| Atrial Fibrillation | 474 (66) |
| Pulmonary Embolism | 110 (15) |
| Deep Vein Thrombosis | 76 (11) |
| Cerebrovascular accident and Transient ischaemic attacks | 44 (6) |
| Mechanical heart valve replacement | 9 (1.3) |
| Myocardial infarction | 3 (0.4) |
| Dilated left atrium | 2 (0.3) |
| Other[b] | 34 (5) |
| **Co-morbidity** | |
| Cardiovascular disease | 576 (81) |
| Musculoskeletal problems | 427 (60) |
| Respiratory disease | 268 (38) |
| Gastrointestinal disease | 254 (36) |
| Neurological disease | 158 (22) |
| Urological condition | 132 (18) |
| Renal disease | 76 (11) |
| History of falls | 58 (8) |
| Hepatic disease | 34 (5) |

BMI: Body Mass Index.

[a] BMI missing for 6 patients.

[b] Other indications include: prevention of clotting in arm for dialysis; systemic lupus erythematosus; anti-phospholipid syndrome; short saphenous vein thrombosis; valvular heart disease; sagittal sinus thrombosis; dilated left ventricle; occluded graft in leg; pulmonary hypertension; apical aneurysm; urticaria with angioedema; femoral embolectomy; aortic and mitral regurgitation; ischaemic colitis; mitral stenosis; and post-surgery.

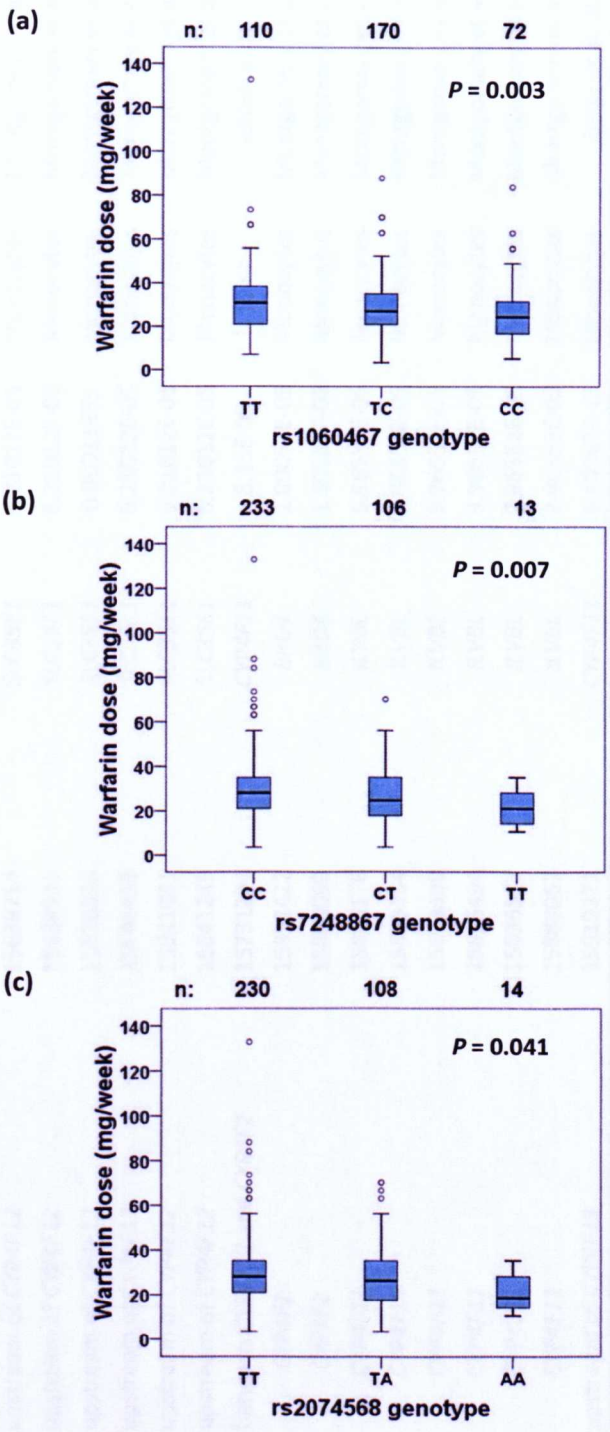



**Figure 3.10. Box and whisker plots showing the distribution of warfarin weekly doses based on genotype groups**. (a) rs1060467; (b) rs7248867; (c) rs2074568. Boxes represent $25^{th}$ – $75^{th}$ percentiles of warfarin doses, whiskers represent $5^{th}$ – $95^{th}$ percentiles, and solid lines represent median dose in each group. Open dots represent outliers.

**Table 3.7. eQTLs in the *CYP4F12-CYP4F2-CYP4F11* gene cluster region.**

| SNP | SNP Localisation | SNP Chromosomal Location | Target eQTL Gene | *P*-Value | Tissue | Study |
|---|---|---|---|---|---|---|
| rs7246556 | 5' upstream of *CYP4F12* | 15637511 | *SLC35E1* | 9.75439E-05 | Monocytes | Montgomery et al, 2010 |
| rs4808351 | 5' upstream of *CYP4F12* | 15638714 | *SLC35E1* | 6.21012E-05 | Monocytes | Montgomery et al, 2010 |
| rs4807967 | 5' upstream of *CYP4F12* | 15638931 | *SLC35E1* | 6.21012E-05 | Monocytes | Montgomery et al, 2010 |
| rs4808352 | 5' upstream of *CYP4F12* | 15638996 | *SLC35E1* | 0.00011855 | Monocytes | Montgomery et al, 2010 |
| rs10409673 | 5' upstream of *CYP4F12* | 15640453 | *SLC35E1* | 6.21012E-05 | Monocytes | Montgomery et al, 2010 |
| rs7251084 | 5' upstream of *CYP4F12* | 15641041 | *SLC35E1* | 6.21012E-05 | Monocytes | Montgomery et al, 2010 |
| rs7259028 | 5' upstream of *CYP4F12* | 15641245 | *SLC35E1* | 6.21012E-05 | Monocytes | Montgomery et al, 2010 |
| rs7248867 | Intergenic, between *CYP4F12* and *CYP4F2* | 15731204 | *CYP4F11* | 8.23E-05 | Liver | Schadt et al, 2008 |
| rs2074901 | *CYP4F2* | 15858422 | *BRD4* | 1.80053E-05 | Monocytes | Montgomery et al, 2010 |
| rs2074902 | *CYP4F2* | 15869099 | *BRD4* | 1.80053E-05 | Monocytes | Montgomery et al, 2010 |
| rs1060463 | *CYP4F11* | 15886176 | *ILVBL* | 9.96552E-05 | Monocytes | Montgomery et al, 2010 |
| rs6512075 | *CYP4F11* | 15899334 | *ILVBL* | 9.96552E-05 | Monocytes | Montgomery et al, 2010 |
| rs3746154 | *CYP4F11* | 15899390 | *ILVBL* | 9.96552E-05 | Monocytes | Montgomery et al, 2010 |
| rs3746156 | *CYP4F11* | 15896494 | *ILVBL* | 9.96552E-05 | Monocytes | Montgomery et al, 2010 |
| rs2219358 | *CYP4F11* | 15896517 | *ILVBL* | 9.96552E-05 | Monocytes | Montgomery et al, 2010 |
| rs2305803 | *CYP4F11* | 15888067 | *ILVBL* | 9.96552E-05 | Monocytes | Montgomery et al, 2010 |
| rs17641483 | 5' upstream of *CYP4F11* | 15919371 | *CYP4F11* | 6.40767E-13 | Monocytes | Zeller et al, 2010 |

Chromosomal positions are given in base pairs from the p-telomere of chromosome 19, as per HapMap Data release 27, NCBI B36 assembly, dbSNP b126.

*BRD4*: Bromodomain containing protein 4

*ILVBL*: Acetolactate synthase-like protein

*SLC35E1*: solute carrier family 35, member E1

### 3.3.8 Interaction between *CYP4F2*, *CYP4F11* and *CYP4F12*

MetaCore™ was used to put *CYP4F2*, *CYP4F11* and *CYP4F12* into a cellular context to evaluate the significance of gene networks that these three genes participate in and to identify regulatory cascades that lead to or from these genes. Figure 3.11 highlights the network interactions of nuclear transcriptional factors relating to the gene expression of *CYP4F2*, *CYP4F11* and *CYP4F12*. Several nuclear transcription factors including pregnane X receptor (PXR), aryl hydrocarbon receptor (AHR), activator protein 1(AP-1), peroxisome proliferator-activated receptor (PPAR), sterol regulatory element-binding protein 1 (SREBP-1), retinoid X receptor (RXR) and retinoic acid receptor (RAR) are known to be involved in the metabolism and clearance of diverse endogenous and exogenous compounds as well as gene activation (Pavek and Dvorak 2008; Zhou *et al.*, 2009).

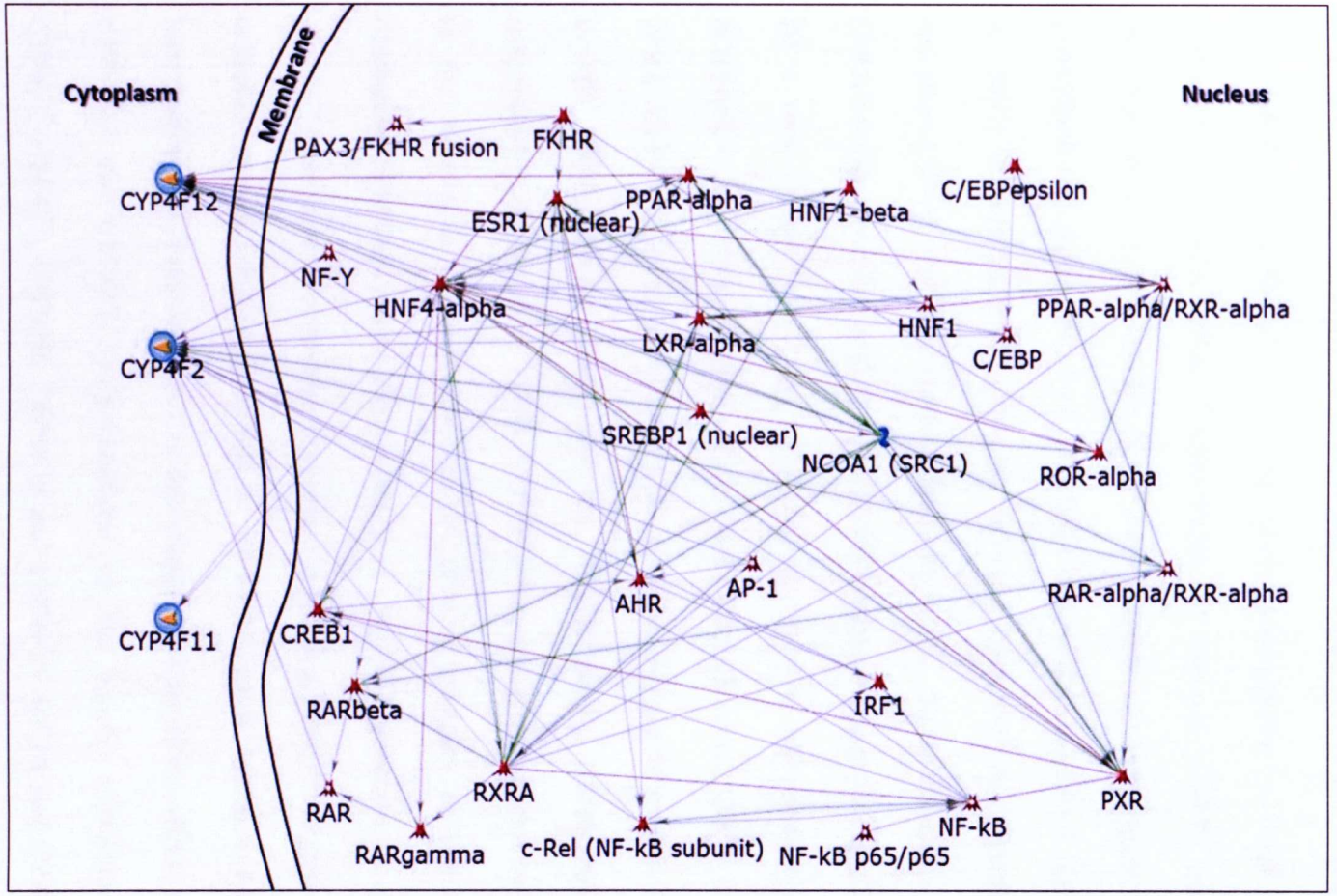



**Figure 3.11. GeneGo graphic illustrating the interaction of nuclear factors with *CYP4F2, CYP4F11* and *CYP4F12* genes.** Each connection represents a direct, experimentally confirmed, physical interaction between the objects.

## 3.4 DISCUSSION

Members of the *CYP4F* subfamily are known for their roles in the metabolism of both endogenous and exogenous compounds. *CYP4F*s are involved in the catabolism of substrates such as fatty acids, prostaglandins, and steroids, and they also catalyse the metabolism of many drugs. For example, *CYP4F2* and *CYP4F3B* have been shown to catalyse the initial O-demethylation of the antiparasitic prodrug pafuramidine by human liver and intestinal microsomes (Wang *et al.*, 2006; Wang *et al.*, 2007). In addition, *CYP4F2* has recently been reported to be a vitamin K oxidase and plays a role in warfarin response (McDonald *et al.*, 2009). Candidate gene(s) and GWAS studies have shown that the *CYP4F2* functional variant, rs2108622, accounts for a small proportion of the variability in warfarin dose requirement (1-7%) (Caldwell *et al.*, 2008; Borgiani *et al.*, 2009; Takeuchi *et al.*, 2009; Pautas *et al.*, 2010). *CYP4F11* is known to be active in the metabolism of several drugs which include erythromycin, benzphetamine, ethylmorphine, chlorpromazine, and imipramine (Kalsotra *et al.*, 2004). *CYP4F12* has also been reported to be involved in the conversion of the antihistaminic prodrug ebastine to the active drug carebastine by hydroxylation (Hashizume *et al.*, 2001; Hashizume *et al.*, 2002).

To elucidate whether the association between genotype and gene expression reflected *cis*-acting regulatory effects on the *CYP4F* gene cluster, a comprehensive investigation has been conducted in this chapter looking at the effects of *CYP4F2*, *CYP4F11* and *CYP4F12* polymorphisms on the hepatic expression levels of *CYP4F2, CYP4F3, CYP4F8, CYP4F11* and *CYP4F12* mRNA and CYP4F2 protein in a Caucasian population. This study reports for the first time that SNPs and extended haplotypes in *CYP4F2*, *CYP4F11* and *CYP4F12* affect the expression

levels of *CYP4F2, CYP4F11* and *CYP4F12* in human liver tissues and that *CYP4F11* plays a role in warfarin response.

The association of rs2108622 with reduced level of human microsomal CYP4F2 protein concentration (McDonald *et al.,* 2009) were confirmed in 26 genotyped microsomal samples. In addition, haplotype 5A harbouring this SNP showed a borderline association with CYP4F2 protein levels in our study. Several other SNPs and haplotypes in the *CYP4F2-CYP4F11* locus also demonstrated associations with CYP4F2 protein expression but their statistical significance did not withstand correction for multiple testing. The small number of available microsomal samples with detectable CYP4F2 protein content (n = 26) may be underpowered to detect significant differences for the numerous polymorphisms tested.

Contrary to McDonald *et al.*'s recent finding (2009), our study observed an association with rs2108622 which explained over 12% of *CYP4F2* mRNA expression in our 149 human liver samples. Likely reasons for this difference in results could be due to the differences in sample size (n = 149 in our study versus n = 59 in McDonald *et al.*'s study) and differences in liver tissue quality and origin. In addition, our data demonstrated that rs2108622 accounts for nearly 14% of *CYP4F11* hepatic mRNA expression. Comparatively, the haplotype harbouring this *CYP4F2* variant also displayed similar associations. Further investigation of SNPs in the *CYP4F11* region revealed a variant in the 3'UTR, rs1060467, to be associated with decreased *CYP4F2* mRNA expression. This SNP explained approximately 7% of *CYP4F2* mRNA expression and the *CYP4F11* haplotype comprising the minor rs1060467 C-allele had a corresponding recessive effect on *CYP4F2* mRNA expression.

Given the LD between rs1060467 and rs2108622 (D' = 1.0), it was hypothesised that rs1060467 may play a role in warfarin stable dose. Using our GWAS data (unpublished) previously conducted in 714 prospective patients on warfarin therapy, rs1060467 explained 2.5% of warfarin dose variability and a difference in mean weekly warfarin dose requirement was observed with the rs1060467 genotype, where our TT patients required a mean value of 4.63 mg/day compared to 3.74 mg/day of our CC patients. Interestingly, the association of rs1060467 with warfarin dose is opposite to that seen with *CYP4F2* rs2108622 as previously reported by other studies (Caldwell *et al.,* 2008; Borgiani *et al.,* 2009; Takeuchi *et al.,* 2009). The minor allele of rs1060467 was associated with reduced warfarin requirements while that of rs2108622 was associated with increased warfarin requirements, suggesting *CYP4F2* and *CYP4F11* may have compensatory effects on each other.

Our results also showed a significant association of rs2189784, a SNP located 30 kb upstream of *CYP4F2*, with differences in mRNA expression of *CYP4F2* and *CYP4F12*. Interestingly, this SNP has been previously reported in chapter 2 to play a role in time taken to achieve therapeutic INR in patients on prospective warfarin therapy (Zhang *et al.,* 2009). Likewise, the haplotype containing the minor A-allele of variant rs2189784 (haplotype 4A) was also significantly associated with increasing *CYP4F2* and decreasing *CYP4F12* mRNA expression. These results suggest that the previously observed association between rs2189784 and time to therapeutic INR may be mediated through an effect on *CYP4F2* and *CYP4F12* mRNA and SNPs in *CYP4F12* may affect *CYP4F2* mRNA expression. Evaluation of variants across the *CYP4F12* region however, did not show any SNPs to be associated with *CYP4F2* mRNA expression.

This study also gained an insight into the gene regulation of the *CYP4F* cluster from publicly available data repositories. *In silico* eQTL analysis revealed the association of *CYP4F11* mRNA expression with an intergenic SNP between *CYP4F12* and *CYP4F2*, rs7248867. This SNP is tagged by a *CYP4F12* intronic SNP (rs2074568) genotyped in our study. These two SNPs were however, not present on the GWAS platform. Imputations were therefore performed and a trend for reduced warfarin stable dose was observed with these two SNPs.

Using the program MetaCore™ as a bioinformatics network hypothesis generating tool, interactions between *CYP4F2*, *CYP4F11* and *CYP4F12* were shown to be interlinked via numerous nuclear transcription factors such as AP-1, RXR, RAR and SREBP. AP-1 and RXR-mediated pathways have been shown to be involved in the regulation of *CYP4F11* expression (Wang *et al.*, 2010) in human keratinocyte-derived HaCaT cells. Using the human hepatoma (HepG2) cell line, studies have shown that retinoic acids and peroxisome proliferators can regulate *CYP4F2* activities, where RXR stimulated whilst RAR repressed *CYP4F2* expression (Zhang *et al.*, 2000; Zhang and Hardwick 2000). Additionally, a report by Hsu and colleagues suggests that SREBP mediates statin induction of *CYP4F2* expression in primary human hepatocytes and HepG2 cells (Hsu *et al.*, 2007). PXR was found to be involved in the regulation of *CYP4F2* expression in healthy human lymphocytes (Siest *et al.*, 2008) and *CYP4F12* expression in primary human hepatocytes (Hariparsad *et al.*, 2009). A very recent study reported that R-warfarin interacts with PXR nuclear receptor and can significantly up-regulate drug-metabolising enzymes such as CYP3A4 and CYP2C9 in the liver (Rulcova *et al.*, 2010). Further work should include some of these transcription factors.

As summarised in Figure 3.12, this study has effectively examined sequence variations across the three *CYP4F* genes – *CYP4F2*, *CYP4F11* and *CYP4F12* and their effect on *CYP4F* gene cluster expression. The expression of these genes in liver could be a useful tool for future pathological and physiological studies in relation to drug intake and metabolic pathways. Our work has also shown that fine mapping via custom design allows greater coverage of SNPs which are not available on the GWAS platform. Out of the 147 SNPs successfully genotyped in this study, only 26 were present in our GWAS. Due to the high degree of LD across the region encompassing *CYP4F2*, *CYP4F11* and *CYP4F12*, it is possible that these three genes play a synergistic role in warfarin response. The selection of all SNPs included in this study was dependent on data submitted to public repositories. As SNPs were selected from the International HapMap project and NCBI Entrez SNP database, it is possible that additional common variants or important functional polymorphisms which were not genotyped in the CEU population may be involved. In order to identify the true causal variant(s), resequencing of these three genes would be required to exhaustively assess all variants spanning these genes.

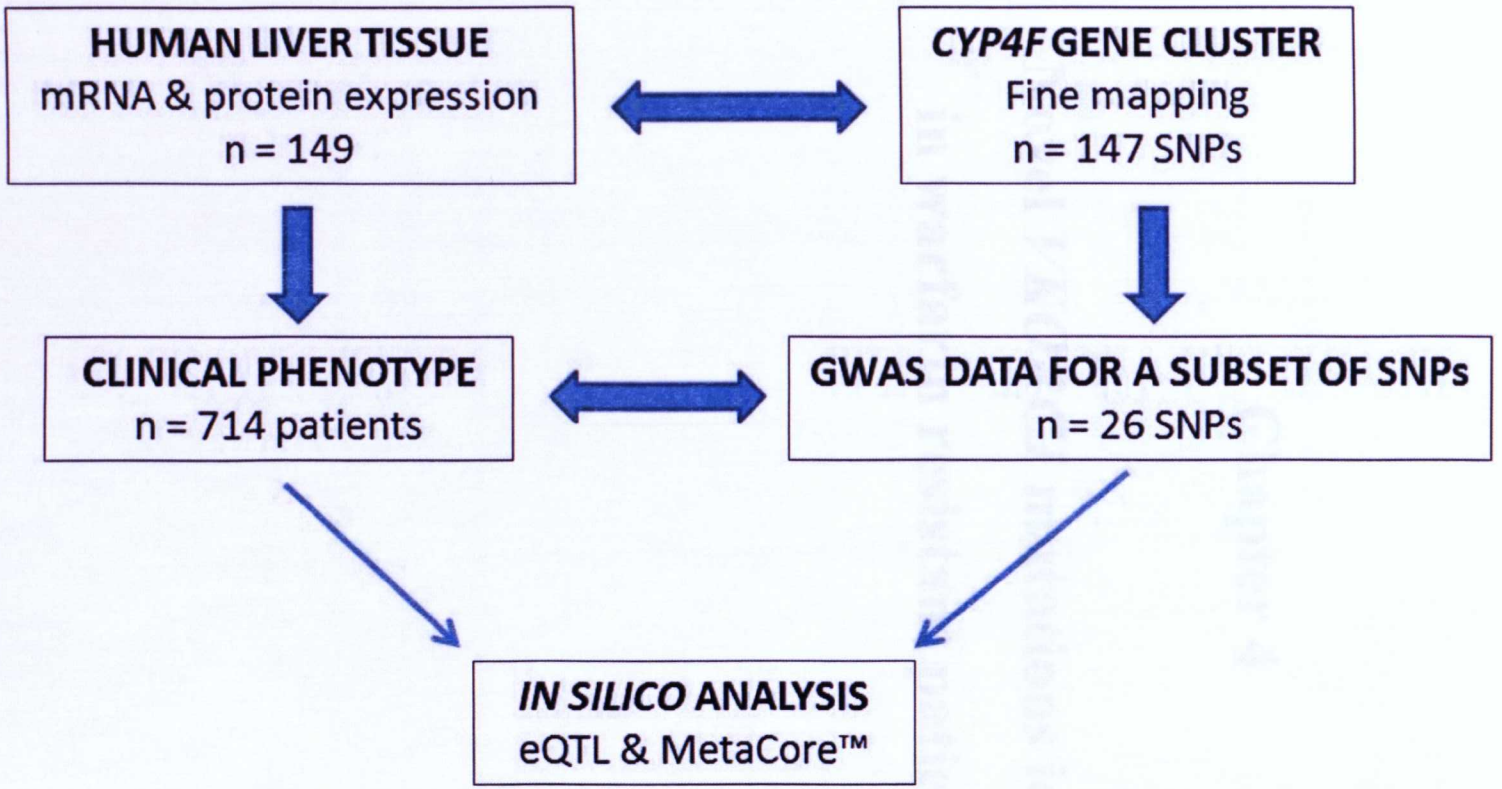



**Figure 3.12. Genotype-phenotype correlation and *in silico* analysis.**

# Chapter 4

# Novel *VKORC1* mutations identified in warfarin resistant patients

# CONTENTS

**4.1 INTRODUCTION** ..... **115**

**4.2 PATIENTS, MATERIALS AND METHODS** ..... **118**

4.2.1 Patient population ..... 118

4.2.2 Warfarin assay ..... 119

4.2.3 Establishing warfarin reference range ..... 120

4.2.4 DNA extraction ..... 120

4.2.5 Amplification of the *VKORC1* gene ..... 123

4.2.6 Sequencing of the *VKORC1* gene ..... 126

4.2.7 Genotyping of *CYP2C9* and *VKORC1* ..... 128

4.2.8 Bioinformatics analysis ..... 128

*4.2.8.1 Prediction of transcription factor binding site* ..... *128*

*4.2.8.2 Prediction of functional consequences of mutations* ..... *128*

*4.2.8.3 Potential splicing aberrations* ..... *130*

**4.3 RESULTS** ..... **132**

4.3.1 Patient population ..... 132

4.3.2 Reference population and warfarin assay ..... 134

4.3.3 Pharmacodynamic warfarin resistant patients ..... 135

4.3.4 Novel *VKORC1* mutations identified ..... 136

4.3.5 Characteristics of patients carrying *VKORC1* mutations ..... 139

4.3.6 Putative transcription factor binding site, protein and splice site analyses ..... 140

**4.4 DISCUSSION** ..... **141**

## 4.1 INTRODUCTION

Currently available pharmacogenetic warfarin dosing algorithms incorporating the common genetic variations *CYP2C9*2* (rs1799853), *CYP2C9*3* (rs1057910) and *VKORC1* -1639 (rs9923231) as well as clinical variables, such as age, weight, and body surface area, have shown to provide better dose prediction compared to clinical algorithms or fixed-dose approaches. (Gage *et al.,* 2008; Klein *et al.,* 2009; Wadelius *et al.,* 2009; Lenzini *et al.,* 2010). However, a large difference in dose requirements between the pharmacogenetic estimation approach and the actual therapeutic dose is observed in patients who required very high doses of warfarin (Klein *et al.,* 2009; Lenzini *et al.,* 2010), indicating that the common SNPs in *CYP2C9* and *VKORC1* do not explain the extremely high warfarin dose requirements in rare individuals with warfarin resistance (James *et al.,* 1992; Oldenburg *et al.,* 2007).

Warfarin resistance has been defined as the inability to prolong INR into the therapeutic range when the drug is given at doses prescribed to the majority of patients (Lefrere *et al.,* 1987). The prevalence of warfarin resistance is rare. Transient causes of acquired resistance to warfarin include poor patient compliance, high vitamin K consumption and concomitant drugs that induce the metabolism of warfarin. Warfarin resistance can be broadly classified into two categories, pharmacokinetic resistance or pharmacodynamic resistance.

Pharmacokinetic resistance can result from diminished absorption or increased metabolism of warfarin, resulting in a rapid decline in plasma warfarin levels. Causes of diminished absorption include emesis, diarrhoea, and malabsorption syndrome, however, since warfarin is absorbed by passive diffusion, malabsorption is extremely uncommon, even in individuals with significant small

bowel disease. Enhanced warfarin metabolism would likely arise through the induction of CYP2C9 and to a lesser extent, CYP1A1, CYP1A2 and CYP3A4. This is commonly due to the copy number of enzyme inducers. Gene duplication or multiplication of *CYP2C9*, *CYP1A1*, *CYP1A2* or *CYP3A4* has not been reported, unlike *CYP2D6*, where gene duplication and multiplication causes ultra-rapid metabolism of debrisoquine, tramadol and codeine (Johansson *et al.,* 1993; Gasche *et al.,* 2004; Kirchheiner *et al.,* 2008).

Pharmacodynamic warfarin resistance is indicated when an elevated circulatory warfarin concentration beyond that of the upper limit of the therapeutic range is achieved through high daily doses of warfarin, while the INR remains consistently sub-therapeutic. The mechanism of pharmacodynamic warfarin resistance in man has not been delineated but it has been postulated that increased affinity of *VKORC1* for vitamin K (O'Reilly *et al.,* 1968; Cain *et al.,* 1998) and decreased *VKORC1* sensitivity to warfarin (Cain *et al.,* 1998) may be important.

Since the discovery of the *VKORC1* gene in 2004 (Li *et al.,* 2004; Rost *et al.,* 2004a), 12 rare missense mutations in *VKORC1* have been reported in patients resistant to warfarin (Table 4.1). The studies however were either isolated case reports (Rost *et al.,* 2004a; Harrington *et al.,* 2005; D'Ambrosio *et al.,* 2007; Loebstein *et al.,* 2007; Ainle *et al.,* 2008) or in a small number of patients (Harrington *et al.,* 2008; Watzka *et al.,* 2011).

The aim of this chapter was therefore to determine if rare mutations are present in the promoter and exon-intron boundaries of the *VKORC1* gene in addition to the exonic regions in our cohort of warfarin resistant patients. In addition, bioinformatics analyses of the novel mutations identified were also undertaken.

**Table 4.1. Previously reported *VKORC1* mutations in warfarin resistant patients.**

| Mutation[a] | Nucleotide[b] | dbSNP rs number | VKOR protein domain[c] | References |
|---|---|---|---|---|
| p.Ala26Pro | c.76G>C | - | TM1 | (Bodin *et al.*, 2005a) |
| p.Val29Leu | c.85G>T | rs104894539 | TM1 | (Rost *et al.*, 2004a) |
| p.Asp36Tyr | c.106G>T | rs61742245 | CL | (Bodin *et al.*, 2005b; D'Ambrosio *et al.*, 2007; Loebstein *et al.*, 2007; Harrington *et al.*, 2008; Watzka *et al.*, 2011) |
| p.Asp36Gly | c.107A>G | - | CL | (Watzka *et al.*, 2011) |
| p.Ala41Ser | c.121G>T | - | CL | (Rieder *et al.*, 2005) |
| p.Val45Ala | c.134T>C | rs104894540 | CL | (Rost *et al.*, 2004a) |
| p.Val54Leu | c.160G>C | - | CL | (Bodin *et al.*, 2005b; Harrington *et al.*, 2008) |
| p.Arg58Gly | c.172A>G | rs104894541 | CL | (Rost *et al.*, 2004a) |
| p.Val66Met | c.196G>A | rs72547529 | CL | (Bodin *et al.*, 2005b; Harrington *et al.*, 2005; Harrington *et al.*, 2008) |
| p.Asn77Tyr | c.229A>T | - | CL | (Watzka *et al.*, 2011) |
| p.Leu128Arg | c.383T>G | rs104894542 | TM3 | (Rost *et al.*, 2004a; Bodin *et al.*, 2005b; Ainle *et al.*, 2008; Harrington *et al.*, 2008) |
| p.Arg151Gln | c.452G>A | - | CT | (Rost *et al.*, 2004a) |

[a] Amino acid numbering is according to reference sequence NP_076869 with A of the ATG initiation codon designated as +1.

[b] Nucleotide numbering is according to reference sequence NM_024006 with A of the ATG initiation codon designated as +1.

[c] Predicted VKOR protein domain according to the model proposed by Tie and colleagues (2005) (see Figure 1.3). TM1: transmembrane domain 1; CL: cytoplasmic loop; TM3: transmembrane domain 3; CT: cytoplasmic tail.

## 4.2 PATIENTS, MATERIALS AND METHODS

### 4.2.1 Patient population

Between November 2004 and September 2007, patients treated with warfarin at two hospitals in Liverpool, UK, the Royal Liverpool and Broadgreen University Hospitals Trust and University Hospital Aintree, were recruited prospectively (n = 1000) and retrospectively (n = 69). Demographics and clinical data for all patients were gathered from their medical notes and clinical records. Written informed consent was obtained from all patients and the study was approved by the Birmingham South Research Ethics Committee. The study design of the prospective cohort has been described in section 2.2.1. The retrospective cohort comprised of patients who were either currently or previously on warfarin therapy, and had warfarin sensitivity or resistance (as defined below and in section 2.2.2 in chapter 2).

Patients showing resistance to warfarin were selected from both the prospective and retrospective cohorts for *VKORC1* sequence analysis. Among the 1000 patients recruited prospectively, 29 showed resistance to warfarin. Of the 69 retrospectively recruited patients, 36 showed warfarin resistance. Resistance to warfarin was defined as daily doses of $\geq$ 10 mg on three successive clinic visits or for more than three weeks. Patients were excluded from the study if (i) they had concomitant treatment with drugs classed as CYP inducers (i.e. carbamazepine, phenytoin, phenobarbitone, and rifampicin) which induce the metabolism of warfarin and hence patients would require higher doses of warfarin, and/or (ii) were noncompliant with warfarin therapy as documented in their diaries and questionnaires. Compliance in the prospectively recruited patients was monitored using patient diaries and questionnaire. Patients were asked to document their daily

dose of warfarin consumption between clinic visits. At each follow-up visit, the patient diaries were assessed by our medical staff for adherence, and the patients needed to answer a medication questionnaire listing the number of doses taken in the last 7 days. In addition, plasma warfarin levels were tested for pharmacodynamic resistance as described in the following sections (sections 4.2.2 and 4.2.3). For patients recruited retrospectively, plasma samples however were not collected and warfarin pharmacodynamic resistance could not be determined in these patients.

### 4.2.2 Warfarin assay

For patients recruited prospectively, heparinised whole blood samples were taken approximately 16 hours after the last warfarin dose at each of the four scheduled visits when possible. Plasma was separated and aliquots were stored at –20°C for determination of plasma warfarin concentrations.

In the first 311 patients recruited prospectively (interim analysis) and warfarin resistant patients recruited prospectively, unbound levels of both R and S-warfarin were simultaneously measured using the chiral high-performance liquid chromatography (HPLC) method as previously described (Naidong and Lee 1993; Lane *et al.,* 2011). Measurements were carried out by Sameh Al-Zubiedi in the Department of Pharmacology, University of Liverpool, and Ellen Hatch and Jeremy Palmer, in Professor Farhad Kamali's lab at the Institute of Cellular Medicine, Newcastle University. Briefly, separation of the warfarin enantiomers was performed on a 250 x 4.6 mm β-cyclodextrin column (Cyclobond I 2000, Advanced Separation Technologies Ltd®) coupled with a 200 x 4 mm β-cyclodextrin guard column, on a Dionex HPLC system consisting of a UVD170U variable wavelength detector set at 303 nm. S-naproxen (80 µg/ml; Sigma-Aldrich, Poole, UK) was used as an internal

standard. The mobile phase used consisted of a mixture of acetonitrile, glacial acetic acid, and triethylamine at 1000:3:2.5 ratio (v/v/v) at a flow rate of 1 ml/min over a total run time of 30 min. Intra- and inter-day coefficients of variation for both R- and S-warfarin were less than 6%. The inter- and intra-assay accuracy (% bias) for all quality control (QC) concentrations was within 15% for both R- and S-warfarin. The assay allowed for the quantification of enantiomers of warfarin over a wide concentration range (100-5000 ng/ml). The limit of quantification was set at 100 ng/ml for each warfarin enantiomer; which is sufficiently low to enable patient samples to be analysed with good accuracy and precision.

### 4.2.3 Establishing warfarin reference range

In order to establish diagnostic criteria for pharmacodynamic warfarin resistance in our cohort of patients, we used a method similar to that previously published by Harrington an colleagues (2005). 145 prospective patients who achieved warfarin stable dose (an unchanged daily dose at $\geq$ 3 consecutive clinic visits within the individual's target INR range) (from the interim analysis described in section 2.2.3) with available plasma warfarin concentration, was used as a reference population to define the therapeutic plasma warfarin reference range (95% confidence interval). Prospectively recruited warfarin resistant patients with plasma warfarin levels beyond this therapeutic reference range were classified as showing pharmacodynamic resistance.

### 4.2.4 DNA extraction

Genomic DNA was isolated from EDTA whole blood using the standard phenol-chloroform method, the E.Z.N.A. kit (Omega Bio-tek Inc., GA, USA), or the

Chemagen Whole Blood DNA Extraction Kit (magnetic beads technology) on the Chemagic Magnetic Separation Module I (Auto-Q Biosciences, Germany), according to manufacturer's instructions. The Chemagic Magnetic Separation Module I is based on the polyvinyl alcohol particles (M-PVA) Magnetic Beads technology. The M-PVA Magnetic Beads exhibit both a hydrophilic surface and low non-specific binding properties, resulting in a unique DNA binding matrix which allows for efficient binding to DNA with high yield and purity. The chemistry of this DNA isolation process is outlined in Figure 4.1.

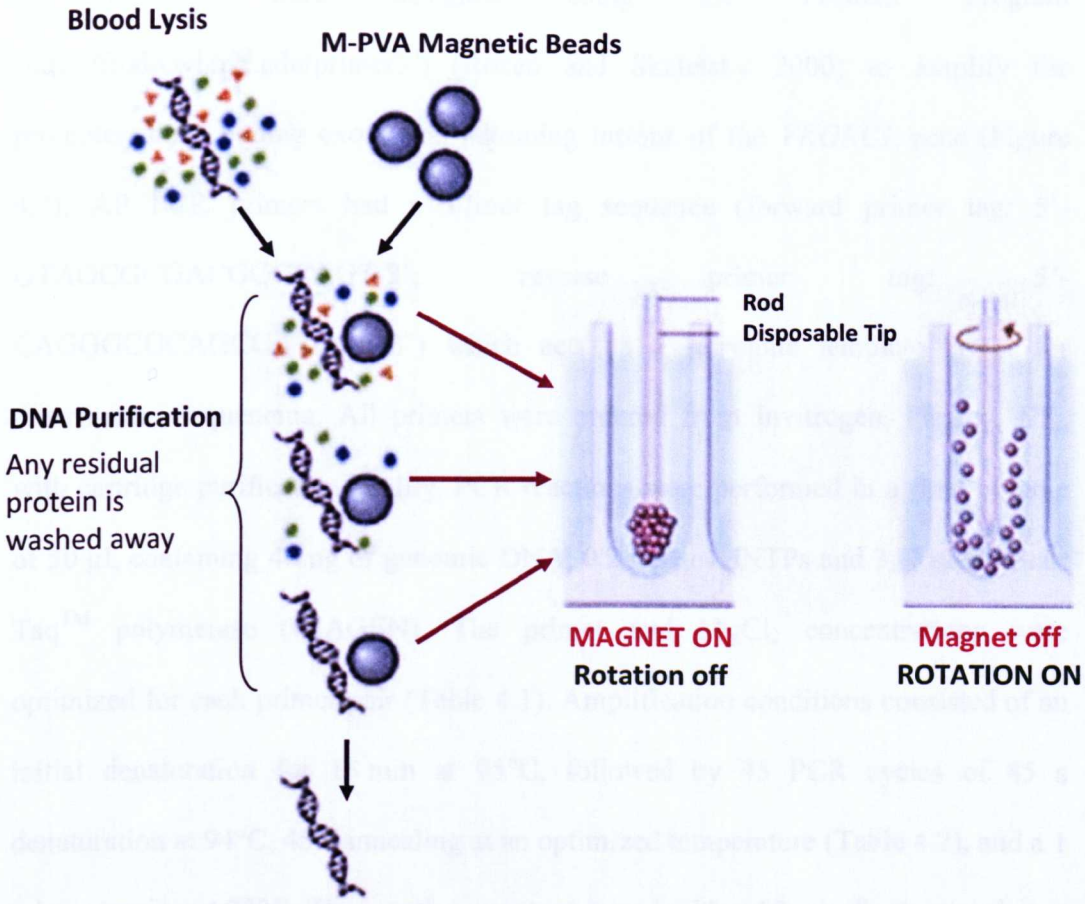



**Figure 4.1 Chemistry of DNA extraction from whole blood using the Chemagic Magnetic Separation Module I.** The lysis of the white and red blood cells is performed in the presence of protease for protein degradation. The isolation of the DNA is achieved through its capture by M-PVA Magnetic Beads. When applying an electromagnetic field, these beads, together with the bound DNA, are attracted to the magnetized metal rods, which can then transfer the DNA from one washing buffer to another. Each rod is covered by a disposable tip to prevent between-run contamination of the rod. At the end of each transfer step, the electromagnet is deactivated and the rotation of the rods is switched on, leading to an efficient and homogeneous resuspension of the particles. In the final step the beads are transferred into elution buffer, which inactivates the interaction between the beads and the DNA. The magnetic beads are then removed, leaving the isolated DNA in suspension.

### 4.2.5 Amplification of the *VKORC1* gene

Primers were designed using the Primer3 program (http://frodo.wi.mit.edu/primer3/) (Rozen and Skaletsky 2000) to amplify the promoter, three coding exons and adjoining introns of the *VKORC1* gene (Figure 4.2). All PCR primers had a 17mer tag sequence (forward primer tag: 5'-GTAGCGCGACGGCCAGT-3'; reverse primer tag: 5'-CAGGGCGCAGCGATGAC-3') which acts as a surrogate template target for downstream sequencing. All primers were ordered from Invitrogen, Paisley, UK, with cartridge purification quality. PCR reactions were performed in a final volume of 50 µl, containing 40 ng of genomic DNA, 0.2 mM of dNTPs and 3 U of HotStart Taq$^{TM}$ polymerase (QIAGEN). The primer and $MgCl_2$ concentrations were optimized for each primer pair (Table 4.1). Amplification conditions consisted of an initial denaturation for 15 min at 95°C, followed by 45 PCR cycles of 45 s denaturation at 94°C, 45 s annealing at an optimized temperature (Table 4.2), and a 1 min extension at 72°C. The reaction was terminated with a 10 min final extension at 72°C. To check the size and specificity of PCR fragments, the PCR products were electrophoresed on a 1.5% ethidium bromide-stained agarose gel in 1x Tris-Borate-EDTA (TBE) buffer for 20 min at 150V, and the gel was visualised under UV light.

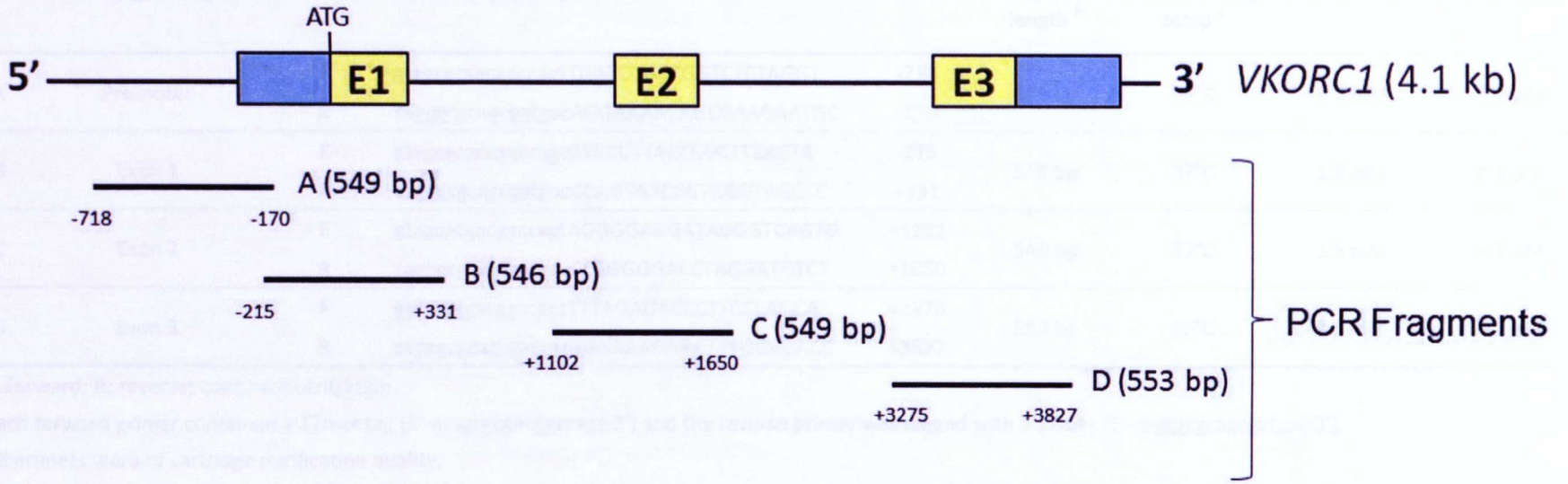



**Figure 4.2. *VKORC1* gene structure and sequencing coverage.** Yellow boxes denote coding regions and blue boxes represent untranslated regions of the exons. Intron regions are indicated by horizontal lines between exons. Four fragments of the *VKORC1* gene spanning nucleotides -718 to +3827 were amplified. Nucleotides are numbered according to reference sequence NM_024006 with the A of the ATG initiation codon designated as +1.

**Table 4.2. Details of PCR primers used for amplifying the *VKORC1* gene for sequencing.**

| | *VKORC1* gene region | Direction | Primer sequence (5' → 3') | Position [a] | Amplicon length [b] | PCR annealing temp [c] | $MgCl_2$ conc. [d] | Primer conc. [e] |
|---|---|---|---|---|---|---|---|---|
| A | Promoter | F | gtagcgcgacggccagtTGATCCGCTGGTCTCTAGGT | -718 | 549 bp | 55°C | 2.5 mM | 0.2 µM |
| | | R | cagggcgcagcgatgacAGGAAAACAGCGAAGAATGC | -170 | | | | |
| B | Exon 1 | F | gtagcgcgacggccagtGTCCCTTACCCGCTTCACTA | -215 | 546 bp | 57°C | 1.5 mM | 0.5 µM |
| | | R | cagggcgcagcgatgacCCAGTATCGCTGGGTAGCTC | +331 | | | | |
| C | Exon 2 | F | gtagcgcgacggccagtAGGGGAGGATAGGGTCAGTG | +1102 | 549 bp | 57°C | 1.5 mM | 0.5 µM |
| | | R | cagggcgcagcgatgacCCTGGGGACCTAGGATGTCT | +1650 | | | | |
| D | Exon 3 | F | gtagcgcgacggccagtTTTAGAGACCCTTCCCAGCA | +3275 | 553 bp | 55°C | 2.5 mM | 0.2 µM |
| | | R | cagggcgcagcgatgacTGGAAAGAGCTTTGGAGACC | +3827 | | | | |

F: forward; R: reverse; conc.: concentration

Each forward primer contained a 17mer tag (5'-gtagcgcgacggccagt-3') and the reverse primer was tagged with a 17mer (5'-cagggcgcagcgatgac-3').

All primers were of cartridge purification quality.

[a] Position of primers relative to ATG translation start site according to reference sequence NM_024006 with the A of the ATG initiation codon designated as +1.

[b] Size of amplified fragments.

[c] Optimised annealing temperature for each set of primers.

[d] Optimised magnesium chloride concentration for each set of primers.

[e] Optimised primer concentration for each set of primers.

### 4.2.6 Sequencing of the *VKORC1* gene

Before identifying the genetic variants in the PCR products by Sanger sequencing, 5 µl of each PCR was incubated with 4 U of ExoSAP-IT® (USB Corporation, Cleveland, Ohio, USA) for 30 min at 37°C in 1.5 x reaction buffer to degrade primers and dephosphorylate dNTPs that were not consumed in the reaction. The reaction was stopped by a 15 min incubation at 80°C.

Sequencing reactions were performed by Kimberley Jones in the genetics core facility at the Liverpool Women's Hospital, Liverpool, UK. PCR products were bidirectionally sequenced using the "UNISEQ" sequencing primers (forward: GTAGCGCGACGGCCAGT; reverse: CAGGGCGCAGCGATGAC). DNA cycle sequencing reactions were carried out in 8 µl reactions using 10 ng of purified PCR product, 200 nM forward and reverse primer, 0.25 µl BigDye® Terminator version 1.1 sequencing buffer (Applied Biosystems) and 3.75 µl dilution buffer (Applied Biosystems) using the following cycling conditions: initial 95°C for 1 min followed by 25 cycles at 96°C for 10 s, 55°C for 10 s and 60°C for 4 min. Sequencing products were then cleaned using the Agencourt CleanSEQ bead system (Beckman Coulter Genomics®, UK) before running on the Applied Biosystems 3730xl DNA analyzer (50 cm capillary, POP7 buffer).

Sequence traces were base called using the Sequence Analysis software version 5.2 (Applied Biosystems) and data quality was checked using the Sequence Scanner freeware version 1.0 (Applied Biosystems). Sequence assembly and analysis were then performed using Mutation Surveyor® software package version 3.24 (SoftGenetics LLC, PA, USA) and were compared to annotated sequences from the human *VKORC1* reference sequence NM_024006.4, as illustrated in Figure 4.3.

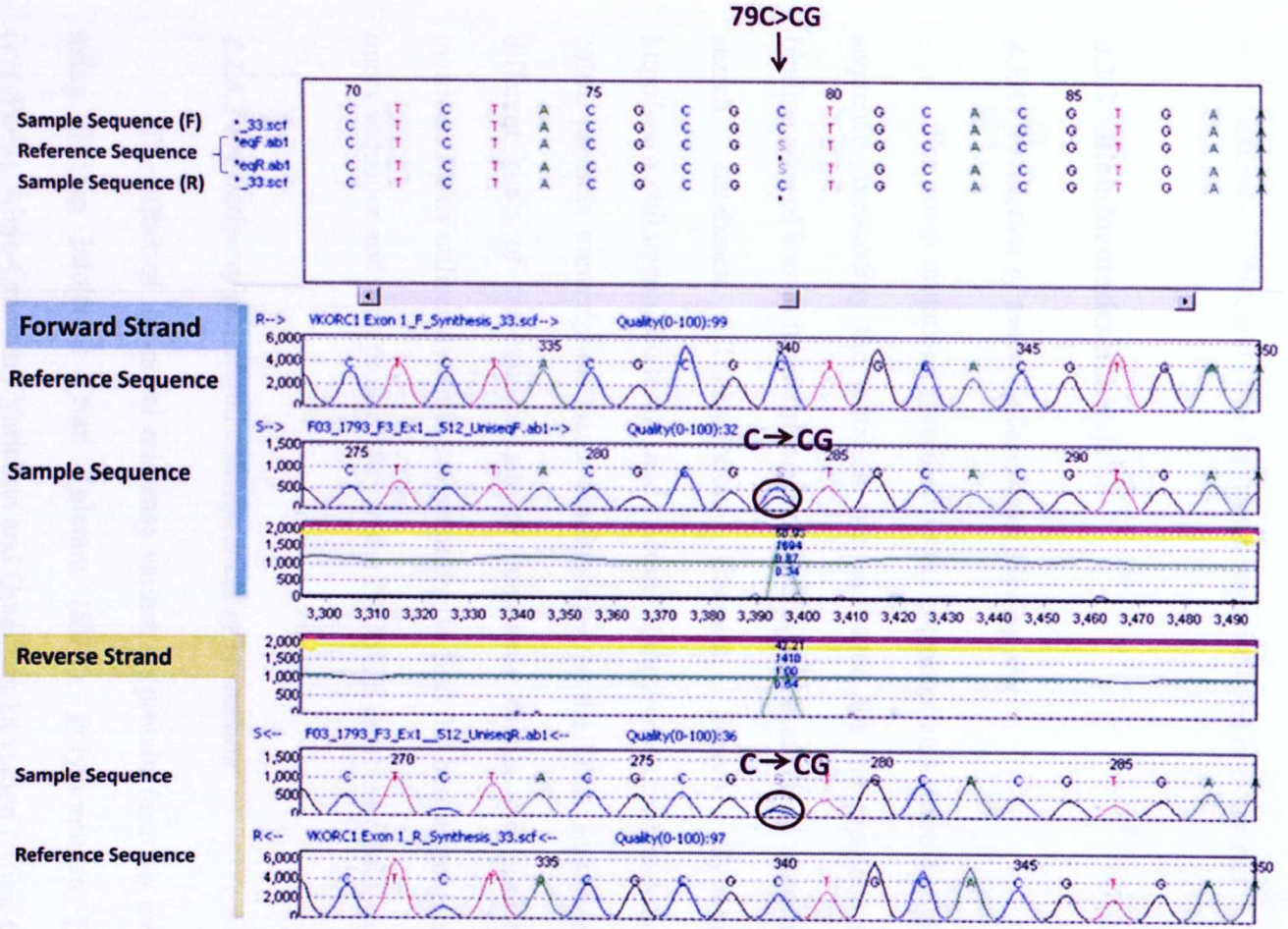



**Figure 4.3 *VKORC1* sequence assembly and analysis using Mutation Surveyor.** Electropherogram showing a region of *VKORC1* exon 1 in a sample heterozygotic at nucleotide position 79. F, forward; R, reverse.

### 4.2.7 Genotyping of *CYP2C9* and *VKORC1*

Genotyping of *CYP2C9*2* (c.430C>T, rs1799853), *CYP2C9*3* (c.1075A>C, rs1057910) and *VKORC1* -1639 (c.-1639G>A, rs9923231) were performed as described in section 2.2.7. As part of quality control, negative controls containing water instead of DNA and 10% duplicates were included in every run

### 4.2.8 Bioinformatics analysis

#### *4.2.8.1 Prediction of transcription factor binding site*

For novel mutations identified in the promoter and intronic regions, a short sequence surrounding the mutation was evaluated for the presence of putative-binding sites of known transcription factors using two established transcription factor search databases: Transcription Element Search System (TESS, http://www.cbil.upenn.edu/tess) and Motif (http://www.genome.jp/tools/motif/). TESS predicts transcription factor binding sites in the DNA sequence using two different kinds of site models, site or consensus strings and positional weight matrices. Motif utilises dynamic programming to find the best alignment between a query sequence and each profile entry in the PROSITE database (Hulo *et al.*, 2006).

#### *4.2.8.2 Prediction of functional consequences of mutations*

The effect of the novel missense variant on protein function was predicted using Sorting Intolerant From Tolerant (SIFT), Polymorphism Phenotyping (PolyPhen), Align-Grantham Variation and Grantham Deviation (Align-GVGD) and I-Mutant2.0.

SIFT (http://sift.jcvi.org/) uses alignments of query sequences to measure conservation of each amino acid between species and calculates whether the biochemical parameters of the exchanged amino acids are similar or disparate (Ng and Henikoff 2001). A SIFT score of less than 0.05 indicates a deleterious amino acid substitution.

PolyPhen (http://genetics.bwh.harvard.edu/pph/) uses conservation of sequences and structural predictions to determine functional consequences of each variant (Sunyaev *et al.,* 2001). PolyPhen scores of less than 1.5 indicate functionally normal variants; scores between 1.5 and 2.0 are categorized as possibly deleterious and greater than 2.0 are categorized as probably deleterious.

Align-GVGD (http://agvgd.iarc.fr/agvgd_input.php) combines the Grantham Variation and Grantham Deviation Scores for variants, comparing biochemical characteristics and conservation of amino acids along the protein sequence (Mathe *et al.,* 2006; Tavtigian *et al.,* 2006). Align-GVGD provides an output score of C(0), C(15) through C(65). A score of C(0) indicates a neutral change, whereas C(15) through C(65) indicates progressively more severe variants.

I-Mutant2.0 (http://gpcr.biocomp.unibo.it/cgi/predictors/I-Mutant2.0/I-Mutant2.0.cgi) is a support vector machine (SVM)-based tool for the prediction of protein stability changes upon single-site mutations (Capriotti *et al.,* 2005). It predicts the direction of protein stability changes by calculating the free energy change of protein stability ($\Delta\Delta G$ value), which is calculated from the unfolding Gibbs free energy value of the mutated protein minus the unfolding Gibbs free energy value of the native type (kcal/mol). I-Mutant2.0 discriminates whether a mutation increases or decreases the protein stability with either a positive or negative $\Delta\Delta G$ value, with a reliability index score from 1 to 9.

#### *4.2.8.3 Potential splicing aberrations*

All intronic and coding mutations were analyzed for their potential effect on splicing (pre-mRNA splicing) using six different splice site prediction algorithms: MaxEntScan (MES), Gene Splicer (GS), Human Splicing Finder (HSF), Splice Site Finder (SSF), splice site prediction by Neural Network (NNSplice, version 0.9), and NetGene2.

MES, GS, HSF, SSF and NNSplice were interrogated simultaneously using the integrated software interface Alamut version 1.5 (Interactive Biosoftware; http://www.interactive-biosoftware.com). Default thresholds were used for all the analyses. These software tools are built upon various types of probabilistic models of human splice sites.

MES (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html) is based on the "maximum entropy principle" and generalises probabilistic models of sequence motifs such as weight matrix models and inhomogeneous Markov models (Yeo and Burge 2004). It assigns a log-odds ratio (MaxEntScan score) to a 9-mer at the 5′ splice site and a 23-mer at the 3′ splice site. The higher the MaxEntScan score, the stronger is the splice site. Use of MaxEntScan requires that the exon–intron boundary is provided, meaning that it is unable to read through a sequence and find the splice site.

GS (http://www.tigr.org/tdb/GeneSplicer/gene_spl.html) uses a combination of Markov modelling techniques and maximal dependence decomposition (MDD) (Pertea *et al.,* 2001), a decision-tree approach that accentuates the strongest dependencies in the early branches of the tree.

HSF (http://splice.uwo.ca) analysis is based on information theory, using weight matrix derived from the nucleotide frequencies at each position of a splice site sequence database (Rogan *et al.,* 1998; Nalla and Rogan 2005).

SSF (http://violin.genet.sickkids.on.ca/ali/splicesitefinder.html) (Shapiro and Senapathy 1987; Senapathy *et al.,* 1990), NNSplice (http://www.fruitfly.org/seq_tools/splice.html) (Reese *et al.,* 1997) and NetGene2 (http://www.cbs.dtu.dk/services/NetGene2) (Brunak *et al.,* 1991; Hebsgaard *et al.,* 1996) are computational classifiers that identify the most probable splice site in a given sequence, and then assign a score of splice site strength. All programs provide similar type of data output, that is, quantitative scores for wild-type and mutant splice site sequences that reflect splice site strength.

The score of the mutant splice site should be at least 10% lower than the score of the corresponding natural splice site for the mutation to be considered deleterious to normal splicing and conclusions should be based upon consistency of the results from at least two software programs.

## 4.3 RESULTS

### 4.3.1 Patient population

A total of 65 warfarin resistant patients requiring daily warfarin doses of ≥ 10 mg were recruited. 29 (45%) were recruited prospectively and 36 (55%) were recruited retrospectively. Their demographic details are summarised in Table 4.3. Patients were all Caucasians (100%), mostly male (58%), mean age of 54 years with a range of 26-78 years. The most common indication for warfarin was pulmonary embolism (38.5%), followed by atrial fibrillation (30%) and deep vein thrombosis (15%). Cardiovascular disease was the most common underlying co-morbidity in the majority of patients (60%).

The target INR range for most patients (n = 64) was 2-3, with the remaining 1 patient suffering from antithrombin III deficiency having a target INR range of 1.5-2.5. The majority of patients (n = 56) achieved therapeutic INR while 6 patients (9%) failed to achieve therapeutic anticoagulation with high doses of warfarin.

**Table 4.3. Clinical profile of 65 warfarin resistant patients.**

| Characteristic | N (%) |
|---|---|
| **Gender - Male** | 38 (58) |
| **Age, mean (range)** | 54 (26-78) |
| **BMI[1], mean (range)** | 31.4 (19-61) |
| **Ethnicity** | |
| White | 65 (100) |
| **Indication for warfarin** | |
| Pulmonary Embolism | 25 (38.5) |
| Atrial Fibrillation | 21 (30.4) |
| Deep Vein Thrombosis | 10 (15.4) |
| Antiphospholipid Syndrome | 3 (4.3) |
| Heart Valve Replacement | 2 (2.9) |
| Cardiomyopathy | 1 (1.4) |
| Heart Failure | 1 (1.4) |
| Other[2] | 2 (2.9) |
| **Co-morbidity** | |
| Cardiovascular disease | 39 (60.0) |
| Respiratory disease | 35 (53.8) |
| Musculoskeletal problems | 33 (50.8) |
| Urological conditions | 21 (32.3) |
| Gastrointestinal disease | 15 (23.1) |
| Neurological disease | 12 (17.4) |
| Renal disease | 9 (13.0) |
| History of falls | 9 (13.0) |
| Hepatic Disease | 7 (10.8) |

BMI: Body Mass Index.

[1] BMI missing for 3 patients.

[2] Other indications include apical mural thrombus and antithrombin III deficiency.

### 4.3.2 Reference population and warfarin assay

In our reference population consisting of 145 patients stably anticoagulated with warfarin, the median plasma total warfarin concentration was 1.27 (± 0.63) μg/ml and the 95% confidence interval in this population, 0.62 – 2.29 μg/ml, was defined as the therapeutic range for plasma warfarin (Figure 4.4). This reference range is similar to that reported by Harrington *et al* (2005) which was 0.7 – 2.3 μg/ml.

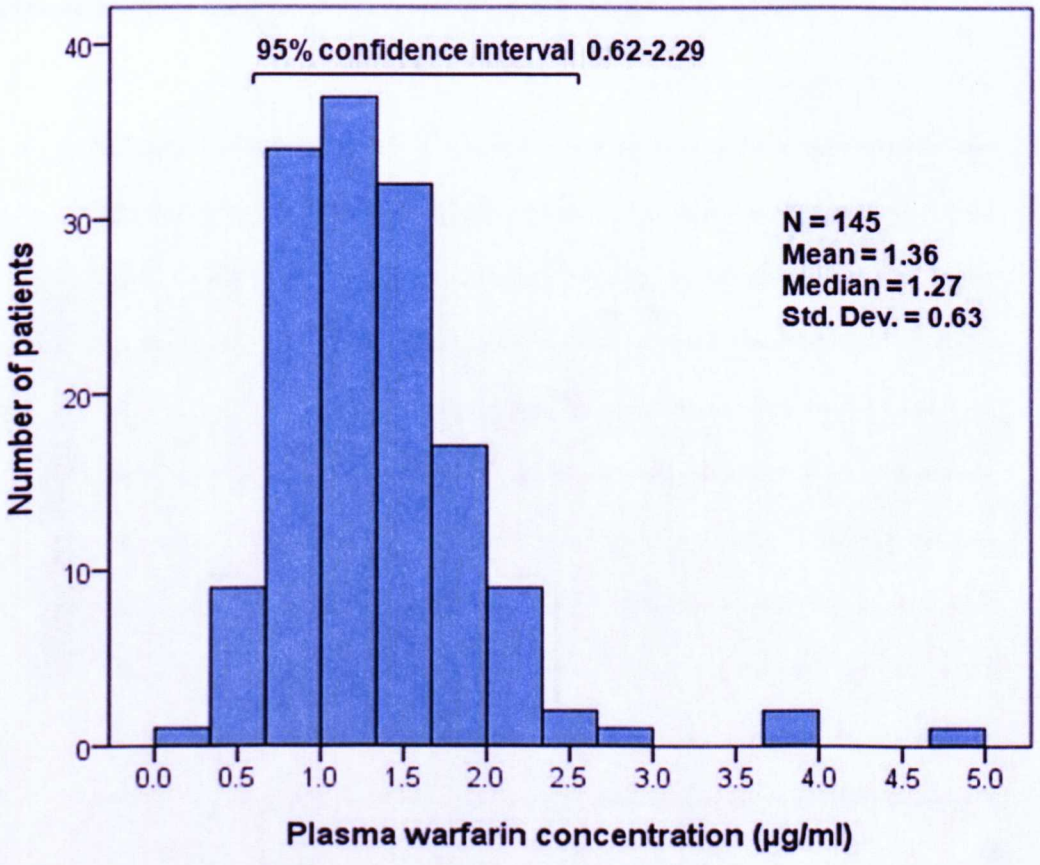



**Figure 4.4. Plasma warfarin concentrations of the reference population.** N: number; Std Dev: standard deviation.

### 4.3.3 Pharmacodynamic warfarin resistant patients

Among the 29 prospectively recruited warfarin resistant patients in our study population, no plasma samples were available for warfarin level measurement in 4 patients. Of the remaining 25 patients, 9 subjects had warfarin concentrations within the therapeutic range (0.62 – 2.29 µg/ml) and 16 patients had plasma warfarin concentrations higher than the therapeutic reference range, > 2.29 µg/ml, and were classified as showing pharmacodynamic resistance (Figure 4.5). The mean prescribed warfarin dose for these 16 patients was 11.8 mg/day (range 10 - 19) and the mean plasma warfarin concentration was 3.18 µg/ml (range 2.35 - 5.66).

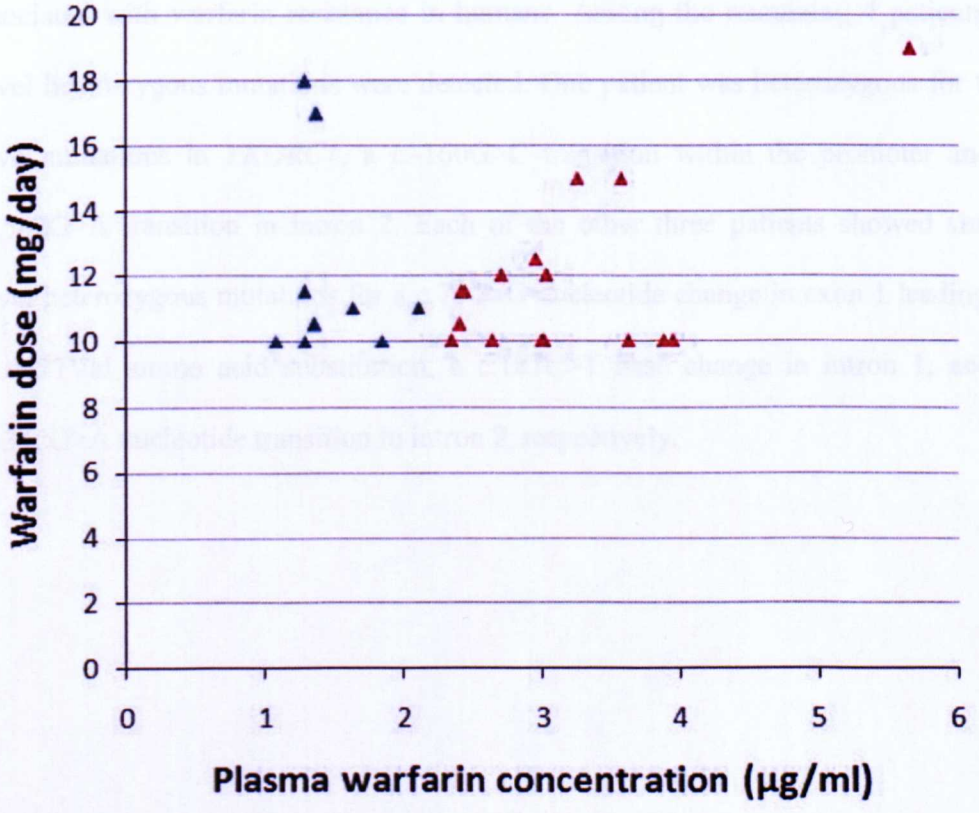



**Figure 4.5. Scattergram showing the plasma warfarin levels in 25 patients with clinical warfarin resistance.** 9 patients had warfarin levels within the therapeutic reference range (represented by blue triangles). 16 patients had warfarin levels greater than the upper therapeutic reference range of 2.29 µg/ml and were classed as showing pharmacodynamic warfarin resistance (represented by red triangles).

### 4.3.4 Novel *VKORC1* mutations identified

As depicted in Figure 4.6, sequencing analysis of the promoter, coding regions and exon-intron junctions of the *VKORC1* gene revealed 3 SNPs and 7 heterozygous mutations in 54 of the 65 patients initially diagnosed as warfarin resistant. The remaining 9 (14%) patients showed wild-type *VKORC1* sequence.

The 7 rare nucleotide substitutions in the *VKORC1* sequence were identified in 8 of the 54 patients, as shown in Table 4.4. c.36G>A (p.Arg12Arg, rs55894764) and c.129C>T (p.Cys43Cys, rs61742233) are both synonymous mutations located in exon 1 of the *VKORC1* gene, leading to no amino acid substitution. Both of these silent mutations were each exhibited in 2 patients and have never been reported to be associated with warfarin resistance in humans. Among the remaining 4 patients, 5 novel heterozygous mutations were detected. One patient was heterozygous for two novel mutations in *VKORC1*, a c.-160G>C transition within the promoter and a c.1542G>A transition in intron 2. Each of the other three patients showed single novel heterozygous mutations for a c.79C>G nucleotide change in exon 1 leading to p.Leu27Val amino acid substitution, a c.181C>T base change in intron 1, and a c.3342G>A nucleotide transition in intron 2, respectively.

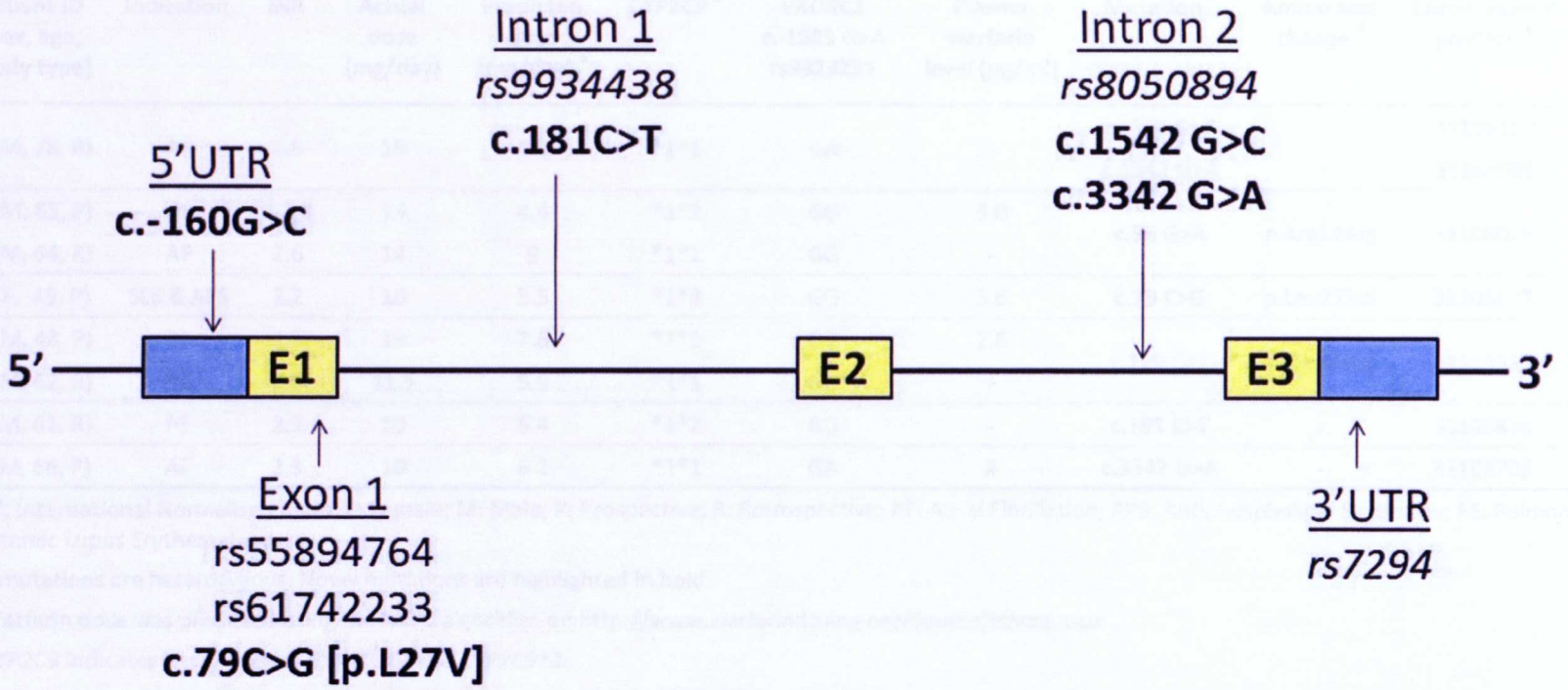



**Figure 4.6. Variants identified in the *VKORC1* gene in 65 warfarin resistant patients via sequencing.** The variant position is relative to the ATG translation start site on exon 1. Novel mutations are highlighted in bold. Common SNPs are in italics.

**Table 4.4. Characteristics of the 8 warfarin resistant patients and details of the *VKORC1* mutations identified.**

| Patient ID (sex, age, study type) | Indication | INR | Actual dose (mg/day) | Predicted dose (mg/day) [a] | *CYP2C9* [b] | *VKORC1* c.-1693 G>A rs9923231 | Plasma warfarin level (µg/ml) | Mutation [c] | Amino acid change [d] | Chromosomal position [e] | Region | dbSNP rs number |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 (M, 78, R) | AF | 2.6 | 15 | 4.3 | *1*1 | GA | - | **c.-160 G>C** | - | **31106210** | **Promoter** | **Novel** |
| | | | | | | | | **c.1542 G>A** | - | **31104509** | **Intron 2** | **Novel** |
| 2 (M, 65, P) | AF | 1.2 ‡ | 15 | 4.4 | *1*2 | GG | 3.0 | c.36 G>A | p.Arg12Arg | 31106015 | Exon 1 | rs55894764 |
| 3 (M, 64, R) | AF | 2.6 | 14 | 9 | *1*1 | GG | - | | | | | |
| 4 (F, 49, P) | SLE & APS | 2.2 | 10 | 5.5 | *1*3 | GG | 3.6 | **c.79 C>G** | **p.Leu27Val** | **31105972** | **Exon 1** | **Novel** |
| 5 (M, 48, P) | PE | 2.2 | 13 | 7.6 | *1*1 | GG | 2.6 | c.129 C>T | p.Cys43Cys | 31105922 | Exon 1 | rs61742233 |
| 6 (M, 62, R) | PE | 2.3 | 11.5 | 5.9 | *1*1 | GG | - | | | | | |
| 7 (M, 61, R) | PE | 2.3 | 10 | 5.4 | *1*2 | GG | - | **c.181 C>T** | - | **31105870** | **Intron 1** | **Novel** |
| 8 (M, 66, P) | AF | 2.3 | 10 | 6.2 | *1*1 | GA | # | **c.3342 G>A** | - | **31102709** | **Intron 2** | **Novel** |

INR: International Normalised Ratio; F: Female; M: Male; P: Prospective; R: Retrospective; AF: Atrial Fibrillation; APS: Antiphospholipid Syndrome; PE: Pulmonary Embolism; SLE: Systemic Lupus Erythematosus.

All mutations are heterozygous. Novel mutations are highlighted in bold.

[a] Warfarin dose was predicted using the IWPC algorithm on http://www.warfarindosing.org/Source/Home.aspx

[b] *CYP2C9* indicates haplotypes of *CYP2C9*2* and *CYP2C9*3*.

[c] Nucleotide numbering is according to reference sequence NM_024006 with A of the ATG initiation codon designated as +1.

[d] Amino acid numbering is according to reference sequence NP_076869 with A of the ATG initiation codon designated as +1.

[e] Chromosomal positions are given in base pairs from the p-telomere of chromosome 16, as per *VKORC1* reference assembly contig NT_010393.16, NCBI build 37 assembly.

‡ Stable anticoagulation was not achieved.

# No plasma sample available for warfarin concentration measurement.

### 4.3.5 Characteristics of patients carrying *VKORC1* mutations

Phenotypic characteristics of the 8 *VKORC1* mutation patients are detailed in Table 4.4. Amongst the 8 patients, 3 showed pharmacodynamic resistance to warfarin with plasma warfarin levels between 2.6 to 3.6 μg/ml. Of the remaining 5 patients, 1 was a prospectively recruited patient but a plasma sample was not available for warfarin level measurement. The other 4 subjects were retrospectively recruited with no plasma sample collected.

Therapeutic INR was achieved in 7 out of the 8 patients carrying rare *VKORC1* nucleotide substitutions with increased warfarin dosages, whereas one subject (patient 2) experienced sub-therapeutic anticoagulation at a high warfarin dose of 15 mg/day for over 3 weeks. This patient later passed away due to heart failure; it is therefore not known whether this patient would have achieved stable anticoagulation with higher doses of warfarin. No patients experienced thrombosis, bleeding or other adverse events during warfarin therapy.

Genotyping of the 3 SNPs known to be responsible for warfarin sensitivity (*CYP2C9*2*, *CYP2C9*3* and *VKORC1* c.-1639) in these 8 warfarin resistant patients indicated predominance of wild-type *CYP2C9* haplotypes (5 of 8 patients) and *VKORC1* c.-1639 genotype (6 of 8 patients). Only 2 patients were heterozygous for *CYP2C9*2*, 1 heterozygote for *CYP2C9*3* and 2 were heterozygous carriers for *VKORC1* c.-1639. Incorporating both genetic (*CYP2C9* and *VKORC1* SNPs) and clinical factors (gender, weight, height, co-morbidity and co-medications), the warfarin doses of these 8 warfarin resistant patients were predicted using the pharmacogenetics dosing algorithm developed by the International Warfarin Pharmacogenetics Consortium (IWPC) (Klein *et al.*, 2009). Similar to previous findings, a large difference between the actual and predicted warfarin dose (mean =

6.3 ±2.8 mg/day) was observed, indicating that warfarin resistant patients are less likely to achieve therapeutic INR if their warfarin doses were predicted using the pharmacogenetics algorithm.

### 4.3.6 Putative transcription factor binding site, protein and splice site analyses

*In silico* predictive analysis of the novel mutations was performed using established databases and bioinformatics tools. Sequences surrounding the promoter c.-160G>C and intronic c.181C>T, c.1542G>A and c.3342G>A mutations were searched for potential transcription binding sites. The TESS program revealed that the mutation of G to C allele at nucleotide position -160 created a potential binding site for Sp1 transcription factor with a log-likelihood score of 12 (score range 1 to 20). No putative predictions were found with the MOTIF program.

PolyPhen, Align-GVGD and SIFT were used to predict the effect of the amino acid substitution at position 27 from leucine to valine due to the exonic c.79C>G mutation. PolyPhen and Align-GVGD classified this mutation as "benign" indicating the mutation was predicted to have little deleterious effect on protein function or structure. SIFT, however predicted this amino acid change to have a deleterious effect on VKORC1 protein function but with low confidence.

The novel exon mutation c.79C>G, rs55894764, rs61742233, and the three novel intronic mutations c.181C>T, c.1542G>A and c.3342G>A were evaluated for possible effects on RNA splicing using six different bioinformatics programs (MES, GS, HSF, SSF, NNSplice and NetGene2) that predict changes of splice site strength or generation of new splice site, or activation of cryptic splice sites. None were predicted to induce splice site changes.

## 4.4 DISCUSSION

Our study cohort consisted of 65 clinically defined warfarin resistant patients who required daily doses of 10 mg of warfarin or more. Of these, 29 were recruited prospectively and 36 were recruited retrospectively. Figure 4.7 shows a flow diagram summarising our cohort of patients with warfarin resistance and the *VKORC1* mutations identified.

Despite exposure to large doses of warfarin, no patients experienced any discernible adverse events such as bleeding complications. Among the 29 prospectively recruited patients, pharmacodynamic resistance to warfarin was determined in 25 patients whose plasma samples were available for warfarin level measurements. 16 patients had plasma warfarin level > 2.29 μg/ml, indicating that the mechanism of warfarin resistance in these patients was a reduction in warfarin's anticoagulant effect (pharmacodynamic resistance). Warfarin resistance was observed clinically in the remaining 9 patients with plasma warfarin levels within the therapeutic range, possible reasons for this include poor compliance, high dietary vitamin K intake, diminished warfarin absorption or increased warfarin metabolism/clearance. These, however, have not been investigated systematically in our study.

Sequencing of the *VKORC1* gene revealed seven heterozygous mutations in 8 of our 65 warfarin resistant patients. None of these mutations have been previously reported to be associated with warfarin resistance in human.

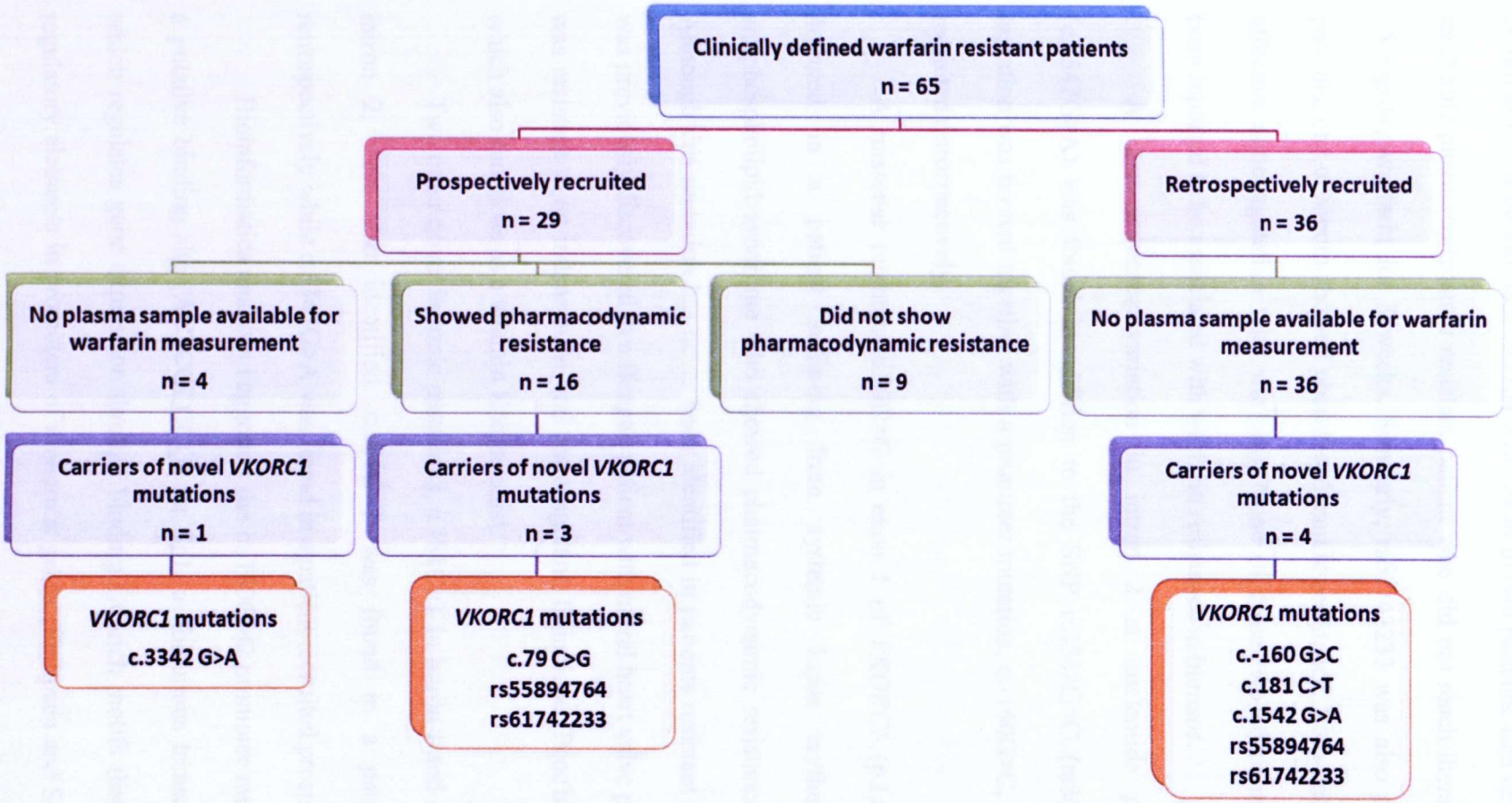



**Figure 4.7. Flow chart summarising our cohort of 65 warfarin resistant patients and the novel *VKORC1* mutations identified.**

The two silent mutations, rs55894764 and rs61742233, were the most common in our cohort. rs55894764 was found in two patients, one of whom was an evaluable pharmacodynamic resistant patient who did not reach therapeutic INR at 15 mg/day warfarin for 3 weeks. Similarly, rs61742233 was also present in two patients, one of whom showed pharmacodynamic resistance. Both patients achieved effective anticoagulation with warfarin. None of these two silent mutations have been reported to be associated with warfarin resistance in humans.

A new sequence variation in intron 2 at nucleotide position 1542 (c.1542G>A) was found in addition to the SNP c.1542G>C (rs8050894). This mutation was present together with a promoter mutation, c.-160G>C, in one patient recruited retrospectively.

A missense mutation c.79C>G in exon 1 of *VKORC1* (p.Leu27Val) was detected in a patient suffering from systemic lupus erythematosus and antiphospholipid syndrome who showed pharmacodynamic resistance to warfarin. Although this mutation has never been identified in patients resistant to warfarin, it was previously discovered in a Bengali patient with mitral heart valve prosthesis who was resistant to an indanedione oral anticoagulant, fluindione (Peoc'h *et al.,* 2009), which also functions as a vitamin K antagonist.

Two other novel intronic mutations, c.181C>T in intron 1 and c.3342G>A in intron 2, were also identified. c.181C>T was found in a patient recruited retrospectively while c.3342G>A was found in a patient recruited prospectively.

Bioinformatics analysis suggested the c.-160G>C promoter mutation creates a putative binding site (ACTCGCCC**C**) for Sp1, a ubiquitous transcription factor which regulates gene expression through binding GC-rich motifs that are common regulatory elements in promoters of numerous genes (Philipsen and Suske 1999). It

is possible that Sp1 binds to *VKORC1* promoter, leading to enhanced *VKORC1* transcription and subsequent increase in active vitamin K dependent clotting factors, leading to higher warfarin dose requirement. Nevertheless, functional work needs to be carried out to confirm this hypothesis.

*In silico* analysis predicted the p.Leu27Val amino acid substitution to be likely deleterious. This substitution is localized within the putative VKOR protein first transmembrane domain (TM1) which is approximately 100 residues away from the putative third transmembrane domain (TM3) (Tie *et al.,* 2005). Although TM3 has been postulated to contain a binding motif for the ring portion of warfarin (Ma *et al.,* 1992), other regions in the VKOR protein have been proposed to possess candidate accessory binding sites (Harrington *et al.,* 2008; Watzka *et al.,* 2011) for the various components at the 3-position (spatial residues) of warfarin which may also affect the affinity and potency of warfarin (Park 1988). One likely hypothesis therefore is that the p.Leu27Val amino acid substitution in TM1 disturbs the spatial constellation of the 3-substituent required for stable warfarin binding, leading to increased conversion of vitamin K epoxide, hence requiring a higher warfarin dose to achieve a therapeutic effect. In addition, this p.Leu27Val mutation is localized near p.Val28Leu, which has been described to cause very severe warfarin resistance in an experimental cellular model (Rost *et al.,* 2004a). On the other hand, given the enhanced oxidative status seen in antiphospholipid syndrome (Iuliano *et al.,* 1997; Ames *et al.,* 1998), a recent report by Ames and colleagues suggests that this oxidation might increase the generation of γ–carboxylated coagulation proteins and consequently more warfarin will be required to counteract the γ–carboxylation process (Ames *et al.,* 2011) in this patient.

Mutations in *VKORC1* may be an important, though rare, cause of warfarin resistance. Further investigation is required to understand the underlying functions of these *VKORC1* mutations in relation to warfarin resistance. Of note, among the 16 patients with pharmacodynamic warfarin resistance, only 3 showed mutations in *VKORC1*, suggesting that other genes may also play a role in warfarin resistance.

Pharmacogenetic algorithms incorporating *CYP2C9* and *VKORC1* polymorphisms are not accurate enough to predict warfarin dose requirements in warfarin resistant patients. The identification of specific *VKORC1* mutations, via Sanger sequencing, next generation sequencing, or third generation sequencing technologies, could potentially lead to successful clinical management and treatment of warfarin resistant patients. Our patient population and other studies published in the literature thus far are underpowered to evidently support this hypothesis. Recruitment of more warfarin resistant patients is currently underway and genotyping of the whole human exome will be carried out in an attempt to identify rare causal mutations of warfarin resistance as other genes could also be responsible for warfarin resistance.

# Chapter 5

# Investigation of the functional consequences of novel *VKORC1* mutations identified

# CONTENTS

**5.1 INTRODUCTION** ..... **148**

**5.2 MATERIALS AND METHODS** ..... **151**

5.2.1 Synthesis of *in vitro* Sp1 protein ..... 151

5.2.2 Bradford Assay ..... 151

5.2.3 Western Blotting ..... 151

5.2.4 EMSA ..... 153

5.2.5 Construction of plasmids ..... 155

*5.2.5.1 Promoter-luciferase reporter plasmid* ..... *155*

*5.2.5.2 Site-directed mutated plasmid* ..... *158*

5.2.6 Sequencing ..... 159

5.2.7 Cell culture ..... 160

5.2.8 Transient transfection of HepG2 cells ..... 160

*5.2.8.1 pGL3 constructs* ..... *160*

*5.2.8.2 pCMV6-XL4 constructs* ..... *161*

5.2.9 Dual-luciferase reporter assay ..... 161

5.2.10 MTT cytotoxicity assay ..... 162

5.2.11 Cell treatment with warfarin and PIVKA-II ELISA ..... 163

5.2.12 Quantitative analysis of DNA methylation ..... 164

5.2.13 Statistical Analysis ..... 169

*5.2.13.1 Assessing association with different genotype groups* ..... *169*

*5.2.13.2 Assessing conformity with Hardy-Weinberg Equilibrium* ..... *169*

*5.2.13.3 Assessing association with methylation levels* ..... *169*

**5.3 RESULTS** ..... **170**

5.3.1 Binding of Sp1 to the *VKORC1* promoter -160 region ..... 170

5.3.2 Effect of *VKORC1* c.-160G>C variant on promoter activity using reporter gene assay . 173

5.3.3 Effect of exonic mutation on PIVKA-II levels in HepG2 cells ..... 175

5.3.4 Effect of *VKORC1* genetic variants on genomic DNA-CpG methylation ..... 177

**5.4 DISCUSSION** ..... **179**

## 5.1 INTRODUCTION

Recent studies of the enzymatic characteristics of vitamin K epoxide reductase in rats and mice have provided insights into some important differences in the function of wild-type and warfarin resistant mutant proteins. Lasseur and colleagues reported diminished VKOR enzymatic activity for warfarin-resistant rats in the presence of the *VKORC1* Tyr139Phe mutation (Lasseur *et al.,* 2005; Rost *et al.,* 2005), and for mice with complete warfarin resistance caused by *VKORC1* p.Trp59Gly mutation (Lasseur *et al.,* 2006). Reduction in VKORC1 kinetic parameters and mRNA levels were also observed in rats carrying the mutation Tyr139Phe (Lasseur *et al.,* 2005).

Using site-directed mutagenesis and heterologous expression of mutant VKORC1 proteins in an embryonic kidney cell line, HEK293 cells, four *VKORC1* mutations namely p.Val29Leu (rs104894539), p.Val45Ala (rs104894540), p.Arg58Gly (rs104894541), and p.Leu128Arg (rs104894542), which caused warfarin resistance in humans have been reported to show reduced VKOR activity (Rost *et al.,* 2004a).

In chapter 4, bioinformatics analyses predicted that the novel *VKORC1* promoter c.-160G>C mutation creates a putative Sp1 transcription binding site, and the exon 1 c.79C>G mutation causes a likely deleterious p.Leu27Val amino acid substitution. Bioinformatics programs however only serve as a prediction tool. The data obtained *in silico* need to be interpreted with caution as they may not accurately reflect whether the sequence variation found in a patient is or is not responsible for the disease (Tchernitchko *et al.,* 2004). As these two mutations have never been reported to be associated with warfarin resistance, the aim of this chapter was to confirm the *in silico* predictions via *in vitro* functional studies, in an attempt to

understand the underlying functional activities of the enzyme in the presence of these mutations.

To gain an insight into the molecular effect of the promoter c.-160G>C mutation on Sp1 protein DNA binding, electrophoretic mobility shift assay (EMSA) was carried out. Furthermore, the effect of the promoter c.-160G>C mutation on *VKORC1* promoter activity was also investigated using luciferase reporter gene assay.

The inhibition of *VKORC1* by warfarin prevents vitamin K recycling (see Figure 1.2), resulting in an accumulation of inactive vitamin K-dependent coagulation factors, which are also known as proteins induced by vitamin K antagonism/absence (PIVKA). Utilising an enzyme-linked immunosorbent assay (ELISA) which uses the inactive precursor of coagulation factor II (PIVKA-II) or des-gamma-carboxy prothrombin (DCP) as a biomarker for the anticoagulation effect of warfarin, the functional effect of exon 1 c.79C>G mutation on PIVKA-II accumulation was evaluated.

Epigenetics refers to the reversible regulation of various genomic functions mediated through partially stable modifications of DNA and chromatin histones, which are essential for normal cellular development and differentiation (Docherty *et al.,* 2010). Of particular interest is DNA methylation occurring at cytosine-phosphate-guanine (CpG) sites, where a methyl group is attached to a cytosine base directly followed by a guanine in the DNA sequence. This is an inheritable DNA modification that does not alter the nucleotide sequence (Holliday and Pugh 1975; Bird 1986). It has been estimated that 80% of CpG sites in the genome are methylated, especially within non-coding DNA, repetitive sequences and potentially active transposable elements, resulting in long-term silencing (Bird 1986; Cross and

Bird 1995; Jones 1999; Jaenisch and Bird 2003). The CpG dinucleotides tend to cluster in regions called CpG islands, defined as regions of more than 200 bases with a G+C content of at least 50% (Portela and Esteller 2010). Approximately 60% of human gene promoters are associated with CpG islands and are usually unmethylated in normal cells, although some (~6%) become methylated in a tissue-specific manner during developmental processes (Straussman *et al.,* 2009). Aberrant methylation signatures have been implicated in a growing number of human pathologies (Robertson and Wolffe 2000; Hatchwell and Greally 2007), including cancer (Jones and Baylin 2007), imprinting disorders (Feinberg 2007), and even complex neuropsychiatric phenotypes such as schizophrenia and bipolar disorder (Mill *et al.,* 2008). Several studies have suggested that variation in DNA methylation levels can significantly contribute to variation in gene expression (Jaenisch and Bird 2003; Manolio *et al.,* 2009), which may in turn affect drug response (Esteller *et al.,* 2002; Shen *et al.,* 2007). The effects of *VKORC1* genetic variants on DNA methylation at specific CpG sites across the *VKORC1* CpG island were therefore also explored in this chapter.

## 5.2 MATERIALS AND METHODS

### 5.2.1 Synthesis of *in vitro* Sp1 protein

The cDNA clone of *Homo sapiens* Specificity Protein 1 (Sp1) was purchased from OriGene Technologies, Rockville, MD, USA (catalogue no. SC101137) which encoded the full-length Sp1 (reference sequence: NM_138473) cloned into the pCMV6-XL6 vector. The *Sp1* gene was *in vitro* transcribed and translated using the TnT® SP6 Quick Coupled Transcription/Translation Systems (Promega, Southampton, UK). The reaction contained 40 µl of Quick TnT® SP6 master mix, 20 µM of methionine, and 1 µg of Sp1 plasmid in a 50 µl reaction. The mixture was incubated at 30°C for 90 min.

### 5.2.2 Bradford Assay

Estimation of *in vitro* transcribed/translated Sp1 protein concentration was performed using the Bradford Reagent (Sigma-Aldrich). The Bradford assay has been previously described in section 3.2.7. Absorbance was measured using the DTX 880 Multimode Detector (Beckman Coulter®, UK) at 570 nm. A standard curve, ranging from 0.0125 – 0.25µg bovine serum albumin (BSA) was used to calculate the protein content.

### 5.2.3 Western Blotting

To check that Sp1 has been successfully transcribed and translated, western blot analysis of the *in vitro* transcribed/translated Sp1 protein was carried out. 5 µl of the *in vitro* Sp1 protein was added to 5 µl laemmli buffer (Sigma-Aldrich) and denatured by incubating at 99°C for 5min. Alongside a molecular weight marker, the

PrecisionPlus Protein Kaleidoscope Standards (Bio-Rad Laboratories Ltd., Hertfordshire, UK), the denatured sample underwent sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) on an 8% SDS-polyacrylamide gel under reducing conditions in running buffer (25 mM Tris, pH 8.3, 192 mM glycine, and 0.1% SDS) at 90 V for 10 min, followed by 60 min at 170 V. Separated proteins were transferred to Hybond nitrocellulose membrane using Bio-Rad Mini Trans-Blot cell, in transfer buffer (25 mM Tris, pH 8.3, 192 mM glycine, with 20% methanol and 0.1% SDS) at 100 V for 60 min. To determine that the transfer process was successful, the membrane was stained with Ponceau S solution for 10 s. The membrane was then blocked at room temperature for 1 h with 5% skimmed milk (Sigma-Aldrich) in tris-buffered saline (TBS; 0.15 M NaCl, 20 mM Trizma base, pH 7.6) containing 0.1% Tween 20. After 3 washes, the membrane was probed with the primary antibody, rabbit anti-human Sp1 (PEP2-X, Santa Cruz Biotechnology, Santa Cruz, CA, USA), at 1 : 5000 dilution in TBS-Tween containing 0.5% skimmed milk, overnight at 4°C. After 3 washes, the membrane was incubated with the secondary antibody horseradish peroxidise-conjugated (HRP) goat anti-rabbit IgG (Amersham Biosciences) in 1: 50 000 dilution (in TBS-Tween containing 0.5% skimmed milk) for 1 h at room temperature. After several washes, the Western Blot was developed using the enhanced chemiluminescence detection system (ECL) according to the instructions of the manufacturer (Amersham Biosciences), and was exposed to Hyperfilm™ ECL (Amersham Biosciences) under darkroom conditions, using a Kodak BioMax MS intensifying screen and developed using Kodak developer and fixer solutions (Sigma-Aldrich).

### 5.2.4 EMSA

The following experiments included the use of annealed oligonucleotides (Invitrogen) which represent the region between -186 and -147 bp of the human *VKORC1* promoter encompassing the wild-type and mutant nucleotide at -160 bp. A consensus Sp-1 binding sequence (Promega) was used as a positive control. These oligonucleotide sequences are summarised in Table 5.1.

End labelling of the annealed oligonucleotides with [$\gamma$-$^{32}$P]ATP (PerkinElmer, UK) and subsequent DNA-protein binding reactions were performed using the Gel Shift Assay System (Promega), following the manufacturer's instructions. The phosphorylation reaction was prepared by mixing 20% annealed probe (1.75 pmol/µl) with 10% T4 Polynucleotide Kinase 10x buffer, 10% [$\gamma$-$^{32}$P]ATP (250 µCi), 50% water and 10% T4 Polynucleotide Kinase (5-10 U/µl). After an incubation at 37°C for 10min, 10% 0.25 M EDTA was added to terminate the reaction. The radiolabelled oligonucleotide was reconstituted in 90% 1x Tris-EDTA buffer and stored at -20°C until use. The DNA-protein reaction mixture consisted of 10% [$\gamma$-$^{32}$P]ATP-labelled oligonucleotide with 5 µg of *in vitro* Sp1 protein, 20% Gel Shift 5x Binding Buffer (Promega), water and 10% poly dI.dC (0.1 µg/µl) (Sigma-Aldrich) to block unspecific binding, and was left at room temperature for 30 min. For the competition assay, the reaction mixture was incubated in the presence of 50-fold molar excess of unlabelled Sp1 consensus oligonucleotide or 50-fold, 500-fold or 1000-fold molar excess of unlabelled promoter -160 wild-type or mutant annealed oligonucleotides. For the supershift assay, the reaction mixture was incubated with 1 µl of antibody against human Sp1 (PEP2X, Santa Cruz Biotechnology) for 30 min at 4°C.

After a 5% polyacrylamide gel was pre-run for 30 min at 160V, the reactions were loaded and subjected to non-denaturing gel electrophoresis using a Hoefer SE 600 Ruby vertical electrophoresis system for 1.5 h at 230V in 0.5x Tris-Borate-EDTA buffer. The gel was then fixed in a mixture containing 10% methanol and 10% acetic acid for 30 min. The fixed gel was transferred onto a sheet of filter paper, covered with Saran wrap and vacuum-dried. Dried gels were subsequently exposed to a Storage Phosphor Screen overnight, before being scanned by a Phosphor Imager. The gels were visualised and bands were quantified using ImageQuant software

**Table 5.1. Oligonucleotides used in EMSA experiments.**

| Name | Position (relative to TSS) | Direction | Primer sequence (5' → 3') |
|---|---|---|---|
| *VKORC1* promoter -160G Wild-type | -186 to -147 bp | F | TTTTCCTAACTCGCCCGCTTGACTAGCGCC |
| | | R | GGCGCTAGTCAAGCGGGCGAGTTAGGAAAA |
| *VKORC1* promoter -160C Mutant | -186 to -147 bp | F | TTTTCCTAACTCGCCCCCTTGACTAGCGCC |
| | | R | GGCGCTAGTCAAGGGGGCGAGTTAGGAAAA |
| Sp1 consensus | None | F | ATT CGA TCG GGG CGG GGC GAG C |
| | | R | TAA GCT AGC CCC GCC CCG CTC G |

F: forward; R: reverse; TSS: translation start site.

Translation start site reference sequence: NM_024006.

The sequences of the *VKORC1* promoter probes are the same except for the nucleotide at -160 (underlined).

## 5.2.5 Construction of plasmids

### *5.2.5.1 Promoter-luciferase reporter plasmid*

Genomic DNA from patients with *VKORC1* promoter wild-type -160 GG and heterozygous -160 GC genotypes were amplified to generate a 202 bp fragment encompassing the region from -327 to -126 bp. The forward primer 5'-GCGCCGCTCGAGCCCGGCCAACAGTTTTTA-3' with a *Xho*I restriction site (underlined) and the reverse primer containing a *Hind*III restriction site (underlined) 5'-GGGCCCAAGCTTCGACCCAAATGGCTGTTC-3' were used. In a final volume of 50 µl, PCR was conducted at 95°C for 15 min, followed by 35 cycles at 94°C for 45 s, 63°C for 45 s, and 72°C for 1 min. A final extension was carried out at 72°C for 10 min. The size and specificity of the PCR products were examined in 1.5% ethidium bromide-stained TBE agarose gel under UV light.

The PCR products were purified using the QIAquick® PCR Purification Kit (QIAGEN) per manufacturer's protocol. Every 10 µl of purified PCR products was then double digested with 30 U each of *Xho*I and *Hind*III restriction enzymes (New England Biolabs, Herts, UK) in 1× NEB buffer 2 for 4 h at 37°C, followed by enzyme inactivation at 65°C for 20 min. On the other hand, each µg of the pGL3-basic vector (Promega) was first singly digested with 30 U of *Hind*III before being digested with 30 U of *Xho*I. Each digestion was incubated for 3 h at 37°C followed by enzyme inactivation at 65°C for 20 min. The pGL3-basic vector is a promoterless eukaryotic expression vector and contains the cDNA encoding for firefly luciferase which when fused with a promoter, can be used to analyze the inserted promoter activity once transfected into mammalian cells.

The digested pGL3-basic vector and the PCR products were purified (QIAquick®, QIAGEN). To prevent the linearised pGL3-basic vector from

self-ligating, dephosphorylation of the 5' phosphate end of the vector was performed using the enzyme Antarctic Phosphatase (New England Biolabs). In a 50 µl reaction, 1 µg of purified linearised vector was incubated with 5 U of Antarctic Phosphatase in 1x Antarctic Phosphatase Buffer (New England Biolabs) at 37°C for 70 min, followed by enzyme inactivation at 65°C for 5 min.

The purified PCR products were subsequently ligated between the *Hind*III and *Xho*I sites upstream of the luciferase gene in the pGL3-basic vector in a 1:3 vector:insert ratio with 20 U T4 DNA ligase (New England Biolabs) and 1x ligase reaction buffer in a final reaction volume of 20 µl, at 16°C for 16 h, followed by ligase inactivation at 65°C for 20 min.

The ligation products were analysed on a 1% ethidium bromide-stained TBE agarose gel and subsequently transformed to *E. Coli*. Transformed bacteria were plated onto agar containing 50 µg/ml ampicillin. After 16 h of incubation at 37°C, five colonies were picked from each ligation and additional colonies were analyzed if necessary. The plasmids were isolated and purified using the GenElute HP Plasmid Miniprep (Sigma-Aldrich), according to the manufacturer's protocol.

A brief depiction of the procedures described above is shown in Figure 5.1

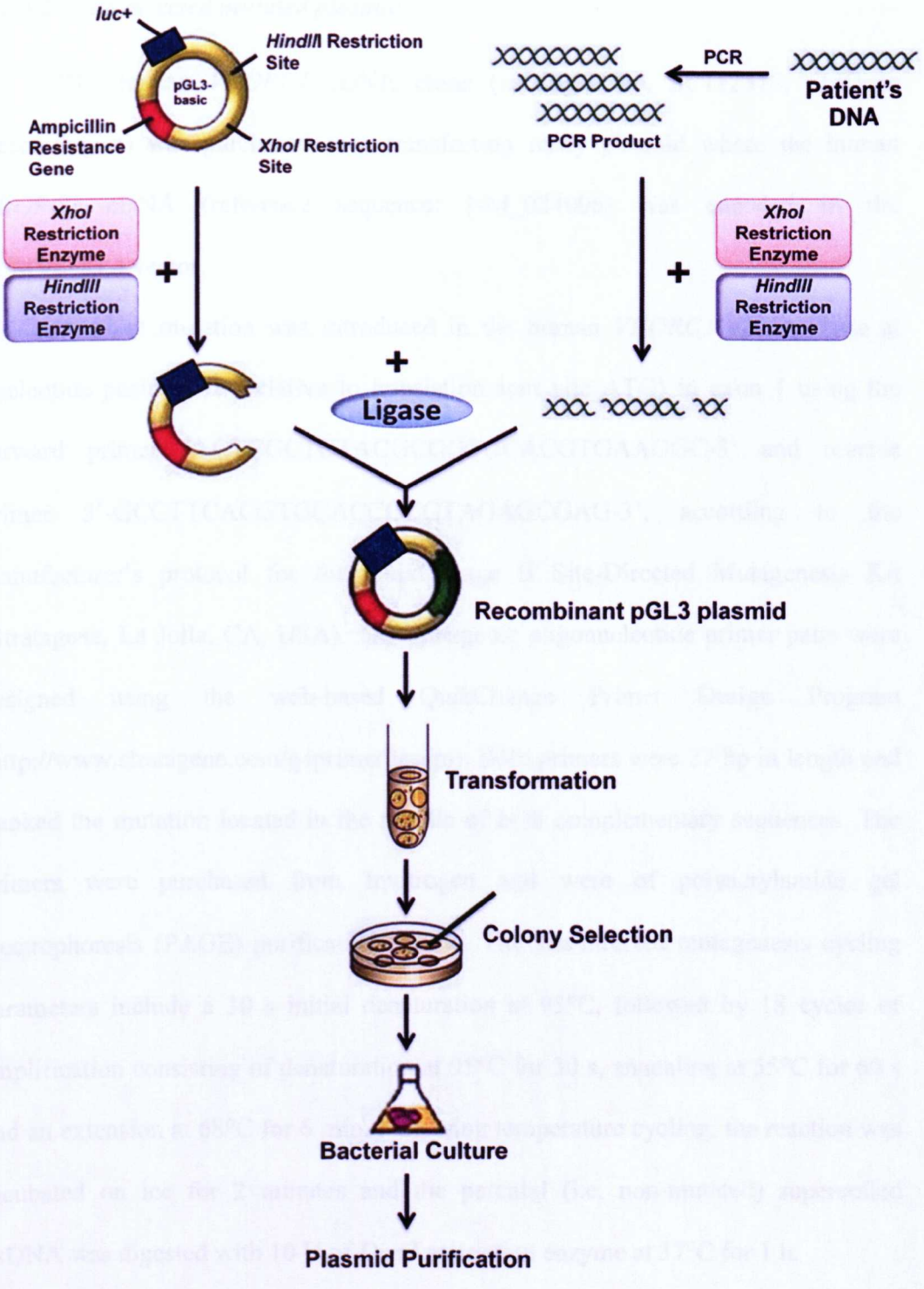



**Figure 5.1. A diagram summarising the steps involved in constructing the recombinant pGL3 reporter plasmid.**

### *5.2.5.2 Site-directed mutated plasmid*

The human *VKORC1* cDNA clone (catalogue no. SC112318; OriGene Technologies) was purchased as a transfection ready plasmid where the human *VKORC1* cDNA (reference sequence: NM_024006) was encoded in the pCMV6-XL4 vector.

A point mutation was introduced in the human *VKORC1* cDNA clone at nucleotide position 79 (relative to translation start site ATG) in exon 1 using the forward primer 5'-CTCGCTCTACGCGGTGCACGTGAAGGC-3' and reverse primer 5'-GCCTTCACGTGCACCGCGTAGAGCGAG-3', according to the manufacturer's protocol for the QuikChange II Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA). The mutagenic oligonucleotide primer pairs were designed using the web-based QuikChange Primer Design Program (http://www.stratagene.com/qcprimerdesign). Both primers were 27 bp in length and flanked the mutation located in the middle of both complementary sequences. The primers were purchased from Invitrogen and were of polyacrylamide gel electrophoresis (PAGE) purification quality. The site-directed mutagenesis cycling parameters include a 30 s initial denaturation at 95°C, followed by 18 cycles of amplification consisting of denaturation at 95°C for 30 s, annealing at 55°C for 60 s and an extension at 68°C for 6 min. Following temperature cycling, the reaction was incubated on ice for 2 minutes and the parental (i.e. non-mutated) supercoiled dsDNA was digested with 10 U of *Dpn* I restriction enzyme at 37°C for 1 h.

The *Dpn* I-treated DNA was then transformed to XL1-Blue supercompetent cells and plated onto agar containing 50 μg/ml ampicillin, 40 mg/ml X-gal solution, and 100 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). Colonies were then picked after 16 h of incubation at 37°C and the plasmids were isolated and purified

using the GenElute HP Plasmid Miniprep (Sigma-Aldrich), according to the manufacturer's protocol.

### 5.2.6 Sequencing

Sequences in the purified plasmid constructs were confirmed to be free of random mutation by Sanger sequencing. The plasmid sequencing procedures were carried out by Kimberley Jones at Liverpool Women's Hospital, Liverpool, UK, using the Applied Biosystems 3730xl DNA analyzer, as described in section 4.2.6. All cycling conditions and procedures were similar except the plasmid DNA cycle sequencing reactions were carried out in 10 µl reactions using 150 ng of purified plasmid, 100 nM forward and reverse primer, 0.4 µl BigDye® Terminator version 1.1 sequencing buffer and 3.75 µl dilution buffer. As listed in Table 5.2, the forward RVprimer3 and reverse GLprimer2 primers were used to obtain bidirectional sequence information of the cloned inserts within the pGL3-basic vector, while the forward VP1.5 and XL39 reverse primers were used to obtain bidirectional sequence information of the site-directed pCMV6-XL4 vector.

**Table 5.2. Primers for sequencing plasmid vectors.**

| Vector | Primer name | Primer sequence (5' → 3') | Primer Direction |
|---|---|---|---|
| pGL3-basic | RVprimer3 | CTAGCAAAATAGGCTGTCCC | F |
| | GLprimer2 | CTTTATGTTTTTGGCGTCTTCCA | R |
| pCMV6-XL4 | VP1.5 | GGACTTTCCAAAATGTCG | F |
| | XL39 | ATTAGGACAAGGCTGGTGGG | R |

F: forward; R: reverse; TSS: translation start site.

All primers were of standard desalted quality.

The recombinant plasmid containing the wild-type *VKORC1* promoter –160 G allele was designated pGL3-G and the vector containing the mutant *VKORC1* promoter –160 C allele was designated pGL3-C. The plasmid containing the wild-type *VKORC1* 79 C allele was designated pCMV6-XL4+79C and the plasmid containing the mutant *VKORC1* 79 G allele was designated pCMV6-XL4+79G.

### 5.2.7 Cell culture

The human hepatoma (HepG2) cell line was cultured in Dulbecco's modified Eagle medium (DMEM; Lonza, UK), supplemented with 10% foetal calf serum (Gibco), 100 U/ml Penicillin, 100 µg/ml Streptomycin and 2 mM L-Glutamine (all from Sigma-Aldrich). The cells were maintained at 37°C in a humidified incubator at 5% $CO_2$. HepG2 cells between passages 11 and 13 were utilized.

### 5.2.8 Transient transfection of HepG2 cells

24 h prior to transfection, HepG2 cells were seeded onto 1.9 $cm^2$ of the Nunclon® 24-well plate at 1 x $10^5$ cells/well. Transfection was performed using Lipofectamine™ 2000 (Invitrogen) according to the manufacturer's protocol.

#### *5.2.8.1 pGL3 constructs*

For each well of cells, 2 µl of Lipofectamine™ 2000 was combined with 48 µl of Opti-MEM® I reduced serum (Invitrogen) in a 0.5 ml eppendorf and incubated at room temperature for no longer than 25 min. In another 0.5 ml eppendorf, 25 ng of the pRL-TK vector and 750 ng of pGL3-G, pGL3-C or pGL3-basic (negative control) were mixed with 50 µl of Lipofectamine™ 2000. The

pRL-TK vector encodes for the *Renilla* luciferase transcribed by a HSV-TK promoter which functions as an internal control to normalize the firefly luciferase expression. The contents from both eppendorfs were mixed and incubated at room temperature for 20 min. All of the mixture was added to the well in a dropwise manner with gentle mixing by gently rocking the plate back and forth. Cells were then incubated at 37°C with 5% $CO_2$ in a humidified incubator for 48 h.

#### *5.2.8.2 pCMV6-XL4 constructs*

Transfection procedures of pCMV6-XL4 constructs were similar to that of pGL3 constructs as described in section 5.2.8.1 except 800 ng of either pCMV6-XL4+79G or pCMV6-XL4+79C plasmid DNA was mixed with 50 µl of Lipofectamine™ 2000.

### 5.2.9 Dual-luciferase reporter assay

The firefly and *Renilla* luciferase activity was measured with the Dual-Luciferase® Reporter Assay System (Promega) according to the manufacturer's recommendation using a DTX 880 Multimode Detector (Beckman Coulter®). 48 h after transfection, growth medium was removed from the cultured cells and to each well, 500 µl of phosphate buffered saline (PBS) was applied gently to wash the cells from the surface of the culture vessel. After three washes were performed, 100 µl of 1x passive lysis buffer was dispensed into each well to completely cover the cell monolayer. This was followed by 30 min incubation at room temperature on an orbital shaker with gentle rocking. To remove cell debris and nuclei, cell lysates were centrifuged at 14,800 rpm, 4°C for 30 s, and the supernatant was retained and kept at -80°C until use. To a well in a 96-well LUMITRAC 200 plate (Greiner Bio-One

GmbH, Germany), 100 µl of Luciferase Assay Reagent II (LAR II) was added followed by 20 µl of the cell lysate with gentle mixing and the activity of firefly luciferase was determined. Then, 100 µl of Stop & Glo Reagent was added into the well to measure the activity of Renilla luciferase in pRL-TK. Firefly luciferase activity was normalised relative to the *Renilla* luciferase activity for each transfection.

For each construct, 2 clones were selected for transfection. A total of 3 independent transfections and triplicate luciferase assays for each transfection were performed for each construct. Because there were no differences between different clones of the same constructs, the results from the 2 clones were combined.

### 5.2.10 MTT cytotoxicity assay

The cytotoxic concentration of racemic warfarin (Sigma-Aldrich) was determined using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (Sigma-Aldrich) according to the manufacturer's instructions. HepG2 cells were seeded at ~ 20, 000 cells/well in a 96-well plate. 100% dimethyl sulphoxide (DMSO) was used as the solvent for warfarin. Final DMSO concentration in each well was < 0.03%. The cells were incubated with warfarin at concentrations of 3 µM, 30 µM, 50 µM, 100 µM, and 500 µM in triplicates, at 37°C in an incubator. Cells without warfarin treatment were included as control. After 24 h of treatment, the medium was removed and 5 mg/ml of MTT reagent (dissolved in Hanks balanced salt solution) was added to each well and incubated at 37°C in an incubator for 4 h. Subsequently, the formazan salts were solubilised in DMSO in the dark, at room temperature overnight. Finally, absorbance values were measured at a wavelength of 570 nm (DTX 880 Multimode Detector, Beckman Coulter®). Absorbance values that

are lower than the control cells indicate a reduction in cell proliferation. The percent viability was calculated as follows: (absorbance of warfarin-treated sample/ absorbance of control) × 100. Our data (Figure 5.2) indicated that no significant cytotoxicity for warfarin was observed at concentrations below 50 μM, the two non-toxic warfarin concentrations, 3 μM and 30 μM, were therefore used in our study.

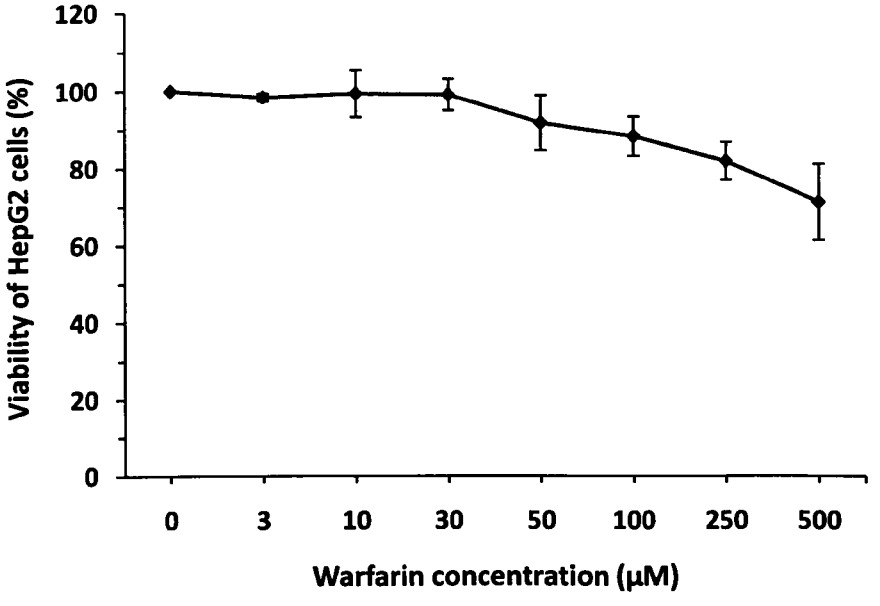



**Figure 5.2. Viability of HepG2 cells after treatment with warfarin.** Data represent mean values from three replicate wells, with the error bars representing standard deviation.

### 5.2.11 Cell treatment with warfarin and PIVKA-II ELISA

24 h after transfection of HepG2 cells with pCMV+79C or pCMV+79G, wells were washed with Hank's balanced saline solution (Sigma-Aldrich) before the addition of 500 μl of DMEM with reduced serum (containing 0.5% serum, without antibiotics) supplemented with racemic warfarin at a concentration of either 3 μM or

30 μM. A no-treatment control was included in each assay. The stock solutions and dilutions for warfarin were freshly prepared on the day of the experiment.

At the end of the 24 h treatment period, without disturbing the adherent cells, 300 μl of supernatant was collected from each culture well and stored at -80°C until use. The levels of PIVKA-II (protein induced in vitamin K absence or antagonism II) antigen in the supernatant was determined using the PIVKA-II ELISA kit (Diagnostica Stago, Asnières-sur-Seine, France) on a LT-4000 Microplate Reader (Labtech International, UK) at 492 nm. On each ELISA plate, four reference standards (2, 57, 111, and 223 ng/ml) provided with the test kit were assayed together with neat experimental samples in duplicate. PIVKA-II production and hence anticoagulant activity was considered negative if the absorbance observed was below that produced by the 2 ng/ml standard.

Two clones were transfected for each construct and each experimental condition was tested in duplicate. A total of 3 independent transfections and duplicate PIVKA-II ELISA were performed for each construct.

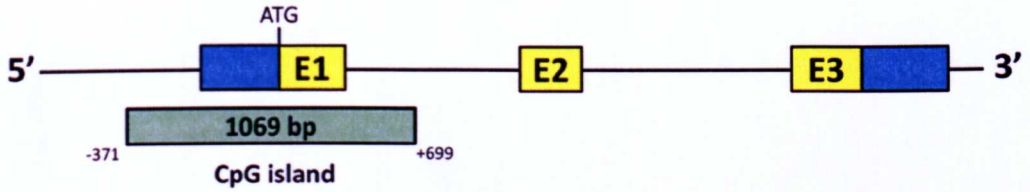
### 5.2.12 Quantitative analysis of DNA methylation

Genomic DNA extracted from peripheral blood was bisulfite converted using the EZ DNA Methylation kit (Zymo Research, CA, USA) according to the manufacturer's protocol. Bisulfite treatment of genomic DNA converts non-methylated cytosine into uracil while methylated cytosine remains unchanged. The CpG island in the *VKORC1* promoter region (+699 to –371) containing 84 CpG sites (Figure 5.3) was amplified by PCR using methylation specific primers (Table 5.3) designed using the EpiDesigner program (www.epidesigner.com). All primers were purchased from Metabion. Each forward primer was tagged with a

10mer (5'–aggaagagag–3') to balance the melting temperature, and the reverse primer had a T7-promoter tag (5'–cagtaatacgactcactatagggagaaggct–3') for *in vitro* transcription.

Quantitative DNA methylation analysis was determined using MALDI-TOF on the Sequenom platform. PCR amplification of 20 ng bisulfite-treated DNA was performed using HotStar Taq Polymerase (QIAGEN) in a 5 µl reaction volume using PCR primers at a 200 nM final concentration. The PCR was performed with a 15 min initial denaturation at 94°C, followed by 45 cycles of amplification (94°C for 20 s, 63°C for 30 s and 72°C for 1 min) and a final extension at 72°C for 3 min. After SAP treatment for 45 min at 37°C, 2 µl of the PCR product was used as a template for *in vitro* transcription with subsequent T-cleavage reaction using RNase A, at 37°C for 3 h. Both methylated and non-methylated regions are cleaved at every T which results in a mass difference of 16 Da per CpG site in the cleavage products having CpG sites. The samples were desalted and spotted on a 384 SpectroCHIP® using a nanodispenser, followed by spectral acquisition on a MassARRAY® Compact MALDI-TOF MS. All experiments were performed in triplicate with 0% and 100% methylated human DNA as control. The resultant methylation calls were performed by EpiTYPER® software version 1.0 to generate quantitative results for each CpG site or an aggregate of multiple CpG sites (Figure 5.4). Since MALDIF-TOF mass methylated peaks do not denote a particular CpG site, but rather correspond to the number of CpG sites methylated within the cleavage fragment, data were presented as average percent methylation of all CpG sites in the bisulfite PCR fragment with standard deviation (Ehrich *et al.*, 2005; Ehrich *et al.*, 2007). CpG residues with non-applicable readings were eliminated in calculation.

(a)




(b)

**chr16: 31106421-31105352, build 37/hg19**
**Position: –371 to +699, relative to ATG start site**

```
-371  5'-CTCGGCCTCCCAAAGTGCTGGGATTACAAGCGTGAGCCACCGTGCCCGGC
-321     CAACAGTTTTTAAATCTGTGGAGACTTCATTTCCCTTGATGCCTTGCAGC
                1  2  3                                4        5  6
-271     CGCGCCGACTACAACTCCCATCATGCCTGGCAGCCGCTGGGGCCGCGATT
                7     8              9                10       11
-221     CCGCACGTCCCTTACCCGCTTCACTAGTCCCGGCATTCTTCGCTGTTTTC
                      12    13             14                  15
-171     CTAACTCGCCCGCTTGACTAGCGCCCTGGAACAGCCATTTGGGTCGTGGA
                   16       17   18          19  20                21 22
-121     GTGCGAGCACGGCCGGCCAATCGCCGAGTCAGAGGGCCAGGAGGGGCGCG
                      23  24   25                26                27 28   29
 -71     GCCATTCGCCGCCCGGCCCCTGCTCCGTGGCTGGTTTTCTCCGCGGGCGC
                   30   31
 -21     CTCGGGCGGAACCTGGAGATAATGGGCAGCACCTGGGGGAGCCCTGGCTG
                32      33                  34                  35       36 37
  30     GGTGCGGCTCGCTCTTTGCCTGACGGGCTTAGTGCTCTCGCTCTACGCGC
                38          39   40 41 42     43       44          45  46 47       48
  80     TGCACGTGAAGGCGGCGCGCGCCCGGGACCGGGATTACCGCGCGCTCTGC
                49          50                  51  52 53
 130     GACGTGGGCACCGCCATCAGCTGTTCGCGCGTCTTCTCCTCCAGGTGTGC
                54                  55              56          57
 180     ACGGGAGTGGGAGGCGTGGGGCCTCGGAGCAGGGCGGCCAGGATGCCAGA
                                                58      59      60      61
 230     TGATTATTCTGGAGTCTGGGATCGGTGTGCCCGGGGAACGGACACGGGGC
                      62 63      64                                65
 280     TGGACTGCTCGCGGGGTCGTTGCACAGGGGCTGAGCTACCCAGCGATACT
                   66                 67
 330     GGTGTTCGAAATAAGAGTGCGAGGCAAGGGACCAGACAGTGCTGGGGACT
                      68          69    70                          71
 380     GGGATTATTCCGGGGACTCGCACGTGAATTGGATGCCAAGGAATAACGGT
                            72                  73                  74
 430     GACCAGGAAAGGCGGGGAGGCAGGATGGCGGTAGAGATTGACGATGGTCT
                      75  76 77                          78          79
 480     CAAGGACGGCGCGCAGGTGAAGGGGGGTGTTGGCGATGGCTGCGCCCAGG
                            80                  81          82
 530     AACAAGGTGGCCCGGTCTGGCTGTGCGTGATGGCCAGGCGTTAGCATAAT
                   83                 84
 580     GACGGAATACAGAGGAGGCGAGTGAGTGGCCAGGGAGCTGGAGATTCTGG

 630     GGTCCAGGGCAAAGATAATCTGCCCCCGACTCCCAGTCTCTGATGCAAAA

 680     CCGAGTGAACCGTTATACCA-3'
```

**Figure 5.3. Location and sequence of *VKORC1* CpG island.** (a) The *VKORC1* promoter CpG island is 1069 bp in length and is highlighted in green. (b) *VKORC1* CpG island sequence, Chr16: 31106421-31105352, build 37/hg19, as defined by the UCSC Genome Browser; position –371 to +699 relative to ATG translation start codon. ATG start site is highlighted in red font. 84 CpGs present in the region are in bold. Mutations detected in our study are underlined.

**Table 5.3. Details of bisulfite-specific PCR primers used for quantitative DNA methylation.**

| Amplicon Name | Amplicon Direction | Primer sequence (5' → 3') | Primer Direction | Position (relative to TSS) | Amplicon length | No. CpGs |
|---|---|---|---|---|---|---|
| *VKORC1* CpG A | F | aggaagagagAGTTAGGGTTTTTTTAGGTGTTGTTT | F | -295 | 323 bp | 31 |
| | | cagtaatacgactcactatagggagaaggctTCATTTCCCTTAATACCTTACAACC | R | +28 | | |
| *VKORC1* CpG B | R | aggaagagagATAATGGGTAGTATTTGGGGGAGT | F | -3 | 378 bp | 36 |
| | | cagtaatacgactcactatagggagaaggctCCCAACACTATCTAATCCCTTACCT | R | +375 | | |
| *VKORC1* CpG C | R | aggaagagagAGGTAAGGGATTAGATAGTGTTGGG | F | +351 | 302 bp | 17 |
| | | cagtaatacgactcactatagggagaaggctACAAATTATCTTTACCCTAAACCCC | R | +652 | | |

F: forward; R: reverse; TSS: translation start site.

Each forward primer was tagged with a 10mer (5'-aggaagagag-3'), and the reverse primer contained a T7-promoter tag (5'-cagtaatacgactcactatagggagaaggct-3').

All primers were of standard desalted quality.

Translation start site reference sequence: NM_024006.


**Figure 5.4. Sequenom MassARRAY analysis of percent DNA methylation in *VKORC1* methylated regions.** The EpiTYPER program from Sequenom MassARRAY analysis provides the results of percent methylation as an epigram. The epigram shows the percent DNA methylation level of each CpG site of the target region across different patients' samples. Different colours display relative methylation changes in 10% increments. The yellow circle indicates 100% methylation and the red circle shows 0% methylation, at each CpG site.

### 5.2.13 Statistical Analysis

All statistical analyses were performed using SPSS (version 18) and a *P*-value < 0.05 was considered significant.

#### *5.2.13.1 Assessing association with different genotype groups*

The non-parametric independent *t*-test (Mann-Whitney U test), was performed for comparison of levels among different genotype groups of plasmid constructs with the continuous variables, relative luciferase activity and relative PIVKA-II levels. Values were expressed as mean with standard deviation.

#### *5.2.13.2 Assessing conformity with Hardy-Weinberg Equilibrium*

Hardy-Weinberg Equilibrium (HWE) was tested for each SNP using the $\chi^2$ test. A *P*-value < 0.001 was assumed to indicate deviation from HWE.

#### *5.2.13.3 Assessing association with methylation levels*

To ensure biological relevance of the methylation differences, CpG residues with methylation signals that could not be reliably observed were excluded from subsequent analysis (Philibert *et al.,* 2010). Any methylation levels that were not normally distributed were transformed. The relationship between age and methylation levels was evaluated using Pearson correlation. The effects of gender and genotype on the level of methylation were tested using one-way analysis of variance (ANOVA). To account for multiple testing in the methylation association analyses, FDR (Benjamini *et al.,* 2001) at the 0.05 level was calculated in addition to the *P*-value for each test of association.

## 5.3 RESULTS

### 5.3.1 Binding of Sp1 to the *VKORC1* promoter -160 region

Sp1 protein was *in vitro* synthesized using rabbit reticulocyte lysate which showed a 106 kDa band on SDS-PAGE (Figure 5.5). EMSA was then performed with the *in vitro* synthesized Sp1 protein, using Sp1 consensus probes and *VKORC1* probes (Figure 5.6a). The *VKORC1* probe comprised of a 30 bp sequence flanking either the *VKORC1* promoter mutant -160C allele or the respective wild-type -160G variant. The EMSA results are shown in Figure 5.6b. As a positive control, the Sp1 protein was mixed with the Sp1 consensus probe and a protein-DNA complex was observed, as evidenced with a band shift. To unambiguously confirm the identity of the bound Sp1 protein in the protein-DNA complex, Sp1 antibody was included in the Sp1 protein-Sp1 consensus probe binding mixture. A supershift with decreased band mobility was observed as the Sp1 antibody was bound to the protein-DNA complex. Sp1 protein, however, did not form a protein-DNA complex with either the mutant or wild-type *VKORC1* promoter oligonucleotides, suggesting that the promoter mutation -160G>C did not create a specific binding site for Sp1. To investigate whether the -160G>C mutation had an effect on the archetypal binding of Sp1 protein to Sp1 consensus oligonucleotide, competition assays were performed using 50-, 500- or 1000-fold excess amounts of unlabelled mutant or wild-type double-stranded oligonucleotides. The intensity of the Sp1 protein/Sp1 DNA band shift diminished when the mixture was incubated in the presence of increasing amounts of mutant competitor but not in the presence of the wild-type competitor, indicating that the *VKORC1* promoter mutant sequence out-competes with the Sp1 consensus probe for the binding of Sp1 protein.

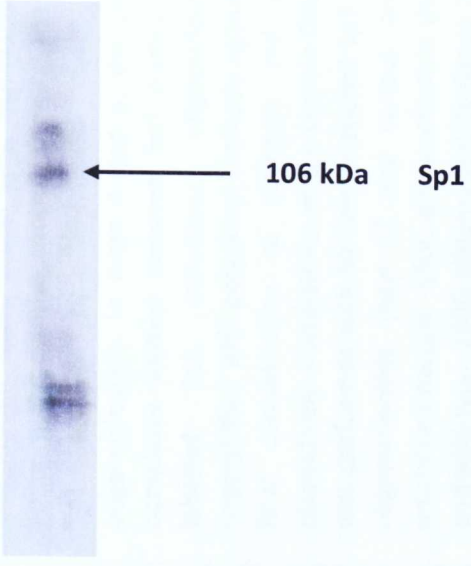



**Figure 5.5. Western blot analysis of *in vitro* synthesised Sp1 protein expression.** The *in vitro* synthesised Sp1 protein was detected by Sp1 antibody as shown by the band of approximately 106 kDa.

(a)

| | |
|---|---|
| **Sp1 consensus oligonucleotide** | 5′-ATTCGATCGGGGCGGGGCGAGC-3′ |
| ***VKORC1* promoter -160 C mutant probe** | 5′-TTTTCCTAACTCGCCCCCTTGACTAGCGCC-3′ |
| ***VKORC1* promoter -160 G wild-type probe** | 5′-TTTTCCTAACTCGCCCGCTTGACTAGCGCC-3′ |

(b)

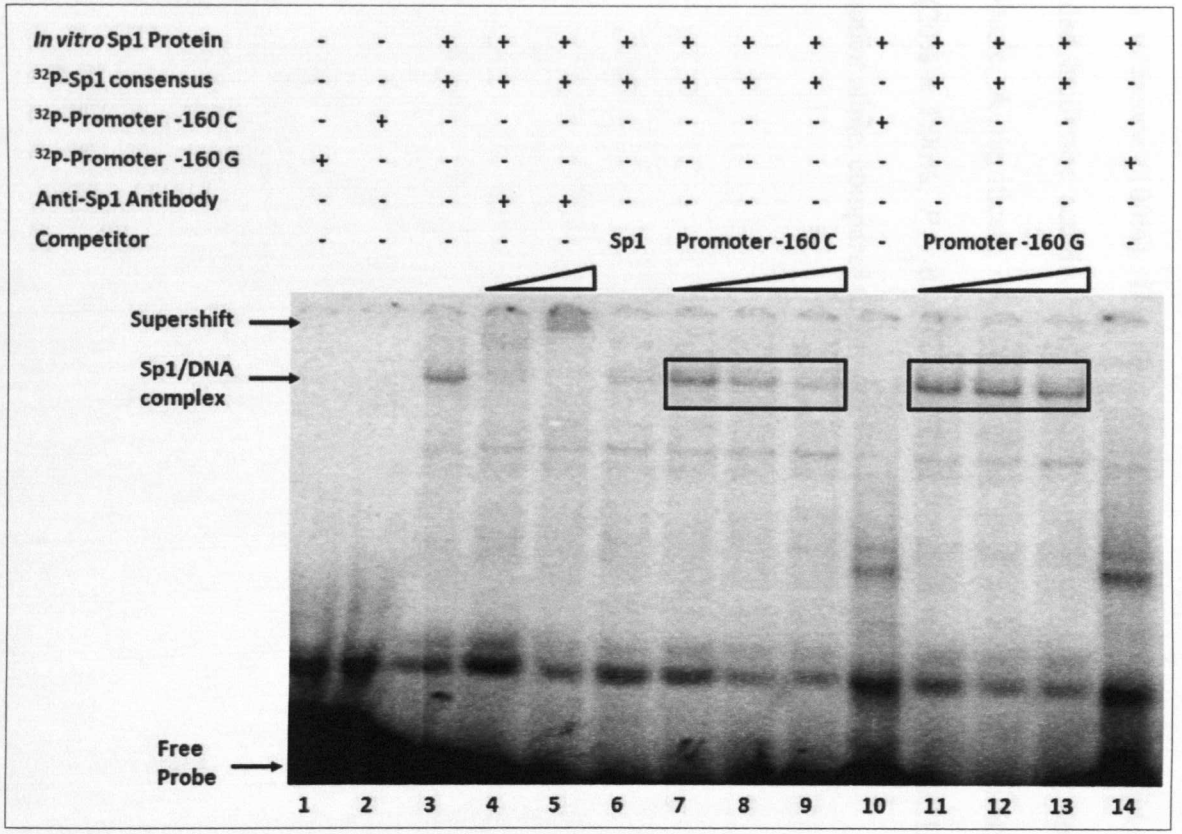



**Figure 5.6. Electrophoretic mobility shift assay of Sp1 consensus, *VKORC1* -160 wild-type and mutant oligonucleotides.** (a) DNA sequences of the Sp1 consensus and *VKORC1* probes. Sp1 binding sites are underlined. Nucleotide at -160 is in bold text and highlighted in yellow. (b) EMSA of Sp1 and *VKORC1* oligos with *in vitro* synthesised Sp1 protein. Lanes 1 and 2 contain $^{32}$P-labelled *VKORC1* wild-type -160G and mutant -160C probe only, respectively. Lane 3 contains the positive control, showing the binding of Sp1 protein to $^{32}$P-labelled Sp1 consensus oligonucleotide. In lanes 4 and 5, the protein-DNA complexes formed between Sp1 protein and 32P-labelled Sp1 consensus oligonucleotide are supershifted with polyclonal antibody against Sp1 at a final concentration of 200ng/µl and 2000ng/µl, respectively. Competition of the protein-DNA complex was performed with 50-fold unlabelled Sp1 consensus oligonucleotide (lane 6) and various amounts of unlabelled mutant -160C oligonucleotide (lanes 7 – 9) and wild-type -160G oligonucleotide (lanes 11 – 13), at 50, 500 and 1000-fold excess. The *in vitro* transcribed/translated Sp1 protein did not bind to *VKORC1* promoter mutant -160C variant (lane 10) or *VKORC1* promoter wild-type -160G variant (lane 14).

### 5.3.2 Effect of *VKORC1* c.-160G>C variant on promoter activity using reporter gene assay

To study the promoter activities of the *VKORC1* c.-160G>C allelic variants, a 202 bp fragment harbouring either the -160 G or C allele was amplified by PCR from patients' DNA and inserted upstream of the luciferase gene in the pGL3-basic vector. The transcriptional activity was analysed by measuring the luciferase expression from the lysate of transfected HepG2 cells. Results of the transient transfection assay are shown in Figure 5.7. The promoterless pGL3-basic vector with negligible amount of luciferase activity was used as the baseline control and pGL3-G luciferase activity was used as a reference (100%). The plasmid pGL3-G containing the wild-type fragment yielded luciferase activity over 100-fold higher than that of the empty vector pGL3-basic. A significant increase in mean level of luciferase activity in HepG2 cells (20% ± 1.96%; $P = 0.003$) was observed with the construct pGL3-C carrying the mutant allele, compared to the wild-type pGL3-G construct.

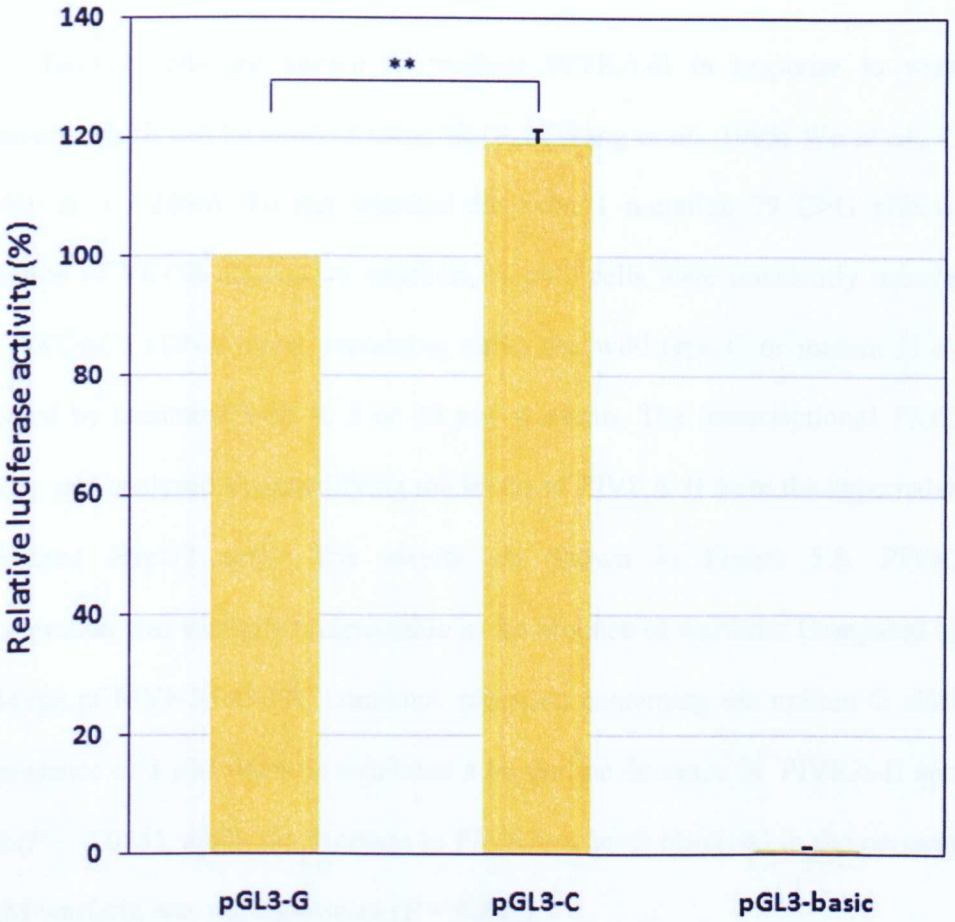



**Figure 5.7. The *VKORC1* c.-160G>C mutation increases promoter activity.** pGL3 luciferase reporter containing either the G (pGL3-G) or the C allele (pGL3-C) at the *VKORC1* promoter -160 locus was transfected into HepG2 cells and the promoter activity was measured by dual-luciferase reporter assay 48 h after transfection. The results were then expressed as relative luciferase activity (firefly luciferase activity in the pGL3 plasmid/*Renilla* luciferase activity in the pRL-TK vector) with the error bars representing standard deviation. Values represent the mean of 18 measurements (three independent transfection experiments performed in duplicate and triplicate luciferase readings were taken for each lysate sample). pGL3-G wild-type luciferase activity was used as a reference (100%). pGL3-basic was a negative control without any promoter sequence inserted. ** $P = 0.009$.

### 5.3.3 Effect of exonic mutation on PIVKA-II levels in HepG2 cells

HepG2 cells are known to produce PIVKA-II in response to warfarin treatment, which can be assayed using ELISA (Wang *et al.*, 1995; Wu *et al.*, 1996; Lawley *et al.*, 2006). To test whether the exon 1 mutation 79 C>G affects the inhibition of VKOR activity by warfarin, HepG2 cells were transiently transfected with *VKORC1* cDNA clone containing either the wild-type C or mutant G allele, followed by treatment with 0, 3 or 30 μM warfarin. The transcriptional *VKORC1* activity was analysed by quantifying the levels of PIVKA-II from the supernatant of transfected HepG2 cells. The results are shown in Figure 5.8. PIVKA-II concentration was virtually undetectable in the absence of warfarin. Compared to the wild-type pCMV6-XL4+79C construct, plasmids containing the mutant G allele in the presence of 3 μM warfarin exhibited a borderline decrease in PIVKA-II antigen level ($P = 0.058$), while the decrease in PIVKA-II level observed in the presence of 30 μM warfarin was not significant ($P = 0.247$).

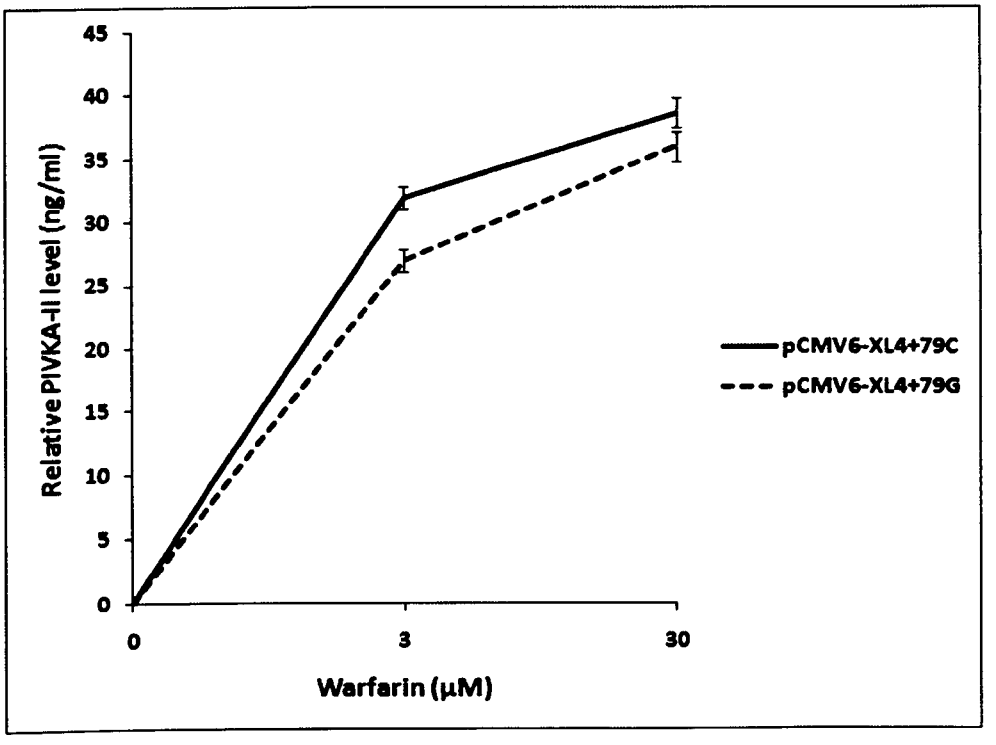



**Figure 5.8. Detection of PIVKA-II in HepG2 supernatants following treatment with serum-free medium supplemented with warfarin.** Supernatants were tested neat using the PIVKA-II ELISA. Data represent mean values from two replicate wells, with the error bars representing standard deviation.

### 5.3.4 Effect of *VKORC1* genetic variants on genomic DNA-CpG methylation

In an attempt to evaluate the effect of *VKORC1* genetic variants on DNA methylation, the extent of methylation in the *VKORC1* promoter CpG island was measured. Out of the 84 CpG sites present, 55 sites yielded distinguishable peaks on the mass array. No association was found between DNA methylation and gender nor age. Methylation levels at several CpG sites, however, showed correlation with *VKORC1* genotypes before FDR correction. The combined methylation status at CpG sites 12 and 13 (nucleotide positions -164 and -160) demonstrated borderline association with the novel intron 2 mutation c.3342G>A, where the heterozygote subject showed higher methylation level ($P$ = 0.045, Figure 5.9a). The silent mutation located in exon 1, c.129C>T (rs61742233), was associated with decreasing methylation level ($P$ = 0.013, Figure 5.9b) at CpG site 49 which was located 3 bp upstream (nucleotide position +132). rs55894764, another silent mutation located in exon 1, and rs7294, an intron 2 SNP, were both associated with lower levels of methylation ($P$ = 0.016, Figure 5.9c; $P$ = 0.026, Figure 5.9d, respectively) at CpG site 44 (nucleotide position +109). However after correction for multiple testing, none of the above associations remained significant.

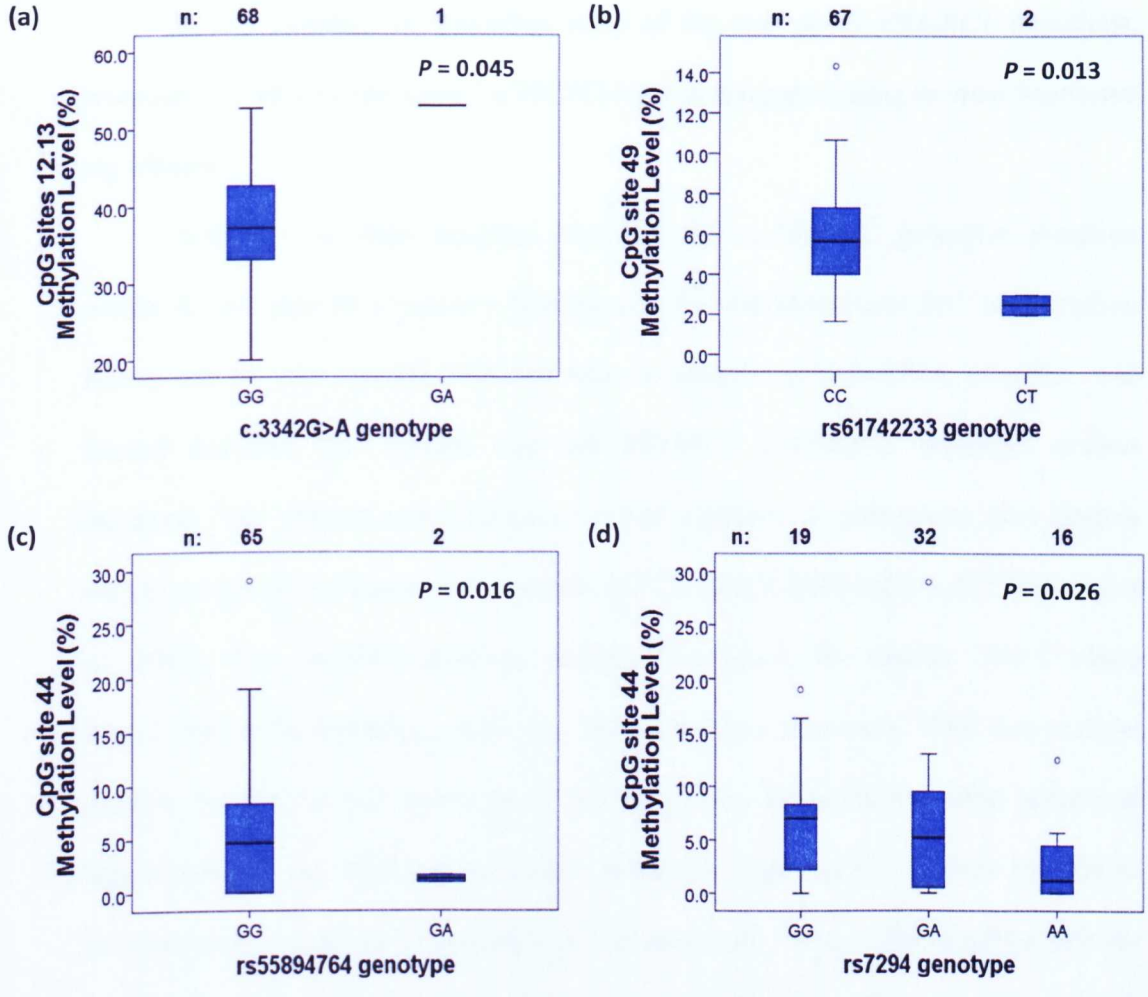



**Figure 5.9. Box and whisker plots showing the levels of methylation at CpG sites based on *VKORC1* genotype groups:** (a) CpG sites 12.13 by intron 2 c.3342G>A, (b) CpG site 49 by rs61742233, (c) CpG site 44 by rs55894764 and (d) rs7294. Boxes represent $25^{th}$-$75^{th}$ percentiles of warfarin doses, whiskers represent $5^{th}$-$95^{th}$ percentiles, and solid lines represent median methylation level in each group. Open dots represent outliers.

## 5.4 DISCUSSION

In this chapter, the predicted roles of the two novel *VKORC1* mutations, promoter c.-160G>C and exon 1 c.79C>G were investigated using *in vitro* functional experiments.

Although *in silico* analysis predicted the c.-160G>C promoter mutation results in the gain of a putative binding site for the ubiquitous Sp1 transcription factor, our *in vitro* results revealed that no specific protein-DNA complex was formed between Sp1 protein and the *VKORC1* c.-160G>C promoter mutant sequence. The DNA-binding domain of Sp1 contains 3 contiguous zinc fingers which recognises the consensus sequence (G/T)GGGCGG(G/A)(G/A)(C/T) (Song *et al.,* 2001). Our *VKORC1* promoter sequence containing the mutant -160 C allele shares only 60% homology with the Sp1 consensus sequence. This may not be specific enough for Sp1 transcription factor binding. Interestingly, when present at high concentration, the promoter mutant sequence displaced Sp1 protein binding to its consensus sequence, suggesting Sp1 protein may have a low binding affinity towards the mutant promoter sequence. It is likely therefore that, through its weak binding to the *VKORC1* c.-160G>C mutant promoter sequence, Sp1 itself or via interaction with co-regulators, may possibly lead to a higher transcriptional activity of the *VKORC1* gene.

The potential functional effect of the *VKORC1* promoter c.-160G>C mutation was also evaluated using a transient transfection assay in HepG2 cell line. Compared to the wild-type promoter -160 G allele, the promoter -160 C mutant allele exhibited approximately 20% significant increase in promoter activity. This increase in promoter activity may partly explain why the patient with the *VKORC1* promoter heterozygous -160 GC genotype require higher warfarin dose. An elevation in

*VKORC1* promoter activity would lead to a rise in *VKORC1* mRNA expression, and subsequently an increased translation of the VKOR protein. This increase in VKOR activity augments the regeneration of reduced vitamin $K_1$, producing more active vitamin K-dependent clotting factors, resulting in higher warfarin dose requirement for its desired anticoagulation effect. It is important to point out that recent studies have suggested that *in vitro* reporter gene assays may not accurately represent gene regulation and expression *in vivo* (Cirulli and Goldstein 2007) and must be viewed with caution. Inconsistent results between chromatin immunoprecipitation (ChIP) analysis and luciferase reporter gene assays have also been reported (Bodin *et al.*, 2005a; Yuan *et al.*, 2005; Bu and Gelman 2007; Kato *et al.*, 2007; Wang *et al.*, 2008), where ChIP is a type of immunoprecipitation method for investigating interactions between specific proteins and a genomic DNA region in intact cells. Therefore, further work using ChIP assay will need to be carried out to confirm our results from the luciferase reporter gene assay.

The missense mutation c.79C>G in exon 1 of *VKORC1* (p.Leu27Val) was predicted to be likely deleterious. Using site-directed mutagenesis and recombinant expression in HepG2 cells, this amino acid substitution demonstrated approximately 10% decrease in PIVKA-II level relative to the wild-type sequence at a warfarin concentration of 3 μM, but this difference in PIVKA-II levels was not statistically significant at the higher warfarin concentration of 30 μM. Nonetheless, the reduced level in PIVKA-II suggests there was an increase in the γ–carboxylation of vitamin K-dependent proteins which could be due to either higher functional efficiency of the VKOR complex or decreased VKOR sensitivity to warfarin.

Inter-individual DNA methylation in relation to age, gender and genotype effects were also investigated. Differential global or loci-specific DNA methylation

patterns have been reported to be age- and/or gender-related (Fuke *et al.*, 2004; Fraga *et al.*, 2005; Martin 2005; El-Maarri *et al.*, 2007; Boks *et al.*, 2009). DNA methylation at the *VKORC1* promoter region, however, was not associated with age or gender in our cohort of warfarin resistant patients. *Cis* effects of several genetic variations on DNA methylation levels were observed with a number of CpG sites but due to the small sample and effect sizes, these associations did not withstand correction for multiple testing. In the majority (73%) of the CpG sites interrogated, no substantial DNA methylation variation was present and this is possibly due to tissue type. Several studies suggest that DNA methylation plays an important role in the specialization of tissues (Nagase and Ghosh 2008), while other studies suggest that the role of DNA methylation in tissue-specific gene expression is confined to tissue-specific differentially methylated regions (T-DMRs) (Song *et al.*, 2005; Kitamura *et al.*, 2007). Given that *VKORC1* is predominantly expressed in the human liver (Rost *et al.*, 2004a), quantification of methylation in liver DNA might have been preferable. However, liver tissues could not be obtained from our patient cohort and peripheral blood mononuclear cells have been used as a non-invasive surrogate. In a recent study by Wang and colleagues, the levels of *VKORC1* DNA methylation in human autopsy liver DNA were quantified but no correlation with *VKORC1* SNPs was observed (Wang *et al.*, 2008), which was similar to our findings in PBMC DNA. Nevertheless, further studies that include several tissue types preferably from single subjects will be essential to provide new insights into the correlation between tissue-specificity and the epigenetic control of *VKORC1*.

This chapter has investigated the roles of *VKORC1* nucleotide substitutions in the pathogenesis of warfarin resistance. Further structural characterization of the

effects of these informative *VKORC1* substitutions on vitamin K metabolism and the interaction with warfarin are now required.

# CHAPTER 6

# Clotting factors and warfarin treatment – genetic, biochemical and clinical variability

# CONTENTS

**6.1 INTRODUCTION** ..... **185**

**6.2 PATIENTS, MATERIALS AND METHODS** ..... **187**

6.2.1 Patients and blood samples ..... 187

6.2.2 Coagulation assays ..... 187

*6.2.2.1 Quality control* ..... *188*

*6.2.2.2 Prothrombin Time (PT) assay* ..... *189*

*6.2.2.3 Activated Partial Thromboplastin Time (aPTT) assay* ..... *190*

*6.2.2.4 Protein C Clotting assay* ..... *190*

*6.2.2.5 Protein S assay* ..... *191*

6.2.3 Genotyping ..... 193

*6.2.3.1 Candidate gene approach* ..... *193*

*6.2.3.2 Genome-wide scan approach* ..... *193*

6.2.4 Statistical analysis ..... 195

*6.2.4.1 Longitudinal data analysis* ..... *195*

*6.2.4.2 Univariate analysis* ..... *196*

*6.2.4.3 Multiple regression analysis* ..... *196*

*6.2.4.4 Correction for multiple testing* ..... *197*

**6.3 RESULTS** ..... **198**

6.3.1 Patient population ..... 198

6.3.2 Plasma levels of clotting factors in patients initiated onto warfarin therapy ..... 200

6.3.3 Correlation between INR and activity levels of clotting factors ..... 204

6.3.4 Association of SNPs with variability in clotting factor activity levels ..... 206

6.3.5 Association of changes in clotting factor activity levels with clinical outcomes ..... 213

6.3.6 Univariate analysis of association between SNPs, reduction in clotting factor levels and clinical outcomes ..... 218

6.3.7 Multivariate model for warfarin stable dose ..... 219

**6.4 DISCUSSION** ..... **222**

## 6.1 INTRODUCTION

The formation of thrombi involves several haemostatic factors including factors II (prothrombin), V, VII, IX and X, proteins C and S, as shown in Figure 1.1 in Chapter 1. Warfarin does not directly antagonise thrombin activity to achieve its net clinical effect of anticoagulation; instead it acts by reducing the enzymatic activities of the vitamin K-dependent procoagulant (factors II, VII, IX and X) and anticoagulant (proteins C and S) proteins, resulting in the suppression of thrombin generation by non-functional prothrombin and factor X (Furie and Furie 1990). Factor V circulates with little or no activity but serves as an essential protein in the coagulation pathway by acting as a cofactor for the conversion of prothrombin to thrombin.

On initiation of warfarin, the rate of change in the vitamin K-dependent proteins will be dependent on their half-lives. Functions of clotting factors with short half-lives is lost acutely while clotting factors with longer half-lives will remain functional for a longer period of time. Table 6.1 lists the half-lives of the six vitamin K-dependent clotting factors. The procoagulant factor VII has a relatively short half-life (4 – 8 h) and high synthesis rate (Dike *et al.*, 1980), and the INR may be prolonged within 24 – 36 h after warfarin administration due to a decrease in factor VII levels. The anticoagulant protein C also has a short half-life (6 – 9 h) (Bertina *et al.*, 1982), and decline in its activity may counteract the effect of the changes in factor VII concentrations. By contrast, factors with a longer half-life and lower synthesis rate (e.g., factor II and protein S) influence the INR later, and may therefore be more important in determining variability in maintenance doses. Although factor IX has a half-life of 21 – 30 h, adequate anticoagulation is not achieved until the levels of biologically active factors II (half-life: 42 – 72 h)

(Breckenridge 1978) and X (half-life: 27 – 48 h) (Trask *et al.,* 2004) are sufficiently decreased, which may require 4 to 6 days after the start of warfarin therapy.

There is some evidence that predisposition to thrombosis is associated with high levels of factors II (Poort *et al.,* 1996) and IX (Vlieg *et al.,* 2000) as well as reduced proteins C and S (Simioni 1999). However, what is unclear is the rate and extent of falls in these factors individually and in combination. Interestingly, two missense mutations at Ala-10 of the factor IX propeptide have been implicated with causing factor IX levels to drop to 3% or less upon warfarinisation and causing severe bleeding despite therapeutic INRs (Chu *et al.,* 1996; Oldenburg *et al.,* 1997).

The variation in activity levels of the vitamin K-dependent clotting proteins (factors II, VII, IX and X, proteins C and S) and the clotting cofactor (factor V) in patients initiated onto warfarin therapy have not been measured in one large cohort previously. The aims of this chapter were therefore to investigate (i) the extent to which these seven proteins decline on warfarin therapy, (ii) whether such variability is related to common polymorphisms in the clotting factors II, V, VII, IX and X, proteins C and S genes (*F2*, *F5*, *F7*, *F9*, *F10*, *PROC*, and *PROS1*), and (iii) how changes in clotting factor levels affect variability in INR, warfarin dose requirements and bleeding complications.

**Table 6.1. Half-lives of vitamin K-dependent clotting factors.**

| Clotting factor | Half-life (h) |
|---|---|
| Procoagulant: | |
| Factor II | 42 - 72 |
| Factor VII | 6 - 8 |
| Factor IX | 18 - 24 |
| Factor X | 27 - 48 |
| Anticoagulant: | |
| Protein C | 6 - 7 |
| Protein S | 30 -60 |

## 6.2 PATIENTS, MATERIALS AND METHODS

### 6.2.1 Patients and blood samples

Patients on prospective warfarin therapy (n = 1000) were recruited as described in section 2.2.1. Sodium citrated blood samples (9 ml) were collected from all patients before warfarin initiation (baseline visit) and at two of the follow-up visits (day 7 and day 28). Centrifugation of blood samples was carried out at 2, 600 *g* for 20 min and plasma was stored in 1 ml aliquots at -80°C until analysis.

Among the 1000 patients recruited, genotype data for the seven clotting factor genes (*F2*, *F5*, *F7*, *F9*, *F10*, *PROC*, and *PROS1*) were available for a total of 745 patients, 31 from a candidate gene approach genotyping and 714 from genome-wide genotyping (of which 280 were also genotyped by a candidate gene approach and genotype data from both genotyping methods were 100% concordant). As part of quality control, patients who missed any warfarin doses the day before blood samples were taken (n = 57), or patients who did not complete the two follow-up visits (n = 69), were excluded and 619 patients remained for downstream statistical analyses.

### 6.2.2 Coagulation assays

Plasma activities of circulating procoagulant (factors II, V, VII, IX and X) and anticoagulant (proteins C and S) proteins were measured using an automated analyzer, Multi-Channel Discrete Analyzer (MDA)-180 (bioMérieux®, Inc., Durham, NC, USA), at the Haematology Laboratory of the Royal Liverpool University Hospital, under the supervision of Professor Cheng Hock Toh. All plasma samples were measured in duplicate, at 1:10 and 1:20 dilutions for the clot-based assays, and at 1:1 and 1:2 dilutions for the protein S immune-turbidimetric assay. Results were

calculated to percent activity by interpolation from the reference curve, and were subsequently multiplied by the respective dilution factor and the mean values of the duplicate readings were reported.

#### *6.2.2.1 Quality control*

Immediately before assaying the patient specimens, the MDA-180 system prepares a reference curve using the reference material plasma, Coagulation Reference (bioMérieux®). The coagulation reference plasma is a pooled plasma control prepared from at least 100 Caucasian healthy volunteers, giving the average presence of all coagulation factors and inhibitors. Using the dilutions specific for each assay as listed in Table 6.2, the reference curve was constructed by fitting a regression model, a piecewise function with two subfunctions each of which used a second order polynomial fit to different subgroups of the reference plasma dilution series. Only reference curves with correlation coefficients > 0.9 were accepted.

**Table 6.2. Dilutions used in constructing coagulation assay reference curve.**

| Assay | Reference curve dilutions |
|---|---|
| PT (for factors II, VII and X) | 1:10, 1:20, 1:40, 1:80, 1:160, 1:320 |
| aPTT (for factors V and IX) | 1:10, 1:20, 1:40, 1:80, 1:160, 1:320, 1:640 |
| Protein C clot-based | 1:10, 1:20, 1:40, 1:80, 1:160 |
| Protein S immune-turbidimetric | 1:1, 1:2, 1:3, 1:6 |

PT: Prothrombin Time; aPTT: Activated Partial Thromboplastin Time

Two levels of controls for each assay were then tested, using a "normal" and "abnormal" control. The "normal" control has characteristics similar to those of fresh normal plasma that may be used as a control to monitor thrombosis and haemostasis assays, while the "abnormal" control is prepared from citrated plasma of healthy

donors where all clotting factor levels were reduced, i.e. abnormally low. The "normal" and "abnormal" controls used for the clot-based assays were MDA Verify 1 and Coagulation Control A (bioMérieux®), respectively; while that for the protein S immune-turbidimetric assay were STA®-LIATest® Control N and Control P (Diagnostica Stago, Asnieres-sur-Seine, France), respectively. Control results had to be within accepted limits (lot specific target values provided by the manufacturer) before assaying the patient specimens.

#### *6.2.2.2 Prothrombin Time (PT) assay*

Factors II, VII and X levels were determined by a one-stage Prothrombin Time (PT) based clotting assay using respective factor depleted plasma (Precision BioLogic Inc., Dartmouth, Canada) and Simplastin® HTF (bioMérieux®) as the activator of the extrinsic coagulation cascade. Simplastin® HTF is a thromboplastin reagent which contains the human thromboplastin tissue factor, phospholipids and calcium ions, and when added to anticoagulated plasma, it triggers the formation of a fibrin clot. The instrument procedure on the MDA-180 system is described in the following paragraph and depicted in Figure 6.2a.

For a 1:10 dilution, 10 µl of specimen or control was pipetted into an optically clear cuvette and was serially diluted with imidazole buffer (bioMérieux®) to a total volume of 50 µl. The instrument then pipetted 50 µl of factor deficient plasma into each cuvette and the fluids were warmed to 37 ± 1°C. To initiate the clot reaction, 100 µl of warmed Simplastin® HTF was pipetted into each cuvette. As a clot forms, the amount of light that passes through a reaction mixture abruptly decreases, decreasing the output of the photodiode detectors. The time required for this decrease to occur was measured.

#### *6.2.2.3 Activated Partial Thromboplastin Time (aPTT) assay*

Levels of factors V and IX were quantitated by a one-stage Activated Partial Thromboplastin Time (aPTT) based clotting assay using factor deficient plasma (Precision BioLogic Inc.) and MDA Platelin LS test kit (bioMérieux®). The MDA Platelin LS test kit consists of the MDA Platelin LS reagent (which contains phospholipids with micronized silica as particulate contact activators) and MDA Platelin LS $CaCl_2$ required for the activation of the intrinsic coagulation cascade. The instrument procedure for the aPTT assay (Figure 6.2b) is similar to that described for PT assay except after the addition of deficient plasma, 50 µl of warmed MDA Platelin LS reagent was mixed with the sample and incubated at 37 ± 1°C for 3 min 40 s. The aPTT activation was then initiated with the addition of 50 µl of warmed MDA Platelin $CaCl_2$ and the timing for clot detection was measured.

#### *6.2.2.4 Protein C Clotting assay*

A functional clotting assay based on the prolongation of the aPTT was utilized to measure protein C activity using Protein C deficient plasma and Protac® (Technoclone GmbH, Vienna, Austria). Protac® contains phospholipids, calcium ions and a highly purified extract of the snake venom from *Agkistrodon Contortrix* which is a quick-acting, direct activator of protein C, converting it from its zymogen to the protease. The active enzyme was then determined indirectly by the prolongation of aPTT (see Figure 6.2c).

### *6.2.2.5 Protein S assay*

Free protein S levels were measured with an immune-turbidimetric assay using the STA®-LIATest® Free Protein S kit (Diagnostica Stago) which consists of the protein S buffer and latex reagent. The latex reagent is a suspension of latex microparticles, coated by covalent bonding with monoclonal antibodies specific for free protein S. When testing the 1:2 dilution, 12 µl of specimen or control was pipetted by the MDA-180 instrument into an optically clear cuvette. 12 µl of imidazole buffer was then added and the fluids were heated to 37 ± 1°C. Following this, 50 µl of warmed protein S buffer was pipetted into the cuvette and incubated at 37 ± 1°C for 3 min 40 s. Finally, 75 µl of warmed latex reagent was added to the mixture and after a 15 s blank interval, the timing for latex agglutination was monitored at 540 nm wavelength for 240 s (Figure 6.2d). The antigen-antibody reaction between the test sample and latex reagent led to an agglutination of the latex microparticles, which induced an increase in turbidity of the reaction medium. The increase in turbidity was reflected by an increase in absorbance and is a function of the free protein S level present in the test sample.

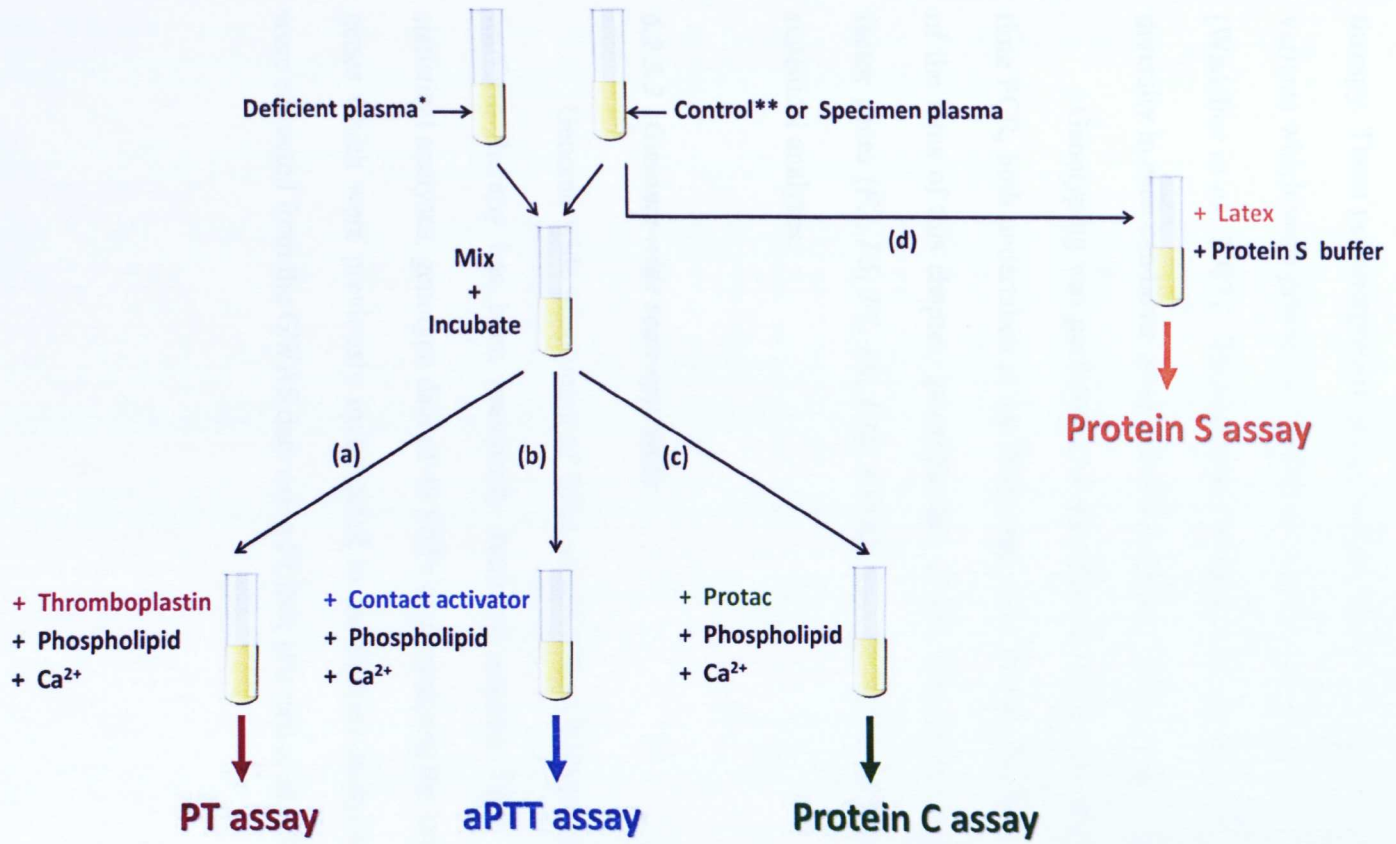



**Figure 6.2. MDA-180 instrument procedures for coagulation assays undertaken in this study.** (a) PT assay, (b) aPTT assay, (c) Protein C assay, and (d) Protein S assay. *Plasma deficient in either factor II, V, VII, IX, X or protein C was used, depending on the assay tested. **Control plasma each coagulation assay was used for each clot-based assay. **Control plasma refers to either the "normal" or "abnormal" plasma being tested as part of quality control.

### 6.2.3 Genotyping

#### *6.2.3.1 Candidate gene approach*

As part of the interim analysis described in section 2.2.3, a total number of 197 SNPs spanning 29 candidate genes involved in the pharmacology of warfarin were selected and investigated in the first 311 patients on prospective warfarin therapy. These polymorphisms were a combination of tagging SNPs and functional variants which were previously identified in a study undertaken in a Swedish cohort (Wadelius *et al.*, 2007). These genetic variants explain at least 95% of the genetic diversity in each candidate gene. Table 6.3 lists the 29 genes investigated.

Genotyping was performed by MALDI-TOF mass spectrometry and by real-time PCR, both undertaken at the Wellcome Trust Sanger Institute. For the purpose of the aims of this chapter, genotype data of 48 SNPs spanning the seven clotting factor genes (*F2*, *F5*, *F7*, *F9*, *F10*, *PROC*, and *PROS1*) were used for downstream statistical analyses.

#### *6.2.3.2 Genome-wide scan approach*

Genome-wide genotyping of DNA samples from 752 patients on prospective warfarin therapy has been previously described section 3.2.9. For downstream statistical analyses, genotype data of 48 SNPs encompassing the seven clotting factor genes which were previously investigated in the interim analysis (section 6.2.3.1) were extracted from the GWAS data using PLINK (Purcell *et al.*, 2007).

**Table 6.3. List of 29 genes investigated in the interim analysis.**

| Gene symbol | Gene name |
|---|---|
| *ABCB1* | ATP binding cassette, subfamily B, member 1 |
| *APOE* | Apolipoprotein E |
| *CALU* | Calumenin |
| *CYP1A1* | Cytochrome P450, family 1, subfamily A, polypeptide 1 |
| *CYP1A2* | Cytochrome P450, family 1, subfamily A, polypeptide 2 |
| *CYP3A4* | Cytochrome P450, family 3, subfamily A, polypeptide 4 |
| *CYP3A5* | Cytochrome P450, family 3, subfamily A, polypeptide 5 |
| *CYP2C8* | Cytochrome P450, family 2, subfamily C, polypeptide 8 |
| *CYP2C9* | Cytochrome P450, family 2, subfamily C, polypeptide 9 |
| *CYP2C18* | Cytochrome P450, family 2, subfamily C, polypeptide 18 |
| *CYP2C19* | Cytochrome P450, family 2, subfamily C, polypeptide 19 |
| *EPHX1* | Epoxide hydrolase 1 |
| *F2* | Coagulation factor 2 |
| *F5* | Coagulation factor 5 |
| *F7* | Coagulation factor 7 |
| *F9* | Coagulation factor 9 |
| *F10* | Coagulation factor 10 |
| *GAS6* | Growth arrest-specific 6 |
| *GGCX* | Gamma-glutamyl carboxylase |
| *NQO1* | NAD(P)H dehydrogenase, quinone 1 |
| *NR1I2* | Nuclear receptor subfamily 1, group I, member 2 |
| *NR1I3* | Nuclear receptor subfamily 1, group I, member 3 |
| *ORM1* | Orosomucoid 1 |
| *ORM2* | Orosomucoid 2 |
| *PROC* | Protein C |
| *PROS1* | Protein S |
| *PROZ* | Protein Z |
| *SERPINC1* | Serpin peptidase inhibitor, clade C (antithrombin), member 1 |
| *VKORC1* | Vitamin K epoxide reductase complex, subunit 1 |

### 6.2.4 Statistical analysis

HWE for each SNP was determined using a computationally efficient exact $P_{HWE}$ test statistic implemented in HaploView (Wigginton *et al.,* 2005). A $P$-value < 0.001 was assumed to indicate deviation from HWE.

All statistical analyses were performed in SPSS version 18. A $P$-value < 0.05 was regarded as statistically significant. To assess the effect of heparin on baseline clotting factor levels, the independent samples $t$-test was performed. For comparisons of clotting factor levels before and after warfarin commencement, the paired samples $t$-test was undertaken. Relationships between INR variability and clotting factor activity levels were examined by Pearson correlation analysis.

#### *6.2.4.1 Longitudinal data analysis*

A repeated measures linear model was used to assess the changes in activity levels of each clotting factor between patients at each study time point (0, 7 and 28 days after warfarin treatment). To examine the SNP-effect on the variation in clotting factor levels, patients were stratified according to the different genotype groups for each SNP and the measurements were compared between strata. Clinical outcomes evaluated included the achievement of stable warfarin dose, achievement of therapeutic INR, and bleeding complications. As the timing of warfarin commencement relative to the index visit could influence the changes in clotting factor levels, the number of days from warfarin initiation was included as a covariate in all above models.

#### *6.2.4.2 Univariate analysis*

To evaluate the individual effect of each SNP and the decline in activity level of each clotting factor (difference between the index visit and the two follow-up time points at day 7 and 28 after warfarin treatment) on each clinical outcome, univariate tests of association were conducted. For the time to event outcomes (time to achieving warfarin stable dose, time to achieving therapeutic INR and time to bleeding complications), the log-rank test for trend was used. As the distribution of warfarin stable dose was skewed, square-root transformation was undertaken to ensure normal distribution. For the continuous outcome of stable warfarin dose, univariate linear regression analysis was undertaken.

#### *6.2.4.3 Multiple regression analysis*

To determine the relative effects of reduction in clotting factors on variability of warfarin stable dose, stepwise linear regression models were fitted and compared using the likelihood ratio test (LRT). In addition, clinical and genetic factors which have been previously reported (D'Andrea *et al.*, 2005; Caldwell *et al.*, 2008; Gage *et al.*, 2008; Borgiani *et al.*, 2009; Klein *et al.*, 2009; Zhang *et al.*, 2009) to affect warfarin dose variability, namely age, BMI, gender, *CYP2C9*2* (rs1799853), *CYP2C9*3* (rs1057910), *VKORC1* -1693 (rs9923231) and *CYP4F2* rs2108622 and rs2189784, were assessed univariately in our cohort of patients. Any factors subsequently found significant ($P < 0.05$) were included in the multiple regression analysis. The proportion of variability explained by the covariates was calculated using Nagelkerke's $R^2$ statistic.

#### *6.2.4.4 Correction for multiple testing*

*P*-values from all association tests undertaken were adjusted for multiple testing using the FDR (Benjamini *et al.*, 2001) in the genetics package of R (http://cran.r-project.org/web/packages/genetics/index.html), version 2.6.2. FDR-corrected *P*-values are denoted as $P_c$-values.

## 6.3 RESULTS

### 6.3.1 Patient population

Demographics of the 619 patients are reported in Table 6.4. Most of the patients were Caucasians (98.5%) and majority were male (57%), with mean age of 69 years. Atrial fibrillation (67%) was the most common indication for warfarin therapy, followed by pulmonary embolism (15%) and deep vein thrombosis (9%). In majority of the patients, cardiovascular disease (81%) was a common underlying co-morbidity, followed by musculoskeletal problems (62%) and gastrointestinal disease (38%).

Among the 619 patients, 59 were on low molecular weight heparin (LMWH) therapy when warfarin treatment was initiated. Given that heparin enhances the adhesion of a serine protease inhibitor, antithrombin III, to serine proteases such as factors II, VII, IX and X, proteins C and S, their activity levels could be inhibited in patients on concomitant heparin and warfarin therapies. We however did not find any significant differences in the mean baseline levels of these clotting factors when patients on overlapping heparin and warfarin therapy were compared to patients on warfarin treatment only.

Of the 619 patients, 149 (24%) patients had their index visit between one and eleven days before they started warfarin, while 378 (61%) patients had their index visit within 2 days of commencing warfarin. Of the remaining patients, 87 (14%) had their index visit on the third day, and 5 (1%) had their index visit on the fourth day after starting warfarin.

**Table 6.4. Clinical profile of 619 warfarin patients.**

| Characteristic | N (%) |
|---|---|
| **Gender - Male** | 355 (57) |
| **Age in years, mean (range)** | 69 (19-95) |
| **BMI[a], mean (range)** | 28 (13-61) |
| **Ethnicity** | |
| White | 612 (98.9) |
| Black Other | 4 (0.6) |
| Black African | 2 (0.3) |
| Other[b] | 1 (0.2) |
| **Indication for warfarin** | |
| Atrial Fibrillation | 412 (67) |
| Pulmonary Embolism | 94 (15) |
| Deep Vein Thrombosis | 59 (9) |
| Cerebrovascular accident and Transient ischaemic attacks | 36 (6) |
| Mechanical heart valve replacement | 14 (2.3) |
| Myocardial infarction | 2 (0.3) |
| Dilated left atrium | 1 (0.2) |
| Other[c] | 29 (5) |
| **Co-morbidity** | |
| Cardiovascular disease | 502 (81) |
| Musculoskeletal problems | 385 (62) |
| Gastrointestinal disease | 236 (38) |
| Respiratory disease | 216 (35) |
| Neurological disease | 136 (22) |
| Urological condition | 113 (18) |
| Renal disease | 56 (9) |
| History of falls | 46 (7) |
| Hepatic disease | 24 (4) |

BMI: Body Mass Index.

[a] BMI missing for 2 patients.

[b] Other self-reported ethnicities include: Mixed race.

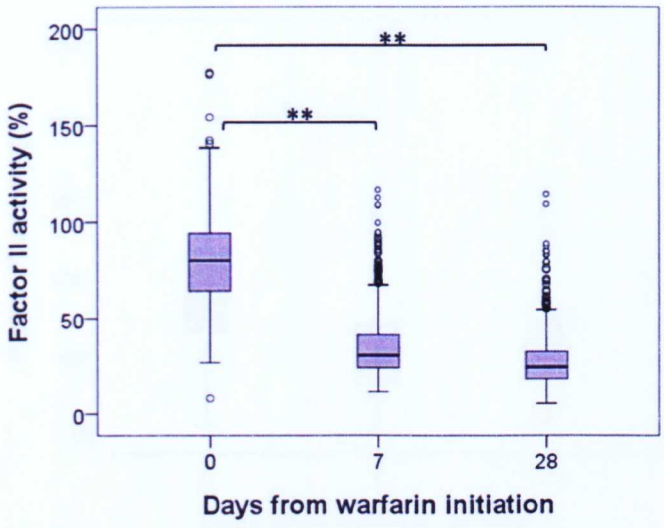
[c] Other indications include: systemic lupus erythematosus; anti-phospholipid syndrome; short saphenous vein thrombosis; valvular heart disease; sagittal sinus thrombosis; dilated left ventricle; occluded graft in leg; apical aneurysm; urticaria with angioedema; aortic and mitral regurgitation; ischaemic colitis; mitral stenosis; and post-surgery.

### 6.3.2 Plasma levels of clotting factors in patients initiated onto warfarin therapy

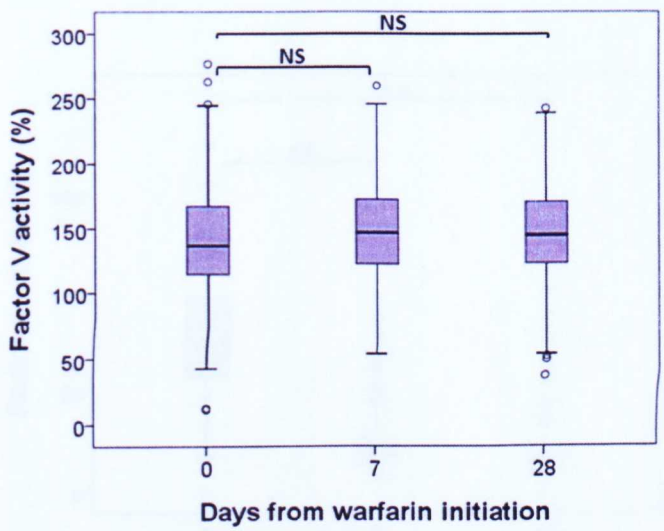
Figure 6.3 displays the variability in plasma levels of clotting factors across the 3 study time points (0, 7, and 28 days after warfarin initiation) in 619 patients. The plasma levels of procoagulation factors II, VII, IX and X, and anticoagulation proteins C and S, were significantly lower at 7 days and 28 days after warfarin treatment, compared to baseline. Factor II levels decreased by 55% (95% CI: 53% – 57%) at day 7 and 65% (95% CI: 64% – 67%) at day 28 (Figure 6.3a). The activity level of factor VII was reduced by 53% (95% CI: 50% – 56%) at day 7 and by 41% (95% CI: 38% – 43%) at day 28 (Figure 6.3c). For factor IX, a 44% (95% CI: 41% – 46%) and 32% (95% CI: 30% – 33%) reduction in plasma activity was observed at day 7 and day 28, respectively (Figure 6.3d). A huge reduction in factor X level (Figure 6.3e) was observed at day 7 (mean: 70%, 95% CI: 69% – 72%), and at day 28 (mean: 79%, 95% CI: 77% – 80%). The activity level of protein C (Figure 6.3f) decreased by 31% both at day 7 (95% CI: 28% – 35%) and at day 28 (95% CI: 28% – 33%). Protein S level was reduced by 47% (95% CI: 45% – 48%) at day 7 and by 50% (95% CI: 55% – 57%) at day 28 (Figure 6.3f). However, very little change was observed with factor V level after warfarin treatment (Figure 6.3b). Only a small 5% increase in activity level was observed at day 7 (95% CI: 3% – 7%) and day 28 (95% CI: 3% – 6%).

After 28 days of warfarin therapy, factor X showed the lowest level at ~17%, intermediate levels were exhibited by factors II (~28%) and VII (~38%), while factor IX, proteins C and S showed the highest activity at 64%, 48% and 58%, respectively.

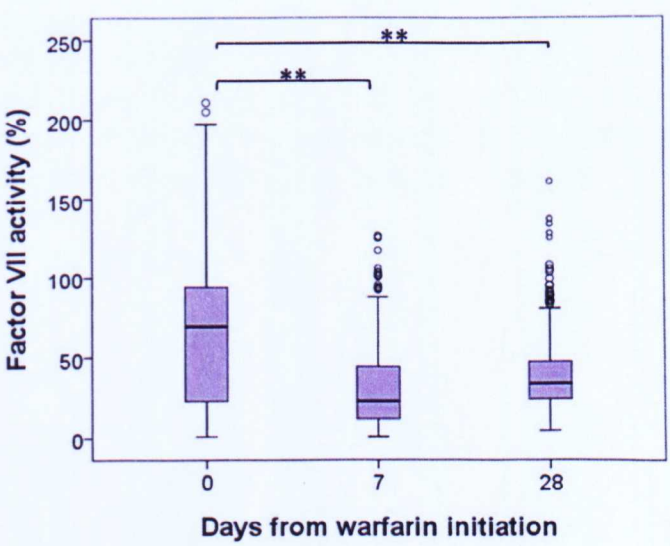
(a)


**
**
Factor II activity (%)
200
150
100
50
0
0
7
28
Days from warfarin initiation

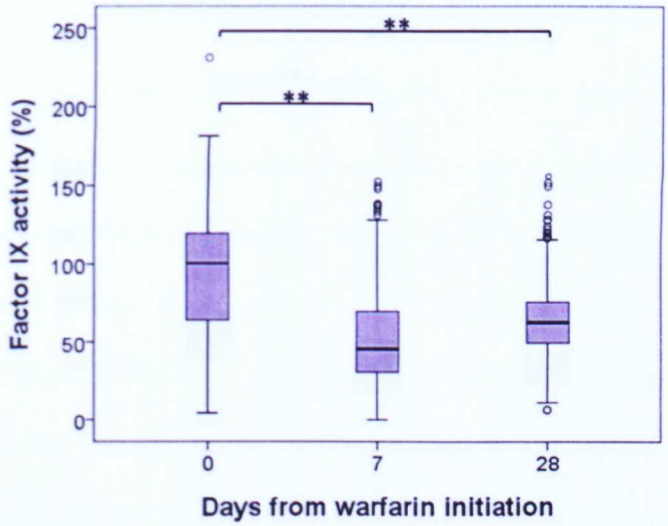
(b)


NS
NS
Factor V activity (%)
300
250
200
150
100
50
0
0
7
28
Days from warfarin initiation

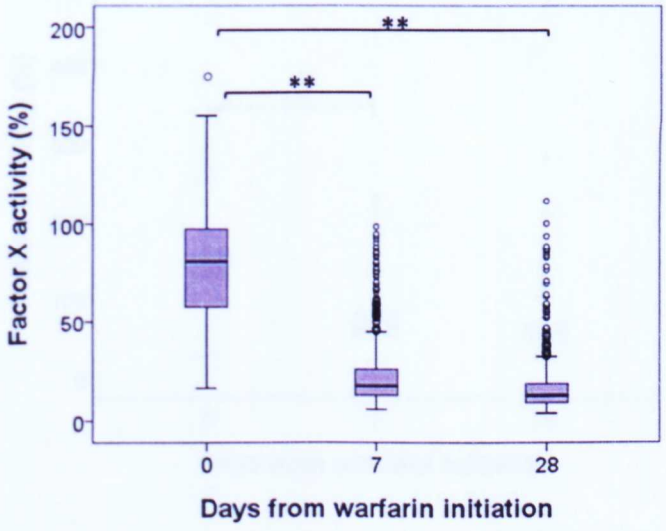
(c)


**
**
Factor VII activity (%)
250
200
150
100
50
0
0
7
28
Days from warfarin initiation

(d)


**
**
250
200
150
100
50
0
Factor IX activity (%)
0
7
28
Days from warfarin initiation

(e)


**
**
200
150
100
50
0
Factor X activity (%)
0
7
28
Days from warfarin initiation

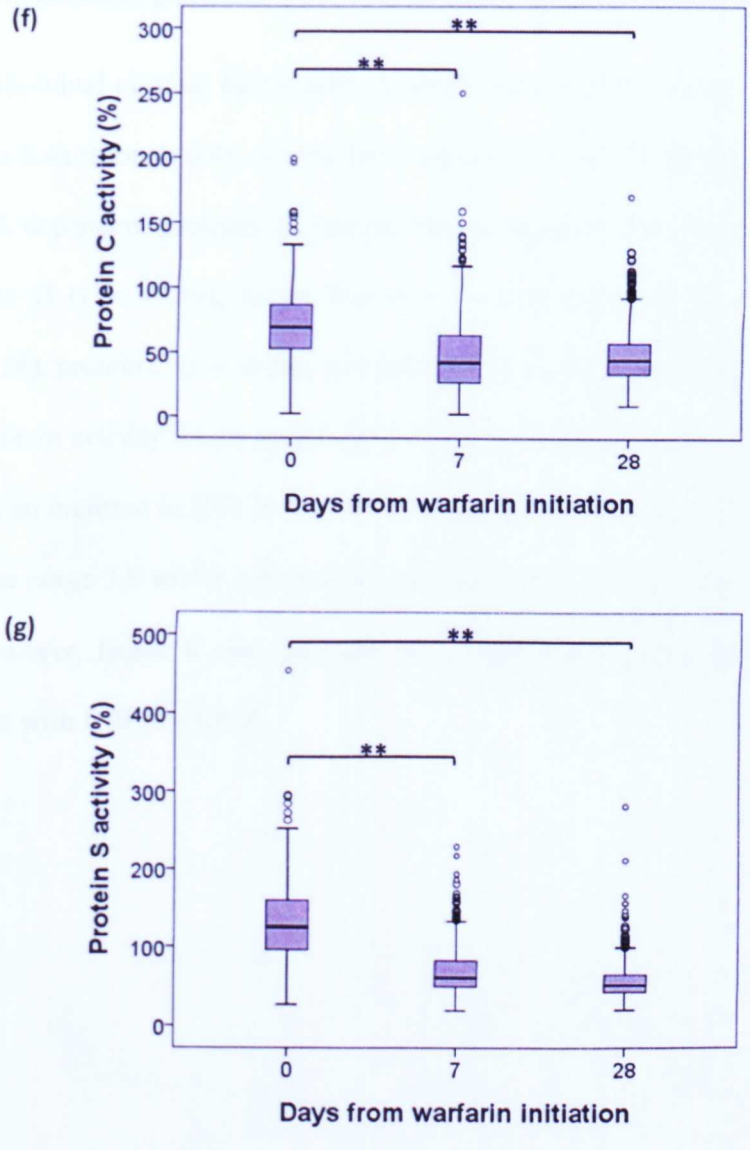



**Figure 6.3. Plasma levels of clotting factors in patients at 0, 7 and 28 days of warfarin treatment.** (a) Factor II, (b) Factor V, (c) Factor VII, (d) Factor IX, (e) Factor X, (f) Protein C, and (g) Protein S. Boxes represent 25th-75th percentiles of clotting factor activity levels, whiskers represent 5th-95th percentiles, solid lines represent median clotting factor level at respective time points, and open dots represent outliers. NS = not significant. ***P*-value < $1 \times 10^{-7}$.

### 6.3.3 Correlation between INR and activity levels of clotting factors

Individual clotting factor activity levels were plotted against the INR values in Figures 6.4a through 6.4g. As the INR value increased, the activity levels of the 6 vitamin-K dependent proteins decreased. Strong negative correlations between INR and factor II ($r = -0.59$), factor VII ($r = -0.67$), factor IX ($r = -0.69$), factor X ($r = -0.58$), protein C ($r = -0.54$), and protein S ($r = -0.53$) activities were observed. The decline in activity levels appeared to follow a triphasic pattern. An initial sharp drop with an increase in INR from 1.0 to 2.0 was followed by a gradual decline with INR in the range 2.0 to 4.0, ultimately reaching a plateau when the INR was 4.0 or more. However, factor V, the cofactor for thrombin formation, did not show any correlation with INR ($r = 0.09$).

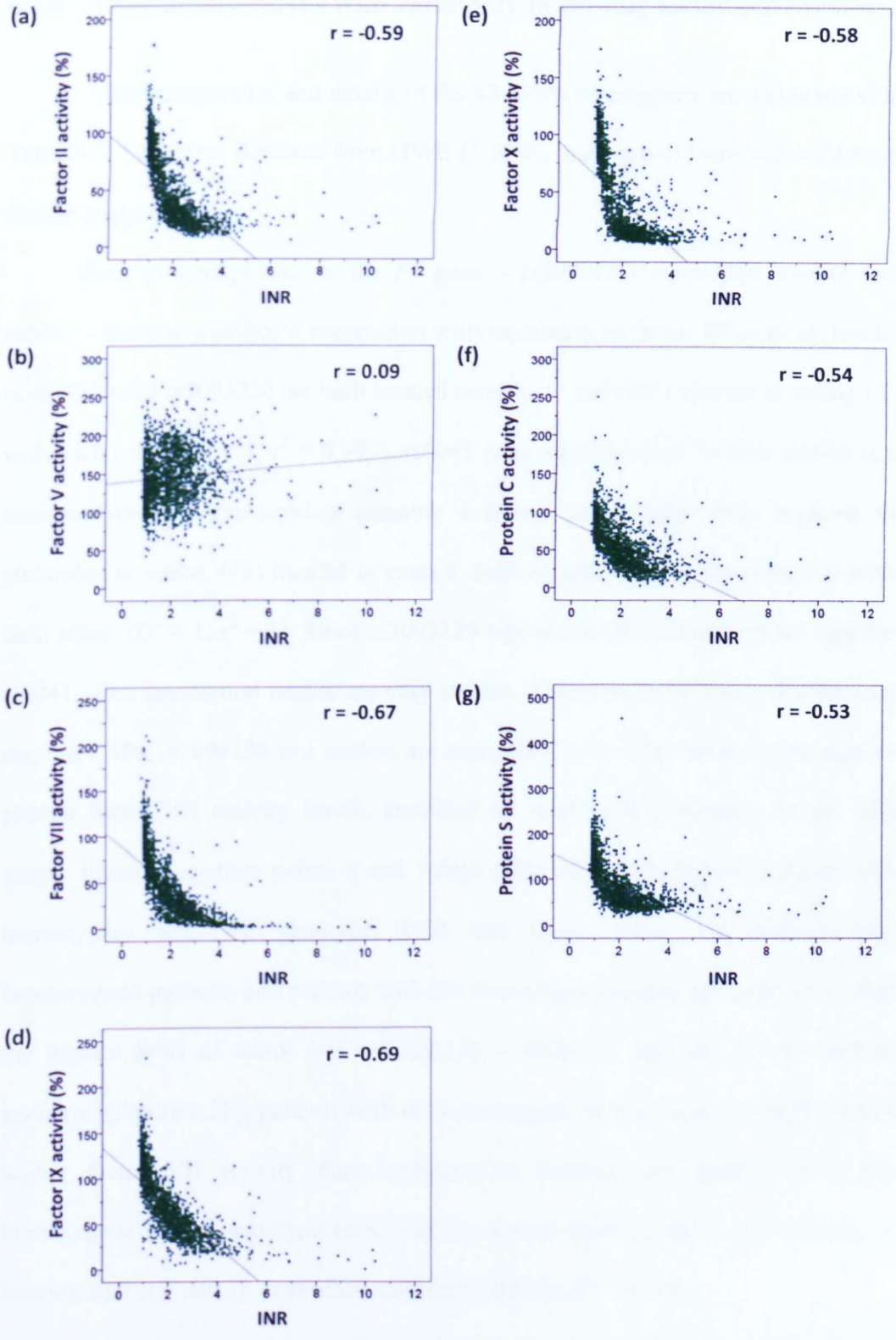



**Figure 6.4. Scatter plots showing correlations between the international normalised ratio (INR) and activity levels in (a) factor II, (b) factor V, (c) factor VII, (d) factor IX, (e) factor X, (f) protein C, and (g) protein S in 619 patients on warfarin.** Values of Pearson correlation coefficient are shown and all *P*-values are < 0.001.

### 6.3.4 Association of SNPs with variability in clotting factor activity levels

Allele frequencies and details of the 48 SNPs investigated are summarised in Table 6.5. SNPs that deviated from HWE (3 SNPs in *F9* gene) were excluded from further analysis.

Four polymorphisms in the *F7* gene – rs3093229, rs3093230, rs6041 and rs6047 – showed significant association with variability in factor VII activity levels. rs3093229 and rs3093230 are both located near the 5' end of *F7* and are in strong LD with each other (D' = 1, $r^2$ = 0.995). rs6041 is located in intron 7 while rs6046 is a nonsynonymous polymorphism (causing a amino acid change from arginine to glutamine at codon 413) located in exon 8, both of which are in complete LD with each other (D' = 1, $r^2$ = 1). Since rs3093229 tags for rs3093230 and rs6046 tags for rs6041, their association results are very similar. Therefore, only results for the two tagging SNPs, rs3093230 and rs6046, are reported. Figure 6.5a shows the changes in plasma factor VII activity levels, stratified by rs3093230 genotype. As the line graphs illustrate, at time points 0 and 7 days after warfarin initiation, patients with homozygous wild-type genotype (GG) had lower factor VII activity than heterozygous patients, and patients with the homozygous mutant genotype (AA) had the highest level of factor VII activity ($P_c$ = 0.02). In contrast, for the rs6046 genotype (Figure 6.5b), patients with the homozygous wild-type genotype (GG) had higher factor VII activity than heterozygous patients, and patients with the homozygous mutant genotype (AA) had the lowest level of factor VII activity, at baseline and at 7 and 28 days after warfarin treatment ($P_c$ = 0.003).

Given that genes *F7* and *F10* are located in close proximity (~ 2000 bp) with each other on chromosome 13q34, and some degree of LD is observed between these 2 genes (see Figure 6.6), the effect of *F7* SNPs on factor X level changes and vice

versa were investigated. No significant association was observed between *F7* SNPs and factor X level but interestingly, an intron 1 genetic variant within *F10*, rs776905, exhibited a borderline association with changes in factor VII activity levels ($P_c$ = 0.04). As shown in Figure 6.7, patients with homozygous wild-type genotype (AA) had higher factor VII activity than heterozygous patients, and patients with the homozygous mutant genotype (CC) had the lowest level of factor VII activity, at baseline and at 7 and 28 days after warfarin commencement.

Interestingly, when compared to baseline, the relative reduction in factor VII activity was not significantly different among the 3 different genotype groups for *F7* SNPs rs3093230 and rs6046 and *F10* rs776905 after warfarin treatment, suggesting that these 3 SNPs do not play a role in the rate of decline in factor VII level during early warfarin treatment.

No significant association was found with variability in activity levels of factors II, V, protein C and protein S, in relation to SNPs in *F2*, *F5*, *PROC* and *PROS1*.

**Table 6.5. Details of 48 coagulation factor SNPs genotyped in 619 warfarin patients.**

| Gene | SNP | Chromosomal Position | SNP Location and Function | SNP Alleles | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| F2 | rs3136435 | Chr 11: 46699002 | Intron 4 | G > A | 1 | 100 | 0.064 | 0.066 | *HapMap R28* |
| F2 | rs2070851 | Chr 11: 46701104 | Intron 4 | T >C | 0.6979 | 99.8 | 0.218 | 0.252 | *NCBI B36* |
| F2 | rs2070852 | Chr 11: 46701501 | Intron 5 | C > G | 0.7681 | 99.8 | 0.285 | 0.257 | *HapMap R28* |
| F2 | rs2282687 | Chr 11: 46707977 | Intron 11 | T > C | 1 | 99.8 | 0.133 | 0.115 | *HapMap R28* |
| F2 | rs3136516 | Chr 11: 46717332 | Intron 11 | A > G | 0.7671 | 98.9 | 0.478 | 0.456 | *HapMap R28* |
| F5 | rs9332618 | Chr 1: 167767105 | Intron 14 | C > T | 0.9602 | 97.7 | 0.112 | 0.119 | *HapMap R28* |
| F5 | rs4656687 | Chr 1: 167771782 | Intron 14 | A > G | 0.2497 | 93.4 | 0.345 | 0.305 | *HapMap R28* |
| F5 | rs1557572 | Chr 1: 167775200 | Intron 13 | A > C | 0.0796 | 99.8 | 0.345 | 0.325 | *HapMap R28* |
| F5 | rs6018 | Chr 1: 167778502 | Exon 13, Missense, Asn817Thr | T > G | 0.3001 | 98.9 | 0.073 | 0.035 | *HapMap R28* |
| F5 | rs6037 | Chr 1: 167780207 | Exon 12, Synonymous, Thr642Thr | G > T | 0.0889 | 98.9 | 0.087 | 0.044 | *HapMap R28* |
| F5 | rs3766110 | Chr 1: 167781807 | Intron 11 | A > C | 0.9853 | 99.4 | 0.23 | 0.228 | *HapMap R28* |
| F5 | rs6025 | Chr 1: 167785673 | Exon 10, Missense, Arg534Gln | C > T | 0.1233 | 100 | 0.011 | 0.022 | *HapMap R28* |
| F5 | rs2298909 | Chr 1: 167787617 | Intron 8 | A > T | 0.9146 | 99.4 | 0.262 | 0.289 | *HapMap R28* |
| F5 | rs721161 | Chr 1: 167788175 | Intron 8 | C > G | 0.4703 | 100 | 0.433 | 0.452 | *NCBI B36* |
| F5 | rs6035 | Chr 1: 167788473 | Exon 8, Synonymous, Lys414Lys | T > C | 1 | 99.2 | 0.081 | 0.088 | *HapMap R28* |
| F5 | rs1894699 | Chr 1: 167793574 | Intron 5 | C > G | 0.5408 | 95.3 | 0.46 | 0.5 | *NCBI B36* |
| F5 | rs6029 | Chr 1: 167796597 | Exon 4, Synonymous, Ala135Ala | C > T | 0.5848 | 99.2 | 0.16 | 0.173 | *HapMap R28* |
| F5 | rs9332504 | Chr 1: 167816960 | Intron 2 | C > G | 1 | 99.8 | 0.095 | 0.034 | *HapMap R28* |
| F5 | rs3753305 | Chr 1: 167820682 | Intron 1 | G > C | 0.0108 | 99.2 | 0.442 | 0.496 | *HapMap R28* |
| F7 | rs3093229 | Chr 13: 112805827 | 5' near gene | C > T | 0.6918 | 99.5 | 0.214 | 0.217 | *NCBI B36* |
| F7 | rs3093230 | Chr 13: 112805969 | 5' near gene | G > A | 0.6757 | 99.2 | 0.215 | 0.217 | *NCBI B36* |
| F7 | rs2774030 | Chr 13: 112808416 | Intron 1 | A > G | 0.6049 | 99.7 | 0.329 | 0.326 | *NCBI B36* |
| F7 | rs6041 | Chr 13: 112820708 | Intron 7 | G > A | 0.1449 | 98.9 | 0.105 | 0.118 | *HapMap R28* |
| F7 | rs6046 | Chr 13: 112821150 | Exon 8, Missense, Arg413Gln | G > A | 0.128 | 99.5 | 0.106 | 0.124 | *HapMap R28* |
| F9 | rs392959[a] | Chr X: 138449920 | Intron 3 | G > A | 2.77E-34 | 99.8 | 0.083 | 0.084 | *HapMap R28* |
| F9 | rs6048[a] | Chr X: 138460946 | Exon 6, Missense, Thr194Ala | T > C | 8.18E-46 | 99.5 | 0.257 | 0.317 | *HapMap R28* |
| F9 | rs413536[a] | Chr X: 138470175 | Intron 6 | T > C | 1.06E-34 | 100 | 0.179 | 0.173 | *HapMap R28* |

**Table 6.5. Details of 48 coagulation factor SNPs genotyped in 619 warfarin patients continued.**

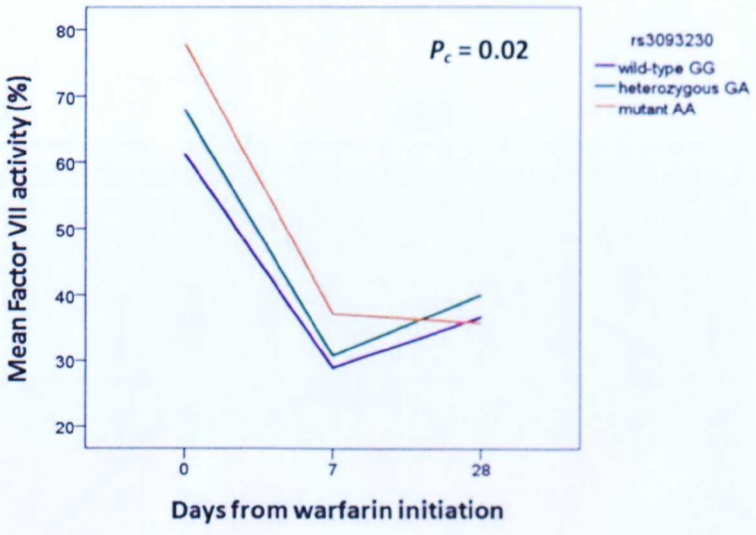
| Gene | SNP | Chromosomal Position | SNP Location and Function | SNP Alleles | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| F10 | rs3093261 | Chr 13: 112824083 | 5' near gene | C > T | 0.7077 | 99.8 | 0.444 | 0.496 | *HapMap R28* |
| F10 | rs563964 | Chr 13: 112824948 | 5' near gene | C > T | 0.2621 | 100 | 0.428 | 0.381 | *HapMap R28* |
| F10 | rs776905 | Chr 13: 112829943 | Intron 1 | A > C | 0.5215 | 98.2 | 0.1 | 0.125 | *HapMap R28* |
| F10 | rs693335 | Chr 13: 112832444 | Intron 2 | C > G | 0.2686 | 95.5 | 0.372 | 0.379 | *NCBI B36* |
| F10 | rs3211764 | Chr 13: 112840391 | Intron 2 | C > G | 0.5776 | 99.8 | 0.443 | 0.5 | *HapMap R28* |
| F10 | rs2251102 | Chr 13: 112840755 | Intron 2 | C > T | 0.2873 | 99.8 | 0.17 | 0.175 | *NCBI B36* |
| F10 | rs2026160 | Chr 13: 112840894 | Intron 2 | A > C | 0.0245 | 95.2 | 0.251 | 0.295 | *HapMap R28* |
| F10 | rs5960 | Chr 13: 112849738 | Exon 6, Synonymous, Thr264Thr | A > G | 0.5368 | 100 | 0.129 | 0.143 | *HapMap R28* |
| PROC | rs2069901 | Chr 2: 127891313 | 5' near gene | T > C | 1 | 99.8 | 0.432 | 0.383 | *NCBI B36* |
| PROC | rs1799809 | Chr 2: 127892345 | 5' near gene | A > G | 0.7464 | 100 | 0.432 | 0.392 | *NCBI B36* |
| PROC | rs2069910 | Chr 2: 127894444 | Intron 2 | C > T | 0.8474 | 98.9 | 0.439 | 0.348 | *NCBI B36* |
| PROC | rs2069919 | Chr 2: 127896023 | Intron 3 | G > A | 0.4773 | 99.8 | 0.303 | 0.317 | *NCBI B36* |
| PROC | rs5936 | Chr 2: 127897342 | Exon 4, Synonymous, Ser141Ser | T > G | 0.455 | 99.5 | 0.306 | 0.375 | *HapMap R28* |
| PROC | rs2069928 | Chr 2: 127900384 | Intron 6 | G > T | 1 | 99.5 | 0.183 | 0.159 | *NCBI B36* |
| PROC | rs2069933 | Chr 2: 127902158 | Intron 7 | C > T | 0.5106 | 99.4 | 0.305 | 0.392 | *HapMap R28* |
| PROS1 | rs9683303 | Chr 3: 95071254 | 3' near gene | A > T | 1 | 99.7 | 0.321 | 0.327 | *HapMap R28* |
| PROS1 | rs4857037 | Chr 3: 95110502 | Intron 4 | A > G | 0.0592 | 99.7 | 0.093 | 0.083 | *HapMap R28* |
| PROS1 | rs4857343 | Chr 3: 95127737 | Intron 2 | C > T | 0.5919 | 100 | 0.191 | 0.179 | *HapMap R28* |
| PROS1 | rs8178610 | Chr 3: 95129086 | Intron 1 | G > A | 0.2329 | 99.2 | 0.432 | 0.438 | *HapMap R28* |
| PROS1 | rs8178607 | Chr 3: 95136682 | Intron 1 | C > T | 0.1695 | 99.7 | 0.223 | 0.227 | *HapMap R28* |
| PROS1 | rs8178592 | Chr 3: 95163717 | Intron 1 | T > A | 1 | 97.9 | 0.116 | 0.093 | *HapMap R28* |

Frequency data were compiled from HapMap and NCBI dbSNP databases.

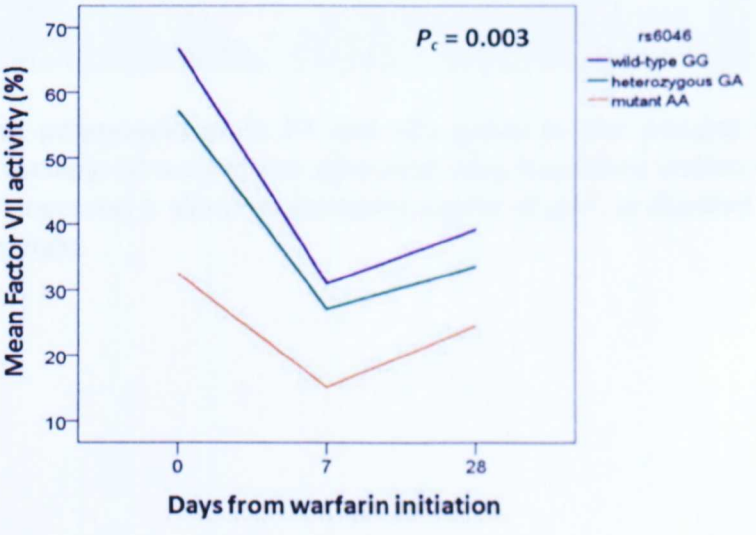
Chromosomal positions are given as per HapMap Data release 28, August 2010, NCBI B36 assembly, dbSNP b126.

[a] SNPs were not in HWE.

(a)




| Days from warfarin initiation | 0 | 7 | 28 |
|---|---|---|---|
| *rs3093230 genotype* | Mean Factor VII activity (%) | | |
| GG (n = 364) | 62 | 29 | 37 |
| GA (n = 205) | 68 | 31 | 40 |
| AA (n = 26) | 77 | 37 | 36 |




| Days from warfarin initiation | 0 | 7 | 28 |
|---|---|---|---|
| *rs6046 genotype* | Mean Factor VII activity (%) | | |
| GG (n = 473) | 67 | 31 | 39 |
| GA (n = 119) | 57 | 27 | 33 |
| AA (n = 3) | 33 | 16 | 25 |

**Figure 6.5. Line graph showing change in factor VII activity level for (a) rs3093230, and (b) rs6046 genotypes.** The tables beneath the graphs show the number of patients in each genotype group and their mean factor VII level at respective time points.

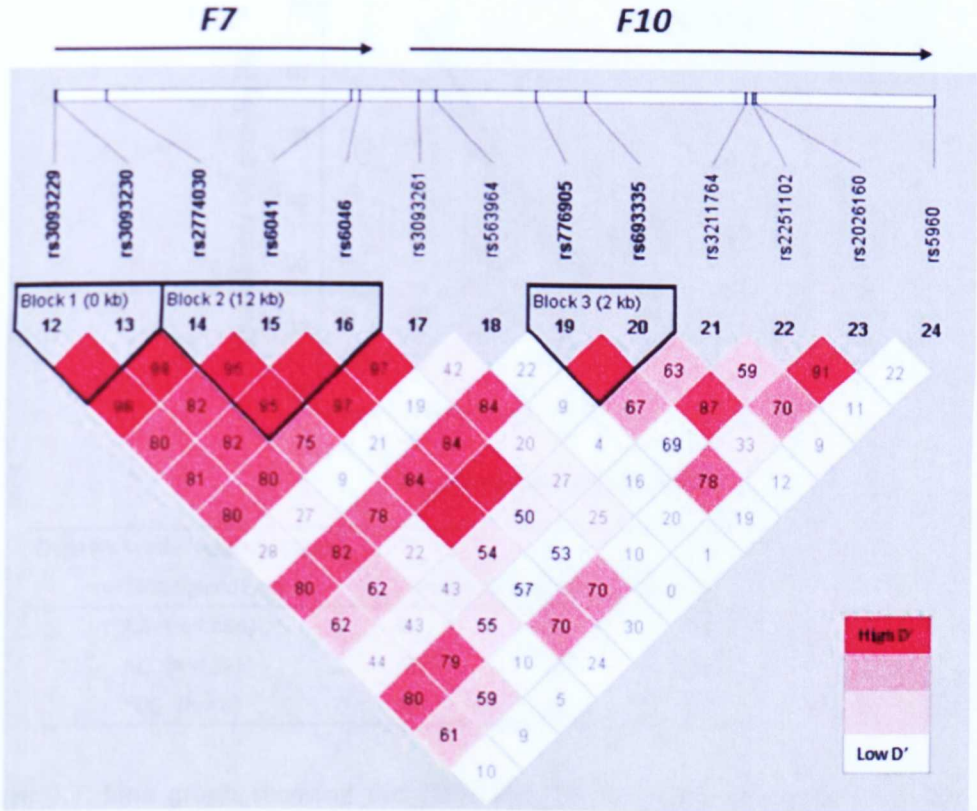
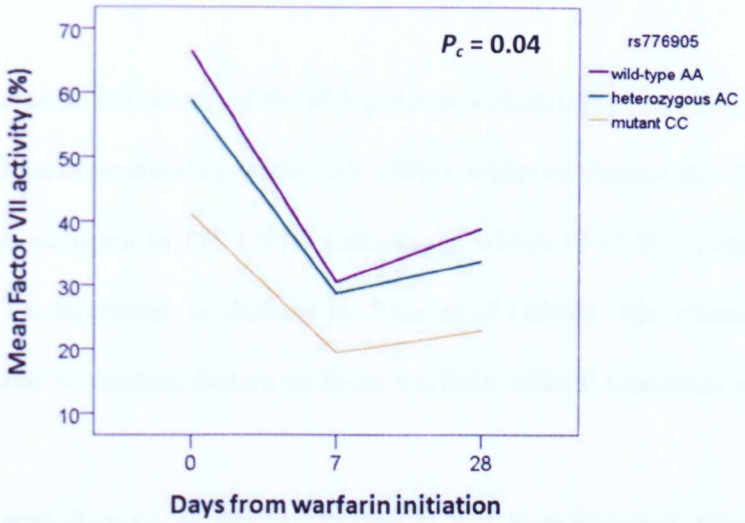



**Figure 6.6. LD among polymorphisms in F7 and F10 genes in 619 patients on prospective warfarin therapy.** LD pattern was generated using HaploView version 4.2 and strength of LD (D' measure) is shown in increasing shades of pink, as depicted by the bars on the bottom right.




| Days from warfarin initiation | 0 | 7 | 28 |
|---|---|---|---|
| rs776905 genotype | Mean Factor VII activity (%) | | |
| AA (n = 484) | 66 | 31 | 39 |
| AC (n = 111) | 60 | 29 | 34 |
| CC (n = 4) | 36 | 20 | 23 |

**Figure 6.7. Line graph showing the effect of *F10* rs776905 genotype on factor VII activity level.** The table beneath the graph shows the number of patients in each genotype group and their mean factor VII level at respective time points.

### 6.3.5 Association of changes in clotting factor activity levels with clinical outcomes

During the course of follow-up, of the 619 patients included in this study, 362 (58%) achieved warfarin dose stability while 558 (90%) achieved therapeutic INR. Bleeding complication occurred in 179 (29%) patients, of which 15 (2.4%) patients experienced a major haemorrhage, as defined by Fihn *et al* (1996). The effects of changes in activity levels of clotting factors on these warfarin clinical outcomes were examined.

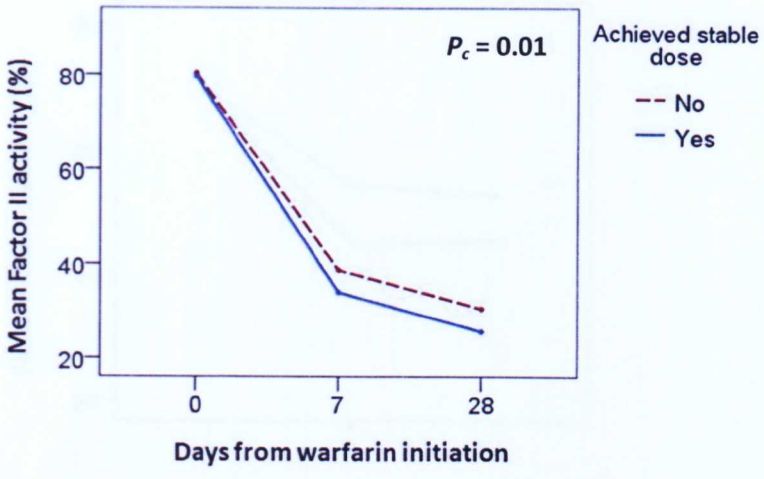
The activity level changes in clotting factors II and X and protein C were associated with the achievement of warfarin stable dose. At time points 7 and 28 days after warfarin initiation, the levels of factors II ($P_c$ = 0.01) and X ($P_c$ = 0.02), and protein C ($P_c$ = 0.003) were significantly lower in patients who achieved warfarin stable dose, compared to patients who did not achieve warfarin stable dose (Figure 6.8). After 7 and 28 days of warfarin treatment, the relative reduction in factors II and X were both ~5% greater among patients who achieved warfarin stable dose compared to those who did not achieve stable dose (Figures 6.8a and 6.8b), while that for protein C was nearly 10% greater (Figure 6.8c).

In patients who achieved therapeutic INR, the levels of factors II ($P_c$ = 0.01) and X ($P_c$ = 0.01) were significantly lower after 7 and 28 days of warfarin treatment (Figure 6.9), and the plasma activity levels of factors II and X both decreased ~10% faster in comparison with patients who did not achieve therapeutic INR (Figures 6.9a and 6.9b).

There was no significant association with any changes in clotting activity levels and bleeding complications.

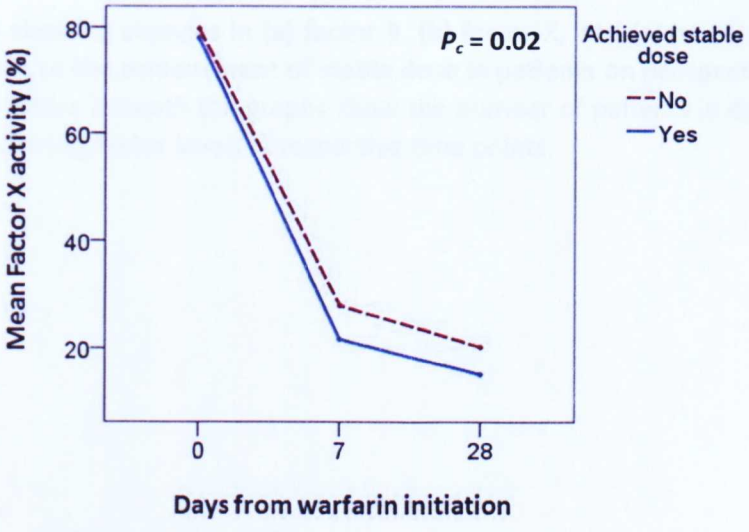
Given that the assessment performed in section 6.3.4 showed that the plasma levels of factors VII and X were different in patients with different genotype groups for *F7* and *F10* SNPs (rs3093230, rs6046 and rs776905), their relationships with clinical outcomes and clotting factor levels were examined. When patients were stratified according to their genotypes for the *F7* and *F10* SNPs, no association with changes in clotting factor levels was observed in the 3 clinical outcomes tested: achievement of warfarin stable dose, achievement of therapeutic INR and bleeding complications.

(a)



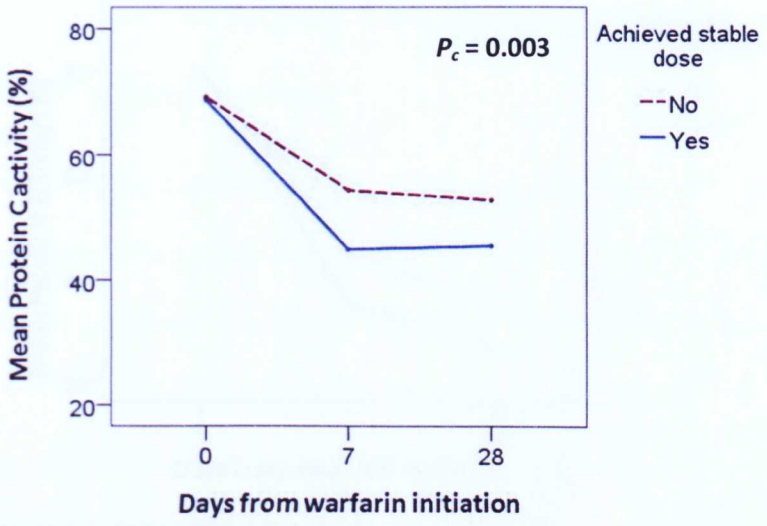

| Days from warfarin initiation | 0 | 7 | 28 |
|---|---|---|---|
| Achieved warfarin stable dose | Mean Factor II activity (%) | | |
| No (n = 244) | 81 | 39 | 30 |
| Yes (n = 356) | 80 | 34 | 26 |

(b)




| Days from warfarin initiation | 0 | 7 | 28 |
|---|---|---|---|
| Achieved warfarin stable dose | Mean Factor X activity (%) | | |
| No (n = 242) | 80 | 26 | 19 |
| Yes (n = 356) | 78 | 22 | 15 |

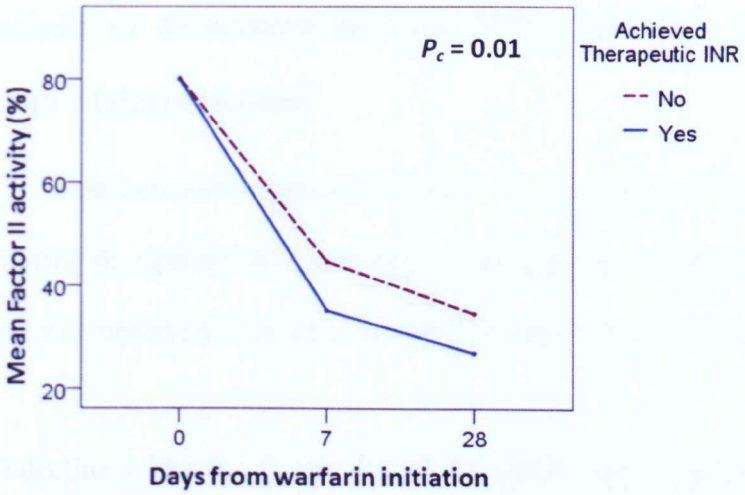
(c)




| Days from warfarin initiation | 0 | 7 | 28 |
|---|---|---|---|
| Achieved warfarin stable dose | Mean Protein C activity (%) | | |
| No (n = 247) | 70 | 52 | 52 |
| Yes (n = 355) | 69 | 45 | 45 |

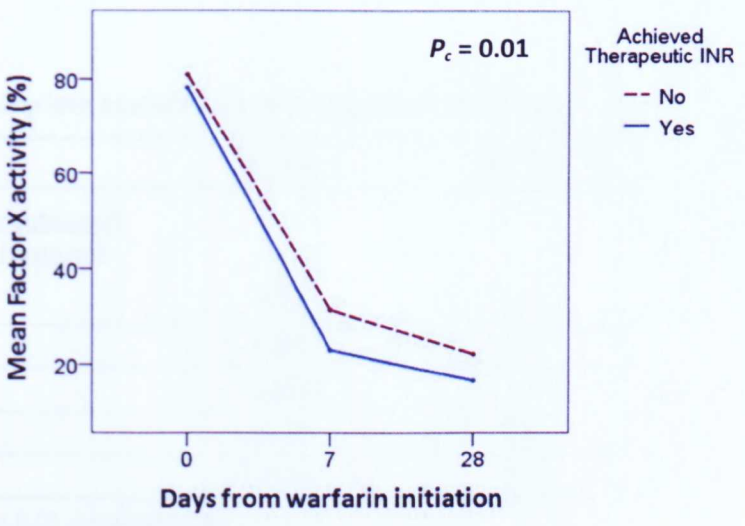
**Figure 6.7. Line graph showing changes in (a) factor II, (b) factor X, and (c) protein C activity levels in relation to the achievement of stable dose in patients on prospective warfarin therapy.** The tables beneath the graphs show the number of patients in each group and their mean clotting factor levels at respective time points.

(a)




| Days from warfarin initiation | 0 | 7 | 28 |
|---|---|---|---|
| Achieved therapeutic INR | Mean Factor II activity (%) | | |
| No (n = 56) | 81 | 46 | 35 |
| Yes (n = 544) | 80 | 35 | 27 |

(b)




| Days from warfarin initiation | 0 | 7 | 28 |
|---|---|---|---|
| Achieved therapeutic INR | Mean Factor X activity (%) | | |
| No (n = 54) | 82 | 31 | 22 |
| Yes (n = 544) | 78 | 23 | 16 |

**Figure 6.8. Line graph showing changes in (a) factor II and (b) factor X activity levels in relation to the achievement of therapeutic INR in patients on prospective warfarin therapy.** The tables beneath the graphs show the number of patients in each group and their mean clotting factor levels at respective time points.

### 6.3.6 Univariate analysis of association between SNPs, reduction in clotting factor levels and clinical outcomes

For the 3 time to event outcomes investigated (time to achieving warfarin stable dose, time to achieving therapeutic INR and time to bleeding complications), no significant association was observed with clotting factor levels and SNPs in the clotting factor genes.

Interestingly, the decline in clotting factors II, VII, IX, and X, and protein S, between 0 and 28 days of warfarin treatment, were significantly associated with the clinical outcome of warfarin stable dose. Table 6.6 details the *P*-values after FDR ($P_c$) and the variability explained by each variable. No clotting factor SNPs, however, showed significant association with the variability in warfarin stable dose.

**Table 6.6. Significant univariate associations with warfarin stable dose.**

| Variable | $P_c$-value | $R^2$ (%) |
|---|---|---|
| Difference in clotting factor between 0 and 28 days after warfarin treatment | | |
| Factor II | 0.02 | 2.2 |
| Factor VII | 0.003 | 4.0 |
| Factor IX | 0.004 | 3.3 |
| Factor X | 0.003 | 3.7 |
| Protein S | 0.02 | 2.2 |

Only those with a $P_c$-value < 0.05 are shown here.

362 patients achieved warfarin stable dose.

### 6.3.7 Multivariate model for warfarin stable dose

Table 6.7 outlines the clinical variables (age, BMI and gender) and genetic factors (*CYP2C9*2*, *CYP2C9*3*, *VKORC1* -1693 and *CYP4F2* rs2189784) which showed significant univariate association with warfarin stable dose requirement. A multiple regression model including these clinical and genetic covariates was built to assess their association with warfarin stable dose (Table 6.8). Models using the LRT both including and excluding covariates to represent the decline in clotting factors II, VII, IX, X and protein S were compared. Among these 5 covariates tested, the inclusion of a covariate representing the decline in clotting factor X explained the largest variability in warfarin stable dose requirement, increasing the $R^2$ value by 1.4% from 49.5% to 50.9%. A summary of the $R^2$ results is presented in Table 6.8.

**Table 6.7. Univariate associations of clinical and genetic variables with warfarin stable dose.**

| Variable | $P_c$ -value | $R^2$ (%) |
|---|---|---|
| Age | $3 \times 10^{-6}$ | 8.2 |
| BMI | $5 \times 10^{-5}$ | 6.6 |
| Gender | 0.01 | 3.0 |
| *CYP2C9*2* (rs1799853) | 0.01 | 3.9 |
| *CYP2C9*3* (rs1057910) | $3 \times 10^{-3}$ | 5.9 |
| *VKORC1* -1639 (rs9923231) | $3 \times 10^{-16}$ | 23.1 |
| *CYP4F2* rs2189784 | 0.01 | 3.8 |

Only those with a $P_c$-value < 0.05 are shown here.

362 patients achieved warfarin stable dose.

**Table 6.8. Contribution of variables to warfarin stable dose requirement.**

| Covariate | $P_c$ -value | $R^2$ (%) |
|---|---|---|
| Clinical only (age, BMI, gender) | $4 \times 10^{-11}$ | 15.1 |
| Genetics only (*CYP2C9*2* + *CYP2C9*3* + *VKORC1* -1639 + *CYP4F2* rs2189784) | $5 \times 10^{-15}$ | 32.7 |
| Clinical + genetics | $3 \times 10^{-23}$ | 49.5 |
| [a] Decline in clotting factors II, VII, IX, X and protein S | 0.028 | 3.4 |
| Clinical + genetics + [a] decline in clotting factors II, VII, IX, X and protein S | $7 \times 10^{-23}$ | 50.2 |
| Clinical + genetics + [a] decline in clotting factor II | $3 \times 10^{-23}$ | 50.4 |
| Clinical + genetics + [a] decline in clotting factor VII | $2 \times 10^{-23}$ | 50.5 |
| Clinical + genetics + [a] decline in clotting factor IX | $4 \times 10^{-23}$ | 50.2 |
| Clinical + genetics + [a] decline in clotting factor X | **$1 \times 10^{-23}$** | **50.9 *** |
| Clinical + genetics + [a] decline in clotting protein S | $6 \times 10^{-23}$ | 49.5 |

362 patients achieved warfarin stable dose.

[a] Decline was the difference in respective coagulation factor levels between day 0 and day 28 after warfarin treatment.

* Largest increase in $R^2$.

## 6.4 DISCUSSION

A clear advantage of our study is that this is the largest prospective study that has evaluated the coagulation profiles in patients on warfarin therapy to date.

In keeping with previous reports (Paul *et al.*, 1987; Kumar *et al.*, 1990; Jerkeman *et al.*, 2000), we found that the levels of the vitamin K-dependent procoagulation factors II, VII, IX and X, and anticoagulation proteins C and S, all declined in our cohort of patients in response to warfarin therapy, while that for factor V did not change significantly. Although the pattern of decline was similar among all 6 vitamin K-dependent clotting factors, the rate and magnitude of decrease was different, with factor X activity showing the greatest rate and magnitude of decline.

Also in agreement with previous studies (Kumar *et al.*, 1990; Lind *et al.*, 1997; Jerkeman *et al.*, 2000; Gulati *et al.*, 2011), our results showed that the levels of all 6 vitamin K-dependent clotting factors declined with increasing INR but again, not the rate and magnitude of decrease. Our observation of an apparent plateau effect at INR values greater than 4.0 is akin to findings of Sarode *et al.* (2006) and Gulati *et al.* (2011), who reported a poor correlation between supratherapeutic INR (> 5.0 or 3.6) and the levels of vitamin K-dependent clotting factors (Sarode *et al.*, 2006; Gulati *et al.*, 2011).

The effects of 48 candidate polymorphisms across the 7 clotting factor genes on the variability as well as the rate and magnitude of decline in clotting factor activity levels were investigated in our cohort of patients on prospective warfarin therapy. Several SNPs were found to contribute to the variability in factor VII plasma activity. The minor allele of the promoter SNP rs3093230 in the *F7* gene was associated with higher factor VII levels before and 7 days after warfarin treatment.

This promoter SNP has been reported to tag for a haplotype in the *F7* gene (Taylor *et al.*, 2011) and showed a similar association in previous studies conducted in healthy populations of European descent (Reiner *et al.*, 2007; Ken-Dror *et al.*, 2010; Taylor *et al.*, 2011), suggesting that rs3093230 constitutes a major determinant of factor VII activity. In addition, we found an association between lower factor VII levels and the minor allele of a nonsynonymous genetic polymorphism in exon 8 of *F7* gene, rs6046, which causes a p.Arg143Gln amino acid substitution. A recently reported meta-analysis in 1781 Caucasians found an association between warfarin dose variability and a synonymous SNP in exon 5 of *F7*, rs6042; this association was further replicated in an independent cohort of 693 Caucasian subjects, accounting for 0.7% of warfarin dose variance (Bourgeois *et al.*, unpublished). Interestingly, rs6042 which is approximately 3 kb upstream of rs6046, is in strong LD with rs6046 (D' = 1.0, $r^2$ = 0.92), suggesting that rs6046 could be the causal variant given its function. Further *in vitro* functional analysis on rs6046 should be carried out to confirm this hypothesis.

A genetic variation within intron 1 of *F10* gene (rs776905) also showed association with factor VII levels in our study, where patients carrying the minor allele had lower factor VII levels. Our results, together with a recent finding by Taylor and colleagues (2011) who discovered an association between factor VII levels and another intron 1 SNP in *F10* (rs3093268) in a healthy European American population (Taylor *et al.*, 2011), suggest that SNPs in *F10* play a role in the variability of factor VII levels.

Our findings indicate that patients who achieved stable dose experienced a faster decline in the anticoagulation factors II and X, and anticoagulant protein C levels, compared to patients who did not achieve stable dose. Additionally, in

patients who achieved therapeutic INR during the study period, a higher reduction rate in the procoagulation factors II and X was observed, when compared to patients who did not achieve therapeutic INR. Although the difference in clotting factor levels between patients who achieved stable dose or therapeutic INR and those who failed to achieve effective anticoagulation was only 5 to 10%, this small difference in levels appears to have a great impact on patients' response to warfarin. Although, the concomitant measurement of INR and factors II and X, and protein C levels, may improve warfarin control, the expense would be prohibitive.

The decline in clotting factors II, VII, IX, X and protein S after 28 days of warfarin treatment were found to be predictors of warfarin maintenance dose. A multiple regression model using the decrease in clotting factor X in addition to the predictors *CYP2C9*2*, *CYP2C9*3*, *VKORC1* -1639, *CYP4F2* rs2189784, age, BMI and gender explained almost 51% of variance in warfarin stable dose, with the decline in clotting factor X accounting for 1.4% of dose variability.

A limitation of our study is that SNPs were selected by candidate gene approach. It is possible that genetic polymorphisms in other genes influence changes in clotting factor levels, as demonstrated in recent studies which reported the associations of hepatocyte nuclear factor 4 (*HNF4*) and protein C receptor (*PROCR*) genes with factor VII levels in healthy European populations (Smith *et al.*, 2010; Taylor *et al.*, 2011). Furthermore, none of the *F9* SNPs selected in our study were included in the association tests due to deviation from HWE (which could reflect genotyping errors or suggest a genuine genetic association). We therefore could not determine if genetic variations in *F9* affect factor IX levels or if they are related to bleeding complications. To further characterise the genotype-phenotype relationships

for all clotting factors investigated in this study, genome-wide scan data should be explored and fine-mapping of any discovered associations should be conducted.

# CHAPTER 7

# Final discussion

# CONTENTS

**7.1 PROGRESS OF WARFARIN PHARMACOGENETICS ..... 228**

**7.2 OUR FINDINGS AND FUTURE DIRECTIONS ..... 232**

7.2.1 Effect of genes in the *CYP4F* cluster on warfarin response ..... 233

7.2.2 Warfarin resistance and *VKORC1* mutations ..... 235

7.2.3 Changes in clotting factor levels during warfarin treatment ..... 237

**7.4 CONCLUSIONS ..... 239**

## 7.1 Progress of warfarin pharmacogenetics

Warfarin was approved for medical use as early as the 1950s (Clatanoff *et al.*, 1954), but it was not until more than 4 decades later that experimental evidence of the effects of *CYP2C9* genetic polymorphisms on warfarin dose requirement and warfarin-related bleeding incidence were published (Furuya *et al.*, 1995; Steward *et al.*, 1997; Aithal *et al.*, 1999). These observations were confirmed by many studies but it quickly became clear that the presence of *CYP2C9* polymorphisms did not explain most of the variation in warfarin dosage.

In 2004, the gene coding for the target enzyme of warfarin, *VKORC1*, was discovered (Li *et al.*, 2004; Rost *et al.*, 2004a) and soon after, SNPs in *VKORC1* were shown to be associated with warfarin dose variance (Rieder *et al.*, 2005). Taken together with the completion of the human genome project in 2001 and the advent of genomic technologies, warfarin has been the "poster child" for integrating pharmacogenetics into clinical practice. A summary of the progress in warfarin pharmacogenetics is illustrated in Figure 7.1.

Many observational studies reported that the combination of both *CYP2C9* and *VKORC1* genotypes explain 30-40% of total variation in warfarin dose. In August 2007, the FDA updated the label of warfarin to include the statement that "lower initiation doses should be considered for patients with certain genetic variations in CYP2C9 and VKORC1 enzymes", and in 2010, introduced dosing tables into the drug label.

A functional SNP, rs2108622, in *CYP4F2* was first reported to affect warfarin dose requirement in 2008 (Caldwell *et al.*, 2008) and CYP4F2 was subsequently shown to catalyse vitamin K oxidation (McDonald *et al.*, 2009).

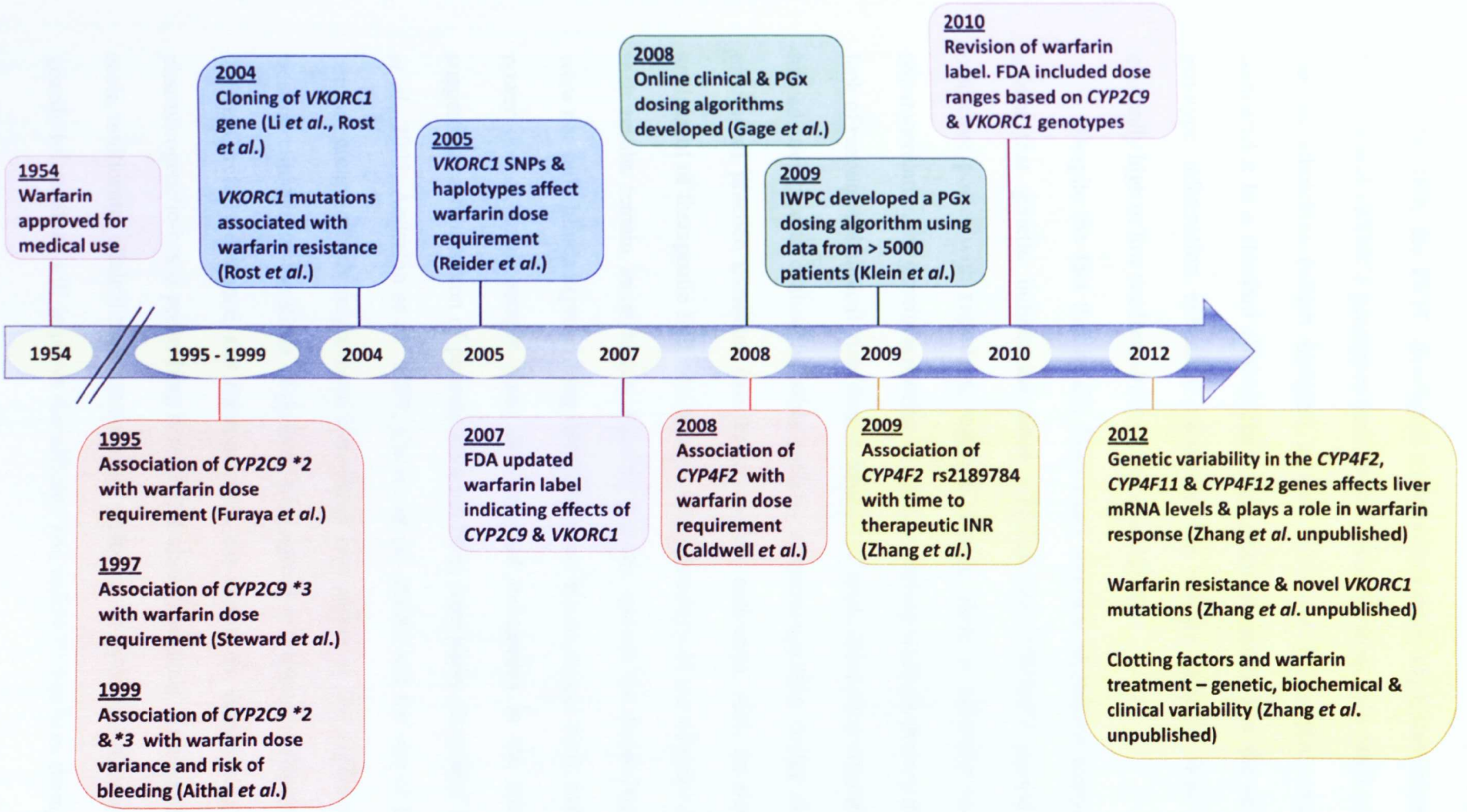



**Figure 7.1 Timeline of warfarin pharmacogenetics.** PGx: Pharmacogenetics; IWPC: International Warfarin Pharmacogenetics Consortium.

In 2009, the IWPC developed a pharmacogenetic algorithm incorporating *CYP2C9* and *VKORC1* genotypes and clinical variables from over 5000 patients of various ethnicities (which included patients from our prospective cohort) and compared it to a standard clinical algorithm. They concluded that the addition of genotype information enhanced outcomes, but not for patients who required unusually high or low warfarin doses (Klein *et al.*, 2009).

Despite the fact that results from many multiple regression analyses have shown that genetic information from *CYP2C9* and *VKORC1* provides good predictive power with regards to warfarin dosage, there is currently no general recommendation for genetic screening of patients starting warfarin therapy due to the lack of randomised clinical trials data. A handful of randomised controlled trials have indeed attempted to evaluate whether applying pharmacogenetics dosing algorithms to clinical practice translates into better clinical outcomes, such as more rapid attainment of therapeutic INR or a reduction in percentage of out-of-range INR, but their results remain inconclusive. Reasons for this include the following: studies were not fully blinded (prescribing clinicians knew the treatment arm), had limited power due to small sample sizes, demonstrated imbalances in the study arms suggesting randomisation of patients have not been completely successful (Hillman *et al.*, 2005; Anderson *et al.*, 2007; Caraco *et al.*, 2008), and the use of historical control groups which might have introduced bias either in the vigilance of the treating clinician or the kinds of patients who agreed to participate (Epstein *et al.*, 2010). Carefully designed and rigorous trials are critical to determining whether pharmacogenetic-based prescribing is ready for clinical practice. At least five large-scale, multicentre clinical trials are currently underway to determine if knowledge of genetic information will improve the efficacy and safety of warfarin therapy. They

include a single-blinded and randomized controlled trial, the European Pharmacogenetics of Anticoagulant Therapy (EU-PACT) study (in which our research team is taking part) (van Schie *et al.*, 2009), the Classification of Optimal Anticoagulation through Genetics (COAG) study (French *et al.*, 2010), and the Genetics Informatics Trial (GIFT) of warfarin to prevent deep vein thrombosis study (Do *et al.*, 2011). Even then, the cost-effectiveness of pharmacogenetic-based prescribing needs to be determined. Some estimates are that genetic-based dosing will be cost-effective, but this needs to be evaluated formally within a randomized trial (You *et al.*, 2004; McWilliam *et al.*, 2006) rather than many of the suppositions that have been made using retrospective data to construct the economic models that have been published. There are also arguments that clinical-based algorithms that do not require genetic information may need to be compared with genetic-based algorithms because, if the former work as well as the latter, they will be much more (You *et al.*, 2004; Kimmel 2008) cost-effective.

## 7.2 Our findings and future directions

Many studies have shown that other genetic and clinical factors play a role in warfarin response, albeit with small effects. Our findings demonstrated that the variability in warfarin response involves a complex interplay of genetic, clinical and biochemical factors (as depicted in Figure 7.2).

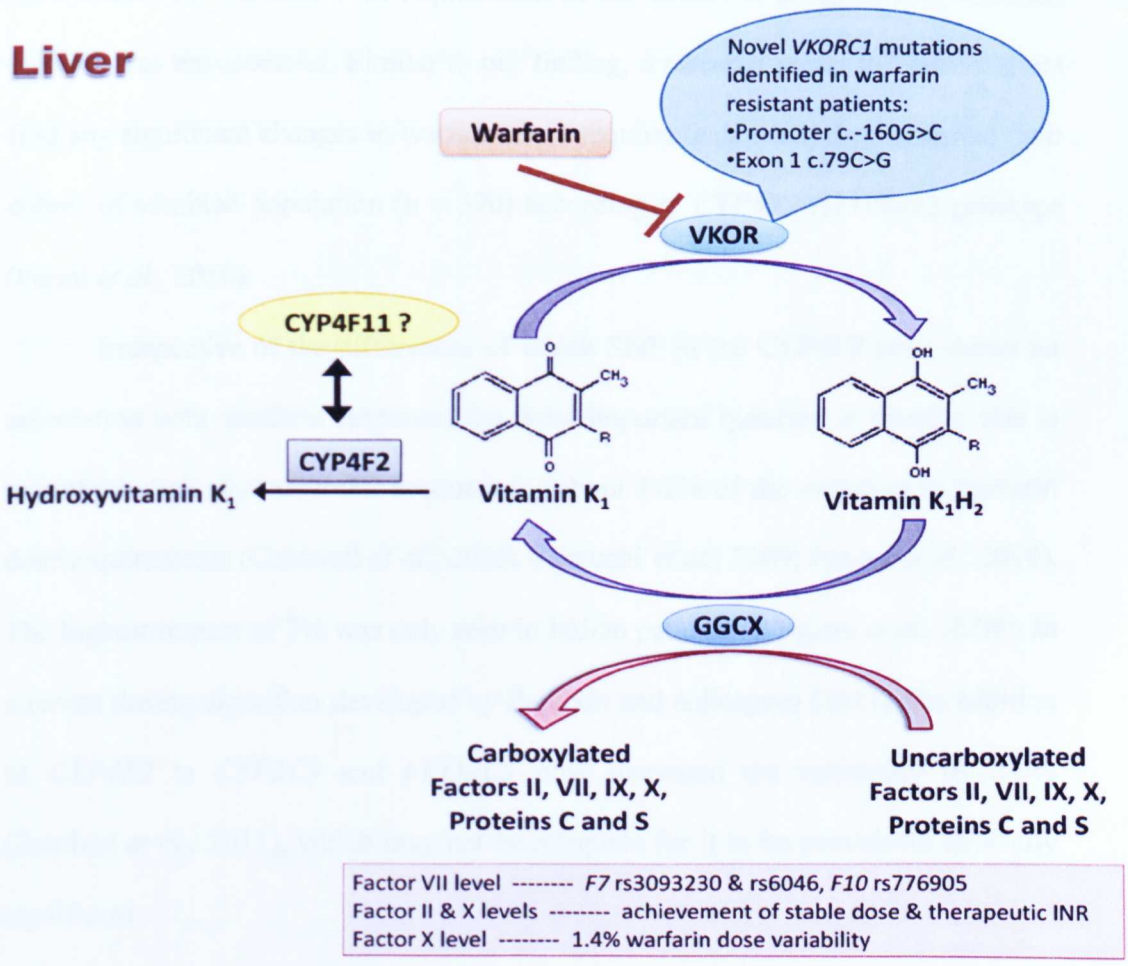



**Figure 7.2. Summary of our findings.**

### 7.2.1 Effect of genes in the *CYP4F* cluster on warfarin response

Although we found an association between a SNP, rs2189784, which is in LD with *CYP4F2* rs2108622 to be associated with time to therapeutic INR, our results in chapters 2 and 3 highlighted that inadequate genetic assessment, which unfortunately is common in the pharmacogenetic literature, can lead to false positive finding, and inability to replicate. In chapter 2, our attempt to replicate the association of *CYP4F2* rs2108622 with warfarin dose requirement in our cohort of prospectively recruited patients was unsuccessful. Similar to our finding, a research group in Brazil did not find any significant changes in warfarin dose requirements when they stratified their cohort of admixed population (n = 370) according to *CYP4F2* rs2108622 genotype (Perini *et al.*, 2010).

Irrespective of the differences of which SNP in the *CYP4F2* gene shows an association with warfarin response, the most important question is whether this is important clinically? *CYP4F2* accounts for about 1-2% of the variation in warfarin dose requirements (Caldwell *et al.*, 2008; Takeuchi *et al.*, 2009; Pautas *et al.*, 2010). The highest impact of 7% was only seen in Italian patients (Borgiani *et al.*, 2009). In a recent dosing algorithm developed by Zambon and colleagues (2011), the addition of *CYP4F2* to *CYP2C9* and *VKORC1* only increased the variability by 2-5% (Zambon *et al.*, 2011), which may not be adequate for it to be considered clinically significant.

The work presented in chapter 3 shows the complexity of gene-gene interactions, where competing effects of different SNPs within the same gene cluster can cancel out the level of *CYP4F2* mRNA and warfarin daily doses required to maintain anticoagulation. Through an international collaboration with a Professor Uli Zanger's research team in Stuttgart, Germany, we observed significant association

between *CYP4F2* SNPs and *CYP4F2* mRNA expression in human liver, a finding which was not previously found by McDonald and colleagues (2009). Our sample size is at least 2.5 times bigger than that by McDonald *et al.* (2009), highlighting the importance of sample size in detecting an effect size. Interestingly, in addition to increasing *CYP4F2* mRNA expression, the rs2108622 T allele was also associated with decreasing *CYP4F11* mRNA expression. Moreover, an association between *CYP4F11* rs1060467 (a SNP in LD with rs2108622) and down-regulation of *CYP4F2* mRNA expression was also observed. By segregating the patients according to their haplotypes for *CYP4F2* rs2108622 and *CYP4F11* rs1060467 as illustrated in Table 7.1, it can be clearly seen that there was minimal dose changes in patients carrying the haplotype consisting of *CYP4F2* rs2108622 homozygous wild-type genotype and *CYP4F11* rs1060467 homozygous mutant genotype and vice versa. Taken together, given the low level of dose variability attributable to SNPs in *CYP4F2* and *CYP4F11*, and their competing effects, the *CYP4F* cluster is highly unlikely to have a significant effect on the actual warfarin dose required by patients. Indeed, unless several of the SNPs in the CYP4F cluster are measured simultaneously, there may be errors in the calculation of warfarin dose requirement.

**Table 7.1. Effects of *CYP4F2* and *CYP4F11* SNPs on stable warfarin dose requirement*.**

| Warfarin dose (mg/day) | | | |
|---|---|---|---|
| ***CYP4F2* rs2108622 genotype** | | ***CYP4F11* rs1060467 genotype** | |
| | | TT (n) | CT and TT (n) |
| CC (n) | 3.7 (166) | 3.9 (18) | 3.7 (148) |
| CT and TT (n) | 4.5 (186) | 4.7 (92) | 4.2 (94) |

* Out of 714 patients we recruited prospectively, 352 achieved warfarin stable dose.

### 7.2.2 Warfarin resistance and *VKORC1* mutations

In chapter 4, although we could not determine the status of pharmacodynamic resistance for all the 65 clinically defined warfarin resistant patients in our cohort, 7 novel heterozygous *VKORC1* mutations were identified via Sanger sequencing. Of these, c.-160G>C and c.79C>G were predicted to have potential functional effects. *In vitro* studies carried out in chapter 5 showed that the promoter c.-160G>C mutation led to 20% increase in promoter activity, which could possibly explain the higher warfarin dose required by the patient to achieve therapeutic anticoagulation. Our EMSA results, however, could not confirm if this promoter mutation creates a binding site for Sp1 transcription factor. Whether this promoter mutation introduces binding sites for other transcription factors remains to be elucidated. Using the PIVKA-II ELISA kit, we were unable to clearly define the functional role of the exon 1 missense mutation c.79C>G. Further experiments will be required and assessment of the VKOR enzymatic activity by HPLC (Wallin and Martin 1985) may be an alternative to clarify the functional implication of this missense mutation.

We have also investigated the role of epigenetics in warfarin resistance. To date, we are the first group who have explored the effect of DNA methylation at *VKORC1* CpG sites in relation to warfarin resistance. Although we are aware of the limitations of our study and that our interpretation of the results is conservative, we can speculate that mutations and SNPs in *VKORC1* down-regulate methylation expression at several CpG sites. It is important to mention that due to the rarity of the *VKORC1* mutations evaluated in chapter 5, the number of patients carrying the heterozygous genotype was very small (n = 1 or 2), making it difficult to explore their association with methylation levels.

Studies on warfarin resistance pharmacogenetics have largely concentrated on the *VKORC1* gene except one recent report which found a coding polymorphism (rs2290228) in the gene *CALU* to be related to exceptionally high warfarin dose (up to 20 mg/day) (Vecsler *et al.*, 2006). *CALU* codes for the endoplasmic reticulum chaperone protein calumenin, and results from rat studies suggest that calumenin binds to the VKOR enzyme complex and inhibits the vitamin K cycle (Wallin *et al.*, 2001; Wajih *et al.*, 2004). It is therefore likely that *CALU* may confer warfarin resistance. Further studies will be required to confirm this. In addition, the relation of other genes in the human genome to warfarin resistance should also be explored.

As none of the previously reported missense mutations were present in our cohort of patients and given the mutations identified in our population appear to be "private" mutations, a larger study will be required to determine the prevalence of these mutations before the clinical utility of these mutations in warfarin resistant patients could be evaluated.

Although we have only focussed on patients resistant to warfarin in this study, it is equally important to investigate patients who are sensitive to warfarin as over-dosing in these patients could lead to bleeding complications. We are currently actively recruiting patients with discordant phenotypes, those who are sensitive and resistant to warfarin. Our initial plan is to perform exome genotyping in search of novel genetic markers associated with warfarin resistance and sensitivity. Further work will include either exome or whole genome sequencing to identify underlying rare mutations, which could hopefully improve the clinical management of patients who require unusually high or low doses of warfarin.

### 7.2.3 Changes in clotting factor levels during warfarin treatment

Our work presented in chapter 6 demonstrated that the interactions between clotting factor genes and their plasma activity levels play a role in warfarin response, where we showed for the first time that SNPs in clotting factor genes *F7* and *F10* affect the variability in factor VII levels in patients on warfarin therapy. It is plausible that rs6046, which is in strong LD with rs6042, could be the true causal variant accounting for the 0.7% variability in warfarin dose requirements reported by a recent meta-analysis (Bourgeois *et al.*, unpublished).

Our findings also highlighted that small biochemical changes have an impact on warfarin response where the change in clotting factor X plasma levels contributed to 1.4% of warfarin dose variance and changes in factor II and X levels affect the achievements of warfarin stable dose and therapeutic INR. Future studies of gene-gene interactions and gene-clinical/biochemical interactions might be the key to improving our ability to accurately predict warfarin doses.

One caveat of our study is that we have only concentrated on selected SNPs in the clotting factor genes. Fine mapping of the clotting factor genes and their haplotype structures may expand our understanding on the genotype-phenotype relationships of variability in clotting factor levels and warfarin clinical outcomes. Further work at the genome-wide level will help in elucidating the relationships of other genes with variability in clotting factor levels before and after warfarin treatment.

One aspect which we did not explore in chapter 6 is whether patients with extremely low clotting factor levels are prone to bleeding complications. The risk of bleeding is higher when INR is over 3, and high INR values are usually associated with extremely low levels of clotting factors. However, bleeding can also occur when

the INR is within the therapeutic range (Fanikos *et al.,* 2005) and we do not know if this is reflected in the levels of clotting factors. With the wealth of clinical and biochemical data we have collected, we will be able to investigate this and determine if the levels of clotting factors can serve as biomarkers to prevent bleeding complications.

One also must not forget that adherence remains, perhaps, the "elephant in the room". Innovative, sustainable, and cost-effective strategies to improve adherence could make warfarin safer and more effective for existing patients and expand the number of patients who could benefit but who do not currently receive the medication due to concerns about adherence. The potential public health impact of a successful intervention to improve warfarin adherence among the millions of patients who require preventive therapy for thromboembolism is enormous.

New oral anticoagulants such as the thrombin inhibitor dabigatran (Connolly *et al.,* 2009) and the factor Xa inhibitor rivaroxaban (Cleland *et al.,* 2011) have been shown to be equally or more effective than warfarin and because their anticoagulant effects are much more predictable, there is no need for monitoring. However, they are much more costly compared to warfarin and there is no available pharmacodynamic biomarker and nor is there an antidote. The latter in particular has led to concerns with dabigatran about bleeding risks (Eikelboom *et al.,* 2011). Whether these new anticoagulants will supersede warfarin is unclear. Perhaps a stratified prescription approach for warfarin, factor Xa inhibitor, thrombin inhibitor, and heparin, may maximise the clinical effectiveness of anticoagulation.

## 7.4 Conclusions

Warfarin, being one of the most widely prescribed drugs and having a narrow therapeutic window, seems to be an ideal candidate for the application of the concept of personalised medicine in which an individual can be prescribed with a specific regimen matching their particular genetic makeup so that the benefit of treatment can be maximised while minimising complications. Better understanding of the individual genetic polymorphisms contributing to the variations in response of warfarin has enabled us to see beyond the mere biochemical aspects of warfarin dosing by trial and error. However, a multitude of factors affect warfarin dosing, some of which still elude our knowledge.

In conclusion, this thesis demonstrates that the control of daily warfarin dose and its anticoagulant effect is complex, and many factors including common and rare genetic variants, together with clinical covariates, will be responsible for the missing heritability. Whether such complexity can ever be incorporated into clinical practice is unclear, and will depend on the advances that are made in sequencing, other –omics technologies, integrative biology and decision support systems. Because there is relatively robust phenotype by which variability in response to warfarin can be assessed, further work on warfarin is still worthwhile, as it will provide valuable lessons for other drugs that are widely used in patients.

# BIBLIOGRAPHY


Absher, R.K., Moore, M.E., Parker, M.H., Rivera-Miranda, G. and Perreault, M.M. (2002). Patient-specific factors predictive of warfarin dosage requirements. *Annals of Pharmacotherapy.* **36** (10), 1512-1517.

Ageno, W., Johnson, J., Nowacki, B. and Turpie, A.G. (2000). A computer generated induction system for hospitalized patients starting on oral anticoagulant therapy. *Thrombosis and Haemostasis.* **83** (6), 849-52.

Ainle, F.N., Mumford, A., Tallon, E., Mccarthy, D. and Murphy, K. (2008). A vitamin K epoxide reductase complex subunit 1 mutation in an Irish patient with warfarin resistance. *Ir J Med Sci.* **177** (2), 159-61.

Aithal, G.P., Day, C.P., Kesteven, P.J. and Daly, A.K. (1999). Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. *Lancet.* **353**, 717-719.

Altshuler, D.L., Durbin, R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Collins, F.S., De La Vega, F.M., Donnelly, P., Egholm, M., *et al.* (2010). A map of human genome variation from population-scale sequencing. *Nature.* **467** (7319), 1061-1073.

Alving, B.M., Strickler, M.P. and Knight, R.D. (1985). Hereditary warfarin resistance. Investigation of a rare phenomenon. *Archives of internal medicine.* **145** (3), 499-501.

Ames, P.R., Nourooz-Zadeh, J., Tommasino, C., Alves, J., Brancaccio, V. and Anggard, E.E. (1998). Oxidative stress in primary antiphospholipid syndrome. *Thromb Haemost.* **79** (2), 447-9.

Ames, P.R., Margaglione, M., Ciampa, A., Colaizzo, D., Ferrara, F., Iannaccone, L. and Vincenzobrancaccio (2011). Increased warfarin consumption and residual fibrin turnover in thrombotic patients with primary antiphospholipid syndrome. *Thromb Res.* **127** (6), 595-9.

Anderson, J.L., Horne, B.D., Stevens, S.M., Grove, A.S., Barton, S., Nicholas, Z.P., Kahn, S.F.S., May, H.T., Samuelson, K.M., Muhlestein, J.B. and Carlquist, J.F. (2007). Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. *Circulation.* **116** (22), 2563-2570.

Anon Health Care Guideline Institute for Clinical Systems Improvement Web site Antithrombotic therapy supplement (guideline) September2007 ed.

Ansell, J., Hirsh, J., Dalen, J., Bussey, H., Anderson, D., Poller, L., Jacobson, A., Deykin, D. and Matchar, D. (2001). Managing oral anticoagulant therapy. *Chest.* **119** (1 SUPPL.), 22S-38S.

Aquilante, C.L., Langaee, T.Y., Lopez, L.M., Yarandi, H.N., Tromberg, J.S., Mohuczy, D., Gaston, K.L., Waddell, C.D., Chirico, M.J. and Johnson, J.A. (2006). Influence of coagulation factor, vitamin K epoxide reductase complex subunit 1, and cytochrome P450 2C9 gene polymorphisms on warfarin dose requirements. *Clinical Pharmacology and Therapeutics.* **79** (4), 291-302.

Assenat, E., Gerbal-Chaloin, S., Larrey, D., Saric, J., Fabre, J.M., Maurel, P., Vilarem, M.J. and Pascussi, J.M. (2004). Interleukin 1β inhibits CAR-induced expression of hepatic genes involved in drug and bilirubin clearance. *Hepatology.* **40** (4), 951-960.

Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics.* **21** (2), 263-265.

Bell, R.G., Sadowski, J.A. and Matschiner, J.T. (1972). Mechanism of action of warfarin. Warfarin and metabolism of vitamin K1. *Biochemistry.* **11** (10), 1959-1961.

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behavioural brain research.* **125** (1-2), 279-84.

Berkner, K.L. (2000). The vitamin K-dependent carboxylase. *Journal of Nutrition.* **130** (8), 1877-1880.

Berkner, K.L. and Runge, K.W. (2004). The physiology of vitamin K nutriture and vitamin K-dependent protein function in atherosclerosis. *Journal of Thrombosis and Haemostasis.* **2** (12), 2118-2132.

Bertina, R.M., Broekmans, A.W., Van Der Linden, I.K. and Mertens, K. (1982). Protein C deficiency in a Dutch family with thrombotic disease. *Thrombosis and haemostasis.* **48** (1), 1-5.

Beyth, R.J., Quinn, L. and Landefeld, C.S. (2000). A multicomponent intervention to prevent major bleeding complications in older patients receiving warfarin: A randomized, controlled trial. *Annals of internal medicine.* **133** (9), 687-695.

Bird, A.P. (1986). CpG-rich islands and the function of DNA methylation. *Nature.* **321** (6067), 209-13.

Bodin, L., Verstuyft, C., Tregouet, D.A., Robert, A., Dubert, L., Funck-Brentano, C., Jaillon, P., Beaune, P., Laurent-Puig, P., Becquemont, L. and Loriot, M.A. (2005a). Cytochrome P450

2C9 (CYP2C9) and vitamin K epoxide reductase (VKORC1) genotypes as determinants of acenocoumarol sensitivity. *Blood.* **106** (1), 135-140.

Bodin, L., Horellou, M.H., Flaujac, C., Loriot, M.A. and Samama, M.M. (2005b). A vitamin K epoxide reductase complex subunit-1 (VKORC1) mutation in a patient with vitamin K antagonist resistance. *J Thromb Haemost.* **3** (7), 1533-5.

Boks, M.P., Derks, E.M., Weisenberger, D.J., Strengman, E., Janson, E., Sommer, I.E., Kahn, R.S. and Ophoff, R.A. (2009). The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. *PLoS One.* **4** (8), e6767.

Booth, S.L. and Al Rajabi, A. (2008). Determinants of vitamin K status in humans. *Vitam Horm.* **78**, 1-22.

Borgiani, P., Ciccacci, C., Forte, V., Sirianni, E., Novelli, L., Bramanti, P. and Novelli, G. (2009). CYP4F2 genetic variant (rs2108622) significantly contributes to warfarin dosing variability in the Italian population. *Pharmacogenomics.* **10** (2), 261-6.

Boulanger, L., Kim, J., Friedman, M., Hauch, O., Foster, T. and Menzin, J. (2006). Patterns of use of antithrombotic therapy and quality of anticoagulation among patients with non-valvular atrial fibrillation in clinical practice. *Int J Clin Pract.* **60** (3), 258-64.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* **72**, 248-54.

Breckenridge, A., Orme, M. and Wesseling, H. (1974). Pharmacokinetics and pharmacodynamics of the enantiomers of warfarin in man. *Clinical Pharmacology and Therapeutics.* **15** (4), 424-430.

Breckenridge, A. (1978). Oral anticoagulant drugs: Pharmacokinetic aspects. *Seminars in Hematology.* **15** (1), 19-26.

Brian, W.R. (1990). Catalytic Activities of Human Liver Cytochrome P-450 IIIA4 Expressed in Saccharomyces cerevisiae. *Biochemistry®.* **29** (51), 11280-11292.

Broze G.J, Jr. (2001). Protein Z-dependent regulation of coagulation. *Thrombosis and Haemostasis.* **86** (1), 8-13.

Brunak, S., Engelbrecht, J. and Knudsen, S. (1991). Prediction of human mRNA donor and acceptor sites from the DNA sequence. *J Mol Biol.* **220** (1), 49-65.

Bu, Y. and Gelman, I.H. (2007). v-Src-mediated down-regulation of SSeCKS metastasis suppressor gene promoter by the recruitment of HDAC1 into a USF1-Sp1-Sp3 complex. *J Biol Chem.* **282** (37), 26725-39.

Budnitz, D.S., Shehab, N., Kegler, S.R. and Richards, C.L. (2007). Medication use leading to emergency department visits for adverse drug events in older adults. *Annals of internal medicine.* **147** (11), 755-65.

Burton, P.R., Clayton, D.G., Cardon, L.R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D.P., Mccarthy, M.I., Ouwehand, W.H., Samani, N.J., Todd, J.A., Donnelly, P., Barrett, J.C., Davison, D., Easton, D., Evans, D., Leung, H.T., Marchini, J.L., Morris, A.P., Spencer, C.C.A. and Tobin, M.D. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* **447** (7145), 661-678.

Bylund, J., Finnstrom, N. and Oliw, E.H. (1999). Gene expression of a novel cytochrome P450 of the CYP4F subfamily in human seminal vesicles. *Biochemical and biophysical research communications.* **261** (1), 169-74.

Bylund, J., Hidestrand, M., Ingelman-Sundberg, M. and Oliw, E.H. (2000). Identification of CYP4F8 in human seminal vesicles as a prominent 19-hydroxylase of prostaglandin endoperoxides. *J Biol Chem.* **275** (29), 21844-9.

Bylund, J., Bylund, M. and Oliw, E.H. (2001). cDNA cloning and expression of CYP4F12, a novel human cytochrome P450. *Biochemical and biophysical research communications.* **280** (3), 892-7.

Cain, D., Hutson, S.M. and Wallin, R. (1997). Assembly of the warfarin-sensitive vitamin K 2,3-epoxide reductase enzyme complex in the endoplasmic reticulum membrane. *Journal of Biological Chemistry.* **272** (46), 29068-29075.

Cain, D., Hutson, S.M. and 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** (1), 128-33.

Calabro, V., Cascone, A., Malaspina, P. and Battistuzzi, G. (1989). Glucose-6-phosphate dehydrogenase (G6PD) deficiency in southern Italy: a case of G6PD A(-) associated with favism. *Haematologica.* **74** (1), 71-3.

Caldwell, M.D., Berg, R.L., Kai, Q.Z., Glurich, I., Schmelzer, J.R., Yale, S.H., Vidaillet, H.J. and Burmester, J.K. (2007). Evaluation of genetic factors for warfarin dose prediction. *Clinical Medicine and Research.* **5** (1), 8-16.

Caldwell, M.D., Awad, T., Johnson, J.A., Gage, B.F., Falkowski, M., Gardina, P., Hubbard, J., Turpaz, Y., Langaee, T.Y., Eby, C., King, C.R., Brower, A., Schmelzer, J.R., Glurich, I., Vidaillet, H.J., Yale, S.H., Qi Zhang, K., Berg, R.L. and Burmester, J.K. (2008). CYP4F2 genetic variant alters required warfarin dose. *Blood.* **111** (8), 4106-12.

Campbell, H.A. and Link, K.P. (1941). STUDIES ON THE HEMORRHAGIC SWEET CLOVER DISEASE. *Journal of Biological Chemistry.* **138** (1), 21-33.

Cappellini, M.D. and Fiorelli, G. (2008). Glucose-6-phosphate dehydrogenase deficiency. *Lancet.* **371** (9606), 64-74.

Capriotti, E., Fariselli, P. and Casadio, R. (2005). I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Research.* **33** (suppl 2), W306-W310.

Caraco, Y., Blotnick, S. and Muszkat, M. (2008). CYP2C9 genotype-guided warfarin prescribing enhances the efficacy and safety of anticoagulation: A prospective randomized controlled study. *Clinical Pharmacology and Therapeutics.* **83** (3), 460-470.

Carlquist, J.F., Horne, B.D., Muhlestein, J.B., Lappé, D.L., Whiting, B.M., Kolek, M.J., Clarke, J.L., James, B.C. and Anderson, J.L. (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. *Journal of Thrombosis and Thrombolysis.* **22** (3), 191-197.

Cha, P.C., Mushiroda, T., Takahashi, A., Kubo, M., Minami, S., Kamatani, N. and Nakamura, Y. (2010). Genome-wide association study identifies genetic determinants of warfarin responsiveness for Japanese. *Human molecular genetics.* **19** (23), 4735-44.

Chen, Y., Ferguson, S.S., Negishi, M. and Goldstein, J.A. (2004). Induction of Human CYP2C9 by Rifampicin, Hyperforin, and Phenobarbital Is Mediated by the Pregnane X Receptor. *Journal of Pharmacology and Experimental Therapeutics.* **308** (2), 495-501.

Choonara, I.A., Malia, R.G., Haynes, B.P., Hay, C.R., Cholerton, S., Breckenridge, A.M., Preston, F.E. and Park, B.K. (1988). The relationship between inhibition of vitamin K1 2,3-epoxide reductase and reduction of clotting factor activity with warfarin. *British Journal of Clinical Pharmacology.* **25** (1), 1-7.

Christmas, P., Jones, J.P., Patten, C.J., Rock, D.A., Zheng, Y., Cheng, S.M., Weber, B.M., Carlesso, N., Scadden, D.T., Rettie, A.E. and Soberman, R.J. (2001). Alternative splicing determines the function of CYP4F3 by switching substrate specificity. *J Biol Chem.* **276** (41), 38166-72.

Chu, K., Wu, S.M., Stanley, T., Stafford, D.W. and High, K.A. (1996). A mutation in the propeptide of Factor IX leads to warfarin sensitivity by a novel mechanism. *The Journal of clinical investigation.* **98** (7), 1619-25.

Cirulli, E.T. and Goldstein, D.B. (2007). In vitro assays fail to predict in vivo effects of regulatory polymorphisms. *Hum Mol Genet.* **16** (16), 1931-9.

Clatanoff, D.V., Triggs, P.O. and Meyer, O.O. (1954). Clinical experience with coumarin anticoagulants warfarin and warfarin sodium. *AMA Arch Intern Med.* **94** (2), 213-20.

Cleland, J.G.F., Coletta, A.P., Buga, L., Antony, R., Pellicori, P., Freemantle, N. and Clark, A.L. (2011). Clinical trials update from the American Heart Association Meeting 2010: EMPHASIS-HF, RAFT, TIM-HF, Tele-HF, ASCEND-HF, ROCKET-AF, and PROTECT. *European Journal of Heart Failure.* **13** (4), 460-465.

Comp, P.C., Nixon, R.R. and Esmon, C.T. (1984). Determination of functional levels of protein C, an antithrombotic protein, using thrombin-thrombomodulin complex. *Blood.* **63** (1), 15-21.

Connolly, S.J., Ezekowitz, M.D., Yusuf, S., Eikelboom, J., Oldgren, J., Parekh, A., Pogue, J., Reilly, P.A., Themeles, E., Varrone, J., Wang, S., Alings, M., Xavier, D., Zhu, J., Diaz, R., Lewis, B.S., Darius, H., Diener, H.C., Joyner, C.D. and Wallentin, L. (2009). Dabigatran versus warfarin in patients with atrial fibrillation. *The New England journal of medicine.* **361** (12), 1139-51.

Crawford, D.C., Ritchie, M.D. and Rieder, M.J. (2007). Identifying the genotype behind the phenotype: A role model found in VKORC1 and its association with warfarin dosing. *Pharmacogenomics.* **8** (5), 487-496.

Cross, S.H. and Bird, A.P. (1995). CpG islands and genes. *Curr Opin Genet Dev.* **5** (3), 309-14.

Cui, X., Nelson, D.R. and Strobel, H.W. (2000). A novel human cytochrome P450 4F isoform (CYP4F11): cDNA cloning, expression, and genomic structural characterization. *Genomics.* **68** (2), 161-6.

Custodio Das Dores, S.M., Booth, S.L., Martini, L.A., De Carvalho Gouvea, V.H., Padovani, C.R., De Abreu Maffei, F.H., Campana, A.O. and Rupp De Paiva, S.A. (2007). Relationship between diet and anticoagulant response to warfarin: a factor analysis. *Eur J Nutr.* **46** (3), 147-54.

D'ambrosio, R.L., D'andrea, G., Cappucci, F., Chetta, M., Di Perna, P., Brancaccio, V., Grandone, E. and Margaglione, M. (2004). Polymorphisms in factor II and factor VII genes modulate oral anticoagulation with warfarin. *Haematologica.* **89** (12), 1510-1516+1540.

D'ambrosio, R.L., D'andrea, G., Cafolla, A., Faillace, F. and Margaglione, M. (2007). A new vitamin K epoxide reductase complex subunit-1 (VKORC1) mutation in a patient with decreased stability of CYP2C9 enzyme. *J Thromb Haemost.* **5** (1), 191-3.

D'andrea, G., D'ambrosio, R.L., Di Perna, P., Chetta, M., Santacroce, R., Brancaccio, V., Grandone, E. and Margaglione, M. (2005). A polymorphism in the VKORC1 gene is associated with an interindividual variability in the dose-anticoagulant effect of warfarin. *Blood.* **105** (2), 645-9.

D'andrea, G., D'ambrosio, R. and Margaglione, M. (2008). Oral anticoagulants: Pharmacogenetics. Relationship between genetic and non-genetic factors. *Blood Reviews.* **22** (3), 127-140.

Dahlbäck, B. (2005). Blood coagulation and its regulation by anticoagulant pathways: Genetic pathogenesis of bleeding and thrombotic diseases. *Journal of Internal Medicine.* **257** (3), 209-223.

Dang, M.T.N., Hambleton, J. and Kayser, S.R. (2005). The influence of ethnicity on warfarin dosage requirement. *Annals of Pharmacotherapy.* **39** (6), 1008-1012.

Davie, E.W. and Ratnoff, O.D. (1964). Waterfall Sequence for Intrinsic Blood Clotting. *Science.* **145** (3638), 1310-2.

Dike, G.W., Griffiths, D., Bidwell, E., Snape, T.J. and Rizza, C.R. (1980). A factor VII concentrate for therapeutic use. *Br J Haematol.* **45** (1), 107-18.

Dimas, A.S., Deutsch, S., Stranger, B.E., Montgomery, S.B., Borel, C., Attar-Cohen, H., Ingle, C., Beazley, C., Gutierrez Arcelus, M., Sekowska, M., Gagnebin, M., Nisbett, J., Deloukas, P., Dermitzakis, E.T. and Antonarakis, S.E. (2009). Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science.* **325** (5945), 1246-50.

Do, E.J., Lenzini, P., Eby, C.S., Bass, A.R., Mcmillin, G.A., Stevens, S.M., Woller, S.C., Pendleton, R.C., Anderson, J.L., Proctor, P., Nunley, R.M., Davila-Roman, V. and Gage, B.F. (2011). Genetics informatics trial (GIFT) of warfarin to prevent deep vein thrombosis (DVT): rationale and study design. *The pharmacogenomics journal.*

Docherty, S.J., Davis, O.S., Haworth, C.M., Plomin, R. and Mill, J. (2010). DNA methylation profiling using bisulfite-based epityping of pooled genomic DNA. *Methods.* **52** (3), 255-8.

Ehrich, M., Nelson, M.R., Stanssens, P., Zabeau, M., Liloglou, T., Xinarianos, G., Cantor, C.R., Field, J.K. and Van Den Boom, D. (2005). Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc Natl Acad Sci U S A.* **102** (44), 15785-90.

Ehrich, M., Zoll, S., Sur, S. and Van Den Boom, D. (2007). A new method for accurate assessment of DNA quality after bisulfite treatment. *Nucleic Acids Res.* **35** (5), e29.

Eichelbaum, M., Spannbrucker, N., Steincke, B. and Dengler, H.J. (1979). Defective N-oxidation of sparteine in man: A new pharmacogenetic defect. *European Journal of Clinical Pharmacology.* **16** (3), 183-187.

Eikelboom, J.W., Wallentin, L., Connolly, S.J., Ezekowitz, M., Healey, J.S., Oldgren, J., Yang, S., Alings, M., Kaatz, S., Hohnloser, S.H., Diener, H.C., Franzosi, M.G., Huber, K., Reilly, P., Varrone, J. and Yusuf, S. (2011). Risk of bleeding with 2 doses of dabigatran compared with warfarin in older and younger patients with atrial fibrillation: an analysis of the randomized evaluation of long-term anticoagulant therapy (RE-LY) trial. *Circulation.* **123** (21), 2363-72.

El-Maarri, O., Becker, T., Junen, J., Manzoor, S.S., Diaz-Lacava, A., Schwaab, R., Wienker, T. and Oldenburg, J. (2007). Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. *Hum Genet.* **122** (5), 505-14.

Epstein, R.S., Moyer, T.P., Aubert, R.E., Dj, O.K., Xia, F., Verbrugge, R.R., Gage, B.F. and Teagarden, J.R. (2010). Warfarin genotyping reduces hospitalization rates results from the MM-WES (Medco-Mayo Warfarin Effectiveness study). *J Am Coll Cardiol.* **55** (25), 2804-12.

Esmon, C.T., Esmon, N.L. and Harris, K.W. (1982). Complex formation between thrombin and thrombomodulin inhibits both thrombin-catalyzed fibrin formation and factor V activation. *J Biol Chem.* **257** (14), 7944-7.

Esteller, M., Gaidano, G., Goodman, S.N., Zagonel, V., Capello, D., Botto, B., Rossi, D., Gloghini, A., Vitolo, U., Carbone, A., Baylin, S.B. and Herman, J.G. (2002). Hypermethylation of the

DNA repair gene O(6)-methylguanine DNA methyltransferase and survival of patients with diffuse large B-cell lymphoma. *J Natl Cancer Inst.* **94** (1), 26-32.

Evans, D.A., Manley, K.A. and Mckusick, V.A. (1960). Genetic control of isoniazid metabolism in man. *British Medical Journal.* **2** (5197), 485-491.

Faber, M.S. and Fuhr, U. (2004). Time response of cytochrome P450 1A2 activity on cessation of heavy smoking. *Clin Pharmacol Ther.* **76** (2), 178-84.

Fanikos, J., Grasso-Correnti, N., Shah, R., Kucher, N. and Goldhaber, S.Z. (2005). Major bleeding complications in a specialized anticoagulation service. *Am J Cardiol.* **96** (4), 595-8.

Feinberg, A.P. (2007). Phenotypic plasticity and the epigenetics of human disease. *Nature.* **447** (7143), 433-40.

Fihn, S.D., Mcdonell, M., Martin, D., Henikoff, J., Vermes, D., Kent, D. and White, R.H. (1993). Risk factors for complications of chronic anticoagulation: A multicenter study. *Annals of internal medicine.* **118** (7), 511-520.

Fihn, S.D., Callahan, C.M., Martin, D.C., Mcdonell, M.B., Henikoff, J.G. and White, R.H. (1996). The risk for and severity of bleeding complications in elderly patients treated with warfarin. The National Consortium of Anticoagulation Clinics. *Annals of internal medicine.* **124** (11), 970-9.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T.D., Wu, Y.Z., Plass, C. and Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A.* **102** (30), 10604-9.

Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M., *et al.* (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature.* **449** (7164), 851-61.

Freeman, B.D., Zehnbauer, B.A., Mcgrath, S., Borecki, I. and Buchman, T.G. (2000). Cytochrome P450 polymorphisms are associated with reduced warfarin dose. *Surgery.* **128** (2), 281-285.

French, B., Joo, J., Geller, N.L., Kimmel, S.E., Rosenberg, Y., Anderson, J.L., Gage, B.F., Johnson, J.A. and Ellenberg, J.H. (2010). Statistical design of personalized medicine interventions: the Clarification of Optimal Anticoagulation through Genetics (COAG) trial. *Trials.* **11**, 108.

Fuke, C., Shimabukuro, M., Petronis, A., Sugimoto, J., Oda, T., Miura, K., Miyazaki, T., Ogura, C., Okazaki, Y. and Jinno, Y. (2004). Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study. *Ann Hum Genet.* **68** (Pt 3), 196-204.

Furie, B. and Furie, B.C. (1990). Molecular basis of vitamin K-dependent gamma-carboxylation. *Blood.* **75** (9), 1753-62.

Furuya, H., Fernandez-Salguero, P., Gregory, W., Taber, H., Steward, A. and Gonzalez, F.J. (1995). Genetic polymorphism of CYP2C9 and its effect on warfarin maintenance dose requirement in patients undergoing anticoagulation therapy. *Pharmacogenetics.* **5**, 389-392.

Gage, B.F. and Eby, C.S. (2003). Pharmacogenetics and anticoagulant therapy. *Journal of Thrombosis and Thrombolysis.* **16** (1-2), 73-78.

Gage, B.F., Eby, C., Milligan, P.E., Banet, G.A., Duncan, J.R. and Mcleod, H.L. (2004). Use of pharmacogenetics and clinical factors to predict the maintenance dose of warfarin. *Thromb Haemost.* **91** (1), 87-94.

Gage, B.F., Eby, C., Johnson, J.A., Deych, E., Rieder, M.J., Ridker, P.M., Milligan, P.E., Grice, G., Lenzini, P., Rettie, A.E., Aquilante, C.L., Grosso, L., Marsh, S., Langaee, T., Farnett, L.E., Voora, D., Veenstra, D.L., Glynn, R.J., Barrett, A. and Mcleod, H.L. (2008). Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. *Clin Pharmacol Ther.* **84** (3), 326-31.

Garrod, A.E. (1909). *Inborn Errors of Metabolism.*

Gasche, Y., Daali, Y., Fathi, M., Chiappe, A., Cottini, S., Dayer, P. and Desmeules, J. (2004). Codeine intoxication associated with ultrarapid CYP2D6 metabolism. *N Engl J Med.* **351** (27), 2827-31.

Geick, A., Eichelbaum, M. and Burk, O. (2001). Nuclear Receptor Response Elements Mediate Induction of Intestinal MDR1 by Rifampin. *Journal of Biological Chemistry.* **276** (18), 14581-14587.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A.J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J.Y.

and Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome biology.* **5** (10), R80.

Ginsburg, G.S. and Willard, H.F. (2009). Genomic and personalized medicine: foundations and applications. *Transl Res.* **154** (6), 277-87.

Goldstein, J.A. and De Morais, S.M. (1994). Biochemistry and molecular biology of the human CYP2C subfamily. *Pharmacogenetics.* **4** (6), 285-99.

Gomes, A.M., Winter, S., Klein, K., Turpeinen, M., Schaeffeler, E., Schwab, M. and Zanger, U.M. (2009). Pharmacogenomics of human liver cytochrome P450 oxidoreductase: multifactorial analysis and impact on microsomal drug oxidation. *Pharmacogenomics.* **10** (4), 579-99.

Gonzalez, F.J., Skoda, R.C., Kimura, S., Umeno, M., Zanger, U.M., Nebert, D.W., Gelboin, H.V., Hardwick, J.P. and Meyer, U.A. (1988). Characterization of the common genetic defect in humans deficient in debrisoquine metabolism. *Nature.* **331** (6155), 442-446.

Goodstadt, L. and Ponting, C.P. (2004). Vitamin K epoxide reductase: homology, active site and catalytic mechanism. *Trends Biochem Sci.* **29** (6), 289-92.

Grossman, S.J., Herold, E.G., Drey, J.M., Alberts, D.W., Umbenhauer, D.R., Patrick, D.H., Nicoll-Griffith, D., Chauret, N. and Yergey, J.A. (1993). CYP1A1 specificity of Verlukast epoxidation in mice, rats, rhesus monkeys, and humans. *Drug Metabolism and Disposition.* **21** (6), 1029-1036.

Gulati, G., Hevelow, M., George, M., Behling, E. and Siegel, J. (2011). International normalized ratio versus plasma levels of coagulation factors in patients on vitamin K antagonist therapy. *Arch Pathol Lab Med.* **135** (4), 490-4.

Gulløv, A.L., Koefoed, B.G. and Petersen, P. (1999). Bleeding during warfarin and aspirin therapy in patients with atrial fibrillation: The AFASAK 2 Study. *Archives of internal medicine.* **159** (12), 1322-1328.

Gurwitz, J.H., Avorn, J., Ross-Degnan, D., Choodnovskiy, I. and Ansell, J. (1992). Aging and the anticoagulant response to warfarin therapy. *Annals of internal medicine.* **116** (11), 901-904.

Haining, R.L., Hunter, A.P., Veronese, M.E., Trager, W.F. and Rettie, A.E. (1996). Allelic variants of human cytochrome P450 2C9: Baculovirus-mediated expression, purification, structural characterization, substrate stereoselectivity, and prochiral selectivity of the wild-type and 1359L mutant forms. *Archives of Biochemistry and Biophysics.* **333** (2), 447-458.

Han, X., Fiehler, R. and Broze, G.J. (1998). Isolation of a protein Z-dependent plasma protease inhibitor. *Proceedings of the National Academy of Sciences.* **95** (16), 9250-9255.

Hariparsad, N., Chu, X., Yabut, J., Labhart, P., Hartley, D.P., Dai, X. and Evers, R. (2009). Identification of pregnane-X receptor target genes and coactivator and corepressor binding to promoter elements in human hepatocytes. *Nucleic acids research.* **37** (4), 1160-73.

Harrington, D.J., Underwood, S., Morse, C., Shearer, M.J., Tuddenham, E.G. and Mumford, A.D. (2005). Pharmacodynamic resistance to warfarin associated with a Val66Met substitution in vitamin K epoxide reductase complex subunit 1. *Thromb Haemost.* **93** (1), 23-6.

Harrington, D.J., Gorska, R., Wheeler, R., Davidson, S., Murden, S., Morse, C., Shearer, M.J. and Mumford, A.D. (2008). Pharmacodynamic resistance to warfarin is associated with nucleotide substitutions in VKORC1. *J Thromb Haemost.* **6** (10), 1663-70.

Hashizume, T., Imaoka, S., Hiroi, T., Terauchi, Y., Fujii, T., Miyazaki, H., Kamataki, T. and Funae, Y. (2001). cDNA cloning and expression of a novel cytochrome P450 (CYP4F12) from human small intestine. *Biochemical and biophysical research communications.* **280** (4), 1135-41.

Hashizume, T., Imaoka, S., Mise, M., Terauchi, Y., Fujii, T., Miyazaki, H., Kamataki, T. and Funae, Y. (2002). Involvement of CYP2J2 and CYP4F12 in the metabolism of ebastine in human intestinal microsomes. *The Journal of pharmacology and experimental therapeutics.* **300** (1), 298-304.

Hatchwell, E. and Greally, J.M. (2007). The potential role of epigenomic dysregulation in complex human disease. *Trends Genet.* **23** (11), 588-95.

Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouze, P. and Brunak, S. (1996). Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information. *Nucleic Acids Res.* **24** (17), 3439-52.

Herman, D., Peternel, P., Stegnar, M., Breskvar, K. and Dolzan, V. (2006). The influence of sequence variations in factor VII, γ-glutamyl carboxylase and vitamin K epoxide reductase complex genes on warfarin dose requirement. *Thrombosis and Haemostasis.* **95** (5), 782-787.

Higashi, M.K., Veenstra, D.L., Midori Kondo, L., Wittkowsky, A.K., Srinouanprachanh, S.L., Farin, F.M. and Rettie, A.E. (2002). Association between CYP2C9 genetic variants and

anticoagulation-related outcomes during warfarin therapy. *Journal of the American Medical Association.* **287** (13), 1690-1698.

Hillman, M.A., Wilke, R.A., Caldwell, M.D., Berg, R.L., Glurich, I. and Burmester, J.K. (2004). Relative impact of covariates in prescribing warfarin according to CYP2C9 genotype. *Pharmacogenetics.* **14** (8), 539-47.

Hillman, M.A., Wilke, R.A., Yale, S.H., Vidaillet, H.J., Caldwell, M.D., Glurich, I., Berg, R.L., Schmelzer, J. and Burmester, J.K. (2005). A prospective, randomized pilot trial of model-based warfarin dose initiation using CYP2C9 genotype and clinical data. *Clinical medicine & research.* **3** (3), 137-145.

Hochstein, P. (1983). Futile redox cycling: implications for oxygen radical toxicity. *Fundam Appl Toxicol.* **3** (4), 215-7.

Holliday, R. and Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. *Science.* **187** (4173), 226-32.

Howie, B.N., Donnelly, P. and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5** (6), e1000529.

Hsu, M.H., Savas, U., Griffin, K.J. and Johnson, E.F. (2007). Regulation of human cytochrome P450 4F2 expression by sterol regulatory element-binding protein and lovastatin. *J Biol Chem.* **282** (8), 5225-36.

Huang, W., Lin, Y.S., Mcconn Ii, D.J., Calamia, J.C., Totah, R.A., Isoherranen, N., Glodowski, M. and Thummel, K.E. (2004). Evidence of significant contribution from CYP3A5 to hepatic drug metabolism. *Drug Metabolism and Disposition.* **32** (12), 1434-1445.

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P.S., Pagni, M. and Sigrist, C.J. (2006). The PROSITE database. *Nucleic Acids Res.* **34** (Database issue), D227-30.

Hylek, E.M., Evans-Molina, C., Shea, C., Henault, L.E. and Regan, S. (2007). Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. *Circulation.* **115** (21), 2689-2696.

Imai, J., Ieiri, I., Mamiya, K., Miyahara, S., Furuumi, H., Nanba, E., Yamane, M., Fukumaki, Y., Ninomiya, H., Tashiro, N., Otsubo, K. and Higuchi, S. (2000). Polymorphism of the cytochrome P450 (CYP) 2C9 gene in Japanese epileptic patients: Genetic analysis of the CYP2C9 locus. *Pharmacogenetics.* **10** (1), 85-89.

Ingelman-Sundberg, M. (2004). Pharmacogenetics of cytochrome P450 and its applications in drug therapy: The past, present and future. *Trends in Pharmacological Sciences.* **25** (4), 193-200.

Iuliano, L., Pratico, D., Ferro, D., Pittoni, V., Valesini, G., Lawson, J., Fitzgerald, G.A. and Violi, F. (1997). Enhanced lipid peroxidation in patients positive for antiphospholipid antibodies. *Blood.* **90** (10), 3931-5.

Jaenisch, R. and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet.* **33 Suppl**, 245-54.

James, A.H., Britt, R.P., Raskino, C.L. and Thompson, S.G. (1992). Factors affecting the maintenance dose of warfarin. *J Clin Pathol.* **45** (8), 704-6.

Jerkeman, A., Astermark, J., Hedner, U., Lethagen, S., Olsson, C.G. and Berntorp, E. (2000). Correlation between different intensities of anti-vitamin K treatment and coagulation parameters. *Thromb Res.* **98** (6), 467-71.

Jin, R., Koop, D.R., Raucy, J.L. and Lasker, J.M. (1998). Role of human CYP4F2 in hepatic catabolism of the proinflammatory agent leukotriene B4. *Archives of biochemistry and biophysics.* **359** (1), 89-98.

Johansson, I., Lundqvist, E., Bertilsson, L., Dahl, M.L., Sjoqvist, F. and Ingelman-Sundberg, M. (1993). Inherited amplification of an active gene in the cytochrome P450 CYP2D locus as a cause of ultrarapid metabolism of debrisoquine. *Proc Natl Acad Sci U S A.* **90** (24), 11825-9.

Jones, M., Mcewan, P., Morgan, C.L., Peters, J.R., Goodfellow, J. and Currie, C.J. (2005). Evaluation of the pattern of treatment, level of anticoagulation control, and outcome of treatment with warfarin in patients with non-valvar atrial fibrillation: A record linkage study in a large British population. *Heart.* **91** (4), 472-477.

Jones, P.A. (1999). The DNA methylation paradox. *Trends Genet.* **15** (1), 34-7.

Jones, P.A. and Baylin, S.B. (2007). The epigenomics of cancer. *Cell.* **128** (4), 683-92.

Jorgensen, A.L., Al-Zubiedi, S., Zhang, J.E., Keniry, A., Hanson, A., Hughes, D.A., Eker, D., Stevens, L., Hawkins, K., Toh, C.H., Kamali, F., Daly, A.K., Fitzmaurice, D., Coffey, A., Williamson, P.R., Park, B.K., Deloukas, P. and Pirmohamed, M. (2009). Genetic and

environmental factors determining clinical outcomes and cost of warfarin therapy: a prospective study. *Pharmacogenetics and genomics.* **19** (10), 800-12.

Kalow, W. and Genest, K. (1957). A method for the detection of atypical forms of human serum cholinesterase; determination of dibucaine numbers. *Canadian journal of biochemistry and physiology.* **35** (6), 339-346.

Kalow, W. (1962). *Pharmacogenetics: Heredity and the Response to Drugs.*

Kalsotra, A., Turman, C.M., Kikuta, Y. and Strobel, H.W. (2004). Expression and characterization of human cytochrome P450 4F11: Putative role in the metabolism of therapeutic drugs and eicosanoids. *Toxicology and applied pharmacology.* **199** (3), 295-304.

Kaminsky, L.S., De Morais, S.M.F., Faletto, M.B., Dunbar, D.A. and Goldstein, J.A. (1993). Correlation of human cytochrome P4502C substrate specificities with primary structure: Warfarin as a probe. *Molecular Pharmacology.* **43** (2), 234-239.

Kaminsky, L.S. and Zhang, Z.Y. (1997). Human P450 metabolism of warfarin. *Pharmacology and Therapeutics.* **73** (1), 67-74.

Kato, M., Zhang, J., Wang, M., Lanting, L., Yuan, H., Rossi, J.J. and Natarajan, R. (2007). MicroRNA-192 in diabetic kidney glomeruli and its function in TGF-beta-induced collagen expression via inhibition of E-box repressors. *Proc Natl Acad Sci U S A.* **104** (9), 3432-7.

Kelly, J.G. and O'malley, K. (1979). Clinical pharmacokinetics of oral anticoagulants. *Clinical Pharmacokinetics.* **4** (1), 1-15.

Ken-Dror, G., Drenos, F., Humphries, S.E., Talmud, P.J., Hingorani, A.D., Kivimaki, M., Kumari, M., Bauer, K.A., Morrissey, J.H. and Ireland, H.A. (2010). Haplotype and genotype effects of the F7 gene on circulating factor VII, coagulation activation markers and incident coronary heart disease in UK men. *J Thromb Haemost.* **8** (11), 2394-403.

Kikuta, Y., Kusunose, E., Endo, K., Yamamoto, S., Sogawa, K., Fujii-Kuriyama, Y. and Kusunose, M. (1993). A novel form of cytochrome P-450 family 4 in human polymorphonuclear leukocytes. cDNA cloning and expression of leukotriene B4 omega-hydroxylase. *J Biol Chem.* **268** (13), 9376-80.

Kikuta, Y., Kato, M., Yamashita, Y., Miyauchi, Y., Tanaka, K., Kamada, N. and Kusunose, M. (1998). Human leukotriene B4 omega-hydroxylase (CYP4F3) gene: molecular cloning and chromosomal localization. *DNA Cell Biol.* **17** (3), 221-30.

Kikuta, Y., Miyauchi, Y., Kusunose, E. and Kusunose, M. (1999). Expression and molecular cloning of human liver leukotriene B4 omega-hydroxylase (CYP4F2) gene. *DNA and cell biology.* **18** (9), 723-30.

Kimmel, S.E., Chen, Z., Price, M., Parker, C.S., Metlay, J.P., Christie, J.D., Brensinger, C.M., Newcomb, C.W., Samaha, F.F. and Gross, R. (2007). The influence of patient adherence on anticoagulation control with warfarin: Results from the international normalized ratio adherence and genetics (IN-RANGE) study. *Archives of internal medicine.* **167** (3), 229-235.

Kimmel, S.E. (2008). Warfarin therapy: in need of improvement after all these years. *Expert Opin Pharmacother.* **9** (5), 677-86.

Kimmel, S.E., Christie, J., Kealey, C., Chen, Z., Price, M., Thorn, C.F., Brensinger, C.M., Newcomb, C.W. and Whitehead, A.S. (2008). Apolipoprotein E genotype and warfarin dosing among Caucasians and African Americans. *Pharmacogenomics J.* **8** (1), 53-60.

Kimura, R., Miyashita, K., Kokubo, Y., Akaiwa, Y., Otsubo, R., Nagatsuka, K., Otsuki, T., Okayama, A., Minematsu, K., Naritomi, H., Honda, S., Tomoike, H. and Miyata, T. (2007). Genotypes of vitamin K epoxide reductase, γ-glutamyl carboxylase, and cytochrome P450 2C9 as determinants of daily warfarin dose in Japanese patients. *Thrombosis Research.* **120** (2), 181-186.

Kirchheiner, J., Keulen, J.T., Bauer, S., Roots, I. and Brockmoller, J. (2008). Effects of the CYP2D6 gene duplication on the pharmacokinetics and pharmacodynamics of tramadol. *J Clin Psychopharmacol.* **28** (1), 78-83.

Kitamura, E., Igarashi, J., Morohashi, A., Hida, N., Oinuma, T., Nemoto, N., Song, F., Ghosh, S., Held, W.A., Yoshida-Noro, C. and Nagase, H. (2007). Analysis of tissue-specific differentially methylated regions (TDMs) in humans. *Genomics.* **89** (3), 326-37.

Klein, T.E., Altman, R.B., Eriksson, N., Gage, B.F., Kimmel, S.E., Lee, M.T., Limdi, N.A., Page, D., Roden, D.M., Wagner, M.J., Caldwell, M.D. and Johnson, J.A. (2009). Estimation of the warfarin dose with clinical and pharmacogenetic data. *N Engl J Med.* **360** (8), 753-64.

Kohlmeier, M., Salomon, A., Saupe, J. and Shearer, M.J. (1996). Transport of vitamin K to bone in humans. *Journal of Nutrition.* **126** (4 SUPPL.), 1192S-1196S.

Kohnke, H., Scordo, M.G., Pengo, V., Padrini, R. and Wadelius, M. (2005). Apolipoprotein E (APOE) and warfarin dosing in an Italian population. *European Journal of Clinical Pharmacology.* **61** (10), 781-783.

Kristensen, S.R. (2002). Warfarin treatment of a patient with coagulation factor IX propeptide mutation causing warfarin hypersensitivity [4]. *Blood.* **100** (7), 2676-2677.

Kroetz, D.L., Pauli-Magnus, C., Hodges, L.M., Huang, C.C., Kawamoto, M., Johns, S.J., Stryke, D., Ferrin, T.E., Deyoung, J., Taylor, T., Carlson, E.J., Herskowitz, I., Giacomini, K.M. and Clark, A.G. (2003). Sequence diversity and haplotype structure in the human ABCB1 (MDR1, multidrug resistance transporter) gene. *Pharmacogenetics.* **13** (8), 481-494.

Kumar, S., Haigh, J.R., Tate, G., Boothby, M., Joanes, D.N., Davies, J.A., Roberts, B.E. and Feely, M.P. (1990). Effect of warfarin on plasma concentrations of vitamin K dependent coagulation factors in patients with stable control and monitored compliance. *Br J Haematol.* **74** (1), 82-5.

Lamon-Fava, S., Sadowski, J.A., Davidson, K.W., O'brien, M.E., Mcnamara, J.R. and Schaefer, E.J. (1998). Plasma lipoproteins as carriers of phylloquinone (vitamin K1) in humans. *American Journal of Clinical Nutrition.* **67** (6), 1226-1231.

Landefeld, C.S. and Beyth, R.J. (1993). Anticoagulant-related bleeding: Clinical epidemiology, prediction, and prevention. *American Journal of Medicine.* **95** (3), 315-328.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature.* **409** (6822), 860-921.

Lane, S., Al-Zubiedi, S., Hatch, E., Matthews, I., Jorgensen, A.L., Deloukas, P., Daly, A.K., Park, B.K., Aarons, L., Ogungbenro, K., Kamali, F., Hughes, D. and Pirmohamed, M. (2011). The Population Pharmacokinetics of R and S-Warfarin: Effect of Genetic and Clinical Factors. *British Journal of Clinical Pharmacology.* no-no.

Lang, T., Klein, K., Fischer, J., Nussler, A.K., Neuhaus, P., Hofmann, U., Eichelbaum, M., Schwab, M. and Zanger, U.M. (2001). Extensive genetic polymorphism in the human CYP2B6 gene with impact on expression and function in human liver. *Pharmacogenetics.* **11** (5), 399-415.

Lasker, J.M., Chen, W.B., Wolf, I., Bloswick, B.P., Wilson, P.D. and Powell, P.K. (2000). Formation of 20-hydroxyeicosatetraenoic acid, a vasoactive and natriuretic eicosanoid, in human kidney. Role of Cyp4F2 and Cyp4A11. *J Biol Chem.* **275** (6), 4118-26.

Lasseur, R., Longin-Sauvageon, C., Videmann, B., Billeret, M., Berny, P. and Benoit, E. (2005). Warfarin resistance in a French strain of rats. *J Biochem Mol Toxicol.* **19** (6), 379-85.

Lasseur, R., Grandemange, A., Longin-Sauvageon, C., Berny, P. and Benoit, E. (2006). Heterogeneity of the coumarin anticoagulant targeted vitamin K epoxide reduction system. Study of kinetic parameters in susceptible and resistant mice (Mus musculus domesticus). *J Biochem Mol Toxicol.* **20** (5), 221-9.

Lawley, W.J., Charlton, A.J., Hughson, E.J., Grundy, H.H., Brown, P.M. and 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** (5), 1588-93.

Lazarou, J., Pomeranz, B.H. and Corey, P.N. (1998). Incidence of Adverse Drug Reactions in Hospitalized Patients. *JAMA.* **279** (15), 1200-1205.

Lee, C.R., Goldstein, J.A. and Pieper, J.A. (2002). Cytochrome P450 2C9 polymorphisms: a comprehensive review of the in-vitro and human data. *Pharmacogenetics.* **12** (3), 251-63.

Lee, S.C., Ng, S.S., Oldenburg, J., Chong, P.Y., Rost, S., Guo, J.Y., Yap, H.L., Rankin, S.C., Khor, H.B., Yeo, T.C., Ng, K.S., Soong, R. and Goh, B.C. (2006). Interethnic variability of warfarin maintenance requirement is explained by VKORC1 genotype in an Asian population. *Clinical Pharmacology and Therapeutics.* **79** (3), 197-205.

Lefevre, C., Bouadjar, B., Ferrand, V., Tadini, G., Megarbane, A., Lathrop, M., Prud'homme, J.F. and Fischer, J. (2006). Mutations in a new cytochrome P450 gene in lamellar ichthyosis type 3. *Human molecular genetics.* **15** (5), 767-76.

Lefrere, J.J., Horellou, M.H., Conard, J. and Samama, M. (1987). Proposed classification of resistances to oral anticoagulant therapy. *J Clin Pathol.* **40** (2), 242.

Lehmann, J.M., Mckee, D.D., Watson, M.A., Willson, T.M., Moore, J.T. and Kliewer, S.A. (1998). The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions. *Journal of Clinical Investigation.* **102** (5), 1016-1023.

Lenzini, P., Wadelius, M., Kimmel, S., Anderson, J.L., Jorgensen, A.L., Pirmohamed, M., Caldwell, M.D., Limdi, N., Burmester, J.K., Dowd, M.B., Angchaisuksiri, P., Bass, A.R., Chen, J., Eriksson, N., Rane, A., Lindh, J.D., Carlquist, J.F., Horne, B.D., Grice, G., Milligan, P.E.,

Eby, C., Shin, J., Kim, H., Kurnik, D., Stein, C.M., Mcmillin, G., Pendleton, R.C., Berg, R.L., Deloukas, P. and Gage, B.F. (2010). Integration of genetic, clinical, and INR data to refine warfarin dosing. *Clin Pharmacol Ther.* **87** (5), 572-8.

Li, T., Chang, C.Y., Jin, D.Y., Lin, P.J., Khvorova, A. and Stafford, D.W. (2004). Identification of the gene for vitamin K epoxide reductase. *Nature.* **427** (6974), 541-544.

Li, T., Lange, L.A., Li, X., Susswein, L., Bryant, B., Malone, R., Lange, E.M., Huang, T.Y., Stafford, D.W. and Evans, J.P. (2006). Polymorphisms in the VKORC1 gene are strongly associated with warfarin dosage requirements in patients receiving anticoagulation. *Journal of Medical Genetics.* **43** (9), 740-744.

Limdi, N.A., Beasley, T.M., Crowley, M.R., Goldstein, J.A., Rieder, M.J., Flockhart, D.A., Arnett, D.K., Acton, R.T. and Liu, N. (2008a). VKORC1 polymorphisms, haplotypes and haplotype groups on warfarin dose among African–Americans and European–Americans. *Pharmacogenomics.* **9** (10), 1445-1458.

Limdi, N.A., Mcgwin, G., Goldstein, J.A., Beasley, T.M., Arnett, D.K., Adler, B.K., Baird, M.F. and Acton, R.T. (2008b). Influence of CYP2C9 and VKORC1 1173C/T genotype on the risk of hemorrhagic complications in African-American and European-American patients on warfarin. *Clinical Pharmacology and Therapeutics.* **83** (2), 312-321.

Limdi, N.A. and Veenstra, D.L. (2010). Expectations, validity, and reality in pharmacogenetics. *J Clin Epidemiol.* **63** (9), 960-9.

Lind, S.E., Callas, P.W., Golden, E.A., Joyner, K.A., Jr. and Ortel, T.L. (1997). Plasma levels of factors II, VII and X and their relationship to the international normalized ratio during chronic warfarin therapy. *Blood Coagul Fibrinolysis.* **8** (1), 48-53.

Lindh, J.D., Holm, L., Andersson, M.L. and Rane, A. (2009). Influence of CYP2C9 genotype on warfarin dose requirements-a systematic review and meta-analysis. *European Journal of Clinical Pharmacology.* **65** (4), 365-375.

Link, K.P. (1959). The Discovery of Dicumarol and Its Sequels. *Circulation.* **19** (1), 97-107.

Loebstein, R., Yonath, H., Peleg, D., Almog, S., Rotenberg, M., Lubetsky, A., Roitelman, J., Harats, D., Halkin, H. and Ezra, D. (2001). Pharmacoepidemiology and drug utilization: Interindividual variability in sensitivity to warfarin - Nature or nurture? *Clinical Pharmacology and Therapeutics.* **70** (2), 159-164.

Loebstein, R., Vecsler, M., Kurnik, D., Austerweil, N., Gak, E., Halkin, H. and Almog, S. (2005). Common genetic variants of microsomal epoxide hydrolase affect warfarin dose requirements beyond the effect of cytochrome P450 2C9. *Clinical Pharmacology and Therapeutics.* **77** (5), 365-372.

Loebstein, R., Dvoskin, I., Halkin, H., Vecsler, M., Lubetsky, A., Rechavi, G., Amariglio, N., Cohen, Y., Ken-Dror, G., Almog, S. and Gak, E. (2007). A coding VKORC1 Asp36Tyr polymorphism predisposes to warfarin resistance. *Blood.* **109** (6), 2477-2480.

Ma, Q., Cui, K., Xiao, F., Lu, A.Y.H. and Yang, C.S. (1992). Identification of a glycine-rich sequence as an NAD(P)H-binding site and tyrosine 128 as a dicumarol-binding site in rat liver NAD(P)H:quinone oxidoreductase by site-directed mutagenesis. *Journal of Biological Chemistry.* **267** (31), 22298-22304.

Macfarlane, R.G. (1964). An Enzyme Cascade in the Blood Clotting Mechanism, and Its Function as a Biochemical Amplifier. *Nature.* **202**, 498-9.

Mahgoub, A., Idle, J.R. and Dring, L.G. (1977). Polymorphic hydroxylation of debrisoquine in man. *Lancet.* **2** (8038), 584-586.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., Mccarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., Cho, J.H., Guttmacher, A.E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C.N., Slatkin, M., Valle, D., Whittemore, A.S., Boehnke, M., Clark, A.G., Eichler, E.E., Gibson, G., Haines, J.L., Mackay, T.F., Mccarroll, S.A. and Visscher, P.M. (2009). Finding the missing heritability of complex diseases. *Nature.* **461** (7265), 747-53.

Manotti, C., Moia, M., Palareti, G., Pengo, V., Ria, L. and Dettori, A.G. (2001). Effect of computer-aided management on the quality of treatment in anticoagulated patients: a prospective, randomized, multicenter trial of APROAT (Automated PRogram for Oral Anticoagulant Treatment). *Haematologica.* **86** (10), 1060-70.

Margaglione, M., Colaizzo, D., D'andrea, G., Brancaccio, V., Ciampa, A., Grandone, E. and Di Minno, G. (2000). Genetic modulation of oral anticoagulation with warfarin. *Thromb Haemost.* **84** (5), 775-8.

Marketos, M. (2004). The top 200 generic drugs in 2003 (by units). *Drug Topics.* **148**, 82.

Marsh, S., King, C.R., Porsche-Sorbet, R.M., Scott-Horton, T.J. and Eby, C.S. (2006). Population variation in VKORC1 haplotype structure [3]. *Journal of Thrombosis and Haemostasis.* **4** (2), 473-474.

Marshall, A. (1997). Genset-Abbott deal heralds pharmacogenomics era. *Nat Biotechnol.* **15** (9), 829-30.

Martin, G.M. (2005). Epigenetic drift in aging identical twins. *Proc Natl Acad Sci U S A.* **102** (30), 10413-4.

Mathe, E., Olivier, M., Kato, S., Ishioka, C., Hainaut, P. and Tavtigian, S.V. (2006). Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. *Nucleic Acids Res.* **34** (5), 1317-25.

Mcdonald, M.G., Rieder, M.J., Nakano, M., Hsia, C.K. and Rettie, A.E. (2009). CYP4F2 is a vitamin K1 oxidase: An explanation for altered warfarin dose in carriers of the V433M variant. *Molecular pharmacology.* **75** (6), 1337-46.

Mcwilliam, A., Lutter, R. and Nardinelli, C. (2006). *Health care savings from personalizing medicine using genetic testing: The case of warfarin.*

Meletis, J. and Konstantopoulos, K. (2004). Favism - From the "avoid fava beans" of Pythagoras to the present. *HAEMA.* **7** (1), 17-21.

Meloni, T., Forteleoni, G., Dore, A. and Cutillo, S. (1983). Favism and hemolytic anemia in glucose-6-phosphate dehydrogenase-deficient subjects in North Sardinia. *Acta Haematol.* **70** (2), 83-90.

Mendel, G. (1866). Versuche über Pflanzenhybriden. *Verh Naturforschung Ver Brünn.* **4**, 3-47.

Meyer, U.A. (2004). Pharmacogenetics - Five decades of therapeutic lessons from genetic diversity. *Nature Reviews Genetics.* **5** (9), 669-676.

Mill, J., Tang, T., Kaminsky, Z., Khare, T., Yazdanpanah, S., Bouchard, L., Jia, P., Assadzadeh, A., Flanagan, J., Schumacher, A., Wang, S.C. and Petronis, A. (2008). Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am J Hum Genet.* **82** (3), 696-711.

Montgomery, S.B., Sammeth, M., Gutierrez-Arcelus, M., Lach, R.P., Ingle, C., Nisbett, J., Guigo, R. and Dermitzakis, E.T. (2010). Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature.* **464** (7289), 773-7.

Moore, T.J., Cohen, M.R. and Furberg, C.D. (2007). Serious Adverse Drug Events Reported to the Food and Drug Administration, 1998-2005. *Arch Intern Med.* **167** (16), 1752-1759.

Morisseau, C. and Hammock, B.D. (2005). pp. 311-333.

Motulsky, A.G. (1957). Drug reactions enzymes, and biochemical genetics. *Journal of the American Medical Association.* **165** (7), 835-837.

Mushiroda, T., Ohnishi, Y., Saito, S., Takahashi, A., Kikuchi, Y., Shimomura, H., Wanibuchi, Y., Suzuki, T., Kamatani, N. and Nakamura, Y. (2006). Association of VKORC1 and CYP2C9 polymorphisms with warfarin dose requirements in Japanese patients. *Journal of Human Genetics.* **51** (3), 249-253.

Myers, A.J., Gibbs, J.R., Webster, J.A., Rohrer, K., Zhao, A., Marlowe, L., Kaleem, M., Leung, D., Bryden, L., Nath, P., Zismann, V.L., Joshipura, K., Huentelman, M.J., Hu-Lince, D., Coon, K.D., Craig, D.W., Pearson, J.V., Holmans, P., Heward, C.B., Reiman, E.M., Stephan, D. and Hardy, J. (2007). A survey of genetic human cortical gene expression. *Nature genetics.* **39** (12), 1494-9.

Nagase, H. and Ghosh, S. (2008). Epigenetics: differential DNA methylation in mammalian somatic tissues. *FEBS J.* **275** (8), 1617-23.

Naidong, W. and Lee, J.W. (1993). Development and validation of a high-performance liquid chromatographic method for the quantitation of warfarin enantiomers in human plasma. *J Pharm Biomed Anal.* **11** (9), 785-92.

Nakagawa, T., Kishino, S., Itoh, S., Sugawara, M. and Miyazaki, K. (2003). Differential binding of disopyramide and warfarin enantiomers to human α1-acid glycoprotein variants. *British Journal of Clinical Pharmacology.* **56** (6), 664-669.

Nalla, V.K. and Rogan, P.K. (2005). Automated splicing mutation analysis by information theory. *Hum Mutat.* **25** (4), 334-42.

Nebert, D.W. (1999). Pharmacogenetics and pharmacogenomics: why is this relevant to the clinical geneticist? *Clin Genet.* **56** (4), 247-58.

Nelsestuen, G.L., Zytkovicz, T.H. and Howard, J.B. (1974). The mode of action of vitamin K. Identification of γ carboxyglutamic acid as a component of prothrombin. *Journal of Biological Chemistry.* **249** (19), 6347-6350.

Nelsestuen, G.L. (1976). Role of γ carboxyglutamic acid. An unusual protein transition required for the calcium dependent binding of prothrombin to phospholipid. *Journal of Biological Chemistry.* **251** (18), 5648-5656.

Ng, P.C. and Henikoff, S. (2001). Predicting deleterious amino acid substitutions. *Genome Res.* **11** (5), 863-74.

O'reilly, R.A., Pool, J.G. and Aggeler, P.M. (1968). Hereditary resistance to coumarin anticoagulant drugs in man and rat. *Ann N Y Acad Sci.* **151** (2), 913-31.

O'reilly, R.A. and Aggeler, P.M. (1968). Studies on coumarin anticoagulant drugs. Initiation of warfarin therapy without a loading dose. *Circulation.* **38** (1), 169-77.

O'reilly, R.A. (1974). Studies on the optical enantiomorphs of warfarin in man. *Clinical Pharmacology and Therapeutics.* **16** (2), 348-354.

O'reilly, R.A. (1976). Vitamin K and the oral anticoagulant drugs. *Annual Review of Medicine.* **27**, 245-261.

Oldenburg, J., Quenzel, E.M., Harbrecht, U., Fregin, A., Kress, W., Muller, C.R., Hertfelder, H.J., Schwaab, R., Brackmann, H.H. and Hanfland, P. (1997). Missense mutations at ALA-10 in the factor IX propeptide: an insignificant variant in normal life but a decisive cause of bleeding during oral anticoagulant therapy. *Br J Haematol.* **98** (1), 240-4.

Oldenburg, J., Bevans, C.G., Fregin, A., Geisen, C., Muller-Reible, C. and Watzka, M. (2007). Current pharmacogenetic developments in oral anticoagulation therapy: the influence of variant VKORC1 and CYP2C9 alleles. *Thromb Haemost.* **98** (3), 570-8.

Osler, W. (2003). The Quotable Osler. In: Silverman, M.E., Murray, T.J. and Bryan, C.S. eds. American College of Physicians, Philadelphia.

Otagiri, M., Maruyama, T. and Imai, T. (1987). A comparative study of the interaction of warfarin with human α1-acid glycoprotein and human albumin. *Journal of Pharmacy and Pharmacology.* **39** (6), 416-420.

Ovesen, L., Lyduch, S. and Ott, P. (1989). A simple technique for predicting maintenance dosage of warfarin - Is it better than empirical dosing? *European Journal of Clinical Pharmacology.* **37** (6), 573-576.

Palareti, G., Leali, N., Coccheri, S., Poggi, M., Manotti, C., D'angelo, A., Pengo, V., Erba, N., Moia, M., Ciavarella, N., Devoto, G., Berrettini, M. and Musolesi, S. (1996). Bleeding complications of oral anticoagulant treatment: An inception-cohort, prospective collaborative study (ISCOAT). *Lancet.* **348** (9025), 423-428.

Park, B.K. (1988). Warfarin: metabolism and mode of action. *Biochem Pharmacol.* **37** (1), 19-27.

Paul, B., Oxley, A., Brigham, K., Cox, T. and Hamilton, P.J. (1987). Factor II, VII, IX and X concentrations in patients receiving long term warfarin. *J Clin Pathol.* **40** (1), 94-8.

Pautas, E., Moreau, C., Gouin-Thibault, I., Golmard, J.L., Mahe, I., Legendre, C., Taillandier-Heriche, E., Durand-Gasselin, B., Houllier, A.M., Verrier, P., Beaune, P., Loriot, M.A. and Siguret, V. (2010). Genetic factors (VKORC1, CYP2C9, EPHX1, and CYP4F2) are predictor variables for warfarin response in very elderly, frail inpatients. *Clinical pharmacology and therapeutics.* **87** (1), 57-64.

Pavek, P. and Dvorak, Z. (2008). Xenobiotic-induced transcriptional regulation of xenobiotic metabolizing enzymes of the cytochrome P450 superfamily in human extrahepatic tissues. *Current drug metabolism.* **9** (2), 129-43.

Peoc'h, K., Pruvot, S., Gourmel, C., Dit Sollier, C.B. and Drouet, L. (2009). A new VKORC1 mutation leading to an isolated resistance to fluindione. *Br J Haematol.* **145** (6), 841-3.

Perez-Andreu, V., Roldan, V., Anton, A.I., Garcia-Barbera, N., Corral, J., Vicente, V. and Gonzalez-Conejero, R. (2009). Pharmacogenetic relevance of CYP4F2 V433M polymorphism on acenocoumarol therapy. *Blood.* **113** (20), 4977-9.

Perini, J.A., Struchiner, C.J., Silva-Assunção, E., Santana, I.S.C., Rangel, F., Ojopi, E.B., Dias-Neto, E. and Suarez-Kurtz, G. (2008). Pharmacogenetics of warfarin: Development of a dosing algorithm for Brazilian patients. *Clinical Pharmacology and Therapeutics.* **84** (6), 722-728.

Perini, J.A., Struchiner, C.J., Silva-Assuncao, E. and Suarez-Kurtz, G. (2010). Impact of CYP4F2 rs2108622 on the stable warfarin dose in an admixed patient cohort. *Clinical pharmacology and therapeutics.* **87** (4), 417-20.

Pertea, M., Lin, X. and Salzberg, S.L. (2001). GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res.* **29** (5), 1185-90.

Petty, G.W., Brown Jr, R.D., Whisnant, J.P., Sicks, J.D., O'fallon, W.M. and Wiebers, D.O. (1999). Frequency of major complications of aspirin, warfarin, and intravenous heparin for secondary stroke prevention: A population-based study. *Annals of internal medicine.* **130** (1), 14-22.

Philibert, R.A., Beach, S.R., Gunter, T.D., Brody, G.H., Madan, A. and Gerrard, M. (2010). The effect of smoking on MAOA promoter methylation in DNA prepared from lymphoblasts and whole blood. *Am J Med Genet B Neuropsychiatr Genet.* **153B** (2), 619-28.

Philipsen, S. and Suske, G. (1999). A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. *Nucleic Acids Res.* **27** (15), 2991-3000.

Pickrell, J.K., Marioni, J.C., Pai, A.A., Degner, J.F., Engelhardt, B.E., Nkadori, E., Veyrieras, J.B., Stephens, M., Gilad, Y. and Pritchard, J.K. (2010). Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature.* **464** (7289), 768-72.

Pirmohamed, M. (2001). Pharmacogenetics and pharmacogenomics. *Br J Clin Pharmacol.* **52** (4), 345-7.

Pirmohamed, M., James, S., Meakin, S., Green, C., Scott, A.K., Walley, T.J., Farrar, K., Park, B.K. and Breckenridge, A.M. (2004). Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ.* **329** (7456), 15-9.

Pirmohamed, M. (2006). Warfarin: almost 60 years old and still causing problems. *Br J Clin Pharmacol.* **62** (5), 509-11.

Pirmohamed, M. (2011). Pharmacogenetics: past, present and future. *Drug Discov Today.* **16** (19-20), 852-61.

Poort, S.R., Rosendaal, F.R., Reitsma, P.H. and Bertina, R.M. (1996). A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood.* **88** (10), 3698-703.

Portela, A. and Esteller, M. (2010). Epigenetic modifications and human disease. *Nat Biotechnol.* **28** (10), 1057-68.

Porter, R.S. and Sawyer, W.T. (1992). Warfarin. *Applied Pharmacokinetics: Principles of Therapeutic Drug Monitoring.* 1-46.

Powell, P.K., Wolf, I., Jin, R. and Lasker, J.M. (1998). Metabolism of arachidonic acid to 20-hydroxy-5,8,11, 14-eicosatetraenoic acid by P450 enzymes in human liver: involvement of CYP4F2 and CYP4A11. *The Journal of pharmacology and experimental therapeutics.* **285** (3), 1327-36.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J. and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics.* **81** (3), 559-75.

Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., Cho, E.K., Dallaire, S., Freeman, J.L., González, J.R., Gratacòs, M., Huang, J., Kalaitzopoulos, D., Komura, D., Macdonald, J.R., Marshall, C.R., Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M.J., Tchinda, J., Valsesia, A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Armengol, L., Conrad, D.F., Estivill, X., Tyler-Smith, C., Carter, N.P., Aburatani, H., Lee, C., Jones, K.W., Scherer, S.W. and Hurles, M.E. (2006). Global variation in copy number in the human genome. *Nature.* **444** (7118), 444-454.

Reese, M.G., Eeckman, F.H., Kulp, D. and Haussler, D. (1997). Improved splice site detection in Genie. *J Comput Biol.* **4** (3), 311-23.

Reiner, A.P., Carlson, C.S., Rieder, M.J., Siscovick, D.S., Liu, K., Chandler, W.L., Green, D., Schwartz, S.M. and Nickerson, D.A. (2007). Coagulation factor VII gene haplotypes, obesity-related traits, and cardiovascular risk in young women. *J Thromb Haemost.* **5** (1), 42-9.

Reitsma, P.H., Van Der Heijden, J.F., Groot, A.P., Rosendaal, F.R. and Buller, H.R. (2005). A C1173T dimorphism in the VKORC1 gene determines coumarin sensitivity and bleeding risk. *PLoS Med.* **2** (10), e312.

Rettie, A.E., Korzekwa, K.R., Kunze, K.L., Lawrence, R.F., Eddy, A.C., Aoyama, T., Gelboin, H.V., Gonzalez, F.J. and Trager, W.F. (1992). Hydroxylation of warfarin by human cDNA-expressed cytochrome P-450: a role for P-4502C9 in the etiology of (S)-warfarin-drug interactions. *Chemical Research in Toxicology.* **5** (1), 54-9.

Rettie, A.E., Wienkers, L.C., Gonzalez, F.J., Trager, W.F. and Korzekwa, K.R. (1994). Impaired (S)-warfarin metabolism catalysed by the R144C allelic variant of CYP2C9. *Pharmacogenetics.* **4** (1), 39-42.

Rieder, M.J., Reiner, A.P., Gage, B.F., Nickerson, D.A., Eby, C.S., Mcleod, H.L., Blough, D.K., Thummel, K.E., Veenstra, D.L. and Rettie, A.E. (2005). Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. *N Engl J Med.* **352** (22), 2285-93.

Robertson, K.D. and Wolffe, A.P. (2000). DNA methylation in health and disease. *Nat Rev Genet.* **1** (1), 11-9.

Roderick, L.M. (1929). The pathology of sweet clover disease in cattle. *J Am Vet Med Assoc.* **74**, 314-325.

Roderick, L.M. (1931). A problem in the coagulation of the blood: Sweet-clover disease of cattle. *Am J Physiol.* **96**, 413-425.

Rogan, P.K., Faux, B.M. and Schneider, T.D. (1998). Information analysis of human splice site mutations. *Hum Mutat.* **12** (3), 153-71.

Ross, D. and Siegel, D. (2004). pp. 115-144.

Rost, S., Fregin, A., Ivaskevicius, V., Conzelmann, E., Hortnagel, K., Pelz, H.J., Lappegard, K., Seifried, E., Scharrer, I., Tuddenham, E.G., Muller, C.R., Strom, T.M. and Oldenburg, J. (2004a). Mutations in VKORC1 cause warfarin resistance and multiple coagulation factor deficiency type 2. *Nature.* **427** (6974), 537-41.

Rost, S., Fregin, A., Koch, D., Compes, M., Müller, C.R. and Oldenburg, J. (2004b). Compound heterozygous mutations in the γ-glutamyl carboxylase gene cause combined deficiency of all vitamin K-dependent blood coagulation factors. *British Journal of Haematology.* **126** (4), 546-549.

Rost, S., Fregin, A., Hunerberg, M., Bevans, C.G., Muller, C.R. and Oldenburg, J. (2005). Site-directed mutagenesis of coumarin-type anticoagulant-sensitive VKORC1: evidence that highly conserved amino acids define structural requirements for enzymatic activity and inhibition by warfarin. *Thrombosis and Haemostasis.* **94** (4), 780-6.

Rowan, A.B. (1859). Curious inaccuracy: De quincy and coleridge-pythagoras on beans. *Notes and Queries.* **s2-VII** (163), 125-126.

Rozen, S. and Skaletsky, H.J. (2000). Primer3 on the WWW for general users and for biologist programmers. *In: Krawetz, S. and Misener, S. eds.* Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 365-386

Rulcova, A., Prokopova, I., Krausova, L., Bitman, M., Vrzal, R., Dvorak, Z., Blahos, J. and Pavek, P. (2010). Stereoselective interactions of warfarin enantiomers with the pregnane X nuclear receptor in gene regulation of major drug-metabolizing cytochrome P450 enzymes. *Journal of thrombosis and haemostasis : JTH.* **8** (12), 2708-17.

Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., Hunt, S.E., Cole, C.G., Coggill, P.C., Rice, C.M., Ning, Z., Rogers, J., Bentley, D.R., Kwok, P.Y., Mardis, E.R., Yeh, R.T., Schultz, B., Cook, L., Davenport, R., Dante, M., Fulton, L., Hillier, L., Waterston, R.H., Mcpherson, J.D., Gilman, B., Schaffner, S., Van Etten, W.J., Reich, D., Higgins, J., Daly, M.J., Blumenstiel, B., Baldwin, J., Stange-Thomann, N., Zody, M.C., Linton, L., Lander, E.S. and Altshuler, D. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature.* **409** (6822), 928-33.

Sanderson, S., Emery, J. and Higgins, J. (2005). CYP2C9 gene variants, drug dose, and bleeding risk in warfarin-treated patients: A HuGEnet™ systematic review and meta-analysis. *Genetics in Medicine.* **7** (2), 97-104.

Sanoski, C.A. and Bauman, J.L. (2002). Clinical observations with the amiodarone/warfarin interaction: Dosing relationships with long-term therapy. *Chest.* **121** (1), 19-23.

Sarode, R., Rawal, A., Lee, R., Shen, Y.M. and Frenkel, E.P. (2006). Poor correlation of supratherapeutic international normalised ratio and vitamin K-dependent procoagulant factor levels during warfarin therapy. *Br J Haematol.* **132** (5), 604-7.

Saupe, J., Shearer, M.J. and Kohlmeier, M. (1993). Phylloquinone transport and its influence on γ-carboxyglutamate residues of osteocalcin in patients on maintenance hemodialysis. *American Journal of Clinical Nutrition.* **58** (2), 204-208.

Schadt, E.E., Molony, C., Chudin, E., Hao, K., Yang, X., Lum, P.Y., Kasarskis, A., Zhang, B., Wang, S., Suver, C., Zhu, J., Millstein, J., Sieberts, S., Lamb, J., Guhathakurta, D., Derry, J., Storey, J.D., Avila-Campillo, I., Kruger, M.J., Johnson, J.M., Rohl, C.A., Van Nas, A., Mehrabian, M., Drake, T.A., Lusis, A.J., Smith, R.C., Guengerich, F.P., Strom, S.C., Schuetz, E., Rushmore, T.H. and Ulrich, R. (2008). Mapping the genetic architecture of gene expression in human liver. *PLoS biology.* **6** (5), e107.

Schiliro, G., Russo, A., Curreri, R., Marino, S., Sciotto, A. and Russo, G. (1979). Glucose-6-phosphate dehydrogenase deficiency in Sicily. Incidence, biochemical characteristics and clinical implications. *Clin Genet.* **15** (2), 183-8.

Schofield, F.W. (1924). Damaged sweet clover. The cause of a new disease in cattle simulating hemorrhagic septicemia and blackleg. *J Am Vet Med Assoc.* **64**, 553-575.

Schroder, A., Klein, K., Winter, S., Schwab, M., Bonin, M., Zell, A. and Zanger, U.M. (2011). Genomics of ADME gene expression: mapping expression quantitative trait loci relevant for absorption, distribution, metabolism and excretion of drugs in human liver. *Pharmacogenomics J.*

Schwarz, U.I. (2003). Clinical relevance of genetic polymorphisms in the human CYP2C9 gene. *European Journal of Clinical Investigation, Supplement.* **33** (2), 23-30.

Schwarz, U.I., Ritchie, M.D., Bradford, Y., Li, C., Dudek, S.M., Frye-Anderson, A., Kim, R.B., Roden, D.M. and Stein, C.M. (2008). Genetic determinants of response to warfarin during initial anticoagulation. *New England Journal of Medicine.* **358** (10), 999-1008.

Sconce, E., Avery, P., Wynne, H. and Kamali, F. (2007). Vitamin K supplementation can improve stability of anticoagulation for patients with unexplained variability in response to warfarin. *Blood.* **109** (6), 2419-2423.

Sconce, E.A., Khan, T.I., Wynne, H.A., Avery, P., Monkhouse, L., King, B.P., Wood, P., Kesteven, P., Daly, A.K. and 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** (7), 2329-2333.

Sconce, E.A., Daly, A.K., Khan, T.I., Wynne, H.A. and Kamali, F. (2006). APOE genotype makes a small contribution to warfarin dose requirements. *Pharmacogenetics and genomics.* **16** (8), 609-611.

Senapathy, P., Shapiro, M.B. and Harris, N.L. (1990). Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. *Methods Enzymol.* **183**, 252-78.

Shapiro, M.B. and Senapathy, P. (1987). RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* **15** (17), 7155-74.

Shapiro, S. (1953). Warfarin sodium derivative: (coumadin sodium); an intravenous hypoprothrombinemia-inducing agent. *Angiology.* **4** (4), 380-90.

Shea, M.K., Booth, S.L., Massaro, J.M., Jacques, P.F., D'agostino, R.B., Sr., Dawson-Hughes, B., Ordovas, J.M., O'donnell, C.J., Kathiresan, S., Keaney, J.F., Jr., Vasan, R.S. and Benjamin, E.J. (2008). Vitamin K and vitamin D status: associations with inflammatory markers in the Framingham Offspring Study. *Am J Epidemiol.* **167** (3), 313-20.

Shea, M.K., Benjamin, E.J., Dupuis, J., Massaro, J.M., Jacques, P.F., D'agostino, R.B., Sr., Ordovas, J.M., O'donnell, C.J., Dawson-Hughes, B., Vasan, R.S. and Booth, S.L. (2009). Genetic and non-genetic correlates of vitamins K and D. *Eur J Clin Nutr.* **63** (4), 458-64.

Shen, L., Kondo, Y., Ahmed, S., Boumber, Y., Konishi, K., Guo, Y., Chen, X., Vilaythong, J.N. and Issa, J.P. (2007). Drug sensitivity prediction by CpG island methylation profile in the NCI-60 cancer cell line panel. *Cancer Res.* **67** (23), 11335-43.

Shikata, E., Ieiri, I., Ishiguro, S., Aono, H., Inoue, K., Koide, T., Ohgi, S. and Otsubo, K. (2004). Association of pharmacokinetic (CY72C9) and pharmacodynamic (factors II, VII, IX, and X; proteins S and C; and γ-glutamyl carboxylase) gene variants with warfarin sensitivity. *Blood.* **103** (7), 2630-2635.

Siest, G., Jeannesson, E., Marteau, J.-B., Samara, A., Marie, B., Pfister, M. and Visvikis-Siest, S. (2008). Transcription Factor and Drug-Metabolizing Enzyme Gene Expression in Lymphocytes from Healthy Human Subjects. *Drug Metabolism and Disposition.* **36** (1), 182-189.

Simioni, P. (1999). The molecular genetics of familial venous thrombosis. *Bailliere's best practice & research. Clinical haematology.* **12** (3), 479-503.

Smith, N.L., Chen, M.H., Dehghan, A., Strachan, D.P., Basu, S., Soranzo, N., Hayward, C., Rudan, I., Sabater-Lleal, M., Bis, J.C., De Maat, M.P., Rumley, A., Kong, X., Yang, Q., Williams, F.M., Vitart, V., Campbell, H., Malarstig, A., Wiggins, K.L., Van Duijn, C.M., Mcardle, W.L., Pankow, J.S., Johnson, A.D., Silveira, A., Mcknight, B., Uitterlinden, A.G., Aleksic, N., Meigs, J.B., Peters, A., Koenig, W., Cushman, M., Kathiresan, S., Rotter, J.I., Bovill, E.G., Hofman, A., Boerwinkle, E., Tofler, G.H., Peden, J.F., Psaty, B.M., Leebeek, F., Folsom, A.R., Larson, M.G., Spector, T.D., Wright, A.F., Wilson, J.F., Hamsten, A., Lumley, T., Witteman, J.C., Tang, W. and O'donnell, C.J. (2010). Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. *Circulation.* **121** (12), 1382-92.

Snyder, L.H. (1932). Studies in human inheritance. IX. The inheritance of taste deficiency in man. *Ohio J Sci.* **32**, 436-440.

Song, F., Smith, J.F., Kimura, M.T., Morrow, A.D., Matsuyama, T., Nagase, H. and Held, W.A. (2005). Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc Natl Acad Sci U S A.* **102** (9), 3336-41.

Song, J., Ugai, H., Ogawa, K., Wang, Y., Sarai, A., Obata, Y., Kanazawa, I., Sun, K., Itakura, K. and Yokoyama, K.K. (2001). Two consecutive zinc fingers in Sp1 and in MAZ are essential for interactions with cis-elements. *J Biol Chem.* **276** (32), 30429-34.

Sontag, T.J. and Parker, R.S. (2002). Cytochrome P450 omega-hydroxylase pathway of tocopherol catabolism. Novel mechanism of regulation of vitamin E status. *The Journal of biological chemistry.* **277** (28), 25290-6.

Stark, K., Bylund, J., Torma, H., Sahlen, G. and Oliw, E.H. (2005a). On the mechanism of biosynthesis of 19-hydroxyprostaglandins of human seminal fluid and expression of cyclooxygenase-2, PGH 19-hydroxylase (CYP4F8) and microsomal PGE synthase-1 in seminal vesicles and vas deferens. *Prostaglandins & other lipid mediators.* **75** (1-4), 47-64.

Stark, K., Wongsud, B., Burman, R. and Oliw, E.H. (2005b). Oxygenation of polyunsaturated long chain fatty acids by recombinant CYP4F8 and CYP4F12 and catalytic importance of Tyr-125 and Gly-328 of CYP4F8. *Archives of biochemistry and biophysics.* **441** (2), 174-81.

Stec, D.E., Roman, R.J., Flasch, A. and Rieder, M.J. (2007). Functional polymorphism in human CYP4F2 decreases 20-HETE production. *Physiol Genomics.* **30** (1), 74-81.

Stenflo, J., Fernlund, P., Egan, W. and Roepstorff, P. (1974). Vitamin K dependent modifications of glutamic acid residues in prothrombin. *Proceedings of the National Academy of Sciences of the United States of America.* **71** (7), 2730-2733.

Stephens, M., Smith, N.J. and Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *American journal of human genetics.* **68** (4), 978-89.

Stephens, M. and Donnelly, P. (2003). A comparison of bayesian methods for haplotype reconstruction from population genotype data. *American journal of human genetics.* **73** (5), 1162-9.

Stephens, M. and Scheet, P. (2005). Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *American journal of human genetics.* **76** (3), 449-62.

Steward, D.J., Haining, R.L., Henne, K.R., Davis, G., Rushmore, T.H., Trager, W.F. and Rettie, A.E. (1997). Genetic association between sensitivity to warfarin and expression of CYP2C9*3. *Pharmacogenetics.* **7** (5), 361-7.

Stranger, B.E., Nica, A.C., Forrest, M.S., Dimas, A., Bird, C.P., Beazley, C., Ingle, C.E., Dunning, M., Flicek, P., Koller, D., Montgomery, S., Tavare, S., Deloukas, P. and Dermitzakis, E.T. (2007). Population genomics of human gene expression. *Nature genetics.* **39** (10), 1217-24.

Straussman, R., Nejman, D., Roberts, D., Steinfeld, I., Blum, B., Benvenisty, N., Simon, I., Yakhini, Z. and Cedar, H. (2009). Developmental programming of CpG island methylation profiles in the human genome. *Nat Struct Mol Biol.* **16** (5), 564-71.

Stroke Prevention in Atrial Fibrillation Investigators (1996). Adjusted-dose warfarin versus low-intensity, fixed-dose warfarin plus aspirin for high-risk patients with atrial fibrillation: Stroke Prevention in Atrial Fibrillation III randomised clinical trial. *Lancet.* **348** (9028), 633-8.

Stubbins, M.J., Harries, L.W., Smith, G., Tarbit, M.H. and Wolf, C.R. (1996). Genetic analysis of the human cytochrome P450 CYP2C9 locus. *Pharmacogenetics.* **6** (5), 429-39.

Sullivan-Klose, T.H., Ghanayem, B.I., Bell, D.A., Zhang, Z.Y., Kaminsky, L.S., Shenfield, G.M., Miners, J.O., Birkett, D.J. and Goldstein, J.A. (1996). The role of the CYP2C9-Leu359 allelic variant in the tolbutamide polymorphism. *Pharmacogenetics.* **6** (4), 341-349.

Sunyaev, S., Ramensky, V., Koch, I., Lathe, W., 3rd, Kondrashov, A.S. and Bork, P. (2001). Prediction of deleterious human alleles. *Hum Mol Genet.* **10** (6), 591-7.

Sussman, N.L., Waltershied, M., Butler, T., Cali, J.J., Riss, T. and Kelly, J.H. (2002). The predictive nature of high throughput toxicity screening using a human hepatocyte cell line. *Cell Notes.* **3**, 7-10.

Suttie, J.W. (1978). Vitamin K. *Handbook of Lipid Research.* 211-277.

Tai, G., Farin, F., Rieder, M.J., Dreisbach, A.W., Veenstra, D.L., Verlinde, C.L.M.J. and Rettie, A.E. (2005). In-vitro and in-vivo effects of the CYP2C9*11 polymorphism on warfarin metabolism and dose. *Pharmacogenetics and genomics.* **15** (7), 475-481.

Takahashi, H., Kashima, T., Nomizo, Y., Muramoto, N., Shimizu, T., Nasu, K., Kubota, T., Kimura, S. and Echizen, H. (1998). Metabolism of warfarin enantiomers in Japanese patients with heart disease having different CYP2C9 and CYP2C19 genotypes. *Clinical Pharmacology and Therapeutics.* **63** (5), 519-528.

Takahashi, H. and Echizen, H. (2003). Pharmacogenetics of CYP2C9 and interindividual variability in anticoagulant response to warfarin. *Pharmacogenomics Journal.* **3** (4), 202-214.

Takahashi, H., Wilkinson, G.R., Nutescu, E.A., Morita, T., Ritchie, M.D., Scordo, M.G., Pengo, V., Barban, M., Padrini, R., Ieiri, I., Otsubo, K., Kashima, T., Kimura, S., Kijima, S. and Echizen, H. (2006). Different contributions of polymorphisms in VKORC1 and CYP2C9 to intra- and inter-population differences in maintenance dose of warfarin in Japanese, Caucasians and African-Americans. *Pharmacogenetics and genomics.* **16** (2), 101-10.

Takeuchi, F., Mcginnis, R., Bourgeois, S., Barnes, C., Eriksson, N., Soranzo, N., Whittaker, P., Ranganath, V., Kumanduri, V., Mclaren, W., Holm, L., Lindh, J., Rane, A., Wadelius, M. and Deloukas, P. (2009). A genome-wide association study confirms VKORC1, CYP2C9, and CYP4F2 as principal genetic determinants of warfarin dose. *PLoS Genet.* **5** (3), e1000433.

Taube, J., Halsall, D. and Baglin, T. (2000). Influence of cytochrome P-450 CYP2C9 polymorphisms on warfarin sensitivity and risk of over-anticoagulation in patients on long-term treatment. *Blood.* **96** (5), 1816-1819.

Tavtigian, S.V., Deffenbaugh, A.M., Yin, L., Judkins, T., Scholl, T., Samollow, P.B., De Silva, D., Zharkikh, A. and Thomas, A. (2006). Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet.* **43** (4), 295-305.

Taylor, K.C., Lange, L.A., Zabaneh, D., Lange, E., Keating, B.J., Tang, W., Smith, N.L., Delaney, J.A., Kumari, M., Hingorani, A., North, K.E., Kivimaki, M., Tracy, R.P., O'donnell, C.J., Folsom, A.R., Green, D., Humphries, S.E. and Reiner, A.P. (2011). A gene-centric association scan for Coagulation Factor VII levels in European and African Americans: the Candidate Gene Association Resource (CARe) Consortium. *Hum Mol Genet.* **20** (17), 3525-34.

Tchernitchko, D., Goossens, M. and Wajcman, H. (2004). In silico prediction of the deleterious effect of a mutation: proceed with caution in clinical genetics. *Clin Chem.* **50** (11), 1974-8.

Tham, L.S., Goh, B.C., Nafziger, A., Guo, J.Y., Wang, L.Z., Soong, R. and Lee, S.C. (2006). A warfarin-dosing model in Asians that uses single-nucleotide polymorphisms in vitamin K epoxide reductase complex and cytochrome P450 2C9. *Clinical Pharmacology and Therapeutics.* **80** (4), 346-355.

The International Hapmap Consortium (2003). The International HapMap Project. *Nature.* **426** (6968), 789-96.

The International Hapmap Consortium (2005). A haplotype map of the human genome. *Nature.* **437** (7063), 1299-1320.

Theofanous, T.G. and Barile, R.G. (1973). Multiple-dose kinetics of oral anticoagulants: methods of analysis and optimized dosing. *Journal of Pharmaceutical Sciences.* **62** (2), 261-266.

Tie, J.K., Nicchitta, C., Von Heijne, G. and Stafford, D.W. (2005). Membrane topology mapping of vitamin K epoxide reductase by in vitro translation/cotranslocation. *Journal of Biological Chemistry.* **280** (16), 16410-16416.

Trask, A.S., Gosselin, R.C., Diaz, J.A. and Dager, W.E. (2004). Warfarin initiation and monitoring with clotting factors II, VII, and X. *Ann Pharmacother.* **38** (2), 251-6.

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D. and Altman, R.B. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics.* **17** (6), 520-5.

Van Den Boom, D. and Ehrich, M. (2007). Discovery and identification of sequence polymorphisms and mutations with MALDI-TOF MS. *Methods in molecular biology.* **366**, 287-306.

Van Der Heijden, J.F., Rekké, B., Hutten, B.A., Van Der Meer, F.J.M., Remkes, M.G.H., Vermeulen, M., Büller, H.R. and Reitsma, P.H. (2004). Non-fatal major bleeding during treatment with vitamin K antagonists: Influence of soluble thrombomodulin and mutations in the propeptide of coagulation factor IX. *Journal of Thrombosis and Haemostasis.* **2** (7), 1104-1109.

Van Schie, R.M., Wadelius, M.I., Kamali, F., Daly, A.K., Manolopoulos, V.G., De Boer, A., Barallon, R., Verhoef, T.I., Kirchheiner, J., Haschke-Becher, E., Briz, M., Rosendaal, F.R., Redekop, W.K., Pirmohamed, M. and Maitland Van Der Zee, A.H. (2009). Genotype-guided dosing of coumarin derivatives: the European pharmacogenetics of anticoagulant therapy (EU-PACT) trial design. *Pharmacogenomics.* **10** (10), 1687-95.

Vecsler, M., Loebstein, R., Almog, S., Kurnik, D., Goldman, B., Halkin, H. and Gak, E. (2006). Combined genetic profiles of components and regulators of the vitamin K-dependent γ-carboxylation system affect individual sensitivity to warfarin. *Thrombosis and Haemostasis.* **95** (2), 205-211.

Veenstra, D.L., You, J.H.S., Rieder, M.J., Farin, F.M., Wilkerson, H.W., Blough, D.K., Cheng, G. and Rettie, A.E. (2005). Association of Vitamin K epoxide reductase complex 1 (VKORC1) variants with warfarin dose in a Hong Kong Chinese patient population. *Pharmacogenetics and genomics.* **15** (10), 687-691.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., *et al.* (2001). The sequence of the human genome. *Science.* **291** (5507), 1304-51.

Veyrieras, J.B., Kudaravalli, S., Kim, S.Y., Dermitzakis, E.T., Gilad, Y., Stephens, M. and Pritchard, J.K. (2008). High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS genetics.* **4** (10), e1000214.

Vlieg, A.V.H., Van Der Linden, I.K., Bertina, R.M. and Rosendaal, F.R. (2000). High levels of factor IX increase the risk of venous thrombosis. *Blood.* ***95*** *(12), 3678-3682.*

Vogel, F. (1959). Moderne probleme der humangenetik. *Ergeb. Inn. Med. Kinderheild.* ***12****, 52-62.*

Voora, D., Mcleod, H.L., Eby, C. and Gage, B.F. (2005). The pharmacogenetics of coumarin therapy. *Pharmacogenomics.* **6** (5), 503-513.

Wadelius, M., Sorlin, K., Wallerman, O., Karlsson, J., Yue, Q.Y., Magnusson, P.K., Wadelius, C. and Melhus, H. (2004). Warfarin sensitivity related to CYP2C9, CYP3A5, ABCB1 (MDR1) and other factors. *Pharmacogenomics J.* **4** (1), 40-8.

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

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

Wadelius, M., Chen, L.Y., Lindh, J.D., Eriksson, N., Ghori, M.J.R., Bumpstead, S., Holm, L., Mcginnis, R., Rane, A. and Deloukas, P. (2009). The largest prospective warfarin-treated cohort supports genetic forecasting. *Blood.* **113** (4), 784-792.

Wajih, N., Sane, D.C., Hutson, S.M. and Wallin, R. (2004). The inhibitory effect of calumenin on the vitamin K-dependent γ-carboxylation system: Characterization of the system in normal and warfarin-resistant rats. *Journal of Biological Chemistry.* **279** (24), 25276-25283.

Wajih, N., Sane, D.C., Hutson, S.M. and Wallin, R. (2005). Engineering of a recombinant vitamin K-dependent γ-carboxylation system with enhanced γ-carboxyglutamic acid forming capacity: Evidence for a functional CXXC redox center in the system. *Journal of Biological Chemistry.* **280** (11), 10540-10547.

Wallin, R. and Hutson, S. (1982). Vitamin K-dependent carboxylation. Evidence that at least two microsomal dehydrogenases reduce vitamin K1 to support carboxylation. *Journal of Biological Chemistry.* **257** (4), 1583-1586.

Wallin, R. and Martin, L.F. (1985). Vitamin K-dependent carboxylation and vitamin K metabolism in liver. Effects of warfarin. *Journal of Clinical Investigation.* **76** (5), 1879-1884.

Wallin, R. and Martin, L.F. (1987). Warfarin poisoning and vitamin K antagonism in rat and human liver. Design of a system in vitro that mimics the situation in vivo. *Biochemical Journal.* **241** (2), 389-396.

Wallin, R., Hutson, S.M., Cain, D., Sweatt, A. and Sane, D.C. (2001). A molecular mechanism for genetic warfarin resistance in the rat. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology.* **15** (13), 2542-2544.

Wang, D., Chen, H., Momary, K.M., Cavallari, L.H., Johnson, J.A. and Sadee, W. (2008). Regulatory polymorphism in vitamin K epoxide reductase complex subunit 1 (VKORC1) affects gene expression and warfarin dose requirement. *Blood.* **112** (4), 1013-21.

Wang, M.Z., Saulter, J.Y., Usuki, E., Cheung, Y.L., Hall, M., Bridges, A.S., Loewen, G., Parkinson, O.T., Stephens, C.E., Allen, J.L., Zeldin, D.C., Boykin, D.W., Tidwell, R.R., Parkinson, A., Paine, M.F. and Hall, J.E. (2006). CYP4F enzymes are the major enzymes in human liver microsomes that catalyze the O-demethylation of the antiparasitic prodrug DB289 [2,5-bis(4-amidinophenyl)furan-bis-O-methylamidoxime]. *Drug metabolism and disposition: the biological fate of chemicals.* **34** (12), 1985-94.

Wang, M.Z., Wu, J.Q., Bridges, A.S., Zeldin, D.C., Kornbluth, S., Tidwell, R.R., Hall, J.E. and Paine, M.F. (2007). Human enteric microsomal CYP4F enzymes O-demethylate the antiparasitic prodrug pafuramidine. *Drug metabolism and disposition: the biological fate of chemicals.* **35** (11), 2067-75.

Wang, Y., Bell, J.C., Keeney, D.S. and Strobel, H.W. (2010). Gene regulation of CYP4F11 in human keratinocyte HaCaT cells. *Drug metabolism and disposition: the biological fate of chemicals.* **38** (1), 100-7.

Wang, Z., Wang, M., Finn, F. and Carr, B.I. (1995). The growth inhibitory effects of vitamins K and their actions on gene expression. *Hepatology.* **22** (3), 876-82.

Watzka, M., Geisen, C., Bevans, C.G., Sittinger, K., Spohn, G., Rost, S., Seifried, E., Müller, C.R. and Oldenburg, J. (2011). Thirteen novel VKORC1 mutations associated with oral anticoagulant resistance: insights into improved patient diagnosis and treatment. *Journal of Thrombosis and Haemostasis.* **9** (1), 109-118.

White, R.H., Beyth, R.J., Zhou, H. and Romano, P.S. (1999). Major bleeding after hospitalization for deep-venous thrombosis. *American Journal of Medicine.* **107** (5), 414-424.

Whitlon, D.S., Sadowski, J.A. and Suttie, J.W. (1978). Mechanism of coumarin action: Significance of vitamin K epoxide reductase inhibition. *Biochemistry.* **17** (8), 1371-1377.

Wiffen, P., Gill, M., Edwards, J. And Moore, A. (2002). Adverse drug reactions in hospital patients. *Bandolier Extra* http://www.jr2.ox.ac.uk/bandolier/Extraforbando/ADRPM.pdf//.

Wigginton, J.E., Cutler, D.J. and Abecasis, G.R. (2005). A note on exact tests of Hardy-Weinberg equilibrium. *American journal of human genetics.* **76** (5), 887-93.

Williams, D.B. and Karl, R.C. (1979). A simple technic for predicting daily maintenance dose of warfarin. *American Journal of Surgery.* **137** (4), 572-576.

Wilting, J., Van Der Giesen, W.F. and Janssen, L.H.M. (1980). The effect of albumin conformation on the binding of warfarin to human serum albumin. The dependence of the binding of warfarin to human serum albumin on the hydrogen, calcium, and chloride ion concentrations as studied by circular dichroism, fluorescence, and equilibrium dialysis. *Journal of Biological Chemistry.* **255** (7), 3032-3037.

Wu, A.H.B., Wang, P., Smith, A., Haller, C., Drake, K., Linder, M. and Valdes Jr, R. (2008). Dosing algorithm for warfarin using CYP2C9 and VKORC1 genotyping from a multi-ethnic population: Comparison with other equations. *Pharmacogenomics.* **9** (2), 169-178.

Wu, S.M., Stafford, D.W., Frazier, L.D., Fu, Y.Y., High, K.A., Chu, K., Sanchez-Vega, B. and Solera, J. (1997). Genomic sequence and transcription start site for the human γ-glutamyl carboxylase. *Blood.* **89** (11), 4058-4062.

Wu, W., Bancroft, J.D. and Suttie, J.W. (1996). Differential effects of warfarin on the intracellular processing of vitamin K-dependent proteins. *Thromb Haemost.* **76** (1), 46-52.

Wysowski, D.K., Nourjah, P. and Swartz, L. (2007). Bleeding complications with warfarin use: a prevalent adverse effect resulting in regulatory action. *Archives of internal medicine.* **167** (13), 1414-9.

Xie, H.G., Prasad, H.C., Kim, R.B. and Stein, C.M. (2002). CYP2C9 allelic variants: Ethnic distribution and functional significance. *Advanced Drug Delivery Reviews.* **54** (10), 1257-1270.

Yeo, G. and Burge, C.B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol.* **11** (2-3), 377-94.

Yin, T. and Miyata, T. (2007). Warfarin dose and the pharmacogenomics of CYP2C9 and VKORC1 - rationale and perspectives. *Thromb Res.* **120** (1), 1-10.

You, J.H.S., Chan, F.W.H., Wong, R.S.M. and Cheng, G. (2004). The potential clinical and economic outcomes of pharmacogenetics-oriented management of warfarin therapy - A decision analysis. *Thrombosis and Haemostasis.* **92** (3), 590-597.

Yuan, H.Y., Chen, J.J., Lee, M.T., Wung, J.C., Chen, Y.F., Charng, M.J., Lu, M.J., Hung, C.R., Wei, C.Y., Chen, C.H., Wu, J.Y. and Chen, Y.T. (2005). A novel functional VKORC1 promoter polymorphism is associated with inter-individual and inter-ethnic differences in warfarin sensitivity. *Human Molecular Genetics.* **14** (13), 1745-51.

Zambon, C.F., Pengo, V., Padrini, R., Basso, D., Schiavon, S., Fogar, P., Nisi, A., Frigo, A.C., Moz, S., Pelloso, M. and Plebani, M. (2011). VKORC1, CYP2C9 and CYP4F2 genetic-based algorithm for warfarin dosing: an Italian retrospective study. *Pharmacogenomics.* **12** (1), 15-25.

Zeller, T., Wild, P., Szymczak, S., Rotival, M., Schillert, A., Castagne, R., Maouche, S., Germain, M., Lackner, K., Rossmann, H., Eleftheriadis, M., Sinning, C.R., Schnabel, R.B., Lubos, E., Mennerich, D., Rust, W., Perret, C., Proust, C., Nicaud, V., Loscalzo, J., Hubner, N., Tregouet, D., Munzel, T., Ziegler, A., Tiret, L., Blankenberg, S. and Cambien, F. (2010). Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PloS one.* **5** (5), e10693.

Zhang, J.E., Jorgensen, A.L., Alfirevic, A., Williamson, P.R., Toh, C.H., Park, B.K. and Pirmohamed, M. (2009). Effects of CYP4F2 genetic polymorphisms and haplotypes on clinical outcomes in patients initiated on warfarin therapy. *Pharmacogenetics and genomics.* **19** (10), 781-9.

Zhang, X. and Hardwick, J.P. (2000). Regulation of CYP4F2 leukotriene B4 omega-hydroxylase by retinoic acids in HepG2 cells. *Biochemical and biophysical research communications.* **279** (3), 864-71.

Zhang, X., Chen, L. and Hardwick, J.P. (2000). Promoter activity and regulation of the CYP4F2 leukotriene B(4) omega-hydroxylase gene by peroxisomal proliferators and retinoic acid in HepG2 cells. *Archives of biochemistry and biophysics.* **378** (2), 364-76.

Zhang, Z., Fasco, M.J., Huang, Z., Guengerich, F.P. and Kaminsky, L.S. (1995). Human cytochromes P4501A1 and P4501A2: R-warfarin metabolism as a probe. *Drug Metabolism and Disposition.* **23** (12), 1339-1345.

Zhou, C., Verma, S. and Blumberg, B. (2009). The steroid and xenobiotic receptor (SXR), beyond xenobiotic metabolism. *Nuclear receptor signaling.* **7**, e001.

Zimmermann, A. and Matschiner, J.T. (1974). Biochemical basis of hereditary resistance to warfarin in the rat. *Biochemical Pharmacology.* **23** (6), 1033-40.

# Appendix 1

# CONTENTS

**APPENDIX 1.1. *CYP4F2* PCR AND EXTENSION PRIMERS FOR MALDI-TOF MS..... 262**

**APPENDIX 1.2. DETAILS AND ALLELE FREQUENCIES OF 80 *CYP4F2* SNPS GENOTYPED IN 311 WARFARIN PATIENTS..... 265**

**APPENDIX 1.3. UNIVARIATE TESTS FOR ASSOCIATION BETWEEN SNPS IN THE *CYP4F2* GENE AND CLINICAL OUTCOMES OF RESPONSE TO WARFARIN..... 269**

**Appendix 1.1. *CYP4F2* PCR and extension primers for MALDI-TOF MS.**

| SNP | Assay | Chromosomal position | Forward Amplification Primer Sequence (5'>3') | Reverse Amplification Primer Sequence (5'>3') | Amplification Product (bp) | Amplification Direction | Extension Primer Sequence (5'>3') |
|---|---|---|---|---|---|---|---|
| rs3093168 | 1 | 15857245 | ACGTTGGATGCTCCAGAGCACAGAAAGATG | ACGTTGGATGTCCGTCCCATCCTTATTCAG | 104 | F | CTCCCACACTGCATT |
| rs3093200 | 1 | 15850589 | ACGTTGGATGAGAGGGCGGACTTTGGCTG | ACGTTGGATGAGAGTGGGTCTCTGCAGAAC | 80 | R | GCAGAACTCAGCTCA |
| rs3761014 | 1 | 15872763 | ACGTTGGATGAGAAACCTAGCATGAGGCAG | ACGTTGGATGAGCAACAAGCTTCCTCTCTG | 113 | F | cCTCCCTCGACTGCTT |
| rs3093128 | 1 | 15866104 | ACGTTGGATGATCCCTGTGTCTCTGATTGC | ACGTTGGATGACAAACTTCTTTGTGCCCCC | 100 | R | GTAGCTCTGGAAGACC |
| rs3093180 | 1 | 15856148 | ACGTTGGATGTGTGGGAGCTGAAAAAGTGG | ACGTTGGATGAATCAACTTCTCCTCACCCC | 100 | F | TCCTCACCCCTCCCTCC |
| rs3093105 | 1 | 15869388 | ACGTTGGATGATGGGCCAGGAGCCAGGAG | ACGTTGGATGATGTCCCAGCTGAGCCTGTC | 119 | F | ccTCCTGGCTGGGCCTC |
| rs3093110 | 1 | 15868784 | ACGTTGGATGTATAAGAGAAAGGTGCCTTC | ACGTTGGATGTGTTATGAGGGTACAGCTGG | 103 | F | acaGTCTCCCACTTCCAC |
| rs3093198 | 1 | 15851048 | ACGTTGGATGCAACCCAACCGTACTCTATG | ACGTTGGATGTTGTAGATGGTCCAAGTTCC | 104 | F | tcCCAAGTTCCAGCTCTC |
| rs3093195 | 1 | 15852433 | ACGTTGGATGGCTCTTTTTCTTTGGATTC | ACGTTGGATGCAACCTGGGCGATAGAGACT | 118 | R | GCGATAGAGACTGCATCT |
| rs1126433 | 1 | 15850405 | ACGTTGGATGAAGACGCTTGTGCGTGAATG | ACGTTGGATGGGGTTGTTTTCATCGTTTGG | 85 | R | tGCTACTGTGAATAGGGC |
| rs1272 | 1 | 15850040 | ACGTTGGATGGGGATGGTGAAAATGTTCCG | ACGTTGGATGTTAGTGCACTCACAGTGTCG | 85 | R | ccccCGTGCTACCTTCGTCA |
| rs3093166 | 1 | 15857349 | ACGTTGGATGCCCTAATTCCTACTCTCCAG | ACGTTGGATGGTTCCGGACAAGAATGGATG | 101 | R | CCCTAGGTCTCCCTTATAAA |
| rs3093150 | 1 | 15862629 | ACGTTGGATGTAGAAACTCTGCTCCACCTC | ACGTTGGATGTGCAGGTCTGAGAGATTAGG | 100 | F | tAGAGATTAGGGAACTAACC |
| rs3093211 | 1 | 15848967 | ACGTTGGATGCCTGCAATTTCAGCTACTCG | ACGTTGGATGTCTTGGCTCACTGAAACCTC | 99 | F | ggtaaTGAAACCTCTGCCTTC |
| rs3093106 | 1 | 15869257 | ACGTTGGATGTGGTGTCCCCAAAACCAGTT | ACGTTGGATGCCTACGCCTTCTATGACAAC | 99 | F | ttctGTGTTTCCCACAACCCCC |
| rs2108622 | 1 | 15851431 | ACGTTGGATGGTGTTTTCGGAACCCATCAC | ACGTTGGATGGCCTTGGAATGGACAAAAAC | 106 | F | ctcaaCTCAGGGTCCGGCCACA |
| rs3093204 | 1 | 15850167 | ACGTTGGATGAAAGCCACCAATCCGCTATG | ACGTTGGATGGGTCACATAGTGTACTGTCC | 101 | F | ctccACGAAATTTCCAGAACAG |
| rs3093194 | 1 | 15852676 | ACGTTGGATGGAAATATGAAGATTGGCATAG | ACGTTGGATGCCTATTAGCTGAGTTTGATG | 118 | F | ggggaTGAGTTTGATGGCACAA |
| rs3093208 | 1 | 15849582 | ACGTTGGATGCCTTCTCAGTGATCTACTTC | ACGTTGGATGGGTGTGGTTAGACACCTTTC | 118 | F | ggTAGACACCTTTCATAGAACAG |
| rs7252046 | 1 | 15832473 | ACGTTGGATGTACAGTTAGGTTCCAGCGAG | ACGTTGGATGCACACACAAACAAGGCTACG | 94 | F | CGTAATGATCAGTTGACAAAATG |
| rs3093141 | 1 | 15863521 | ACGTTGGATGTCCTCCCTATGATTGTCTCC | ACGTTGGATGGCTAAGCCTTGTCTTACTCC | 104 | R | cccaCTTGGAACCCTGCCTGCATT |
| rs3093209 | 1 | 15849420 | ACGTTGGATGTCCTCACATCGATCTTACAG | ACGTTGGATGTGCCAACTGCTCAGATTCAC | 113 | R | ggacaCTGACTCCACCCTTCACTGA |
| rs2079288 | 1 | 15825203 | ACGTTGGATGTTTCAGAGCCCTGAATGTGC | ACGTTGGATGGGGAAAGTCATACAGTTGGC | 107 | F | cttaACTCCTTCTGAGTAGGCTTAG |
| rs3093135 | 1 | 15865371 | ACGTTGGATGACTTCCACCTGTCTAACCTC | ACGTTGGATGAGACCTGGCCATATAACAGC | 100 | R | tggaCATATAACAGCCATTGTAGAG |
| rs3093201 | 1 | 15850454 | ACGTTGGATGTTCACGCACAAGCGTCTTTC | ACGTTGGATGCACAGTCCTGTATTCCATCC | 100 | F | aATCCTAGATATCTACTCAAAATAAT |
| rs2189784 | 1 | 15820200 | ACGTTGGATGTATGTCTGATGCAGTAACTC | ACGTTGGATGCAAGATGAGATCCTCTGCCC | 107 | F | taaaTCCTGGTGTTTTTTTTGTTCTC |
| rs736089 | 1 | 15866300 | ACGTTGGATGGGAAAAGGCCACCTTGATAG | ACGTTGGATGACAAGGTCACTGGAAATCTG | 100 | R | ccccgCTCTTGGCTTTAGCTTTCTCCA |
| rs2886296 | 1 | 15855478 | ACGTTGGATGCCCAGAAATTCCACTACTGA | ACGTTGGATGACGGAAGTATAGATATCTC | 110 | F | cCGGAAGTATAGATATCTCTTCAACAG |
| rs3093092 | 1 | 15870127 | ACGTTGGATGGGATGGATACATTAGGTAG | ACGTTGGATGTAGCTATCTATCCATCTGTC | 93 | F | aTATCCATCTGTCTAATATATCCATTAT |
| rs17756654 | 1 | 15849725 | ACGTTGGATGGAGATCTCTACCCCAAAAGT | ACGTTGGATGAGTGTTTTGTCTCACTTTG | 120 | R | ccctTTTTGTCTCACTTTGTAAGAATTA |
| rs3093116 | 1 | 15866574 | ACGTTGGATGATGGTCTCTGTAGGATGCTG | ACGTTGGATGTTCCTTGACCCCTGTGCTTC | 98 | F | ggaaTGTGCTTCCCATCGTTGGACGGGC |

**Appendix 1.1. *CYP4F2* PCR and extension primers for MALDI-TOF MS continued.**

| SNP | Assay | Chromosomal position | Forward Amplification Primer Sequence (5'>3') | Reverse Amplification Primer Sequence (5'>3') | Amplification Product (bp) | Amplification Direction | Extension Primer Sequence (5'>3') |
|---|---|---|---|---|---|---|---|
| rs2074901 | 1 | 15858422 | ACGTTGGATGTTACTGGCAACCTGAGTTAC | ACGTTGGATGTGTGGGCAAATCTTTGAAGC | 107 | R | ggggTAATTTCCAAGGGAAATTGCCATT |
| rs3093097 | 2 | 15869690 | ACGTTGGATGTAAAGTCCAGAAAGGCCCAG | ACGTTGGATGCCGAGTAACTGAAACCAAGC | 86 | F | CCCCTGGCAACCTCA |
| rs3093113 | 2 | 15867545 | ACGTTGGATGGAATAGATTAGACACACCCAG | ACGTTGGATGCTAGGTGTCTACTTGACTGG | 98 | R | GGATGCCCAGGTAGT |
| rs3093098 | 2 | 15869512 | ACGTTGGATGACCTCCAGGGACAGTGGAGA | ACGTTGGATGCCTTTTTACCCATCCCTGAG | 98 | F | GCTCTACCTCCGGCAC |
| rs3093167 | 2 | 15857292 | ACGTTGGATGATGCAGTGTGGGAGTGTTTG | ACGTTGGATGCATCCATTCTTGTCCGGAAC | 94 | F | TCCGGAACACTTGCAC |
| rs3093114 | 2 | 15867413 | ACGTTGGATGTCCCATCCAGACCTTAAAGC | ACGTTGGATGAGAGGAGGGCATGAGAGTTC | 90 | F | ggGACTCAGCTGGTGGC |
| rs984692 | 2 | 15867074 | ACGTTGGATGGGGATATCTCTATCCAGGTC | ACGTTGGATGAACCCTAAGAGACCTCAACC | 101 | R | GACCTCAACCCAAGGCCC |
| rs2215092 | 2 | 15866521 | ACGTTGGATGCATCCTACAGAGACCATGGC | ACGTTGGATGAATAACCAGAGACAGCTGGG | 96 | F | AGAGTCACTCAGCATTTT |
| rs3093103 | 2 | 15869434 | ACGTTGGATGACAGGCTCAGCTGGGACATC | ACGTTGGATGTTCTCTGGGCCTCAGGATCT | 86 | F | cccaTCACCGTCCATCCCG |
| rs3093212 | 2 | 15848966 | ACGTTGGATGCCTGCAATTTCAGCTACTCG | ACGTTGGATGTCTTGGCTCACTGAAACCTC | 99 | F | cccttAAACCTCTGCCTTCC |
| rs3093210 | 2 | 15849401 | ACGTTGGATGTCAAGGTCACTCAGTGAAGG | ACGTTGGATGACGATAAGGCATGACCTCTG | 93 | R | ttgcgGCTCAGATTCACCCC |
| rs3093129 | 2 | 15866019 | ACGTTGGATGAAAGTCACTTGCCCACAGTC | ACGTTGGATGTTAAGGATTCAGGCCATGGG | 104 | F | GTGCAGACAGACCTAGTCTT |
| rs3093163 | 2 | 15857497 | ACGTTGGATGGTTATAGGCCCTTAGGACAG | ACGTTGGATGGCTGAACACGTGCTAAGTGG | 112 | R | gggGTGCTAAGTGGCGTTAT |
| rs3093121 | 2 | 15866277 | ACGTTGGATGTTCTGGGACTCATCGTTGAC | ACGTTGGATGCTCTTGGCTTTAGCTTTCTC | 100 | F | gaGCTTTTCCTATCAAGGTGG |
| rs3093169 | 2 | 15857008 | ACGTTGGATGGTCTTCTTGACCTAGAGCAC | ACGTTGGATGAAGTGTCTCAGTTGAGAGCC | 101 | F | AGTAGATAATTGGGTACAGTT |
| rs2074902 | 2 | 15869099 | ACGTTGGATGATCTGGGGTGGCAGAGAAGC | ACGTTGGATGGTAGGCACCTCACAGAAATG | 93 | F | ggAGCAGAGGAATGAGTAAGA |
| rs3093173 | 2 | 15856788 | ACGTTGGATGGATAAAACCGGAGATGGTAG | ACGTTGGATGATGGCTCCCTATTGCCATTG | 114 | R | ACTGAAGCCATTAACATTTTCC |
| rs2016503 | 2 | 15866185 | ACGTTGGATGATCAGAGACACAGGGATTGG | ACGTTGGATGCCCTAGAGTTCCCAGAAATG | 108 | R | agatTTCCCAGAAATGGCCCATG |
| rs3093192 | 2 | 15853319 | ACGTTGGATGCCTATCTACTATCTATATATC | ACGTTGGATGGAGAGATAGAGAATAGACAG | 98 | R | CATACATAGATGATAGAAACAGA |
| rs3093134 | 2 | 15865396 | ACGTTGGATGTGGACTTCCACCTGTCTAAC | ACGTTGGATGAGACCTGGCCATATAACAGC | 103 | R | ccctCCCCAGACATGAGCAAACAG |
| rs3093216 | 2 | 15848737 | ACGTTGGATGCACCCAGCCCCAAGTTTTTA | ACGTTGGATGCACCATTACTTCTCACTGGG | 111 | R | agaggGAATGTACTTCTGGTCAAC |
| rs3093158 | 2 | 15861166 | ACGTTGGATGTCAATCACCTTCCATGGCTC | ACGTTGGATGGTAGAAGGGAGCTTCATGTG | 105 | F | gTCATGTGAAATGTCAGATGAAAG |
| rs3093193 | 2 | 15852914 | ACGTTGGATGGCCACATACACATTGATGGG | ACGTTGGATGTTCCTCCTTTTGTGTAGCTG | 99 | F | TTGTGTAGCTGTATAGTATTCCATC |
| rs3093091 | 2 | 15870133 | ACGTTGGATGGGATGGATACATTAGGTAG | ACGTTGGATGTAGCTATCTATCCATCTGTC | 93 | F | CTATCTATCCATCTGTCTAATATATC |
| rs3093144 | 2 | 15863297 | ACGTTGGATGGGGAAGAATTGTGGCAAAGG | ACGTTGGATGGAAACAGTAACAAGCTAAC | 105 | R | AGTTAAAAAAAAAATCCTAGATACTT |
| rs1558139 | 2 | 15858564 | ACGTTGGATGTGGTGTGCTTTGTTCTGGAG | ACGTTGGATGGAGACGTTTGGGTTGTAAGG | 100 | F | aataGACTATGGCATCTCACTGAAGA |
| rs3093203 | 2 | 15850184 | ACGTTGGATGAAAGCCACCAATCCGCTATG | ACGTTGGATGGGTCACATAGTGTACTGTCC | 101 | F | ctTCACATAGTGTACTGTCCTTTTATA |
| rs3093197 | 2 | 15851128 | ACGTTGGATGTTTTATTCCCTTCTCGGCAG | ACGTTGGATGTTTTCCCCACTGTCATGCCC | 119 | R | gccccGACCCCTGCACCCATCAACCCGT |
| rs3093156 | 2 | 15861609 | ACGTTGGATGAATCTCATGCTGATCCCCAC | ACGTTGGATGCAGGATCTACCTCCTATCAG | 101 | R | gagCTACCTCCTATCAGTAAACAGAGTA |
| rs2365178 | 3 | 15866003 | ACGTTGGATGTTAAGGATTCAGGCCATGGG | ACGTTGGATGAAAGTCACTTGCCCACAGTC | 104 | R | CACAGTCCCAGACCA |
| rs3093100 | 3 | 15869469 | ACGTTGGATGTCTCCACTGTCCCTGGAGGT | ACGTTGGATGATGGACGGTGAGATCCTGAG | 80 | R | CTGAGGCCCAGAGAA |

**Appendix 1.1. *CYP4F2* PCR and extension primers for MALDI-TOF MS continued.**

| SNP | Assay | Chromosomal position | Forward Amplification Primer Sequence (5'>3') | Reverse Amplification Primer Sequence (5'>3') | Amplification Product (bp) | Amplification Direction | Extension Primer Sequence (5'>3') |
|---|---|---|---|---|---|---|---|
| rs3093115 | 3 | 15866615 | ACGTTGGATGTCTCCAAGAGGAGACTCCAC | ACGTTGGATGTCCAACGATGGGAAGCACAG | 87 | R | GGGGTCAAGGAAAGGG |
| rs3093170 | 3 | 15856974 | ACGTTGGATGTTCCTCCCTCTCTCTTCCTC | ACGTTGGATGGATAATTGGGTACAGTTGGG | 103 | F | GTCAAGAAGACTGCCTG |
| rs3093184 | 3 | 15855730 | ACGTTGGATGGTGAAACCCCAGCCCTATTA | ACGTTGGATGTATTCTTCTGCCTCAGCCTC | 116 | F | ttTGGGTGCACACCACCA |
| rs2074900 | 3 | 15857820 | ACGTTGGATGAGTGGTCTCTCCTGGGTCCT | ACGTTGGATGAGAAGTTCTTGCACCTCCTG | 103 | F | ggtTGGGTGCTTTGCAAG |
| rs3093126 | 3 | 15866160 | ACGTTGGATGAAGACTGGGCAATCAGAGAC | ACGTTGGATGGGCCCATGATTTACTTCTCC | 100 | F | ttcgCCTGGCCTGGATCAC |
| rs3093124 | 3 | 15866232 | ACGTTGGATGTCAACGATGAGTCCCAGAAG | ACGTTGGATGAATCATGGGCCATTTCTGGG | 99 | R | gaagGGTGGCCTTTCCCAT |
| rs3093199 | 3 | 15850871 | ACGTTGGATGGCTTGCAGGATATGCAAGCC | ACGTTGGATGCCCACCTAGAGCAGTTTGAG | 99 | R | cGGAATAGAGAAGGGGGTA |
| rs3093207 | 3 | 15849658 | ACGTTGGATGCTTCATCAAGTAACTTTTGGG | ACGTTGGATGAAGGTGTCTAACCACACCTG | 112 | R | CTATTGCTGTGAAAATGAGA |
| rs3093226 | 3 | 15853780 | ACGTTGGATGTCCCAAAGTGCTGGGATTAC | ACGTTGGATGGCTCATAGAAACAGGAGTGG | 84 | R | ccaccGGAGTGGAGGTCGGG |
| rs3093160 | 3 | 15857907 | ACGTTGGATGAAGGCAGGCTTGGGTCTCT | ACGTTGGATGAATGCATGGGTGCTGTCTAC | 80 | F | cggttCTGTCTACCTTCGGGT |
| rs3093112 | 3 | 15867611 | ACGTTGGATGGGTCAAACATTATTTCTGGG | ACGTTGGATGATGGATAGATGGGTAGGCAG | 112 | F | gggcTGGGTAGGCAGATAGAT |
| rs2006193 | 3 | 15865800 | ACGTTGGATGTGCATGGAGGATGTGAGGAG | ACGTTGGATGCTCTCCCTGAGAATTTCAAG | 99 | F | aTGAGAATTTCAAGCAGCATCT |
| rs12984060 | 3 | 15853553 | ACGTTGGATGGGTGCAAGAAGAATCTGTTC | ACGTTGGATGGGCTGAATGGTATTCCATTG | 111 | R | CATGGAAAATGAGCATTAAATTA |
| rs12610189 | 3 | 15839641 | ACGTTGGATGGGCTAGAACTTGAGCAGTTG | ACGTTGGATGACTTCCTTATGACAACCACG | 107 | F | ttcaTTATGACAACCACGTTAAAA |
| rs3093227 | 3 | 15852520 | ACGTTGGATGGAATCCAAAGAAAAAGAGC | ACGTTGGATGGGTTGAACATCTCTTTCACG | 105 | F | cGTTTGTTTCTTTTTATGGAAGTG |
| rs3093122 | 3 | 15866276 | ACGTTGGATGCTCTTGGCTTTAGCTTTCTC | ACGTTGGATGTTCTGGGACTCATCGTTGAC | 100 | R | ggagaGTTGACATCATGAGGAAAA |
| rs3093145 | 3 | 15862852 | ACGTTGGATGAGGACTCAACGAAGGACTAC | ACGTTGGATGATGTTGGATTTCCCCGTCTG | 104 | R | cttATTTCAAAAATATCCCTGTTTT |
| rs3093206 | 3 | 15849753 | ACGTTGGATGGCTAATTCTTACAAAGTGAG | ACGTTGGATGCCTCAAAATGTGTGTGTTAAG | 120 | F | AAAAAAAGACATCATATTAGAAAATAA |

Chromosomal positions are given in base pairs from the p-telomere of chromosome 19, as per HapMap Data release 23, March 2008, NCBI B36 assembly, dbSNP b126.

MALDI-TOF MS analysis was performed with 3 different assays for genotype and haplotype determination.

All amplification primers have a tag sequence ACGTTGGATG that permits more efficient amplification.

Nucleotides not capitalized are mismatches introduced by the software MassARRAY Assay Design to enhance detection by MS.

F: forward orientation; R: reverse orientation.

**Appendix 1.2. Details and allele frequencies of 80 *CYP4F2* SNPs genotyped in 311 warfarin patients.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| - | rs2189784 | 15820200 | G > A | Downstream of CYP4F2 | 0.655 | 100 | 0.445 | 0.381 | *HapMap R23* |
| - | rs2079288 | 15825203 | T > C | Downstream of CYP4F2 | 0.993 | 100 | 0.238 | 0.217 | *HapMap R23* |
| - | rs7252046 | 15832473 | T > C | Downstream of CYP4F2 | 0.4 | 100 | 0.243 | 0.325 | *HapMap R23* |
| - | rs12610189 | 15839641 | T > G | Downstream of CYP4F2 | 0.335 | 100 | 0.375 | 0.300 | *HapMap R23* |
| *CYP4F2* | rs3093216 | 15848737 | T > C | 3` near gene | 0.251 | 100 | 0.371 | 0.300 | *NCBI B36* |
| *CYP4F2* | rs3093212[b] | 15848966 | A > A | 3` near gene | 0.139 | 1.9 | 0.5 | 0.475 | *NCBI B36* |
| *CYP4F2* | rs3093211[c] | 15848967 | C > T | 3` near gene | 0 | 0 | 0 | 0.200 | *NCBI B36* |
| *CYP4F2* | rs3093210[c] | 15849401 | G > A | 3` near gene | 0 | 100 | 0.121 | 0.142 | *NCBI B36* |
| *CYP4F2* | rs3093209 | 15849420 | C > G | 3` near gene | 0.251 | 100 | 0.371 | 0.347 | *NCBI B36* |
| *CYP4F2* | rs3093208[a] | 15849582 | T > C | 3` near gene | 1 | 100 | 0.003 | 0.000 | *HapMap R23* |
| *CYP4F2* | rs3093207 | 15849658 | A > G | 3` near gene | 0.251 | 100 | 0.371 | 0.300 | *HapMap R23* |
| *CYP4F2* | rs17756654 | 15849725 | C > A | 3` near gene | 1 | 100 | 0.05 | 0.033 | *NCBI B36* |
| *CYP4F2* | rs3093206 | 15849753 | AA > DEL | 3` near gene | 0.029 | 100 | 0.383 | 0.217 | *NCBI B36* |
| *CYP4F2* | rs1272 | 15850040 | G > C | Exon 13 | 0.193 | 100 | 0.167 | 0.180 | *HapMap R23* |
| *CYP4F2* | rs3093204 | 15850167 | G > A | Exon 13 | 0.924 | 100 | 0.071 | 0.065 | *NCBI B36* |
| *CYP4F2* | rs3093203 | 15850184 | C > T | Exon 13 | 1 | 98.7 | 0.337 | 0.292 | *NCBI B36* |
| *CYP4F2* | rs1126433[c] | 15850405 | A > G | Exon 13 | 0 | 100 | 0.423 | 0.477 | *NCBI B36* |
| *CYP4F2* | rs3093201[a] | 15850454 | T > C | Exon 13 | 1 | 100 | 0 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs3093200 | 15850589 | C > A | Exon 13, Missense, Met519Leu | 0.453 | 100 | 0.068 | 0.083 | *HapMap R23* |
| *CYP4F2* | rs3093199[b] | 15850871 | A > T | Intron 12 | 0 | 0 | 0 | 0.284 | *NCBI B36* |
| *CYP4F2* | rs3093198 | 15851048 | C > T | Intron 12 | 0.548 | 100 | 0.244 | 0.340 | *NCBI B36* |
| *CYP4F2* | rs3093197[a] | 15851128 | A > T | Intron 12 | 1 | 100 | 0 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs2108622 | 15851431 | C > T | Exon 11, Missense, Met433Val | 0.636 | 100 | 0.33 | 0.233 | *HapMap R23* |
| *CYP4F2* | rs3093195 | 15852433 | G > A | Intron 9 | 1 | 100 | 0.338 | 0.325 | *NCBI B36* |
| *CYP4F2* | rs3093227 | 15852520 | G > T | Intron 9 | 1 | 100 | 0.002 | 0.023 | *NCBI B36* |

**Appendix 1.2. Details and allele frequencies of 80 *CYP4F2* SNPs genotyped in 311 warfarin patients continued.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| *CYP4F2* | rs3093194 | 15852676 | G > A | Intron 9 | 1 | 100 | 0.339 | 0.332 | *HapMap R23* |
| *CYP4F2* | rs3093193 | 15852914 | G > C | Intron 9 | 0.251 | 100 | 0.371 | 0.333 | *NCBI B36* |
| *CYP4F2* | rs3093192[a] | 15853319 | C > T | Intron 9 | 1 | 100 | 0.002 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs12984060 | 15853553 | G > A | Intron 9 | 1 | 100 | 0.291 | 0.375 | *NCBI B36* |
| *CYP4F2* | rs3093226[b] | 15853780 | T > T | Intron 9 | 0 | 0 | 0 | 0.071 | *NCBI B36* |
| *CYP4F2* | rs2886296 | 15855478 | A > T | Intron 9 | 1 | 100 | 0.035 | 0.065 | *NCBI B36* |
| *CYP4F2* | rs3093184[b] | 15855730 | C > T | Intron 9 | 1 | 57.2 | 0.048 | 0.329 | *NCBI B36* |
| *CYP4F2* | rs3093180 | 15856148 | C > T | Intron 9 | 1 | 100 | 0.177 | 0.153 | *NCBI B36* |
| *CYP4F2* | rs3093173 | 15856788 | G > A | Intron 9 | 1 | 100 | 0.177 | 0.095 | *NCBI B36* |
| *CYP4F2* | rs3093170 | 15856974 | T > G | Intron 9 | 0 | 100 | 0.341 | 0.432 | *NCBI B36* |
| *CYP4F2* | rs3093169 | 15857008 | G > A | Intron 9 | 0 | 99 | 0.159 | 0.139 | *NCBI B36* |
| *CYP4F2* | rs3093168 | 15857245 | C > T | Intron 9 | 0.952 | 100 | 0.259 | 0.345 | *HapMap R23* |
| *CYP4F2* | rs3093167 | 15857292 | T > C | Intron 9 | 1 | 100 | 0.177 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093166 | 15857349 | A > T | Intron 9 | 0.779 | 100 | 0.257 | 0.345 | *HapMap R23* |
| *CYP4F2* | rs3093163[a] | 15857497 | G > A | Intron 9 | 1 | 100 | 0 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs2074900 | 15857820 | G > A | Exon 9, Synonymous, His343His | 0 | 100 | 0.288 | 0.325 | *NCBI B36* |
| *CYP4F2* | rs3093160 | 15857907 | G > A | Intron 8 | 1 | 100 | 0.177 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs2074901 | 15858422 | A > C | Intron 7 | 1 | 100 | 0.177 | 0.100 | *HapMap R23* |
| *CYP4F2* | rs1558139 | 15858564 | A > G | Intron 7 | 0 | 100 | 0.341 | 0.433 | *HapMap R23* |
| *CYP4F2* | rs3093158[b] | 15861166 | A > G | Intron 8 | 0 | 46.9 | 0.267 | 0.286 | *NCBI B36* |
| *CYP4F2* | rs3093156 | 15861609 | A:T | Intron 6 | 0.868 | 100 | 0.441 | 0.433 | *HapMap R23* |
| *CYP4F2* | rs3093150 | 15862629 | C > T | Intron 5 | 1 | 100 | 0.042 | 0.045 | *NCBI B36* |
| *CYP4F2* | rs3093145 | 15862852 | C:A | Intron 5 | 0.868 | 100 | 0.441 | 0.441 | *NCBI B36* |
| *CYP4F2* | rs3093144 | 15863297 | G > A | Intron 5 | 0.882 | 100 | 0.22 | 0.224 | *NCBI B36* |
| *CYP4F2* | rs3093141[a] | 15863521 | C > T | Intron 5 | 1 | 100 | 0.005 | 0.000 | *HapMap R23* |

**Appendix 1.2. Details and allele frequencies of 80 *CYP4F2* SNPs genotyped in 311 warfarin patients continued.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| *CYP4F2* | rs3093135 | 15865371 | T > A | Intron 3 | 1 | 100 | 0.178 | 0.118 | *HapMap R23* |
| *CYP4F2* | rs3093134 | 15865396 | G > C | Intron 3 | 0.885 | 100 | 0.182 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs2006193 | 15865800 | C > A | Intron 3 | 0 | 100 | 0.162 | 0.142 | *NCBI B36* |
| *CYP4F2* | rs2365178 | 15866003 | G > A | Intron 3 | 0.885 | 100 | 0.182 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs3093129 | 15866019 | A > G | Intron 3 | 0.83 | 100 | 0.18 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs3093128 | 15866104 | G > A | Intron 3 | 0.83 | 100 | 0.18 | 0.142 | *NCBI B36* |
| *CYP4F2* | rs3093126[a] | 15866160 | A > G | Intron 3 | 1 | 100 | 0.002 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs2016503 | 15866185 | T > C | Intron 3 | 0.235 | 100 | 0.08 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093124 | 15866232 | T > C | Intron 3 | 1 | 99 | 0.18 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093122[b] | 15866276 | C > G | Intron 3 | 1 | 100 | 0.178 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093121 | 15866277 | C > A | Intron 3 | 1 | 100 | 0.178 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs736089[b] | 15866300 | A > G | Intron 3 | 0.209 | 53.7 | 0.027 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs2215092 | 15866521 | G > T | Intron 3 | 1 | 100 | 0.042 | 0.043 | *NCBI B36* |
| *CYP4F2* | rs3093116 | 15866574 | G > A | Intron 3 | 1 | 100 | 0.177 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093115 | 15866615 | T > A | Intron 3 | 1 | 100 | 0.178 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs984692 | 15867074 | T > A | Intron 3 | 0.832 | 100 | 0.227 | 0.208 | *NCBI B36* |
| *CYP4F2* | rs3093114 | 15867413 | C > T | Exon 3, Synonymous, Ala82Ala | 1 | 100 | 0.178 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs3093113[a] | 15867545 | A > G | Intron 2 | 1 | 100 | 0.002 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs3093112 | 15867611 | A > G | Intron 2 | 1 | 100 | 0.178 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs3093110 | 15868784 | T > C | Intron 2 | 0.83 | 100 | 0.18 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs2074902[c] | 15869099 | C > T | Intron 2 | 0 | 100 | 0.486 | 0.180 | *HapMap R23* |
| *CYP4F2* | rs3093106 | 15869257 | A > G | Exon 2, Synonymous, Pro55Pro | 0.941 | 100 | 0.183 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093105[c] | 15869388 | G > T | Exon 2, Missense, Gly12Trp | 0 | 0 | 0 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093103 | 15869434 | T > C | Intron 1 | 1 | 99.7 | 0.177 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093100 | 15869469 | G > C | Intron 1 | 0.83 | 100 | 0.18 | 0.139 | *NCBI B36* |

**Appendix 1.2. Details and allele frequencies of 80 *CYP4F2* SNPs genotyped in 311 warfarin patients continued.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE P-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| *CYP4F2* | rs3093098 | 15869512 | T > C | Intron 1 | 0.941 | 100 | 0.183 | 0.128 | *HapMap R23* |
| *CYP4F2* | rs3093097 | 15869690 | CA > DEL | Intron 1 | 0.872 | 100 | 0.359 | NA | - |
| *CYP4F2* | rs3093092[c] | 15870127 | C > A | 3` near gene | 0 | 95.2 | 0.235 | 0.142 | *NCBI B36* |
| *CYP4F2* | rs3093091[a] | 15870133 | C > T | 3` near gene | 0 | 99 | 0.013 | 0.222 | *NCBI B36* |
| *CYP4F2* | rs3761014 | 15872763 | C > G | 3` near gene | 1 | 100 | 0.23 | 0.207 | *HapMap R23* |

Frequency data were compiled from HapMap and NCBI dbSNP databases.

Chromosomal positions are given in base pairs from the p-telomere of chromosome 19, as per HapMap Data release 23, March 2008, NCBI B36 assembly, dbSNP b126.

[a] SNPs were monomorphic.

[b] SNPs gave <90% call rate.

[c] SNPs were not in HWE.

**Appendix 1.3. Univariate tests for association between SNPs in the *CYP4F2* gene and clinical outcomes of response to warfarin.**

| INR > 4 | | | Major Bleeds | | | Time to Therapeutic INR | | | Time to Stable Dose | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SNP | n | *P*-value | SNP | n | *P*-value | SNP | n | *P*-value | SNP | n | *P*-value |
| rs12610189 | 311 | 0.024 | rs12610189 | 311 | 0.015 | rs12610189 | 310 | 0.004 | rs12984060 | 310 | 0.007 |
| rs2108622 | 311 | 0.022 | rs2074900 | 311 | 0.008 | rs12984060 | 310 | 0.036 | rs2189784 | 310 | 0.016 |
| rs3093193 | 311 | 0.032 | rs2108622 | 311 | 0.015 | rs2079288 | 310 | 0.004 | rs3093193 | 310 | 0.043 |
| rs3093194 | 311 | 0.042 | rs3093097 | 311 | 0.015 | rs2108622 | 310 | 0.003 | rs3093206 | 310 | 0.037 |
| rs3093195 | 311 | 0.037 | rs3093166 | 311 | 0.033 | rs2189784[a] | 310 | 0.00028 | rs3093207 | 310 | 0.043 |
| rs3093203 | 307 | 0.036 | rs3093168 | 311 | 0.043 | rs3093193 | 310 | 0.003 | rs3093209 | 310 | 0.043 |
| rs3093206 | 311 | 0.023 | rs3093193 | 311 | 0.015 | rs3093206 | 310 | 0.005 | rs3093216 | 310 | 0.043 |
| rs3093207 | 311 | 0.032 | rs3093194 | 311 | 0.025 | rs3093207 | 310 | 0.003 | rs7252046 | 310 | 0.032 |
| rs3093209 | 311 | 0.032 | rs3093195 | 311 | 0.025 | rs3093209 | 310 | 0.003 | | | |
| rs3093216 | 311 | 0.032 | rs3093203 | 307 | 0.026 | rs3093216 | 310 | 0.003 | | | |
| | | | rs3093206 | 311 | 0.005 | rs7252046 | 310 | 0.027 | | | |
| | | | rs3093207 | 311 | 0.015 | | | | | | |
| | | | rs3093209 | 311 | 0.015 | | | | | | |
| | | | rs3093216 | 311 | 0.015 | | | | | | |

| Stable Warfarin Dose | | | Sensitive | | | Resistant | | |
|---|---|---|---|---|---|---|---|---|
| SNP | n | *P*-value | SNP | n | *P*-value | SNP | n | *P*-value |
| rs2079288 | 204 | 0.01 | rs17756654 | 273 | 0.025 | rs2189784 | 273 | 0.019 |
| rs2189784 | 204 | 0.017 | rs2079288 | 273 | 0.017 | | | |

Only those with a *P* value <0.05 are shown here.

[a]Association of rs2189784 with time to target INR remained significant after FDR was calculated ($P_c$ = 0.03).

# Appendix 2

# CONTENTS

**APPENDIX 2.1. *CYP4F11* AND *CYP4F12* PCR AND EXTENSION PRIMERS FOR MALDI-TOF MS. ..... 272**

**APPENDIX 2.2. DETAILS AND ALLELE FREQUENCIES OF 210 SNPS ACROSS *CYP4F2*, *CYP4F11* AND *CYP4F12* GENOTYPED IN 149 HUMAN LIVER SAMPLES..... 277**

**APPENDIX 2.3: TAGGING SNPS ACROSS THE *CYP4F12-CYP4F2-CYP4F11* LOCUS ($R^2 \geq$ 0.9)..... 285**

**Appendix 2.1. *CYP4F11* and *CYP4F12* PCR and extension primers for MALDI-TOF MS.**

| SNP | Assay | Chromosomal position | Forward Amplification Primer Sequence (5'>3') | Reverse Amplification Primer Sequence (5'>3') | Amplification Product (bp) | Amplification Direction | Extension Primer Sequence (5'>3') |
|---|---|---|---|---|---|---|---|
| rs4808362 | 1 | 15664892 | ACGTTGGATGACAGAGACTCCATTCCCT | ACGTTGGATGTAGGAGCACACCATCTACTG | 91 | F | CTCCATTCCCTCCCCCC |
| rs9305063 | 1 | 15646525 | ACGTTGGATGCTCATTCACTGATTCCTCT | ACGTTGGATGAAAGGTTTGGCATGGAACTG | 100 | F | CTGATTCCTCTACCCAC |
| rs10420354 | 1 | 15657946 | ACGTTGGATGAGGATGGAGTTGAGAGCAAG | ACGTTGGATGCACCAGATCTGGCTTGAAAC | 99 | F | GAGCAAGAGAAGCATGG |
| rs11666521 | 1 | 15722672 | ACGTTGGATGTAAAAACAGAAAGCCTCAG | ACGTTGGATGATTGGTAAGAATTTCTTCC | 98 | F | GAAAGCCTCAGACCCAGA |
| rs688755 | 1 | 15668305 | ACGTTGGATGAGAGAACAGCAAGGGGAGGT | ACGTTGGATGTCAGACACAGGGCGCTCTTA | 98 | F | CACCTCTGGCTTTTATTCC |
| rs688468 | 1 | 15651752 | ACGTTGGATGGCAGACAGGGTGTCCGAGG | ACGTTGGATGCTGCAACAACTGAACACAGC | 88 | F | AGGGTGTCCGAGGTCCTTT |
| rs616745 | 1 | 15656420 | ACGTTGGATGCTCTCTTGTTCACCATCACC | ACGTTGGATGTGTCACCAGGGTCCAAAGAG | 101 | F | aacAGGGCTTCCCCCAGATC |
| rs609636 | 1 | 15650098 | ACGTTGGATGACTTGCTTTCAGATCACTCC | ACGTTGGATGTAAAGCCCTGGGAATAGGTG | 99 | R | ctcCAGAGGAGGGCTTGAAG |
| rs11879787 | 1 | 15662454 | ACGTTGGATGTGTCCTTTAGGTTCTTGCAG | ACGTTGGATGGCTAGTCAGCCTTAAAACAG | 86 | F | cctcTCTTGCAGATTGCCACA |
| rs10410357 | 1 | 15661331 | ACGTTGGATGATCCTTATCGTTTTCCACAG | ACGTTGGATGGTGGATGTATTGAAAAGTG | 95 | R | TATCGTTTTCCACAGAGGTTA |
| rs10409750 | 1 | 15652606 | ACGTTGGATGTCCCTGTAAACTCAAGAGAG | ACGTTGGATGGCCAATGAGTTCTGAAGTG | 100 | F | AATGAAGATTGTGTAGTAGTC |
| rs3826950 | 1 | 15906891 | ACGTTGGATGAGGTGGCTGTAAACTTTGGG | ACGTTGGATGAACATCTCACACCTGAGTGG | 98 | F | gaAGAGTCACCTTTGACTCTTC |
| rs12459933 | 1 | 15884116 | ACGTTGGATGGTTCCATTGACCTATTTCTC | ACGTTGGATGGACTTACCGTGAAGCTACAG | 93 | R | cTCTCTTTTCTTTTACCAATACC |
| rs3810428 | 1 | 15907478 | ACGTTGGATGCATTTTCCTGTGATCCCTCC | ACGTTGGATGTGGACTCAGGCTTCCCATCT | 100 | F | cctaTAAGGCCTTATCTGACATC |
| rs11085969 | 1 | 15653546 | ACGTTGGATGTGAGAAATGGCTCAGGTAAG | ACGTTGGATGACTTCTTCATCAGCCCTGAG | 101 | F | cATGGCTCAGGTAAGAACATCAC |
| rs3813137 | 1 | 15643397 | ACGTTGGATGGTGCCAGATGTTCCACATAA | ACGTTGGATGGTCAAGTATGTTGCACCGAG | 104 | R | agtgCCACATAAAATGCTTCTCCC |
| rs7249167 | 1 | 15898359 | ACGTTGGATGGTTTCAAAGTTCAACCAGGC | ACGTTGGATGTGGAGACTGTTTTATGACCC | 81 | R | gtccAAGTTCAACCAGGCTTAGGT |
| rs2116951 | 1 | 15908806 | ACGTTGGATGTCCTCCCACACCCCGGATG | ACGTTGGATGGGCAATTCTGGACAGGATAC | 100 | F | aacccCACCCCGGATGCTCCACCCA |
| rs2079234 | 1 | 15718510 | ACGTTGGATGCCCCATACCTTACTGCCATA | ACGTTGGATGCTTGATACCCAGACATGAAG | 100 | R | acGGCTTATTTAAAACAGTTCTTTT |
| rs594255 | 1 | 15668984 | ACGTTGGATGTGAAGCCTCCCAGTCCCCAA | ACGTTGGATGTAAAACGGTGCTGTCACCTC | 94 | R | gcAGCCTCCCAGTCCCCAAGAGCCCC |
| rs615392 | 1 | 15663595 | ACGTTGGATGGGGGTGCTTTTGGTTAATTT | ACGTTGGATGGGAAATAGCTGGGTAGTGTG | 96 | F | TTAATTTTTATGAAGTTACATTTGTT |
| rs675326 | 1 | 15651144 | ACGTTGGATGGCTGAATCCAGGTGTCAATG | ACGTTGGATGAGTGATGCAGAGAAATGGTG | 95 | F | ggagaGTGTCAATGAACATTATCAGG |
| rs17641489 | 1 | 15920483 | ACGTTGGATGCATGTTTTCTAGAACTGGCAC | ACGTTGGATGACTTGGAGCCATTAGAGACC | 96 | F | ccccTTTCTAGAACTGGCACATTATCA |
| rs2176917 | 2 | 15644236 | ACGTTGGATGTCAGCTCAAGGCCCCTCTCA | ACGTTGGATGCAAACCTAGAAGCTGGGCTC | 98 | R | AGCCTCATCCCCTGTGC |
| rs11670533 | 2 | 15911160 | ACGTTGGATGGTACCACTTGAACCCTCTTG | ACGTTGGATGCCATCATAGCTGGGAACTCA | 91 | F | TGAACATGGAAACCCAA |
| rs672645 | 2 | 15653127 | ACGTTGGATGATCCTGACCCCCAATTCTTC | ACGTTGGATGTCTCTTAATGGGAGTTGCTG | 99 | F | TCACTTCTCCCTGTATTC |
| rs12976669 | 2 | 15671312 | ACGTTGGATGAAAACCATCCCCTACCTTCG | ACGTTGGATGCAGTCTCCATCTTTTTTGGC | 99 | R | cCGTCCATGGGAAAATCG |
| rs611848 | 2 | 15669379 | ACGTTGGATGGCTCTCAGGAGTTCTGTTTC | ACGTTGGATGGGGTCACAAAAGCGATCAAC | 99 | F | CTGTTTCCTTTCTTGCTAA |
| rs10423986 | 2 | 15727655 | ACGTTGGATGAGTTTGTATGCATAGAGGAG | ACGTTGGATGATGGAGGCATTGCAACTGAC | 97 | F | ATGCATAGAGGAGTTCATA |
| rs12460703 | 2 | 15650776 | ACGTTGGATGACATCCACCTATGATGGGTC | ACGTTGGATGGAGGAGAGCAGGAAACTTAG | 94 | F | ccaATGGGTCAGTTCCTTCT |
| rs609290 | 2 | 15650140 | ACGTTGGATGTGTCAGGGTGGCATAAAACG | ACGTTGGATGAGATGTCGGCCACCTATTCC | 101 | F | GATGATGGGACCCAGCCATA |

**Appendix 2.1. *CYP4F11* and *CYP4F12* PCR and extension primers for MALDI-TOF MS continued.**

| SNP | Assay | Chromosomal position | Forward Amplification Primer Sequence (5'>3') | Reverse Amplification Primer Sequence (5'>3') | Amplification Product (bp) | Amplification Direction | Extension Primer Sequence (5'>3') |
|---|---|---|---|---|---|---|---|
| rs11086012 | 2 | 15893643 | ACGTTGGATGGATGGTTTGCAGGCTTTTGG | ACGTTGGATGAAGTGGCATCATCTGCCTTC | 95 | R | gaTGGGCTGCTGGGAGAATT |
| rs2285890 | 2 | 15654415 | ACGTTGGATGGGAGGAATGAGCAAAGTAAC | ACGTTGGATGTTCCTCGACCTTCAAAGCAC | 95 | F | AAAGTAACCAGAAGTACCTTT |
| rs1060467 | 2 | 15885538 | ACGTTGGATGTTTCTGGGACTCTACAGAGG | ACGTTGGATGCCCTGGGTGCGAACTCACA | 83 | R | ctctGTGGGTGGGTGGGTAGG |
| rs624512 | 2 | 15669890 | ACGTTGGATGTGTGGAGAACTCAGAACTTG | ACGTTGGATGGAGGAGGTAATAATGGCCAC | 91 | F | caAACTCAGAACTTGTTGCCTG |
| rs688256 | 2 | 15668175 | ACGTTGGATGGGATTTCCTATGTCCAGCAG | ACGTTGGATGGGAAAATCCAACATCACCTC | 92 | R | gtcaAGCAGCCTTGGAGAGACA |
| rs2018460 | 2 | 15873757 | ACGTTGGATGGTAAGATTCACCTAACTACAC | ACGTTGGATGAACATGTTTATGCCATAAG | 87 | R | CACTTTTCATATTGTCCTTCAAT |
| rs8102331 | 2 | 15883144 | ACGTTGGATGCTCTTCCATTAATGGGTCCT | ACGTTGGATGGGGCAAAAAAGCTATGATCC | 99 | F | gggtgATGGGTCCTTGCTGACCT |
| rs631279 | 2 | 15659334 | ACGTTGGATGTGGAAGGACAGAATAGTGTG | ACGTTGGATGTTTCCTGTTAAGTCCCTGTG | 100 | R | gaggGTGTAGAGATTCACAGTGT |
| rs10410157 | 2 | 15646622 | ACGTTGGATGCCACCTCAGCCTGCTTGGA | ACGTTGGATGACAACCATACAGGAGCCTCT | 99 | F | aacaTCAGCCTGCTTGGATCCTTT |
| rs7248804 | 2 | 15644051 | ACGTTGGATGTCCCTCTTTTTGGAGCTCCC | ACGTTGGATGGGGAAATGGGATGTCAGATG | 98 | F | ttcgTCTTAGGAGCCTTCCTCCGTC |
| rs651109 | 2 | 15665158 | ACGTTGGATGGCCCCACTGGAAAACGTTTA | ACGTTGGATGGCCACTGCAAAAGAAGAAAG | 104 | F | AAAACGTTTATAAACATTTAAGATC |
| rs3810427 | 2 | 15907650 | ACGTTGGATGGCACATGGTAGGTGTGCTTT | ACGTTGGATGCTCTGCTCTTTGCAAAAGAC | 99 | F | aggtcGAGTCCAGTACACAGTAAAA |
| rs16980800 | 2 | 15706047 | ACGTTGGATGGTGAGCCACATTGCTACAAG | ACGTTGGATGGGATGTAATGATGAATGGGC | 88 | R | cccccGCTACAAGAGTTTCTGAAATC |
| rs7255335 | 2 | 15672079 | ACGTTGGATGTTGAGATGGAATTTCACTT | ACGTTGGATGGAGGCGAAGATTTCAGTGAG | 96 | R | gggGAGATGGAATTTCACTTTTCTAG |
| rs642322 | 2 | 15659173 | ACGTTGGATGGTTCTTATAAAGTGGGCATC | ACGTTGGATGTCACCCTAGTAGTTTCTACC | 100 | F | gggtgATAAAGTGGGCATCAAGATAG |
| rs2305800 | 3 | 15906294 | ACGTTGGATGATCTCAGAGCCTTCCCTGAT | ACGTTGGATGCCCAGGAAGCTCCAAGGAC | 96 | F | TGCCCATCCTTGCCCCT |
| rs16995376 | 3 | 15645377 | ACGTTGGATGTGCAGGATGTCGCTGCTGAG | ACGTTGGATGACAACCAGCAGCAGGAGTAG | 103 | F | CTGGCTGGGCCTCAGAC |
| rs688231 | 3 | 15668157 | ACGTTGGATGGGCAGGGTTTTGATGAGGAA | ACGTTGGATGTATGTCCAGCAGCCTTGGAG | 100 | F | AAAATCCAACATCACCTC |
| rs12151032 | 3 | 15653566 | ACGTTGGATGGGCTCAGGTAAGAACATCAC | ACGTTGGATGAGAGGCACTTCTTCATCAGC | 99 | F | gCTTGCATTGGTGGCATA |
| rs10854148 | 3 | 15643025 | ACGTTGGATGGCCTCAGTTTGCTTTTGTG | ACGTTGGATGAACCCTGGGCATTGTTGAAG | 103 | F | CTCCTCTGGATGCCATCAC |
| rs2116952 | 3 | 15913255 | ACGTTGGATGAGTATGTGAAAACTCCTGTC | ACGTTGGATGCATTTGTTTATCCCCTTTC | 82 | R | GAAAACTCCTGTCAGAAAA |
| rs7245534 | 3 | 15643098 | ACGTTGGATGCCTGCTTGTATCAGATGCTC | ACGTTGGATGTTCTGGCCTAGGGAGCTGTT | 100 | R | ctcTGCTCACCTGCATCACA |
| rs593818 | 3 | 15668884 | ACGTTGGATGAAAAAACAGGTGGATGGGTC | ACGTTGGATGGGCTTTGGCTGCGGGTGGA | 90 | R | GTCAGAAAGTCACTGCAAGC |
| rs628603 | 3 | 15655595 | ACGTTGGATGCCTGACATTTGACATGAAGC | ACGTTGGATGGCTTCTGCTGAGCAAGGTAG | 96 | R | AGCTCTATTCTCTCACCTAAA |
| rs2139749 | 3 | 15644378 | ACGTTGGATGGGCTGGATTCTTTCTTCCTG | ACGTTGGATGAGTTCAAGTCCTGCCTCCTG | 98 | R | aaaTCAAGCAGTCAGGCCTCA |
| rs659825 | 3 | 15657584 | ACGTTGGATGATGAATTTTCCCTTCACTCC | ACGTTGGATGCAGGAAGTAAAGATGTTTG | 92 | R | gCCCTTCACTCCTTTCTTGTTT |
| rs677139 | 3 | 15651547 | ACGTTGGATGCGTGATTAGAATGGATTCCG | ACGTTGGATGCTAGGGTCTCTTTGGAATTG | 96 | R | ggattGATTCCGAAACCTGGCT |
| rs610652 | 3 | 15653900 | ACGTTGGATGTATTGCCCTCCTCTAGGT | ACGTTGGATGTGGTTTATCTGGCCCCAAAG | 97 | R | CCACCTGTCAAATATTCAATAAC |
| rs7253051 | 3 | 15888131 | ACGTTGGATGCTCTAAACTTTTTTGGGAAG | ACGTTGGATGGCTTTTCTCTAATATTTTAT | 93 | F | AGAAAAACATCATAAAGAGCTAG |
| rs627971 | 3 | 15664094 | ACGTTGGATGTTAATATCTGGCAGAGTGGC | ACGTTGGATGAACATACACTTTATCACAC | 97 | F | aTTTTTTTGTTGTTAAATCAACCC |
| rs659447 | 3 | 15657662 | ACGTTGGATGCTCTTCTGTTCCCTAATTCC | ACGTTGGATGCAGGTCTCCCTTAGGACTC | 80 | R | gTCTGTTCCCTAATTCCTACCCTCC |

**Appendix 2.1. *CYP4F11* and *CYP4F12* PCR and extension primers for MALDI-TOF MS continued.**

| SNP | Assay | Chromosomal position | Forward Amplification Primer Sequence (5'>3') | Reverse Amplification Primer Sequence (5'>3') | Amplification Product (bp) | Amplification Direction | Extension Primer Sequence (5'>3') |
|---|---|---|---|---|---|---|---|
| rs12610962 | 3 | 15884377 | ACGTTGGATGCGATGCAACATCCCTTAGGC | ACGTTGGATGGAAACCTCAGTCAGTGTCTC | 99 | F | cccCAACATCCCTTAGGCTATGAAA |
| rs629748 | 3 | 15655843 | ACGTTGGATGGGTGCATCTTTTATCATAG | ACGTTGGATGGTGTGTTAGTCTTTCACAT | 102 | F | cctcGATACAGACATAGACAAAGAC |
| rs2886476 | 3 | 15672717 | ACGTTGGATGAATGATGGTGAAAGAACTG | ACGTTGGATGGAATATGGTAAGAGATAAGC | 89 | R | tATGGTGAAAGAACTGAATGTGTAC |
| rs680681 | 3 | 15663148 | ACGTTGGATGTGTGTGTACAGTGAGCATTG | ACGTTGGATGCTTTTGGGAATTTGTCCTGC | 81 | R | ccctgGAGCATTGCTCATATGTATAA |
| rs12327750 | 3 | 15916687 | ACGTTGGATGTCCTCCCCAGTGAGAGACAT | ACGTTGGATGTGGGAGTGATCTTGAGGCTG | 99 | R | ggaAGTGAGAGACATGCTGGCCCCGA |
| rs59302543 | 3 | 15667846 | ACGTTGGATGTGAGGTTACATCCCCCAGC | ACGTTGGATGATGACTCGGCCATCTGGGA | 96 | F | cTCATCTCCCGATGCTGCACCCAGGAC |
| rs17682485 | 4 | 15650698 | ACGTTGGATGAATAGCAGCCCAGATTCTCC | ACGTTGGATGGCACCAGCCTTGCCCATC | 94 | R | GCCACAAATCCTCCTAC |
| rs10414690 | 4 | 15645195 | ACGTTGGATGTTGTTGGTTTTGGTTGGGAC | ACGTTGGATGCTCAGGGGTGAGTAAACGAG | 83 | F | TTGGTTGGGACTTTCTG |
| rs595381 | 4 | 15652809 | ACGTTGGATGTGCTGATCCTGGCTCAGTTC | ACGTTGGATGTGTCACAGCTGAGCCCATC | 98 | F | CTTCAGTACCTTCAGACT |
| rs7252197 | 4 | 15696098 | ACGTTGGATGGTAGGTGTGTATACTTATGG | ACGTTGGATGCCCTGATGTAATTATTACTC | 93 | F | GGGGTACATGAGCTGTTT |
| rs12971888 | 4 | 15887204 | ACGTTGGATGAAGCAAACAATTCAGCTCCC | ACGTTGGATGAAAACAGAACACTTCACTGG | 91 | R | TCCCTTACTGGGTCTTAAA |
| rs2305804 | 4 | 15899365 | ACGTTGGATGGAAGGACTTTGGGTATGGAG | ACGTTGGATGTCTTCCTGCTATGCTGGGTG | 98 | R | TTGGGTATGGAGGAATGCT |
| rs687774 | 4 | 15668062 | ACGTTGGATGCCACTAGGAACCTTGGAATT | ACGTTGGATGTGCTGCCTTCCCCATTCAC | 100 | R | gAAAGTAGAGGATGGGGGT |
| rs12978309 | 4 | 15892188 | ACGTTGGATGCTCTGGCAGTTTCCATCATC | ACGTTGGATGTGTGTTCATATACAGTGGGC | 99 | R | TTTCTGTGTGTCCTAGAAAA |
| rs16980968 | 4 | 15907925 | ACGTTGGATGGTAGGTGCACTTGACATAG | ACGTTGGATGCAGATGTTCTATGTAGAGC | 81 | R | cttcCATAGATCACCAGGCAC |
| rs590828 | 4 | 15674919 | ACGTTGGATGAGCATAAATTGTGGCAGCCC | ACGTTGGATGAGTCTCAGAGAGCCAGACTA | 100 | F | agaaCAAGAGGAGCCGCAGCC |
| rs57578760 | 4 | 15655463 | ACGTTGGATGTCAATACCCTGAGTGGGGAG | ACGTTGGATGGCTTCCACAGGGCCTGCCG | 108 | R | cccaGACGCCGCTCCCGGATGA |
| rs2305801 | 4 | 15906141 | ACGTTGGATGATCCCCGTGGCTGCTTCTG | ACGTTGGATGCATAGAAGGTGTAGGTCCAG | 100 | F | gtCTGCTTCTGCTGCTGGTTGG |
| rs3765070 | 4 | 15901292 | ACGTTGGATGACTACCACAAGCTCTGCACG | ACGTTGGATGCCTCCTCATTTTATGCCACC | 98 | R | gGGCACTGGTGATAGGCCGGAT |
| rs629651 | 4 | 15664417 | ACGTTGGATGGAACATTATCTTGTCATAAG | ACGTTGGATGGGCAACTGGAAAGCCTTCTG | 91 | R | ATCTTGTCATAAGATATGAGAAC |
| rs4808409 | 4 | 15894379 | ACGTTGGATGTGAAGAAGCCAGCGGGAAG | ACGTTGGATGCCCATCATGCACTCCTAGAT | 98 | R | ggtgGGGCTCTGATGGTGGTGAC |
| rs673020 | 4 | 15653081 | ACGTTGGATGTTTCAGCAACTCCCATTAAG | ACGTTGGATGCAAAAGTCCAGTACAGACTC | 86 | F | ccatCAACTCCCATTAAGAGATGC |
| rs630133 | 4 | 15655769 | ACGTTGGATGCACAGGGCACTGCTAATTCT | ACGTTGGATGGTGAAAGACTAACACACTAAC | 99 | R | cggcgGCTAATTCTGAAACTGTGA |
| rs1471112 | 4 | 15891458 | ACGTTGGATGGAGAACAGATTCTTCTTGGAC | ACGTTGGATGGGAATGTCAAGGTTAAACCC | 86 | R | ccctAGATTCTTCTTGGACTCAATG |
| rs4433928 | 4 | 15699690 | ACGTTGGATGCCTGCACAAAGCACTTAAGG | ACGTTGGATGTACTCCATCTGCAGAAAGGC | 99 | R | TTAAGGTGAGTGATCTTATTTAAAC |
| rs4347744 | 4 | 15643859 | ACGTTGGATGAAGTGCACTTACTGTGTATC | ACGTTGGATGTGTGGCATACCTGGCACATA | 86 | F | cccgaACTTACTGTGTATCAGATACT |
| rs17682497 | 4 | 15657542 | ACGTTGGATGAACAAGAAAGGAGTGAAGGG | ACGTTGGATGATATGAAGGCAGACGATGAC | 88 | R | caatGGAAAATTCATTGACTGAACAT |
| rs2074569 | 4 | 15655727 | ACGTTGGATGTGCTTGAGAAAGGGAAGGTC | ACGTTGGATGAGAATTAGCAGTGCCCTGTG | 98 | F | ggcGATAGGTTTTAGAGATGACTTTA |
| rs12985091 | 5 | 15906749 | ACGTTGGATGTGGTCGAGCCAATCCCCAT | ACGTTGGATGCCTACGCCAGGTTTTGGTTG | 90 | F | CCAATCCCCATGGACTG |
| rs16995378 | 5 | 15645386 | ACGTTGGATGACAACCAGCAGCAGGAGTAG | ACGTTGGATGTGCAGGATGTCGCTGCTGAG | 103 | R | GGAGTAGCCATGGGGAC |
| rs11085971 | 5 | 15657813 | ACGTTGGATGTCCATAGGCAGGGTTCTTTC | ACGTTGGATGACATGCCTTTCCATTGCCAG | 96 | R | CTTTCCAGGTCTCCAGAG |

**Appendix 2.1. *CYP4F11* and *CYP4F12* PCR and extension primers for MALDI-TOF MS continued.**

| SNP | Assay | Chromosomal position | Forward Amplification Primer Sequence (5'>3') | Reverse Amplification Primer Sequence (5'>3') | Amplification Product (bp) | Amplification Direction | Extension Primer Sequence (5'>3') |
|---|---|---|---|---|---|---|---|
| rs12977005 | 5 | 15643668 | ACGTTGGATGCTTGATGCAGAGTAGGTGCG | ACGTTGGATGAGGTGCTCCACATTGAGCAT | 91 | F | CTGTGGGTCACCAGGTAC |
| rs8104361 | 5 | 15895714 | ACGTTGGATGTGAGGAAATCATCAATACCC | ACGTTGGATGTGCACGACTTCACAGATGCC | 95 | F | CCTGAGTGGGGAGGGTGC |
| rs3765071 | 5 | 15901034 | ACGTTGGATGGAGGACAGGGCCCTAAATTG | ACGTTGGATGTCTCTGCCTTTGTGTTCTGG | 85 | R | GGGCCCTAAATTGAGGCTT |
| rs4375789 | 5 | 15912804 | ACGTTGGATGCCTACCCAGAAACCATACTC | ACGTTGGATGATAGAAGAGAGGTTGCCCAG | 97 | R | cccTACTCCAGCGTCAGCCA |
| rs12977516 | 5 | 15886919 | ACGTTGGATGTAGCTGGCTTCAGAGTGTTG | ACGTTGGATGACCAGAAAAAAATGGTGCCG | 89 | R | TCAGAGTGTTGTGTTTTCCT |
| rs2240228 | 5 | 15713872 | ACGTTGGATGCCTGATAGGCTGTTTGTTCC | ACGTTGGATGCTCAGCAGAAGGAATCCTCA | 99 | F | tTCCTCTCCTTTGTCTTCATT |
| rs670476 | 5 | 15674467 | ACGTTGGATGGGTGGTAGCAGTGGAATTTG | ACGTTGGATGAGACGCGAGAGTACCTCCT | 80 | R | CAGTGGAATTTGTGTCTGTCT |
| rs2365177 | 5 | 15880408 | ACGTTGGATGTGGGGACTTCCAGTTTTTC | ACGTTGGATGTGACTTGAAATTCATTCCC | 101 | F | ACTTCCAGTTTTTCTTCACTTT |
| rs12460831 | 5 | 15908758 | ACGTTGGATGTGTGGGAGGAGGAGCCAGGT | ACGTTGGATGTCTCAGGCAATGTCATGGTC | 95 | R | tttttAGCCAGGTGCTGATTCA |
| rs1064796 | 5 | 15885662 | ACGTTGGATGTTTCCTGCGGGGTTCAGTGT | ACGTTGGATGTGGCTGAGATGAAGGTGGTC | 100 | R | tcatGATGCGGAAGTGCAGCAG |
| rs1060463 | 5 | 15886176 | ACGTTGGATGGTGACCTCTCCTTGATGTTC | ACGTTGGATGAGCCTTCTTGGTCTCACCTG | 92 | R | cccaTCCTTGATGTTCTCTTGGT |
| rs11879253 | 5 | 15906491 | ACGTTGGATGAAAAGGCCCAGCCTGAATC | ACGTTGGATGGCCTGGGAGACTGGTGATTT | 91 | F | gttggATCACTGGGAGGTTGCCA |
| rs4109350 | 5 | 15661921 | ACGTTGGATGTAATGTGCAGATCTGTCTC | ACGTTGGATGCATCCATGTAACCAAACCAC | 99 | F | cgggGCAGATCTGTCTCTTCTTTA |
| rs10500211 | 5 | 15919348 | ACGTTGGATGAAGTGCAATCTGGTTATTT | ACGTTGGATGCCTTTAGCGTGTGTGATTTC | 104 | R | gggtCTGTGAAGAAAAATGTTGTT |
| rs6512074 | 5 | 15887894 | ACGTTGGATGCACCATAGCAACAAAATATC | ACGTTGGATGAAGCTTTTAATAGAGCTTAA | 107 | F | ccACAAAACATGCATATGATATCTA |
| rs16980704 | 5 | 15651531 | ACGTTGGATGTGGGAGCTAGGGTCTCTTTG | ACGTTGGATGAGAATGGATTCCGAAACCTG | 95 | F | tttcgCTTTGGAATTGTCAGAGTTA |
| rs672555 | 5 | 15653185 | ACGTTGGATGTGCATGAGACAGTAACAAGC | ACGTTGGATGGGTCAGGATTTTGTGAAGTG | 97 | F | caaaTTGTCAGTTTGCATATTTAAAA |
| rs16980695 | 5 | 15643275 | ACGTTGGATGGTGTGCCTTATTGTGTGCCA | ACGTTGGATGAGTAGGTACACTTTATGCGG | 100 | R | ggtgATTGTGTGCCAGATATGCTACA |
| rs613503 | 5 | 15663113 | ACGTTGGATGCATATGAGCAATGCTCACTG | ACGTTGGATGGGCTTGTTTCATCCTTTTGC | 92 | R | cccgcGCAATGCTCACTGTACACACAT |
| rs640169 | 6 | 15655885 | ACGTTGGATGATAATTAATGGGCTGCTGTG | ACGTTGGATGGTCTATGTCTGTATCTATG | 85 | R | CTGCTGTGATCCAGGGG |
| rs61391486 | 6 | 15668720 | ACGTTGGATGACAGTCACAGTCCCCACTC | ACGTTGGATGCAGGACCACTTTCATCTCCG | 98 | F | CCACTCCCGCCTGCAGGA |
| rs2074568 | 6 | 15652132 | ACGTTGGATGGGTGAGTACCTGCAGGTGA | ACGTTGGATGAAGAGCAGCACTTCCCCTAC | 88 | F | ACCTGCAGGTGAAAGGGG |
| rs3746152 | 6 | 15895519 | ACGTTGGATGTCTCAAACCTATCTCTGACC | ACGTTGGATGTTGGACTTGATCCAAAGGGC | 93 | F | CTGTCTTTTTCAAGCATGT |
| rs627753 | 6 | 15660130 | ACGTTGGATGTGTTTTGCCCTGTGAGAGTC | ACGTTGGATGCAGAGGCAGTTATACTCACC | 98 | F | ggCTGTGAGAGTCAGAGGG |
| rs633929 | 6 | 15657441 | ACGTTGGATGTAGTACTAGGAGCAGAGCAC | ACGTTGGATGGTCTGAAAACTTTCCCCTGC | 90 | F | GAGGCAGAACTTGGTATCCA |
| rs665872 | 6 | 15664106 | ACGTTGGATGCCCTCAATCCAAATAAAAC | ACGTTGGATGTTTGTTGTTAAATCAACCC | 89 | F | AACATACACTTTATCACACAA |
| rs10415331 | 6 | 15652617 | ACGTTGGATGGGCTCTAAGCCAATGAGTTC | ACGTTGGATGTGAAGATTGTGTAGTAGTC | 82 | R | CAATGAGTTCTGAAGTGGTTT |
| rs2072269 | 6 | 15885739 | ACGTTGGATGATGGGGGTCATGGGCACAG | ACGTTGGATGTGCCCGATGCAGTTTCTGG | 80 | R | ccctcCATGGGCACAGTCAGAT |
| rs17641206 | 6 | 15645152 | ACGTTGGATGCTCTTTCCCTTTCTGTGTGG | ACGTTGGATGCAGAAAGTCCCAACCAAAAC | 92 | F | gattgCTGTGTGGTTACCAGGT |
| rs4808369 | 6 | 15691869 | ACGTTGGATGGGAGATGGAAATTGTGTGTC | ACGTTGGATGGATAGTGAACAAGCCATGGG | 94 | R | gGTGTGTCAACTTTTAATGAGTC |
| rs10405688 | 6 | 15643494 | ACGTTGGATGCACACAGCAAGCATGGTTTA | ACGTTGGATGTGTCTACTGTATGCTGGGAG | 101 | F | gaagGGTTTATATAAAGGGCCAG |

**Appendix 2.1. *CYP4F11* and *CYP4F12* PCR and extension primers for MALDI-TOF MS continued.**

| SNP | Assay | Chromosomal position | Forward Amplification Primer Sequence (5'>3') | Reverse Amplification Primer Sequence (5'>3') | Amplification Product (bp) | Amplification Direction | Extension Primer Sequence (5'>3') |
|---|---|---|---|---|---|---|---|
| rs6512075 | 6 | 15888067 | ACGTTGGATGGAGTAAGTTCATCAAGTTC | ACGTTGGATGCAAGAGAATATATTCAGTGC | 84 | R | ctcAAGTTCATCAAGTTCCTAAAA |
| rs4331436 | 6 | 15643897 | ACGTTGGATGATCTGATACACAGTAAGTGC | ACGTTGGATGGACTGTACGGAAAGCACTAC | 99 | R | acGTAAGTGCACTTTATATAGTGA |
| rs2285888 | 6 | 15654235 | ACGTTGGATGTTCCCTTCTCTGGCCAGGA | ACGTTGGATGTGAGGCTGATGTGCTCAAAC | 104 | F | gggcgGGCCTCAGAGGGCAGCAGT |
| rs2163859 | 6 | 15914143 | ACGTTGGATGCCCATTAAAAATTTTAATTTC | ACGTTGGATGGAAATACTTTTGCACTATTAG | 80 | F | AAAAATTTTAATTTCTACCATTCTAC |
| rs10854147 | 6 | 15642995 | ACGTTGGATGAGCCTCAGTTTGCTTTTGTG | ACGTTGGATGCTGGGCATTGTTGAAGAGAC | 100 | F | agccgCAGTTTGCTTTTGTGTTAAAT |
| rs610142 | 6 | 15649930 | ACGTTGGATGCTGTGAGGGTGTTTGAGAAG | ACGTTGGATGACTCACTCATTCCTCTGCAC | 97 | F | cctgaTGTTTGAGAAGAGATTAGCAT |

Chromosomal positions are given in base pairs from the p-telomere of chromosome 19, as per HapMap Data release 27, February 2009, NCBI B36 assembly, dbSNP b126.

MALDI-TOF MS analysis was performed with 6 different assays for genotype and haplotype determination.

All amplification primers have a tag sequence ACGTTGGATG that permits more efficient amplification.

Nucleotides not capitalized are mismatches introduced by the software MassARRAY Assay Design to enhance detection by MS.

F: forward orientation; R: reverse orientation.

**Appendix 2.2. Details and allele frequencies of 210 SNPs across *CYP4F2*, *CYP4F11* and *CYP4F12* genotyped in 149 human liver samples.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| *CYP4F12* | rs10854147 | 15642995 | A > G | 5' near gene | 0.798 | 100 | 0.077 | 0.116 | *HapMap R27* |
| *CYP4F12* | rs10854148 | 15643025 | C > T | 5' near gene | 0.798 | 100 | 0.077 | 0.116 | *HapMap R27* |
| *CYP4F12* | rs7245534 | 15643098 | T > C | 5' near gene | 0.798 | 100 | 0.077 | 0.083 | *HapMap R27* |
| *CYP4F12* | rs16980695 | 15643275 | A > T | 5' near gene | 0.095 | 93.3 | 0.241 | 0.263 | *HapMap R27* |
| *CYP4F12* | rs3813137 | 15643397 | C > T | 5' near gene | 1.000 | 99.3 | 0.041 | 0.054 | *HapMap R27* |
| *CYP4F12* | rs10405688[b] | 15643494 | A > G | 5' near gene | 0.000 | 0 | 0.000 | 0.089 | *HapMap R27* |
| *CYP4F12* | rs12977005 | 15643668 | A > G | 5' near gene | 0.736 | 92.6 | 0.083 | 0.116 | *HapMap R27* |
| *CYP4F12* | rs4347744 | 15643859 | C > G | 5' near gene | 1.000 | 96.6 | 0.069 | 0.127 | *HapMap R27* |
| *CYP4F12* | rs4331436 | 15643897 | C > T | 5' near gene | 0.595 | 98.7 | 0.088 | 0.116 | *HapMap R27* |
| *CYP4F12* | rs7248804[c] | 15644051 | A > G | 5' near gene | 7.584E-42 | 95.3 | 0.500 | 0.117 | *HapMap R27* |
| *CYP4F12* | rs2176917 | 15644236 | A > G | 5' near gene | 0.731 | 100 | 0.081 | 0.130 | *HapMap R27* |
| *CYP4F12* | rs2139749[b] | 15644378 | C > T | 5' near gene | 1.682E-24 | 56.4 | 0.500 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs17641206[b] | 15645152 | T > C | Intron 1 | 1.000 | 83.2 | 0.077 | 0.216 | *HapMap R27* |
| *CYP4F12* | rs10414690 | 15645195 | C > G | Intron 1 | 0.670 | 100 | 0.091 | 0.102 | *HapMap R27* |
| *CYP4F12* | rs16995376 | 15645377 | C > T | Exon 2, Missense, Pro13Leu | 0.106 | 100 | 0.232 | 0.250 | *HapMap R27* |
| *CYP4F12* | rs16995378[b] | 15645386 | T > C | Exon 2, Missense, Met16Thr | 0.407 | 61.1 | 0.126 | 0.096 | *HapMap R27* |
| *CYP4F12* | rs9305063 | 15646525 | T > C | Intron 2 | 0.798 | 100 | 0.077 | 0.102 | *HapMap R27* |
| *CYP4F12* | rs10410157 | 15646622 | C > T | Intron 2 | 0.798 | 100 | 0.077 | 0.112 | *HapMap R27* |
| *CYP4F12* | rs610142[a] | 15649930 | A > T | Intron 2 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs609636 | 15650098 | C > T | Exon 3, Missense, Asp76Asn | 0.776 | 97.3 | 0.079 | 0.117 | *HapMap R27* |
| *CYP4F12* | rs609290 | 15650140 | C > T | Exon 3, Missense, Val90Ile | 0.798 | 100 | 0.077 | 0.110 | *HapMap R27* |
| *CYP4F12* | rs17682485 | 15650698 | G > A | Intron 3 | 0.890 | 100 | 0.419 | 0.398 | *HapMap R27* |
| *CYP4F12* | rs12460703 | 15650776 | A > G | Intron 3 | 0.963 | 100 | 0.490 | 0.500 | *HapMap R27* |
| *CYP4F12* | rs675326 | 15651144 | G > A | Intron 3 | 1.000 | 98.7 | 0.003 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs16980704[b] | 15651531 | T > C | Intron 3 | 0.221 | 80.5 | 0.242 | 0.267 | *HapMap R27* |
| *CYP4F12* | rs677139[a] | 15651547 | A > G | Intron 3 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs688468[a] | 15651752 | A > G | Intron 3 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs2074568 | 15652132 | T > A | Intron 4 | 0.792 | 100 | 0.218 | 0.164 | *HapMap R27* |

**Appendix 2.2. Details and allele frequencies of 210 SNPs across *CYP4F2*, *CYP4F11* and *CYP4F12* genotyped in 149 human liver samples continued.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| *CYP4F12* | rs10409750 | 15652606 | A > G | Intron 5 | 0.569 | 100 | 0.433 | 0.397 | *HapMap R27* |
| *CYP4F12* | rs10415331 | 15652617 | C > T | Intron 5 | 0.709 | 97.3 | 0.083 | 0.121 | *HapMap R27* |
| *CYP4F12* | rs595381[a] | 15652809 | C > G | Intron 5 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs673020[b] | 15653081 | G > A | Intron 5 | 1.000 | 75.8 | 0.018 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs672645[a] | 15653127 | A > G | Intron 5 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs672555[b] | 15653185 | C > A | Intron 5 | 1.000 | 88.6 | 0.057 | 0.008 | *HapMap R27* |
| *CYP4F12* | rs11085969 | 15653546 | G > A | Intron 5 | 0.870 | 99.3 | 0.236 | 0.203 | *HapMap R27* |
| *CYP4F12* | rs12151032 | 15653566 | C > T | Intron 5 | 0.156 | 100 | 0.242 | 0.268 | *NCBI B36* |
| *CYP4F12* | rs610652[a] | 15653900 | C > T | Intron 5 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs2285888[b] | 15654235 | C > T | Exon 6, Missense, Arg188Cys | 0.009 | 83.9 | 0.360 | 0.483 | *NCBI B36* |
| *CYP4F12* | rs2285890 | 15654415 | C > T | Intron 6 | 1.000 | 100 | 0.030 | 0.061 | *HapMap R27* |
| *CYP4F12* | rs57578760[a] | 15655463 | G > C | Exon 7, Missense, Val270Leu | 1.000 | 100 | 0.000 | NA | *NCBI B36* |
| *CYP4F12* | rs628603[a] | 15655595 | A > G | Intron 7 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs2074569[a] | 15655727 | C > T | Intron 7 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs630133[a] | 15655769 | C > T | Intron 7 | 1.000 | 96 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs629748[a] | 15655843 | A > G | Intron 7 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs640169[a] | 15655885 | A > G | Intron 7 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs616745[a] | 15656420 | C > T | Intron 7 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs633929[b] | 15657441 | A > G | Intron 9 | 1.000 | 6.7 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs17682497[b] | 15657542 | C > T | Intron 9 | 0.094 | 87.9 | 0.206 | 0.200 | *HapMap R27* |
| *CYP4F12* | rs659825[a] | 15657584 | A > T | Intron 9 | 1.000 | 100 | 0.000 | 0.018 | *HapMap R27* |
| *CYP4F12* | rs659447 | 15657662 | C > T | Intron 9 | 1.000 | 96.6 | 0.031 | 0.051 | *HapMap R27* |
| *CYP4F12* | rs11085971 | 15657813 | G > T | Intron 9 | 1.000 | 92.6 | 0.152 | 0.067 | *HapMap R27* |
| *CYP4F12* | rs10420354 | 15657946 | C > T | Intron 9 | 0.252 | 100 | 0.158 | 0.085 | *HapMap R27* |
| *CYP4F12* | rs642322 | 15659173 | C > G | Intron 9 | 0.188 | 100 | 0.238 | 0.250 | *NCBI B36* |
| *CYP4F12* | rs631279 | 15659334 | T > C | Intron 9 | 0.188 | 100 | 0.238 | 0.268 | *NCBI B36* |
| *CYP4F12* | rs627753[b] | 15660130 | T > C | Intron 9 | 0.617 | 61.1 | 0.192 | 0.248 | *HapMap R27* |

**Appendix 2.2. Details and allele frequencies of 210 SNPs across *CYP4F2*, *CYP4F11* and *CYP4F12* genotyped in 149 human liver samples continued.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| *CYP4F12* | rs10410357 | 15661331 | A > G | Intron 9 | 0.252 | 100 | 0.158 | 0.083 | *HapMap R27* |
| *CYP4F12* | rs4109350 | 15661921 | C > T | Intron 9 | 0.234 | 98 | 0.236 | 0.264 | *NCBI B36* |
| *CYP4F12* | rs11879787 | 15662454 | A > G | Intron 9 | 0.676 | 100 | 0.242 | 0.186 | *HapMap R27* |
| *CYP4F12* | rs613503[a] | 15663113 | C > T | Intron 9 | 1.000 | 96.6 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F12* | rs680681 | 15663148 | T > C | Intron 9 | 0.188 | 100 | 0.238 | 0.242 | *NCBI B36* |
| *CYP4F12* | rs615392[c] | 15663595 | C > T | Intron 9 | 9.050E-05 | 91.3 | 0.346 | 0.264 | *NCBI B36* |
| *CYP4F12* | rs627971 | 15664094 | T > C | Intron 9 | 0.778 | 100 | 0.490 | 0.491 | *NCBI B36* |
| *CYP4F12* | rs665872 | 15664106 | C > T | Intron 9 | 0.649 | 97.3 | 0.493 | 0.470 | *NCBI B36* |
| *CYP4F12* | rs629651[b] | 15664417 | G > T | Intron 9 | 7.171E-05 | 83.9 | 0.376 | 0.458 | *NCBI B36* |
| *CYP4F12* | rs4808362 | 15664892 | G > A | Intron 9 | 0.902 | 98.7 | 0.497 | 0.483 | *HapMap R27* |
| *CYP4F12* | rs651109 | 15665158 | G > A | Intron 9 | 1.000 | 100 | 0.067 | 0.067 | *HapMap R27* |
| *CYP4F12* | rs59302543[a] | 15667846 | A > G | Exon 10, Missense, Ile406Val | 1.000 | 100 | 0.000 | NA | *NCBI B36* |
| *CYP4F12* | rs687774[b] | 15668062 | C > T | Intron 11 | 0.665 | 65.8 | 0.235 | 0.225 | *NCBI B36* |
| *CYP4F12* | rs688231 | 15668157 | A > G | Intron 11 | 0.188 | 100 | 0.238 | 0.225 | *NCBI B36* |
| *CYP4F12* | rs688256 | 15668175 | G > C | Intron 11 | 0.188 | 100 | 0.238 | 0.242 | *NCBI B36* |
| *CYP4F12* | rs688755 | 15668305 | T > C | Exon 12, Synonymous, Pro460Pro | 0.188 | 100 | 0.238 | 0.258 | *HapMap R27* |
| *CYP4F12* | rs61391486[a] | 15668720 | A > G | Exon 13, Missense, Asn467Ser | 1.000 | 100 | 0.000 | NA | *NCBI B36* |
| *CYP4F12* | rs593818 | 15668884 | G > A | Exon 13, Missense, Gly522Ser | 0.844 | 99.3 | 0.486 | 0.492 | *HapMap R27* |
| *CYP4F12* | rs594255 | 15668984 | G > A | 3' UTR | 0.188 | 100 | 0.238 | 0.242 | *NCBI B36* |
| *CYP4F12* | rs611848 | 15669379 | C > T | 3' near gene | 0.778 | 100 | 0.490 | 0.486 | *NCBI B36* |
| *CYP4F12* | rs624512 | 15669890 | G > A | 3' near gene | 0.188 | 100 | 0.238 | 0.258 | *HapMap R27* |
| *CYP4F12* | rs12976669 | 15671312 | A > C | 3' near gene | 0.283 | 100 | 0.272 | 0.250 | *HapMap R27* |
| *CYP4F12* | rs7255335 | 15672079 | G > C | 3' near gene | 0.283 | 100 | 0.272 | 0.228 | *HapMap R27* |
| *CYP4F12* | rs2886476 | 15672717 | T > G | 3' near gene | 0.283 | 100 | 0.272 | 0.254 | *HapMap R27* |
| *CYP4F12* | rs670476 | 15674467 | A > G | 3' near gene | 0.118 | 95.3 | 0.268 | 0.322 | *HapMap R27* |
| *CYP4F12* | rs590828 | 15674919 | T > C | 3' near gene | 1.000 | 98.7 | 0.480 | 0.450 | *HapMap R27* |
| - | rs4808369 | 15691869 | T > G | Intergenic, between CYP4F12 and CYP4F2 | 0.283 | 100 | 0.272 | 0.233 | *HapMap R27* |
| - | rs7252197 | 15696098 | T > G | Intergenic, between CYP4F12 and CYP4F2 | 1.000 | 100 | 0.087 | 0.110 | *HapMap R27* |

**Appendix 2.2. Details and allele frequencies of 210 SNPs across *CYP4F2*, *CYP4F11* and *CYP4F12* genotyped in 149 human liver samples continued.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| - | rs4433928 | 15699690 | G > A | Intergenic, between CYP4F12 and CYP4F2 | 0.744 | 99.3 | 0.470 | 0.500 | *HapMap R27* |
| - | rs16980800 | 15706047 | A > G | Intergenic, between CYP4F12 and CYP4F2 | 0.546 | 100 | 0.285 | 0.297 | *HapMap R27* |
| - | rs2240228 | 15713872 | G > A | Intergenic, between CYP4F12 and CYP4F2 | 0.001 | 94.6 | 0.241 | 0.300 | *HapMap R27* |
| - | rs2079234 | 15718510 | G > T | Intergenic, between CYP4F12 and CYP4F2 | 0.691 | 100 | 0.466 | 0.392 | *HapMap R27* |
| - | rs11666521 | 15722672 | T > C | Intergenic, between CYP4F12 and CYP4F2 | 0.966 | 100 | 0.268 | 0.233 | *HapMap R27* |
| - | rs10423986 | 15727655 | C > T | Intergenic, between CYP4F12 and CYP4F2 | 0.731 | 100 | 0.081 | 0.092 | *HapMap R27* |
| - | rs2189784 | 15820200 | G > A | Downstream of CYP4F2 | 0.216 | 100 | 0.433 | 0.381 | *HapMap R27* |
| - | rs2079288 | 15825203 | T > C | Downstream of CYP4F2 | 0.240 | 100 | 0.205 | 0.217 | *HapMap R27* |
| - | rs7252046 | 15832473 | T > C | Downstream of CYP4F2 | 0.957 | 100 | 0.336 | 0.325 | *HapMap R27* |
| - | rs12610189 | 15839641 | T > G | Downstream of CYP4F2 | 0.881 | 100 | 0.359 | 0.300 | *HapMap R27* |
| *CYP4F2* | rs3093216 | 15848737 | T > C | 3` near gene | 0.708 | 100 | 0.362 | 0.300 | *NCBI B36* |
| *CYP4F2* | rs3093212[b] | 15848966 | A > A | 3` near gene | 0.000 | 0 | 0.000 | 0.475 | *NCBI B36* |
| *CYP4F2* | rs3093211[c] | 15848967 | C > T | 3` near gene | 6.069E-44 | 100 | 0.500 | 0.200 | *NCBI B36* |
| *CYP4F2* | rs3093210[c] | 15849401 | G > A | 3` near gene | 1.718E-07 | 100 | 0.238 | 0.142 | *NCBI B36* |
| *CYP4F2* | rs3093209 | 15849420 | C > G | 3` near gene | 0.708 | 100 | 0.362 | 0.347 | *NCBI B36* |
| *CYP4F2* | rs3093208[a] | 15849582 | T > C | 3` near gene | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |
| *CYP4F2* | rs3093207 | 15849658 | A > G | 3` near gene | 0.708 | 100 | 0.362 | 0.300 | *HapMap R27* |
| *CYP4F2* | rs17756654 | 15849725 | C > A | 3` near gene | 1.000 | 100 | 0.044 | 0.033 | *NCBI B36* |
| *CYP4F2* | rs3093206 | 15849753 | AA > DEL | 3` near gene | 0.095 | 100 | 0.383 | 0.217 | *NCBI B36* |
| *CYP4F2* | rs1272 | 15850040 | G > C | Exon 13 | 0.581 | 100 | 0.275 | 0.180 | *HapMap R27* |
| *CYP4F2* | rs3093204 | 15850167 | G > A | Exon 13 | 1.000 | 100 | 0.057 | 0.065 | *NCBI B36* |
| *CYP4F2* | rs3093203 | 15850184 | C > T | Exon 13 | 0.791 | 99.3 | 0.260 | 0.292 | *NCBI B36* |
| *CYP4F2* | rs1126433[c] | 15850405 | A > G | Exon 13 | 2.236E-28 | 100 | 0.453 | 0.477 | *NCBI B36* |
| *CYP4F2* | rs3093201[a] | 15850454 | T > C | Exon 13 | 1.000 | 100 | 0.000 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs3093200 | 15850589 | C > A | Exon 13, Missense, Met519Leu | 1.000 | 100 | 0.057 | 0.083 | *HapMap R27* |
| *CYP4F2* | rs3093199[b] | 15850871 | A > T | Intron 12 | 0.000 | 0 | 0.000 | 0.284 | *NCBI B36* |
| *CYP4F2* | rs3093198 | 15851048 | C > T | Intron 12 | 1.000 | 100 | 0.332 | 0.340 | *NCBI B36* |
| *CYP4F2* | rs3093197[a] | 15851128 | A > T | Intron 12 | 1.000 | 100 | 0.000 | 0.000 | *NCBI B36* |

**Appendix 2.2. Details and allele frequencies of 210 SNPs across *CYP4F2*, *CYP4F11* and *CYP4F12* genotyped in 149 human liver samples continued.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| *CYP4F2* | rs2108622 | 15851431 | C > T | Exon 11, Missense, Met433Val | 1.000 | 100 | 0.289 | 0.233 | *HapMap R27* |
| *CYP4F2* | rs3093195 | 15852433 | G > A | Intron 9 | 0.856 | 100 | 0.262 | 0.325 | *NCBI B36* |
| *CYP4F2* | rs3093227 | 15852520 | G > T | Intron 9 | 1.000 | 100 | 0.010 | 0.023 | *NCBI B36* |
| *CYP4F2* | rs3093194 | 15852676 | G > A | Intron 9 | 0.856 | 100 | 0.262 | 0.332 | *HapMap R27* |
| *CYP4F2* | rs3093193 | 15852914 | G > C | Intron 9 | 0.708 | 100 | 0.362 | 0.333 | *NCBI B36* |
| *CYP4F2* | rs3093192[a] | 15853319 | C > T | Intron 9 | 1.000 | 100 | 0.000 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs12984060 | 15853553 | G > A | Intron 9 | 0.884 | 100 | 0.376 | 0.375 | *NCBI B36* |
| *CYP4F2* | rs3093226[b] | 15853780 | T > T | Intron 9 | 0.000 | 0 | 0.000 | 0.071 | *NCBI B36* |
| *CYP4F2* | rs2886296 | 15855478 | A > T | Intron 9 | 0.493 | 100 | 0.094 | 0.065 | *NCBI B36* |
| *CYP4F2* | rs3093184[b] | 15855730 | C > T | Intron 9 | 0.022 | 19.5 | 0.328 | 0.329 | *NCBI B36* |
| *CYP4F2* | rs3093180 | 15856148 | C > T | Intron 9 | 1.000 | 100 | 0.171 | 0.153 | *NCBI B36* |
| *CYP4F2* | rs3093173 | 15856788 | G > A | Intron 9 | 1.000 | 100 | 0.168 | 0.095 | *NCBI B36* |
| *CYP4F2* | rs3093170 | 15856974 | T > G | Intron 9 | 0.002 | 100 | 0.487 | 0.432 | *NCBI B36* |
| *CYP4F2* | rs3093169 | 15857008 | G > A | Intron 9 | 0.032 | 100 | 0.154 | 0.139 | *NCBI B36* |
| *CYP4F2* | rs3093168 | 15857245 | C > T | Intron 9 | 0.644 | 100 | 0.393 | 0.345 | *HapMap R27* |
| *CYP4F2* | rs3093167 | 15857292 | T > C | Intron 9 | 1.000 | 100 | 0.168 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093166 | 15857349 | A > T | Intron 9 | 0.644 | 100 | 0.393 | 0.345 | *HapMap R27* |
| *CYP4F2* | rs3093163[a] | 15857497 | G > A | Intron 9 | 1.000 | 100 | 0.000 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs2074900 | 15857820 | G > A | Exon 9, Synonymous, His343His | 0.002 | 100 | 0.225 | 0.325 | *NCBI B36* |
| *CYP4F2* | rs3093160 | 15857907 | G > A | Intron 8 | 1.000 | 100 | 0.168 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs2074901 | 15858422 | A > C | Intron 7 | 1.000 | 100 | 0.168 | 0.100 | *HapMap R27* |
| *CYP4F2* | rs1558139 | 15858564 | A > G | Intron 7 | 0.002 | 100 | 0.487 | 0.433 | *HapMap R27* |
| *CYP4F2* | rs3093158[b] | 15861166 | A > G | Intron 8 | 6.418E-06 | 48.3 | 0.417 | 0.286 | *NCBI B36* |
| *CYP4F2* | rs3093156 | 15861609 | A:T | Intron 6 | 0.709 | 98.7 | 0.442 | 0.433 | *HapMap R27* |
| *CYP4F2* | rs3093150 | 15862629 | C > T | Intron 5 | 0.867 | 100 | 0.074 | 0.045 | *NCBI B36* |
| *CYP4F2* | rs3093145 | 15862852 | C:A | Intron 5 | 0.593 | 100 | 0.443 | 0.441 | *NCBI B36* |
| *CYP4F2* | rs3093144 | 15863297 | G > A | Intron 5 | 1.000 | 100 | 0.185 | 0.224 | *NCBI B36* |
| *CYP4F2* | rs3093141[a] | 15863521 | C > T | Intron 5 | 1.000 | 100 | 0.000 | 0.000 | *HapMap R27* |

**Appendix 2.2. Details and allele frequencies of 210 SNPs across *CYP4F2*, *CYP4F11* and *CYP4F12* genotyped in 149 human liver samples continued.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| *CYP4F2* | rs3093135 | 15865371 | T > A | Intron 3 | 1.000 | 100 | 0.168 | 0.118 | *HapMap R27* |
| *CYP4F2* | rs3093134 | 15865396 | G > C | Intron 3 | 1.000 | 100 | 0.168 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs2006193 | 15865800 | C > A | Intron 3 | 0.040 | 100 | 0.151 | 0.142 | *NCBI B36* |
| *CYP4F2* | rs2365178 | 15866003 | G > A | Intron 3 | 1.000 | 100 | 0.168 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs3093129 | 15866019 | A > G | Intron 3 | 1.000 | 100 | 0.168 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs3093128 | 15866104 | G > A | Intron 3 | 1.000 | 100 | 0.168 | 0.142 | *NCBI B36* |
| *CYP4F2* | rs3093126[a] | 15866160 | A > G | Intron 3 | 1.000 | 100 | 0.000 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs2016503 | 15866185 | T > C | Intron 3 | 0.890 | 100 | 0.181 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093124 | 15866232 | T > C | Intron 3 | 1.000 | 99.3 | 0.169 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093122[b] | 15866276 | C > G | Intron 3 | 1.000 | 3.4 | 0.100 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093121 | 15866277 | C > A | Intron 3 | 1.000 | 100 | 0.168 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs736089[b] | 15866300 | A > G | Intron 3 | 0.588 | 43 | 0.133 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs2215092 | 15866521 | G > T | Intron 3 | 0.867 | 100 | 0.074 | 0.043 | *NCBI B36* |
| *CYP4F2* | rs3093116 | 15866574 | G > A | Intron 3 | 1.000 | 100 | 0.168 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093115 | 15866615 | T > A | Intron 3 | 1.000 | 100 | 0.168 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs984692 | 15867074 | T > A | Intron 3 | 0.676 | 100 | 0.242 | 0.208 | *NCBI B36* |
| *CYP4F2* | rs3093114 | 15867413 | C > T | Exon 3, Synonymous, Ala82Ala | 1.000 | 100 | 0.168 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs3093113[a] | 15867545 | A > G | Intron 2 | 1.000 | 100 | 0.000 | 0.000 | *NCBI B36* |
| *CYP4F2* | rs3093112 | 15867611 | A > G | Intron 2 | 1.000 | 100 | 0.168 | 0.114 | *NCBI B36* |
| *CYP4F2* | rs3093110 | 15868784 | T > C | Intron 2 | 1.000 | 100 | 0.168 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs2074902[c] | 15869099 | C > T | Intron 2 | 8.325E-38 | 100 | 0.487 | 0.180 | *HapMap R27* |
| *CYP4F2* | rs3093106 | 15869257 | A > G | Exon 2, Synonymous, Pro55Pro | 1.000 | 100 | 0.168 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093105[c] | 15869388 | G > T | Exon 2, Missense, Gly12Trp | 8.325E-38 | 100 | 0.487 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093103 | 15869434 | T > C | Intron 1 | 1.000 | 100 | 0.168 | 0.109 | *NCBI B36* |
| *CYP4F2* | rs3093100 | 15869469 | G > C | Intron 1 | 1.000 | 100 | 0.168 | 0.139 | *NCBI B36* |
| *CYP4F2* | rs3093098 | 15869512 | T > C | Intron 1 | 0.871 | 100 | 0.171 | 0.128 | *HapMap R27* |
| *CYP4F2* | rs3093097 | 15869690 | CA > DEL | Intron 1 | 0.748 | 100 | 0.282 | NA | - |
| *CYP4F2* | rs3093092[c] | 15870127 | C > A | 3` near gene | 1.000E-04 | 96.6 | 0.212 | 0.142 | *NCBI B36* |

**Appendix 2.2. Details and allele frequencies of 210 SNPs across *CYP4F2*, *CYP4F11* and *CYP4F12* genotyped in 149 human liver samples continued.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| *CYP4F2* | rs3093091[a] | 15870133 | C > T | 3` near gene | 1.000 | 100 | 0.000 | 0.222 | *NCBI B36* |
| *CYP4F2* | rs3761014 | 15872763 | C > G | 3` near gene | 1.000 | 100 | 0.185 | 0.207 | *HapMap R27* |
| *CYP4F11* | rs2018460 | 15873757 | G > A | 3` near gene | 1.000 | 100 | 0.185 | 0.217 | *HapMap R27* |
| *CYP4F11* | rs2365177 | 15880408 | T > G | 3` near gene | 1.000 | 95.3 | 0.035 | 0.042 | *HapMap R27* |
| *CYP4F11* | rs8102331[b] | 15883144 | G > A | 3` near gene | 1.000 | 82.6 | 0.069 | 0.333 | *HapMap R27* |
| *CYP4F11* | rs12459933 | 15884116 | T > C | 3` near gene | 0.460 | 99.3 | 0.365 | 0.308 | *NCBI B36* |
| *CYP4F11* | rs12610962 | 15884377 | G > A | 3' UTR | 0.268 | 100 | 0.500 | 0.483 | *HapMap R27* |
| *CYP4F11* | rs1060467 | 15885538 | T > C | 3' UTR | 0.692 | 99.3 | 0.334 | 0.425 | *HapMap R27* |
| *CYP4F11* | rs1064796 | 15885662 | G > C | Exon 12, Synonymous, Thr485Thr | 0.524 | 94 | 0.282 | 0.283 | *HapMap R27* |
| *CYP4F11* | rs2072269 | 15885739 | G > A | Intron 11 | 0.203 | 100 | 0.497 | 0.467 | *HapMap R27* |
| *CYP4F11* | rs1060463[b] | 15886176 | G > A | Exon 11, Missense, Asn446Asp | 8.314E-06 | 67.8 | 0.381 | 0.392 | *HapMap R27* |
| *CYP4F11* | rs12977516 | 15886919 | G > C | Intron 8 | 0.450 | 96.6 | 0.420 | 0.491 | *HapMap R27* |
| *CYP4F11* | rs12971888 | 15887204 | G > A | Intron 8 | 0.782 | 100 | 0.440 | 0.417 | *HapMap R27* |
| *CYP4F11* | rs6512074 | 15887894 | A > G | Intron 8 | 0.496 | 92.6 | 0.471 | 0.361 | *NCBI B36* |
| *CYP4F11* | rs6512075 | 15888067 | C > T | Intron 8 | 0.677 | 99.3 | 0.449 | 0.382 | *HapMap R27* |
| *CYP4F11* | rs7253051 | 15888131 | G > C | Intron 8 | 0.600 | 100 | 0.426 | 0.486 | *NCBI B36* |
| *CYP4F11* | rs1471112 | 15891458 | T > G | Intron 8 | 1.000 | 100 | 0.265 | 0.242 | *NCBI B36* |
| *CYP4F11* | rs12978309 | 15892188 | T > A | Intron 8 | 0.782 | 100 | 0.440 | 0.361 | *NCBI B36* |
| *CYP4F11* | rs11086012 | 15893643 | C > A | Intron 8 | 0.966 | 100 | 0.268 | 0.283 | *HapMap R27* |
| *CYP4F11* | rs4808409[b] | 15894379 | G > T | Intron 6 | 0.008 | 73.8 | 0.445 | 0.492 | *NCBI B36* |
| *CYP4F11* | rs3746152 | 15895519 | T > A | Intron 6 | 1.000 | 100 | 0.037 | 0.025 | *HapMap R27* |
| *CYP4F11* | rs8104361[b] | 15895714 | A > G | Exon 6, Missense, Arg276Cys | 1.000 | 48.3 | 0.014 | 0.258 | *HapMap R27* |
| *CYP4F11* | rs7249167 | 15898359 | G > A | Intron 4 | 0.603 | 100 | 0.174 | 0.167 | *HapMap R27* |
| *CYP4F11* | rs2305804 | 15899365 | G > C | Intron 2 | 0.094 | 96.6 | 0.125 | 0.121 | *HapMap R27* |
| *CYP4F11* | rs3765071 | 15901034 | G > A | Intron 2 | 1.000 | 96.6 | 0.441 | 0.361 | *NCBI B36* |
| *CYP4F11* | rs3765070[b] | 15901292 | C > T | Exon 2, Synonymous, Ile106Ile | 0.318 | 83.9 | 0.428 | 0.397 | *HapMap R27* |
| *CYP4F11* | rs2305801[b] | 15906141 | G > A | Exon 1, Synonymous, Gly26Gly | 0.207 | 74.5 | 0.135 | 0.250 | *HapMap R27* |
| *CYP4F11* | rs2305800 | 15906294 | T > C | 5' UTR | 0.807 | 100 | 0.262 | 0.190 | *HapMap R27* |

**Appendix 2.2. Details and allele frequencies of 210 SNPs across *CYP4F2, CYP4F11* and *CYP4F12* genotyped in 149 human liver samples continued.**

| Gene | SNP | Chromosomal Position | SNP Alleles | SNP Location and Function | HWE *P*-Value | % Call Rate | MAF (this study) | MAF (public database) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| *CYP4F11* | rs11879253[b] | 15906491 | G > C | 5' UTR | 0.163 | 68.5 | 0.147 | 0.250 | *NCBI B36* |
| *CYP4F11* | rs12985091 | 15906749 | G > A | 5' near gene | 0.417 | 96 | 0.392 | 0.481 | *HapMap R27* |
| *CYP4F11* | rs3826950 | 15906891 | C > T | 5' near gene | 0.871 | 99.3 | 0.260 | 0.192 | *HapMap R27* |
| *CYP4F11* | rs3810428 | 15907478 | C > A | 5' near gene | 0.019 | 99.3 | 0.240 | 0.145 | *HapMap R27* |
| *CYP4F11* | rs3810427 | 15907650 | G > T | 5' near gene | 0.509 | 100 | 0.406 | 0.492 | *HapMap R27* |
| *CYP4F11* | rs16980968 | 15907925 | C > T | 5' near gene | 1.000 | 89.9 | 0.015 | NA | *NCBI B36* |
| *CYP4F11* | rs12460831 | 15908758 | T > C | 5' near gene | 0.818 | 96 | 0.266 | 0.186 | *HapMap R27* |
| *CYP4F11* | rs2116951 | 15908806 | A > G | 5' near gene | 1.000 | 100 | 0.040 | 0.018 | *HapMap R27* |
| *CYP4F11* | rs11670533 | 15911160 | G > A | 5' near gene | 0.337 | 99.3 | 0.341 | 0.308 | *HapMap R27* |
| *CYP4F11* | rs4375789 | 15912804 | G > A | 5' near gene | 1.000 | 98.7 | 0.265 | 0.173 | *HapMap R27* |
| *CYP4F11* | rs2116952 | 15913255 | C > T | 5' near gene | 0.015 | 99.3 | 0.101 | 0.125 | *HapMap R27* |
| *CYP4F11* | rs2163859 | 15914143 | C > A | 5' near gene | 1.000 | 100 | 0.265 | 0.192 | *HapMap R27* |
| - | rs12327750[c] | 15916687 | T > C | Upstream of CYP4F11 | 1.784E-41 | 98.7 | 0.497 | 0.195 | *HapMap R27* |
| - | rs10500211 | 15919348 | A > T | Upstream of CYP4F11 | 0.895 | 96 | 0.164 | 0.125 | *HapMap R27* |
| - | rs17641489[c] | 15920483 | A > G | Upstream of CYP4F11 | 2.411E-43 | 98.7 | 0.500 | 0.052 | *HapMap R27* |

Frequency data were compiled from HapMap and NCBI dbSNP databases.

Chromosomal positions are given in base pairs from the p-telomere of chromosome 19, as per HapMap Data release 27, February 2009, NCBI B36 assembly, dbSNP b126.

[a] SNPs were monomorphic.

[b] SNPs gave <90% call rate.

[c] SNPs were not in HWE.

**Appendix 2.3: Tagging SNPs across the *CYP4F12-CYP4F2-CYP4F11* locus ($r^2 \geq 0.9$).**

| Tagging SNP | Tagged SNPs | Gene | Chromosomal Location | Localisation |
|---|---|---|---|---|
| rs3093209 | rs3093209 | *CYP4F2* | 15849420 | 3` near gene |
| | rs12610189 | *CYP4F2* | 15839641 | Downstream of CYP4F2 |
| | rs3093216 | *CYP4F2* | 15848737 | 3` near gene |
| | rs3093207 | *CYP4F2* | 15849658 | 3` near gene |
| | rs3093206 | *CYP4F2* | 15849753 | 3` near gene |
| | rs3093193 | *CYP4F2* | 15852914 | Intron 9 |
| rs3093173 | rs3093173 | *CYP4F2* | 15856788 | Intron 9 |
| | rs3093180 | *CYP4F2* | 15856148 | Intron 9 |
| | rs3093167 | *CYP4F2* | 15857292 | Intron 9 |
| | rs3093160 | *CYP4F2* | 15857907 | Intron 8 |
| | rs2074901 | *CYP4F2* | 15858422 | Intron 7 |
| | rs3093135 | *CYP4F2* | 15865371 | Intron 3 |
| | rs3093134 | *CYP4F2* | 15865396 | Intron 3 |
| | rs2365178 | *CYP4F2* | 15866003 | Intron 3 |
| | rs3093129 | *CYP4F2* | 15866019 | Intron 3 |
| | rs3093128 | *CYP4F2* | 15866104 | Intron 3 |
| | rs3093124 | *CYP4F2* | 15866232 | Intron 3 |
| | rs3093121 | *CYP4F2* | 15866277 | Intron 3 |
| | rs3093116 | *CYP4F2* | 15866574 | Intron 3 |
| | rs3093115 | *CYP4F2* | 15866615 | Intron 3 |
| | rs3093114 | *CYP4F2* | 15867413 | Exon 3, Synonymous, Ala82Ala |
| | rs3093112 | *CYP4F2* | 15867611 | Intron 2 |
| | rs3093110 | *CYP4F2* | 15868784 | Intron 2 |
| | rs3093106 | *CYP4F2* | 15869257 | Exon 2, Synonymous, Pro55Pro |
| | rs3093103 | *CYP4F2* | 15869434 | Intron 1 |
| | rs3093100 | *CYP4F2* | 15869469 | Intron 1 |
| | rs3093098 | *CYP4F2* | 15869512 | Intron 1 |

**Appendix 2.3: Tagging SNPs across the *CYP4F12-CYP4F2-CYP4F11* locus ($r^2 \geq 0.9$), continued.**

| Tagging SNP | Tagged SNPs | Gene | Chromosomal Location | Localisation |
|---|---|---|---|---|
| rs3093169 | rs3093169 | *CYP4F2* | 15865800 | Intron 3 |
| | rs2006193 | *CYP4F2* | 15857008 | Intron 9 |
| rs12977516 | rs12977516 | *CYP4F11* | 15886919 | Intron 8 |
| | rs7253051 | *CYP4F11* | 15888131 | Intron 8 |
| rs627971 | rs627971 | *CYP4F12* | 15664094 | Intron 9 |
| | rs665872 | *CYP4F12* | 15664106 | Intron 9 |
| | rs4808362 | *CYP4F12* | 15664892 | Intron 9 |
| | rs593818 | *CYP4F12* | 15668884 | Exon 13, Missense, Gly522Ser |
| | rs611848 | *CYP4F12* | 15669379 | 3' near gene |
| rs2886476 | rs2886476 | *CYP4F12* | 15672717 | 3' near gene |
| | rs12976669 | *CYP4F12* | 15671312 | 3' near gene |
| | rs7255335 | *CYP4F12* | 15672079 | 3' near gene |
| | rs4808369 | - | 15691869 | Intergenic, between CYP4F12 and CYP4F2 |

Chromosomal positions are given in base pairs from the p-telomere of chromosome 19, as per HapMap Data release 27, NCBI B36 assembly, dbSNP b126.