CHAPTER I

INTRODUCTION

1.1 PHARMACOGENETICS AND PHAMARCOGENOMICS

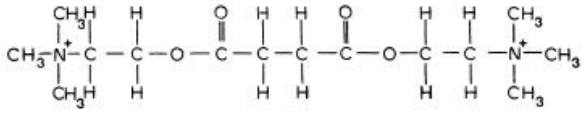
Each human being carries its own unique genetic makeup. In concert with the environmental factors to which he / she is exposed, this may result in differing reactions to external treatments including drugs.

1.1.1 Introduction

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

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

A



B

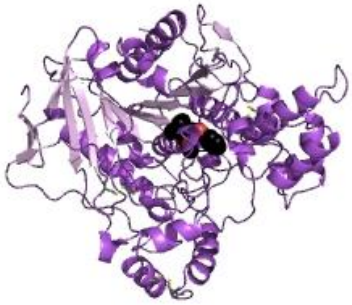


Figure 1.1. Structures for (A) Suxamethonium; (B) Human cholinesterase. Cartoon diagram of figure 1.1B is modified from Protein Data Bank (PDB) 1P0I.

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

1.1.2 Benefit of PGx

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

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

1. *Right drug and right dose.*

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

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

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

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

3. Powerful medicine and drug discovery

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

1.1.3 Challenges of PGx practice

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

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

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

1.1.4 Personalised medicine

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

1.2 DRUG METABOLISM AND ADVERSE DRUG RESPONSE (ADR)

1.2.1 Drug metabolism

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

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

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

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

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

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

Table 1.1 Hepatic enzymes involved in drug metabolism.

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

1.2.2 ADR caused by drug metabolising genes

1. N-acetyltransferase

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

2. UDP-glucuronosyltransferases

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

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

3. Glucose-6-phosphate dehydrogenase

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

4. Cytochrome P450 2D6

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

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

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

1.3 INTERNATIONAL EFFORT ON DRUG SAFETY

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

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

1.3.1 ADR and cost paid by society

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

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

1.3.2 Efforts coordinated in USA

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

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

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

1.3.3 Efforts coordinated in Europe

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

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

1.3.4 Important issues

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

(a) characterisation of ADR cases and controls;

(b) required contact of an optimal genotyping panel;

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

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

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

1.4 HUMAN GENOME

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

1.4.1 The material of inheritance

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

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

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

1.4.2 The Human Genome Project

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

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

1.4.3 Human genetic variation

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

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

1.5 HUMAN SEQUENCE VARIATION

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

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

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

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

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

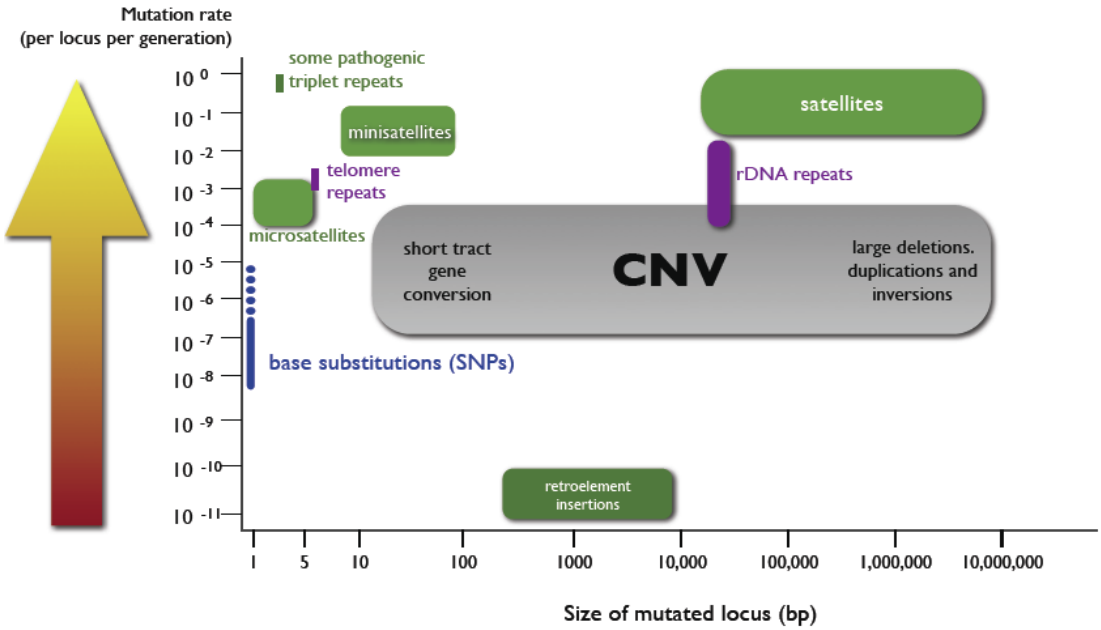


Figure 1.2. The mutated size and rate of genetic polymorphisms. Various sequence variations have been identified during the use of DNA sequencing technology. Smaller satellites (minisatellites and microsatellites) have been extensive applied in genetic disease association and DNA profiling, such as forensic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

1.5.1 Variable number of tandem repeat (VNTR)

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

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

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

1.5.2 Copy number variation (CNV)

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

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

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

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

1.5.3 Single nucleotide polymorphism (SNP)

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

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

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

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

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

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

1.6 LINKAGE DISEQUILIBRIUM & ASSOCIATION STUDY

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

1.6.1 The HapMap Project

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

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

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

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

1.6.2 Haplotype Tag SNPs

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

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

1.6.3 Association Studies

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

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

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

1.7 COAGULATION AND ANTICOAGULANTS

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

1.7.1 Coagulation

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

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

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

bleeding (Riddel, Aouizerat et al. 2007).

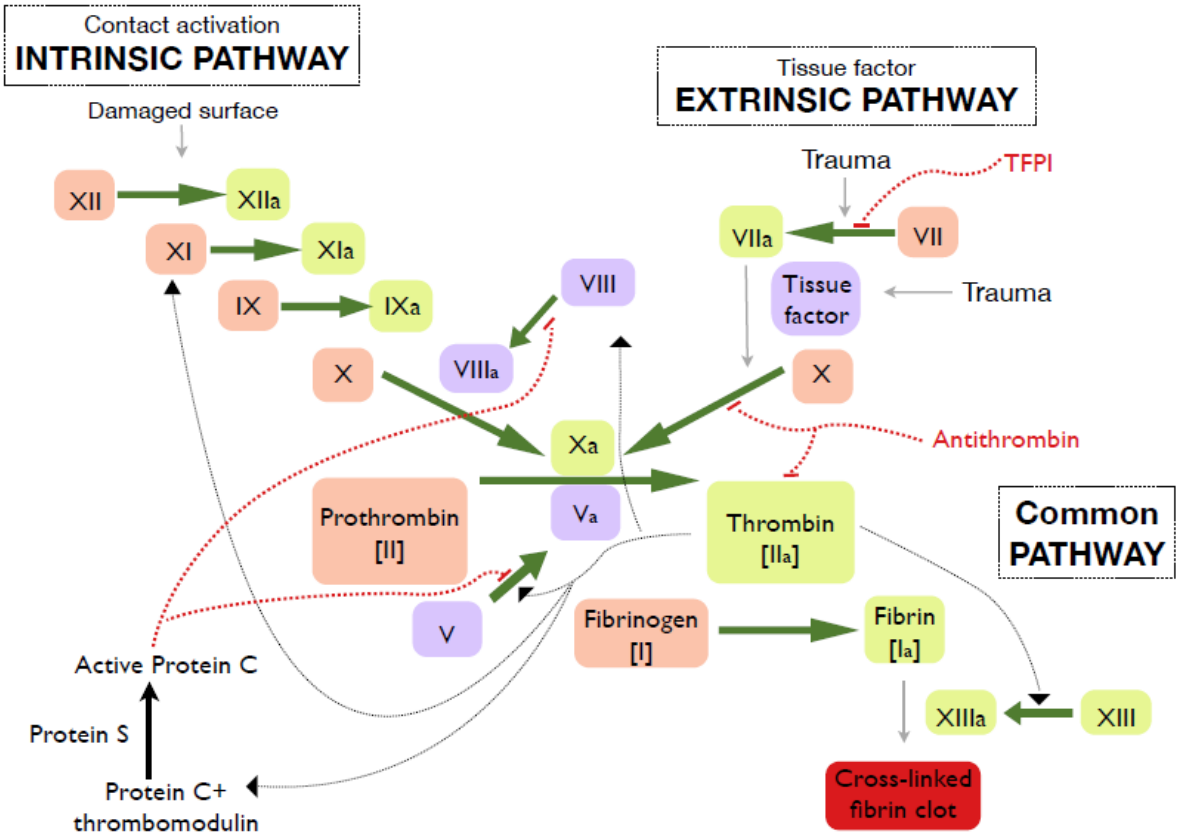


Figure 1.3. The coagulation cascade. The cascade is involved in the phase II coagulation including intrinsic, extrinsic and common pathways. All these coagulation factors exist in inactive form. Protein C (with help from Protein S), Antithrombin, and TFPI act as natural anticoagulant to stop fibrin clot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

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

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

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

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

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

1.7.2 Anticoagulant Drugs

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

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

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

Table 1.2. Prescribed anticoagulants and antiplatelet drugs.

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

1.7.3 Example of drug safety of an anticoagulant

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

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

1.8 WARFARIN

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

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

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

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

1.8.1 Introduction

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

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

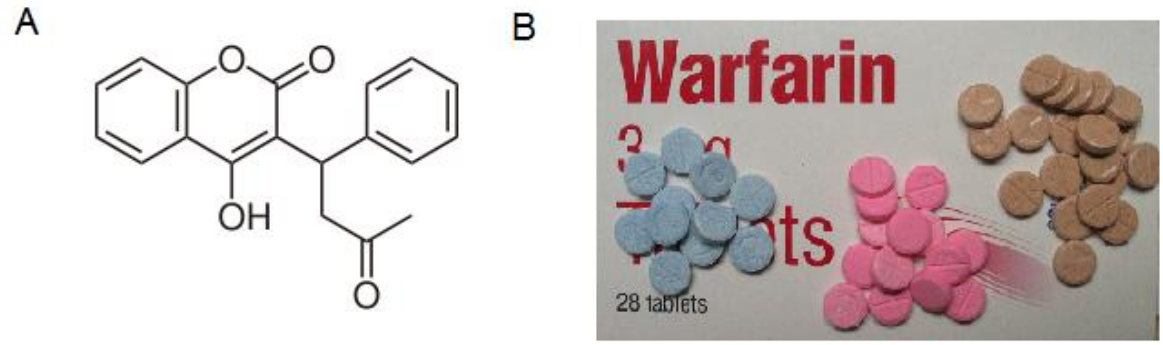


Figure 1.4. Warfarin (A) chemical structure; (B) tablets used in UK (Photo by Gonegonegone). Warfarin ($C_{19}H_{16}O_4$) have a molecular weight of 308.33 g/mol. The prescription warfarin contains mixtures of R- and S-4-hydroxy enantiomers. Different colours are used to distinguish different doses.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

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

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

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

1.8.2 Warfarin and coumarin derivatives

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

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

1.8.3 Side effects and adverse reactions

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

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

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

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

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

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

1.8.4 Drug interaction and effective therapeutic dose

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

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

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

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

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

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

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

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

1.9 THIS THESIS

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

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

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

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

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

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

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