

Genetics of the anticoagulant drug warfarin

Yen-Yu Chen

This dissertation is submitted for the degree of Doctor of Philosophy

St Edmund's College
University of Cambridge

30th November 2007

PREFACE

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

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

ABSTRACT

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

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

The two studies were analysed separately for genetic association to warfarin dose requirement (single marker and haplotypic tests). Common SNPs in the vitamin K epoxide reductase gene (*VKORC1*) are significantly associated with dose in the Uppsala and WARG studies ($p = 1.9 \times 10^{-15}$ and 6.5×10^{-100} , respectively). Cytochrome P450 2C9 (*CYP2C9*) has been known to affect dose requirement and was confirmed in both Swedish cohorts ($p = 2.3 \times 10^{-5}$ and 4.9×10^{-32}). The two genes together explain ~40% of warfarin dose variation. SNPs in microsomal epoxide hydrolase (*EPHX1*) and orosomucoid 1 (*ORM1*) genes do not show a broad effect but are associated with dose in both studies. Genes encoding PROC, APOE, CALU, PDIA2 and GGCX showed nominal association with dose in the Uppsala study. Likewise, *PROS1*, *CYP1A1*, *CYP3A4*, *PDIA5*, *PDIA3* and *F10* showed nominal association to dose in the WARG study. Most of these minor effects, if real, are most likely to be population/treatment specific. A model taking in to account genetic factors (*VKORC1* and *CYP2C9**2 / *3) and non genetic factors (age, gender and drug interaction) together explained more than 50% inter-individual dose variance.

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

To my dear parents

Mr Chi-Hwa Chen

Mrs Huei-Wan Liu-Chen

ACKNOWLEDGEMENTS

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

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

TABLE OF CONTENTS

Preface		II
Abstract		III
Acknowledgement		V
Table of contents		VI
List of Tables		XI
List of Figures		XIV
List of Abbreviations		XVII
Publications arising from this work		XXI
Chapter I	Introduction	1
1.1	Pharmacogenetics and pharmacogenomics	2
1.1.1	Introduction	2
1.1.2	Benefit of PGx	3
1.1.3	Challenges of PGx practice	6
1.1.4	Personalised medicine	7
1.2	Drug metabolism and adverse drug response (ADR)	8
1.2.1	Drug metabolism	8
1.2.2	ADR caused by drug metabolising genes	10
1.3	International effort on drug safety	13
1.3.1	ADR and cost paid by society	13
1.3.2	Efforts coordinated in USA	14
1.3.3	Efforts coordinated in Europe	15
1.4	Human genome	17
1.4.1	The material of inheritance	17
1.4.2	The Human Genome Project	18
1.4.3	Human genetic variation	19
1.5	Human sequence variation	20
1.5.1	Variable number of tandem repeat (VNTR)	21
1.5.2	Copy number variation (CNV)	22
1.5.3	Single nucleotide polymorphism (SNP)	23
1.6	Linkage disequilibrium & association study	25
1.6.1	The HapMap Project	25
1.6.2	Haplotype Tag SNPs	26
1.6.3	Association Studies	27
1.7	Coagulation and anticoagulant	28
1.7.1	Coagulation	28
1.7.2	Anticoagulant Drugs	30

1.7.3	Example of drug safety of an anticoagulant	32
1.8	Warfarin	33
1.8.1	Introduction	33
1.8.2	Warfarin and coumarin derivatives	35
1.8.3	Side effect and adverse reactions	35
1.8.4	Drug interaction and effective therapeutic dose	37
1.9	This thesis	39
Chapter II	Materials and Methods	41
2.1	Patients	42
2.1.1	Uppsala study	43
2.1.2	WARG study	44
2.1.3	Patients subject to severe bleeding	44
2.1.4	DNA preparation	44
2.1.4.1	<i>Uppsala study</i>	45
2.1.4.2	<i>WARG study</i>	45
2.2	Genetic marker selection	46
2.2.1	Single nucleotide polymorphism (SNP)	46
2.2.2	Microsatellite repeat marker	46
2.3	Mass spectrometry genotyping	47
2.3.1	Assay design	47
2.3.2	PCR amplification of SNP loci	47
2.3.2.1	<i>MassEXTEND</i>	47
2.3.2.2	<i>iPLEX</i>	48
2.3.3	Shrimp alkaline phosphatase treatment	48
2.3.4	Oligo extension	49
2.3.4.1	<i>MassEXTEND</i>	49
2.3.4.2	<i>iPLEX</i>	50
2.3.5	Desalting	50
2.3.6	Sample spotting and analysis	50
2.4	Other genotyping	52
2.4.1	Taqman genotyping	52
2.4.2	Microsatellite genotyping for GGCX	53
2.5	Exon resequencing	55
2.5.1	Introduction	55
2.5.2	Primer design	55
2.5.3	PCR and sequencing	56
2.5.4	Sequence analysis	57
2.6	VKORC1 expression in human liver	58
2.6.1	RNA extraction	58

2.6.2	Assay of VKORC1 mRNA	59
2.7	Computational analysis	61
2.7.1	Data processing	61
2.7.2	Genotype analysis	61
2.7.3	Multiple sequence alignment	62
2.7.4	Evolutional conserver region (ECR) analysis	63
2.8	Statistical analysis	64
2.8.1	Statistics in Uppsala study	64
2.8.2	Statistics in WARG study	64
2.8.3	Statistics in Case/control association	65
Chapter III	Selection of candidate gene and construction of LD maps	66
3.1	Introduction	67
3.2	Genotyping Study Design	70
3.3	Candidate Genes in Warfarin Transportation	74
3.3.1	ORM1 and ORM2	75
3.3.2	ABCB1 (MDR1)	77
3.4	Candidate Genes in Warfarin Metabolism	80
3.4.1	S-warfarin metabolism (CYP2C8, CYP2C9, CYP2C18, and CYP219)	81
3.4.2	R-warfarin metabolism (CYP1A1, CYP1A2, CYP3A4 and CYP3A5)	83
3.4.3	P450 inducibility	86
3.5	Candidate Genes in Vitamin K Intake and Recycling	91
3.5.1	APOE	92
3.5.2	VKORC1	92
3.5.3	EPHX1	94
3.5.4	NQO1	96
3.5.5	GGCX	97
3.5.6	CALU	99
3.5.7	Protein Disulfide Isomerase (PDI)	101
3.6	Vitamin K Dependent Proteins	108
3.6.1	Gene cluster on chromosome 13 (F7, F10, PROZ)	109
3.6.2	Other VKD genes (F2, F9, PROC, PROS1 and GAS6)	113
3.7	Other Coagulation Factors	118
3.7.1	Anti-thrombin III	119
3.7.2	Coagulation factor V	119
3.8	Conclusion	123
Chapter IV	An investigation of genetic determinants of warfarin dose requirement in 201 Swedish patients (Uppsala cohort)	124

4.1	Patient recruitment	125
4.2	Univariate association analysis	128
4.3	<i>VKORC1</i> and <i>GGCX</i>	133
4.3.1	<i>VKORC1</i>	133
4.3.2	Fine mapping across the 500 kb LD block harbouring <i>VKORC1</i> in Caucasians	136
4.3.3	The search of causative variants in <i>VKORC1</i>	140
4.3.4	<i>GGCX</i>	144
4.4	CYP2C cluster	148
4.5	Other nominally associated genes	153
4.6	Predictive Model for warfarin dose	156
4.7	Conclusions	159
Chapter V	An investigation of genetic determinants of warfarin dose requirement in 1500 Swedish patients (WARG cohort)	160
5.1	The National Warfarin Genetic (WARG) Study	162
5.2	Genotyping approach	167
5.2.1	SNP selection	167
5.2.2	Taqman genotyping CYP2C9*2 and *3	168
5.2.3	Genotyping summary	168
5.3	Dose variation	170
5.3.1	Gender, age, and dose	170
5.3.2	Univariate regression	171
5.3.3	Comparison of results in the Uppsala and WARG studies	177
5.3.4	Multivariate	180
5.4	Regression model	181
5.5	Therapeutic stabilisation	183
5.5.1	INR stabilisation	183
5.5.2	Over-anticoagulation	186
5.6	Conclusion	189
Chapter VI	Warfarin and adverse drug reaction	190
6.1	Introduction	191
6.2	Patients	192
6.3	Statistical power	195
6.4	Case-control analysis - All bleeders	198
6.4.1	Single marker	198
6.4.2	Two- and three-marker sliding window haplotype	199
6.4.2.1	<i>Two marker haplotype</i>	200
6.4.2.2	<i>Three marker haplotype</i>	202

6.4.3	Summary of findings for all bleeder groups	203
6.4.3.1	<i>CYP2C9</i>	204
6.4.3.2	<i>VKORC1</i>	205
6.5	Single study case-control analysis	207
6.5.1	Bleeders in the Uppsala study	208
6.5.2	Bleeders in the WARG study	209
6.5.3	Summary of findings for separate bleeder groups	210
6.5.3.4	<i>Cytochrome P450 2C9</i>	211
6.6	The protein disulfide isomerase gene family	213
6.6.1	Three SNPs in trend and Fisher's exact tests on P4HB, PDIA4 and PDIA6	215
6.6.2	The gastrointestinal bleeding sub-phenotype	218
6.6.3	Predictors of gastrointestinal bleeding	220
6.6.4	Predictors of non-gastrointestinal bleeding	222
6.7	Re-sequencing	225
6.7.1	Candidate genes for bleeding	225
6.7.2	Sequencing results	226
6.8	Summary	231
Chapter VII	Summary and Discussion	232
7.1	This thesis	233
7.2	Molecular mechanism	238
7.3	Concluding remarks	240
Reference		241
Appendix 1	PRIMERS FOR 216 TAG SNPS	i
Appendix 2	UNIVARIATE RESULT OF DOSE ASSOCIATION (UPPSALA STUDY)	xi
Appendix 3	UNIVARIATE RESULT OF DOSE ASSOCIATION (WARG STUDY)	xx
Appendix 4	TREND TEST RESULT IN BLEEDING (ALL BLEEDERS)	xxv
Appendix 5	TREND TEST RESULT IN BLEEDING (UPPSALA BLEEDERS)	xxx
Appendix 6	TREND TEST RESULT IN BLEEDING (WARG BLEEDERS)	xxxv
Appendix 7	SEQUENCE VARIANTS IDENTIFIED WITH EXON RE-SEQUENCING	xl

LIST OF TABLES

Table 1.1.	Haepatic enzymes involved in drug metabolism.	9
Table 1.2.	Prescribed anticoagulants and antiplatelet drugs.	31
Table 1.3.	Interacting drugs with increased risk of severe bleeding in warfarin treatment.	45
Table 2.1	Oligo sequence for amplying microsatellite in intron 6 of <i>GGCX</i> .	53
Table 2.2.	Thermo cycling conditions of <i>GGCX</i> microsatellite PCR.	53
Table 2.3.	<i>GGCX</i> microsatellite genotype and expected amplicon.	54
Table 2.4.	Primer sequences for <i>VORCI</i> and <i>GAPDH</i> gene expression.	60
Table 2.5	Scripts used for data processing.	61
Table 2.6	SNP exclusion criteria in Haploview.	62
Table 3.1.	Summary of iterative genotyping.	71
Table 3.2.	Summary of SNP genotyping for each of the candidate genes.	73
Table 3.3.	Candidate genes in warfarin transportation	74
Table 3.4.	Genotyping summary of <i>ORM1</i> and <i>ORM2</i> .	76
Table 3.5.	Genotyping summary of <i>ABCB1</i> .	78
Table 3.6.	Candidate genes in warfarin metabolism.	80
Table 3.7.	Genotyping summary for candidate genes in S-warfarin metabolism.	82
Table 3.8.	Genotyping summary for candidate genes in R-warfarin metabolism.	83
Table 3.9.	Candidate genes of regulating warfarin metaboliser.	87
Table 3.10.	Genotyping summary of <i>NR1I2</i> and <i>NR1I3</i> .	88
Table 3.11.	Candidate genes in vitamin K intake and recycling.	91
Table 3.12.	Genotyping summary of <i>APOE</i> .	92
Table 3.13.	Genotyping summary of <i>VKORC1</i> and its nearby flanking region.	93
Table 3.14.	Genotyping summary of <i>EPHX1</i> .	95
Table 3.15.	Genotyping summary of <i>NQO1</i> .	96
Table 3.16.	Genotyping summary of <i>GGCX</i> .	98
Table 3.17.	Genotyping summary of <i>CALU</i> .	100
Table 3.18.	Genotyping summary of <i>PDI</i> family.	104
Table 3.19.	Candidate genes of vitamin K dependent.	108
Table 3.20.	Genotyping summary of <i>F7</i> , <i>F10</i> and <i>PROZ</i> cluster.	111
Table 3.21.	Genotyping summary of other VKD candidate genes.	113
Table 3.22.	Other candidate genes in this study.	118
Table 3.23.	Genotyping summary of antithrombin III and factor V.	118

Table 4.1.	Medical information statistics of patients in Uppsala cohort.	126
Table 4.2.	Univariate analysis in regression model.	129
Table 4.3.	Two or three marker haplotype association in dose variation.	132
Table 4.4.	Significance of pairwise haplotype comparison.	136
Table 4.5.	Nominally significant result in multivariate regression model for SNPs in <i>CYP2C</i> gene cluster.	151
Table 4.6.	Association test of non-genetic factors.	156
Table 4.7.	Predictors in final multiple model of warfarin dose variation.	157
Table 5.1.	Medical information statistics of patients in WARG cohort.	165
Table 5.2.	Male/female ratio in a selection of warfarin genetic studies.	166
Table 5.3.	Gender effect in warfarin dose association.	170
Table 5.4.	Patient in the analyses with different inclusion criteria.	172
Table 5.5.	SNPs showing significant association after Meff correction.	174
Table 5.6.	Comparison between Uppsala and validation studies.	179
Table 5.7.	Other SNPs showing nominal significant association with dose.	179
Table 5.8.	Predictors in multiple regression model.	181
Table 5.9.	SNPs associated with over anti-coagulation in warfarin treatment.	186
Table 6.1.	Age distribution of case/control in the Uppsala and WARG studies.	194
Table 6.2.	SNPs nominally associated with bleeding complication in Cochran-Armitage trend test and Fisher's exact test.	199
Table 6.3.	P-value of two marker sliding window.	200
Table 6.4.	P-value of three marker sliding window.	202
Table 6.5.	Bleeding association of single marker, 2 SNPs haplotype and 3 SNP haplotype.	203
Table 6.6.	Bleeding phenotype of patients in Uppsala and WARG studies.	207
Table 6.7.	Association of bleeders in Uppsala study.	208
Table 6.8.	Association of bleeders in WARG study.	210
Table 6.9.	Association of 2-SNP haplotype (rs10272564-rs10269104) in <i>PDIA4</i> .	210
Table 6.10.	Summary of single study analysis.	211
Table 6.11.	Result of <i>CYP2C9</i> *2 and *3 allele.	212
Table 6.12.	The most significant SNPs in <i>PDIA4</i> , <i>P4HB</i> and <i>PDIA6</i> .	215
Table 6.13.	Population frequency of genotypic combination of the three SNPs in <i>PDIA6</i> (rs1686482), <i>PDIA4</i> (rs4727005) and <i>P4HB</i> (rs1799919).	217
Table 6.14.	Bleeding sub-phenotype stratification with genotype combinations in WARG bleeders.	219
Table 6.15.	Association of gastrointestinal bleeding in the WARG cohort.	220
Table 6.16.	Prediction of gastrointestinal bleeding using rs1198873 in the WARG study.	221

Table 6.17.	Association of gastrointestinal bleeding based on recessive action assumption.	222
Table 6.18.	Fisher's exact test on non-gastrointestinal bleeding in the WARG study.	223
Table 6.19.	Association of non-gastrointestinal bleeding based on recessive action assumption.	224
Table 6.20.	Sequencing summary of 11 candidate genes.	227
Table 6.21.	Consequence of SNPs identified in warfarin-treated patients.	228
Table 7.1.	Predictors for warfarin dosing algorithm.	236

LIST OF FIGURES

Figure 1.1	Structures for (A) Suxamethonium; (B) Cholinesterase.	3
Figure 1.2	The mutated size and rate of genetic polymorphisms	21
Figure 1.3.	The coagulation cascade.	29
Figure 1.4.	Warfarin (A) chemical structure; (B) tablet used in UK (Photo by Gonegonegone).	34
Figure 3.1.	Genes selected in this study.	69
Figure 3.2.	cDNA sequence alignment of ORM1 and ORM2.	74
Figure 3.3.	(A) Swedish LD structure; (B) LD structure from Hapmap project phase II result.	76
Figure 3.4.	LD structure from (A) 201 Swedish and (B) Hapmap CEU.	79
Figure 3.5.	LD structure of CYP1A1 and CYP1A2 in (A) Swedish and (B) Hapmap CEU populations.	84
Figure 3.6.	Pairwise cDNA sequence alignment of <i>CYP1A1</i> and <i>CYP1A2</i> .	85
Figure 3.7.	LD structure of <i>CYP3A4</i> and <i>CYP3A5</i> in (A) Swedish and (B) Hapmap CEU populations.	86
Figure 3.8.	LD structure of (A) <i>NR1I2</i> and (B) <i>NR1I3</i> .	89
Figure 3.9.	Hapmap CEU result for <i>NR1I2</i> region.	90
Figure 3.10.	LD map of <i>VKORC1</i> in Hapmap CEU panel.	94
Figure 3.11.	LD plots for <i>EPHX1</i> region in (A) Uppsala study and (B) Hapmap CEU panel.	95
Figure 3.12.	Genetic architecture of the <i>NQO1</i> locus in (A) Swedish and (B) Hapmap Caucasians.	97
Figure 3.13.	Genomic architecture of GGCX region in Hapmap CEU panel.	99
Figure 3.14.	Genomic architecture of <i>CALU</i> in (A) Swedish and (B) Hapmap CEU.	101
Figure 3.15.	Hypothetical model of protein disulfide isomerase functionality.	102
Figure 3.16	Genomic architectures of (A) <i>PDIA2</i> , (B) <i>PDIA3</i> , (C) <i>PDIA4</i> , (D) <i>PDIA5</i> and (E) <i>PDIA6</i> in Hapmap.	106
Figure 3.17.	LD structures of six genes in PDI family A.	107
Figure 3.18.	The coagulation cascade.	110
Figure 3.19.	Genomic architecture and LD organisation of the chromosome 13 locus harbouring <i>F7</i> , <i>F10</i> and <i>PROZ</i> in (A) Swedish and (B) Hapmap Caucasians.	112
Figure 3.20.	LD architecture of (A) F2 and (B) F9.	115
Figure 3.21.	Genomic architecture of protein C and protein S in Swedish.	116
Figure 3.22.	Genomic structure of growth arrest-specific gene 6 (GAS6) peptide in (A) Swedish and (B) Hapmap CEU panel.	117
Figure 3.23.	Genomic architecture of <i>SERPINC1</i> in Swedish.	119
Figure 3.24.	Genomic architecture of factor V in Swedish.	120
Figure 3.25.	Genomic architecture of F5 region in Hapmap (A) Caucasian, (B) Han-Chinese and (C) Japanese.	121

Figure 4.1.	Mean weekly dose of rs2359612 genotype in 201 Swedish.	134
Figure 4.2.	Genomic structure, haplotype, and linkage disequilibrium of <i>VKORC1</i> .	135
Figure 4.3.	Genomic architecture of <i>VKORC1</i> region in Hapmap CEU panel (Caucasian).	137
Figure 4.4.	SNP selection for fine mapping <i>VKORC1</i> locus.	138
Figure 4.5.	Association result of fine mapping <i>VKORC1</i> locus.	139
Figure 4.6.	<i>VKORC1</i> mRNA expression assays.	140
Figure 4.7.	Electrophoresis of RNA samples prepared from the liver biopsies.	141
Figure 4.8.	Genomic sequence alignment in <i>VKORC1</i> locus of human, chimpanzee, mouse, rat, dog and chicken.	142
Figure 4.9.	Alternative splicing variants of <i>VKORC1</i> .	143
Figure 4.10.	Genomic structure, haplotype, and linkage disequilibrium of <i>GGCX</i> .	145
Figure 4.11.	Mean weekly dose of rs12714145 genotype in 201 Swedish.	145
Figure 4.12.	Individuals are divided into 4 groups.	146
Figure 4.13.	Genomic architecture of CYP2C gene cluster on chromosome 10.	148
Figure 4.14.	Genomic architecture of CYP2C gene cluster in 201 Swedish for (A) Pairwise D prime and (B) Pairwise r^2 calculation.	150
Figure 4.15.	Multiple models of dose explained by genetic and non-genetic factors.	157
Figure 5.1.	Age distribution of patients in the Uppsala and WARG studies.	162
Figure 5.2.	The number of PT INR recorded to each patient.	164
Figure 5.3.	Minor allele frequency distribution of 216 SNPs genotyped in WARG study.	169
Figure 5.4.	Age effect on warfarin dose association in (A) male and (B) female patients Patient's age and maintenance dose is plotted and a linear trend line is shown accordingly.	171
Figure 5.5.	Quantile-quantile plot for univariate analysis in dose association.	173
Figure 5.6.	Mean weekly dose of rs9923231 genotype in WARG Swedish.	176
Figure 5.7.	Mean weekly dose of rs9923231 genotype in WARG patients.	177
Figure 5.8.	Multiple regression model developed accordingly in the WARG study.	182
Figure 5.9.	Lowess smoothed plot of PT INR values of patients treated with warfarin.	184
Figure 5.10.	Lowess smoothed plot of PT INR values of patients treated with warfarin.	185
Figure 5.11.	Survival (Kaplan-Meier) curve of cumulative probability in patients with PT INR > 4 related to (A) <i>VKORC1</i> rs9923231 and (B) <i>CYP2C9</i> *2 and *3.	188
Figure 6.1.	Age of each bleeding patient collected in the Uppsala and WARG studies.	192
Figure 6.2.	Age distribution of bleeders in the Uppsala and WARG studies.	193

Figure 6.3.	Statistical power calculation of cases required to achieve 80% power based on the assumption of 2% bleeding prevalence in population of causative variants of (A) 0.01; (B) 0.05; (C) 0.1; (D) 0.25; (E) 0.5 in MAF.	196
Figure 6.4.	Bleeding association of individual haplotype in (A) <i>F2</i> , (B) <i>CYP2C8</i> , (C) <i>ABCB1</i> and (D) <i>NR1I2</i> .	201
Figure 6.5.	Single marker and 2- / 3- marker haplotype analysis of <i>CYP2C9</i> (A) *2 and (B) *3 alleles.	205
Figure 6.6.	Single marker and 2- / 3- marker haplotype analysis of <i>VKORC1</i> .	206
Figure 6.7.	In silico analysis of mRNA expression using Unigene database.	214
Figure 6.8.	All 27 combinations of the top associated SNPs in <i>P4HB</i> (rs1799919), <i>PDIA4</i> (rs4727005) and <i>PDIA6</i> (rs1686482).	216
Figure 6.9.	Ratio of rs1198873 genotype in the gastrointestinal (GI) bleeding and control patients in the WARG study.	221
Figure 6.10.	Candidate genes for exon sequencing.	226
Figure 6.11.	Minor allele frequency of novel SNPs found specific in Swedish (blue) and in CEPH (red).	229
Figure 6.12.	Allele frequency differences of SNPs found in both Swedish patients and CEPH Caucasians.	229

LIST OF ABBREVIATIONS

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

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

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

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

PUBLICATIONS ARISING FROM THIS WORK

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

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

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

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