CHAPTER VI

WARFARIN AND ADVERSE DRUG REACTION

6.1 INTRODUCTION

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

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

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

6.2 PATIENTS

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

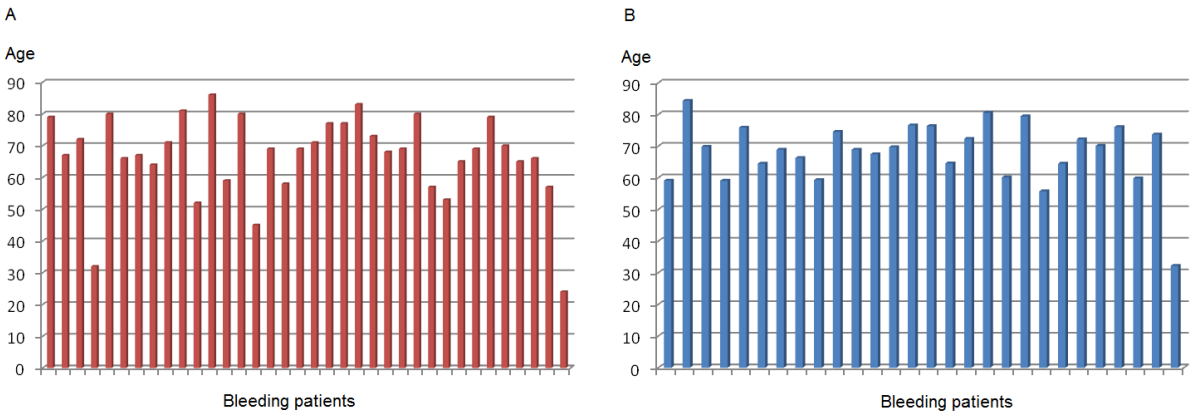

Figure 6.1. Age of each bleeding patient collected in the Uppsala and WARG studies.

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

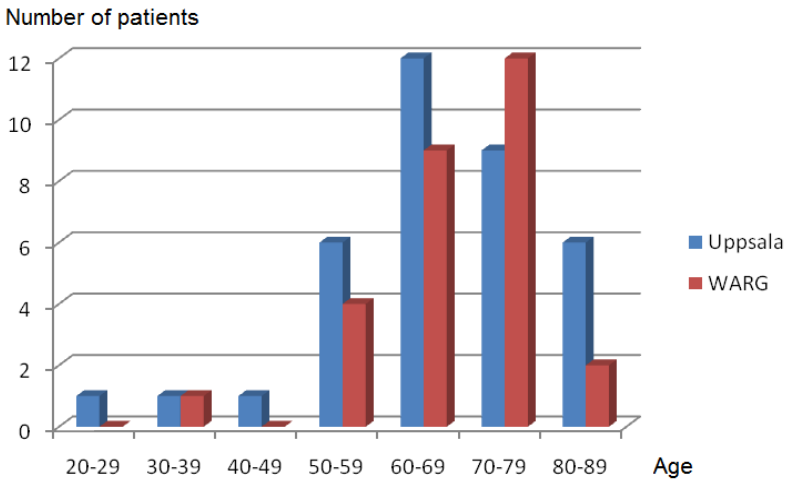

Figure 6.2. Age distribution of bleeders in the Uppsala and WARG studies.

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

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

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

6.3 STATISTICAL POWER

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

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

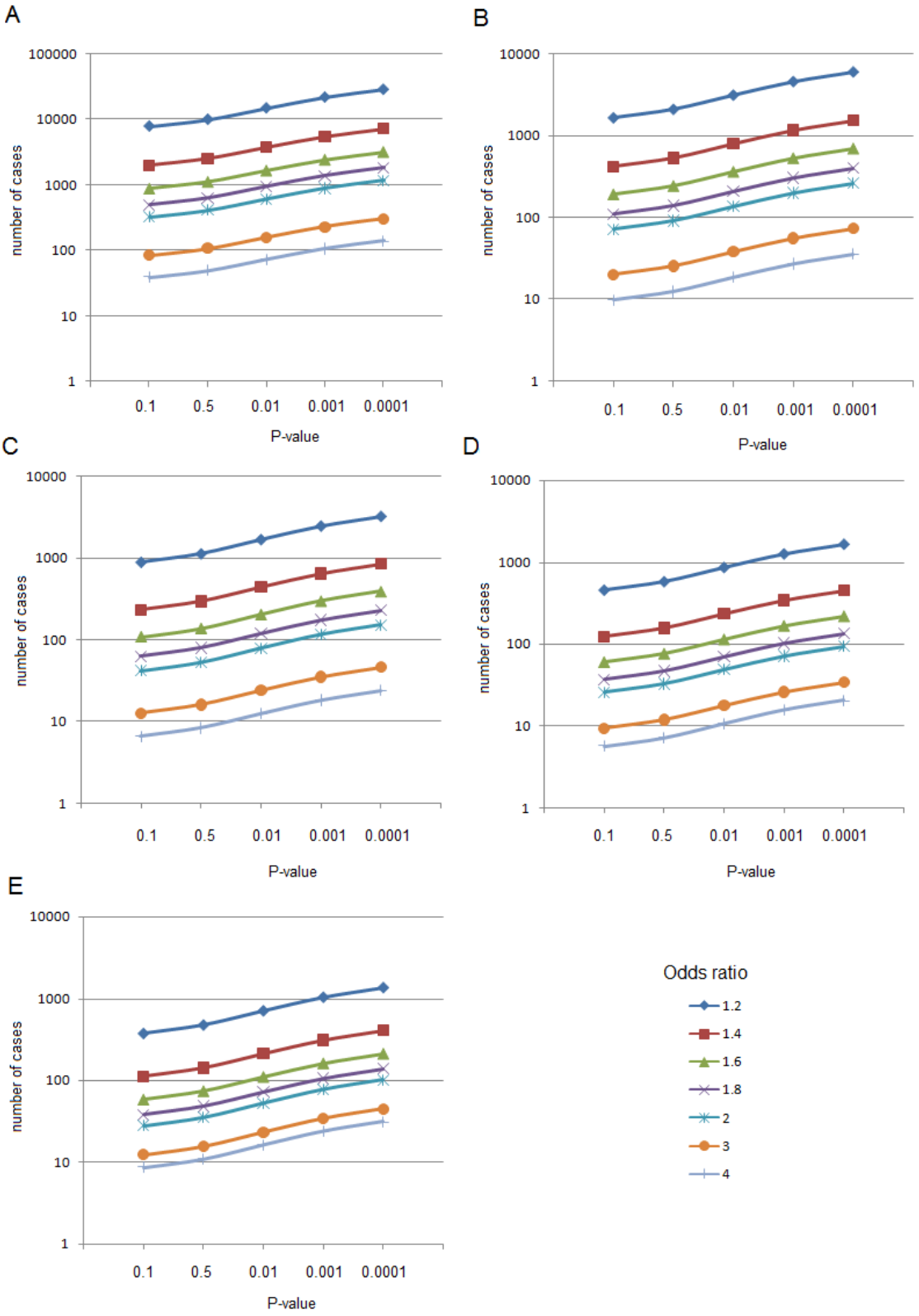

Figure 6.3. Statistical power calculation of cases required to achieve 80% power based on the assumption of 2% bleeding prevalence in population of causative variants of (A) 0.01; (B) 0.05; (C) 0.1; (D) 0.25; (E) 0.5 in MAF.

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

6.4 CASE-CONTROL ANALYSIS - ALL BLEEDERS

6.4.1 Single marker

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

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

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

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

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

6.4.2 Two- and three-maker sliding window haplotype

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

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

6.4.2.1 Two marker haplotype

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

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

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

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

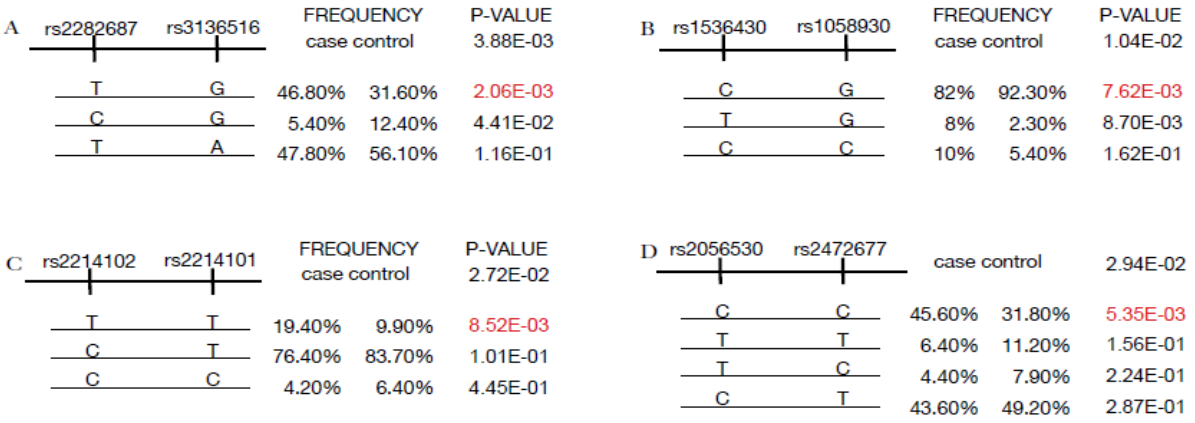

Figure 6.4. Bleeding association of individual haplotype in (A) *F2*, (B) *CYP2C8*, (C) *ABCB1* and (D) *NR1I2*. Omnibus p-value was listed on the top row for each gene whilst p-values for each haplotype in the window were noted next to corresponding haplotype and frequencies in case and control groups.

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

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

6.4.2.2 Three marker haplotype

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

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

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

6.4.3 Summary of findings for all bleeder groups

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

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

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

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

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

6.4.3.1 CYP2C9

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

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

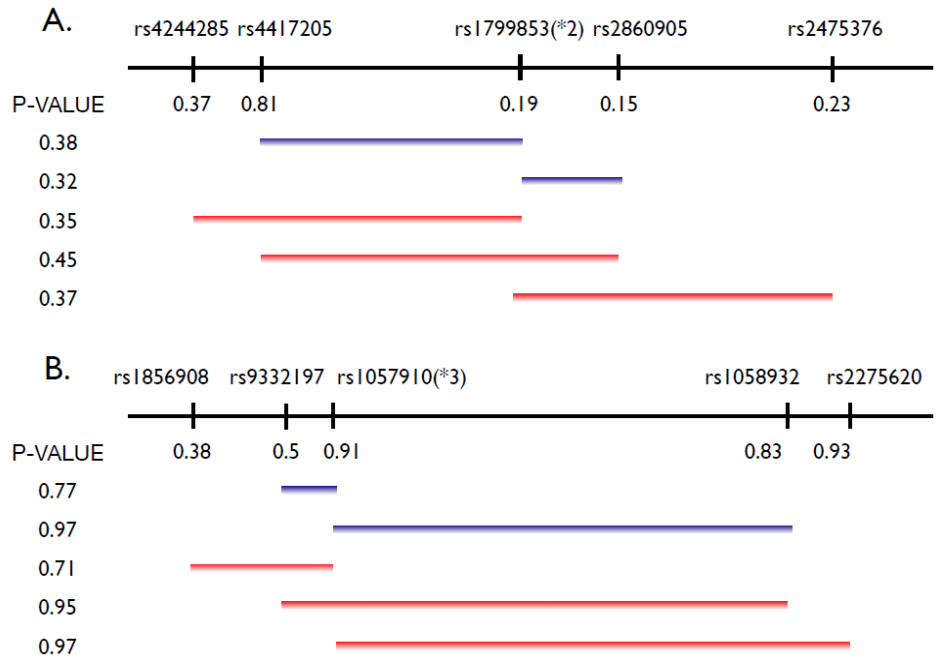

Figure 6.5. Single marker and 2- / 3- marker haplotype analysis of *CYP2C9* (A) *2 and (B) *3 alleles. P-values of single marker trend test are noted below each SNP. The omnibus p-values are noted on the left of each sliding window.

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

6.4.3.2 VKORC1

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

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

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

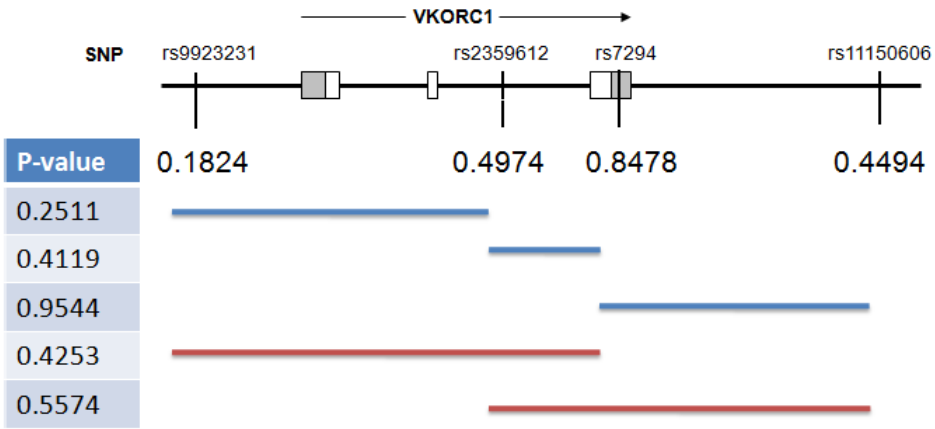

Figure 6.6. Single marker and 2- / 3- marker haplotype analysis of *VKORC1*. P-values from the single marker trend test are noted below each SNP. The omnibus p-values are noted on the left of each sliding window

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

6.5 SINGLE STUDY CASE-CONTROL ANALYSIS

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

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

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

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

6.5.1 Bleeders in the Uppsala study

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

Table 6.7. Association of bleeders in Uppsala study.

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

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

6.5.2 Bleeders in the WARG study

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

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

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

Table 6.8. Association of bleeders in WARG study.

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

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

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

6.5.3 Summary of findings for separate bleeder groups

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

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

Table 6.10. Summary of single study analysis.

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

6.5.3.4 Cytochrome P450 2C9

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

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

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

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

6.6 THE PROTEIN DISULFIDE ISOMERASE GENE FAMILY

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

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

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

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

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

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

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

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

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

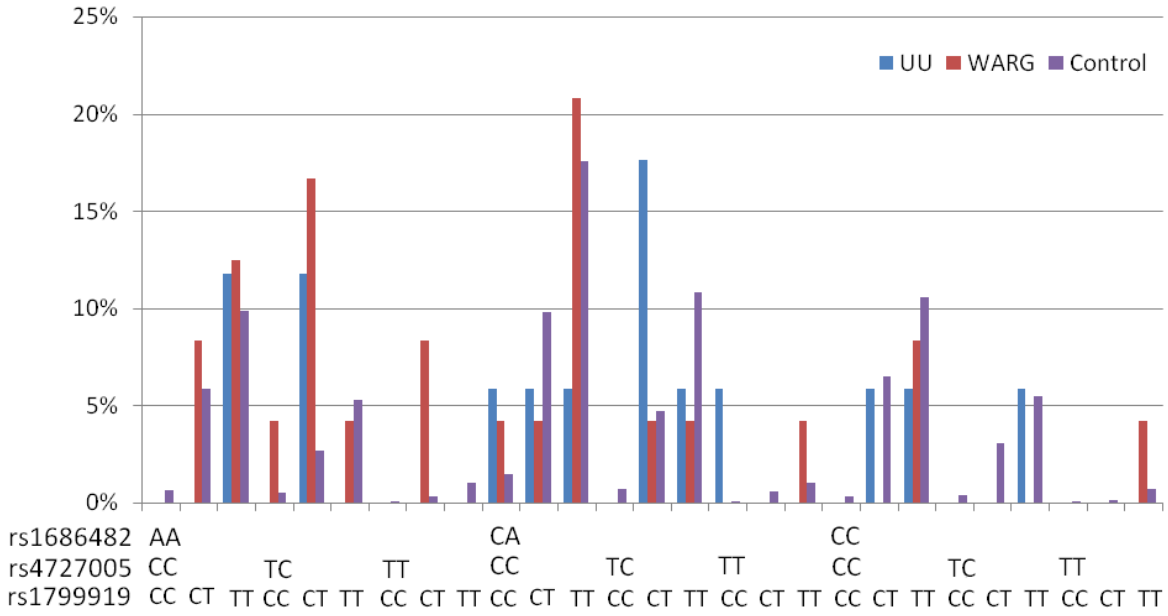

Figure 6.8. All 27 combinations of the top associated SNPs in *P4HB* (rs1799919), *PDIA4* (rs4727005) and *PDIA6* (rs1686482). Colour of the columns represents the proportion of genotype combination in different cohorts of Uppsala bleeder (blue), WARG bleeder (red) and all controls (Uppsala and WARG).

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

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

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

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

6.6.2 The gastrointestinal bleeding sub-phenotype

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

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

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

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

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

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

6.6.3 Predictors of gastrointestinal bleeding

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

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

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

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

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

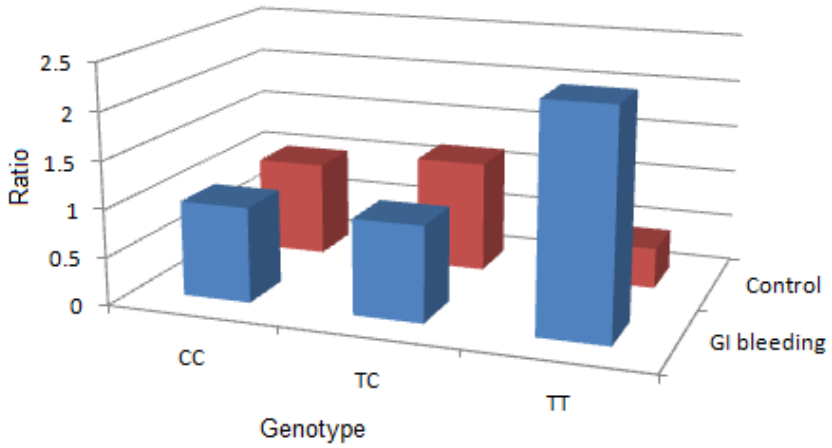

Figure 6.9. Ratio of rs1198873 genotype in the gastrointestinal (GI) bleeding and control patients in the WARG study. The ratio of TC and TT genotypes were presented accordingly to CC genotype in both groups. A significant enrichment of TT genotyped is found in GI bleeders.

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

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

		WARG	
		Case	Control
Prediction	Positive	7	233
	Negative	6	1191

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

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

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

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

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

6.6.4 Predictors of non-gastrointestinal bleeding

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

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

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

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

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

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

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

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

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

6.7 RE-SEQUENCING

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

6.7.1 Candidate genes for bleeding

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

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

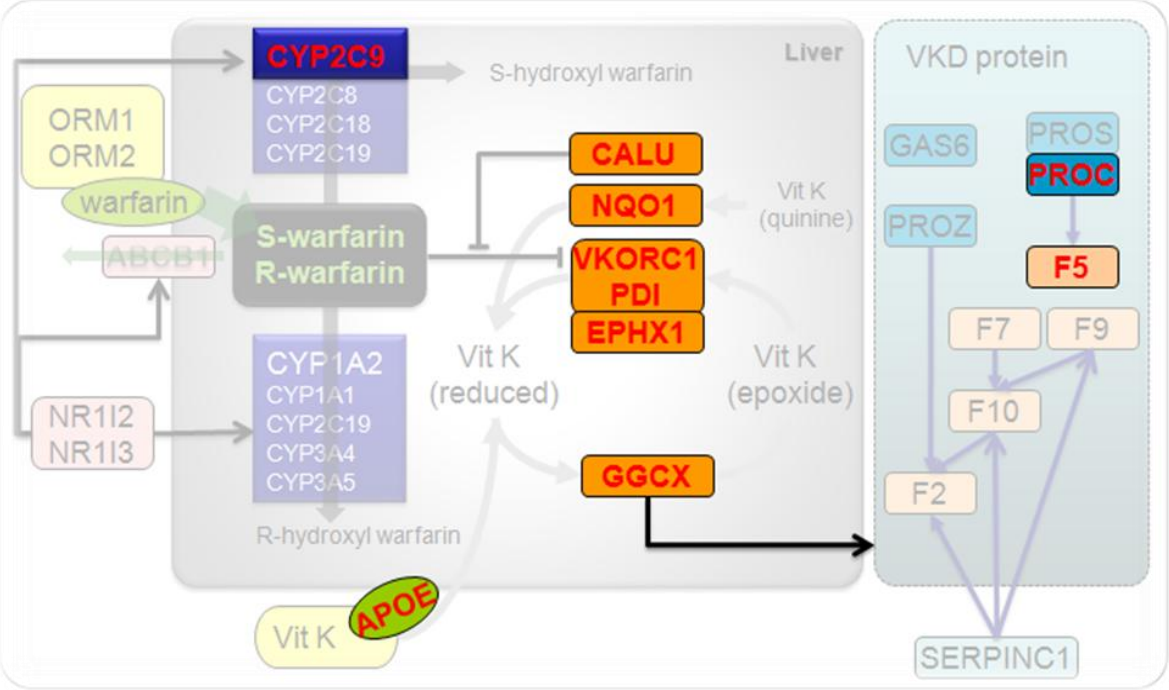

Figure 6.10. Candidate genes for exon sequencing. Among the 35 candidate genes, used for searching for the genetic determinants for associations with warfarin dose and bleeding complications using genotyping technology, a subset of 11 genes were selected to discover common and rare variants using sequencing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

6.7.2 Sequencing results

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

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

Table 6.20. Sequencing summary of 11 candidate genes.

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

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

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

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

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

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

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

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

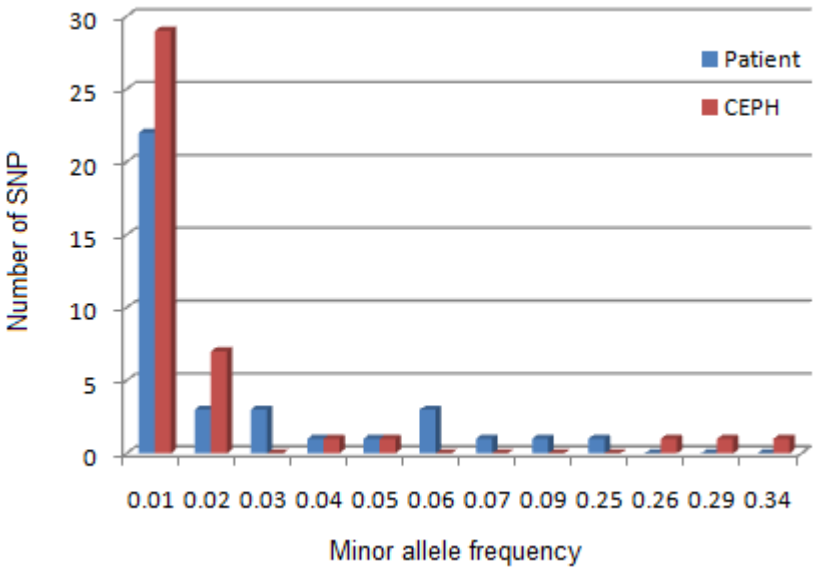

Figure 6.11. Minor allele frequency of novel SNPs found specific in Swedish (blue) and in CEPH (red).

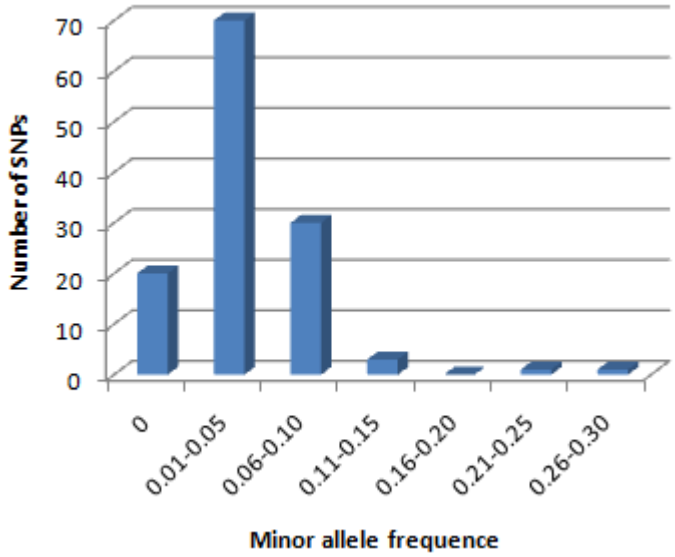

Figure 6.12. Allele frequency differences of SNPs found in both Swedish patients and CEPH Caucasians.

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

6.8 SUMMARY

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

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