UNIVERSITY OF Southampton

University of Southampton Research Repository
ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

http://eprints.soton.ac.uk

UNIVERSITY OF SOUTHAMPTON
FACULTY OF MEDICINE

Study to identify the associations between polymorphisms in pharmacogenetic loci, mycophenolic acid precursors (mofetil or sodium) and clinical outcomes in renal transplant recipients using array based exome SNP sequencing

Natalie Borman

25192671

Thesis for the degree of Doctor of Medicine
October 2014

ABSTRACT

Title: A Study to identify the association between polymorphisms in pharmacogenetic loci, mycophenolic acid precursors (mofetil/sodium) and clinical outcomes in renal transplant recipients using array based exome SNP genotyping

Introduction: Mycophenolic Acid precursors (MPAP) are widely used in transplantation. Adverse drug reactions are dose dependent and usually improve with reduction or cessation, but with increased rejection risk and poorer long-term graft survival. Individuals respond in different ways to immunosuppression and genetic variability accounts for 20-90%. Given these challenges, there is growing interest in the role of pharmacogenetics in individualising drug regimens.

Aim: To identify and investigate the association of SNP's with clinical response to MPAP in renal transplant recipients.

Methods: 287 RTR were studied for primary outcome measures of biopsy proven acute rejection (BPAR), leucopenia (wcc<3), anaemia (Hb<10), gastro-intestinal side effects (GISE), infection, dose reduction or cessation in the first year post transplantation. Secondary outcome measures of time to event were also analysed.

Array based exome SNP genotyping was carried out using Illumina Human exome Beadchip v1.1. Associations were sought between SNP's and primary outcome measures. Extensive clinical data was collected.

Quality control of the data was carried out prior to statistical analysis. PLINK genome association tool kit was used for statistical analysis to seek associations between genotypic data and primary outcomes and logistical regression for genotypic confounders. SPSSv21 was used for logistical regression of phenotypic confounders as well as time to event analysis and Cox regression.

Results: All participants received CNI (64.1% Ciclosporine A, 35.9% Tacrolimus) and MPAP (91.3% MMF 8.7% Myfortic) 60.9% male, 93.4% Caucasian, mean age at transplantation 47 years (range 17-79). Frequency of primary outcome events was between 10.8 % and 34.5%.

Significant associations were seen with novel SNPs in UGT1A9 and SLCO1B1 and the primary outcome measure of Leucopenia, anaemia and BPAR and were supported by time to event analysis. Further associations were seen when a Caucasian only subset were analysed. A number of genes with no known involvement in MPA metabolism have also been identified as potential candidates for future research.

Conclusion: This study has shown several SNP's in genes known to be associated with MPA metabolism or excretion to have a significant and important impact on clinical outcomes in the first year following renal transplantation. It has demonstrated the use of a fast and efficient genotyping technique which can be applied to future research in this field.

Presentations/Publications

American Society of Nephrology conference 2013:

Influences of single nucleotide polymorphisms (SNP) on mycophenolic acid tolerance and side effects in renal transplant recipients (RTR)
Natalie Borman[(1)], Anthony Marinaki [(2)] and G Venkat-Raman [(1)]

(1) Wessex renal and transplantation unit, Portsmouth Hospitals NHS Trust
(2) Purine Laboratory, Guy's and St Thomas NHS Trust

A single nucleotide polymorphism (SNP) significantly reduces mycophenolic acid (MPA) associated leucopenia in renal transplant recipients (RTR)
Natalie Borman[(1)], Anthony Marinaki [(2)] and G Venkat-Raman [(1)]

(1) Wessex renal and transplantation unit, Portsmouth Hospitals NHS Trust
(2) Purine Laboratory, Guy's and St Thomas NHS Trust

Southwest and East Kidney Society Conference 2013:

A single nucleotide polymorphism (SNP) in SLCO1B1 gene which encodes for OATP1B1 Transporter, significantly increases Biopsy proven acute rejection in renal transplant recipients (RTR) receiving mycophenolic acid (MPA).
Natalie Borman[(1)], Anthony Marinaki [(2)] and G Venkat-Raman [(1)]

(1) Wessex renal and transplantation unit, Portsmouth Hospitals NHS Trust
(2) Purine Laboratory, Guy's and St Thomas NHS Trust

British Transplant Society and Renal Association Conference 2013: (Poster Prize)

Mycophenolic acid withdrawal in the first 12 months post renal transplantation predicts worse graft outcome at 24 months
Natalie Borman [(1)], Scott Harris[(2)],and Dr G Venkat-Raman [(1)]

(1) Wessex renal and transplantation unit, Portsmouth Hospitals NHS Trust

(2) Medical Statistician, University of Southampton

Academic Thesis: Declaration of Authorship

I, NATALIE BORMAN
declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

A study to identify the association between polymorphisms in pharmacogenetic loci, mycophenolic acid precursors (mofetil/sodium) and clinical outcomes in renal transplant recipients using array based exome SNP genotyping.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;

2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

3. Where I have consulted the published work of others, this is always clearly attributed;

4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

5. I have acknowledged all main sources of help;

6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

7. Either none of this work has been published before submission, or parts of this work have been published as: [please list references below]:

Signed: …….

Date: ……

Acknowledgements

This thesis is dedicated the patients of the Wessex renal and transplantation unit, without their willingness to participate in research for the benefit of others, studies such as this could not be completed.

I would like to offer my sincere thanks to my Clinical Supervisor Dr Venkat-Raman for giving me this opportunity and for his immense support and belief in me through this process, without which I would not be where I am today. I would like to thank Dr Tony Marinaki for his support and guidance through this project and his expertise in the field. I would also like to extend my thanks to Lynette, Monica, Adele and Paul in the Purine research laboratory for their support and shared knowledge. Thank you also to Dr Cathryn Lewis and Jemma Walker for their statistical expertise which was essential to the success of this work.

I would like to thank Professor Chris Byrne for his academic supervision and support and Sara Campbell for her laboratory support and guidance.

Thank you to the Wessex renal research charity and Guy's and St Thomas's research Charity for providing the funding for this project.

I would like to thank my husband Ross who has stood by me throughout this process, without his patience, support and encouragement I would not have made it this far.

Finally I would like to thank my Parents for always believing in me and giving me so many opportunities in life to allowing me to full fill my potential, I hope in return I have made you proud.

List of abbreviations

AcMPAG	acyl-glucuronide
AGBL4	ATP/GTP binding protein-like 4
AHSG	Alpha-2-HS-glycoprotein gene
ALPK3	Alpha kinase-3
ARHGEF15	Rho GTPase activating protein 15
ARRDC4	Arrestin domain containing 4 gene
ASAP1	ArfGAP with SH3 domain, ankyrin repeat and PH domain 1 gene
AUC	Area under curve concentration
BCRP	breast cancer resistance protein
BF	Bonferoni
BKV	BK Virus
BOD1L	Biorientation of chromosomes in cell division 1-Like 1
BPAR	Biopsy Proven Acute Rejection
BRS	British renal society
CAD	Chronic allograft dysfunction
CDC42BPA	CDC42 binding protein kinase alpha
CES	Carboxylesterases
CI	Confidence interval
CMV	Cytomegalovirus
CNI	Calcinurine inhibitor
CYA	Cyclosporine A
DAB2	Dab, mitogen-responsive phosphoprotein, homolog 2
DBD	Deceased brain stem death donor
DCD	Deceased cardiac death donor
DEAH	DEAH (Asp-Glu-Ala-His) box polypeptide 16
DKKL1	Dickkopf-like 1
DPCR1	Diffuse panbronchiolitis critical region 1
EC-MPS	Enteric coated- Mycophenolic sodium
EDTA	Ethylenediamine tetra-acetic acid
eGFR	Estimated glomerular filtration rate
ESRF	End stage renal failure
FDR	False discovery rate
GI	gastrointestinal
GISE	Gastrointestinal Side effects
GMP	Guanosine monophosphate
GSTS	Guys and St Thomas
GWAS	Genome-wide association studies
Hb	Haemoglobin
HCP5	HLA complex P5
HD	Heamodialysis
HLA	Human leukocyte antigen
HLA MM	Human leukocyte antigen miss match
HSV	Herpes simplex virus
HWE	Hardy Weinberg equilibrium
IBD	Identical By Descent
IMP	inosine-5'-monophosphate
IMPDH	Inosine monophosphate dehydrogenase

KM	Kaplan Meier
KRUK	Kidney research UK
LD	Linkage Disequilibrium
LKDT	Live Kidney Donor
MAF	Minor Allele Frequency
MMF	Mycophenolate Mofetil
MPA	mycophenolic acid
MPAG	Hydroxyphenyl-β-glucuronic acid
MPAP	Mycophenolic Acid precursors
MPS	Mycophenolate Sodium
MRP	Multidrug resistance proteins
MUC22	Mucin 22
NEU1	Sialidase 1
NUP153	Nuclear pore complex protein 153
OAT	Organic anion transporter polypeptides
PC	Principle components
PCR	Polymerase chain reaction
PD	Peritoneal dialysis
PDHX	Pyruvate dehydrogenase complex
PKD	Polycystic kidney disease
pmp	Per million population
PRA	Panel reactive antibodies
PSORS1C1	Psoriasis susceptibility 1 candidate 1
QC	Quality control
RCT	Randomised controlled trial
rtPCR	Real time Polymerase chain reaction
RTR	Renal Transplant Recipient
RVD	Renovascular disease
SETD5	SET domain containing 5
SFTA2	Surfactant associated protein 2
SNP	Single Nucleotide Polymorphism
TAC	Tacrolimus
TMEM108	Transmembrane protein 108
TMEM70	Transmembrane protein 70
UGT's	Uridine diphospate-glucuronosyltransferases
VCAM-1	Vascular cell adhesion molecule 1
VSV	Varicella zoster virus
VWF	Von Willebrand factor
WCC	White cell count
WRTU	Wessex Renal and Transplantation unit
XMP	xanthosine-5’-monophosphate
ZNF605	Zinc finger protein 605

Contents

List of Tables

List of figures

1 Chapter 1: Introduction and Background

1.1 Introduction

Mycophenolate Mofetil (MMF) has been widely used in transplantation since its introduction in the mid-1990's, and has contributed significantly to excellent 1 year graft survival (2, 3). The efficacy of MMF in the prevention of acute rejection in transplantation has been shown to be superior to its predecessor, azathioprine (Aza) in clinical practice (4), with acute rejection rates of up to 46% in 1990 reducing to 23% in 2004 (5). Despite this, many patients suffer intolerable side effects to MMF therapy, particularly gastro-intestinal upset and leucopenia, while a smaller number of patients suffer graft rejection episodes despite a combination of immunosuppressive agents (6). Enteric coated mycophenolate sodium (MPS) is a more recently introduced alternative to MMF, with similar immunosuppressive properties, but fewer gastrointestinal side-effects (3), which may lead to fewer dose reductions or discontinuations (7). Adverse drug reactions related to these two mycophenolic acid precursors (MPAP) are dose dependent, with dose reduction or withdrawal of MPA usually reversing the side effects, but with increased risk of rejection and poorer long-term graft survival (8, 9). A 15 year follow up study of Aza versus MMF found that 42% of patients had switched to Aza from MPA due to intolerance (10). Although 1 year graft survival now exceeds 90%, long term graft survival has not improved as much as anticipated, despite the introduction of newer more effective drugs. There has been a small improvement in the graft loss per year after the first year of transplant (Tx) from 4-6% for those transplanted in 1989, to 2-4% for those transplanted in 2005 (5). Graft loss remains one of the commonest reasons for starting dialysis. The transplantation of increasingly older and higher risk individuals, and the use of extended criteria donors to address the shortage of organs (11), goes some way to account for this, but it is increasingly recognised that early graft factors such as rejection dramatically influence long term outcome (4), and that achieving therapeutic levels of immunosuppressive drugs in the immediate transplant period is key.

There is also concern regarding the high incidence of cardiovascular complications and malignancy in renal transplant recipients (RTR), with cardiovascular disease accounting for around 30% of those deaths with a functioning graft, after the first year (11, 12). Chronic allograft dysfunction (CAD) remains a significant issue and emerging problems such as polyomavirus associated nephropathy have further added to the problem (11, 12). These are all, in part, felt to be related to the use of immunosuppression with either suboptimal or excessive exposure.

Different transplant recipients respond in different ways to immunosuppression and achieving optimal immunosuppression remains a challenge. Therapeutic drug monitoring is widely used to

help achieve this and the majority of individuals achieve steady state by 2 weeks post- transplant (13). Unfortunately there is no definitive way of measuring the immunological status of the patient (12). Studies have shown that adequate exposure to MPA during the first week is an important factor (11, 14) but in reality, it is not usually known if this is achieved. Many factors influence the individual's response to drug therapy and fully understanding the origin of this variation has proven difficult, but genetic variability is thought to account for between 20% to 90% of the variability in a drugs disposition and its effects (12, 15). Given these challenges, there is growing interest in the role of pharmacogenetics and it's potential to aid in individualising drug regimens and a move towards 'personalized medicine'. The hope is to provide the physicians with tools to prescribe the right drug, at the right dose to achieve maximal therapeutic impact with minimal adverse effects.

1.2 Background

1.2.1 Pharmacogenomics and Pharmacogenetics

The way in which drugs are absorbed, distributed and eliminated within the body is affected by their bioavailability and pharmacokinetics, but genetic variability in enzymatic action involved in drug metabolism is a key factor in determining the therapeutic levels achieved by drugs, drug interactions and side effects (16, 17). This forms the basis of pharmacogenetic research.

Pharmacogenetics refers to the study of genetic differences, which affect an individual's response to drugs and the use of genetic information about patients to allow individualization of drug therapy (18). Pharmacogenomics uses the impact of the whole genome across groups of individuals to illustrate inter-individual variation in response to drugs on the basis of inherited differences (17, 19). These terms are often used interchangeably in the field of research. Broadly they involve the discovery of markers that can predict the outcome of drug therapies before they are administered.

The science of pharmacogenetics is not a new phenomenon, with the earliest publications of inherited differences in drug effects appearing in the late 1950's largely driven by Werner Kalow (19, 20).

Pharmacogenomics aims to use genetic information to tailor drug regimens, improve response and reduce adverse drug reactions (ADR) (21). Medications that have been proven efficacious in large rigorous trials still fail to produce an adequate response in some individuals with others suffering intolerable side effects (20). Adverse drug reactions are a significant cause of morbidity, hospitalisation and in extreme cases mortality and they are thought to be the cause of around 7% of UK hospital admission (18). These obviously have a negative impact on quality of life and

potentially reduce adherence (22). Many drugs work within a therapeutic window and individual variations in the absorption or metabolism of drugs can result in some patients exceeding this therapeutic window whilst others fall short of it. Individual variation in drug handling can significantly alter the dose response curve and clinical outcome will be affected if the dosing is not altered appropriately (20).

Many factors which influence drug response will alter throughout the individuals lifetime such as weight, organ function, body composition, concomitant drug therapy, drug interactions and the nature of the underlying disease (19, 23) but genetic factors remain fixed and the consequent drug effects are likely to be permanent (20).

The use of genetic profiling increases the information that is available about the individual (24) and can help guide the dose and frequency of medication and in some cases warrant the use of alternative drugs. The aim is to improve both safety and efficacy of therapy (25).

1.2.2 Single nucleotide polymorphisms (SNP)

The human genome consists of 23 pairs of chromosomes, consisting of 6 billion base pairs. The human genome has now been successfully mapped and sequenced, vastly increasing our knowledge of genetic variability.

The development of the international Hapmap project produced publicly available genome-wide database of human genetic variation to aid the development of genome wide association studies (GWAS) (26).

A single nucleotide polymorphism (SNP) is a variation in a single base present in DNA, and is the most common genetic variation in the human genome (27, 28).

More than 90% of human genes contain at least one SNP (20). Fourteen million SNPs have now been identified following the human genome project, with 60,000 in the coding regions (19, 20). It is estimated that there may be a variation of 2.5 million SNPs between any two individuals (18).

The presence of SNPs is thought to have a significant impact on inter-individual variability of drug response and the proximity of the SNP to the coding regions will determine if that gene or protein will function normally (29). The structure or target of the protein may be altered, the metabolizing enzymes or drug transporters may change which will subsequently affect absorption, distribution, metabolism and elimination of the drugs (20). The effect may be enhanced or reduced by the SNP, with corresponding differing effects on the overall impact of the drug. Detecting SNPs which alter

gene or protein function is likely to have the greatest impact in the move towards individualized dosing (29).

Each individual has two copies of each gene and SNPs can be inherited in one copy (heterozygous) or both copies (homozygous) (29). The effect can also be recessive, requiring the SNP to be present in both genes or dominant, when the SNP only needs to be present in one gene copy in order to exert its effect.

It is also now understood that individuals may need to inherit a combination of SNPs in a haplotype (sometimes in more than one gene) to have an effect. Studies combining the impact of more than one gene are few in numbers (30).

With the extensive data available from GWAS using large, highly phenotyped cohorts, as well as Hapmap and the 1000 genomes projects, it is now possible to identify many of the SNP variants (21, 26, 28). It is therefore not surprising that the study of SNPs forms the basis of the majority of pharmacogenetic research.

1.2.3 Advances in pharmacogenomics research

Until fairly recently, discovering pharmacogenetic markers involved a detailed knowledge of how specific drugs are metabolized in the body and then selecting "candidate" genes associated with this metabolism. The sequence of these "candidate" genes is then compared between individuals who have poor responses to therapies and those who have the desired responses to see if genetic variation is linked to drug response. This candidate gene approach has the advantage of focussing resources, but is painstakingly slow and limited by the need to understand the metabolic pathway of the drug (19), as well as knowledge of the specific single nucleotide polymorphisms (SNP).

Originally, studies only focussed on drug metabolizing enzymes, but as knowledge has increased, studies have expanded to include drug transporters, drug targets (18, 19) and other proteins which have been shown to alter drug response. Many studies are now adopting a broad panel approach allowing the identification of SNP in genes that were not previously known to be involved in drug metabolism.

Advances in technology, including whole genome SNP genotyping using microarrays (26, 31), and more recently whole exome sequencing (32), have facilitated major advances in the field of pharmacogenetics (and the genetics of disease in general).

These next generation or 'massive parallel' sequencing platforms are now widely available. These high-throughput platforms provide powerful molecular methods for the simultaneous probing of

genes across the whole genome(23), and are particularly efficient at genotyping SNPs, making them ideal tools for pharmacogenomic studies, allowing thousands of SNPs to be simultaneously studied (26). In particular, it has made it unnecessary to understand the specific metabolic properties of a drug, and isolate candidate genes. Instead the technology tests variations in DNA right across the human genome to determine if any specific variation can be linked to an altered response to drug therapy.

Whilst readily available, sequencing the whole genome generates enormous amounts of data making interpretation a challenge. The protein coding region of the DNA accounts for around 1% of the total and is composed of regions termed the exons. It is thought that around 85% of clinically significant genetic mutations occur within the protein-encoding exons of the human genome and hence selectively sequencing these coding regions would seem an easier and cheaper alternative to sequencing the whole genome, although still remains an expensive option for large pharmacogenomics projects (32-34). Technological advances in the development of the next generation sequencers means they are now both readily available and becoming cheaper (31). With such rapid advances in the field it is likely that pharmacogenetics will soon become a standard tool in clinical practice and drug development (18).

With the advances in technology new methods are still evolving. This includes the development of array based exome SNP genotyping. This is the basis for the Human exome beadchip developed by illumina ® which provides extensive coverage of >240,000 functional exonic variants selected from over 1200 individuals across diverse populations(1). The beadchip is specifically designed to provide coverage of both common and rare variants and provides rapid SNP genotyping at a fraction of the price of exome sequencing making it a useful research tool for large studies.

With the production of such extensive quantities of data, rapid development of bioinformatics tools has taken place in parallel. This intricate coupling of computer software and statistical programmes for the handling, comparison and assembly of this sequencing data has been key to further progression in this field (28, 35).

1.2.4 Pharmacogenomics and organ transplantation

A combined immunosuppressive approach is common practice following organ transplantation to prevent recognition of the donor organ as ‘foreign’ by the recipient’s immune system and resultant organ rejection.

Prescribing of immunosuppressive drugs is currently done in a fairly rigid and stereotyped manner with most drugs being given at either a fixed dose or according to the patient’s weight, following

protocols based on the latest evidence. Decision on the immunosuppression regimen does take into account predetermined factors relating to 'immunological risk' including genetic testing to determine how well matched the donated organ and the recipient are (HLA mismatch)(12), as well as the number of previous transplants the patient has received and the recipient's underlying medical condition.

Different patients respond in different ways to immunosuppression with some suffering from acute rejection episodes, whilst others suffer side effects either directly related to the drug or due to over immunosuppression. The variable pharmacokinetics and narrow therapeutic window of immunosuppressive drugs (13, 36) make it difficult to achieve optimal immunosuppression in many individuals. The general toxicity profile of immunosuppressive drugs as well as drug specific side effects are often intolerable for the patient (13, 17).

Therapeutic drug monitoring (TDM) is common practice for some immunosuppressive agents like calcineurine inhibitors (CNI's, namely ciclosporine and tacrolimus) and mTOR inhibitors (sirolimus). There is good clinical evidence for a correlation between blood concentrations and therapeutic / toxic effects. TDM is available for MPA but it is not used in routine practice for reasons which will be discussed later in this chapter. The main limitation of TDM is that it is not of use until the drug has been administered and reaches a steady state which is usually 72 hours after administration (37). In this time there is the chance of under immunosuppression potentially exposing the individual to acute rejection during this critical period. The option of pre-testing the drug to establish the correct dose prior to transplantation is flawed due to the underlying organ dysfunction (38). This is particularly the case in liver or renal failure as the organs play a vital role in drug metabolism and excretion and hence the required dose will alter significantly following correction of the organ failure by transplantation.

It is also well recognised that rejection or toxicity can still occur even when the immunosuppression falls within the acceptable "therapeutic range", which may reflect abnormal binding of the drug to its target or abnormal intracellular responses (39).

The potential use of pharmacogenetics in the world of transplantation is becoming increasingly acknowledged. The narrow therapeutic index of immunosuppressant drugs means the ability to predetermine the genotype of patients and individualise immunosuppressive regimens may be a key breakthrough. Pharmacogenetics may aid in the choice of drug, initial drug dosing, reaching therapeutic levels rapidly, and reducing both rejection and adverse drug effects which produce important morbidity and in some cases mortality (12, 39-41). It may prove particularly useful when TDM is not routinely undertaken, such as in MPA treatment, when pharmacogenetics may

have a strong influence on decision making (13). Although many factors influence drug response, individual genetic determinants will remain stable throughout the person's life time(19).

Organ transplantation is a planned procedure and recipients already undergo extensive work up including routine genetic tissue typing. It would therefore seem reasonable to incorporate pharmacogenetic testing into this routine. Immunosuppression regimens can then take into account this information to increase the chance of getting the drug and dose correct (40) and limiting short and long term complications related to the immunosuppression.

Several pharmacogenomics studies to date have looked at azathioprine, CNI's, mTOR's and MPA in relation to solid organ transplantation with some interesting results. The studies will not be discussed in detail with the exception of MPA which forms the basis of this research.

1.2.5 Pharmacogenetics in clinical practice

The translation of pharmacogenetic findings into clinical practice is challenging (23).

There have been several areas of medicine where pharmacogenomics have been successfully incorporated into practice. The clinical application in transplantation has been hampered by several factors including research with small study populations, lack of consistency between studies (13) and paucity of positive replication. The generation of vast quantities of information about individuals often makes it difficult to determine what is important and translate this into clinical practice. The mathematical models which are produced generally predict average or population outcome whereas in clinical practice, the focus is on the individual patients (23).

Barriers also exist to the routine application of genetic profiling into clinical practice with patients' concerns about what might be found when sequencing the genome, and doctors' concern that adding genetics into the picture will lead to further complexity in the prescribing(18).

The rapid advances in the field of human genomics has led to an increased understanding of the genetics of disease and drug interactions (20). Pharmacogenetics research and clinical practice have grown in response to this knowledge and technological advances (22), which are becoming economically viable and making individualized dosing more of a reality (40). There is now a publicly available web-based resource (42) which aims to assist future research and clinical application by enhancing the understanding of pharmacogenomics and drug metabolic pathways. The hope is that pooling information from various studies into a single resource will support the implementation of personalized medicine in the future (42, 43). The field of pharmacogenetics has been referred to as a 'revolution that is occurring medicine'(21).

1.3 Metabolism of MMF/MPS

Even with whole genome approaches to genotyping, an understanding of drug metabolic pathways is fundamental before considering pharmacogenomic studies as it will provide a starting point for analysis.

Figure 1.1 shows a schematic representation of the current knowledge of mycophenolic acid metabolism including the known genes involved. Copyright for this figure belongs to PharmGKB, and permission has been given for use of this figure by PharmGKB and Stanford University (42).

Key for figure 1-1 ABCB1 = ATP-binding cassette subfamily B member 1. ABCC2 ATP-binding cassette subfamily C member 2. Ac-MPAG = Acyl glucuronide MPA. CES1 = Carboxylesterase 1. CES2 = Carboxylesterase 2. DM-MPA = 6-0-desmethyl-MPA. GI- Gastrointestinal. GMP= Guanosine monophosphate. IMP = inosinemonophosphate. IMPDH1 = inosinemonophosphate dehydrogenase 1. IMPDH2 = inosinemonophosphate dehydrogenase 2. MMF= Mycophenolate mofetil. MPA= Mycophenolic acid. MPAG= MPA-7-0-glucuroide. SLCO1B1 = Solute carrier organic anion transporter family member 1B1. SLCO1B3 = Solute carrier organic anion transporter family member 1B3. XMP= Xanthine monophosphate.

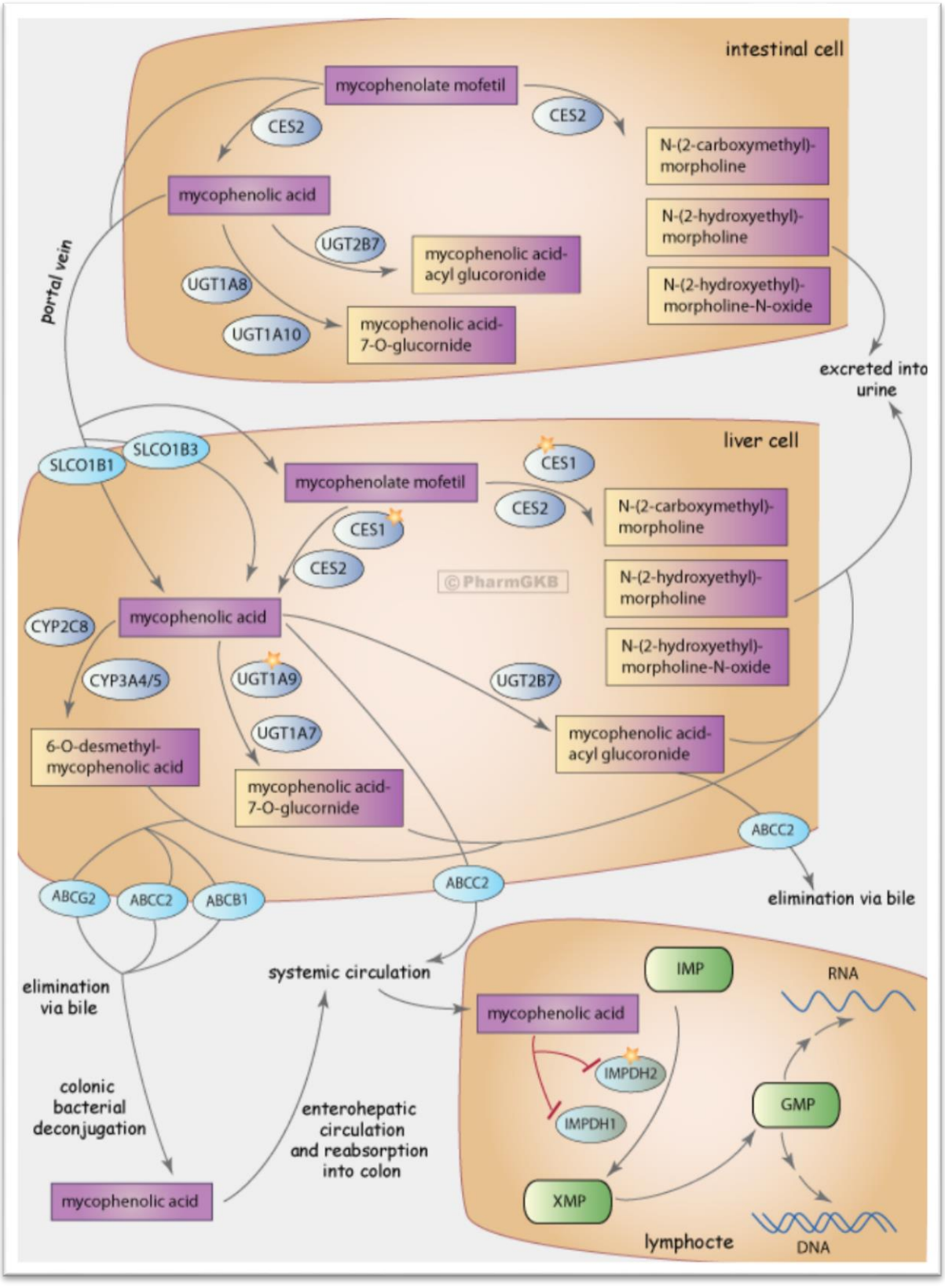


Figure 1-1: Schematic representation of mycopneolic acid metabolism (42)

This figure shows the metabolic pathway of Mycophenolic acid and the known genes involved in the pathway. A full detailed explanation of this pathway is given on the following pages.

The immunosuppressive drugs MMF and MPS are the prodrugs of mycophenolic acid (MPA), a selective transition state analogue inhibitor of inosine monophosphate dehydrogenase (IMPDH). These prodrugs are rapidly and extensively hydrolysed to active MPA within the intestine by tissue and plasma esterase's (15, 44). MPA is predominantly protein bound in the plasma. Only free MPA (<0.1% of the total) is pharmacologically active (45).

Carboxylesterases (CES1 and CES2) play a major role in the hydrolysis of MMF to MPA, N-(2-carboxymethyl)- morpholine, N-(2-hydroxyethyl)-morpholine, and the N-oxide of N-(2-hydroxyethyl)-morpholine (42). CES1 and CES2 are found in the liver but only CES2 is abundant in the intestine. Hydrolysis initially takes place in the intestine with any MMF that enters the portal vein being hydrolysed in the liver (42, 46).

MPA by inhibiting IMPDH, has a powerful effect in inhibiting or blocking immune responses. The two IMPDH enzymes Type I and Type II catalyse the conversion of inosine-5'-monophosphate (IMP) to xanthosine-5'-monophosphate (XMP), the rate limiting step in the biosynthesis of guanosine nucleotides including guanosine monophosphate (GMP) (46, 47). The enzymes share 85% homology at the amino acid level and are similar kinetically (48). IMPDH I is expressed at low levels in most cell types. In contrast, expression of IMPDH II is up regulated in proliferating cells. T- and B- lymphocytes are critically dependent for their proliferation on the de novo synthesis of purines, whereas other cell types are able to utilise salvage pathways (49, 50). MPA thus selectively inhibits proliferation of these cells by restricting guanine nucleotide pools (dGTP and GTP) necessary for DNA and RNA synthesis in rapidly dividing cells (2). IMPDHI and II are therefore candidate genes to explain variability in clinical response to therapy. The expression of several adhesion receptors including vascular cell adhesion molecule 1 (VCAM-1), is also suppressed by the depletion of guanosine nucleotides, interfering with leucocyte attachment to endothelial cells and prevention of lymphocyte recruitment.(2) Thus MPA has pleotropic effects on the immune response.

The metabolism of MPA is primarily by glucuronidation of the phenolic hydroxyl group by uridine diphosphate-glucuronosyltransferases (UGTs) to an inactive glucuronide metabolite (MPAG) - known as hydroxyphenyl-β-glucuronic acid ,(15) - and to acyl-glucuronide (AcMPAG) which is pharmacologically active (51).

Glucuronidation to the water soluble and inactive form MPAG, mainly occurs in the liver, predominantly by UGT1A9 and to a lesser extent in the gut by UGT1A8. UGT2B7 is the predominant isoform responsible for the formation of AcMPAG which occurs in both the intestine and hepatic cells (39, 49).

MPAG is taken up into hepatocytes via the portal vein by OATP1B1 and OAT1B3 transporters (encoded for by the SLCO gene). OATP1B1 is a hepatic transporter expressed at the basolateral membrane of hepatocytes and represents a crucial step in the hepatic clearance of several drugs, including MPA, from the circulation (8).

The metabolite 6-0-desmethyl-MPA is formed by the hepatic CYP450 enzymes CYP3A4, 3A5 and 2C8, this forms a very minor fraction of MPA (42)

MPAG is then excreted from the hepatic cells into the bile by MRP2 encoded by ABCC2 gene and to a lesser extent breast cancer resistance protein (BCRP) (44, 49) encoded by *ABCG2,* and P-glycoprotein encoded by *ABCB1* (42).

MPA and MPAG are subject to extensive enterohepatic recirculation. MPAG is de-glucuronidated back to MPA by gut bacteria, which is absorbed in the colon. Several transporter mechanisms, including organic anion transporter polypeptides (OAT), multidrug resistance proteins (MRP) and UGT's are all involved in MPA/MPAG biliary excretion and reuptake (12).

The principal MPA elimination mechanism is renal excretion as MPAG, in part by Multidrug resistance protein 2 (MRP2) in the proximal tubules (44, 52). AcMPAG constitutes only around 5% of the total metabolic elimination pathway and is also excreted via the kidneys. It has been shown to bind to proteins and macromolecules, this is thought to be one explanation for the myelotoxicity and the resultant leucopenia associated with this drug (53).

Given the narrow therapeutic window of MPA, the resulting high concentrations of MPA in the gut are probably responsible for the damage to the intestinal epithelium and the consequent gastrointestinal (GI) side effects such as nausea and vomiting, diarrhoea and GI tract bleeding. The glucuronosyltransferase UGT1A1 family of enzymes are very similar in structure to the GI expressed UGT1A8 and UGT1A10 and the liver expressed UGT1A9 (15) and must be considered to be important pharmacogenetic candidates to explain adverse drug reactions to MPA. The reliance of MPA metabolism and excretion on a number of transporters including OAT's MRP-2 and BCRP, means that these must also be considered when exploring pharmacogenetic candidates for this drug.

It is also important to consider an extended candidate approach including looking at other genes involved in the *de novo* synthesis of purine nucleotides. It is unknown if these have any involvement in the metabolic pathway of MPA but it will form part of the analysis in this study.

Whilst understanding the metabolic pathways provides some understanding of MPA pharmacokinetics, side effects and metabolism it is unlikely to be the full explanation for the large variability seen between individuals. It is likely that the pharmacogenetic influences on MPA metabolism are far more complex than those that can be predicted from the current understanding of metabolic pathways and likely to involve intricate interactions in many other genes, and for these reasons we need to go beyond a candidate gene approach.

1.4 MPA therapeutic drug monitoring

Whilst therapeutic drug monitoring is routinely used for calcineurine inhibitors (CNI's), which are widely used in combination with MPA, this is not performed in routine practice for MPA as single MPA plasma concentrations do not correlate with MPA area under the curve concentration (AUC). Studies to date have shown a high degree of inter and intra-individual variability in MPA pharmacokinetics in transplant recipients (54). Whole AUC measurements are required for accurate MPA monitoring and hence it would be challenging to routinely measure in clinical practice. Although some argue it has a place in high risk renal transplant recipients (8, 45) a 2006 systematic review found a lack of correlation between MPA plasma concentrations and adverse or therapeutic effects in the majority of studies(55). Since then two randomised trials have set out to study the value of MPA monitoring in renal transplantation. Both studies contained a standard group and an intensive monitoring group with MPA alteration based on TDM (56). The multicentre French trial (APOMYGRE) suggested a significant benefit in terms of treatment failure and acute rejection (57), but a larger study (FDCC) failed to show any benefit (58) and therefore there is no strong evidence to suggest that MPA TDM currently has a place in routine post-transplant care. A group in 2009 conducted a randomised crossover study to assess potential differences in pre-dose concentrations of MMF and EC-MPS, but again found that trough levels were of no clinical benefit (59). There is currently no clear evidence to support MPA monitoring in renal transplantation, and given the high cost involved, it is not routinely undertaken.

With MPA acting by potent and selective IMPDH inhibition, thereby limiting the *de novo* pathway for guanine nucleotide production, there has been some interest in the potential of measuring GTP and dGTP in MPA treated patients (60, 61). Intracellular GTP levels are reduced by MPA and moreover, addition of GTP to MPA treated cells in vitro has been shown to reverse MPA effects (60). In vivo studies have shown a less dramatic effect, suggesting that salvage pathways may prevent massive depletion of the GTP pool in patients chronically treated with MPA (61). The situation is of interest in renal disease as red cell GTP levels have been found to be elevated in chronic renal disease, with normalization of levels following successful transplantation and rising during rejection episodes, although these findings were reported prior to the introduction of MPA

(62). A 2004 study looked at a comparison between GTP levels in renal transplant patients treated with MPA versus azathioprine and found significantly higher erythrocyte GTP levels (63) and sustained lower GTP concentration in mononuclear leucocytes (64) in the MPA treated patients, although the studies were limited by significantly poorer graft function in the MPA group, making interpretation difficult. Although the literature is sparse, there is still interest in the potential of GTP as a marker of MPA compliance, a surrogate marker of MPA levels or indeed as a marker of impending graft loss.

1.5 Pharmacogenomics and MPA: Review of the literature

Several studies to date, have attempted to look for associations between SNPs and inter-individual variability in response to MPA. The vast majority have focused on a candidate gene approach limited to small numbers of SNPs in one or two candidate genes which have previously been found to be associated with MPA metabolism, absorption and excretion or those involved in the immunomodulatory effects of MPA. Details of the studies conducted to date are tabulated and summarised in Table 1 and will be discussed.

The most extensively studied genes are the UGT family with a number of studies looking at UGT1A8, UGT1A9 and UGT2B7.

UGT1A9 -275A>T and -2152T>C SNPs were found to be associated with lower MPA levels in RTR's taking 2g MMF per day in a 2005 study of 95 patients by Kuypers et al (39, 51). Similar results were reported in Levesque et al's 2007 study of 52 healthy volunteers (39, 53). van Schaik et al's 2009 study of 338 RTR found these SNPs were significantly associated with biopsy-proven acute rejection (BPAR) (65). Johnson et al's 2008 study only found the association to exist in those individuals taking tacrolimus and not in those taking ciclosporine (66). These results have not been consistently proven with Sanchez-Fuictuoso et el finding no difference in MPA area under the curve with -2152T>C or -275A>T but greater incidence and severity of GI side effects (16). Similarly Baldelli et al's 2007 study of 40 RTR's reported no effect of these two SNPs on MPA levels, but they identified higher MPA levels with C-440T and T-331C SNPs (52). Prousa et al's study of paediatric RTR's found -331T>C to be significantly associated with an increased risk of developing adverse effects, though the study contained just 38 participants (67).

UGT1A8 SNPs -999C>T, 255A>G and 277G>A were found to be associated with more infective episodes in Brazillian RTR's in a 2008 study (68). Johnson et al found UGT1A8*2 (173G>A) to be an important predictor of MPA dose corrected trough concentrations, but this association was not found in Kagaya et al 2009 study of Japanese RTR's (69), although the different ethnicity may account for this.

UGT2B7 C802T variant appeared to be protective against GI side effects in Yang et al's study of 67 RTR's (70), but this has not been reproduced in any similar studies to date. Kagaya et al (69) and Van Agteren et al (71) studied UGT2B7 SNPs but found no association with MPA levels, leucopenia or diarrhoea.

The most recent study looked at 32 paediatric RTR and found that a combination of UGT1A9-440C>T, UGT2B7-900A>G and MRP2-24T>C polymorphisms were important predictors of interindividual variability in MPA exposure (72).

Whilst the results of the studies to date do not give us conclusive or reproducible evidence that polymorphisms in the UGT family have clinical relevance in RTR, they do appear to have an impact on MPA metabolism with some evidence for an impact on side effects and rejection. It is not yet clear from these studies how MPA dosing could be individualised to provide clinical benefit. The UGT family remain of interest and require further investigation and validation.

IMPDH1 and 2 were first studied in relation to MPA and transplantation in 2008 by Wang et al who looked at numerous SNPs in 191 RTR. They found a significant increase in BPAR in individuals with variations in rs2278293 and rs2278294 SNPs. There was no significant association of leucopenia with any of the allelic variations (48). A study in 2010 of healthy volunteers found that MPA had significantly less antiproliferative effect on lymphocytes in individuals with IMPDH1 rs11706052 SNP (50). This would potentially have important implications for rejection and required further clinical investigation in the transplant population. Kagaya (73) went on to find no association between IMPDH1, the SNPs rs2278293 or rs2278294 and acute rejection. They found some association between rs2278293 and MPA pharmacokinetics but this was not statistically significant nor did it translate into increased rejection rates (73). Gensburger et al's study of 456 RTR's found no association between IMPDH1 rs4974081 or rs11706052 and side effects or rejection. The IMPDH1 rs2278294 variant was significantly associated with a lower risk of BPAR and increased risk of leucopenia (74). Sombogoard et al looked at IMPDH2 rs11706052 in 101 RTR and found the 12 hour MPA concentrations to be elevated in the 3757C>C group compared to the control group (75).

These studies do suggest that IMPDH1 may be important in MPA metabolism and immunomodulation. They also suggest that genetic polymorphisms may have important clinical effects. They warrant further research as the results will need to be reproducible before they are applicable in clinical practice.

CES2 was studied in 80 Japanese RTR by Fujiyama et al in 2008 in relation to MPA pharmacokinetics. Three different SNPs were studied but none were found to be associated with inter-individual variations in MPA concentrations (76).

Naesens et al 2006 studied SNPs in MRP2 in 95 RTR looking at the association with MPA pharmacokinetic and laboratory data. They found MRP2 24C>T allelic variation to be associated with higher trough levels, a lower oral clearance of MPA and an increase in reported diarrhoeal episodes. The 3972C>T allelic variation was associated with higher MPA dose and the 4544G>A was associated with greater MPA exposure at all time points. There was no association with any of the allelic variations and infection episodes or white cell count (WCC) (77). Yang et al also found individuals with MRP2 24C>T had an increased tendency towards GI side effects but it did not reach statistical significance (70)

Michelon et al 2010 was the first study to look at multiple SNPs in several different genes; they studied the effects of 14 SNPs across 7 different genes involved in various stages of MPA metabolism and excretion. The study of 239 RTR looked at the adverse effects of MPA and BPAR. Adverse effects were defined as first occurrence of leucopenia (Total WCC <4), anaemia (Hb <120g/L), thrombocytopenia <150x10^9/L, diarrhoea (greater than two episodes per day), nausea, vomiting or infections, requiring temporary or permanent MPA dose reduction or interruption. The only allelic variant found to be associated with adverse effects was SLCO1B1 521T>C with a significant increase in adverse effects in those carrying wild type 521T compared to 521C allelic variant. They found no association with BPAR or adverse effects with any of the other SNPs. The study lacked statistical power to detect differences in other SNPs with lower frequency such as SLCO1B1 11187G>A. It is the first study which attempted to look at a number of different genes in MPA pathway, which is fundamental to this work (8).

The potential importance of SLCO genes and the importance of multiple gene analysis was further highlighted in Miura et al's study looking at SLCO1B1, 1B3 and 2B1 genes as well as ABCC2 (MRP2) association with pharmacokinetics and clinical factors including diarrhoea, nausea, vomiting or abdominal pain in 87 Japanese RTR. SLCO1B3 334T>G and 699G>A variants were found to be in complete linkage disequilibrium with each other. The 334 TT and GG genotypes showed no significant difference in MPA levels, but the GG variant showed significantly increased MPA AUC at 6-12 hours. There was no association with diarrhoea. The SLCO2B1 and ABCC2 allelic variations were not associated with significant variations in MPA pharmacokinetics. However individuals with both SLCO1B3 334 TT and ABCC2 24 TT genotypes had significantly lower MPA concentrations. They concluded that ABCC2 C24T may be co-associated with SLCO1B3 T334G for

the clearance of oral MPA (78). A further study published in 2012 looked at 4 SNPs in ABCB1 gene, 2 in SLCO1B1 and 2 in SLCO1B3 in 338 RTR and did not find significant association with any of the SNPs and dose adjusted exposure to MPA, MPAG or AcMPA, nor with the incidence of diarrhoea or leucopenia (6).

Both SLCO and MRP2 seem to play a role in MPA levels and this appears to have clinical relevance as outlined in the studies above. Both genes require further pharmacogenetic investigation.

The major limitation of all of these studies is the use of a candidate gene approach. The majority of studies are underpowered and the effect of multiple SNPs in several genes has not been comprehensively studied. These studies do not seek to look beyond what is known about MMF metabolism and transport and hence, are setting out to prove assumed associations and not to discover the new or unknown.

Another apparent limitation when reviewing the available literature thus far, is that the studies are not consistent, although some have looked at the same genes they have chosen different end points and others have looked at different SNPs within these genes. While this has provided an expanding knowledge base, there has been minimal replication which is essential before clinical application can be considered.

With the rapidly expanding field of pharmacogenomics and the development of new genetic techniques, we now have the ability to go far beyond what has been achieved in the studies so far. While the improvements in methodology will enhance the future understanding of pharmacogenomics and bring more prospects for the development of individualized dosing, it remains important to understand what is known to date and these studies form the basis of our current knowledge. They also provide important foundations on which to build for the future.

The first study that has adopted these newer methods in relation to MPA and renal transplantation was by Jacobson et al, in a 2011 study that used a broad panel SNP chip looking at 2724 SNPs in 978 RTR's. The study looked for association in multiple SNPs (many of which had previously been found to be associated with drug metabolism, absorption and excretion or immunomodulation) and mycophenolate related anaemia or leucopenia. Anaemia occurred in 15% (Hb<10g/dL or haematocrit <30%) of individuals within 6 months while leucopenia occurred in 22.9% (wcc <3000 cells/mm3) (79).

They found an increased hazard ratio (HR) for anaemia with Interleukin (IL)12A (rs568408). The presence of one A allele conferred an HR of 1.98 (95% CI 1.39-2.82) and 2 A alleles an HR of 3.93 (1.95-7.95) compared to non-carriers. CYP2C8 SNP (rs11572076) also increased the HR for

anaemia by 3.24 (1.7-6.2). The presence of HUS1 SNP (rs2037483) reduced the HR for anaemia 0.54 (0.39-0.74), these results took into account a 20% false detection rate (79).

The study also found a number of SNPs to be associated with the development of leucopenia (wcc<3000 cells/mm^3) but none remained significant when accounting for the 20% false detection rate. Of those identified SNPs in the vascular cell adhesion molecule VCAM gene (rs1041163 and rs2392221) and SLCO1B1 gene (rs4149056) were found to be the most promising (79).

This large study was the first to adopt multiple SNP panels to look for potential pharmacogenetic determinants of MPA associated anaemia and leucopenia. The study included a large number of participants but did not look at GI side effects or rejection rates, both clinically important. The study found important associations with 4 genes (IL, VCAM, HUS1 and CYP2C8) that have not been previously studied in relation to genetic determinants of MPA-related side effects or metabolism. This highlights the importance of going beyond the candidate gene approach as significant predictors of clinical outcomes may not exist in expected genes.

A 2012 publication(80) attempted to replicate these results in 338 RTR's and found a similar association with CYP2C8 SNP rs11572076 and an increased risk of anaemia and leucopenia. But they did not find any significant associations with IL12A rs568408 or HUS1 rs1056663 SNPs and these outcomes (80). It was argued that differences in the study population, treatment regimens and statistical methods could explain this lack of association (81)

The literature to date thus lacks a well powered reproducible replication study with direct correlation to clinical outcome, and hence the potential to provide sufficient evidence for clinical application; this study aims to produce that.

1.6 Summary

This chapter has given an in depth overview of MPA metabolism and pharmacogenetic/genomics. It is clear that the field of pharmacogenomics research is rapidly developing and that a move towards the use of genetic profiling to individualised medicine is taking shape. It is also evident that pharmacogenetics has a potential role in transplantation, particularly with regards to MPA for which there is currently no clear benefit of TDM. This study aims to extend the knowledge about pharmacogenomics of MPA, using array based exome chip genotyping as a novel approach to provide whole exome profiling, to identify pharmacogenetic variants that influence tolerance of MPAP in renal transplant recipients.

Table 1-1: Summary of MPA pharmacogenomics studies to date

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
Michelon et al 2010 (8)	ABCB1 (MRP1)	-3435C>T	rs1045642		239	Renal Transplant recipients 83.9% Caucasian 12.4% African 3.2% Asian	81.7% Tacrolimus 11% Ciclosporin 0.5% MTOR 6.9% steroids only	ADR: Leucopenia (WCC<4) Anaemia (Hb<120g/L) Thrombocytopenia (Plts <150) GI Side effects (Diarrhoea, Vomiting, Nausea) Infections BPAR: BANF 2005 Criteria	Patients with SLCO1B1 521C have significantly less ADR's than those with 521T (P=0.002)	Lack of statistical power. No explanation why these SNPs chosen. First study to look at a large number of genes.
	ABCC2 (MRP2)	-24C>T -1249G>A -3972C>T -1446C>G	rs717620 rs2273697 rs374006 no rs							
	UGT2B7	-802C>T	rs7439366							
	UGT1A9	-275T>A -98T>C	No rs rs72551330							
	SLCO1B1	-11187G>A -388A>G -521T>C	rs4149015 rs2306283 rs4149056							
	SLCO1B3	-334T>G	rs4149117							
	IMPDH1	-106G>A -125G>A	rs2278294 rs2278293							

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
Fujiyama et al 2009 (76)	CES2	-1548A>G -4549A>G -8271C>T	rs3890213 rs2303218 rs2241409		80	Renal Transplant recipients Japanese	Not Specified	MPA and MPAG Area under Curve	No CES2 allelic variations ass with inter-individual MPA concentrations	Small number of patients Cannot extrapolate to wider population
Naesens et al 2006 (77)	MRP2	-1549G>A -1023G>A -1019A>G -24C>T -1249G>A -3972C>T 4544G>A	rs717620 rs2273697 rs374006		95	Renal Transplant recipients Caucasian	Tacrolimus	MPA AUC RBC, ACC, Creat, LFT, Alb, Creat Clearance, 24 Hour protein, Diarrhoea Infection	MRP2 24C>T increased dose corrected MPA, increased Diarrhoea 3972T>A increased MPA dose 24C>T and 3972C>T protective against drop in MPA when ass Liver dysfunction	Small numbers.
	UGT1A9	2152C>T 278T>A 98T>C	rs17868320 rs72551330							

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
Wang et al 2008 (48)	IMPDH1	109A>T 462T>C 354A>T 227C>T 169T>C 125G>A 106G>A 1572C>T 898G>A 1552G>A	rs2288553 rs11770116 rs2288548 rs2288549 rs4731448 rs2278293 rs2278294 rs2228075 rs2288550		191	Renal Transplant recipients	Tacrolimus or Cyclosporin Plus Prednisolone	Leucopenia Acute biopsy proven rejection (ABPR)	IMPDH1 rs2278293 and rs 2278294 associated with increased incidence of ABPR P<0.03 in first 1 year No associations with SNPs and leucopenia	Good number of patients but still underpowered. Good number of SNPs
	IMPDH2	787C>T 3757T>C	 rs11706052							

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
Miura et al 2007 (78)	SLCO1B1				87	Renal Transplant Recipients Japanese	Tacrolimus and Steroids	MPA AUC Diarrhoea, Nausea, Vomiting or abdominal pain	SLCO1B3 334GG associated with Significantly increased AUC 6-12 but not associated with GI side effects Presence of ABCC2 C24T and SLCO1B3 associated with significantly lower MPA (P=0.001)	Small study. Cannot extrapolate to wider population
	SLCO1B3	334T>G 699G>A	rs4149117 rs7311358							
	SLCO2B1									
	ABCC2	24C>T	rs717620							
Kuypers et al 2005 (51)	UGT1A9	2152C>T 975T>A	rs17868320		95	Renal Transplant Recipients	Tacrolimus and Prednisolone	MPA plasma concentration Leucopenia Diarrhoea	275T>A and 2152C>T were associated with lower MPA exposure in patients given 2g MMF	No information on power of the study
Djebli et al 2007 (82)	UGT2B7	842G>A	rs7438135		92	Renal Transplant Recipients	Cyclsporine, Tacrolimus or Sirolimus		842AA was associated with higher MPAG AUC when MMF given in combination	No information on patient ethnicity No power information

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
									with sirolimus	
Baldelli et al 2007 (52)	UGT1A9	2152C>T 1887T>G 665C>T 440C>T 331T>C 275T>A 33M>T	rs17868320 rs2741045 rs2741046		40	Renal Transplant Recipients Caucasian	Cyclosporin	MPA pharmacokinetics	440C>T and 331T>C associated with higher MPA AUC	Small Study Caucasian only population.
Levesque et al 2006 (53)	UGT1A8 UGT1A9 UGT2B7	277C>Y 173A>G 2152C>T 275T>A *1*2*3	Rs17863762 rs17868320		52	Healthy Volunteers	N/A	MPA pharmacokinetics	275T>A and 2152C>T were associated with lower MPA exposure UGT2B7*2 associated with higher MPA exposure than UGT2B7*1	Small study Not validated in renal transplant population.
Van Schaik et al 2009 (65)	UGT1A8	518C>G 830G>A	rs1042597 rs17863762		338	Renal Transplant Recipients 88% Caucasian 3% Blacks 4% Asian	Cyclosporine or tacrolimus	MPA Pharmacokinetics and ABPR	518GG associated with ↑ MPA AUC-12 vs to CC (P=0.03). UGT1A9 275T>A and 2152C>T associated with	Good number of patients Did not consider clinical outcomes beyond
	UGT1A9	2152C>T 275T>A 98T>C 1399C>T	rs17868320 rs725513							

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
			30 rs2741049						↓ MPA AUC0-12 in tac treated patients associated with ABPR (P=0.042). UGT1A9 98T>C associated with ↑MPA AUC0-12 in tac and cyA. UGT2B7 842G>A associated with ↓MPA AUC0-12 in cyA (P=0.09 ns)	ABPR.
	UGT2B7	842G>A 79G>A 802C>T	rs7438135 rs73823859 rs7439366							
	MRP2	24C>T 3972C>T	rs717620 rs3740066							
Johnson et al 2008 (66)	UGT1A8	518C>G 173A>G 830G>A 277C>Y	rs1042597 rs17863762		117	93 Renal 11 Pancreas 13 Kidney and pancreas transplants 92.35 Caucasian 2.5% African American 2.5% Asian 2.5% Other	Cyclosporine or tacrolimus	MPA trough concentration Alb, Hb, AST,ALT, Creat, Bili	UGT1A9 275T>A/2152C>T in combination with CyA led to reduced MPA Conc (P=0.008) UGT1A8*2 important predictor of MPA dose corrected trough conc (P=0.001)	No power calculations. Good attempt to consider effect of other immunosuppression.
	UGT1A9	8G>A 3C>Y 98T>C 2152C>T 275T>A	rs145084767 rs72551330 rs17868320 rs6714489							

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
Sanchez-Fuctusos et al 2009 (16)	UGT1A9	98T>C 2152C>T 275T>A	Rs72551330 rs17868320		133	Caucasian RTR	Tacrolimus	30 patients (15 with SNPs 15 without) had MPA pharmacokinetics GI side effects	SNPs had no effect on AUC0-6 but decreased AUC6-12 P<0.04 UGT1A(275T>A/2152C>T great incidence and severity of GI side effects	Only looked at AUC in 30 patients
Prausa et al 2008(67)	UGT1A8	830G>A	rs17863762		38	Paediatric patients	Tacrolimus (73%) Cyclosporine (9%)	Two groups 1.Adverse effects GI or Leucopenia requiring MMF reduction or stopping 2. Tolerated MMF	Increased incidence of: UGT1A9 331T>C in AE group (P=0.04) UGT2B7 in AE group (P=0.08 ns)	Small number of children. Not validated for adult population
	UGT1A9	275T>A 331T>C 2152C>T 98 T>C	rs2741046 rs17868320 rs72551330							
	UGT2B7	900A>G								
Betonico et al 2008 (68)	UGT1A8	999C>T 255A>G 277 G>A			74	Brazillian RTR	Tacrolimus Cyclosporine Sirolimus	Side effects: Diarrhoea Infection Blood disorders	UGT1A8 227A increased incidence of infection compared to 227G 999C/255T in	Small numbers. Not validated outside Brazilian population.

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
									combination with 277A- more infective episodes with 2g MMF (P<0.02)	
Yang et al 2009 (70)	MRP2	C24T	rs717620		67	Predominantly Hispanic RTR	Tacrolimus or Cyclosporine	GI side effect using Gastrointestinal system rating scale (GSRS)	UGT2B7 C802T variant seemed to protect against GI side effects. Higher diarrhoea GSRS in cyclosporine than tacrolimus group.	Small no of patients. Predominantly Hispanic = not validated outside this population.
	UGT2B7	C802T	rs7439366							
Van Agteren et al 2008 (71)	UGT2B7	840G>A			332	RTR		MPA, MPAG and Acryl MPAG concentrations Diarrhoea and leucopenia	No significant associations found	Good size study but only one SNP in one gene looked at.
Kagaya et al 2007 (69)	UGT1A8	*2			72	Renal Transplant Recipients Japanese	Tacrolimus	Day 28 MPA AUC	No significant difference in MPA AUC amongst different	Small Study. Cannot extrapolate to wider
	UGT2B7	*2								
	UGT1A9	-275T>A -2152C>T	rs178683							

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
			20						genotypes	population
Kagaya et al 2010 (73)	IMPDH1	125G>A 106G>A	rs2278293 rs2278294		82	RTR Japanese	Tacrolimus	Day 28 MPG BPAR	No association between SNPs and MPA No clinical rejection in Study	Small Study. Not validated outside the Japanese
Winnicki et al 2010 (50)	IMPDH2	3757T>C	rs11706052		100	Healthy volunteers	N/A	IMPDH activity and lecopenis	Significantly less antiporliferative effect of MPA on leucocytes in healthy individuals with rs11706052 SNP	Not validated in transplant population
Gensburger et al 2010 (74)	IMPDH2	3642A>G 3757T>C 787C>T	rs4974081 rs11706052		456	RTR	Tacrolimus or cyclosporine	BPAR Leucopenis Cytomegalovirus Infection Diarrhoea	IMPDH1 rs2278294 significantly associated with lower risk of BPAR (P=0.0075)	Good size study powered to detect significance Potential

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
	IMPDH1	125G>A 106G>A	rs2278293 rs2278294						and higher risk of leucopenia	clinical importance
Sombogoard et al 2009 (75)	IMPDH2	3757T>C	Rs11706052		101	RTR	Tacrolimus and pred	IMPDH activity MPA AUC	IMPDH1 3757T>C polymorphism is associated with increased IMPDH activity in MMF treated patients	Underpowered study, set out to look for 8 different SNP and could only identify one in their treatment group.
Jacobson et al 2011 (79)	Broad panel SNP chip	2724 SNPs			978	RTR or SPKT 76.5% Caucasian 17.5% African American 3% Asian 3% other	All received CNI, not clear % that received CYA vs Tac	Leucopenia (Wcc<3000 cells/mm3) Anaemia (Hb<10g/dL or Haemocrit <30%)	ILI2A (rs568408) increased HR for Anaemia: 1 Allele- HR 1.98(1.39-2.82) 2 Allele – HR 1.95-7.95) HUS1 (rs2037483)	Good number of participants. Multiple SNPs Did not look at GI side effects

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
									decreased HR for anaemia: 0.54(0.39-0.74) CYP2C8 Increased HR for anaemia: 3.24 (1.7-6.2) No SNP significantly associated with leucopenia after adjustment for 20% FDR but VCAM and SLCO1B1 most promising.	or rejection which are clinically relevant.
Bouamar et al 2012 (6)	ABCB1	C1236T G2677A G2677T C3435T			338	RTR participating in 'FDCC trail' 88% Caucasian 3% Black 4% Asian 5% Unknown	46.4% Tacrolimus 50.9% CYA	MPA-AUC Diarrhoea Leucopenia	RR Diarrhoea found to be 1.8 fold higher in patients co-treated with tacrolimus than CYA No significant associations found with SNPs and any of the outcome	Good number of participants Side effects not well defined. Used participants from a study looking at fixed dose v
	SLCO1B1	388A>G 521T>G	Rs2306283 Rs4149056							
	SLCO1B3	334T>G 699G>A	Rs4149117 Rs73113							

Reference	Gene	SNP	Rs Number	Caucasian Frequency	Patient number	Patient/ Transplant characteristics	Immunosuppression Co-treatment	Outcome measured	Main Results	Limitations / Comments
			58						measures	concentration controlled MPA so dose being adjusted as part of this trail
Bouamar et al 2012(80)	CYP2C8		Rs11572076		338	RTR participating in 'FDCC trail' 88% Caucasian 3% Black 4% Asian 5% Unknown	46.4% Tacrolimus 50.9% CYA	Leucopenia Anaemia	CYP2C8 rs 11572076 was found to be assicoated with Anaemia (P=0.021) and Leucopenia (P=0.007) in the fixed dose group (178 participants) No association with IL12A or HUS1 SNPs with these outcomes.	Good no of participants Used participants from a study looking at fixed dose v conc controlled MPA so dose adjusted as part of this trail
	IL12A		Rs568408							
	HUS1		Rs1056663							

This table summarises the mycophenolate related pharmacogenomics studies in renal transplant recipients to date, it includes details of samples size and a critique of the study

2 Chapter 2: Methods

2.1 Introduction

The aim of this research project is to identify and investigate genetic single nucleotide polymorphisms SNP (both candidate and novel), which are likely to be associated with a patient's response to the immunosuppressive drug Mycophenolate (mofetil or sodium) following renal transplantation. Ultimately it aims to contribute to the future development of a simple genetic test that will allow us to predict a patient's response to Mycophenolate (mofetil or sodium) before they receive an organ transplant and hence allow us to develop individualised dosing of this drug to optimise immunosuppression, with improved graft function and reduced intolerance and toxicity. Indeed, the principle can equally well be applied to non-transplant immunosuppression in situations like vasculitis and other immune-mediated diseases where MPA is currently used with good outcomes.

This chapter describes the overall study design and methodology used. It will also describe the recruitment process, data collection, laboratory methods and results analysis.

2.2 Research question

Does the presence of genetic single nucleotide polymorphisms, either candidate or novel, predict individual tolerance and efficacy of MPA in RTR and does this:

1) Influence the rate of rejection in the early post-transplant period
2) Predict adverse drug side effects
3) Predict drug tolerance
4) Influence the predilection for infection post transplantation

2.3 Study Design

A cohort study has been chosen as it is currently the most valid and effective study design in pharmacogenetic research (25) and is suitable for studying multiple outcomes. This is a non-interventional cohort study.

2.3.1 Primary outcomes

The primary outcomes measure is to describe 1) the frequency of genetic polymorphisms, both candidate and novel and 2) study to association of these polymorphisms with the following clinical outcomes post transplantation.

Alteration to MPA dose (Stop, dose reduction, Change of preparation)

Drug side effects (Leucopenia, Anaemia, Gastrointestinal)

Incidence of biopsy proven transplant rejection

Incidence of infection post transplantation

Definitions of 'events' in this study are given in table 2.4.

2.3.2 Secondary outcomes

The secondary outcome measures will be looking at the associations between:

Time to event (anaemia, leucopenia, BPAR) analysis

MPA dose alteration and Transplant graft function

2.3.3 Sample Size

An example of a conservative power calculation based on a genotype frequency of just 10% is given. Professor Cathryn Lewis from Guy's & St Thomas' Trust (GSTS) has provided expert statistical help to ensure correct sample size calculation for this study. With 30% of 285 patients, developing one or more of the above side effects, there would be: 80% power (at the 5% significance level) to detect a difference between a genotype frequency of 10% in tolerant patients, and 24% in patients with side effects.

2.3.4 Setting and population

Recruitment for this study took place at a single transplant centre the Wessex Renal and Transplantation Unit (WRTU). WRTU is based in the Queen Alexandra Hospital in Portsmouth, UK and is part of the Portsmouth Hospitals NHS Trust. It is one of the larger regional renal units in the UK, providing renal services to an adult population of 2.2 million. The unit covers the majority of Hampshire and Isle of Weight as well as parts of the adjoining counties of Wiltshire, West Sussex, and Surrey. The unit currently (October 2013) has 620 patients on dialysis and 750 patients with a functioning renal transplant. It performs approximately 70 renal transplants per year, including a living donor programme.

The region covered by WRTU has an end stage renal failure (ESRF) incidence rate of 110 per million population (pmp) per year for new end stage renal failure. This is comparable to the overall UK incidence rate of 107pmp. There is a marked gender difference in incident ESRF population in the UK with males at 136 pmp and females 79pmp (83).

The primary renal diagnosis in patients with ESRF at WRTU is comparable to the rest of the UK, this is shown in Table 2.1 which uses data from 14th UK Renal registry report (83).

Table 2-1: Percentage distribution of primary renal diagnosis in the 2010 incident cohort WRTU and UK

	Percentage							
	Unknown	DM	GN	BP	PKD	PN	RVD	Other
WRTU	10.3	25.5	8.3	11	5.5	12.4	8.3	18.6
UK	19.8	24.2	11.6	6.7	6.6	7.4	7.5	16.2

Table to show the cause of renal failure (expressed as a percentage of the total number of individuals with ESRF) in the UK population as reported in the 2010 renal registry and the WRTU as reported in 2010. GN –glomerulonephritis DM-Diabetes BP- Hypertension PKD-Polycystic kidney disease PN- Pyelonephritis RVD- Renovascular disease

The WRTU has a higher percentage of patients with a working transplant as their modality of ESRF treatment in comparison to the rest of the UK. This is shown in table 2.2 and data is taken from 14th UK Renal Registry report (84).

Table 2-2: Percentage distribution of 2010 prevalent RRT cohort by modality WRTU and UK

	Percentage by Modality of RRT		
	Haemodialysis	Peritoneal dialysis	Transplantation
WRTU	36	8	56
UK	44	8	49

The percentage of patients on each modality of treatment for ERF in the UK as reported in the 2010 renal registry and in WRTU as reported in 2010.

The population of the south coast of England and the Counties which are served by the WRTU have a lower ethnic diversity when compared to the rest of the UK. The WRTU population is predominantly Caucasian and this will need to be taken into account when considering the validity and generalizability of results. The ethnic origin of both incident and prevalent ESRF population in WRTU and the UK is shown in Table 2.3 and uses data from 14th UK Renal Registry report (83, 85).

Table 2-3: Percentage distribution of ethnicity in 2010 incident and prevalent cohort WRTU and UK

	Percentage (Incident/ Prevalent)				
	White	**Black**	**South Asian**	**Chinese**	**Other**
WRTU	90.5 / 92.7	2.7 / 1.2	5.4 / 3.0	0.7 / 0.7	0.7 / 1.1
UK	79.8 / 69.7	6.8 / 6.5	11.3 / 9.3	0.4 / 0.6	1.7 / 1.3

Table to show the ethnic diversity within the UK ESRF population and the WRTU as reported in 2010 (expressed as a percentage incident/prevalent)

It was felt that the WRTU would provide a large transplant cohort which is fairly representative of the UK ESRF population. The lack of ethnic diversity is the main factor which differs from the UK average but it is felt that this provides an advantage in this pharmacogenic study due to the differing minor allele frequency (MAF) of SNPs within the different ethnicities.

2.3.5 Subject selection

A combination of MPAP, CNI, Steroids and an induction agent such as Basiliximab constitutes the standard induction therapy in renal transplant recipients in WRTU. RTR are then maintained long term on a combination of CNI and MPA and they will constitute the cohorts of this study.

The study consists of already transplanted patients who received MPA either in the form of MMF or MPS at full dose at the time of transplantation as well as patients transplanted during the study period, who fit the same criteria but were enrolled at time of transplantation.

The WRTU database (Proton) was interrogated to identify patients who had received a renal transplant and received MPA after the year 2000 (in combination with CNI at the time of transplantation). A monthly update of new transplants was also obtained to allow recruitment of new transplant patients. The subject identification flow chart is shown below Figure 2.1.

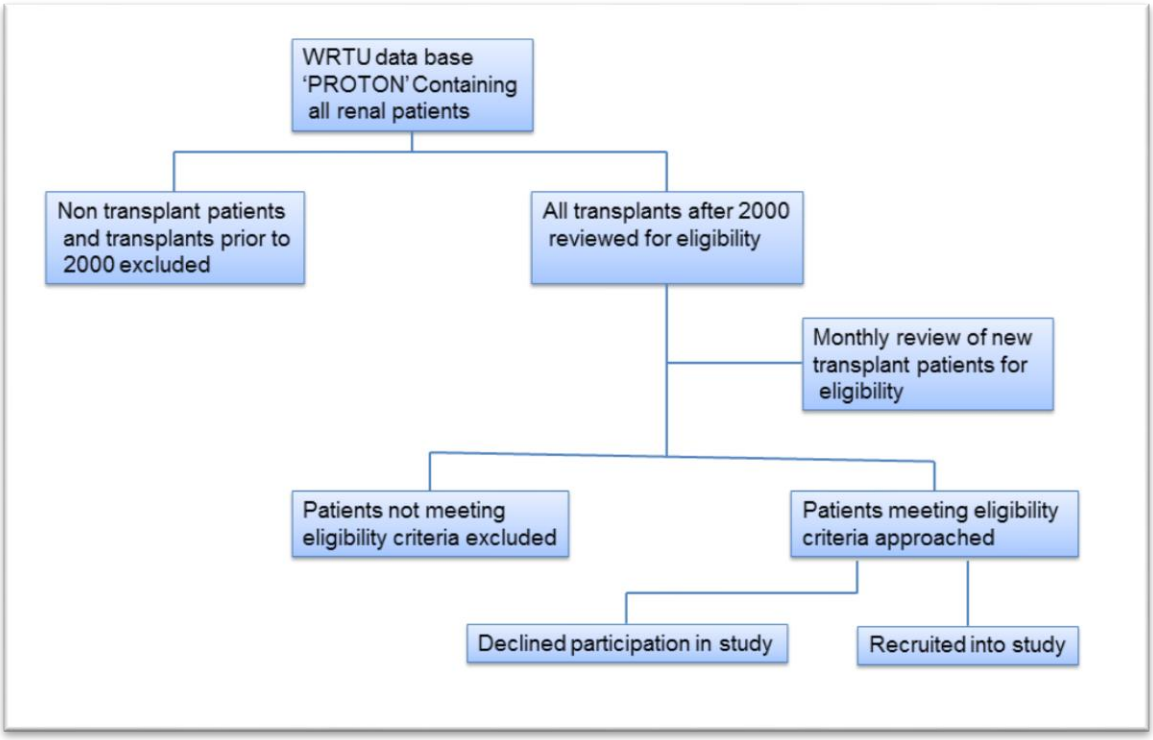


Figure 2-1: Flow chart of subject indentification for the study.
Subjects transplanted after 2000 were identified using the WRTU system proton at the beginning of the study. New transplants taking place during the study were also reviewed for eligibility. All eligible subjects were approached for participation.

2.3.6 Eligibility

The eligibility criteria in this study were set out to provide an inclusive study with very few individuals being excluded from participation. The aim is that the study will reflect the UK transplantation population as a whole and selection bias should be minimal. The criteria are defined below.

2.3.7 Inclusion criteria

Renal Transplant recipient.

Received full dose Mycophenolate Mofetil (2g per day) or Sodium (1440 mg per day) at the time of renal transplantation.

Age 17 years or older at the time of transplantation.

2.3.8 Exclusion Criteria

Simultaneous transplantation of any other organ at the time of renal transplantation.

Did not receive Mycophenolate Mofetil or Sodium at the time of renal transplantation.

Age less than 17 years at the time of transplantation.

Pregnant women.

Due to the non-interventional nature of this study participants were allowed to be enrolled in other studies provided they still fitted the above criteria.

2.4 Ethical approval

Ethical approval for the study was sought and obtained from the Southampton and South West Hampshire Research Ethics Committee A. The standard application for ethical approval in the UK is via the Integrated Research Application System 'IRAS', and this on line application system was used for this study. Approval was granted on 28th January 2011 (Appendix 1). A substantial amendment was approved on 2nd February 2012 (Appendix 2). Approval was also granted at a local level on 1st February 2011 for conduct of this study within the Portsmouth Hospitals NHS Trust (Appendix 3).

2.5 Peer Review

Peer review by an expert in the field of transplantation was sought prior to applying for ethical approval and commencing this study. The review is included in appendix 4.

The study was also patient reviewed by a renal transplant recipient to assess the acceptability to patients and a favourable opinion was given.

2.6 Subject recruitment and Consent

Patients identified as eligible were approached to take part in this study. They were approached at an outpatient appointment or whilst on the ward following transplantation. Details of the study and information sheets (Appendix 5) were supplied. Patients willing to participate were asked to give written consent on a consent form supplied (Appendix 6) once they had sufficient time to consider participation, ask any questions and consult with others should they wish.

Once consented patients had a blood sample (5 ml in EDTA) taken for genetic analysis, the blood sample was taken at the same time as their routine clinic or ward bloods.

At entrance into study all participants were allocated a study number which was subsequently used for all laboratory samples and to anonymise all the data collected. The study numbers with patient details were stored on a secure NHS computer.

2.7 Discussion on recruitment

In the initial study protocol the aim was to recruit 450 patients into the study. This would then allow for the identification of rarer SNPs with low minor allele frequency (MAF) within the population. It became apparent that this number would not be achievable within the study time period at a single centre, even with a combination of previously transplanted patients and new

transplants. The ability to recruit these numbers was also effected by the recruitment site enrolling in another national transplant study. This meant that a number of previously eligible subjects were now not suitable for recruitment due to alteration in their baseline immunosuppression.

All subject identification and recruitment was done by me as the primary investigator and it is recognised that having a single recruiter is also a limiting factor in reaching the initial target.

It was therefore decided that a realistic number of recruits would be 300, which would be achievable within the time constraints of the study, and still statistically viable. This was felt to be a good number of patients given the technique being used. This was discussed with the genetic statistician and it was felt that the study still remained adequately powered for SNPs with higher MAF frequencies within the population but would not be sufficient for some of the rarer SNPs. This was felt to be acceptable as the aim of the study and outcomes measures relate to clinical significant side effects and hence to have the potential for clinical applications the SNPs would need a MAF of at least 10%.

With ethical approval still in place to recruit a further 200 patients there is the potential to continue recruitment and add to this work without a separate ethical approval and provide a repetition study to support the outcome from this research.

There has also been a separate ethical approval to produce a repetition cohort in subjects with lupus nephritis who are treated with MPA. This work will not be discussed further here as it is beyond the scope of the methods chapter of this thesis and will not form part of this DM.

2.8 Control group

A separate control group was not recruited for this study. The participants within the cohort will act as the controls as the comparison is between wild, heterozygotes and homozygotes for the SNP of interest and individuals with or without the outcome of interest.

A number of pharmacogenomic studies, including 2400 participants, looking at different agents in specific populations, are being conducted through the Purine Laboratory at GSTS using the same sequencing method. These studies will not be discussed in this thesis but the overall SNP MAF for participants in all these studies will be quoted to provide a comparison to those found in the MPA cohort. The genetic data produced on all 2400 patients will also be used for quality control of the data that will be discussed later in the chapter.

2.9 Data collection

Extensive clinical data was collected from the patients' case notes, clinical letters, transplant cards, 'Apex' laboratory results system and 'Proton' renal data base system. The principal investigator (myself) remained blinded to the results of the pharmacogenetic analysis during collection of the patient data so as to eliminate bias. Due to the use of multiple sources to collect and cross reference the data the final data set was >99.95% complete. The collection of extensive data ensures information is available on numerous potential confounding factors which can then be corrected for in the analysis.

The following data was collected and recorded on each patient at weekly intervals up to 12 months post transplantation

Patient demographics

- Age
- Sex
- Ethnicity
- Cause of renal failure
- Previous renal transplants
- Type of transplant received (deceased / live donor)
- HLA mismatch
- CMV status of donor and recipient
- Donor age, gender, cold ischaemic time
- Induction therapy and dosage at time of transplantation
- Maintenance immunosuppression
- Delayed graft function
- Episodes of biopsy proven rejection
- Serum creatinine, eGFR, CNI drug concentrations, haemoglobin, white cell counts
- Drug side effects, with particular focus on gastro intestinal symptoms
- Episodes of infection post transplantation
- Any dose adjustments, preparation change or cessation of MPA and reasons for this

All the collected data was recorded on a data collection sheet (Appendix 7). Data was then entered on a Microsoft Excel$^{(R)}$ spread sheet. Collected data was double checked and double entered to minimise error.

As well as documenting data at specified time intervals all renal biopsy reports were reviewed and all measures of haemoglobin, white cell count, creatinine and eGFR taken in the first 12 months

post transplantation were reviewed to ensure no primary outcome events were missed, and allowing observation of trends.

Once data collection was complete data was coded for demographics, baseline characteristics and outcome measures and recorded in a separate Excel spread sheet. Data was again double checked and double entered to minimise error.

2.10 Definitions used

Events were defined prior to commencing the study based on clinical practice within the recruiting unit and on evidence from the literature (Table 2.4).

Table 2-4: Definitions used to define events in study

Event	Definition of event	Justification
Biopsy proven acute rejection (BPAR)	Evidence of acute rejection on transplant biopsy reported by a renal histopathologist using the BANF criteria guideline and unaware of the study.	BANFF universally recognised criteria (86)
vascular or cellular rejection	Evidence of acute vascular or cellular rejection on transplant biopsy reported by a renal histopathologist using the BANF criteria guidelines and unaware of the study.	BANFF universally recognised criteria (86)
Any Change to MPA	Any change in dose, frequency or preparation of MPA for any reason	Suggest intolerance or side effects
Stopped MPA	Patients stopped MPA completely and it was not successfully reintroduced	Suggests severe intolerance or side effects
Any Upper GI side effects	Any upper GI symptoms felt by the treating clinician to be related to MPA and requiring alteration to MPA	Must be clearly documented in case notes or patient letters
Any Lower GI Side effects	Any Lower GI symptoms felt by the treating clinician to be related to MPA and requiring alteration to MPA	Must be clearly documented in case notes or patient letters
Any GISE	Any upper or Lower GI symptoms felt by the treating clinician to be related to MPA and requiring alteration to MPA	Must be clearly documented in case notes or patient letters

Event	Definition of event	Justification
Leucopenia WCC<3.0 x10^9/L	A reduction in wcc to <3.0 on two or more consecutive blood test or a single test if this resulted in a change in MPA dose	The definition of leucopenia varies in the literature {Bouamar, 2012 #12
Anaemia Hb <10 g/dL	A reduction in Hb to <10g/dL after day 30 on two or more consecutive blood test	Different transplant pharmacogenomics studies in the literature have used variable Hb to define anaemia (8, 66, 79, 80). The majority use a predefined Hb level after day 30 post-transplant to account for surgical blood loss and graft function. Hb<10g/dL is the level at which action would be taken at the recruiting centre.
Infective episode	Recurrent or severe infection felt by the treating clinician to be as a result of the patient's immunosuppressed state (Excluding urinary tract infection)	Based on practice in the recruiting centre.

This table shows the definitions of events used for this study along with a justification for the definitions used.

2.11 Laboratory methods

2.11.1 DNA Extraction

Peripheral whole blood was collected from all patients recruited into the study in an EDTA tube. Blood samples were immediately frozen and stored at -20^{o}C. Samples were then defrosted at room temperature for 2 hours prior to use.

DNA was extracted from the whole blood using QIAamp® DNA blood Midi kits (Protocol 1 Appendix 8 for full methods used).

Extracted DNA was then stored in DNA free ependorf tubes at -60^{o}C, all stored DNA samples were identifiable only by the allocated study number.

DNA samples were then transferred to the' Purine Laboratory at St Thomas' Hospital, London (PLSTTH) for further analysis; a material transfer agreement was put in place at the time of ethics approval.

DNA to be used for rtPCR and PCR required no further preparation but 30µL of each sample was transferred to 96 well plates with strict templates logging each sample position.

DNA was also required for exome sequencing and this was carried out using Illumina® Human Exome Bead Chip and will be discussed in detail later in this section. DNA was required to be at an exact concentration of 50ng/mL for exome sequencing.

2.11.2 DNA concentration

DNA concentration of each sample was measured using the Qubit® machine. (Protocol 2, Appendix 9 for full methods used). To ensure accuracy, the DNA concentration of each sample was measured twice with an average of the two samples being taken. Both samples were required to read within 5ng/mL of each other or they were repeated with a newly prepared sample following the protocol.

The average sample concentration for each DNA specimen was then calculated. Average concentrations below 48ng/mL required sample concentration to ensure they reached the required level for exome sequencing. DNA purification/concentration was carried out for those samples with a concentration below this level (Protocol 3, Appendix 10).

Once any samples had been concentrated the concentration was rechecked using the Qubit ® (Protocol 2, Appendix 9).

DNA samples were then diluted with buffer to provide a concentration of 50 ng/mL with a sample volume of 10µL. This was calculated for all samples and they were diluted appropriately in a 96 well plate (see Appendix 11 for example calculation).

2.12 Choice of technology for genotyping

When this study was in the initial planning stages, genome wide association studies (GWAS) using microarray SNP chips were felt to be the best available method for pharmacogenetic studies of this sort. The SNPs represented on standard GWAS chips are typically intron located, are based on information from the HapMap and 1000 genomes project (28), with the linkage signals generated pointing to a genomic region and not necessarily directly to the causative gene or polymorphism (31).

Protein coding regions constitute about 1% of the human genome and are located in the exon regions of the genome (32). These coding region variants are considered to have high pharmacogenetic relevance, and hence we considered exome sequencing as a method for generating genome-wide coding region data (87). Discussions were had with several companies providing different GWAS and exome sequencing chips. An acceptable sequencing coverage is generally considered to be >80% of the exome covered >20 fold. This means that homozygous, heterozygous and wild type genotypes for a sequence variant can be called with a high degree of certainty. The Illumina HiSeq next generation exome sequencers were then tried by the GSTS research team multiplexing 6 patients per sequencing lane, and they found that 40% of the genome was covered >20 fold raising the possibility that for 60% of polymorphisms typed variant genotypes may not be called correctly if at all. Nevertheless, >22,000 variants per sample were called which is in the range of what a typical European exome yields. They concluded that we would need to run samples at lower than 4 plex. As each sequencing lane costs £1,000 to run, this would have pushed the cost of exome sequencing per patient beyond the reach of our budget.

Illumina then developed an exome beadchip containing >240,000 predominantly exon located and near gene polymorphisms identified from 12,000 exome sequences and SNP data bases with genotype frequencies down to 1:3,000 represented (1) (Table 2-5) Nonsynonymous SNPs contained on the beadchip had to have been observed in at least two separate studies on three or more occasions with the splicing and stop-altering variants on at least two occasions (88).

Table 2-5: Human Exome Beadchip content (1)

Marker Categories	Values
Total markers	> 240,000
Number of unique RefSeq entries covered by at least one probe	> 20,000
Nonsynonymous SNPs (NCBI)	219,621
SNPs in splice sites	10,675
Stop variants	5,637
SNPs in promoter regions	7,012
SNPs in extended MHC region	5,158
GWAS tag markers*	4,761
HLA tags	2,061
Ancestry informative markers	3,468
Identity by descent markers	3,369
X / Y / mitochondrial	470 / 101 / 177
Indels	180

This table shows the specification of the Human Exome beadchip version 1.1 that was used in this study, outlining the number of SNPs included in different genomic regions.

After discussion with the genetic statisticians it was felt that this chip would give equivalent, and in some respects, better coverage than exome sequencing. The inclusion of rare variants on the chip would also allow the effects of multiple rare SNPs within a candidate gene to be analysed (89). The disadvantage of these chips is that the non-coding region or promoter region variants are under-represented and some SNPs identified by exome sequencing are not on the chip as allele specific probes could not be synthesised.

The chips at a pre-release price of £26.00 per patient also represented exceptional value for money, compared to the £500+ per patient for standard exome sequencing. All patients recruited to the study have been genotyped using the Illumina Human Exome Beadchip v1.1.

2.13 Genotyping

Genotyping was carried out using Illumina Infinium Human Exome BeadChips V1.1 which consists of over 240,000 SNPs with focused coverage of the exonic region as outlined above.

For the exome sequencing illumina bead chips, silca beads are etched into microwells. These beads are coated with a specific oligonucleotide with a probe which targets a specific locus (SNP) within the genome(90).

The Chip wells were loaded with a volume of 5µL of DNA at a concentration of 50ng/mL with a separate well for each patient. A strict log of each sample position was kept to allow reference back to patient data.

DNA then fragments and each probe binds to the complementary base pairs in the DNA. Allele specificity is then conferred by a single base pair extension (figure 2.2 and 2.3).

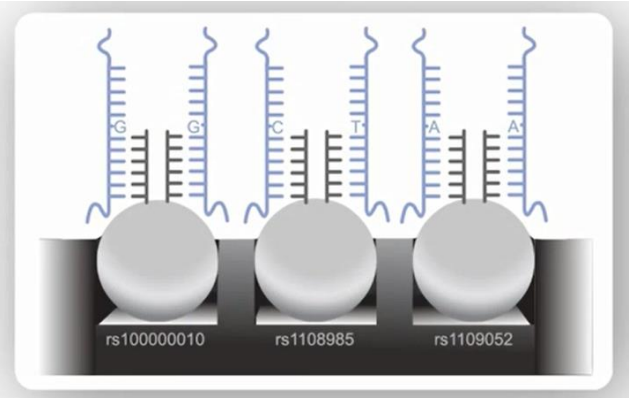


Figure 2-2: Illumina beads in individual wells, with attached specific oligonucleotide (shown in blue) and fragmented DNA (shown in black)(90)

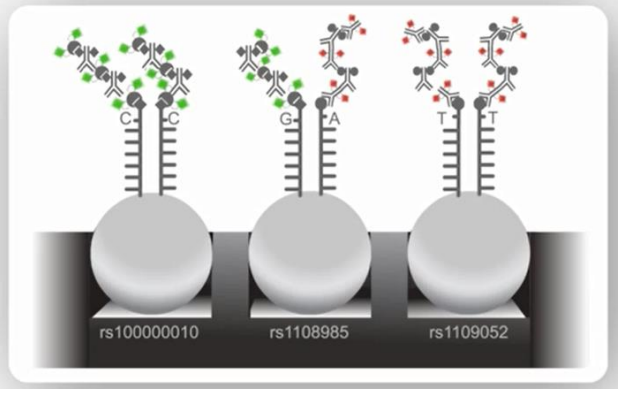


Figure 2-3: Single base pair extensions form on the fragmented DNA, with bases containing different florescent nucleotide labels.(90)

The Bead Chips were then run on the Genome analyser at Guy's Hospital London. A laser is then used to floresce the nucleotide lable which generates a two colour readout. Each predetermined base emits a different colour.

For example an A may emit green whilst a G may emit red. The intensity of the colour emssion is then read by the scanner which will in turn provide the genotype eg full green signal equal AA, full red signal equals GG, green and red signals of equal intensity equals AG shown in figure 2.4 (90)

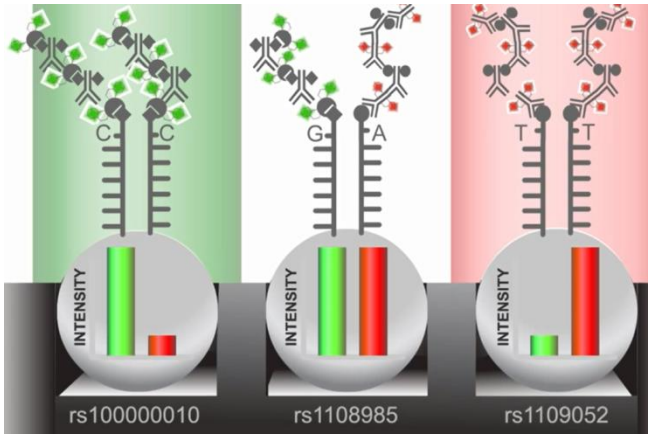


Figure 2-4: Example of colour intensity signals based on a single base present (90)

Results were exported to the computer programme Genome Studio data analysis software 'Gen-call' designed for use with all Illumina products. Samples were identified and labelled with the allocated study number and gender was assigned. Gen-call uses a set of customised clustering algorithms to determine genotypes from intensity clouds from a combination of all individuals results producing a plot generally containing three clusters one for AA individuals, one for AG individuals and one for GG individuals (90, 91). An example plot is shown in figure 2.5

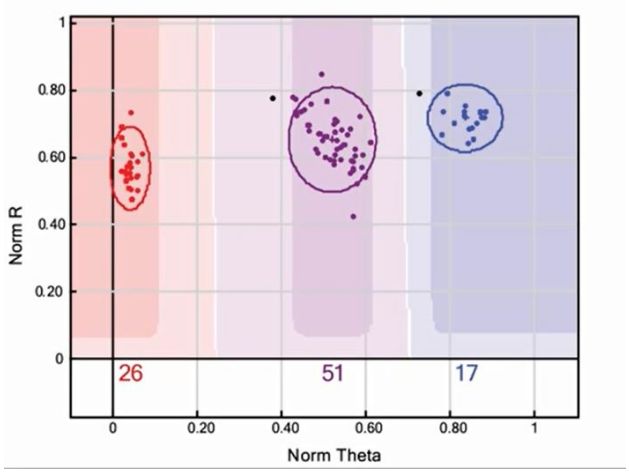


Figure 2-5: Example genotype plot for individual SNP produced by Gen-call (90)

This genotype plots shows the SNP results for all individuals groups as Wild type (red), Heterozygote for the SNP (purple) and Homozygote for the SNP (blue)

Individual SNPs results can then be reviewed and analysed as well as grouped data. Results can also be directly transferred from this programme to PLINK for statistical analysis, this will be discussed later. Results directly read from Genome studio had not been subjected to quality control (QC) measures which will be discussed in subsequent sections of this chapter.

The loading of the Exome BeadChips with the pre-diluted DNA, running of the Beadchips on the genome analyser and transfer of the data to Genome studio were all carried out by a skilled, trained operator and not by myself. This was the only part of the process to this stage that I did not personally complete, as it requires a much greater expertise.

2.14 Real time PCR

Real time PCR (rtPCR) was used to look at some specific candidate SNPs of interest not covered by the Exome BeadChips and also for in house validation of the results of the Exome Beadchip results.

30µl of DNA, which was extracted from whole blood samples (as per protocol 1 Appendix 8) was placed into 96 well plates for use. This allowed for multiple rtPCR tests to be carried out with rtPCR requiring 1µL of DNA per test. The technique for rtPCR is outlined in protocol 4 appendix 12.

2.15 Data analysis

The methodology for analysis of the data will be discussed here. There are several distinct stages to the analysis of the data in this exome genotyping study.

A p value of <0.05 was taken to be significant for candidate gene analyses in this study. Results were subject to Bonferroni correction for multiple comparisons. To accurately estimate significance thresholds when many genetic markers are tested it is essential to control for multiple testing. This will reduce the number of false positive results. Significance thresholds that are accepted as definitive vary in the literature, but most journals accept either 5×10^{-7} prior to correction factors (92). The use of Manhattan plots and quantile-quantile plots (Q-Q plots) allow graphical representation of p values. These allow identification of p values that have deviated from the null hypothesis and hence may be suggestive of a true association (92).

Manhattan plots and Q-Q plots were used to look at the unsupervised analysis where all SNPs on the chips will be analysed.

2.15.1 Descriptive statistics

Descriptive statistics were used to summarize demographic and clinical data collected including event rates for the primary and secondary outcome measure.

Categorical variables are expressed as frequencies and percentages whilst continuous variables are expressed using the central tendency and variability of spread. The mean and standard deviation have been used for normally distributed data and median and range for skewed data.

Bar charts and pie charts have been used to display the data graphically.

2.15.2 Candidate gene analysis

The first review of the results was carried out prior to quality control of the data. This analysis looked only at common SNPs within the cohort and SNP results were taken directly from the genome studio produced by 'Gen-call'. Standard statistical packages were then used to look at the association outlined below. This step was carried out to gain an understanding of the data and analysis steps as well as to gain a general 'feel' for the results. It is however widely accepted that QC is a crucial step in the analysis of such studies and this data will also be analysed following QC of the dataset, with any conclusions being drawn from the post QC data only.

The initial analysis of the candidate genes required binary coding of the data and production of 2 by 2 frequency tables. Three separate frequency tables were produced per SNP for each outcome measure. Results were analysed first in an allelic association model (analysis of allele frequency) and then a genotypic association model (looking at dominant and recessive models).

GraphPad Prism 6 statistical package was used for analysis of the contingency tables. Fisher's exact test was used for the preliminary analysis of this binary data.

All the results that reached statistical significance ($P<0.05$) and those near significance ($P<0.1$) had univariate and multivariate binary logistic regression analysis using SPSS 21 statistical software. Logistic regression is concerned with producing the true probability of the outcome for any given combination of explanatory variables. It was used here to correct for potential confounding factors that may have caused or contributed to the primary outcome events. The results of binary Logistic regression have been expressed as the 'odds' of an event and the 95% confidence interval (CI).

Candidate SNPs were also analysed for time to event data, for the primary outcome events of leucopenia, anaemia and BPAR, with censoring for those individuals who did not have an event by the specified end point of 12 months post transplantation and for death during this study period.

SPSS 21 was used to produce Kaplan-Meier survival plots and non-parametric Log Rank test to compare survival between the groups.

Cox regression analysis was carried out on the time to event data to adjust for confounding variables.

Due to the near complete data (>99.95%) available in the study cohort statistical analysis has not had to account or correct for missing data in the analysis stage.

2.15.3 Quality control of the data set

Both GWAS and exome sequencing studies produce vast quantities of data with challenging statistical analysis. The large numbers of SNPs mean that both genotyping and sequencing errors are common and therefore Quality control (QC) and 'cleaning' steps should be applied to the genotyped data prior to association analysis with measured outcomes(89, 93).

Quality control of the data involves several steps to detect poor quality genotyping, check for gender mismatching, Hardy-Weinberg equilibrium (HWE), related individuals and SNPs within the data set below a pre-specified MAF (below which it is felt that no meaningful association can be sought) (94). Individuals and SNPs that do not pass QC are then removed prior to further analysis (93-95). This will reduce false positive results and improve the statistical power of the study.

Quality control of the genotype data for this study was carried out by a genetic-statistician from GSTS. The QC steps applied will be outlined here. These steps were applied to the entire dataset of over 2400 subjects of which the MPA cohort formed a part. It was felt that applying QC to the entire dataset would reduce genotyping errors more effectively than treating the cohorts separately.

PLINK is a publicly available software for whole genome association analysis designed to handle large datasets (96). It is compatible with genome studio software and was used for a number of steps in both the QC and subsequently association analysis of the data.

The first QC step applied looked at gender mismatch. PLINK offers a gender test using the X-chromosome specific interbreeding coefficient 'F'. This then detects problems when two sexes do not match or if the SNP or pedigrees are ambiguous with regards to sex. The individual is called as male if the X-chromosome 'F' coefficient is more than 0.8 and female if it is less than 0.2, hence females should be close to zero and males close to 1 (96, 97).

Next the genotyped data was investigated for minor allele frequency (MAF) threshold. Using PLINK SNPs with a MAF frequency below a pre-determined level were removed. As this study aims to look at some rare variants with the cohorts a MAF of 0.05 was chosen, this was based on a similar level recommended in comparable studies (88).

Data was then interrogated for 'Missingness' thresholds in relation to call rates and is an informative indicator of sample quality (94). This looks at the fraction of missing calls per SNP and the fraction of missing SNP per individual sample (94). This was done using genome studio and PLINK. A threshold level of >97% call rate for all SNPs per individual and >99% for SNPs with a MAF <0.05 was used, these cuts off for call rates were based on recommendations from The Wellcome Trust case control consortium (98).

Data was then analysed for Hardy-Weinberg Equilibrium (HWE) which states that allele and genotype frequencies remain constant. SNPs out of HWE often indicate a genotyping problem. Genotype clusters that are not well defined often lead to all the individuals being called as the same genotype for that SNP and hence a large departure from HWE. Data was adjusted for HWE using a cut off of $P<10^{-06}$ as this was the level used in similar studies (88). This was applied to the dataset using a command in PLINK (96, 97).

Population stratification was then carried out looking at related individuals. This looks at the average proportion of allele shared as 'Identical by state' (IBS) between two particular individuals. Plink was used to analyse this data which clusters individuals and preformed multidimensional scaling to provide quantitative indices of population genetic variations (96, 97). In large datasets where most individuals are unrelated but belong to roughly homogenous populations, an IBS of approximately 1 indicates a sample duplication or monozygotic twins, 0.5 indicates a first degree relative, 0.25 indicates a secondary degree relative and 0.125 a 3rd degree relative. A threshold of >0.1 was used as this was recommended in similar studies (88). This also tests for and ensures removal of duplicate samples.

2.15.4 Linkage disequilibrium (LD)

Linkage disequilibrium (LD) is the non-random association or correlation between neighbouring alleles (99). SNPs can be genotyped either directly when the causative SNP is studied or indirectly when a SNP which is in LD with a genotyped SNP is the actual causative SNP (92).

SNPs studied in the candidate gene approach were checked for LD using PLINK. SNP analysed in this study were also compared to previously studied SNP not present on the exome chip to look for LD and potential indirect genotyping using pairwise LD in the programme 'SNP Annotation and Proxy search (SNAP) pairwise LD analysis. LD plots were also generated using Haploview, however this relies on the SNP being present in the Hapmap data which was not the case for the majority of the SNPs.

2.15.5 Genotype and SNP calling

The Exome beadchips produced vast quantities of data relating to multiple SNPs. Accurate calling of these SNP is essential to allow robust associations to be sought. This requires sophisticated SNP and genotype calling algorithms using bioinformatics software to reduce and quantify uncertainty (100). The first step in this study uses GenCall developed by illumina, this was predominantly designed to call common variations and hence calling of many rare variants may be limited (91, 101, 102).

One of the advantages of using the exome beadchips is that it allows sequencing of a large number of rare variants and it is recognised that many rare variants can have significant and in some cases profound clinical impact (89). It is therefore important that additional algorithms are applied to aim to correctly call the genotypes. Two options were considered for use in this study, Opticall and ZCall. Both of these calling algorithms have been designed for calling of rare variant, low-frequency genotypes and SNP.

Z call is designed to be applied as a post-processing step after the standard calling algorithm using a linear regression model and has been shown to improve the performance of Gencall by 7%. (101). Opticall uses a combination of within and across sample intensity data to call genotypes across the minor allele frequency spectrum (91). After considering both options it was decided that Zcall would be the most suitable for use in this study. The steps involving Z call were applied by a genetic statistician.

2.15.6 Principal components adjustment

Population stratification refers to systematic differences between individuals that occur as a result of ancestry. If these are not taken into account they can produce spurious associations and hence should be considered a confounder in exome sequencing and GWAS studies (89, 103). Principal component analysis (PCA) is a mathematical procedure that uses orthogonal transformation to convert a set of observations of possibly correlated variables into the continuous linearly uncorrelated variable called principal components. The programme algorithm EIGENSTRAT was applied to all common SNPs (MAF >0.01) to calculate the principal components (PC) accounting for the structure of genetic variation across race. This algorithm has three stages and explicitly models ancestry differences along an axis removing all correlations to ancestry (103). All logistic regression models were adjusted for the first four principle components using PLINK.

2.15.7 Association analysis

Following QC of the data PLINK was used to look for associations between SNP and outcome measures. All outcomes measures were binary coded.

This was carried out for candidate genes, extended candidate genes and for an unsupervised analysis looking at all SNPs.

Significant SNPs were checked for linkage disequilibrium, as described previously, using PLINK.

The commands used in PLINK are shown in Appendix 13 protocol 5 which shows a full run through of the stages involved in data analysis in this study.

2.15.8 Logistic regression

Logistic regression of the data required a two phase approach as it must be considered whether the confounders that are to be adjusted for are relevant to the genotype or the outcome measure (phenotype).

The first phase was to adjust for the PC of the population. This was carried out in PLINK as this confounder relates to genotype. PC analysis was run for the entire cohort and then for a Caucasian only cohort (see appendix 13). An allele frequency model, dominant model and recessive model were run in each case.

Following logistic regression for PC's, the significant results were then subjected to Logistic regression for patient related factors. This used SPSS V21 linear regression as the factors relate to patient and clinical outcomes and not the genotype.

2.15.9 Time to event analysis

Time to event analysis was carried out using Kaplan Meier curves and Cox regression analysis using SPSS V21. This analysis was only applied to the primary outcome measures of Anaemia, Leucopenia and BPAR for SNPs that were significant for that outcome following Logistic regression.

2.16 Data storage and protection

The patient data were stored on a secure password-protected NHS computer accessible to the investigators of the study only. All blood samples were anonymised but allocated a unique number, the details of which were securely stored on the computer and backed up to the secure Trust network. Confidential patient data were not stored on laptops or portable storage devices such as memory sticks. Where it was necessary to transfer data between the clinical care and research team, this was done by email over the Trust's secure Intranet or NHS network.

Consent forms which included patient identification but no clinical details, and data collection sheets including clinical and patient details were stored in a locked filing cabinet in the Wessex Renal Unit. All other data were securely stored on the NHS computer.

DNA, blood samples and confidential patient information will be stored for 5 years after the conclusion of the study and will then be destroyed. Anonymous patient and laboratory data will be stored for 10 years after conclusion of the study.

2.17 Time frame

At the beginning of this study a three year time frame was set for this project. This was defined by a contracted out of programme time for myself (the primary investigator), a university time frame for completion of a DM and funding, which will not extend beyond this period. It was a realistic time frame to achieve the recruitment targets allowing time for completion of data collection, laboratory and statistical analysis.

2.18 Funding

Funding for this project was secured both for the wage of the primary investigator, laboratory techniques and equipment. Funding was provided through charitable funds dedicated to research into kidney disease and pharmacogenetics. Both The Wessex Renal and Transplant Research Fund and Guy's and St Thomas's Research Charity have contributed hugely to meet the cost of this study. Further external funding was sought from Kidney Research UK (KRUK) and the British Renal Society (BRS) but was unfortunately unsuccessful although the feedback on the study from these sources was favourable.

2.19 Declaration

All aspects of this study from application for ethical approval, subject identification, recruitment, data collection, laboratory work (with the exception of Exome BeadChip analysis), analysis of data, statistical analysis, (including analysis using PLINK) were carried out by myself the Primary Investigator and the author of this thesis. Quality control of the data set, use of the rare variant caller software Z call and Opticall and calculation of the population principle components were conducted on the entire GSTS patient set of 2400 individuals by expert statisticians. Appropriate support was sought and provided at various stages by a number of skilled individuals to teach and train me in the various techniques required to complete this research (see Acknowledgements).

3 Chapter 3: The study population, baseline characteristics and event rates

3.1 Introduction

This chapter will outline the study population included in this research. It will present and discuss the baseline characteristics of the study population and the observed event rates relating to the primary and secondary outcome measures.

3.2 The study population

This study was designed to identify and investigate genetic SNPs, both candidate and novel, and study their association with patients' response to the immunosuppressive drug mycophenolate (mofetil or sodium) following renal transplantation. The study was designed to be as inclusive as possible with few exclusion criteria. This aimed to recruit a study population that was representative of the routine recruitment in this unit and ultimately the UK transplant population. Due to the non-interventional nature of this study enrolment in other trials was not an exclusion criterion. The design of the trial should produce results that have validity and hence are generalizable, which are important factors in the potential clinical application of results.

3.3 Screening and recruitment

At the start of the study 593 transplant patients had been identified and these were subsequently screened for inclusion and exclusion criteria. During the 18 month recruitment period a further 111 new transplant patients were screened for eligibility. Recruitment stopped in July 2012 to allow a 12 month follow up period for all recruits. Whilst transplantation in a different renal unit or outside of the UK were not part of the initial exclusion criteria these individuals were not approached for participation as vital clinical data required for analysis would not be available to the researcher and hence the data set would be incomplete which could affect the power and validity of the results. The flowchart of screening and recruitment is outlined in figure 3.1. There was a very high recruitment rate for subjects invited to participate with only 3 subjects (1%) declining to take part. Two individuals declined as they did not wish to take part in any form of research and one was deemed not to have capacity to comprehend the available information due to learning difficulties and it was felt unethical to proceed to recruitment in this case. The final study population consisted of 287 subjects.

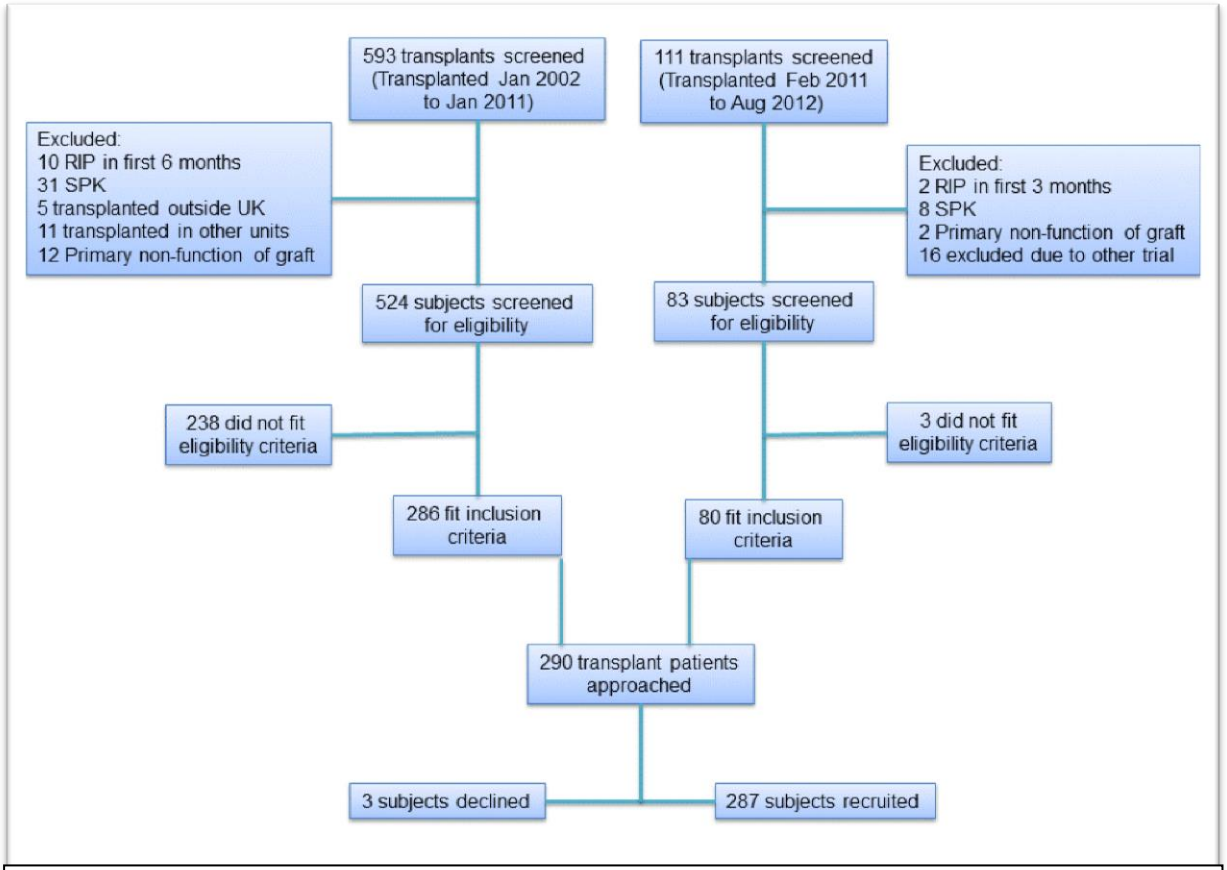


Figure 3-1: Flow chart of screening and recruitment of the final study population

The left hand side of the flowchart shows those patients identified at the beginning of the study as potentially eligible, the right hand side includes those patients newly transplanted during the study period. 241 individual did not meet the eligibility criteria as they did not receive full dose MPA at the time of transplantation. A total of 290 individuals were approached to take part in the study, with only 3 declining to participate. The total number recruited was 287.

3.4 Discussion

The final numbers recruited were smaller than the target, but the study remains powered as discussed in the methods chapter. The number of individuals declining to take part was very low as discussed above. As outlined in figure 1 there were 366 patients identified as eligible for recruitment into this study but only 287 recruited. There were a number of issues around recruitment of the remaining patients which meant the full 366 could not be achieved within the time available. Twelve patients in the retrospective cohort died between the time they were identified as eligible and recruited. These participants were all greater than 1 year post transplant and hence it was not felt that this would introduce bias into the results. The recruiting unit covers a large geographical area as outlined in the methods chapter. Once patients are greater than 3

months post transplantation they return to peripheral clinics within their locality. There are 6 peripheral clinics including the Isle of Wight and due to the time available and a single recruiter (myself), it was not possible to capture all individuals in peripheral clinics. Every effort was made to include as many of these individuals as possible with each peripheral clinic being visited on at least 2 occasions when there were potential recruits attending. Whilst it is recognised that the recruitment from peripheral clinics was incomplete, the likelihood of a bias is low, as the majority of individuals transplanted during the study period were captured in the main unit during the first 3 months. The final study population was therefore representative of the recruiting renal unit and is outlined in this Chapter.

3.5 Baseline characteristics

The baseline characteristics of the final study population comprising 287 renal transplant recipients transplanted at the Wessex Renal and Transplantation Unit between January 2002 and July 2012 are summarised in table 3.1.

Table 3-1: Baseline characteristics of the study population

All participants N=287			
		Number	Percentage
Gender	Male	175	61%
	Female	112	39%
Age when transplanted	Mean	47.1 years	
	Median	47 Years	
	Range	17 to 79 years	
Ethnicity	Caucasian	268	93.4%
	Other	19	6.6%
Type of transplant	Live donor	74	25.8%
	Cadaveric donor	213	74.2%
Cause of ESRF	Diabetes	22	7.7%
	Glomerular Nephritis (GN)	74	25.8%
	Autoimmune/systemic disease	14	4.9%
	Hypertension	18	6.3%
	Polycystic kidney disease (PKD)	41	14.3%
	Congenital/ Familial/Metabolic	60	20.9%
	Obstructive	11	3.8%
	Tubular and interstitial disease	6	2.1%
	Renovascular (RVD)/Ischaemia	5	1.7%
	Other	36	12.5%
Number of previous transplants	None	212	73.9%
	1	55	19.2%
	2	17	5.9 %
	3 or 4	3	1.1%
HLA mismatches	None	56	19.5%
	1 or 2	87	30.3%
	3 or 4	118	41.1%
	5 or 6	23	8.0%
	Data Missing	3	1.1%
CMV status (Donor/Recipient)	neg/neg	87	30.3%
	neg/pos	71	24.7%
	pos/neg	55	19.2%
	pos/pos	73	25.4%
	Data Missing	1	0.4%
Induction agent received	None	84	29.3%
	Basiliximab	196	68.3%
	Other	5	1.7%
	Data Missing	2	0.7%
Initial CNI	Tacrolimus	103	35.9%
	Ciclosporine A	184	64.1%
Initial MPA	MMF	262	91.3%
	Myfortic	25	8.7%
Valgancyclovir prophylaxis	Received Valgancyclovir prophylaxis	59	20.6%
	No Valgancyclovir prophylaxis	228	79.4%

Descriptive statics using SPSS version 21 have been used for analysis of the baseline characteristics.

3.5.1 Year of Transplantation

The years in which the subjects received their transplantation is shown in figure 3.2.

It is important to take this into consideration due to the changing trends in immunosuppression, rejection rates and cause of ESRF over the last decade.

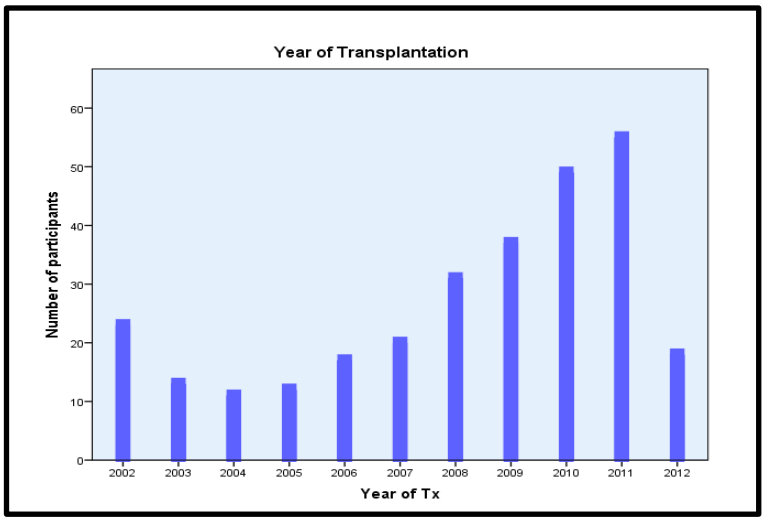


Figure 3-2: Histogram to show year of transplantation

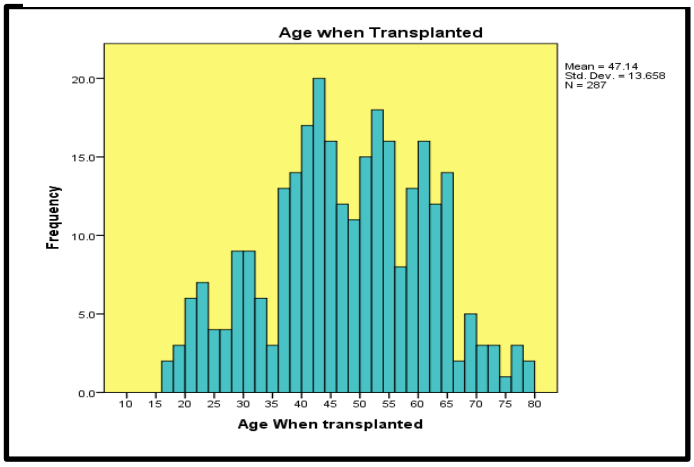
3.5.2 Age and gender of the study population

The median age of the study population at the time of transplantation is 47 years (range 17 to 79 years). This is summarised in table 2.2 and is also displayed graphically in a histogram to show the spread of data figure 3.3.

Figure 3-3: Histogram to shown age when transplanted

Table 3-2: Age when transplanted

Number	Valid	287
	Missing	0
Mean		47.14
Median		47.00
Mode		40
Std. Deviation		13.658
Range		62
Minimum		17
Maximum		79



The study population comprised 61% male and 39% female.

The recruited population in this study have a similar median age at transplantation and male to female ratio as the incident UK transplant population. The similarity of these characteristics will help with generalisation to the UK population. They are also similar to other MPA pharmacogenomics studies in renal transplant receipts (6, 74, 79, 104).

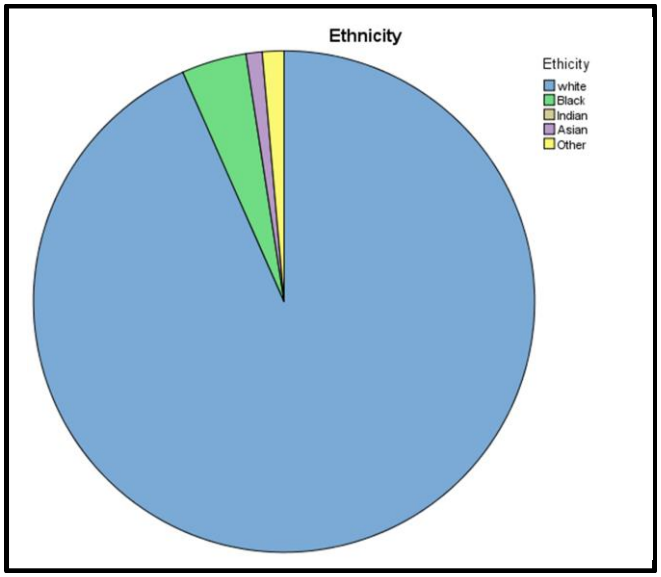
Table 3.3 compares the median age at transplantation and the male to female ratio in the study population compared to the UK incident transplant population 2011 using data from the 15th Renal Registry report (105).

Table 3-3: Age when transplanted and male: Female ratio in study population and UK incident transplant population

	Median age when transplanted	Male: Female ratio
Study population	47 years	1.56
UK (2011)	49 years	1.7

3.5.3 Ethnicity

The ethnicity of the study population was predominantly Caucasian the overall break down of ethnicity is shown in figure 3.4.



	Frequency	Percent
white	268	93.4
Black	12	4.2
Asian	3	1.0
Other	4	1.4
Total	287	100.0

Figure 3-4: Ethnicity of study group

It can be seen that the vast majority of individuals recruited in this study were of Caucasian origin, reflective of the population within the recruiting centre.

The study population has a lower ethnic diversity than the incident UK transplant population. This reflects the ethnicity within the recruitment area as discussed in the methods chapter. Whilst this

will affect generalizability of the study results to the UK population in its entirety, it is actually beneficial in this pharmacogenomic study due to the variability of SNP MAF in differing ethnic populations.

Table 3.4 below compares the ethnicity of the study population compared to the UK incident transplant population 2011 using data from the 15th renal registry report (105).

Table 3-4: Ethnicity of the study population compared to the UK incident transplant population 2011

	White	South Asian	Black	Other	Unknown
Study Population	93.4%	1%	4.2%	1.4%	N/A
UK	72.6%	9.3%	6.6%	2.1%	9.5%

3.5.4 Cause of ESRF

The cause of ESRF leading to the need for transplantation is displayed in figure 3.5.

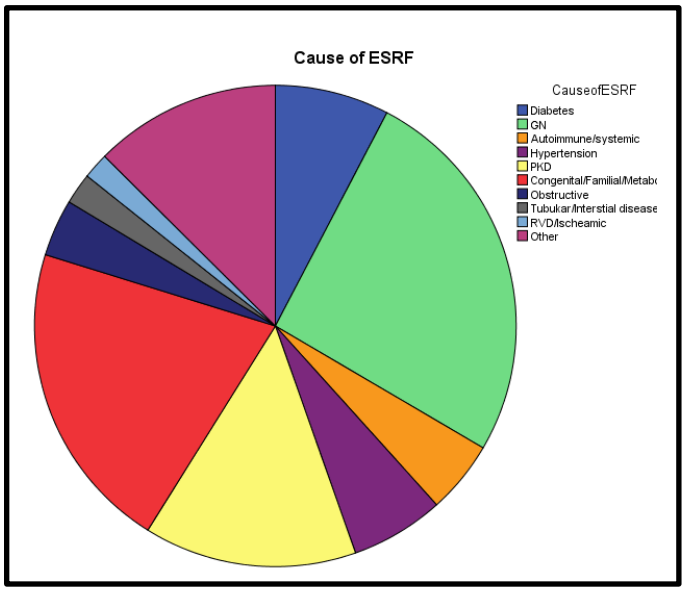


Figure 3-5: Cause of ESRF

	Frequency	Percent
Diabetes	22	7.7
GN	74	25.8
Autoimmune/systemic	14	4.9
Hypertension	18	6.3
PKD	41	14.3
Congenital/Familial/ Metabolic	60	20.9
Obstructive	11	3.8
Tubular/Interstitial disease	6	2.1
RVD/Ischaemic	5	1.7
Other	36	12.5
Total	287	100.0

The cause of end stage renal failure in the individuals recruited into this study in shown in this pie chart, with the exact numbers shown in the table.

Table 3.5 compares the cause of ESFR in the study population and incident UK transplant population 2011 using data from the 15th Renal Registry report (105)

Table 3-5: Cause of ESFR in the study population and incident UK transplant population

	Diabetes	GN	PKD	RVD
Study Population	7.7%	25.8%	14.3%	8.01%
UK (2011)	11.9%	23%	12.2%	6.5%

The cause of ESRF for the 4 most common categories have been presented here. The study population has very similar rates to the incident UK transplant population. The rate of diabetes is slightly lower in the study population and this is likely to reflect changing trends over the last 10 years with diabetes as a cause of ESRF in the incident transplant population rising from 9% in 2002 to 11.9% in 2011(105, 106).

3.5.5 Type of transplant received, number of HLA mismatches and previous transplantation

The donor type and compatibility of the transplant received has a significant impact on factors both in the early days post transplantation and the long term outcome of the graft. This has been demonstrated yearly in the UK renal registry report (105). Transplant recipients receive one of three broad types of transplantation:

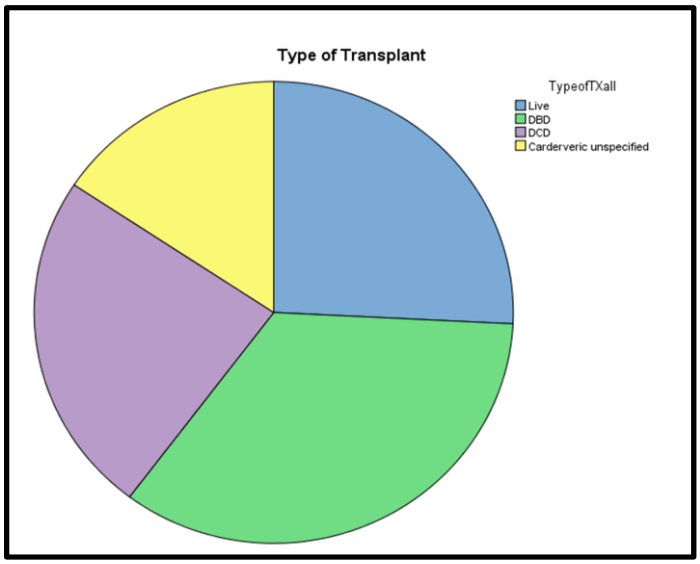
Living donor transplant (LDT) - this may be from a known related or unrelated donor, an altruistic donor or from a pooled matching scheme.

Deceased after cardiac death (DCD) - Previously referred to non-heart beating donors.

Deceased after brain stem death (DBD) - Previously referred to as heart beating donors.

It is well understood that LDT has the best graft and patient outcomes with 5 year graft survival rates of >92% (107) compared to around 70% for DCD and DBD(108). Delayed graft function (DGF) is also much lower in LDT at around 10% (109, 110) compared to 42-51% for DCD and 24% for DBD(108). The number of HLA miss-matches (HLA-MM) has also been well documented to be associated with acute rejection rates and poorer graft outcome (111). The donor type and match of transplant received is likely to be an important confounding factor and so will be accounted for in the analysis.

The type of transplant received by the study population is outlined in figure 3.6:



	Frequency	Percent
Live	74	25.8
DBD	99	34.5
DCD	69	24.0
Cadaveric unspecified	45	15.7
Total	287	100.0

Figure 3-6: Type of transplant received

The type of transplant received by the individual sin this study are shown in this pie chart, the blue segment shows the living related transplants, the remaining transplants were cadaveric the actual numbers are shown in the table.

Table 3.6 compares the donor type in the study population to the UK incident population 2011 using data from Data from 15th renal registry report chapter 3 (105).

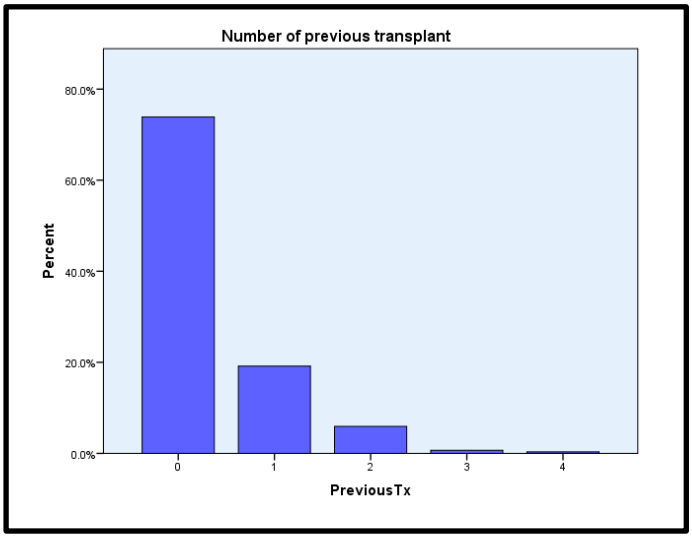
Table 3-6: Type of kidney donor in study population and UK incident population

	% Cadaveric kidney donor			% Live kidney donor
	DBD	DCD	Unknown	
Study population	34.5%	24%	15.7%	25.8%
UK (2011)	40%	23.1		39.9%

The recruited population in this study show a higher percentage of cadaveric versus live kidney donors than the 2011 UK incident transplant population. This is likely to be due to the fact that the study recruits were transplanted in the last 10 years and during this time there have been an increasing number of live donations with a 65% increase from 2000 to 2010. This is due to the growth of the living donor programme including pooled/paired donations and HLA or ABO incompatible transplants and an increase in altruistic donations (112, 113). It is not felt that this will affect the validity of the results but the type of transplant will be taken into consideration as a potential confounding factor when results are analysed. The type of transplant received is frequently not reported in similar pharmacogenomic studies but one large North American study

reported a much higher rate of living donors at 59% (79) with a European study reporting living donor rates of 32.5%, similar to the UK (6).

The number of previous transplants in the study population ranged from 0 to 4, with the majority of recruits (73.9%) being first transplants, similar rates have been published in other studies (66) but one large study with 86% being first transplants (79). The actual data for the study population are shown in figure 3.7.

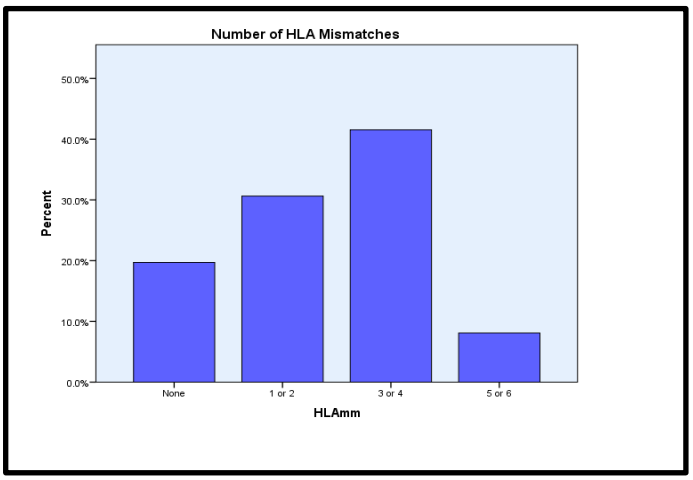


	Frequency	Percent
0	212	73.9
1	55	19.2
2	17	5.9
3	2	.7
4	1	.3
Total	287	100.0

Figure 3-7: Number of previous transplants

The number of previous transplants received by individuals is shown in this bar chart

The number of HLA MM between the donor and recipient are shown in figure 3.8. Comparable data has been presented in the literature in similar studies (114).



	Frequency	Percent
None	56	19.5
1 or 2	87	30.3
3 or 4	118	41.1
5 or 6	23	8.0
Total	284	99.0
Missing	3	1.0
Total	287	100.0

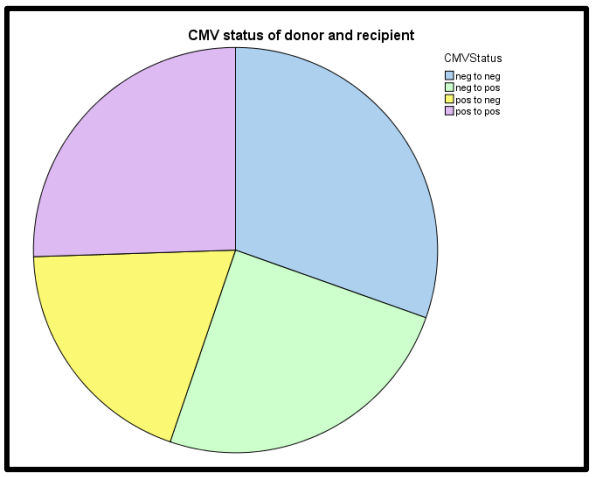
Figure 3-8: Number of HLA miss-matches

The number of HLA mismatches between the recipient and the donor kidney is shown in this bar chart.

3.5.6 CMV status and Prophylaxis

The CMV status of donor and recipient are always checked prior to transplantation. There is substantial evidence that the development of CMV infection post transplantation is associated with increase rejection rates, worse graft outcome and significant mortality (115-117). A recent Cochrane review of 37 studies concluded that CMV prophylaxis should be used in all CMV positive recipients and in all CMV negative recipients in whom the donor is CMV positive(118). The current practice in the recruiting unit is to give Valganciclovir CMV prophylaxis to CMV negative recipients of a CMV positive kidney.

The CMV status of the donor and recipient are displayed in figure 3.9. 20.6% of patients in this study received CMV prophylaxis.



	Frequency	Percent
neg to neg	87	30.3
neg to pos	71	24.7
pos to neg	55	19.2
pos to pos	73	25.4
Total	286	99.7
Missing	1	.3
Total	287	100.0

Figure 3-9: CMV status of donor and recipient

All subjects recruited into the study received a CNI, MPA, steroids and in some cases an induction agent at the time of transplantation. 64.1% of subjects received Ciclosporine A and 35.9% received tacrolimus, 91.3% received MMF with 8.7% receiving MPS, 68.3% received basiliximab induction agent (see table 3.1). The type of CNI and MPA used varies throughout similar studies in the literature (6, 8, 66, 71, 119). This represents changing practice as increasing evidence suggest that tacrolimus significantly reduces graft loss (120) and that MPS may reduce the burden of GISE compared to MMF (121). It is recognised that the trend in immunosuppression use has altered throughout the inclusion period with the recruiting unit now using more tacrolimus and MPS than previously. An induction agent is often used at the time of transplantation; this practice is also becoming more common. These factors need to be considered when analysing results.

3.6 Summary

The baseline characteristics of the study population are comparable to both the UK population and to those of similar studies reported, meaning the study has validity within the transplant population. Variations and potential reasons for these have been discussed. All baseline characteristics will be taken into account when analysing the study results aiming to adjust for confounding variables.

3.7 Event rates

The primary and secondary outcome event rates observed in the study group in the first 12 months post transplantation are summarised in table 3.7.

Table 3-7: Observed event rates in the study subjects including the definition used to define these events.

Event	Definition of event	Number patients with an observed event (287 patients total)	Percentage of study population event observed in
Biopsy proven rejection	Evidence of acute rejection reported on transplant biopsy	54	18.8%
Vascular or cellular rejection	Evidence of acute vascular or cellular rejection reported on transplant biopsy	33	11.5%
Any change to MPA	Any change in dose, frequency or preparation of MPA for any reason	166	57.8%
Stopped MPA	Patients stopped MPA completely and it was not successfully reintroduced	27	9.4%
Any GISE	Any upper or Lower GI symptoms felt by the treating clinician to be related to MPA and requiring some alteration to MPA	84	29.3%
Any Upper GI side effects	Any upper GI symptoms felt by the treating clinician to be related to MPA and requiring alteration to MPA	31	10.8%
Any Lower GI Side effects	Any Lower GI symptoms felt by the treating clinician to be related to MPA and requiring some alteration to MPA	65	22.7%
Leucopenia	A reduction in wcc to <3.0 on two or more consecutive blood test or a single test if this resulted in a change in MPA dose	74	25.8%
Anaemia	A reduction in Hb to <10g/dL after day 30 on two or more consecutive blood test	99	34.5%
Infective episode	Recurrent or severe infection felt by the treating clinician to be as a result of the patients immunosuppressed state (Excluding recurrent UTI)	62	21.6%

3.8 Mycophenolate reduction or withdrawal

The first primary outcome measure discussed here is reduction or withdrawal of MPA. 166 patients (57.8%) had a change to MPA during the first year post-transplant. The percentage that stopped MPA completely was 9.1% (26 patients). Individuals that temporarily stopped MPA but tolerated a reintroduction were counted as a reduction rather than withdrawal. Individuals that had a stepwise reduction in MPA and then permanent withdrawal were counted as stopping MPA.

17 patients had more than one change to MPA, this included a reduction and change in preparation, or a change in preparation and then stopping MPA. The number of changes made in shown in figure 3.10 and the proportional changes made in figure 3.11.

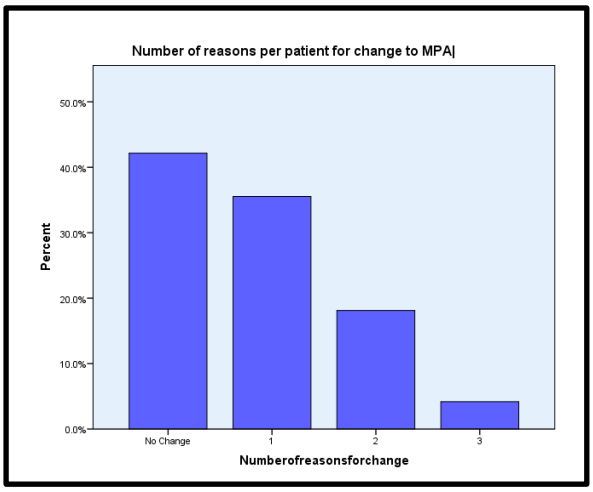


Figure 3-10: No of reasons for change to MPA

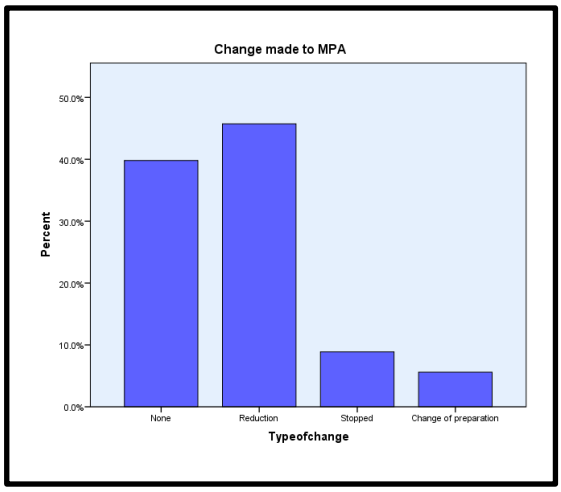
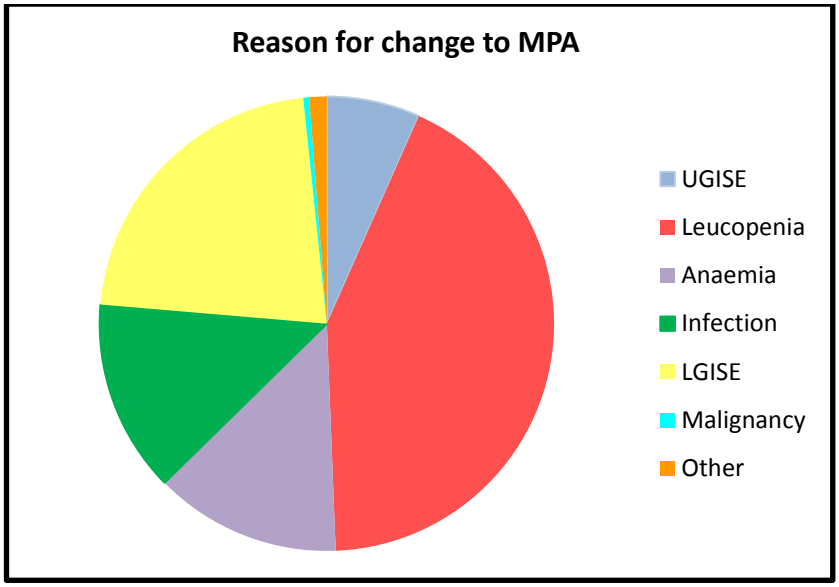


Figure 3-11: Change made to MPA

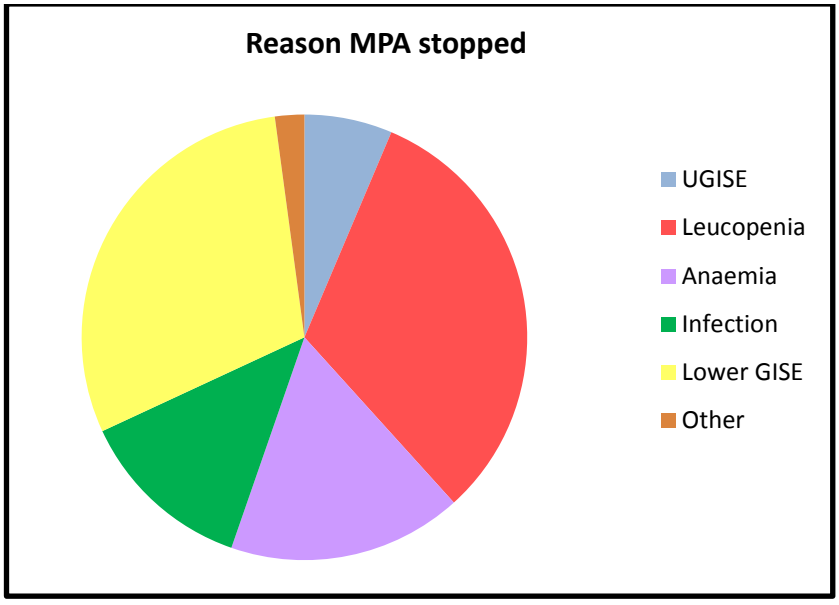
The reasons why patients had MPA changes are outlined in figure 3.12. The commonest reason was leucopenia accounting for 42.7% of changes with GISE being the second most common reason accounting for 28.6% and infections 13.4%. A 2013 study of reasons for dose reduction in MPA in the first year post transplant reported 48.7% of subjects had a dose reduction, 48.7% due to haematological toxicity, 16.1% due to infection and 12.3% due to GISE (122).



	Frequency
Upper GISE	16
Leucopenia	103
Anaemia	32
Infection	33
Lower GISE	53
Malignancy	1
Other	3

Figure 3-12: Reasons for change to MPA are shown in this pie chart.

The reasons for stopping MPA are outlined in figure 3.13, with leucopenia being the commonest reason to withdraw MPA at 31.9% and lower GISE being the second commonest at 29.8%.



	Frequency
Upper GISE	3
Leucopenia	15
Anaemia	8
Infection	6
Lower GISE	14
Other	1

Figure 3-13: Reason for stopping MPA is shown in this pie chart

3.9 Leucopenia

Leucopenia is common in the MPA treated transplant population and can predispose to serious infections. It is the commonest reason for reduction or withdrawal of MPA. Leucopenia was observed in 74 individuals (25.8%).

Both the definition of leucopenia and the event rate post transplantation varies throughout the literature. Table 1-8 below compares the results in this study with similar published MPA

pharmacogenomics studies. The events rates observed in this study are most similar to the large 2011 study by Jacobson et al (79) and share the same definition of leucopenia.

Table 3-8: Leucopenia event rates comparison

Study	Number	Definition	Event Rate
This study	287	A reduction in Wcc to <3.0x10^9/L on two or more consecutive blood test or a single test if this resulted in a change in MPA dose	25.8%
Wang et al 2008(48)	191	Wcc to <3.0x10^9/L	32%
Prausa et al 2009(67)	38 Paediatric	Wcc to <2.5x10^9/L or a steady decline towards this level	42% leucopenia and/or diarrhoea
Michelon 2010(8)	218	Wcc<4x10^9/l and relieved with MPA reduction or interruption	38.1%
Bouamar 2012(80)	332	Wcc<3x10^9/L	16.9%
Jacobson et al 2011(79)	978	Wcc to <3.0x10^9/L in first 6 months post tx	22.9%

The time to event of leucopenia in days is shown in the Kaplan-Meier survival curve in figure 3.14. This shows that more leucopenia was observed in the first 4 months (160 days). Individuals were censored for death or if they reached 12 months without developing leucopenia.

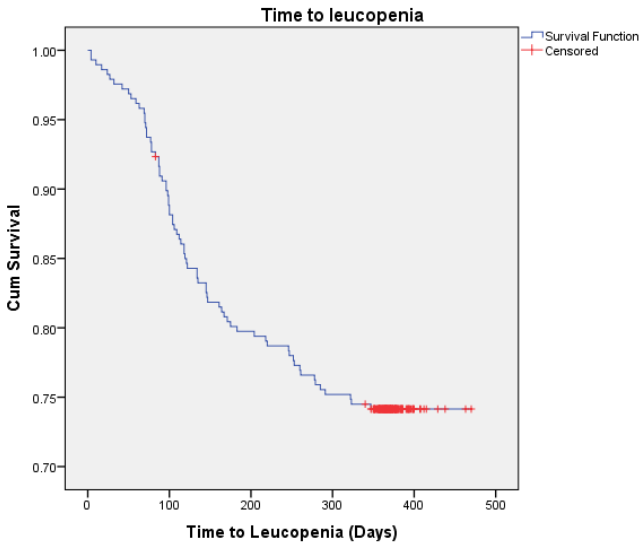


Figure 3-14: Kaplan-Meier survival curve time to Leucopenia

This Km curve shows the time to leucopenia, subjects were censored for death or reaching the end of study without development of leucopenia.

3.10 Anaemia

Anaemia was observed in 99 individuals (24.5%).

Both the definition of anaemia and the event rate post transplantation varies throughout the literature. A large 2011 study found the prevalence of post-transplant anaemia to be 52.7% when the standard definition of anaemia was used (Hb<120g/L for women and <13g/dL for men) and 24.4% with a cut off <110g/L (123). The difficulties of defining anaemia in the renal transplant population and deciding on cut off levels for treatment have been well documented (124, 125) the multifactorial nature of the anaemia causes further difficulty. Table 3.9 compares the results in this study with similar MPA pharmacogenomics studies in the literature. The event rates observed in this study are most similar to 2012 study by Bouamar et al (80) but the cut off defined as anaemia is different. A low level of Hb was used in the definition for this study, a level below which intervention and investigation would certainly take place. It was felt that using this level would pick up the more severe cases that would have clinical relevance. The event rates observed in this study are significantly higher than those reported in Jacobson et al 2011 study (79) Confounding factors for anaemia have been included in the data collection and will be used for statistical analysis.

Table 3-9: Anaemia event rates comparison

Study	Number of recruits	Definition	Event Rate
This study	287	A reduction in Hb to <100g/L after day 30 on two or more consecutive blood test	34.5%
Michelon 2010(8)	218	Hb <120g/l and relieved with MPA reduction or interruption	12.5%
Bouamar 2012(80)	332	Hb<113g/L after day 28 post-transplant	38%
Jacobson et al 2011(79)	918	Hb<100g/L resulting in MPA dose reduction or discontinuation or Epo therapy after day 30	9.5%

Time to event of anaemia is shown in the Kaplan-Meier survival curve in figure 3.15. Anaemia is the first 30 days post-transplant was presumed to be a combination of blood loss from the surgical procedure and in some cases, withdrawal of EPO therapy prior to full graft function. The majority of anaemia events occurred between day 30 and day 120. Individuals were censored for death or reaching 12 months without developing anaemia.

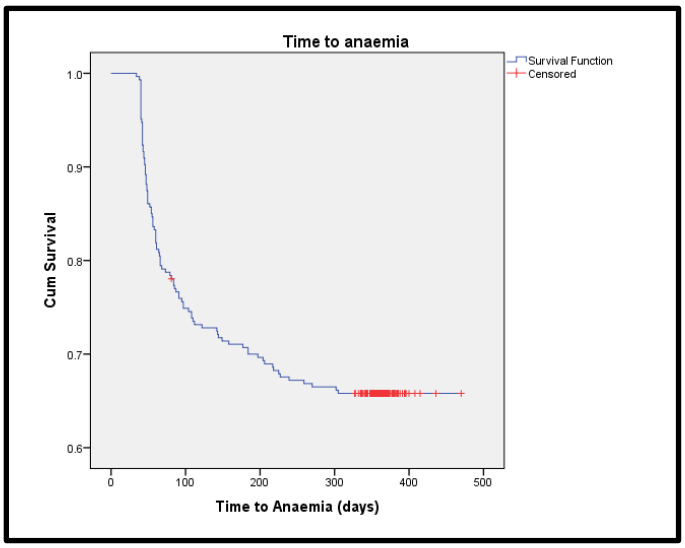


Figure 3-15: Kaplan-Meier survival curve time to Anaemia

This Km curve shows time to anaemia, patients were censored for death or reaching the end of the study event free.

3.11 Biopsy proven acute rejection BPAR

Biopsy proven acute rejection BPAR was sub-categorised in the study group as 'All' (any reported rejection including borderline changes) and 'Vascular or cellular rejection' (reported as evidence of vascular or humoral rejection). Although the reporting histopathologists in the recruiting unit work according to Banff criteria classification of renal allograph pathology (Appendix 14) (86) they do not use the Banff grading in the reports. Many studies use a Banff cut off criteria to define clinical or subclinical rejection and give a grade of rejection. This was not possible in this study as the unit biopsy reports did not allow for this. The decision was made to include any rejection in the 'All' group, this included cases of 'borderline changes' (Banff 3) these cases are often excluded (when occurring alone) from studies where the Banff criteria is used. It is recognised that some of these cases are of clinical significance and hence they have been included in this group. Patients who were reported as having evidence of cellular/humoral or vascular rejection (Banff 2 or 4) on biopsy were subcategorised and labelled as 'vascular or cellular rejection'. The histopathologists were not aware of the study and all biopsy reports were reviewed retrospectively. The protocol at the recruiting unit is to biopsy all patients with delayed graft function at 7 days post transplantation and then weekly until the graft functions, other patients are biopsied only when there is a clinical need such as an increasing creatinine.

BPAR in the first 12 months was observed in 54 (18.8%) of subjects in the 'All' category and 33 (11.5%) in the 'vascular or cellular rejection' category (a subset of BPAR all as discussed in the methods chapter).

Table 3.10 compares the results in this study with similar MPA pharmacogenomics studies in the literature. The BPAR for all biopsies were similar to the rates observed in Michelon et al 2010 study (8). Rejection rates have improved considerably with the introduction of newer immunosuppressive regimens. The current acute rejection rate is between 10 and 15%. This changing reject rate may account for some of the variation in observed event rates.

Table 3-10: BPAR event rates comparison

Study	Number of recruits	Definition	Event Rate
This study	287	BPAR reported by a renal histopathologist using Banff criteria and unaware of the study. **All**- any including borderline change **Vasc or cell**- evidence of vascular or cellular rejection	All- 18.8% Vasc or cell- 11.5%
Wang et al 2008(48)	191	Biopsy proven acute rejection unspecified	15%
Kagaya et al 2010 (73)	82	Subclinical BPAR- Banff 1A or greater on day 29 biopsy	25.6%
Van Schaik 2009(65)	338	BPAR at 1 and 12 months post-transplant.	1 month- 9.2% 12 months- 15%
Michelon 2010(8)	218	Any BPAR in first year according to Banff 2005 criteria	20.5%

Time to rejection episodes are shown in Kaplan-Meier survival curves in figure 3.16. The majority of rejection episodes occur in the early post-transplant period with few episodes occurring after the first 2 months (60 days).

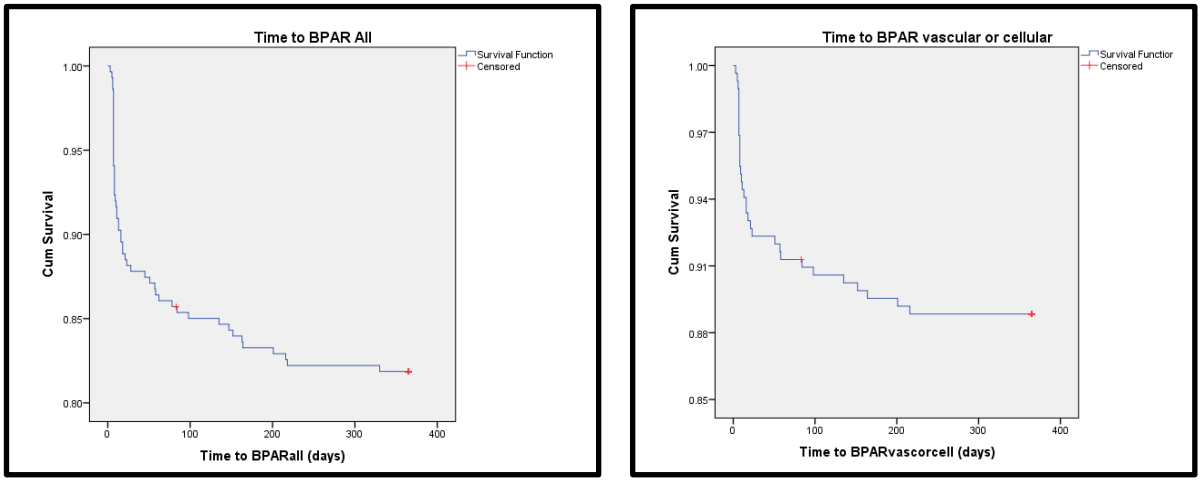


Figure 3-16: Kaplan-Meier survival curve time to BPAR

These KM curves show time to BPAR all (left hand curve) Vascular or cellular subset (Right hand curve), Individuals were censored for death.

3.11.1 Gastro-intestinal side effects (GISE)

Gastrointestinal side effects are common in patients treated with MPA occurring in up to 45% of MMF treated patients, with 20-40% requiring dose reduction or withdrawal due to GISE (2, 3). The type of GI toxicity varies, but diarrhoea, nausea, bloating and gastritis are the most common.

Enteric coated mycophenolate sodium (EC-MPS) has been shown to produce less GISE without compromising patient and graft survival (126, 127) and is increasingly used.

GISE occurred in 29.3% of the study population. GISE were then subcategorised into upper GISE (nausea/ vomiting/ reflux/ gastritis) shown in figure 28 and Lower GISE (diarrhoea/ bloating/ pain).

Table 3-11 compared the frequency of GISE with other similar studies. The frequency of GISE within the study population was similar to those reported in other MPA pharmacogenetic studies and to the rates observed in the initial tolerability studies as stated above.

Table 3-11: GISE event rates comparison

Study	Number of recruits	Definition	Event Rate
This study	287	Any upper GI symptoms felt by the treating clinician to be related to MPA and requiring alteration to MPA **Upper**- Nausea/ vomiting/ reflux/ gastritis **Lower**- Diarrhoea/ Bloating/ Pain	All 29.3% Upper GISE 10.8% Lower GISE 22.6%
Prausa et al 2009(67)	38 Paediatric	Several loose stool a day defined by patient/ family as a change	42% leucopenia and/or diarrhoea
Woillard et al 2010(104)	256	Diarrhoea, abdominal pain, nausea/vomiting and anorexia	All GISE – 35.1% Diarrhoea- 27.7% Ando pain – 11% Upper GISE – 12.5%
Michelon 2010(8)	218	Diarrhoea more than two episodes a day other causes excluded and relieved after MPA reduction or interruption	33.1%
Bouamar 2012(6)	338	More than 4 loose stools per day which was a change from baseline	23%

3.11.2 Infections

Infections remain an important cause of morbidity and mortality in the transplant population. Although infection rates have improved over the years the potent immunosuppressive regimens means they still account for numerous inpatient hospital days. The type of infection varies depending on the time since transplantation, with hospital acquired or donor derived infections being common in the first month, and viral, chest and urinary infections becoming more likely thereafter (128). In the first year post transplantation up to 45% of patients will suffer a significant

infective episode with urinary tract infection being the most common, viral infection the second most common and pneumonia accounting for 70% of infection related deaths (129, 130).

Infections were observed in 21.6% of participants in this study during the first 12 months.

204 (71.1%) patients had no significant infection episodes documented during the first year post-transplantation. Three (1%) patients had two different significant infective episodes and 1(0.35%) patient had 3 different infective episodes during the first year. If the same infection recurred within a short time frame without complete improvement in between it was not counted as a separate episode in this analysis. 22 (7.7%) of patients suffered from recurrent urinary tract infections, whilst these were noted they were not counted in the analysis as 'significant infective events' unless these resulted in a septic episode. There were 7 septicaemia episodes during the first year and 9 episodes of pneumonia. 28 (9.8%) of patients developed CMV infection requiring treatment. Two individuals (0.7%) developed Varicella Zoster virus (VZV) and 9 (3.1%) developed significant Herpes simplex virus (HSV) although none of these episodes were systemic. The infective episodes are shown in figure 3.17.

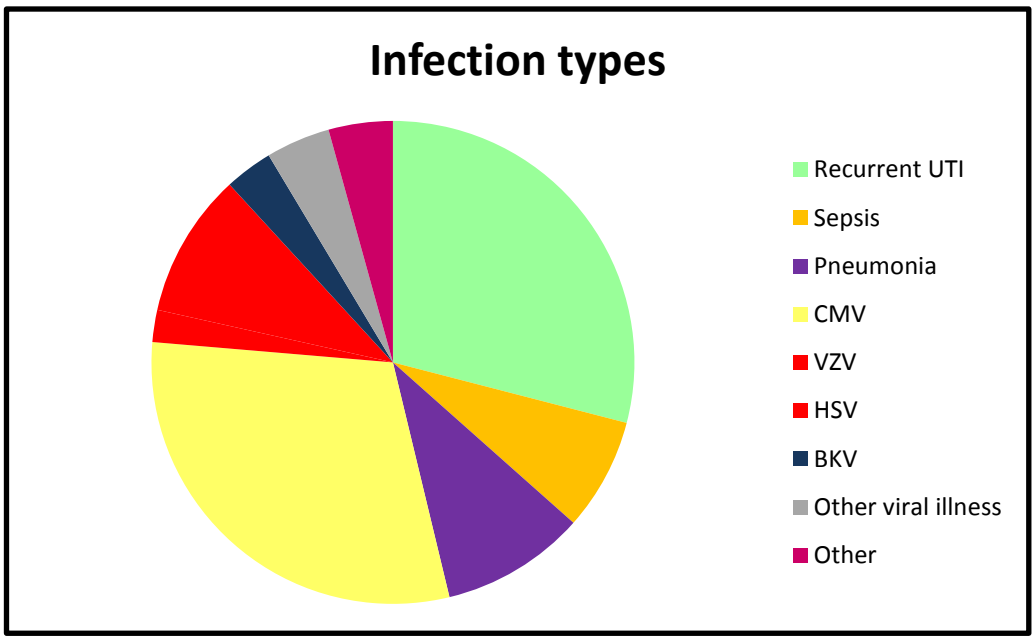


Figure 3-17: Infective episodes

This pie chart shows all episodes of infection documented during the study period.

Infection post-transplant has only been looked at in one other MPA pharmacogenetic study (8) and the event rates were lower than those observed in this study (table 3.12). Although not

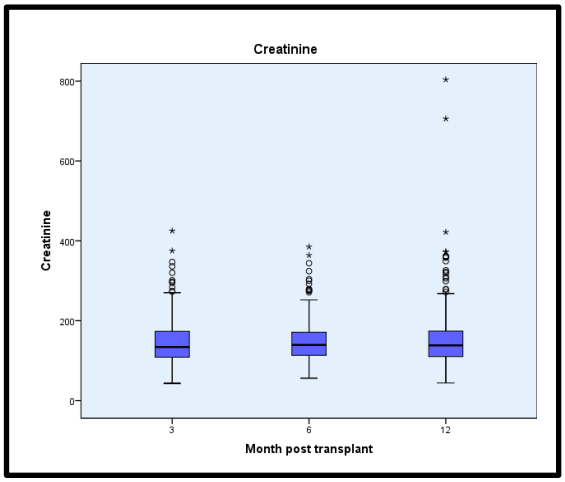
specific to MPA treated transplant patients, a study published in 2013 found an infection rate of 45% with a CMV rate of 6%, pneumonia 12% and UTI of 34.6% (130).

Table 3-12: Infection event rate

Study	Number of recruits	Definition	Event Rate
This study	287	Recurrent or severe infection felt by treating clinician to be as a result of immunosuppressed state (Excluding recurrent UTI)	21.6%
Michelon 2010(8)	218	Infection in first 12 months attributed by treating physician to MPA use	13.1%

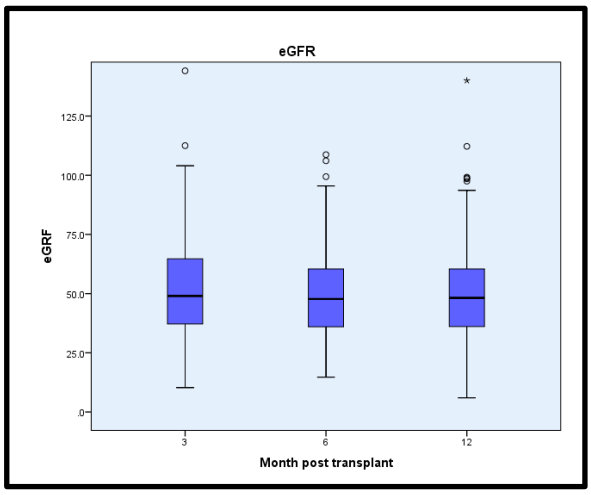
3.11.3 Graft function

The secondary outcome measure of graft function in terms of creatinine and eGFR is shown in figures 3.18 and 3.19 respectively. It can be seen that the mean creatinine and eGFR remained stable across in the first 12 months post transplantation.



		Creat3M	Creat6M	Creat12M
Valid		287	286	285
Missing		0	1	2
Mean		146.15	148.09	154.45
Median		134.00	139.00	138.00
Mode		90	133	133
Std. Deviation		57.332	52.810	80.142
Range		382	329	760
Minimum		43	56	44
Maximum		425	385	804
Percentiles	25	108.00	112.75	110.00
	50	134.00	139.00	138.00
	75	174.00	171.00	174.50

Figure 3-18: Creatinine µmol/L at 3,6 and 12 months post-transplant



		eGFR3M	eGFR6M	eGFR12M
Valid		287	286	285
Missing		0	1	2
Mean		51.02	49.13	49.60
Median		49.0	47.75	48.20
Mode		70.0	34.5	30.0
Std. Deviation		19.65	17.57	19.73
Range		133.8	94.0	134.0
Minimum		10.3	14.7	6.0
Maximum		144.1	108.7	140.0
Percentiles	25	37.1	35.98	35.95
	50	49.0	47.75	48.2
	75	64.7	60.45	60.4

Figure 3-19: eGFR at 3,6 and 12 months post-transplant

The mean creatinine and eGFR were compared using ANOVA one way statistical analysis. This confirmed that there was no statistical difference between the mean creatinine or eGFR between 3, 6 and 12 months post-transplant with P values of 0.28 and 0.47 respectively.

3.11.4 Change in creatinine according to MPA alteration

As discussed above there was no statistical difference in the mean graft function of the study group between months 3,6 and 12 post-transplant. It has been reported in several studies that maintaining patients on MPA post transplantation reduces the risk of rejection episodes as well as improving long term graft outcome, and potentially has more positive impact long term than CNI therapy (131, 132). The study data has been analysed to see if there is any difference in graft function between the participants that continued on full dose MPA compared to those that had dose reduction or withdrawal. Figures 3.20 and 3.21 below show the change in serum creatinine (µmol/l) from month 3 to 12 and month 6 to 12 according to change in MPA.

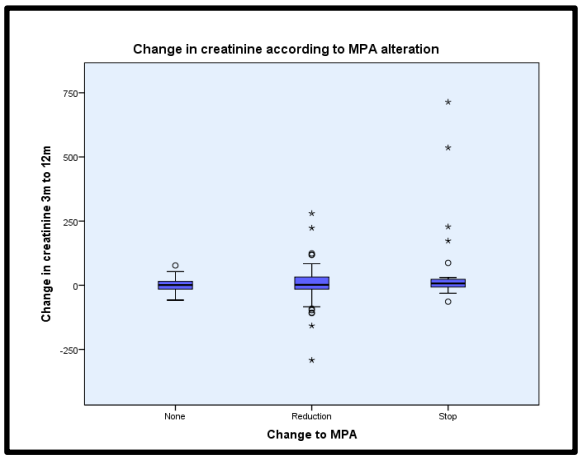


Figure 3-21: Change in creatinine 3 to 12 months according to MPA change

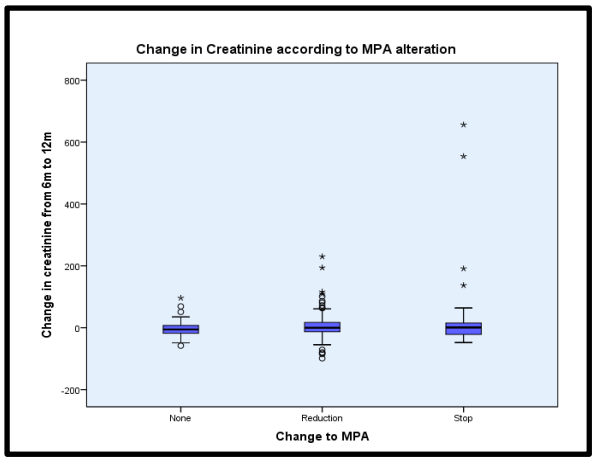


Figure 3-20: Change in creatinine 6 to 12 months according to MPA change

The mean changes in creatinine, between the MPA groups, were then compared using one way ANOVA analysis. This showed a statistically significant difference between MPA withdrawal versus full dose or reduced dose with a P value of <0.0001 for both change in creatinine between 3 and 12 months and 6 and 12 months post transplantation. This supports the evidence in the literature that MPA withdrawal is associated with worse graft outcome.

3.12 Completeness of data set

Information was available for all recruits for the primary and secondary outcome measures up to 1 year. Two subjects died during the first year post transplant 1 at 3 months and 1 at 10 months post-transplant. Data has been included on these individuals up to the time of death. Death occurred due to sudden cardiac death (confirmed at post mortem as myocardial infarction) at home in the patient at 3 months and following prolonged hospital stay for ischaemic bowel and resection in the subject who died at 10 months. In the time to event analysis the data has been censored for death. There was no loss to follow up for any other reason with sufficient clinical and laboratory data available in all subjects.

3.13 Summary

The key findings outlined in this chapter were that the study population is comparable to the UK transplant population with the exception of ethnic diversity, which can be seen as advantageous for genetic studies. The event rates seen in the study population were similar to those observed in other studies, with some variations due in the most part to differing definitions used. These findings are important as they give the study external validity when considering the UK transplant population.

The phenotypic outcome data presented here has been analysed for associations with SNP's and results are presented in chapter 4, demographic and baseline data will also be considered in the analysis.

4 Chapter 4: Results 2

4.1 Introduction

This chapter will present the results of the genetic analysis, using Illumina Human Exome Bead chip V1.1 and the associations with the primary and secondary outcome measures including quality control of the data.

This chapter will first present the results of the quality control of the data which was carried out prior to the genetic analysis for the primary and secondary outcome measures.

Results of a candidate gene approach are then presented including an analysis of SNPs in genes that are known to be involved in the metabolic pathway of mycophenolic acid (133). These include 38 SNPs in 12 genes across 5 chromosomes; the details of these SNPs are shown in Table 1. Testing for association between SNPs and clinical outcomes is presented, followed by Logistic regression for both phenotypic and genotypic variables. Variants that showed an association with the clinical outcomes of leucopenia, anaemia, and BPAR were then analysed for the secondary outcome measure of time to event.

Next the results of an unsupervised analysis are presented. This included all SNPs on the Illumina Human Exome Bead chip V1.1 that remained suitable for analysis following quality control of the data. The association between variants across the exome and the primary outcome measures of leucopenia, anaemia, BPAR and GISE are presented and further analysed by Logistic regression. Secondary outcomes were not included in the unsupervised analysis.

The analysis of these results has produced vast quantities of data. To provide an understanding of the results produced at each stage of the analysis the results will be presented in full for a single primary outcome measure of leucopenia. Results for the remaining outcomes will be summarized here with full results available in the Appendix 15.

4.2 Quality control of the data

Quality control is a fundamental step in the analysis of the genotyping data. The methodology and rationale for the quality control steps have been discussed and explained in Chapter 2: Methods. This stage of the analysis was conducted by a specialised genetic statistician and not by me although I collected and coded all clinical data relating to my cohort. Quality control was carried out on a larger cohort of 2400 individuals who had all been genotyped using the Illumina Human Exome Beadchip V1.1 of which my MPA study cohort formed a subset.

4.2.1 Gender mismatch

Plink was used to confirm gender. This uses a gender coefficient F based on the results of SNPs tested. A male call is made if F is more than 0.8; a female call is made if F is less than 0.2.

The histogram (figure 4-1) shows the X-chromosome specific interbreeding coefficient F, a measure of departure from Hardy-Weinberg Equilibrium. Females should have F close to 0, and males close to 1. Those individuals that show departure from Hardy-Weinberg Equilibrium (HWE), falling between the red lines were highlighted as 'problem' cases and were then further investigated for the F coefficient by gender. Individuals that appeared to be mis-classified according to gender were removed from the data set. This meant that 13 individuals who were recorded as female but had F> 0.5 and 19 individuals recorded as male but with F< 0.3 were excluded from further analysis.

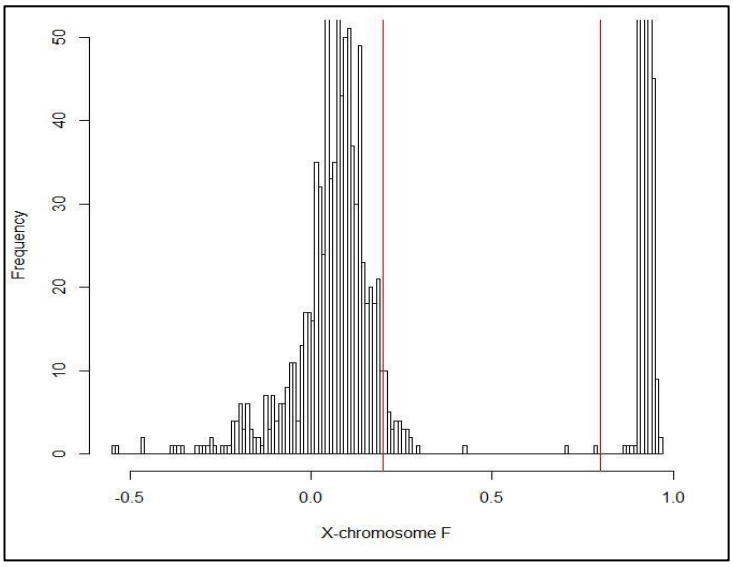


Figure 4-1: Histogram of the X chromosome specific interbreeding coefficient F

This histogram shows the X-chromosome specific interbreeding coefficient F which is a measure of departure from the HWE. Females should have a value close to zero and males should be close to 1. Values that fall within the red lines show a departure from HWE and have been further analysis.

4.2.2 Minor allele frequency and missingness thresholds

The next stage in the analysis was to remove SNPs that were either present at very low frequencies within the cohort (MAF <0.05), or that did not reach a pre specified call rate.

Following investigation for a SNP MAF >0.05 (5%) in the cohort, 110,001 SNP out of the 242,901 SNP included on the beadchip (45%) remained.

Results were then analysed for call rate with a cut off of >97% for individuals (meaning that >97% of SNP analysed in that individual were called as wild type (aa), heterozygote for the SNP (aA) or homozygote for the SNP (AA)), and >99% for SNPs (meaning that 99% of the time the individual

SNP was called as either aa, aA or AA in all individuals) with MAF <0.05. 102 (0.0004) SNP had a call rate <99% and were hence excluded.

Figure 4.2 shows the Q-Q plot of the cumulative call rate for individuals and the cumulative call rate for individual SNP.

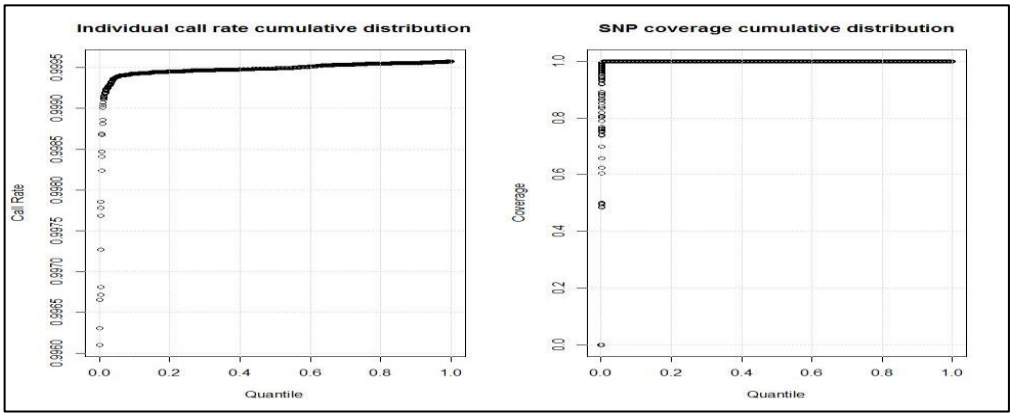


Figure 4-2: Cumulative call rate for individuals and SNPs

The Q-Q plots are shown here for the call rate of SNPs within the cohort. The left hand plot shows the call rates within individual patients, it shows that in all individuals >99% of SNPs were called as either aa, aA or AA. The right hand plot shows the call rate for the SNPs, demonstrating the frequency with which each SNP was called as either aa, aA or AA within the individuals, those that were called <97% of the time were excluded.

Following the application of MAF and missingness thresholds, 1,766 individuals (278 from MPA cohort) and 109,877 SNPs were carried through to the next analysis step.

Hardy Weinberg equilibrium

The remaining SNPs were then investigated for HWE as SNPs out of HWE often indicate a genotyping problem. An HWE cut off of P<10-06 was used for this data, which will exclude SNPs that show a significant departure from HWE. Following investigation for HWE 3,410/109,877 (0.03) of SNPs had a p-value<10-06 and so were removed from the analysis. The Q-Q plot of the HWE p values is shown in figure 4.3.

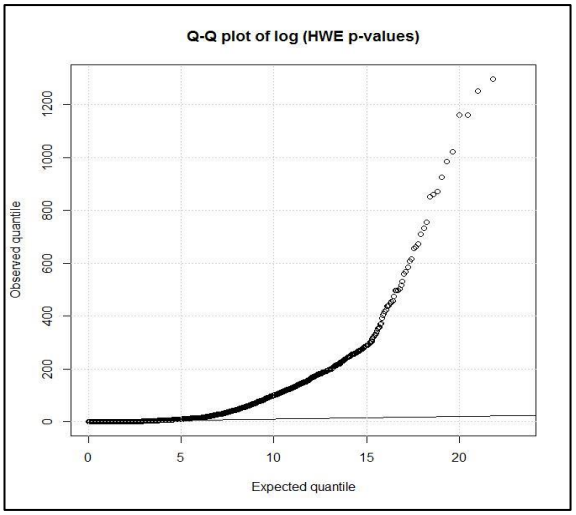


Figure 4-3: Q-Q plot of HWE p values
This Q-Q plot shows the HWE p-values for all the SNP genotyped within the cohort. Deviation from HWE represents a genotype calling problem when SNP called as aa, aA and AA cannot be accurately differentiated from one another. When plotted in this manor results which fall to the right of the plot suggest a large departure from HWE and hence are removed from further analysis.

4.3 Cryptic relatedness

The data was analysed to look for related individuals in the cohort. A cut off IBD (see Methods) value of >0.1 was used meaning individuals above this levels were removed.

The IBD plot for the whole cohort is shown in figure 4.4. There were very few individuals with an IBS value >0.1 (those to the right of the red line) and so the majority remained in the data set.

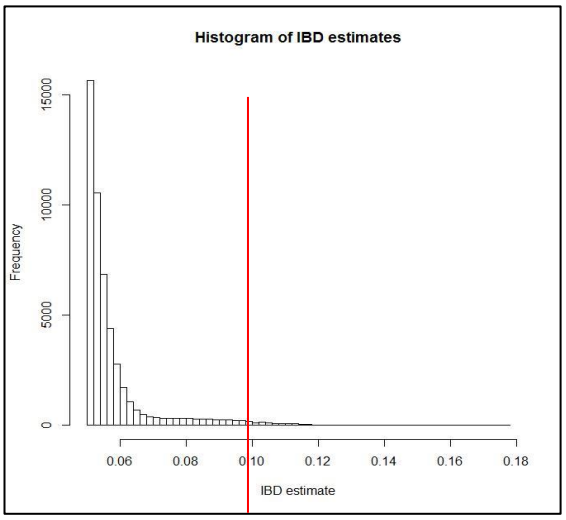


Figure 4-4: Histogram of IBD estimates
This histogram shows the estimated IBD for related individuals. Individuals with an IBS >0.1 have some degree of relatedness (those to the right of the red line). This value will be >0.125 if any individuals were 3rd relatives with closer relatives having a higher IBD value. Very few individuals showed an estimated IBD>0.1 and none were >0.125 so few individuals were removed at this stage.

Following QC steps the final data set contained 1766 individuals and 108,111 SNP.

4.3.1 Principle Components

Principle components analysis was then analysed this produced. The PC cloud for the entire cohort is shown in figure 4.5.

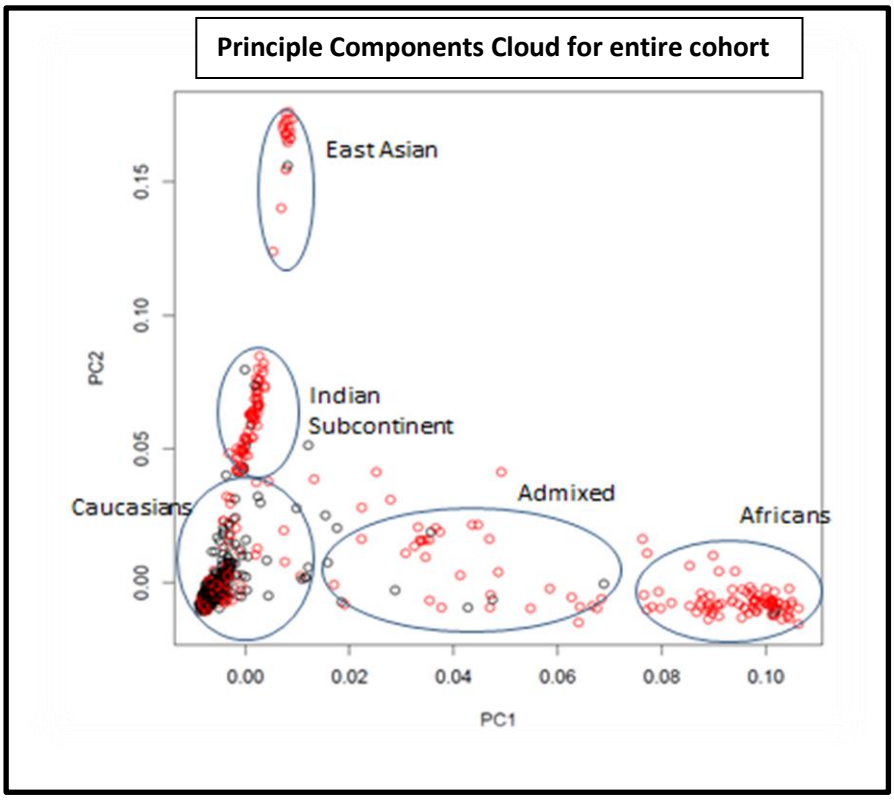


Figure 4-5: Principle components cloud for entire cohort

This PC cloud explicitly models ancestry differences between individuals. The aim is to remove spurious associations due to ancestry.

4.3.2 Samples removed during QC stage

A total of 9 individuals from my MPA Cohort were removed during the QC stage. Two samples were removed due to gender miss-match, when this was further investigated the samples had been incorrectly labelled within genome studio meaning no true miss-match existed. Two individuals were removed due to IBS, these were known siblings. The remaining five individuals were removed due to poor call rates, which may be due to sample quality.

4.4 Validation of the Illumina Beadchip V1.1

The illumina BeadChips have undergone rigorous functional testing by Illumina ensuring the results are sensitive with high call rates and are reproducible (1). However it is important to seek validation within this MPA study cohort and so the Bead Chip V1.1 results were compared between a single SNP on the Bead Chip and the same SNP tested using rtPCR. The results for 255 patients were compared for SNP rs1127354 (*ITPA94*). The call rate for all the patients on the Bead Chip in this SNP were >99.9% and the observed MAF in this cohort was 0.08/255. There was 99.6% concordance between the Bead Chip results and rtPCR results. A single patient was called

differently between the two methods with the Bead Chip calling the patient as homozygote and the rtPCR as heterozygote. It was felt that this was within the limits of acceptability in this cohort.

4.5 Candidate gene analysis

The results presented here represent associations between candidate genes and the primary and secondary outcome measures. Candidate genes were chosen based on evidence within the literature that they were involved in the metabolic pathway of MPA (133). A full list of the SNPs included in this analysis, including the gene that they relate to is shown in table 4.1.

A Bonferroni (BF) correction has been applied to the results to account for multiple testing. The BF used in the candidate gene approach takes into consideration the total number of SNPs analysed across all the candidate genes after taking into account SNPs in complete LD. The rationale for BF correction will be discussed further in Chapter 5: Discussion and Conclusions.

Table 4-1: SNPs analysed in the candidate gene approach

Chrom	Gene	Beadchip identifier	SNP	Position
7	ABCB1	exm2266441	rs3789243	87220886
7	ABCB1	exm631775	rs2032582	87160618
7	ABCB1	exm631843	rs2229109	87179809
7	ABCB1	exm631879	rs9282564	87229440
10	ABCC2	exm848442	rs56131651	101557063
10	ABCC2	exm848464	rs2273697	101563815
10	ABCC2	exm848522	rs17222617	101578952
10	ABCC2	exm848539	rs41318029	101590486
10	ABCC2	exm848562	rs45441199	101591737
10	ABCC2	exm848601	rs17222723	101595996
10	ABCC2	exm848653	rs8187710	101611294
4	ABCG2	exm412774	rs34783571	89013496
4	ABCG2	exm412870	rs2231137	89061114
10	CYP2C8	exm844097	rs10509681	96798749
10	CYP2C8	exm844133	rs1058930	96818119
10	CYP2C8	exm844152	rs41286886	96824658
10	CYP2C8	exm-rs1934951	rs1934951	96798548
7	CYP3A5	exm638602	rs28365083	99250236
7	CYP3A5	exm638644	rs6977165	99269397
12	SLCO1B1	exm2271695	rs2291075	21331625
12	SLCO1B1	exm988933	rs2306283	21329738
12	SLCO1B1	exm988936	rs11045819	21329813
12	SLCO1B1	exm988942	rs4149056	21331549
12	SLCO1B1	exm-rs4363657	rs4363657	21368722
12	SLCO1B1	exm989046	rs34671512	21391976
12	SLCO1B1	exm-rs4149032	rs4149032	21317791
2	UGT1A1	exm-rs887829	rs887829	234668570
2	UGT1A1	exm-rs4148325	rs4148325	234673309
2	UGT1A1	exm-rs6742078	rs6742078	234672639
2	UGT1A5	exm277212	rs3755321	234621825
2	UGT1A8	exm276956	rs17862841	234526784
2	UGT1A8	exm-rs11892031	rs11892031	234565283
2	UGT1A8	exm-rs2602381	rs2602381	234584324
2	UGT1A6	exm277163	rs6759892	234601669
2	UGT1A6	exm277187	rs2070959	234602191
2	UGT1A6	exm277188	rs1105879	234602202
2	UGT1A9	exm277410	rs45449995	234638580
2	UGT1A9	exm277431	rs34622615	234652308
4	UGT2B7	exm403192	rs61361928	69962375

SNPs highlighted in the same colour are in complete LD with each other (R^2=1)

4.6 Leucopenia

4.6.1 Results (pre- QC) of the data

The first results presented were prior to QC of the data set. Only SNPs with a MAF >10% were included in the pre QC analysis. Results were analysed using the Fisher's exact test for allele frequency, dominant and recessive models. These results have been included as they may highlight interesting SNPs that are removed in the QC process and may warrant further investigation. Results were corrected for the potential confounding factors of gender, ethnicity, type of CNI at baseline, type of MPA at baseline, use of induction agent at baseline, use of valganciclovir prophylaxis, type of transplant donor and HLA-MM, using binary Logistic regression. Logistic regression for genotypic factors has not been included here.

Table 4-2: SNP associated with leucopenia prior to QC of the data

Chr	SNP	gene	rs number	MAF	Model	P value	P value after Log reg	Exp (β)	95% CI for Exp (β)
16	exm1241669	CES1	rs62028647	0.36/287	Dom	0.02	0.008	0.46	0.26-0.82
16	exm1241669	CES1	rs62028647	0.36/287	Rec	0.015	0.038	0.34	0.12-0.94
2	exm277187	UGT1A6	rs2070959	0.34/287	Dom	0.006	0.02	0.51	0.28-0.9
2	exm277187	UGT1A6	rs2070959	0.34/287	Rec	0.002	0.14	0.42	0.13-1.3

The genotype allele intensity plot for exm277187 is shown in figure 4.6. The genotypes are well clustered for the wild (red) and homozygotes (blue) with some spread for the heterozygote (Purple) individuals. This SNP passed the QC stage and hence will be analysed further.

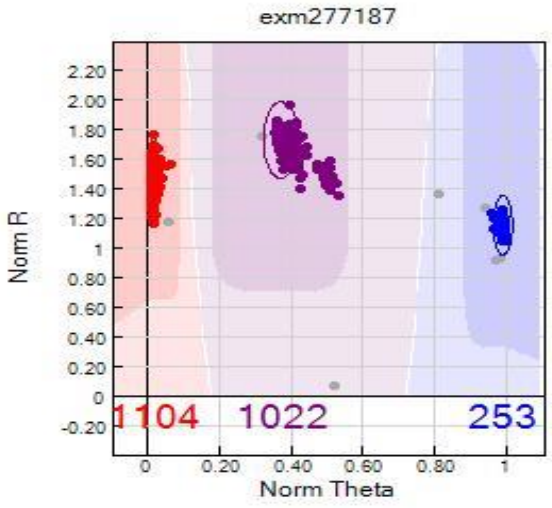


Figure 4-6: Genotype allele intensity plot exm277187.

This plot for exm277187 shows individuals clustered as wild (red) homozygotes (blue) and heterozygote (purple) for the SNP.

During the QC stage SNP exm1241669 was removed due to poor calling of the results across the whole data set of 2135 individuals. The genotype allele intensity plot for exm1241669 is shown in figure 4.7. It can be seen that there is poor differentiation between all genotypes.

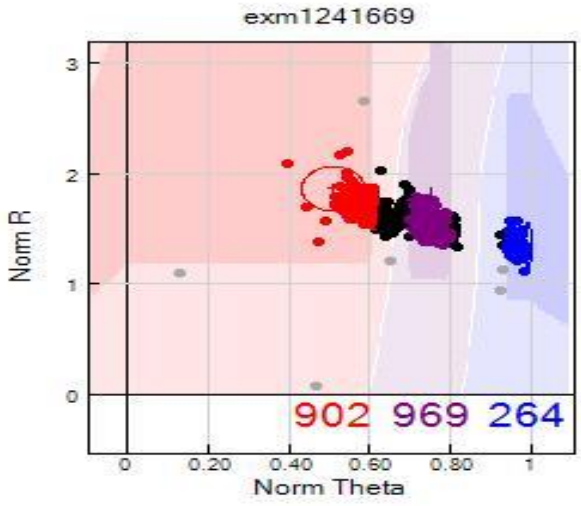


Figure 4-7: Genotype allele intensity plot exm1241669

This plot for exm1241669 showing poor calling of this SNP evidenced by overlapping of the red, purple and blue 'clouds' representing the three genotypes.

When this was analysed using Z call the SNP was removed as it could not differentiate between wild/ heterozygotes and homozygote patients. The individuals within the MPA cohort appeared to be within the more differentiated areas of the allele intensity plot in the majority of individuals although not within the entire cohort. The results were then run through a second programme 'Opticall' which appeared to be able to differentiate the genotypes with more certainty. Due to the poor calling within the entire cohort but potential functional importance within MPA cohort a decision was made to re-genotype the cohort using rtPCR.

Real time PCR data did not match the results from Gencall for exm1241669 (rs62028647). There were no homozygotes identified and all patients were called as 'heterozygotes'. The most likely explanation for this is the presence of a pseudogene variant with an identical genomic context which means that the SNP probe is detecting 4 alleles, two of which are derived from the pseudogene and carry the variant, instead of two from the authentic gene, leading to an inability to call the SNP. This means that the SNP has been correctly removed from further analysis during the QC stage as associations cannot be accurately drawn.

4.6.2 Results for leucopenia following QC of the data

Once the data had been adjusted for QC, Plink was used to look for associations between SNPs in the candidate genes and leucopenia. Fishers exact test was used to test for significance using an allele frequency model looking at whether the individual alleles (e.g. a =wild/reference allele A = alternative allele) are associated with leucopenia. This model does not account for the fact that individuals inherit two copies of each gene and will hence be either wild type (aa), heterozygote for the SNP (aA) or homozygote for the SNP (AA) or that the effect of the SNP may be dominant or recessive. For this reason Fisher exact test was used then used and associations were sought using both a dominant and recessive model. The results of analysis within these three models are presented throughout this results section at each stage of the analysis.

4.6.3 Linkage disequilibrium (LD)

All the SNPs in table 4.1 were analysed using PLINK to look for LD with other SNP on the beadchip. The SNPs rs6742078, rs4148325 and rs887829 which are all located in UGT1A1 gene are in complete LD with one another (R^2 =1). LD was also assessed using 'SNAP' with an R^2 threshold of 0.8, meaning that SNP that exceeded an R value of 0.8 were taken to be in LD, LD plots were produced to show this graphically. Although it cannot be concluded which one of these SNP has the true effect the results for rs4148325 only will be shown here as the analysis has produced identical results for all 3 SNPs.

LD plots were produced for SNP rs6759892 and rs3789243 to include SNPs within 10kbp of the SNP and are shown in figures 4-8 and 4-9 respectively, with table 4.3 showing the SNP's in LD with rs6769892.

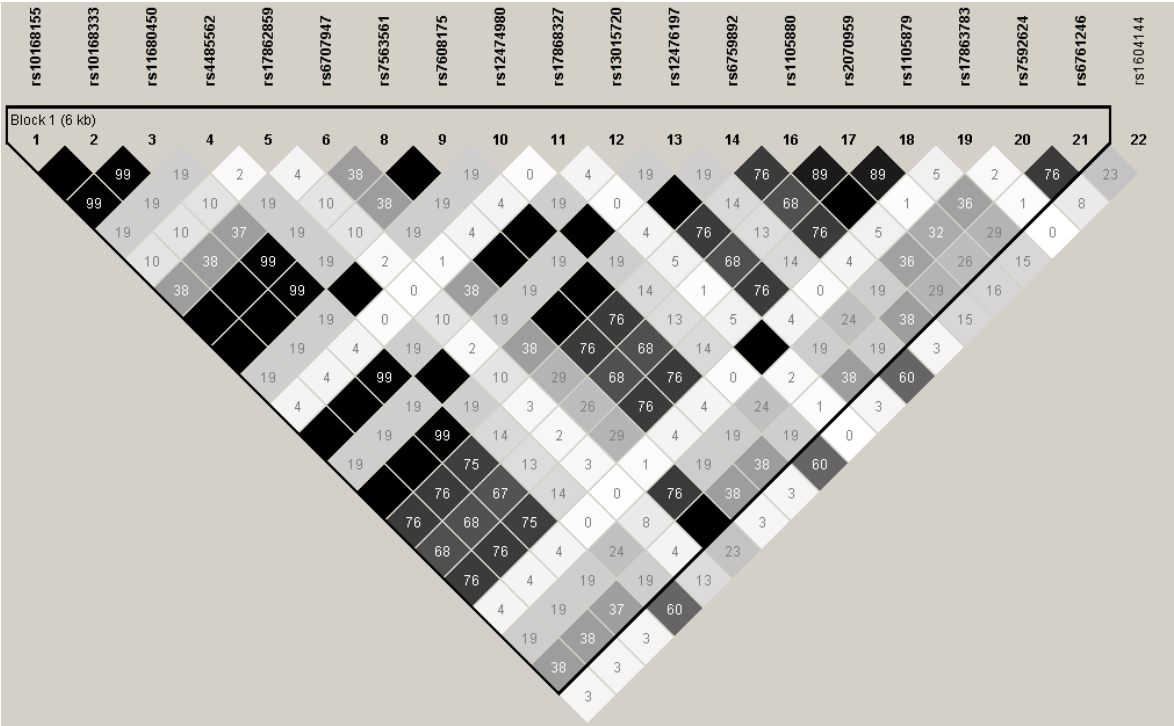


Figure 4-8: LD plot for SNP within 10Kbp of rs6759892.

This shows SNPs that are in LD with rs6759892 with the R^2 value shown on each square, the black squares which represent an R^2 value of 1.

Table 4-3: Summary of R^2 value for SNPs within 10kbp of rs6759892, all the SNP shown in this table are in the UGT1A9 gene.

SNP	R^2
rs7608175	1
rs2563561	1
rs6736508	1
rs10197460	0.9
rs13015720	1
rs11680450	0.99
rs10168333	1
rs10168155	1

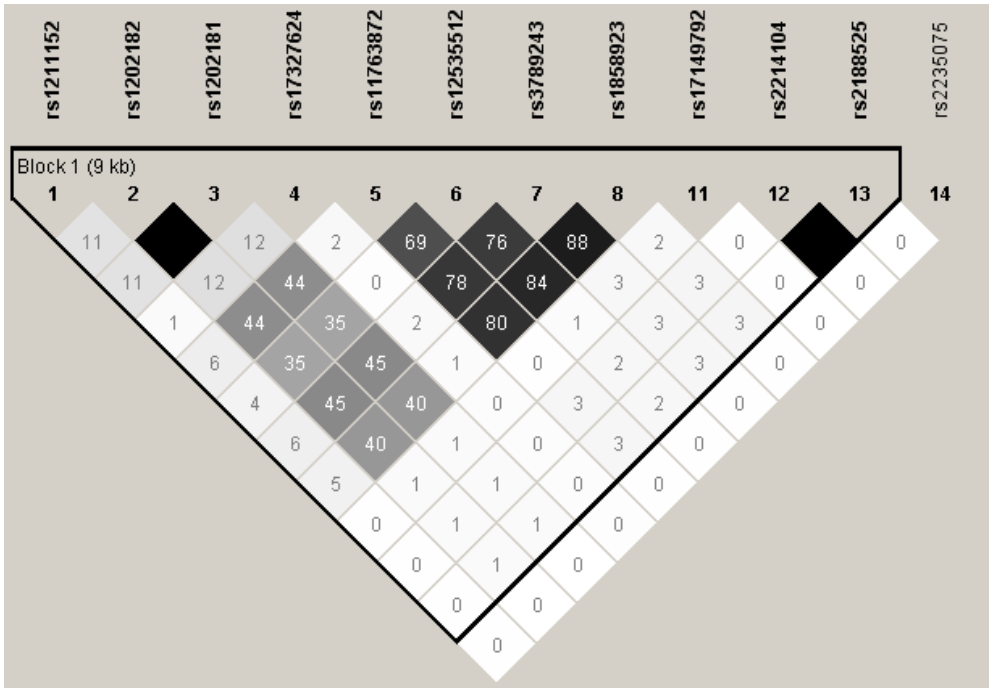


Figure 4-9: LD plot for SNPs within 10kbp of rs3789243

This shows SNPs that are in LD with rs3789243 with the R^2 value shown on each square, the black squares which represent an R^2 value of 1.

All SNPs shown in table 4.1 were included in the analysis but only SNPs with a P value <0.05 prior to BF correction have been presented. A more detailed analysis including Logistic regression and time to event analysis will be presented later in this chapter. Table 4.4 shows the results for the allele frequency model. This model found associations between two UGT1A6 SNPs, one UGT1A1 SNP and one ABCC2 SNP.

Table 4-4: Allele frequency model: SNPs associated with leucopenia

CHR	gene	SNP	rs number	P value	BF
2	UGT1A6	exm277163	rs6759892	0.019	0.63
2	UGT1A6	exm277187	rs2070959	0.028	0.93
2	UGT1A1	exm-rs4148325	rs4148325	0.029	0.96
10	ABCC2	exm848442	rs56131651	0.055	1

The results were then analysed using a dominant model (Table 4.5) and all three variants showed significant association with leucopenia although significance was lost after BF correction for multiple testing.

Table 4-5: Dominant model SNPs associated with leucopenia

CHR	Gene	SNP	rs number	P value	BF
2	UGT1A6	exm277163	rs6759892	0.012	0.4
7	ABCB1	exm2266441	rs3789243	0.015	0.5
2	UGT1A6	exm277187	rs2070959	0.041	1

The results of a recessive model are presented in table 4.6. Only the UGT1A1 gene SNP rs4148325 retained marginal significance before BF correction for multiple testing

Table 4-6: Fishers exact Recessive model SNPs associated with leucopenia

CHR	Gene	SNP	rs number	P value	BF
2	UGT1A1	exm-rs4148325	rs4148325	0.045	1

4.6.4 Logistic regression for genotypic confounders

The next stage was Logistic regression for the first 4 principal components (PC) of the population. PLINK was used for this regression stage as it is regressing against the genotype. Regression was carried out for the whole MPA cohort, with a separate analysis for a pure Caucasian cohort based on the principal components. This separate cohort analysis was carried out to account for any concentration of SNPs within ethnic minority groups. Allele frequency, dominant and recessive models were also incorporated in the analysis (table 4.7-4.8 PC all, tables 8-19 PC Caucasian).

Table 6 shows the results for the allele frequency model following Logistic regression for PC of the population for the entire cohort (N=278). This was carried out for all SNP presented in Table 4.1 but only SNPs with a p value <0.05 prior to BF correction have been presented here. In the allele frequency model, the two UGT1A6 SNPs and the UGT1A1 SNP showed an association prior to BF correction.

Table 4-7: Allele frequency model following Logistic regression for PC of entire cohort (N=278) for leucopenia

CHR	Gene	SNP	rs number	OR	SE	95% CI	P value	BF
2	UGT1A1	exm-rs4148325	rs4148325	0.59	0.23	0.38-0.91	0.019	0.63
2	UGT1A6	exm277163	rs6759892	0.64	0.21	0.42-0.97	0.037	1
2	UGT1A6	exm277187	rs2070959	0.63	0.22	0.41-0.98	0.042	1

The SNPs in Table 4.1 were then analysed for Logistic regression for PC in the dominant model and results are presented in Table 4.7. All the SNPs that showed an association in the allele frequency model continued to show an association in the dominant model. A further UGT1A6 SNP showed an association, although this SNP did not reach pre- BF significance at previous stages of

the analysis, it has been included in the next stage of the analysis to see if a true association may exist.

Table 4-8: Dominant model following Logistic regression for PC of entire cohort (N=278) for leucopenia

CHR	Gene	SNP	rs number	OR	SE	95% CI	P	BF
2	UGT1A6	exm277163	rs6759892	0.45	0.29	0.26-0.8	0.007	0.23
2	UGT1A6	exm277188	rs1105879	0.55	0.28	0.32-0.96	0.035	1
2	UGT1A6	exm277187	rs2070959	0.56	0.28	0.32-0.97	0.038	1
2	UGT1A1	exm-rs4148325	rs4148325	0.57	0.28	0.33-0.99	0.047	1

The next stage of the analysis looked at SNP associations with leucopenia following Logistic regression for the PC of the entire cohort applying the recessive model. No SNPs showed a significant association in the recessive model therefore no results have been shown here. Following Logistic regression for PC of a Caucasian only cohort. (N=233) revealed significant associations for the UGT1A1 (rs4148325) and UGT1A6 (rs6759892) SNPs (Table 1.9). The SNP ABCB1 (rs3789243) also showed an association but the p value does not quite reach pre BF significance at P=0.056.This SNP was previously identified in the dominant model.

Table 4-9: Allele frequency model following Logistic regression for PC of Caucasian only cohort (N=233) for leucopenia

CHR	Gene	SNP	rs number	OR	SE	95% CI	P value	BF
2	UGT1A1	exm-rs4148325	rs4148325	0.61	0.25	0.37-0.99	0.049	1
2	UGT1A6	exm277163	rs6759892	0.60	0.23	0.38-0.96	0.032	1
7	ABCB1	exm2266441	rs3789243	1.49	0.21	0.99-2.24	0.056	1

Table 4.10 shows the SNPs which remain significant pre BF correction following Logistic regression for PC of Caucasian only cohort in the dominant model. The UGT1A6 SNP (rs6759892) seen in the allele frequency model has remained significant. The ABCB1 SNP (rs3789243) which did not reach significance in the allele frequency model has become significant when analysed in the dominant model. The further UGT1A6 SNP included in table 4.10 was also significant in the dominant model following Logistic regression for PC of the entire cohort.

Table 4-10: Dominant model following Logistic regression for PC of Caucasian only cohort (N=233) for leucopenia

CHR	Gene	SNP	rs number	OR	SE	95% CI	P value	BF
7	ABCB1	exm2266441	rs3789243	2.61	0.40	1.2-5.69	0.016	0.53
2	UGT1A6	exm277163	rs6759892	0.49	0.32	0.26-0.91	0.024	0.79
2	UGT1A6	exm277187	rs2070959	0.55	0.31	0.30-0.999	0.05	1

After Logistic regression for PC of the Caucasian only cohort using a recessive model no SNPs remained significant for leucopenia and so have not been presented here.

4.6.5 Logistic regression for phenotypic confounders

Following the above steps any significant SNPs (prior to BF correction) were analysed using Logistic regression to adjust for confounding factors. Results were corrected for gender, ethnicity, type of CNI at baseline, type of MPA at baseline, use of induction agent at baseline, use of valganciclovir prophylaxis, type of transplant donor and HLA-MM. The results of Logistic regression for these confounders are shown below. SNPs were analysed by allele frequency, dominant and recessive models as outlined previously. Only SNPs that had shown significance following the above steps (Logistic regression or PC) were included here as this Logistic regression analysis will adjust for phenotypic factors whereas the previous step had corrected for genotypic factors.

Table 4.11 shows the results of the Logistic regression for phenotypic confounders in the entire cohort (N=278). The analysis was carried out for the allele frequency, dominant and recessive models as outlined previously. The results in table 4.11 present the combined results of these models. It includes candidate SNPs analysed, which reached a p value <0.05 prior to BF correction have been highlighted. At this stage in the analysis the 3 UGT1A6 genes that have shown associations throughout the steps of analysis presented above have again shown associations with leucopenia.

Table 4-11: Results following Logistic regression for phenotypic factors in the entire cohort (N=278) for leucopenia

Chr	Gene	EXM	rs number	Model	P value	Exp (β)	95% CI for Exp (β)	BF
2	UGT1A6	exm277163	rs6759892	Allele freq	0.018	0.46	0.24-0.87	0.59
2	UGT1A6	exm277163	rs6759892	Dom	0.008	0.44	0.24-0.81	0.26
2	UGT1A6	exm277187	rs2070959	Allele freq	0.072	0.33	0.1-1.1	1
2	UGT1A6	exm277187	rs2070959	Dom	0.034	0.53	0.29-0.95	1
2	UGT1A6	exm277188	rs1105879	Dom	0.042	0.54	0.3-0.98	1
2	UGT1A1	exm-rs6742078	rs6742078	Allele freq	0.064	0.32	0.1-1.07	1
2	UGT1A1	exm-rs4148325	rs4148325	Dom	0.1	0.61	0.34-1.1	1

Table 4.12 shows the results of the Logistic regression for phenotypic confounders in the Caucasian only cohort (N=233). The analysis was carried out for allele frequency, dominant and recessive models outlined previously. The results in table 4.11 present the combined results of these models. It includes SNPs that following analysed had a p value <0.05 prior to BF correction.

At this stage in the analysis, the two UGT1A6 SNPs (rs6759892 and rs2070959) and the ABCB1 SNP (rs3789243) have shown associations throughout the previous analysis steps have again shown associations with leucopenia.

Table 4-12: Results following Logistic regression for phenotypic factors in the Caucasian only cohort (N=233) for leucopenia

Chr	Gene	EXM	rs number	Model	P value	Exp (β)	95% CI for Exp (β)	BF
7	ABCB1	exm2266441	rs3789243	Dom	0.024	2.73	1.14-6.53	0.79
2	UGT1A6	exm277163	rs6759892	All freq	0.019	0.43	0.21-0.87	0.63
2	UGT1A6	exm277163	rs6759892	Dom	0.018	0.33	0.13-0.83	0.59
2	UGT1A6	exm277187	rs2070959	Dom	0.031	0.49	0.25-0.94	1
2	UGT1A1	exm-rs4148325	rs4148325	All freq	0.01	2.65	1.25-5.61	0.33

4.6.6 Secondary outcome analysis: Time to leucopenia

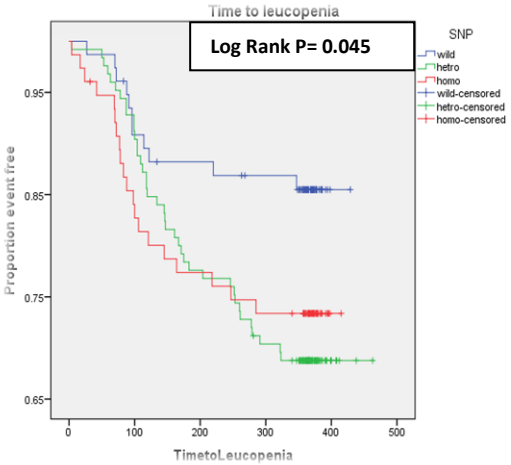
The next stage of the analysis looked at time to event. If having that SNP means that the outcome will occur sooner in those individuals, then association is likely to be of greater clinical significance.

Only SNPs with significant associations (pre BF correction p value <0.05) following all of the above steps were analysed for time to leucopenia. This analysis used Kaplan Meier (KM) survival curves with the Log Rank test. Significant KM results were then analysed using Cox regression to adjust for the confounding variables of gender, ethnicity, type of CNI at baseline, type of MPA at baseline, use of induction agent at baseline, use of valganciclovir prophylaxis, type of transplant donor and HLA-MM.

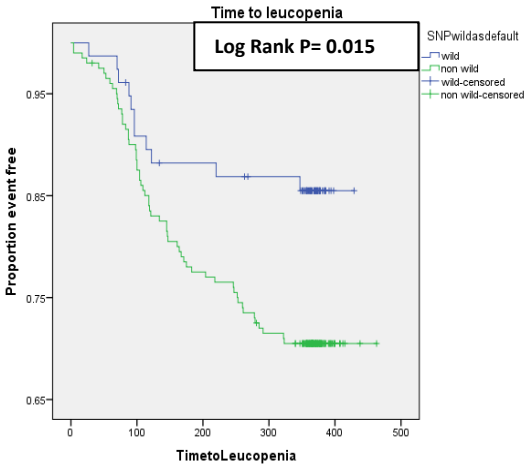
Table 4.13 shows the results of KM and Cox regression for time to leucopenia for the entire cohort (N=278). Only the SNPs that showed significant results have been included in the table. The results were analysed looking at the effect of all 3 genotypes (i.e. aa, aA and AA), this is referred to as 'genotypic' in the table, for the dominant model (Dom) comparing reference homozygote 'wild' to heterozygote or alternative homozygote 'non-wild' (i.e. aa compared to aA or AA) and for a recessive model where non-homozygotes were compared to homozygotes for the SNP (i.e. aa or aA compared to AA). The UGT1A6 SNP (rs6759892) and ABCB1 SNP (rs3789243) were the only two SNP that showed a significant time to leucopenia association and are the results presented here. KM curves are shown in figure 4.10.

Table 4-13: Significant KM and Cox regression results for time to leucopenia for the entire cohort (N=278)

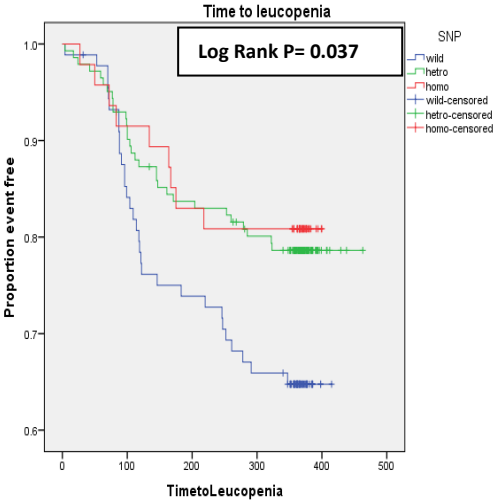
Chr	Gene	EXM	rs number	Model	KM log rank	KM curve	Cox reg P value	RR	95% CI
7	ABCB1	exm2266441	rs3789243	Genotypic	0.045	1 (a)	0.013	2.45	1.2-4.97
7	ABCB1	exm2266441	rs3789243	Dom	0.015	2 (b)	0.014	2.35	1.19-4.66
2	UGT1A6	exm277163	rs6759892	genotypic	0.037	3 (c)	0.01	0.5	0.3-0.85
2	UGT1A6	exm277163	rs6759892	Dom	0.011	4(d)	0.004	0.48	0.29-0.8



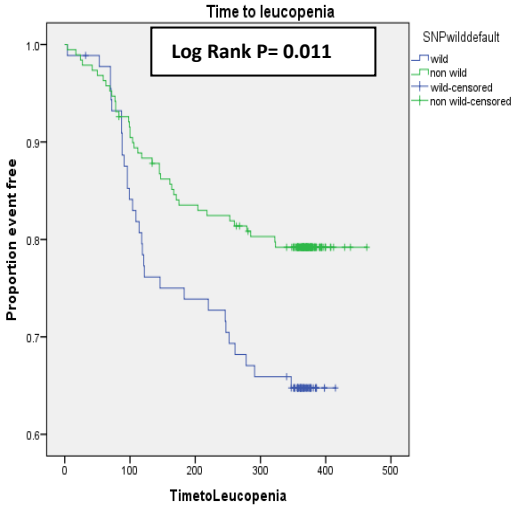
a: KM 1



b: KM 2



c: KM 3



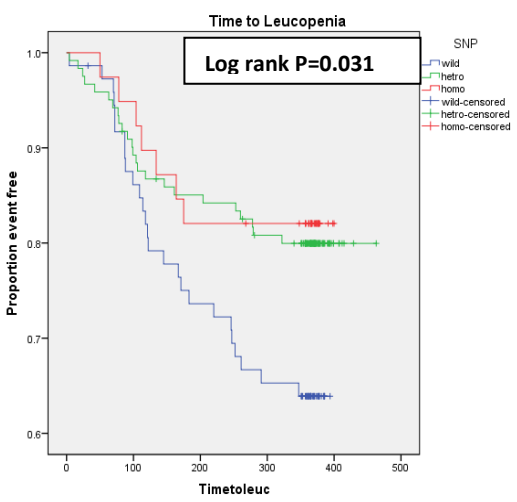
d: KM 4

Figure 4-10: KM curves 1 to 4 for time to leucopenia for analysis of the entire MPA cohort.
a: **KM Curve 1 time to leucopenia for SNP rs3789243 (ABCB1 gene) showing wild type (blue), heterozygotes (green) and Homozygotes (red). The curve clearly shows that a greater proportion of the wild type remain event (leucopenia) free at 1 year. b: KM Curve 2 time to leucopenia for SNP rs3789243 (ABCB1 gene) showing wild type (blue) and non-wild (green). The curve again demonstrates that a greater proportion of the wild type remain event (leucopenia) free at 1 year. It suggests the effect of the SNP is dominant.** c: **KM curve 3 time to leucopenia for SNP rs6759892 (UGT1A6 gene) showing wild type (blue), heterozygotes (green) and homozygotes (red). The curve clearly shows that a greater proportion of the wild type develop the event (leucopenia) during 1 year.** d: **KM Curve 2 time to leucopenia for SNP rs6759892 (UGT1A6 gene) showing wild type (blue) and non-wild (green). The curve again demonstrates that a greater proportion of the wild type develop the event (leucopenia) by 1 year. It suggests a protective effect of the SNP which is dominant.**
(Note that in the KM Survival curve graphs the x axis does not start at zero. As the overall event rate for leucopenia in this cohort was 25.8% the axis has been altered to reflect this).

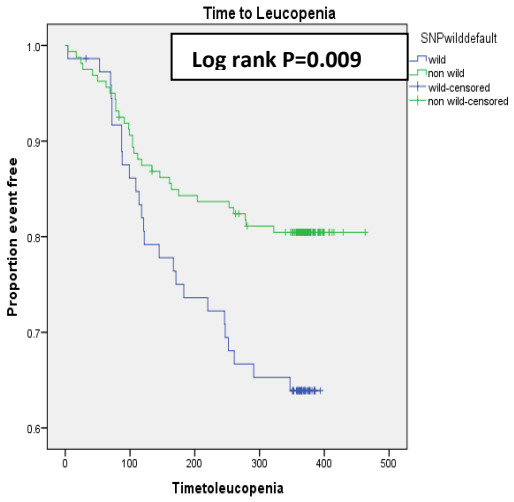
Table 4.14 shows the results of KM and Cox regression for time to leucopenia for the Caucasian only cohort (N=233). Only the SNPs that showed significant results have been included in the table. The results were analysed as outlined above. The UGT1A6 SNP (rs6759892) showed a significant time to leucopenia association for the Caucasian cohort and is presented here; the remaining SNP did not show a time to leucopenia association.

Table 4-14: Significant KM and Cox regression results for time to leucopenia for the Caucasian only cohort (N=233)

Chr	Gene	EXM	rs number	Model	KM log rank	KM curve	Coxreg P value	RR	95% CI
2	UGT1A6	exm277163	rs6759892	All	0.031	5 (e)	0.02	0.5	0.28-0.90
2	UGT1A6	exm277163	rs6759892	Dom	0.009	6 (f)	0.013	0.499	0.29-0.86



e: KM 5



f: KM 6

Figure 4-11: KM curves 5 and 6 showing time to leucopenia for analysis of the Caucasian only MPA cohort.

e: **KM curve 5 time to leucopenia for SNP rs6759892 (UGT1A6 gene) showing wild type (blue), heterozygotes (green) and Homozygotes (red). The curve clearly shows that a greater proportion of the wild type develop the event (leucopenia) during 1 year.** f: **KM Curve 2 time to leucopenia for SNP rs6759892 (UGT1A6 gene) showing wild type (blue) and non-wild (green). The curve again demonstrates that a greater proportion of the wild type develop the event (leucopenia) by 1 year. It suggests a protective effect of the SNP which is dominant.**
(Note that in the KM Survival curve graphs the x axis does not start at zero. As the overall event rate for leucopenia in this cohort was 25.8% the axis has been altered to reflect this).

4.7 Candidate gene analysis summary

The full results for the remaining outcomes of interest of anaemia, biopsy proven acute rejection (BPAR), gastrointestinal side effects (GISE), infection, MPA dose alteration and MPA stop are shown in the appendix 15. A summary of the significant findings for these outcomes is given here. The following tables will present only the SNPs with significant associations (P<0.05) prior to BF correction after Logistic regression for PC of the population. Some SNPs did not remain significant through all stages of the analysis. The results were produced following the same steps as presented for leucopenia in this chapter. At the start of analysis all SNP included in table 1 were analysed for association with clinical outcomes by allele frequency, and by dominant and recessive models. All SNPs in table 4.1 were also include in the Logistic regression for the genotypic factors of PC with results for the entire cohort and pure Caucasian cohort analysed separately and will be presented in separate tables here. Only SNPs that showed a significant association following Logistic regression for PC where analysed for phenotypic confounders. The exact phenotypic confounders included in this stage of analysis varied slightly as per the outcome of interest and these have been included in Appendix 15. At all stages allele frequency, dominant and recessive models were analysed but only results of the dominant and recessive models are included in the summary table (allele frequency model results are available in Appendix 15)

The outcomes of leucopenia, anaemia and BPAR were also analysed for the secondary outcome measure of time to event. The only SNPs that were included in this stage of the analysis were those that had continued to show an association with a p value <0.05 pre BF correction after Logistic regression for both PC and phenotypic confounders. This stage of the analysis was included to look for strength of association as outlined earlier in this chapter.

Table 14 shows a summary of significant results following Logistic regression for PC of the entire population (N=278) and for possible phenotypic factors. The table includes the observed MAF in this cohort. The Exp (β) and 95% CI relate to Logistic regression or phenotypic confounders and odds ratios are presented. The SNPs that are of most interest are those which continue to show significance across all steps and these are highlighted in the table. The results that are not highlighted did not remain significant following Logistic regression for phenotypic confounders. The BF corrections have not been included in this table as no results remained significant following BF correction, this will be further discussed in Chapter 5: Discussion and Conclusions. The clinical significance of the results summarised here will also be discussed in Chapter 5: Discussion and conclusions.

Table 4.15 shows a summary of significant results as outlined for table 4.14 but relate to the Caucasian only cohort.

Table 4.16 shows a summary of results of the secondary outcome measures of time to event for anaemia, leucopenia, and BPAR. This table includes the analysis for both the entire cohort 'All' and the Caucasian only cohort 'Caucasian', the analysis for these cohorts was carried out separately as outlined above for leucopenia. The corresponding KM curves for these results are presented above for leucopenia and in Appendix 15 for the other outcomes. Results that have shown an association which remains following Cox regression analysis are highlighted. Relative risk (RR) and 95% CI for Cox regression are included.

Table 4-15: Summary of significant associations between SNP and outcome of interest following Logistic regression for entire cohort (N=278)

Outcome	Gene	Beadchip identifier	SNP	MAF in this cohort	Model	P value Log reg for PC	P value Log reg for phenotypic confounders	Odds ratio	95% CI
Anaemia	SLCO1B1	exm2271695	rs2291075	0.383/278	Dom	0.014	0.006	0.42	0.23-0.78
Anaemia	SLCO1B1	exm988933	rs2306283	0.414/278	Dom	0.038	0.031	0.51	0.28-0.94
BPAR	ABCC2	exm848464	rs2273697	0.203/278	Dom	0.023	0.103	0.55	0.27-1.13
BPAR	SLCO1B1	exm988942	rs4149056	0.138/278	Dom	0.036	0.04	0.39	0.16-0.96
BPARVor C	ABCC2	exm848464	rs2273697	0.203/278	Dom	0.033	0.25	0.59	0.25-1.4
BPARVor C	SLCO1B1	exm988936	rs11045819	0.151/278	Dom	0.034	0.047	2.27	1.01-5.1
BPARVor C	CYP2C8	exm844097	rs10509681	0.104/278	Dom	0.039	0.74	1.18	0.44-3.21
Leucopenia	UGT1A6	exm277163	rs6759892	0.424/278	Dom	0.007	0.008	0.44	0.24-0.81
Leucopenia	UGT1A6	exm277188	rs1105879	0.344/278	Dom	0.035	0.042	0.54	0.3-0.98
Leucopenia	UGT1A6	exm277187	rs2070959	0.311/278	Dom	0.038	0.034	0.53	0.29-0.95
Leucopenia	UGT1A1	exm-rs6742078	rs6742078	0.318/278	Dom	0.047	0.1	0.61	0.34-1.1
GISE	ABCG2	exm412774	rs34783571	0.007/278	Dom	0.059	0.068	3.08	0.92-10.35
LGISE	SLCO1B1	exm988942	rs4149056	0.138/278	Dom	0.024	0.256	0.65	0.31-1.36
LGISE	ABCB1	exm631879	rs9282564	0.117/278	Dom	0.030	0.479	0.77	0.36-1.61
LGISE	SLCO1B1	exm988933	rs2306283	0.414/278	Dom	0.037	0.157	0.65	0.35-1.18

Outcome	Gene	Beadchip identifier	SNP	MAF in this cohort	Model	P value Log reg for PC	P value Log reg for phenotypic confounders	Odds ratio	95% CI
LGISE	SLCO1B1	exm989046	rs34671512	0.059/278	Dom	0.046	0.755	0.86	0.32-2.28
Infection	ABCC2	exm848522	rs17222617	0.023/278	Dom	0.035	0.068	3.08	0.92-10.35
Infection	SLCO1B1	exm988933	rs2306283	0.414/278	Dom	0.040	0.038	0.52	0.28-0.96
MPA change	SLCO1B1	exm-rs4363657	rs4363657	0.16/278	Dom	0.017	0.058	1.75	0.98-3.1
MPA change	UGT1A9	exm277410	rs45449995	0.027/278	Dom	0.050	0.043	0.28	0.08-0.96
MPA stop	SLCO1B1	exm988933	rs2306283	0.414/278	Dom	0.015	0.257	0.609	0.27-1.44
MPA stop	ABCG2	exm412774	rs34783571	0.007/278	Dom	0.029	0.292	3.61	0.33-39.29
MPA stop	CYP2C8	exm844097	rs10509681	0.104/278	Rec	0.018	0.297	3.98	0.3-53.5

Table 4-16: Summary of significant associations between SNP and outcome of interest following Logistic regression for the Caucasian only cohort (N=233)

Outcome	Gene	Beadchip identifier	SNP	MAF in this cohort	Model	Log reg for PC	Log reg for phenotypic confounders	Odds ratio	95%CI
Anaemia	SLCO1B1	exm2271695	rs2291075	0.395/233	Dom	0.042	0.006	0.4	0.22-.76
Anaemia	SLCO1B1	exm988933	rs2306283	0.408/233	Dom	0.041	0.031	0.49	0.26-.93
Anaemia	UGT1A9	exm277163	rs6759892	0.427/233	Dom	0.044	0.037	0.49	0.26-0.96
Anaemia	CYP2C8	exm844097	rs10509681	0.105/233	Dom	0.031	0.094	1.93	0.9-4.14
BPAR	SLCO1B1	exm989046	rs34671512	0.056/233	Dom	0.003	0.008	3.53	1.382-9.0
BPAR	UGT1A1	exm-rs6742078	rs6742078	0.309/233	Dom	0.014	0.011	2.65	1.25-5.6
BPAR	SLCO1B1	exm988936	rs11045819	0.163/233	Dom	0.026	0.053	2.03	0.99-4.14
BPAR	UGT1A5	exm277212	rs3755321	0.114/233	Dom	0.029	0.019	2.5	1.17-5.39
BPAR	SLCO1B1	exm-rs4149032	rs4149032	0.343/233	Dom	0.029	0.045	2.14	1.02-4.51
BPAR	ABCC2	exm848464	rs2273697	0.206/233	Dom	0.039	0.045	0.44	0.2-0.98
BPAR	SLCO1B1	exm-rs4149032	rs4149032	0.343/233	Rec	0.007	0.005	3.77	1.5-9.5
BPARvorc	SLCO1B1	exm989046	rs34671512	0.056/233	Dom	0.003	0.014	3.7	1.3-10.48
BPARvorc	SLCO1B1	exm988942	rs4149056	0.15/233	Dom	0.027	0.024	0.17	0.035-0.79
BPARvorc	SLCO1B1	exm988936	rs11045819	0.163/233	Dom	0.029	0.055	2.33	0.98-5.5
BPARvorc	SLCO1B1	exm-rs4363657	rs4363657	0.165/233	Dom	0.038	0.035	0.24	0.07-0.9
BPARvorc	UGT1A5	exm277212	rs3755321	0.114/233	Dom	0.041	0.028	2.8	1.12-7.0
BPARvorc	SLCO1B1	exm-rs4149032	rs4149032	0.343/233	Rec	0.007	0.004	4.7	1.6-13.7

Outcome	Gene	Beadchip identifier	SNP	MAF in this cohort	Model	Log reg for PC	Log reg for phenotypic confounders	Odds ratio	95%CI
Leucopenia	ABCB1	exm2266441	rs3789243	0.496/233	Dom	0.016	0.024	2.73	1.14-6.53
Leucopenia	UGT1A6	exm277163	rs6759892	0.427/233	Dom	0.024	0.018	0.33	0.13-0.83
Leucopenia	UGT1A6	exm277187	rs2070959	0.307/233	Dom	0.050	0.031	0.49	0.25-0.94
GISE	UGT1A9	exm-rs2602381	rs2602381	0.457/233	Dom	0.027	0.035	2.17	1.06-4.47
LGISE	UGT1A9	exm-rs2602381	rs2602381	0.457/233	Dom	0.031	0.062	2.15	0.96-4.83
UGISE	UGT1A6	exm277163	rs6759892	0.427/233	Dom	0.026	0.018	0.33	0.13-0.83
UGISE	ABCB1	exm631843	rs2229109	0.028/233	Dom	0.042	0.021	4.65	1.26-17.17
MPA change	UGT1A6	exm277187	rs2070959	0.307/233	Dom	0.024	0.03	0.53	0.30-0.94
MPA change	UGT1A6	exm277188	rs1105879	0.335/233	Dom	0.037	0.031	0.53	0.30-0.94
MPA change	UGT1A6	exm277163	rs6759892	0.427/233	Dom	0.038	0.056	0.55	0.3-1.01

Table 4-17: Summary of significant associations of SNP for time to event analysis, including both analysis for entire cohort and Caucasian only cohort

Outcome	Cohort	Chr	EXM	rs number	Gene	Model	KM log rank	Cox reg p Value	RR	95% CI
Anaemia	All	12	exm2271695	rs2291075	SLCO1B1	All geno	0.008	0.004	0.49	0.30-0.8
Anaemia	All	12	exm2271695	rs2291075	SLCO1B1	Dom	0.002	0.009	0.55	0.35-0.86
Anaemia	All	12	exm988933	rs2306283	SLCO1B1	All geno	0.003	0.004	0.49	0.30-.79
Anaemia	All	12	exm988933	rs2306283	SLCO1B1	Dom	0.002	0.014	0.59	0.38-0.9
Anaemia	Caucasian	10	exm844097	rs10509681	CYP2C8	Dom	0.028	0.136	1.49	0.88-2.52
BPAR	All	12	exm988942	rs4149056	SLCO1B1	Dom	0.03	0.03	0.38	0.16-0.91
BPAR	Caucasian	12	exm989046	rs34671512	SLCO1B1	Dom	0.048	0.212	1.66	0.75-3.68
BPAR	Caucasian	12	exm-rs4149032	rs4149032	SLCO1B1	All geno	0.005	0.001	4.34	1.83-10.29
BPAR	Caucasian	12	exm-rs4149032	rs4149032	SLCO1B1	Dom	0.034	0.021	2.27	1.13-4.56
BPAR	Caucasian	12	exm-rs4149032	rs4149032	SLCO1B1	Rec	0.002	0.002	3.20	1.54-6.64
BPAR	Caucasian	12	exm988936	rs11045819	SLCO1B1	All geno	0.002	<0.0001	8.31	2.63-26.21
BPAR	Caucasian	12	exm988936	rs11045819	SLCO1B1	Dom	0.007	0.01	2.27	1.22-4.24
BPAR	Caucasian	12	exm988936	rs11045819	SLCO1B1	Rec	0.003	0.001	6.81	2.23-20.83
BPAR	Caucasian	2	exm-rs6742078	rs6742078	UGT1A1	All geno	0.041	0.11	1.73	0.88-3.39
BPAR	Caucasian	2	exm-rs887829	rs887829	UGT1A1	Dom	0.041	0.11	1.73	0.88-3.40
BPAR	Caucasian	2	exm-rs4148325	rs4148325	UGT1A1	Dom	0.041	0.11	1.73	0.88-3.41
BPAR vor c	All	12	exm988936	rs11045819	SLCO1B1	Dom	0.048	0.059	2.03	0.975-4.22
BPAR vor c	Caucasian	12	exm989046	rs34671512	SLCO1B1	All geno	0.001	0.004	3.39	1.48-7.78

Outcome	Cohort	Chr	EXM	rs number	Gene	Model	KM log rank	Cox reg p Value	RR	95% CI
BPAR vor c	Caucasian	12	exm989046	rs34671512	SLCO1B1	Dom	<0.0001	0.001	3.69	1.65-8.25
BPAR vor c	Caucasian	12	exm-rs4149032	rs4149032	SLCO1B1	All geno	0.001	<0.0001	7.22	2.48-21.02
BPAR vor c	Caucasian	12	exm-rs4149032	rs4149032	SLCO1B1	Rec	<0.0001	<0.0001	5.47	2.27-13.19
BPAR vor c	Caucasian	12	exm988936	rs11045819	SLCO1B1	All geno	0.001	0.004	3.39	1.48-7.78
BPAR vor c	Caucasian	12	exm988936	rs11045819	SLCO1B1	Dom	<0.0001	0.001	3.69	1.65-8.25
BPAR vor c	Caucasian	2	exm277212	rs3755321	UGT1A5	All geno	0.024	0.021	2.71	1.16-6.29
BPAR vor c	Caucasian	2	exm277212	rs3755321	UGT1A5	Dom	0.018	0.011	2.89	1.28-6.51
Leucopenia	All	7	exm2266441	rs3789243	ABCB1	All geno	0.45	0.013	2.45	1.2-4.97
Leucopenia	All	7	exm2266441	rs3789243	ABCB1	Dom	0.015	0.014	2.35	1.19-4.66
Leucopenia	All	2	exm277163	rs6759892	UGT1A6	All geno	0.037	0.01	0.50	0.3-0.85
Leucopenia	All	2	exm277163	rs6759892	UGT1A6	Dom	0.011	0.004	0.48	0.29-0.8
Leucopenia	Caucasian	2	exm277163	rs6759892	UGT1A6	All geno	0.031	0.02	0.50	0.28-0.9
Leucopenia	Caucasian	2	exm277163	rs6759892	UGT1A6	Dom	0.009	0.013	0.50	0.29-0.86

4.8 Extended candidate gene approach

A further 6 genes (GMPS, ATIC, RRM1, RRM2, PNP and ITPA) were considered in analysis of an 'extended candidate gene' pathway for association with anaemia, leucopenia, and BPAR. This included genes known to be involved in the purine pathway but not known to be directly involved in the MPA pathway. Fifty SNPs in these genes were analysed by allele frequency for the primary outcome events in PLINK as well as dominant and recessive models. No significant associations were found and these results will not be discussed further.

4.9 Unsupervised analysis

All SNPs remaining after QC of the data were then tested for association with the primary outcome measures of leucopenia, anaemia, BPAR and GISE in an unsupervised analysis. Unadjusted p values are presented along with the Bonferroni correction for the unsupervised analysis this corrected for all SNPs (108,111) included in the analysis and hence no results remained statistically significant following BF. The top 5 results will be presented in this thesis. To allow a full understanding of the results produced and methodology used at each step, a complete run through of the results will be presented here for one outcome of interest (leucopenia). Results for the other outcomes of interest will be summarised here with full results shown in Appendix 3.

4.10 Leucopenia

4.10.1 Fisher's exact test following QC of the data

Association between SNPs and leucopenia were analysed by three models: allele frequency, dominant and recessive, with significance calculated using Fishers exact test. No association remained significant after BF correction and only the top five SNP associations with the lowest p-values in each model are shown.

The allelic model (Table 4.18) for the 5 SNPs with the most significant association with leucopenia.

Table 4-18: Results for the five SNPs with the most significant association with leucopenia for the allele frequency model

CHR	Gene	SNP	rs number	P value	Bon f
6	HCP5	exm-rs2596472	rs2596472	0.00002	0.69
6	PPP1R18	exm-rs3129996	rs3129996	0.00009	1
8	CSMD1	exm2273442	rs583087	0.0001	1
19	FBXO46	exm1481621	rs11537711	0.0002	1
17	FOXN1	exm1305650	rs61749867	0.0002	1

For the dominant model the five SNPs with the most significant association with leucopenia are shown in Table 4.19.

Table 4-19: Results for the 5 SNPs with the most significant association with leucopenia for the dominant model

CHR	Gene	SNP	rs number	P value	Bonf
6	HCP5	exm-rs2596472	rs2596472	0.000011	0.35
6	SFTA2	exm-rs3131786	rs3131786	0.000027	0.82
15	ALPK3	exm1184775	rs3803405	0.00008	1
6	DPCR1	exm-rs3132571	rs3132571	0.00011	1
6	PPP1R18	exm-rs3129996	rs3129996	0.00013	1

Table 4.20 shows the results of the recessive model for the five SNPs with the most significant association with leucopenia. It can be seen that two of the top five SNPs are located on chromosome 6. Three of the SNPs, including the two on chromosome 6 are located in between genes and so their functional relevance is difficult to interpret although SNPs up or down stream of genes can alter regulatory elements and so they are still potentially relevant.

Table 4-20: Results for the 5 SNPs with the most significant association with leucopenia using fisher's exact test for the recessive model

CHR	Gene	SNP	rs number	P value	Bonf
9		exm-rs10818918	rs10818918	0.00009	1
17		exm2264634		0.0002	1
6	Between ZFP57 and ZDHHC20P1	exm-rs3131888	rs3131888	0.0004	1
6		exm-rs3131886	rs3131886	0.0005	1
5	DAB2	exm-rs11959928	rs11959928	0.0008	1

4.10.2 Manhattan Plot

A Manhattan plot has been produced for leucopenia from all SNPs on the chip and is shown in Figure 4.12. P- values for the SNPs along the y-axis with the genomic co-ordinates of the SNP on the x-axis and is a useful visual representation of the SNP-event associations across the exome .Several SNPs located within genes on chromosome 6 stand out and correspond to the SNPs reported in tables 4.18-4.20 above. The SNP that appears to be of most significance for association with leucopenia is rs2596472 in gene HCP5 (a human endogenous retrovirus gene). This SNP was shown as the most significantly associated with leucopenia both in the allele frequency and dominant models but this variant did not remain significant following BF correction.

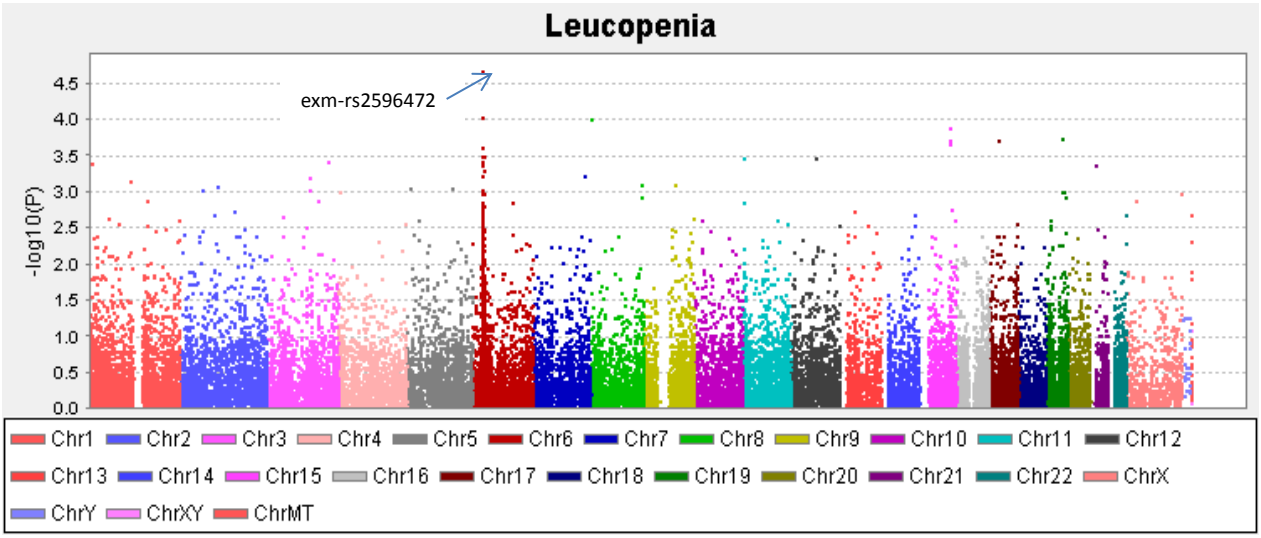


Figure 4-12: Manhattan plot of unsupervised analysis for leucopenia

This Manhattan plot shows a number of SNPs located within genes on chromosome 6 to be associated with leucopenia. The SNP that appears to be of most significance is rs2596472 in gene HCP5. No SNPs remained significant following bonferroni correction.

4.10.3 Logistic regression for genotypic confounders

All the SNPs on the chip remaining after QC of the data were then analysed for genotypic confounders using Logistic regression for PC of the entire cohort. In contrast to the analysis of the candidate gene SNPs a separate analysis for a Caucasian only cohort was not carried out.

A significant p value for unsupervised analysis (prior to BF correction) was taken to be $P<5x10^{-7}$ which is in keeping with that suggested in the literature (92). No results reached this level of significance. Results for Logistic regression are presented in Tables 4.21-4.23. Only the top five most significant SNPs are included here for each of the models. SNPs within genes located on chromosome 6 appear to predominate.

Table 4.21 shows the most significant SNPs associated with leucopenia following Logistic regression for PC of the entire cohort (N=278) using the allele frequency model. Two of the SNPs are located on chromosome 6 but they are not the same SNP reported in Table 4.18. However, the SNP rs3803403 in the ALPK3 gene located on Chromosome 15 was shown to be associated with leucopenia following in the dominant model (Table 4.19).

Table 4-21: Results for the 5 SNPs with the most significant associations with leucopenia following Logistic regression for PC, allele frequency model

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value
6	between MUC22 and HCG22	exm-rs2508015	rs2508015	All F	2.14	0.21	1.4-3.24	0.00032
15	ALPK3	exm1184743	rs3803403	All F	2.14	0.22	1.4-3.27	0.00041
11	OR52E2	exm882093	rs61746343	All F	2.25	0.24	1.4-3.6	0.00081
9	Unspecified	exm-rs10746839	rs10746839	All F	0.5	0.21	0.33-0.76	0.0013
6	Unspecified	exm-rs3131622	rs3131622	All F	1.97	0.22	1.29-3.02	0.0017

Table 4.22 shows the most significant SNPs associated with leucopenia following Logistic regression for PC of the entire cohort (N=278) using the dominant model. Four of the SNPs are located on chromosome 6 again suggesting possible importance of genes on this chromosome. Three of these (rs3131786 in SFTA2 gene, rs3132571 in DPCR1 gene and rs2596472 in HCP5 gene) were also shown to be potentially associated with leucopenia in the dominant model. The SNP rs3803403 in the ALPK3 gene located on Chromosome 15 was shown to be associated with leucopenia also in the dominant model and has been reported here as well. The potential significance of these SNPs will be discussed in chapter 5: Discussion and conclusions.

Table 4-22: Results for the five SNPs with the most significant association with leucopenia following Logistic regression for PC, Dominant model

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value
6	SFTA2	exm-rs3131786	rs3131786	DOM	4.12	0.34	2.1-7.96	0.00002
15	ALPK3	exm1184743	rs3803403	DOM	3.12	0.29	1.8-5.5	0.00006
6	DPCR1	exm-rs3132571	rs3132571	DOM	3.55	0.33	1.9-6.7	0.00007
6	HCP5	exm-rs2596472	rs2596472	DOM	0.23	0.34	0.1-0.46	0.00009
6	NEU1	exm-rs9267649	rs9267649	DOM	0.29	0.39	0.1-0.62	0.0015

Table 4.23 shows the most significant SNPs associated with leucopenia following Logistic regression for PC of the entire cohort (N=278) using the recessive model. These were the same SNP reported as potentially significant in the recessive model above. The potential significance of these SNP will be discussed in Chapter 5: Discussion and conclusions.

Table 4-23: Results for the five SNPs with the most significant associations with leucopenia following Logistic regression for PC, recessive model

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value
17	Unspecified	exm2264634		REC	5.64	0.42	2.5-12.96	0.00004
9	Unspecified	exm-rs10818918	rs10818918	REC	7.66	0.5	2.9-20.6	0.00005
6	between ZFP57 ZDHHC20P1	exm-rs3131888	rs3131888	REC	3.75	0.37	1.8-7.7	0.0003
6	DHX16	exm-rs1076829	rs1076829	REC	4.3	0.41	1.9-9.6	0.0004
5	DAB2	exm-rs11959928	rs11959928	REC	3.79	0.38	1.8-7.9	0.0004

As the unsupervised analysis aims to highlight potential SNPs of interest rather than prove association, a Logistic regression for phenotypic outcome has not been included. This will be discussed in Chapter 5: Discussion and Conclusion.

4.11 Unsupervised analysis summary

Table 4.24 shows a summary of significant results, for all outcomes of interest (anaemia, leucopenia, GISE and BPAR) following Logistic regression for PC of the entire population. A P value<5x10-7 is considered significant prior to BF corrections. Whilst the results presented here are of interest and may suggest areas for further investigation no results reach this level of significance or remain significant following BF correction.

Results are presented in chromosomal order and include the odds ratio and 95% CI.

Table 4-24: Summary of SNPs with the most significant results following logistic regression for Pc of entire cohort

CHR	Gene	SNP	rs number	Outcome	Model	OR	95% CI	P value
1	CDC42BPA	exm154259	rs1929860	BPARall	REC	5.75	2.4-13.9	0.0001
1	between IGSF21 and KLHDC7A	exm-rs3007729	rs3007729	BPARall	All F	2.2	1.4-3.4	0.0007
1	between SRRM1 and CLIC4	exm-rs4601530	rs4601530	BPARall	REC	5.16	2-13.3	0.0006
1	AGBL4	exm2264835	rs657452	Anaemia	REC	3.6	1.8-7.5	0.0004
2	ABCG8	exm190428	rs6544718	Anaemia	All F	0.45	0.3-0.73	0.0013
2	Unspecified	exm-rs7584993	rs7584993	BPARall	REC	5.78	2.5-13.3	3.67E-05
2	Unspecified	exm2254828		BPARall	DOM	0.33	0.2-0.6	0.0006
3	SETD5	exm287880	rs11542009	GISE	DOM	3.6	1.9-7.2	0.0002
3	CLSTN2	exm-rs11708189	rs11708189	Anaemia	All F	1.92	1.3-2.85	0.001
3	TMEM108	exm2255702	rs1197314	GISE	REC	3	1.6-5.5	0.0004
3	AHSG	exm370881	rs4917	Anaemia	REC	3.5	1.7-7.1	0.0007
3	AHSG	exm370882	rs4918	Anaemia	REC	3.5	1.7-7.1	0.0007
4	BOD1L	exm390066	rs3733557	GISE	All F	2.7	1.6-4.5	0.0001
4	BOD1L	exm390066	rs3733557	GISE	DOM	3	1.7-5.3	0.00015
5	DAB2	exm-rs11959928	rs11959928	Leucopenia	REC	3.79	1.8-7.9	0.0004
6	DHX16	exm-rs1076829	rs1076829	Leucopenia	REC	4.3	1.9-9.6	0.0004
6	NUP153	exm-rs12199222	rs12199222	BPARall	REC	4.3	1.8-9.9	0.0007
6	PSORS1C1	exm-rs1265100	rs1265100	GISE	All F	2.6	1.6-4.2	9.00E-05
6	PSORS1C1	exm-rs1265100	rs1265100	GISE	DOM	2.8	1.6-4.9	0.0002
6	NUP153	exm518663	rs2228375	BPARall	REC	4.16	1.9-9.3	0.0005
6	between MUC22 and HCG22	exm-rs2508015	rs2508015	Leucopenia	All F	2.14	1.4-3.24	0.00032
6	MUC22	exm-rs2517554	rs2517554	GISE	DOM	0.3	0.2-0.6	0.0001
6	HCP5	exm-rs2596472	rs2596472	Leucopenia	DOM	0.23	0.1-0.46	9.00E-05
6	Unspecified	exm-rs2621338	rs2621338	Anaemia	DOM	0.35	0.2-0.6	6.15E-05

CHR	Gene	SNP	rs number	Outcome	Model	OR	95% CI	P value
6	Unspecified	exm-rs2621367	rs2621367	Anaemia	All F	0.44	0.3-0.66	6.00E-05
6	Unspecified	exm-rs2621367	rs2621367	Anaemia	DOM	0.35	0.2-0.6	6.15E-05
6	Unspecified	exm-rs3131622	rs3131622	Leucopenia	All F	1.97	1.29-3.02	0.0017
6	SFTA2	exm-rs3131786	rs3131786	Leucopenia	DOM	4.12	2.1-7.96	2.00E-05
6	between ZFP57 and ZDHHC20P1	exm-rs3131888	rs3131888	Leucopenia	REC	3.75	1.8-7.7	0.0003
6	DPCR1	exm-rs3132571	rs3132571	Leucopenia	DOM	3.55	1.9-6.7	7.00E-05
6	NEU1	exm-rs9267649	rs9267649	Leucopenia	DOM	0.29	0.1-0.62	0.0015
7	EGFR	exm-rs11979158	rs11979158	BPARall	All F	2.8	1.5-5	0.0007
7	between CNTNAP2 and MIR548T	exm2270711		GISE	REC	2.8	1.6-4.9	0.0005
7	Unspecified	exm2270569		BPARall	DOM	3.78	1.8-8.0	0.0005
8	TMEM70	exm706306	rs1053077	Anaemia	All F	2.99	1.76-5.1	1.00E-05
8	TMEM70	exm706306	rs1053077	Anaemia	DOM	3.21	1.9-5.5	3.07E-05
8	TMEM70	exm706302	rs1053079	Anaemia	All F	2.99	1.76-5.1	1.00E-05
8	TMEM70	exm706302	rs1053079	Anaemia	DOM	3.21	1.9-5.5	3.07E-05
8	ASAP1	exm720843	rs966185	Anaemia	REC	3.3	1.7-6.3	0.0003
9	Unspecified	exm-rs10746839	rs10746839	Leucopenia	All F	0.5	0.33-0.76	0.0013
9	Unspecified	exm-rs10818918	rs10818918	Leucopenia	REC	7.66	2.9-20.6	5.34E-05
10	MCM10	exm810209	rs2274110	GISE	All F	2.6	1.6-4.5	0.0003
11	OR10A2	exm887083	rs10839632	BPARall	All F	2.3	1.4-3.6	0.0005
11	PDHX	exm900157	rs11539202	BPARall	DOM	3.3	1.7-6.5	0.0005
11	OR52E2	exm882093	rs61746343	Leucopenia	All F	2.25	1.4-3.6	0.00081
12	VWF	exm976501	rs35335161	BPARall	All F	5.4	2.2-13.5	0.0003
12	VWF	exm976501	rs35335161	BPARall	DOM	7.4	2.8-19.9	6.05E-05
12	ZNF605	exm2271816	rs7778	GISE	REC	5.6	2.1-14.8	0.0005
15	ARRDC4	exm1192081	rs2130882	GISE	DOM	0.3	0.2-0.6	0.0002
15	ALPK3	exm1184743	rs3803403	Leucopenia	All F	2.14	1.4-3.27	0.00041

CHR	Gene	SNP	rs number	Outcome	Model	OR	95% CI	P value
15	ALPK3	exm1184743	rs3803403	Leucopenia	DOM	3.12	1.8-5.5	6.00E-05
17	Unspecified	exm2264634	unknown	Leucopenia	REC	5.64	2.5-12.96	3.94E-05
17	ARHGEF15	exm1292049	rs3744647	GISE	REC	3.2	1.7-6.2	0.0005
18	between CDH2 and ARIH2P1	exm-rs11083271	rs11083271	BPARall	All F	0.37	0.2-0.6	0.0004
18	between CDH2 and ARIH2P1	exm-rs11083271	rs11083271	BPARall	DOM	0.24	0.1-0.5	2.98E-05
18	Unspecified	exm2253444		Anaemia	REC	0.25	0.1-0.6	0.0008
18	Unspecified	exm2268111		GISE	All F	0.4	0.3-0.7	0.0001
19	NLRP2	exm1507217	rs1043673	GISE	All F	2.1	1.4-3.2	0.0002
19	DKKL1	exm1490415	rs2288481	GISE	REC	7.7	2.5-23.2	0.0003
21	between KRTAP7-1 and KRTAP11-1	exm-rs7283316	rs7283316	Anaemia	DOM	0.36	0.2-0.6	9.68E-05

4.12 Summary

The most significant findings in the candidate gene approach are those that are supported by the time to event analysis which strengthens the associations seen. The associations supported by time to event analysis were predominantly within the SLCO1B1 and UGT1A genes. This includes the association of rs2291075 and rs2306283, both in the SLCO1B1 gene, which were associated with anaemia in the dominant model. The SNPs rs4149056, rs4149032 and rs11045819, also in the SLCO1B1 gene that were associated with BPAR in the dominant model with rs4149032 and rs11045819 also associated with BPAR vascular or cellular subgroup, as was rs34671512 also in SLCO1B1. The SNP rs3789243 in UGT1A5 was similarly associated with BPAR v or c with the SNP rs6759892 in UGT1A6 associated with leucopenia, both in the dominant model.

Other associations that were seen without support of a time to event association remain of interest.

The results of the unsupervised analysis provide interesting potential associations between SNPs and clinical outcomes, these findings need to be taken in context of the function of the gene within which they are located and will be discussed at length in Chapter 5.

5 Chapter 5: Discussion and conclusions

5.1 Introduction

This chapter will discuss the results presented in chapter 4. The study population and observed phenotypic event rates have already been covered in chapter 3. The candidate gene analysis will be reviewed first followed by the unsupervised analysis. The strengths and limitations of the study will also be discussed. Conclusions and future direction of research will be outlined in Chapter 6.

5.2 Correction for PC and multiple testing

The main strength of population based genetic association studies, such as this one, is that they can test for a large number of SNPs with a high degree of accuracy and do not require knowledge of specific causative or candidate genes. This also gives rise to the two main pitfalls of such studies which are population stratification which can confound associations and multiple testing which reduces the power to identify associations (92, 134, 135).

Population stratification has been corrected for in this study using Logistic regression for the PC of the population to remove individuals exhibiting divergent ancestry. A pure Caucasian cohort has also been analysed separately, this aims to further reduce confounding by analysing a population with pure ancestry as in some cases individuals with the same ethnic origin may be grouped phenotypically and lead to false associations being drawn (92, 134). The main issue with the analysis of a pure Caucasian population in this study is, although they form a majority group, the overall sample size is then reduced in size (N=233).The study population becomes smaller than the sample size calculation and making it difficult to draw robust conclusion, although associations can still be sought.

The second problem of multiple testing gives rise to the issue of selecting the most probable associations. The most effective way is to carry out a single test per SNP but this limits the study to small numbers of likely candidate SNPs. Correction for multiple testing is essential in GWAS studies as multiple testing of the null hypothesis leads to an increase in type 1 error and a high chance of rejecting the null hypotheses when it is in fact true (134). It is also well recognised that the application of approaches for controlling for multiple testing that are too stringent will prevent the detection of some true associations where small genetic effects exist (135). The majority of GWAS to date have either used Bonferoni (BF) correction which multiplies nominal p values or the False discovery rate (FDR) which controls for expected false discoveries amongst the rejected hypotheses (134, 135). The BF correction is recognised as being the more stringent method.

The first MPA pharmacogenetic study published in the literature using a customized chip to simultaneously genotype 2700 SNPs used a FDR of 20% to correct for multiple testing(79). In the study presented in this thesis a BF correction has been applied accounting for full multiple testing in the unsupervised analysis, this subsequently means that no results have remained significant. This is reasonable in an unsupervised review of associations as the main aim is to detect potential associations that may require further investigation. In the candidate gene approach the results have been corrected for the total number of SNP in the candidate gene with which associations have been sort, after accounting for those in complete LD. This less stringent correction was applied to the candidate analysis as the SNPs were carefully selected with pre-existing knowledge of the gene's involvement in MPA metabolism. Despite this no results remains statistically significant following BF correction within the candidate gene model but these results remain of interest.

5.3 Candidate Gene approach

The results of the candidate gene approach will be discussed below. The genes included in this stage of the analysis have known biological plausibility as they are part of the MPA metabolic pathway as discussed in chapter 1. The SNPs included in analysis are summarised in table 4.1 chapter 4. No results remained statistically significant after BF correction for multiple testing. Table 5.1 shows the six most significant outcomes from the candidate gene approach. They relate only the allele frequency model and will be discussed here.

Table 5-1: Summary of candidate gene results

Chr	Gene	rs number	Cohort	Outcome	P value	KM (Log Rank P value)	Cox reg (P value)
2	UGT1A6	rs6759892	All	Leucopenia	0.018	0.037	0.01
2	UGT1A6	rs6759892	Caucasian	Leucopenia	0.019	0.031	0.02
12	SLCO1B1	rs2291075	All	Anaemia	0.006	0.008	0.004
12	SLCO1B1	rs4149032	Caucasian	BPAR (all)	0.003	0.005	0.001
12	SLCO1B1	rs34671512	Caucasian	BPAR (C or V)	0.01	0.001	0.004
12	SLCO1B1	rs414032	Caucasian	BPAR (C or V)	0.006	0.001	<0.0001

This table shows the most significant results from candidate gene analysis; it includes the P value following logistic regression for PC and phenotypic confounders, KM log rank value for time to event analysis and Cox regression analysis p value. Results do not include BF correction for multiple testing.

5.3.1 Leucopenia

The SNP rs6759892 in the UGT1A6 gene was found to be associated with leucopenia both for the entire cohort and for the Caucasian only cohort. This association was seen in the allelic model analysis and in the dominant model. The association showed a benefit of being either heterozygote or alternative homozygote when compared to the reference homozygote for the SNP with the odds of leucopenia reduced by a factor of 2.3 (56%). The association was further supported by association with the secondary outcome measure of time to event analysis.

UGT1A genes area is a locus complex encoded for by UDP-glucuronosyltransferase genes (136, 137). UGT1A8, 1A9, 1A10 and to a lesser extent 1A1 have all been well documented as playing a crucial role in the metabolism of MPA (12, 39, 49, 51, 68)as discussed in chapter 1. To date there have been no documented associations with UGT1A6 and MPA metabolism but studies have demonstrated UGT1A6 SNP to have an impact on glucoronidation (138).

Whilst it is feasible that there is a true association between rs6759892 SNP in UGT1A6 and MPA associated leucopenia, the SNP was further investigated for LD with other UGT1A SNP reported in the literature and within the gene region. This SNP was found to be in complete LD with three SNP in UGT1A9 and 90% LD with a further UGT1A9 SNP. It is therefore more likely that the true association is with the UGT1A9 SNPs. This would be in keeping with the literature reporting SNPs in UGT1A9 being associated with adverse outcomes or variability in pharmacokinetics in MPA treated individuals (16, 51, 53, 65, 67, 68, 72, 139) although there are no studies reporting these particular SNPs. The SNP rs6759892 was part of the 2011 study by Jacaboson et al that used a broad panel of SNPs to look at the association between MPA and leucopenia but the results for this SNP are not presented in the paper and hence it is assumed that they did not find a significant(79).

UGT's are the predominant genes involved in the metabolism of MPA by gluoronidation to the inactive form MPAG, and biliary excretion and reuptake (12), with UGT1A9 being responsible for the majority of glucoronidation within the liver (68). SNP's in UGT1A9 such as those described above are likely to alter the function of this gene. This may reduce UGT1A9 activity leading to a reduction MPA metabolism, increasing IMPDH inhibition and restricting lymphocyte proliferation leading to leucopenia (50). SNP's may also have the opposite effect and increase UGT1A9 activity and hence reduce the likelihood of leucopenia, this is likely to be the case with the SNP's found to be associated with leucopenia in this study as being non-wild type for the SNP reduced the likelihood of leucopenia.

The remaining SNPs that were associated with leucopenia in this study did not remain significant for all steps of the analysis, whilst some were supported by time to leucopenia analysis robust conclusions cannot be drawn.

The results for the SNP rs6759892 discussed here are of significant interest, biologically plausible and represent novel findings that have not previously been reported but require further validation.

5.3.2 Anaemia

The SNP rs2291075 in the SLCO1B1 gene was found to be associated with anaemia in the entire cohort. This association was seen in the allelic model analysis and in the dominant model. This association showed a benefit of being heterozygote or alternative homozygote for this SNP, with the odds of anaemia reduced by a factor of 2.38 (58%). The association was further supported by association with the secondary outcome measure of time to event analysis. This finding could have clinical application in terms of identifying those who have innate protection against anaemia.

The LD plot for this SNP showed that is in LD with rs17329885, rs6487213, rs2306283 and rs6487213 which are also within the SLCO1B1 gene, meaning the true causative SNP may be one of these.

Solute carrier organic anion transporter family member 1B1 (SLCO1B1) encodes for organic anion transporter polypeptide 1 (OATP1) and its main role in MPA metabolism is hepatic uptake of MPA which is a crucial step in hepatic clearance as discussed in chapter 1. The SLCO1B1 SNPs seen to be associated with a reduced risk of anaemia are likely to cause an increase in activity within the gene, increasing hepatic uptake and more rapid clearance of MPA from the system. This is also supported by the association of rs2306283 SNP and a reduction in LGISE, reduction in post-transplant infection episodes and a reduced likelihood of needing to stop MPA for any reason.

Several studies have investigated the association with SNPs in SLCO1B1 and both pharmacokinetics and clinical outcome in MPA treated patients (6, 8, 78, 79). Bouamar et al 2011 studied 4 SNPs in SLCO1B1 including rs2306283 but did not find any significant association with MPA AUC or SE of MPA(6). Jacobson et al 2011 looked at 4 SNPs in SLCO1B1 including both rs2291075 and rs2306283 in relation to MPA induced anaemia but did not report any findings for these SNPs and so presumably significance was not seen (79). Michelon et al 2010 did report a significant association between rs4149056 with homozygotes for the SNP being protected from MPA- related adverse drug reactions (OR 0.15)(8). Although reported in a different SNP in

SLCO1B1, these results also suggest a benefit for the non-wild genotypes in relation to MPA induced anaemia as were seen in this study.

The remaining SNPs that were associated with anaemia in this study were not supported at all stages of analysis but remain of interest but require further validation

5.3.3 Biopsy Proven Acute Rejection

Biopsy proven acute rejection (BPAR) was categorised as 'ALL' (BPARall) which included any reported rejection at biopsy and 'Vascular or Cellular' (BPARvorc) when including cases at the more extreme end of the spectrum with evidence of vascular or cellular rejection on biopsy, but does not include those reported as borderline rejection.

The SNPs rs34671512 and rs41490320 in the SLCO1B1 gene were found to be associated with BPAR in the Caucasian only cohort.

The SNP rs4149032 showed a benefit of being alternative homozygote with the odds of BPAR (all) increased by a factor of 3.77 for either reference homozygote or heterozygote. The time to event analysis supported this association. This SNP showed similar associations for BPAR (v or c) with the odds of BPAR where increased by a factor of 4.7 for either reference homozygote or heterozygotes versus alternative homozygotes. This association was again supported by time to event analysis including Cox regression.

Significant associations seen both for a reduction in BPAR (all) and BPAR (v or c) in individuals who were alternative homozygote for SNP rs4149032 could have important clinical impact. This may be particularly relevant for individuals who are relatively frail or those with a history of malignancy, in whom a lower level of immunosuppression could be used.

The SNP rs34671512 showed an increased odds of BPAR (V or C) by a factor of 3.7 in individuals that were either alternative homozygote or heterozygote compared to reference homozygote. The time to event and cox regression analysis supported the association.

The SLCO genes encode for the OATP transporters which play a role in MPA metabolism the uptake of MPAG into the hepatocytes (49, 133) as discussed in chapter 1. The SNP rs4149032 may lead to reduced function in this gene as it is associated with a decreased likelihood of BPAR, suggesting a reduction in MPA clearance. The SNP rs34671512 seems to have the reverse effect with an increased association with BPAR suggesting more rapid metabolism of MPA its inactive form. Several studies have looked at SNPs in SLCO1B1 and both MPA pharmacokinetic and clinical outcomes (6, 8, 78, 79) with variable associations reported. A single published study by Michelon et al 2010 (8) has looked specifically at SLCO1B1 in relation to BPAR. This study looked

at 3 SNPs rs4149015, rs2300283 and rs4149056. Michelon et al did not find any significant associations with these SNPs and BPAR but they did report a statistically significant association with MPA-related ADR and rs4149015, although the individual aspects of the ADR were not reported separately. The study also looked at MPA-AUC measurements, but did not find any significant associations (8). No other reported studies have looked at SLCO1B1 in relation to MPA and BPAR and the findings here are therefore novel.

The remaining SNPs that showed some association with BPAR were not supported at all a=stages of analysis and hence do not allow robust conclusions to be drawn, but they should be considered for further analysis.

5.3.4 Gastrointestinal side-effects, MPA cessation or reduction and infection

A number of SNPs showed some association with the remaining outcomes of interest of GISE, MPA cessation or dose reduction and post-transplant infection but these associations did not remain at all stages of the analysis.

The UGT1A9 SNP rs2602381 showed some association with GISE suggesting a benefit in those with reference homozygotes for the SNP with the odds of GISE being increased by a factor of 2.17 compared to heterozygote or alternative homozygote for the SNP. Similar results were seen when analysed for associations with lower GISE only. The UGT1A6 SNP rs6758992 showed a reduction in UGISE by a factor of 3 for those individuals with heterozygote or alternative homozygote when compared to reference homozygote. As discussed for leucopenia this SNP is in complete LD with a number of UGT1A9 SNPs and it is more likely that the true association is with a UGT1A9 SNP.

UGT1A9 is the main enzyme involved in glucuronidation of MPA to MPAG and is predominantly found in the liver (49, 133). The UGT's SNP presented above may alter gene function leading to altered MPA metabolism. Decreased UGT1A9 will reduce MPA clearance increasing the likelihood of increased toxicity and side effect, whilst increased UGT1A9 function may increase gut exposure to MPA potentially increasing GI side effects. As discussed in chapter 1, the UGT genes have been investigated in several MPA related pharmacogenomics studies looking at both pharmacokinetics and clinical outcomes (16, 52, 65, 67, 72, 104, 140). Woillard et al 2009 studied several SNP including one in UGT1A9 and risk of diarrhoea, but they did not find the UGT1A9 SNP be associated (104). A small paediatric study showed a significant association with UGT1A9 SNP rs2741046 and adverse events including diarrhoea but the number of events were small (67). While there is much assumption that SNP in UGT genes are associated with GISE in MPA treated

RTR there is no conclusive studies in the literature. The SNPs found in this study are previously unreported and hence present new findings.

These findings can be of potential benefit in identifying patients at risk of GISE when given MPAP for transplantation or indeed, other indications.

Whilst some possible associations were seen with SNPs and MPA dose reduction or cessation and with infection they were not supported at all stages of analysis and so whilst they remain of interest they do not add anything conclusive to the body of evidence.

5.4 Unsupervised analysis

The methodology used in this study has the advantage of looking beyond a candidate gene approach, potentially allowing identification of novel associations. This type of unsupervised analysis must be subject to stringent QC and multiple testing correction as there is high probability of type 1 error (92). As a result of this no SNP in the unsupervised analysis remained statistically significant following the Bonferroni correction for multiple testing. Whilst this means that robust conclusions cannot be drawn there is also the risk of disregarding SNPs that are plausible. SNPs with a tendency towards significance merit further investigation or if the results suggest involvement of specific genes in the metabolic pathway of MPA that have not previously been considered.

The top 5 SNPs for the phenotypes leucopenia, anaemia, BPAR and GISE have been reported in results chapter 4. Here the literature relating to these SNPs that have a tendency towards an association and the genes that they relate to will be reviewed. The plausibility of each in the context of MPA and the phenotype will be discussed with the aim to guide future research.

Each of these SNPs has also been investigated for LD with SNPs in candidate genes known to be involved in MPA metabolism. No SNPs showed significant LD with SNPs in candidate genes and it is therefore assumed that any tendency towards an association is not reflecting the effect of a candidate SNP inherited in conjunction with a SNP in a novel gene.

The results discussed below relate the SNPs and results summarised in table 4.24 in chapter 4.

5.4.1 Leucopenia

Surfactant associated protein 2 (*SFTA2*)

Surfactant associated protein 2 (*SFTA 2*) located on chromosome 6 is also referred to as surfactant protein G (SP-G). It is a recently investigated novel protein which is highly expressed in the lungs with wide spread expression in other tissues at lower levels including lymphocytes, bone marrow, kidneys and ureter (141, 142). There have been no reported associations of SNP in this gene

reported in the literature to date. The SNP rs3131622 is a common variant with a reported MAF = 0.43 in a study of 935 individuals. Whilst the lack of literature makes it difficult to speculate as to the significance of the association seen with leucopenia in MPA treated patients, the present of SP-G protein in lymphocytes and the bone marrow means that SNPs in this gene could be considered in future research as they may lead to functional change.

Alpha kinase-3 (*ALPK3*)

Alpha kinase-3 (*ALPK3*) is located on chromosome 15. There is little reported in the literature relating to this SNP with the exception of a GWAS study looking at severe neutropenia or leucopenia induced by chemotherapeutic agents in Japanese population. A SNP in ALPK3 (rs12900463) was found to be associated with leucopenia in this study(143). The SNP rs3803403 occurs with a MAF= 0.16/357. There are no reported studies in the literature including this SNP, but the association of another SNP in the gene with drug induced leucopenia means it should not be disregarded.

Diffuse panbronchiolitis critical region 1 (*DPCR1*)

Diffuse panbronchiolitis critical region 1 (*DPCR1*) gene is a well-known genetic marker for diffuse panbronchiolitis. Studies have looked at associations with SNPs in this gene and bronchiolitis (144) but there are no studies looking at leucopenia. The SNP rs3132571 is common with a MAF=0.42/914. This gene is an unlikely candidate for MPA induced leucopenia.

HLA complex P5 (*HCP5*)

HLA complex P5 is a human endogenous retrovirus that has become a part of the human genome. The reports in the literature relating to polymorphisms in this gene focus on hypersensitivity to Abacavir for the treatment of HIV and have found associations with SNP rs2395029 (145, 146). These are no reported studies looking specifically at drug related leucopenia or reporting rs2596472, which has a MAF=0.23/492. There is no evidence within the literature to suggest that this is a likely candidate for MPA induced leucopenia.

NEU1

NEU1 gene encodes for lysosomal neuraminidase enzyme Sialidase 1 (lysosomal sialidase), also known as NEU1. This gene has been shown to have a negative effect on the regulation of lysosomal exocytosis(147) and NEU1 is present on the surface of activated T cells with a significant effect on macrophage function(148). It is felt to have an effect on immune function and a study has shown it to be involved in airway epithelial response to inflammation (149).

A large 2009 study which looked at panel of SNP in relation to development of SLE included rs9267649 but no significant associations were seen (150).

Whilst there is no evidence in the literature that SNP in this gene are related to MPA induced leucopenia, the presence of the protein encoded for by this gene on T cells and potential immune modulatory effect means that this gene would warrant further consideration.

DEAH (Asp-Glu-Ala-His) box polypeptide 16 (*DHX16*)

DHX16 (DEAH (Asp-Glu-Ala-His) box polypeptide 16) is a protein-coding gene which contributes to pre-mRNA splicing(151, 152). There are no published studies to date looking at SNP in this gene and hence evidence to support an expected association with MPA induced leucopenia.

Dab, mitogen-responsive phosphoprotein, homolog 2 (*DAB2*)

DAB2 (Dab, mitogen-responsive phosphoprotein, homolog 2) codes for a cytoplasmic adaptor protein expressed in renal proximal tubular cells. Rs11959928 SNP has been shown in a large replicated study to be associated with chronic kidney disease (CKD) (153). A further SNP in the gene was studies with relation to MPA related leucopenia by Jacobson et al in 2011 but no results of this SNP were published (79). There is no evidence in the literature to support an association with MPA induced leucopenia but there is significant evidence linking it to CKD and hence it should be considered in further GWAS studies relating to renal disease.

5.4.2 Anaemia

Transmembrane protein 70 (*TMEM70*)

Transmembrane protein 70 gene encodes a mitochondrial membrane protein. Deficiency in *TMEM70* underlies most cases of ATP synthase deficiency (154, 155). There are no SNP studied in this gene in relation to pharmacogenomics or anaemia. There were two separate SNPs in this gene that have a tendency towards significance for MPA associated anaemia (rs1053079 and rs1053077) which would suggest a true association, however when investigated further they are in complete LD with each other with MAF = 0.25/540. There is no evidence to support an association between SNP in this gene and MPA induced anaemia.

ArfGAP with SH3 domain, ankyrin repeat and PH domain 1 gene (*ASAP1*)

ArfGAP with SH3 domain, ankyrin repeat and PH domain 1 gene (*ASAP1*) encodes an ADP-ribosylation factor (ARF) GTPase-activating protein that functions on membrane surfaces to catalyse the hydrolysis of GTP bound to Arf (156, 157). There have been no published studies that look at SNPs in the ASAP1 gene and drug induced anaemia. However MPA is a selective reversible inhibitor of lymphocyte IMPDH and decreases lymphocyte GTP concentration with an elevation in

red cell GTP(64, 158). One study concluded that intracellular GTP acts as an antagonist to MPA by directly binding to IMPDH (60). Whilst these studies do not directly suggest a link between GTPase- activating proteins (such as that encoded by ASAP1) and MPA metabolism, they should be investigated further given the potential for SNPs in this gene to alter enzymatic function which would impact on hydrolysed GTP.

ATP/GTP binding protein-like 4 (*AGBL4*)

ATP/GTP binding protein-like 4 is a protein coding gene. There have been no recent publications relating to this gene and it has not been studied in the context of MPA or anaemia. The SNP rs657452 is common with a MAF 0.48/1043. Whilst there is no supporting evidence in the literature to suggest that there is a true relationship between SNPs in this gene and MPA associated anaemia, the link to GTP means that it should not be discounted as discussed for ASAP1 gene previously.

Alpha-2-HS-glycoprotein gene (*AHSG*)

Alpha-2-HS-glycoprotein gene (*AHSG*) encodes a protein known as fetuin-A. Fetuin A is a negative acute phase reactant which is considered to be a major inhibitor in arterial calcification. SNP in this gene (including rs4917) have been studied in relation to fetuin-A levels and arterial calcification with variable findings (159, 160). There have been no studies in the transplant population or relating to MPA. Two SNPs had a tendency towards significance for MPA associated anaemia in this study rs4017 and rs4918, these were in complete LD with each other and have a MAF=0.29/641. The mechanism of anaemia in the post-transplant population is complex and it is feasible that a negative acute phase reactant protein may play a role, but it is unlikely that this is related to MPA metabolism.

5.4.3 BPAR

Von Willebrand factor (*VWF*)

Von Willebrand factor gene codes for VWF protein involved in haemostasis by promoting platelet adhesion and aggregation at the site of vascular injury. SNPs this gene have not been studied in relation to MPA or post-transplant rejection. It would be feasible to see a relationship with SNP in this gene and anaemia but there is no evidence to suggest that it would have an impact either on MPA metabolism or rejection rates. The SNP rs35335161 has a MAF=0.026/56 and hence has not been extensively studied.

Pyruvate dehydrogenase complex (*PDHX*)

Pyruvate dehydrogenase complex (*PDHX)* gene encodes pyruvate dehydrogenase binding protein E3. Mutations in this gene are a well-known cause of metabolic disturbance in infants (161). There are no studies of SNPs in this gene and either MPA metabolism or BPAR and the function of the resultant protein does not suggest a link. The SNP rs11539202 has a MAF =0.15/328.

CDC42 binding protein kinase alpha (*CDC42BPA*)

CDC42 binding protein kinase alpha gene and is also known as myotonic dystrophy kinase-related CDC42-binding protein kinase alpha due to its proven involvement in this condition (162). It has also been shown to be associated with endotheliitis but there have been no studies of SNP in this gene and relation to BPAR in transplant recipients. Two SNP in this gene including rs1929860 were studied by Jacobson et al 2011 in relation to MPA associated leucopenia and anaemia (but not BPAR) there were no published associations with these SNPs (79). The SNP rs1929860 has a MAF =0.33/724. Although there is no published literature supporting the association of this SNP with BPAR the implication of this gene in endothiliitis means that it should be considered further.

Nuclear pore complex protein 153 (*NUP153*)

Nuclear pore complex protein 153 gene encodes for NUP153 protein involved in the transport of macromolecules between cell nucleus and cytoplasm. A 2012 GWAS study by Datta et al looked at SNP in various genes that were related to bilirubin conjugation. They found that a SNP in *NUP153* (rs2328136) was associated with raised unconjugated bilirubin (UCB) levels (163). It is well recognised that raised UCB levels can result in altered drug metabolism. Given that biliary excretion of MPA is a major part of its metabolic pathway it is feasible that SNPs in NUP153 gene may result in alterations in MPA metabolism. Two SNPs in this gene rs2228375 and rs12199222 had a tendency towards a significant association with BPAR in this MPA cohort, both in the recessive model. The SNPs are not in LD with each other. The reported MAF for these SNPs are MAF=0.14/308 MAF= 0.27/589 respectively. The association of two different SNPs and the proven association with SNPs in NUP153 and altered UCB levels suggest that there is a potential relationship that should be further investigated.

5.4.4 GISE

Mucin 22 (*MUC22*)

Mucin 22 gene is synonymous with pancbronchiolitis related mucin-like 1 gene which is associated with inflammatory respiratory conditions. There have been no SNP studies in the literature with relation to this gene and MPA or gastrointestinal disturbance. The SNP rs2517554

has a reported MAF=0.496/1080. There is no evidence to suggest a true association exists between this SNP, MPA and GISE.

Biorientation of chromosomes in cell division 1-Like 1 (*BOD1L*)

Biorientation of chromosomes in cell division 1-Like 1 codes for protein involved in DNA binding. There have been no reported studies in the literature of associations with SNP in this gene. There is no literature to support an association with MPA and GISE. The SNP rs3733557 has a reported MAF=0.11/244.

SET domain containing 5 (*SETD5*)

SET domain containing 5 genes is a protein coding gene. There have been no studies published in the literature of associations with SNPs in this gene and no evidence to support associations with MPA and GISE. The SNP rs11542009 has MAF=0.045/99.

Psoriasis susceptibility 1 candidate 1 (*PSORS1C1*)

Psoriasis susceptibility 1 candidate 1 gene, previously known as SEEK1, has been shown to be involved in psoriasis. SNP studies in relation to this gene have largely focused on psoriasis and inflammatory bowel disease but have not conclusively proven an association between SNP in this gene and either phenotypic outcome (164, 165). There are no studies to support a potential association with either MPA or GISE in the literature. The SNP rs1265100 has MAF=0.22/479

Arrestin domain containing 4 gene (*ARRDC4*)

Arrestin domain containing 4 genes, which has been linked to the development of congenital diaphragmatic hernias (166). There have been no studies published in the literature of associations with SNPs in this gene and no evidence to support associations with MPA and GISE. It may suggest individuals with SNP in this gene are more likely to have diaphragmatic herniation, predisposing them to UGISE but this would not relate to the use of MPA. The SNP rs11542009 has MAF=0.045/99.

Dickkopf-like 1 (*DKKL1*)

Dickkopf-like 1 gene plays an important role in testicular development and spermatogenesis (167). There have been no studies published in the literature of associations with SNPs in this gene and no evidence to support associations with MPA and GISE. The SNP rs2288481 has MAF =0.23/489.

Transmembrane protein 108 (*TMEM108*)

Transmembrane protein 108 gene encodes a mitochondrial membrane protein. There have been no studies published in the literature of associations with SNPs in this gene and no evidence to support associations with MPA and GISE. The SNP rs1197314 has MAF =0.39/582.

Rho GTPase activating protein 15 (*ARHGEF15*)

Rho GTPase activating protein 15 gene which in laboratory studies has been shown to have an inhibitory effect on angiogenesis (168). There have been no studies published in the literature of associations with SNPs in this gene and no evidence to support associations with MPA and GISE. The SNP rs3744647 has MAF=0.39/857.

Zinc finger protein 605 (*ZNF605*)

Zinc finger protein 605 gene is a protein coding gene. There have been no studies published in the literature of associations with SNPs in this gene and no evidence to support associations with MPA and GISE. The SNP rs7778 has MAF=0.46/1001.

5.5 Unsupervised analysis discussion

In addition to a review of the genes in which SNPs have shown a possible association with outcome of interest it is important to study the manhattan plots as areas that show clustering of SNPs that near or reach the significance line may suggest important potential associations and warrant further analysis.

Both the manhattan plot for leucopenia (Figure 4-12) and anaemia (Figure 7-1) showed clustering of SNPs with a trend towards significance within chromosome 6. These SNPs were then reviewed to see if they were within the HLA complex as numerous studies have shown associations between HLA alleles and drug hypersensitivity (169, 170), however none of the SNPs within these clusters were within the HLA complex. Further analysis of these clusters was not carried out in this study but it would be interesting to carry out clump analysis (96, 171) of these clusters to see if they are of overall significance

Having reviewed the literature there are a number of genes in which SNPs in this study suggested an association and whilst none are conclusive they should be considered in further MPA pharmacogenomics research. There are a number of genes that have little biological plausibility with either MPA or the phenotypic outcome of interest and at present no evidence to take them forward into future studies. In those with some biological plausibility evidence to suggest a true association will be strengthened if similar results are shown in a replication study.

5.6 Strengths and weaknesses

This study has a number of strengths both in design and execution. The design of this study means that it is inclusive and therefore representative of the transplant population as discussed in validity. The data collection was comprehensive, limiting the chance of missing vital information which would impact on the results. It has looked at several clinically important outcome measures making it clinically relevant and applicable, although the use of several outcomes makes it difficult to correctly power the study.

The genetic methodology used allowed rapid genotyping of thousands of SNP in a cheap and effective way allowing vast quantities of genetic information to be produced. Whilst future research could include a CHIP which is drug tailored to capture more relevant SNPs, this study highlights that fast and effective screening is available and that clinical application is both feasible and affordable.

A limitation of this study (and of pharmacogenomics studies in general) is the failure to yield statistically significant results due to the explorative nature of the study design and testing of multiple variants. Clinical studies usually have clear guidelines for interpretation and statistical analysis with standard values, such as the p value, that is widely accepted. The analysis of pharmacogenomics studies is far more complicated, requiring a multistage analysis process with several quality control and correction steps. This limits the likelihood of achieving statistical significance (92). This is particularly true of the application of P value which infers a distinct cut off between significance and not significant which in GWAS is often not the case. The dependence of power on minor allele frequency and the impact of sample size on interpretation are also challenging (172). With the addition of correction factors the chance of finding a significant test is very low particularly in studies, such as this one where large numbers of SNPs are simultaneously studied. Drawing robust conclusions are therefore difficult without very large numbers of subjects, a very large effect size or a high MAF. Whilst this is a limitation in finding a conclusive result, studies such as this are still hugely valuable in both adding to the body of knowledge shaping the future direction of research.

Evidence based medicine applies guidelines which have been rigorously developed from the outcomes of well conducted randomised controlled trials (RCTs) and it has been argued that 'personalised medicine' may require an alternative approach (173). RCTs usually rely on a well-defined selected group of individuals as therefore results can be construed to the 'average' population. Tailoring therapy to the individual requires a unique approach to each patient taking into account their individual differences which set them aside from the 'average'. Whilst evidence

based guidelines are essential to ensure patient safety and efficacy they should not limit the application of a personalised medicine or the use of pharmacogenetics to assist clinical decision making

One limitation of using exome beadchip for genotyping as that it will miss a fraction of variants(102) and coverage maybe limited in some genes. Variants in promoters and other intron located regulatory elements are not included on the chip. This is particularly noted in this study as there is limited coverage of SNPs in IMPDH1 and IMPDH2 genes location on chromosomes 7 and 3 respectively (74, 174). A number of reported SNPs of interest in these genes are located in the translated intronic, proximal promoter region and the 5'/3' untranslated regions which were not covered by the methodology used. IMPDH1 and 2 are important genes in the metabolic pathway of MPA. The lack of representation on the exome beadchip means that no associations could be sought between SNPs in these genes and the phenotypic outcomes of interest.

A further limitation is the dramatic reduction in SNPs included in analysis following quality control of the data. The number of SNPs that passed QC and were carried forward for analysis reduced from 242,901 to 108,111 in this study as presented in Chapter 4. This had a dramatic impact as the number of SNP in the MPA candidate genes reduced from 303 prior to QC of the data down to 39, significantly limiting the breadth of coverage. The requirement to QC data in GWAS studies is well recognised and widely accepted (93-95) but its requirement for common variants has been debated (95). The majority of SNP removed in the QC process were due to MAF and a small number due to poor call rate. The importance of identifying these SNPs is highlighted in Chapter 4 with regards to CES1 rs62028647 where pre QC results suggested a significant association in the MPA cohort but the SNP has very poor call rates across the entire cohort of 2400 samples included. When this was further investigated using rtPCR to check results it is likely that a pseudogene was present leading to double calling of two SNPs, meaning it was appropriately removed during QC. If QC had not been carried out incorrect associations may have been reported.

This highlights a limitation of the technique used, as ideally extended coverage of SNPs in the candidate genes using conventional GWAS technique would yield more clinically feasible results, especially in those with high MAF. The downside to using a conventional GWAS technique is the lack of coverage of rare variants which was provided by the exom beadchip, although larger cohorts would be required to allow associations with these rare variants to be proven.

5.7 Validity, confounding and bias

The study was carefully planned to be inclusive and hence representative of the transplant population within the recruiting centre. The study population has been compared to the UK transplant population in detail in chapter 3: Results 1. The main deviation from the UK transplant population was the large majority of Caucasian individuals and hence under representation of other ethnic groups. This has the benefit of allowing a substantial size sub group analysis for a Caucasian only cohort, although the number of individuals in this group is less than the sample size calculation.

It is felt that the results of this study are generalizable to the UK transplant population.

One problem with pharmacogenomics studies in acute transplantation is that immunosuppressive regimes change as newer drugs become available or new evidence is published suggesting superiority of certain drugs. MPA is currently in almost universal use for acute transplantation across the UK with proven benefits, there has also been the introduction of MPS which is said to have fewer GISE side effects when compared to MMF. Furthermore over the past decade there has been an increased use of tacrolimus in preference to cyclosporine A as the CNI of choice. The study population recruited received cyclosporine A in 64.1 % of cases compared to tacrolimus in 35.9%. Protocols within the recruiting unit have now changed so that all new transplants receive tacrolimus. More recently the use of once a day slow release tacrolimus has been introduced with increasing popularity. This is also the case with induction agents (like basiliximab) with 70% of the recruited subjects receiving an induction agent, the use of which has increased dramatically in the last 2 years. The preparation of MPA (MMF or MPS), type of CNI (Cyclosporine A or tacrolimus) and the use of an induction agent have been included in the Logistic regression for confounding factors to account for these differences. However it does leads to some difficulty in application of these results to the new transplant population in the recruiting centre as the outcomes may have differed if all individuals had received tacrolimus. This further supports the need for a repetition study for validation of the results found, this with be discussed further in future research.

Bias is an important consideration in all research studies and in a complex study such as this vigilance in the design phase has been crucial in reducing bias, although complete elimination of bias is not realistic.

All data was collected by the Principal Investigator of this study (myself), who remained blinded to the results of the genetic testing during the data collection, this will minimise observer bias and the outcomes were not being reported for the study but rather collected retrospectively. The use of a single person collecting the entire data set also maintains consistency with identical decision

relating to outcome events being made. Observer bias cannot be completely removed as both patients and clinicians may be more likely to report certain outcomes, such as GI side effects, or make dose alterations because of them due to knowledge that MPAP may cause GI upset.

Reporting bias is an important factor to consider especially with the retrospective nature with which much of the data was collect. Some of the data collected was reliant on good documentation and accurate reporting by the clinicians involved in the patient care. This was particularly apparent when collecting data relating to GISE as it was entirely reliant on what was documented in clinical notes. A further level of difficultly here is variability in reporting as individuals put differing emphasis on symptoms depending on their level of personal tolerance or perceived normality. It is impossible to fully exclude reporting bias either from the subject or the attending clinicians but it is felt that the use of several sources to collect the data and the blind nature in which it was collected (neither the subjects nor the clinicians were asked to provide this data for the study) should minimise this.

Data relating the leucopenia and anaemia was based on laboratory results and therefore greatly reducing reporter or observer bias, a level of bias may still be introduced as either side effect may be attributed to MPA by the clinician often without further investigations.

The data collected in relation to renal biopsy reporting was done so in a blinded fashion with the individual reporting the biopsy being unaware of the study and the individual reviewing the reports blinded to the genetic results which should eliminate bias. One difficulty with this approach is the non-standardised way in which the biopsies were reported by 4 separate histopathologists and the lack of reporting a BANF criteria which means that they are open to interpretation bias.

Recall bias was not felt to be a feature of this study as all data was collected from electronic or clinical records and participants or clinicians were not required to provide any information so no recollection was required.

Ascertainment bias is important to consider especially in genetic studies where family pedigree can distort data. The extensive QC of the genetic data, including cryptic relatedness should significantly reduce this. Ascertainmemt bias may still be at play in this cohort as individuals receiving a renal transplant are intensively surveyed for the primary outcome measures of this study as part of their routine post-transplant care which may lead to over reporting of some side effects. This should not be the case for the laboratory measured outcomes or BPAR, although as outlined previously these could be incorrectly attributed to MPA.

The nature of the study population of acute RTR means that several confounding factors exist. RTR receive several new drugs, have differing levels of graft function and a variety of underlying medical conditions and medications. It is impossible in this study population, regardless of the study design, to account for all potential confounding factors as many are unknown. An additional level of complexity is adding by the need to consider both genetic and phenotypic confounders with the application at a different part of the analysis stage. Great lengths have been taken in both the design and analysis stages of this study to account for both genetic and phenotypic confounding factors. It is appreciated that confounders that have not been considered, such as prior use of immunomodulatory drugs before transplantation, certain co-morbidities or other drugs, may be at play in this study population. This further highlights the need to validate this research with a repetition cohort which will greatly reduce any the impact of confounding factors.

Another potential confounding factor is that of gut microbiome which may alter drug metabolism and influence outcomes(175). This is particularly important in MPA metabolism as it undergoes extensive enterohepatic recirculation with deconjugation of MPAG back to active MPA by colonic bacteria and reabsorption (42). The degree to which gut bacteria contribute to overall MPA metabolism is unknown but the 'double peek' in MPA levels suggests a significant role. The human gut contains many trillion bacteria and levels are known to be significantly higher in patients with chronic kidney disease (176), it is therefore possible that individuals with differing levels of renal transplant function may have variable MPA metabolism due to differing gut microbiota.

A further confounding factor which has not been considered in this study is that of epigenetics. Epigenetics factors are heritable changes that may lead to alterations in gene expression but are not actually changes in the DNA sequence within the individual (177-179). Epigenetics is accepted as the reason for altered gene expressions within cells and for environmental adaption. Epigenetics is widely accepted as an important aspect of disease development but more recently there is growing interest in its role in drug metabolism. (179, 180). It is felt that an understanding of both phamacogenetic and epigenetic factors is important in understanding interindividual variability in drug response (180). This is perhaps the most important confounding factor within this study.

5.8 Summary

This study has demonstrated some interesting and previously unreported outcome sin the field of mycophenolate pharmacogenetics in RTR. A number of SNPs in candidate genes have shown associations with the primary and secondary outcome measures which are supported by biological plausibility. The study has demonstrated the use of a novel approach to array based exome SNP genotyping using Illumina Human exome Beadchip v1.1. The strengths of this technique have been highlighted along with the limitations. This study has significantly contributed to the rapidly growing body of evidence in the field of mycopenolate pharmacogenomics in the renal transplant population and provides a valuable step towards individualisation of transplant immunosuppression. The genetic methodology has proven simple and affordable and the clinical outcome measures relevant and easy to attain. The results of this study demonstrate the need for further research in this area particularly in the form of a large prospective multi cantered trial.

6 Chapter 6: Summary and future research

6.1 Introduction

This chapter will summarise the most significant outcome from this study and discuss the future direction of research in the field of transplant pharmacogenomics.

6.2 Summary of results

The results that are of most significance, and of potential clinical importance, will be summarised here.

The SNPs rs2291075 and rs2306283 both in the *SLCO1B1* gene were associated with anaemia in the dominant model. These SNPs were found to be protective against anaemia with those individuals with non-wild genotype. This association was seen for both the entire cohort and when analysed for the subgroup of Caucasians only. The association was further supported by the time to event analysis.

Two SNPs rs4149056 and rs11045819 in the *SLCO1B1* gene were associated with BPAR in the dominant model. The SNP rs4149056 was associated with a reduced risk of BPAR (all) in those with the non-wild genotype and was supported by time to event analysis. The same SNP also showed a protective effect for BPAR (v or c) in the Caucasian only subgroup. The SNP rs11045819 was associated with an increased risk of BPAR (v or c) in those with non-wild genotype which was supported by time to event analysis in the Caucasian subgroup. Several other SNPs in the SLCO1B1 gene also showed an association with BPAR in the Caucasian subgroup with rs4149032 and rs34671512, both showing an increased risk of BPAR in those with non-wild genotype and supported by the time to event analysis.

The SNP rs3755321 in the UGT1A5 gene was associated with BPAR in the Caucasian subgroup, with individuals who were non-wild genotype being more like to have BPAR, time to event analysis also supported this association.

Three SNPs rs6759892, rs2070959 and rs1105879 all in the *UGT1A6* gene were associated with leucopenia, with individuals with non-wild phenotype being less like to develop this outcome. Rs6759892 and rs2070959 showed similar results in the Caucasian only subgroup however only rs6759892 was supported by time to event analysis. This SNP was also in complete LD with a number of SNPs in the UGT1A9 gene which may represent the true causative variant.

All these variants are potentially useful, pointing to the use of an alternative immunosuppressant namely azathioprine in patients with deleterious genotypes. However, before these markers can be brought into clinical practice, replication of associations is necessary. Once this is done, an interventional clinical study with genotype directing immunosuppressant regime choice would be necessary.

The unsupervised analysis produced a number of results and the potential significance of these has been discussed in Chapter 5: Discussion and Conclusions, the genes that should be considered in future MPA pharmacogenomics research are:

The gene *ASAP1* and *AGBL4* and potential association with the development of MPA induced anaemia.

The genes *SFTA2*, *ALPK3*, *NEU1* and *DAB2* and potential association with the development of MPA induced leucopenia.

The gene *CDC42BPA* and *NUP153* and potential association with the development of BPAR.

The clinical significance of the associations postulated remains tenuous, and replication in a second cohort is necessary.

6.3 Future research

The completion of a research study such as this one should not be considered to be an 'end point' but rather the start of the next phase into continuing development and understanding in this field.

Future research should aim to both substantiate associations that have been discovered in this study and explore beyond what is currently known. Ultimately, the aim is to provide better treatment choices and better outcomes for transplant patients.

in a well powered, preferably prospectively recruited cohort. Measurement of MPA levels in the early post-transplant period should be included in the study design as individual variability in these levels may correlate with glucuronidase and carrier genotypes. Ethical approval remains in place for the recruitment and study of a replication cohort, including the measurement of MPA blood levels within the recruiting centre.

A large multicentre study is necessary for power to identify new associations and will be key to moving forward in this field. This would provide consistency across a large cohort applying the same methodology, studying the same SNP's and the same clinical outcomes. This lack of

consistency and replication across pharmacogenetic studies in transplantation has been one of the significant barriers to implementation of such data into clinical practice (13). The study design should involve a two staged approach with the recruitment of a replication cohort for validation of results as outlined above. As the clinical data collected in this study is universally available in all RTR the only additional requirement in these subjects is the collection of a DNA sample which can be done as part of the routine tests in the work up for transplantation. This would minimise the impact of the study both to the recruited subject and to the transplant teams. The non-interventional nature of this study means that it could be done alongside other research in this population.

The future research should seek to look at other SNP's in the candidate genes not covered in this study and consider including coverage of intronic regions in genes such as IMPDH1 and 2 which are central to MPA mode of action. Studies should continue to look beyond known candidate genes, as has been done in this study, to allow the discovery of new genes that could have a fundamental role in metabolism of and response to MPA but are, as yet, unknown. It is vital that rare variants continue to be studied as it is possible that these rarer SNPs may have profound clinical impact. Research must not only aim to clarify what is known but also seek new associations and continue to grow the body of knowledge.

It is likely that the presence of several SNPs together may have an additive effect or indeed counteract each other if they have opposing actions. Future research should seek to explore these interactions.

As this study highlights the transplant population require treatment with a combination of complex drugs to prevent organ rejection with an extensive side effect profile and potential for drug-drug interaction. Future research in this field should consider looking at the effect of co-drugs used in acute transplantation to enhance clinical application.

With the aim for the future being one of individualising therapy in transplantation, research in the field of transplant pharmacogenomics should also seek to evaluate the utility and cost effectiveness (20) of incorporating genetic profiling into the clinical setting. This will be crucial for acceptance within clinical practice. Ultimately, replicated predictive pharmacogenetics markers must be tested within the context of an interventional clinical trial.

6.4 Summary

Research, such as that carried out in this study, continues to be crucial in understanding the role of genetic variations on drug metabolism and potential clinical implications.

Continuation of research in this field will be essential to improve transplant outcomes in the future. The immunosuppressant drugs now used in transplantation are effective in preventing acute rejection in the majority of individuals but adverse drug effects remain a significant problem accounting for preventable morbidity and cost (181). Incorporating pharmacogenomics into both drug development and therapeutic decision making will be key (25). The use of genetic profiling to aide clinical decision making and tailor treatment to the individual should be the aim for all those receiving an organ transplant in the future. This is particularly important as drug level monitoring does not always equate to side effects. Additionally once adverse effects or lack of efficacy become apparent, the individual has often suffered significant morbidity or transplant function has been irreversible compromised. As pharmacogenomics techniques continue to develop and the cost continues to fall the use of pharmacogenomics as an aid to individualised prescribing could revolutionise the future of organ transplantation.

7 APPENDICIES

7.1 APPENDIX 1: Ethical approval from South central- Southampton A REC

REC reference 10/H0502/81

Study to identify the association between polymorphisms in a candidate pharmacogenetic locus predicted to influence the metabolism of mycophenolic acid precursors (mofetil / sodium) and clinical outcomes in renal transplant recipients

1O/H0502/81

The Research Ethics Committee reviewed the above application at the meeting held on 09 November 2010. Thank you for attending to discuss the study.

Ethical opinion

1. The Committee asked Dr Borman about the recruitment of participants and from which locations this would be done. Dr Borman confirmed that all participants would come from the Portsmouth unit and therefore all medical files would already be available on site.

2. The Committee expressed its concern at the legibility of the PIS and noted the duplication of points 4 and 5 in the consent form. Dr Borman agreed for the need to be consistent and confirmed that the duplication was an error.

3. The Committee requested clarification of the recruitment process. Dr Borman explained the WRTU central patient database would be used and a list generated of regular patient attendees. Patients would be approached at standard clinical appointments and given at least an hour to decide if they wish to participate. Patients will be asked by their consultant or research nurse who are all aware of the study and willing to make an initial approach. They will then feedback to Dr Borman.

4. The Committee asked for more details on the transportation of blood samples to London. Dr Borman confirmed the unit has a material transfer agreement with Guys and St Thomas' NHS Foundation Trust and that transportation would be via freezer compartment, transported by Dr Borman herself via train or car. This is comparable to current standard sample transportation, usually by courier.

5. The Committee wondered about the arrangements for comprehension by non- English speakers and those unable to fully understand and consent. Dr Borman explained it was important for patients to understand there was no benefit to them directly. There is a very small population of non-English speakers in Portsmouth; Dr Borman confirmed they would seek translators in this case but obtain clarification from the translator that they themselves understood and that the participant adequately understood to provide consent.

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting_ documentation, subject to the conditions specified below. –

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

For NHS research sites only, management permission for research ("R&D approval'? should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where the only involvement of the NHS organisation is as a Participant Identification Centre (PIG), management permission for research is not required but the R&D office should be notified of the study and agree to the organisation's involvement. Guidance on procedures for PICs is available in IRAS. Further advice should be sought from the R&D office where necessary.

Sponsors are not required to notify the Committee of approvals from host organisations. Other conditions specified by the REC:

1. Amended PIS to be submitted to the REC, using consistent terminology for genetic or DNA sequencing *I* testing *I* analysis. Amended consent form with removal of duplicated text also to be submitted.

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

You should notify the REC in writing once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers.

Approved documents

The documents reviewed and approved at the meeting were:

Documents	Version	Date
REC application		20 September 2010
Protocol	1.0	22 September 2010
Participant Information Sheet	1.0	22 September 2010
Referees or other scientific critique report	Peter Friend	28 July 2010
Investigator CV	Natalie Borman	15 January 2010
Investigator CV	Dr Venkat- Raman	04 August 2010
Participant Consent Form	1.0	22 September 2010
Covering Letter		11 October 2010

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

Notifying substantial amendments

Adding new sites and investigators

Progress and safety reports

Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

With the Committee's best wishes for the success of this project

Yours sincerely

Dr Iain Macintosh

This Research Ethics Committee is an advisory committee to South Central Strategic Health Authority

The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England

7.2 APPENDIX 2: Ethical approval of substantial amendment

Ethical approval of substantial amendment from South central- Southampton A REC

NHS
Health Research Authority

NRES Committee South Central - Southampton A

Study title:	**A study to identify the association between polymorphisms in a candidate pharmacogenetic locus predicted to influence the metabolism of mycophenolic acid precursors (mofetil / sodium) and clinical outcomes in renal transplant recipients**
REC reference:	**10/H0502/81**
Amendment number:	**1**
Amendment date:	**02 February 2012**

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Participant Information Sheet: updated	4.0	02 February 2012
Participant Information Sheet	3.0	01 February 2011
Protocol	1.0 & 2.0	02 February 2012
Notice of Substantial Amendment (non-CTIMPs)	1	02 February 2012

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

10/H0502/81: Please quote this number on all correspondence

Yours sincerely

Dr Iain MacIntosh Chair

E-mail: scsha.berksrec@nhs.net

7.3 APPENDIX 3: Portsmouth Hospital NHS Trust R and D approval

REC reference 10/H0502/81

Portsmouth Hospitals NHS
NHS Trust

Dear Dr N Borman

Re: NHS Organisational Permission - Non CTiMP research

Study Title: Pharmacogenetic Determinants of Mycophenolic Acid Metabolism

Research Office No: PHT/2010/40

Sponsor: Portsmouth Hospitals NHS Trust

Chief Investigator: Dr Venkat-Raman, Wessex Renal Unit, Queen Alexandra Hospital, Portsmouth

I have received confirmation that the above study has been processed through the Portsmouth Research Office. The Office has reviewed your submission and confirms that it meets the requirements of the Trust and Research Governance Framework.

On behalf of Portsmouth Hospitals NHS Trust I therefore give NHS organisational permission for the above named project to commence.

Conditions of approval

1. That you accept the responsibility of Principal Investigator as defined in the current Research Governance Framework and as you have declared in your signed SSIF.
2. Submit any changes in accordance with IRAS guidance to the study documentation before implementation for confirmation of continued NHS Organisational Permission
3. Ensure all study personnel, not employed by Portsmouth Hospitals NHS Trust, hold either honorary research contracts/ letters of access with this Trust, before they have access to any facilities, patients, staff, their data, tissue or organs.
4. Submit copies of Serious Adverse Events involving subjects from this Trust to the R&D Department.
5. Complete R&D Research Governance interim and final reports as requested.
6. Maintain an Investigator Site File (ISF) within your department containing essential study documentation for the governance and management of your study. Your ISF must be available at all times for monitoring purposes and you must inform the Research Office of the ISF location at commencement of the project by e-mail to: research.office@porthosp.nhs.uk .
7. Enter recruitment data onto the Portsmouth Hospitals EDGE database in accordance with local research governance. If you do not have access to EDGE, please contact the Research Office; access and training will be arranged.
8. Ensure that research protocol exposures are accurately identified, i.e. all those that are part of normal clinical practice, as well as those additional to normal practice. All referrals to Radiology, Nuclear Medicine, Radiotherapy or Medical Physics must be clearly

a. Identified as within a research project using the GREEN IRMER stickers. For further detail s please refer to the IRMER procedure via the R&D Trust intranet site.

9. Agree to conduct this research project in accordance with the conditions of this approval.

Additional approvals

1. No samples to be transferred until a signed MTA has been received by the R&D Department.

Please ensure we are copied in to all correspondence and reporting requirements of the National Research Ethics Service (NRES). This includes annual reports submitted by Chief Investigator and the end of study declaration. We should also be informed of any publications or conference presentations resulting from this research.

Should you find yourself unsure of any of the above requirements please do not hesitate to contact the R esearch Office for support.

Yours sincerely

Jo Newbury

Portsmouth Hospitals NHS Trust

7.4 APPENDIX 4: Peer review of the study

Peter J. Friend, MD,
FRCS Professor of
Transplantation

Nuffield Department of Surgical
Sciences University of Oxford

Oxford Transplant
Centre Churchill
Hospital

Oxford OX3 7LJ

Tel: +44 (0)1865 223872 Fax: +44 (0)1865 223872

peter.friend@nds.ox.ac.uk

July 28th 2010

re: Pharmacogenetic determinants of mycophenolic acid metabolism

Dear Venkat

Thank you for asking me to review this proposed clinical study. I note that the primary aims of the study are:

- To identify polymorphisms predicted to influence the metabolism of mycophenolic acid precursors in renal transplant recipients.
- To study clinical outcomes after renal transplantation including rejection and drug toxicity
- To establish associations between the outcomes and specific polymorphisms in the candidate pharmacogenetic loci.

It is proposed to study 400 patients as part of a two cohort design – the first 200 being part of a retrospective study in patients who have already received transplants and the second cohort of 200 being studied prospectively. The interventions for patients will be restricted to blood samples for genetic analysis and mycophenolate levels, in addition to testing already being carried out for routine clinical purposes.

Justification for the study: The use of mycophenolate mofetil (MMF) and mycophenolate sodium (MPS) has been shown to reduce rejection rates in renal transplantation and these drugs have become part of the routine medication in kidney transplant recipients in the majority of transplant

units in the UK and elsewhere. However, the cost and toxicity of these drugs remain important issues. It is recognised that patients respond in an unpredictable way with respect to toxicity and protection from rejection (as well as infection) and that this is related to genetic factors. The published literature is not conclusive and there is no generally agreed method whereby the heterogeneity of response can be predicted in such a way (using either genetic analysis or drug level monitoring) as to tailor the treatment of the individual patient.

This study will generate data which may enable this class of drugs to be used in a much more individualized way – in order to maximise the benefit (freedom from rejection) and minimise the detriment (toxicity, infection). It is also proposed to incorporate a pharmaco-economic evaluation within the design of the study – this is a clear opportunity to measure the cost-benefit of a more intensive and individualized method of drug use.

Statistical design and endpoints: The study will recruit 400 patients of which 200 will be retrospective and 200 prospective. This is an elegant design feature insofar as it enables the total number of patients (400) to be used to answer the correlation of genetic data with outcome parameters whilst a smaller prospective cohort of 200 patients will be studied with the additional information of drug levels.

Although not part of this study, it is clear that a positive result from this trial might lead on to a future randomised trial in which 50% patients were managed with and 50% without the addition of genetic analysis and individualised drug treatment.

The statistical design seems to be based on realistic (perhaps even conservative) assumptions and this has clearly been established with the involvement of appropriate statistical expertise.

Ethical considerations: There are no concerns with respect to ethical considerations. The only additional intervention for patients who agree to take part in the trial is some additional blood sampling. There is no reason to believe that the analysis of the genes in question should have untoward implications.

Costs: The costing appears to be very reasonable for a trial of this nature. I note that the costs of genetic analysis will be provided independently. The study will be run on a day-to-day basis by a half-time research fellow. There is no allowance for any costs for statistical analysis (although this may well be within the remit of the MRC statistics unit).

Conclusions: I regard this to be a well-designed trial which is intended to address an important clinical question. The trial is capable (depending upon its outcome) of altering clinical practice in an important way, to the benefit of future patients.

Yours

Professor Peter J Friend

7.5 APPENDIX 5: Patient information sheet

Ref No: 10/H0502/81 Version 5.0 7/12/2012

Portsmouth Hospitals NHS Trust

Wessex Renal & Transplant Unit
Queen Alexandra Hospital
Portsmouth PO6 3LY
Dr Natalie Borman MBBCh MRCP
Renal Registrar
Renal Unit, G Level, QAH
Tel No: 02392286000

Natalie.borman@porthosp.nhs.uk

PATIENT INFORMATION SHEET

1. **Study title**

Pharmacogenetics of mycophenolate in patients receiving a kidney transplant

2. **Invitation paragraph**

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with friends, relatives and your GP if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

3. **What is the purpose of the study?**

It is currently accepted that a combination of mycophenolate (mofetil or sodium) plus either cyclosporine or tacrolimus, is the best treatment to prevent rejection of your kidney. However, in about 10% of patients the treatment may have to be changed because of side effects or because of problems with rejection. The reasons why some patients do not respond to therapy is not known. We would like to invite you to take part in a study which will help us to determine whether genetic markers, part of the normal variation found between people, predict response to mycophenolate treatment. It has recently become evident that differences in the way many drugs are handled by the body may be determined by the genetic make-up. Looking at genetic differences and the way people respond to a particular drug is called pharmacogenetics. We will study specific genetic markers which may predict how you respond to treatment, as well as whole genome scanning. In addition to this we will look at biochemical markers of response to see how this compares with either side effects attributable to the mycophenolate or episodes of rejection. These words and meanings are quite complex, so please don't hesitate to ask if you want a more detailed explanation!

4. **Why have I been chosen?**

You have been chosen for this study because you have either had a kidney transplant and been prescribed mycophenolate, or you are on the transplant waiting list and will be prescribed these drugs on receiving a renal transplant. Many other patients also fit these criteria and will also be asked to participate in the study

5. **Do I have to take part?**

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. This will not affect the standard of care you receive. If you decide not to take part your normal care will not be affected in any way.

6. **What will happen to me if I take part?**

If you decide to participate in the study we ask that the next time blood is drawn as part of your usual medical examination, an extra 5 ml (one teaspoonful) of blood be taken for our study. We may ask you to donate one further 5ml blood sample 1 year after your transplant. If you have already received a kidney transplant this is all that we need for our study and this is the only time you will need to give any blood for the study.

If you are on the transplant waiting list we will ask that the next time blood is drawn as part of your usual medical examination, an extra 5 ml of blood be taken for our study. This blood will only be tested once you receive a kidney transplant. Following your kidney transplant we may then be asked for 2 extra blood samples of 5ml of blood to be taken on 5 separate occasions during the first month after your transplant at a time when you are having routine blood tests.

DNA will be extracted from your blood and will be stored in Queen Alexandra hospital; it will then be taken to the Purine Research Laboratory at Guy's Hospital in London. The specimen will be labelled with a laboratory number so that you cannot be identified and access to your sample will be strictly controlled. We will be studying genetic markers which may predict how you respond to treatment. We will keep your sample for two years and at the end of this period your sample will be destroyed.

Five additional blood tests taken in those individuals who receive a new kidney transplant during the study will be sent to London to measure drug levels in the blood to provide additional information. This will not be required if you have already had a recent renal transplant at the time the study is commenced.

7. **What do I have to do?**

We ask that you give an extra 5 ml or around one teaspoonful of blood for the study the next time blood is taken from you as part of your normal medical examination. Followed by up to a further 5 blood samples taken in the first 12 month after transplant, at a time you are having routine blood test, if you have not yet had your kidney transplant.

8. **What are the possible benefits of taking part?**

You will not be told the results of any of these genetic tests. However in the future, the results of the study may lead to useful genetic tests to predict how people will respond to treatment. This could lead to improved treatment dosing.

9. **What happens when the research study stops?**

The study does not have a formal closing date, the aim is to continue until data is available on 450 patients to given ample information to prove any associations are not due to chance or coincidence.

10. **What if something goes wrong?**

As participation in the study is extremely unlikely to cause harm, any incidental harm that occurs during this research project, will not be covered by any special compensation arrangements. If you are harmed due to someone's negligence, then you may have grounds for a legal action. Regardless of this, if you wish to complain about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanisms will be available to you.

11. **Will my taking part in this study be kept confidential?**

All information which is collected about you during the course of the research will be kept strictly confidential. Any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised from it.

12. **What will happen to the results of the research study?**

The result of the study will be published in the scientific literature. You will not be identified in any publication.

13. **Who has reviewed the study?**

The study has been reviewed by external sources – Prof P.Friend, Consultant Transplant Surgeon at Oxford. The study is being conducted in collaboration with Guy's & St Thomas's Hospitals Trust, and approved by the Southampton Research Ethics committee and Guy's Research Ethics Committee.

The study will be supervised and monitored by Dr G.Venkat-Raman, Consultant Nephrologist, Queen Alexandra Hospital, who will be available for any queries or advice.

Please note that you may withdraw from the study at any time.

14. **Contact for Further Information**

Dr Natalie Borman
Wessex Renal Unit
Queen Alexandra Hospital
Southwick hill Road
Portsmouth
PO6 3LY

7.6 APPENDIX 6: Study Consent form

Ref No: 10/H0502/81 Version 5.0 7/12/2012

Portsmouth Hospitals NHS
NHS Trust

Wesses Renal Unit
Queen Alexander Hospital
Cosham
Dr Natalie Borman MBBCh MRCP
Renal Registrar
Renal Unit, G Level, QAH
Tel No: 02392286000

Natalie.borman@porthosp.nhs.uk

LREC Study Number:
Patient Identification Number for this trial:

CONSENT FORM

Title of Project: **Pharmacogenetics of mycophenolate**
(Renal: MMF/MPS)
Name of Researcher: Dr Natalie Borman (Renal Unit),

Please initial box

1. I confirm that I have read and understand the information sheet (Version 5 7/12/2012) for the above study and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that sections of any of my medical notes may be looked at by responsible individuals from the Renal Unit where it is relevant to my taking part in research. I give permission for these individuals to have access to my records. ☐
4. I understand that my blood sample will be stored and used for research into the way patients respond to mycophenolate treatment. ☐
5. I agree to take part in the above study. ☐

_____	_____	_____
Name of Patient	Date	Signature
_____	_____	_____
Name of Person taking consent (if different from researcher)	Date	Signature
_____	_____	_____
Researcher	Date	Signature

7.7 APPENDIX 7: Data collection sheet

Data Collection Sheet

Patient Name

Date of birth

Date of Tx

Type of TX

Previous Tx

Dialysis modality and duration pre transplant

HLA MM

CMV status

Donor age and sex

Cold IT

Warm IT

Induction therapy

CMV prophylaxis?

Initial immunosuppression:

Drug	Dose
MMF/ myfortic	
Tac	
Cya	
Pred	

Biopsies:

Treatment for rejection

Side effects and what happened to MMF at this time

MMF dose adjustment and why

Infection episodes

Date												
	Day 1	Day 7	Day 14	Day 21	Day 28	Month 2	Month 3	Month 4	Month 5	Month 6	Month 9	Month 12
Creat												
eGFR												
CNI Level												
CNI dose												
MMF /myf Dose												
Hb												
WCC												
CRP												
Alb												
Glucose												
BMI												
EPO												

Date	Leucopenia ? (wcc<3.5)	Any changes to MPA	Anaemia (Hb<10.5)	Any changes to MPA	GI side effects	Define		

7.8 APPENDIX 8: Experimental Protocol 1: DNA purification from whole blood using QIAamp DNA Blood Midi Kits.

Material required

- QIAamp DNA blood Midi Kit (available in 20 or 100 kits). This contains all necessary Buffers and Protease and QIAamp Midi Columns.
- EDTA 2ml Blood samples, can be used fresh or frozen (defrosted and equilibrated to room temperature prior to use)
- 96-100% Ethanol
- Phosphate-buffered saline (PBS) may be required for some samples

Equipment required

- 1 additional 15ml Centrifuge tube per sample
- Water bath heated to 70oC
- Timer
- Positive displacement Pipets
- Pipet Tips
- Mixing Vortex
- Centrifuge capable of attaining 4500 x g (5000rpm) with a swing out rota and buckets that can accommodate 15ml Centrifuge tubes.
- Small Eppendorf receiver tubes (3 per sample).

Methods

Reagent Preparation:

The reagents come with the QIAamp DNA kits but require preparation:

- QIAGEN Protease Stock solution S
  Midi 20 Kits- 4.4ml distilled water
  Midi 100 kits- 5.5ml distilled water
  Store at 2-8oC
- Buffer AW1: Add ethanol (96-100%) before using the kit for the first time.
  Midi 20 kits- 25ml of ethanol, final Volume AW1 44ml
  Midi 100 Kits – 125ml ethanol, final Volume AW1 220ml
  Store at room temperature (15-25oC), stable for 1 year.
- Buffer AW2: Add ethanol (96-100%) before using the kit for the first time.

Midi 20 kits- 40ml ethanol, final volume AW2 57ml

Midi 100 kits- 150ml ethanol, final volume AW2 216ml

Store at room temperature (15-25oC), stable for 1 year.

DNA purification

1. Pipet 200µl QIAGEN Protease into 15ml centrifuge tube
2. Add 2ml blood and mix briefly (if necessary bring volume of sample up to 2ml with PBS before adding to centrifuge tube)
3. Add 2.4ml Buffer AL, invert tube 15 times, then vigorous shaking for at least 1 minute
4. Incubated at 70oC for 10 minutes
5. Add 2ml ethanol (96-100%), invert tube 10 times, then vigorous shaking for 1 minute
6. Carefully transfer half of solution onto QIAamp Midi column placed in 15ml centrifuge tube (do not moisten rim), Close cap and centrifuge at 1850 x g (300rpm) for 3 minutes
7. Remove QIAamp Midi column, discard filtrated, return column to centrifuge tube, transfer remaining solution from step 5 onto QIAamp Midi column, close cap and centrifuge again at 1850 x g (3000rpm) for 3 minutes
8. Remove column, discard filtrate, wipe off any spillage from the threads of the centrifuge tube and return column to centrifuge tube.
9. Carefully, without moistening rim, add 2ml Buffer AW1 to the QIAamp Midi Column. Close cap and centrifuge at 4500 x g (5000rpm) for 1 minute
10. Carefully, without moistening rim, add 2ml Buffer AW2 to the QIAamp Midi Column. Close cap and centrifuge at 4500 x g (5000rpm) for 15 minute
11. If centrifuge force below 4000 x g, incubate for 10 minutes at 70oC to evaporate residual ethanol
12. Place QIAamp midi column into a clean 15ml centrifuge tube, discard any filtrate. Clean any spillage off column with a wet tissue first.
13. Pipet 300µl Buffer AE or distilled water directly into the membrane of the QIAamp Midi column.
14. Incubate at room temperature for 5 minutes
15. Centrifuge at 4500 x g (5000rpm) for 2 minutes
16. For high concentration reload the elute containing the DNSA onto the QIAamp Midi column
17. Incubate at room temperature for 5 minutes
18. Centrifuge at 4500 x g (5000rpm) for 2 minutes.

19. Pipet 65µ of DNA containing elute into small elute receiver tube, repeat twice giving 3 separate elute samples.

A maximum of 8 samples can be prepared at any one time.

It will take approximate 1 hour 30 minutes.

Storage

- 1 Eppendorf receiver tube stored at -20oC
- 2 Eppendorf receiver tubes stored between -60oC and -80oC

Hazards

- Gloves should be worn throughout to prevent contamination of samples
- Laboratory coat should be worn at all times

7.9 APPENDIX 9: Experimental protocol 2: DNA concentration measurement using QUBIT

Material required

QUBIT BR Buffer

QUBIT BR Dye (Dye is photosensitive so must be kept in the dark)

QUBIT Standards 0 and 100

DNA

Equipment required

QUBIT Machine

QUBIT test tubes

Mixing Vortex

Micro centrifuge

Pipettes

Pipette tips

Methods

Samples are run in duplicate and then an average of the two taken.

The two readings must be within 5ηg/mL of each other to be acceptable, if there are

not it must be repeated.

Preparation:

To run 12 samples in duplicate and two standards:

1. Add 199µL of QUBIT BR buffer per sample (account for 28 sample sot allow for pipetting error) = 5572µL of BR buffer
2. Add 1µL of QUBIT BR dye per sample = 28µL of dye
3. Vortex and place in dark whilst setting up testing tubes

4. Set out 26 QUBIT testing tubes and label (including standards 0 and 100)- care not to touch the sides of the tubes as this can effect reading
5. Vortex and spin all DNA samples well to ensure good mixing
6. Vortex and spin standards
7. Place 190µL in the two standards tubes and 199µL in all remaining tubes
8. Add 10µL of standards 0 and 100 to the two respective standards tubes and vortex
9. Add 1µL of DNA sample to the correct corresponding tubes and vortex NOTE: pipetting must be very accurate otherwise results will not be valid

Using QUBIT:

1. Switch on machine
2. Select 'DNA(BR)'
3. Select 'dsDNA(BR)'
4. Select 'YES' to run standards
5. Place 0 standard in first and press 'read' the place 100 standard in and press read- should show a linear graph with two points
6. Then put first sample in and press 'read'
7. When reading comes up press 'Calculate stock' to get in ng/mL
8. Then place next sample in and repeat
9. Then calculate average concentration of the sample

7.10 APPENDIX 10: Experimental protocol 3: DNA Purification/ Concentration

Material required

DNA sample

Sodium acetate 3M pH4.6

100% Ethanol

70% Ethanol

TE Buffer

Equipment used

Microcentrifuge capable of 12000rpm
Positive displacement pipets
Eppendorf tubes
Mixing Vortex

Method

For 100µL of DNA (quantaties can be altered depending on amount of DNA sample available)

1. To 100µL of DNA in a clean eppendorf tube add 10µL Sodiumacetate 3M 9pH4.6) and vortex

2. Add 220µL of 100% ethanol and vortex

3. Ensuring eppendorf lid is closed place sample in freezer at -20˚C for 30 minutes

4. Remove from freezer and gentle swirl sample- do not vortex

5. Centrifuge at 12000g for 10 minutes

6. Remove carefully taking care not to shake

7. There will be a small pellet suspended in the supernate

8. Remove the supernate and discard- taking care not to disturb pellet

9. Wash pellet with 500µL of 70% ethanol and centrifuge at 12000g for 2 minutes

10. Remove supernate and discard

11. Leave eppendorf with lid open to dry on the bench for 20 minutes

12. Resuspend pellet in required amount of TE buffer- care not to add too much as will reduce concentration further.

13. Recheck DNA concentration to ensure it has reached desired level.

Storage

Restore DNA sample at -80˚C pending use.

Hazards

- Gloves should be worn throughout to prevent contamination of samples
- Laboratory coat should be worn at all times

7.11 APPENDIX 11: Example DNA concentration calculation

Example DNA concentration calculation and dilution for EXome sequencing preparation

for DNA sample X Qubit ® concentrations

Sample Number	Concentration reading 1 (ηg/mL)	Concentration reading 2 (ηg/mL)
X	105	108

Average concentration = (Reading 1 + reading 2)/2

= (105 + 108)/2

= 106.5 ηg/mL

Amount of DNA per µL to provide concentration of 50 ηg/mL

= 50/average concentration

= 50/106.5

= 0.469 µL required

Required amount of DNA per 10 µL = 10 x amount required per µL

= 10 x 0.469

= 4.69 µL of DNA

Amount of buffer required per 10 µL = 10 – DNA amount

= 10 – 4.69

= 5.31 µL of buffer

Sample number	DNA required per 10 µL for a concentration of 50 ηg/mL	Buffer required per 10 µL for a concentration of 50 ηg/mL
X	4.69	5.31

7.12 APPENDIX 12: Experimental protocol 4: Real-time PCR using DNA for SNP detection

Material required

DNA samples

Master Mix

TE buffer

RNA free water

Probes for SNP to be tested

Equipment required

Real-time PCR machine

MxPro software

96 Well rPCR plates

96 Well rPCR lids

Centrifuge for 96 well plates

Mixing Vortex

Positive displacement pipets

Multichannel pipette

Pipette tips

Multiple well pipette

Methods

Machine set up:

1. Log onto computer system attached to rPCR machine (lamp needs 20 minutes to warm up so set up machine prior to plate set up)
2. Select MxPro programme
3. Select 'Allele discrimination SNP real time'
4. Highlight wells that will be used to plate including blank
5. Select 'unknown' in well type
6. Select dyes in use ('FAM', 'HEX' and 'ROX')

7. Select ‘ROX’ as reference dye.

8. Select ‘Thermal plate’

9. Select ‘Fast step 2’

10. Select ‘Save as’ and create a file in RT results folder

11. Once plate is ready (see plate preparation protocol below) place in machine close door and select ‘RUN’

Plate set up for rPCR

1. Required SNP probe ordered prior to use via rs number

2. Add 100µL of Probe to 100µL of TE buffer in a small sterile tube and vortex (Probe at 40 times concentration and needs dilution to 20 times concentration)- Any that is not used should be stored at -20˚C

3. For 104 samples (ie 96 well plate plus additional for error) add in a sterile tube:
    a. 520µL pcr Master mix
    b. 52µL Probe
    c. 364µL water (RNA free)

4. Vortex and spin

5. Using multiple pipette technique add 9µL of mix into each well in 96 well plate- always add blank first to avoid any chance of contamination

6. Ensure DNA samples are spun and mixed prior to use

7. Add 1µL DNA to each well as per template (use a pre-set out template so you know where each sample is)

8. Cover wells with lids taking care not to touch as any marks or dirt will prevent accurate measurement by the machine

9. Spin plate at 2500g for a few second

10. Load into rPCR machine
11. Select ‘RUN’

12. Cycle will take 1 hour 13 minutes

Reading results:

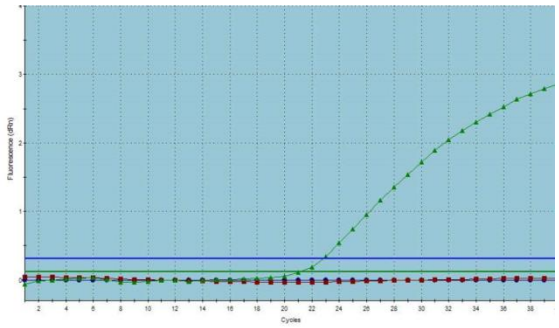
1. When cycles are complete select each well in turn to read results

2. Check blank first (which should show no rise in either dye) to ensure it has worked correctly

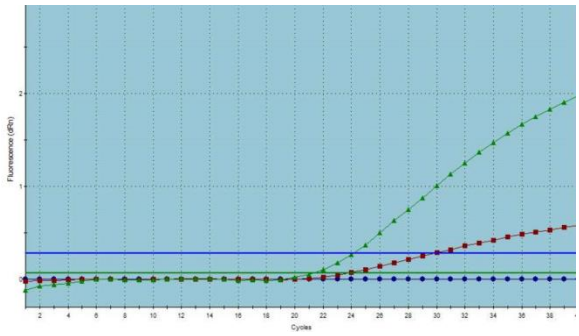
3. Results will show a graph for each sample

Example graphs:

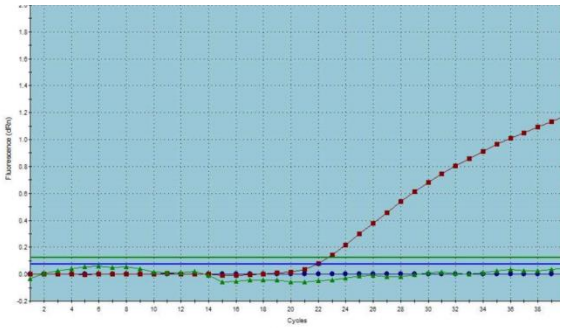
Wild type (homozygote for the common allele)



Hetrozygote



Homozygote (Homozygote for the uncommon allele)



Hazards

- Gloves should be worn throughout to prevent contamination of samples
- Laboratory coat should be worn at all times

7.13 APPENDIX 13: Protocol 5: Analysis of the data and PLINK instruction

This section gives details of the order in which the various stages of the analysis were applied to the data and the instructions for using PLINK to obtain these results.

1) Data on common variants in candidate gene was reviewed in genome studio, this data was analysed pre QC as outlined in this chapter
2) QC of the entire cohort of 2400 individuals
3) Application of Z call algorithm to entire cohort of 2400 individuals
4) Creation of a file of data for all individuals post QC and z call – PostQCall
5) PC were calculated for all individuals and for Caucasian only patients. Two files were produced PCall.txt and PCcauc.txt, these were saved as tab delimited files as this is required by PLINK
6) A PLINK file was created on the C drive of the computer, this is essential as PLINK need to be directed to the files.
7) Exm numbers for candidate genes were places in an excel file and saved as tab delimited file. This was repeated for extended candidates and saved as:
    i. MPAcad.txt
    ii. MPAexcad.txt
8) The ID of all patients in MPA cohort were also saved in a separate Tab delimited file MPApat.txt
9) PLINK opened and directed to the correct place in the computer by the commands
    i. C:\users\xxx>cd\ (then enter)
    ii. C:\>cd(space)plink (enter)
10) MPA patient data was then extracted from the entire cohort to allow analysis of this cohort only
    i. plink --file PostQCall ---keep MPApat.txt --make-bed --recode --out MPA
11) This will create a new BIN and FAM file just for MPA cohort only
12) Create a phenotype file with all outcome in it and save as phenotype.txt
13) Then run the following steps

Unsupervised analysis

Reduce to MAF frequency of 0.005 (ie 0.5%)

plink --bfile MPApat --maf 0.005 --make-bed --recode --out mpamaf

then run the fisher exact allele freq (step 1)

Plink --bfile mpamaf --pheno phenotype.txt –all-pheno --fisher --adjust

Then run the covar for principle components (step 2)

All PC

plink --bfile mpamaf --pheno phenotype.txt --all-pheno --logistic --covar pcall.txt --covar-number 1-4 --adjust --hide-covar --ci 0.95 --out mpamafpcall

Cauc only

plink --bfile mpamaf --pheno phenotype.txt --all-pheno --logistic --covar pccauc.txt --covar-number 1-2 –adjust --hide-covar --ci 0.95 --out mpamafpccauc (add—hide-covar to get rid of extra coulmns)

Then look for models of dom and rec

Dominant for Pc all

plink --bfile mpamaf --pheno phenotype.txt --all-pheno --logistic --dominant --covar pcall.txt --covar-number 1-4 --hide-covar --adjust --ci 0.95 --out mpamafdom

Recessive for PCall

plink --bfile cadsnpmaf --pheno phenotype.txt --all-pheno --logistic --recessive --covar pcall.txt --covar-number 1-4 --hide-covar --adjust --ci 0.95 --out mpamafrec

Repeat all for PCcauc

Results then pruned for SNP in LD using the following command

Sigsnp.txt is the txt file for candidate snps that had significant results

Check if any of these are in LD with each other

First produce a file of significant SNPs –sigsnp.txt

Plink --bfile mpab --extract sigsnp.txt --make-bed --recode --out sigsnp

Then check if in LD

Plink --bfile sigsnp --indep-pairwise 200 5 0.99

This will generate a plink.prune.in and a plink.prune.out file.

To exclude the pair of any in LD use plink.prune.in command

Plink --bfile sigsnp --extract plink.prune.in --make-bed --out sigsnpld

Pulling individual SNP results out

Place each SNP into a separate .txt file and save

Plink --bfile mpab --extract snpx.txt --make-bed --recode --out snpx

This will then produce a binary file of that SNP which can be used in SPSS for analysis.

14) Significant results (pre BF correction) then run in SPSS for logistic regression
15) Repeat process for candidate genes and extended candidate genes, to do this they are extracted from the MPApat file by the following command and then the steps outlined above are the same substituting MPA file for MPAcadsnp and MPAexcadsnp

plink --bfile mpab --extract MPAcad.txt --make-bed --recode --out MPAcadsnp

16) After logistic regression for candidate SNP Time to event analysis was also carried out

7.14 APPENDIX 14: Banff 97 diagnostic categories for renal allograft biopsies

1. Normal

2. Antibody-mediated changes (may coincide with categories 3, 4 and 5 and 6)

Due to documentation of circulating anti donor antibody, and C4d3 or allograft pathology

C4d deposition without morphologic evidence of active rejection

C4d+, presence of circulating anti donor antibodies, no signs of acute or chronic TCMR or ABMR

Cases with simultaneous borderline changes or ATN are considered as indeterminate

Acute antibody-mediated rejection4

C4d+, presence of circulating anti donor antibodies, morphologic evidence of acute tissue injury, such as (Type/Grade):

I. ATN-like minimal inflammation

II. Capillary and or glomerular inflammation (ptc/g >0) and/or thromboses

III. Arterial—v3

Chronic active antibody-mediated rejection4

C4d+, presence of circulating anti donor antibodies, morphologic evidence of chronic tissue injury, such as glomerular double

contours and/or peritubular capillary basement membrane multi layering and/or interstitial fibrosis/tubular atrophy and/or fibrous

intimal thickening in arteries

3. Borderline changes: 'Suspicious' for acute T-cell-mediated rejection (may coincide with categories 2 and 5 and 6)

This category is used when no intimal arteritis is present, but there are foci of tubulitis (t1, t2 or t3) with minor interstitial infiltration (i0

or i1) or interstitial infiltration (i2, i3) with mild (t1) tubulitis

4. T-cell-mediated rejection (TCMR, may coincide with categories 2 and 5 and 6)

Acute T-cell-mediated rejection (Type/Grade:)

IA. Cases with significant interstitial infiltration (>25% of parenchyma affected, i2 or i3) and foci of moderate tubulitis (t2)

IB. Cases with significant interstitial infiltration (>25% of parenchyma affected, i2 or i3) and foci of severe tubulitis (t3)

IIA. Cases with mild-to-moderate intimal arteritis (v1)

IIB. Cases with severe intimal arteritis comprising >25% of the luminal area (v2)

III. Cases with 'transmural' arteritis and/or arterial fibrinoid change and necrosis of medial smooth muscle cells with accompanying

lymphocytic inflammation (v3)

Chronic active T-cell-mediated rejection

'chronic allograft arteriopathy' (arterial intimal fibrosis with mononuclear cell infiltration in fibrosis, formation of neo-intima)

5. Interstitial fibrosis and tubular atrophy, no evidence of any specific aetiology

(may include nonspecific vascular and glomerular sclerosis, but severity graded by tubulointerstitial features)

Grade

I. Mild interstitial fibrosis and tubular atrophy (<25% of cortical area)

II. Moderate interstitial fibrosis and tubular atrophy (26–50% of cortical area)

III. Severe interstitial fibrosis and tubular atrophy/ loss (>50% of cortical area)

6. Other: Changes not considered to be due to rejection—acute and/or chronic (for diagnoses see Table 14 in (42); may include isolated

g, cg or cv lesions and coincide with categories 2, 3, 4 and 5)

7.15 APPENDIX 15: Results:

SNP analysed in candidate gene approach

Chrom	Gene	EXM	SNP	Position
7	ABCB1	exm2266441	rs3789243	87220886
7	ABCB1	exm631775	rs2032582	87160618
7	ABCB1	exm631843	rs2229109	87179809
7	ABCB1	exm631879	rs9282564	87229440
10	ABCC2	exm848442	rs56131651	101557063
10	ABCC2	exm848464	rs2273697	101563815
10	ABCC2	exm848522	rs17222617	101578952
10	ABCC2	exm848539	rs41318029	101590486
10	ABCC2	exm848562	rs45441199	101591737
10	ABCC2	exm848601	rs17222723	101595996
10	ABCC2	exm848653	rs8187710	101611294
4	ABCG2	exm412774	rs34783571	89013496
4	ABCG2	exm412870	rs2231137	89061114
10	CYP2C8	exm844097	rs10509681	96798749
10	CYP2C8	exm844133	rs1058930	96818119
10	CYP2C8	exm844152	rs41286886	96824658
10	CYP2C8	exm-rs1934951	rs1934951	96798548
7	CYP3A5	exm638602	rs28365083	99250236
7	CYP3A5	exm638644	rs6977165	99269397
12	SLCO1B1	exm2271695	rs2291075	21331625
12	SLCO1B1	exm988933	rs2306283	21329738
12	SLCO1B1	exm988936	rs11045819	21329813
12	SLCO1B1	exm988942	rs4149056	21331549
12	SLCO1B1	exm-rs4363657	rs4363657	21368722
12	SLCO1B1	exm989046	rs34671512	21391976
12	SLCO1B1	exm-rs4149032	rs4149032	21317791
2	UGT1A1	exm-rs887829	rs887829	234668570
2	UGT1A1	exm-rs4148325	rs4148325	234673309
2	UGT1A1	exm-rs6742078	rs6742078	234672639
2	UGT1A5	exm277212	rs3755321	234621825
2	UGT1A8	exm276956	rs17862841	234526784
2	UGT1A8	exm-rs11892031	rs11892031	234565283
2	UGT1A8	exm-rs2602381	rs2602381	234584324
2	UGT1A6	exm277163	rs6759892	234601669
2	UGT1A6	exm277187	rs2070959	234602191
2	UGT1A6	exm277188	rs1105879	234602202
2	UGT1A9	exm277410	rs45449995	234638580
2	UGT1A9	exm277431	rs34622615	234652308
4	UGT2B7	exm403192	rs61361928	69962375

NB: SNP's highlighted in the same colour are in LD with each other

Results Appendix: Candidate gene results Anaemia

Table 1: Results prior to QC of the data

Chr	gene	SNP	rs number	Model	P value	P value after Log reg	Exp β	95% CI exp (β)
2	UGT1A6	exm277163	rs6759892	Dom	0.048	0.052	1.676	0.99-2.82
12	SLCO1B1	exm-rs4363657	rs4363657	DOM	0.04	0.075	0.574	0.31-1.06
12	SLCO1B1	exm988933	rs2306283	Allele F	0.044	0.001	2.635	1.47-4.7
12	SLCO1B1	exm988933	rs2306283	Dom	0.003	0.003	2.268	1.3-3.9

Table2-3 Results of Fishers exact test for anaemia following QC of the data

Table 2: Allele frequency model

CHR	Gene	SNP	rs number	Model	P value	Bonf
10	ABCC2	exm848442	rs56131651	Allele F	0.01	0.33
12	SLCO1B1	exm2271695	rs2291075	Allele F	0.02	0.66
2	UGT1A1	exm-rs4148325	rs4148325	Allele F	0.045	1
2	UGT1A1	exm-rs6742078	rs6742078	Allele F	0.045	1

Table 3: Dominant model

CHR	Gene	SNP	rs number	Model	P value	Bonf
10	ABCC2	exm848442	rs56131651	Dom	0.04	1
12	SLCO1B1	exm2271695	rs2291075	Dom	0.003	0.1
12	SLCO1B1	exm988933	rs2306283	Dom	0.006	0.2
2	UGT1A6	exm277163	rs6759892	Dom	0.02	0.66

Table 4: Results following logistic regression for PC of entire cohort (N=278)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
12	SLCO1B1	exm2271695	rs2291075	DOM	0.53	0.26	0.3-0.9	0.014	0.46
12	SLCO1B1	exm988933	rs2306283	DOM	0.58	0.26	0.35-0.97	0.038	1

Table 5: Results following logistic regression for PC of Caucasian cohort (N=233)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
10	CYP2C8	exm844097	rs10509681	Alle F	2.04	0.31	1.1-3.8	0.02	0.66
10	CYP2C8	exm844097	rs10509681	DOM	2.09	0.34	1.1-4.1	0.03	0.99
12	SLCO1B1	exm988933	rs2306283	DOM	0.55	0.29	0.3-0.97	0.04	1
12	SLCO1B1	exm2271695	rs2291075	DOM	0.54	0.30	0.3-0.98	0.04	1
2	UGT1A6	exm277163	rs6759892	DOM	0.55	0.30	0.3-0.99	0.04	1

Following analysis for LD, this found that the following two SNP's were in 99% LD with each other. These SNP were not significant following logistic regression for PC.

exm-rs4148325
exm-rs6742078

The LD plot for rs2291075 and rs2306283 is shown in figure 1.

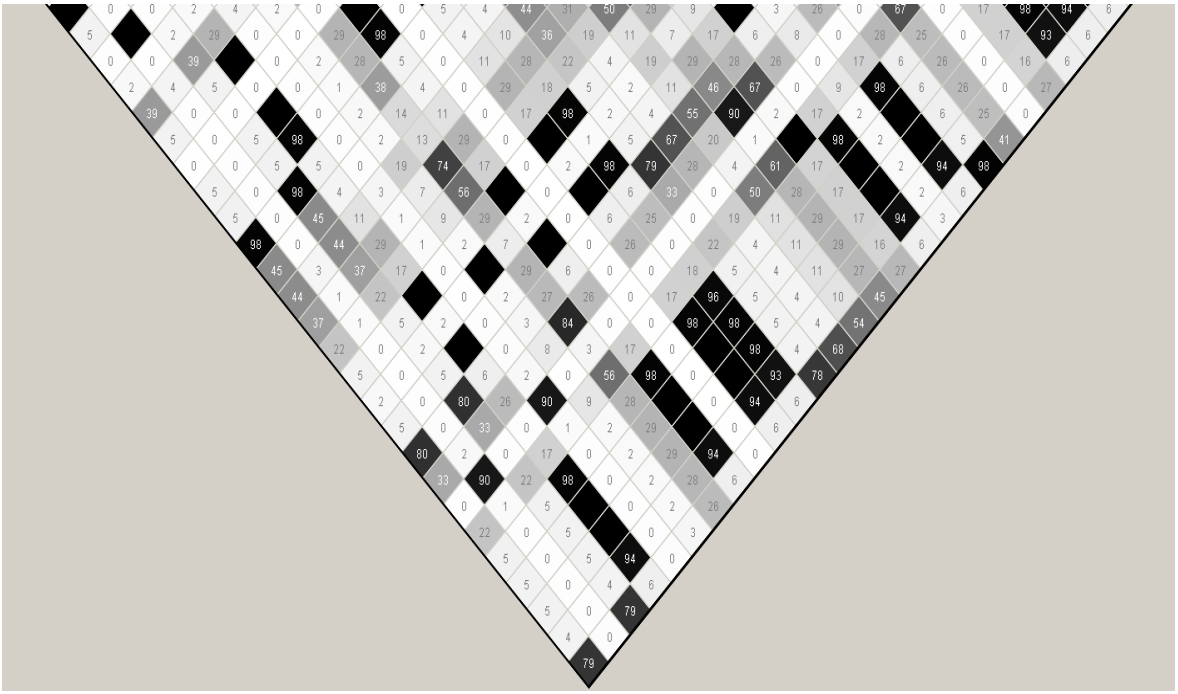


Figure 1: LD plot for SNP within 10kpb of rs2291075 and rs2306283. It shows that rs2291075 is in complete LD with rs17329885, and significant LD with rs6487213, rs2306283 and rs6487213 which are also within the gene.

Results were corrected using logistic regression in SPSS for gender, ethnicity, type of CNI at baseline, type of MPA at baseline, use of induction agent at baseline, use of valganciclovir prophylaxis, type of transplant donor, HLA-MM and eGFR at 3,6 and 12 months post-transplant.

Table 6-7: Results of logistic regression for these confounders.

Table 6: Logistic regression for entire cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
12	SLCO1B1	exm2271695	rs2291075	DOM	0.006	0.2	0.42	0.2-0.78
12	SLCO1B1	exm988933	rs2306283	DOM	0.031	1	0.51	0.28-0.94

Table7: Logistic regression for Caucasian cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
10	CYP2C8	exm844097	rs10509681	Alle F	0.1	1	1.94	0.9-4.3
10	CYP2C8	exm844097	rs10509681	DOM	0.09	1	1.93	0.9-4.1
12	SLCO1B1	exm988933	rs2306283	DOM	0.029	0.96	0.49	0.3-0.9
12	SLCO1B1	exm2271695	rs2291075	DOM	0.005	0.17	0.4	0.2-0.8
2	UGT1A6	exm277163	rs6759892	DOM	0.037	1	0.49	0.3-0.96

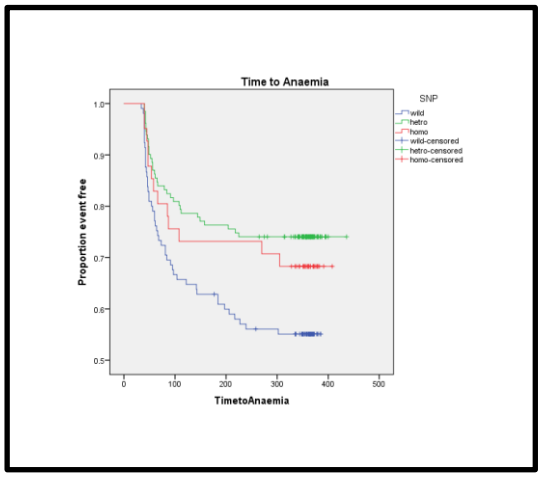
Time to event analysis

SNP's with significant results were then analysed for time to event analysis using Kaplan Meier (KM) survival curves with Log Rank test. Significant KM results were then analysed using cox regression analysis to adjust for the confounding variables gender, ethnicity, type of CNI at baseline, type of MPA at baseline, use of induction agent at baseline, use of valganciclovir prophylaxis, type of transplant donor, HLA-MM and eGFR at the time of event.

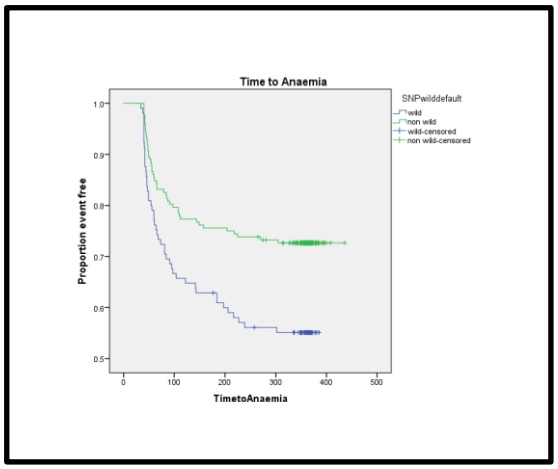
Table 8: Time to anaemia (days) entire cohort

Chr	Gene	EXM	rs number	Model	KM log rank	KM curve	Cox reg p Value	RR	95% CI
12	SLCO1B1	exm2271695	rs2291075	All F	0.008	1	0.004	0.49	0.30-0.8
12	SLCO1B1	exm2271695	rs2291075	DOM	0.002	2	0.009	0.55	0.35-0.86
12	SLCO1B1	exm988933	rs2306283	DOM	0.003	3	0.004	0.49	0.3-0.79
12	SLCO1B1	exm988933	rs2306283	DOM	0.002	4	0.014	0.59	0.38-0.9

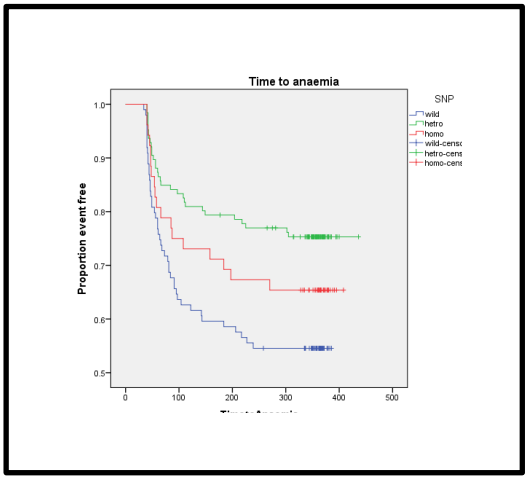
KM Curve 1 (p=0.008)



KM Curve 2 (p=0.002)



KM Curve 3 (p=0.003)



KM Curve (p=0.014)

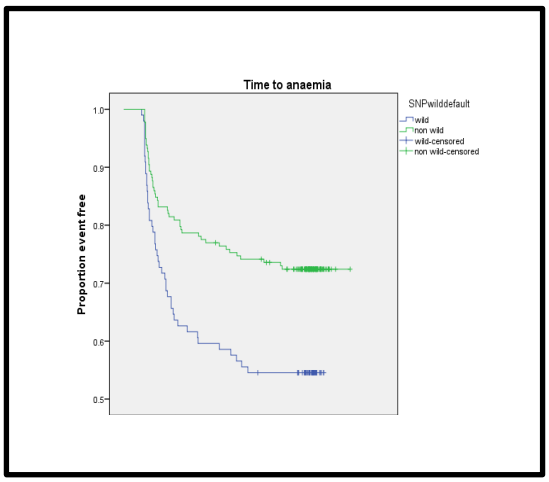
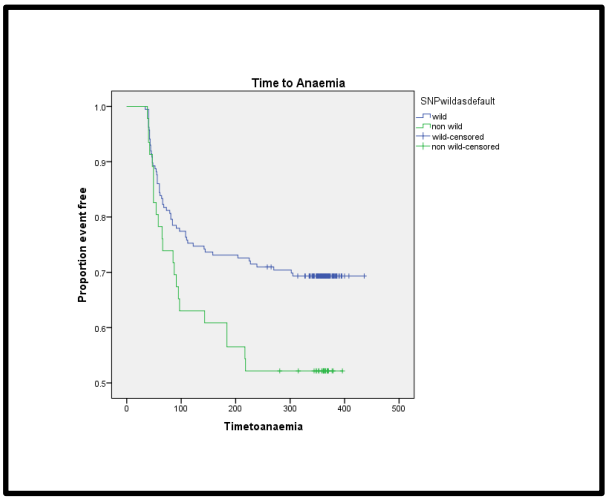


Table 8: Time to event Caucasian cohort

Chr	Gene	EXM	rs number	Model	KM log rank	KM curve	Cox reg p Value	RR	L95
10	CYP2C8	exm844097	rs10509681	AllelF	0.07	na	na	na	na
10	CYP2C8	exm844097	rs10509681	DOM	0.03	5	0.136	1.49	0.9-2.5
12	SLCO1B1	exm988933	rs2306283	DOM	0.3	na	na	na	na
12	SLCO1B1	exm2271695	rs2291075	DOM	0.2	na	na	na	na
2	UGT1A6	exm277163	rs6759892	DOM	0.65	na	na	na	na

KM Curve 5 (p=0.03)



Biopsy proven acute rejection

Table 9: Results of Fisher's exact test prior to QC of the data set

Chr	gene	SNP	rs number	Model	P value	P value after Log reg	Exp (β)	95% CI exp(β)
12	SLCO1B1	exm-rs4149032	rs4149032	Rec	0.005	<0.0001	4.37	1.3-9.9
12	SLCO1B1	exm988936	rs11045819	Allele F	0.004	0.013	6.7	1.5-29.9
12	SLCO1B1	exm988936	rs11045819	Dom	0.016	0.024	2.14	1.1-4.1
12	SLCO1B1	exm988936	rs11045819	Rec	0.044	0.002	5.73	1.3-25.2
12	SLCO1B1	exm988942	rs4149056	Allele F	0.043	0.10	0.49	0.21-1.2
12	SLCO1B1	exm989046	rs34671512	Allele F	0.038	0.05	2.4	1.0-5.7
12	SLCO1B1	exm989046	rs34671512	Dom	0.032	0.05	2.4	1.0-5.7

Table 10-12: Results of Fisher's exact test for BPAR (all) following QC of the data

Table 10: Allele frequency model

CHR	Gene	SNP	rs number	model	P value	Bonf
12	SLCO1B1	exm988936	rs11045819	Alle F	0.005	0.17
12	SLCO1B1	exm-rs4149032	rs4149032	Alle F	0.006	0.2
12	SLCO1B1	exm988942	rs4149056	Alle F	0.026	0.86
7	ABCB1	exm631743	rs55852620	Alle F	0.033	1
12	SLCO1B1	exm989046	rs34671512	Alle F	0.034	1
2	UGT1A5	exm277212	rs3755321	Alle F	0.05	1

Table 11: Dominant model

CHR	Gene	SNP	rs number	model	P value	Bonf
12	SLCO1B1	exm988936	rs11045819	Dom	0.023	0076
12	SLCO1B1	exm989046	rs34671512	Dom	0.029	0.96
7	ABCB1	exm631743	rs55852620	Dom	0.03	0.99
12	SLCO1B1	exm988942	rs4149056	Dom	0.048	1

Table 12: Recessive model

CHR	Gene	SNP	Rs number	Model	P value	Bonf
12	SLCO1B1	exm-rs4149032	Rs4149032	Rec	0.007	0.23

Table 13: Results following logistic regression for PC of entire cohort (N=278)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
10	ABCC2	exm848464	rs2273697	Alle f	0.48	0.33	0.25-0.9	0.026	0.86
12	SLCO1B1	exm988942	rs4149056	Alle f	0.4	0.41	0.2-0.9	0.027	0.89
10	ABCC2	exm848464	rs2273697	DOM	0.43	0.37	0.2-0.89	0.02	0.66
12	SLCO1B1	exm988942	rs4149056	DOM	0.4	0.44	0.17-0.9	0.036	1

Table 14: Results following logistic regression for PC of Caucasian cohort (N=233)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
12	SLCO1B1	exm989046	rs34671512	Alle f	3.78	0.44	1.6-8.95	0.003	0.1
12	SLCO1B1	exm-rs4149032	rs4149032	Alle f	2.06	0.25	1.3-3.4	0.004	0.13
12	SLCO1B1	exm988936	rs11045819	Alle f	2.05	0.29	1.1-3.6	0.014	0.46
2	UGT1A5	exm277212	rs3755321	Alle f	2.19	0.34	1.1-4.3	0.02	0.66
12	SLCO1B1	exm988933	rs2306283	Alle f	1.67	0.25	1.0-2.7	0.039	1
10	ABCC2	exm848464	rs2273697	Alle f	0.5	0.35	0.3-0.98	0.04	1
12	SLCO1B1	exm989046	rs34671512	DOM	3.78	0.44	1.6-8.95	0.003	0.1
2	UGT1A1	exm-rs887829	rs887829	DOM	2.44	0.36	1.2-4.96	0.01	0.3
2	UGT1A1	exm-rs6742078	rs6742078	DOM	2.44	0.36	1.2-4.96	0.01	0.3
2	UGT1A1	exm-rs4148325	rs4148325	DOM	2.44	0.36	1.2-4.96	0.01	0.3
12	SLCO1B1	exm988936	rs11045819	DOM	2.17	0.35	1.1-4.3	0.026	0.86
2	UGT1A5	exm277212	rs3755321	DOM	2.24	0.37	1.1-4.6	0.03	0.99
12	SLCO1B1	exm-rs4149032	rs4149032	DOM	2.2	0.37	1.1-4.5	0.03	0.99
10	ABCC2	exm848464	rs2273697	DOM	0.44	0.39	0.2-0.96	0.04	1
12	SLCO1B1	exm-rs4149032	rs4149032	REC	3.2	0.43	1.4-7.5	0.007	0.23

Significant SNPs were checked for Linkage disequilibrium and the following three SNP's are in 99% LD with each other.

SNP
exm-rs6742078
exm-rs4148325
exm-rs887829

The LD plot for rs4149032 (figure 2) shows it to by in complete LD with rs4149034 and significant LD with rs2199766.

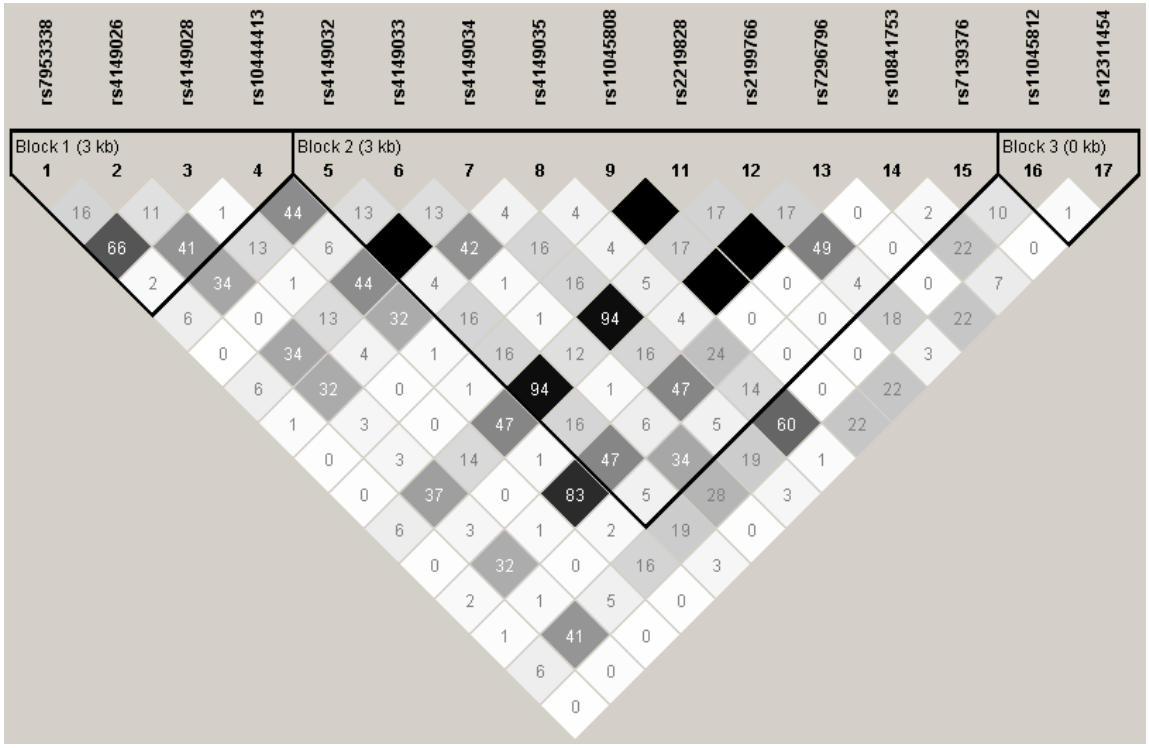


Figure 2: LD plot for SNP within 10kbp of rs4149032

The LD plot for rs11045819 (Figure 3) shows it to be in significant LD with a number of other SNPs summarised in the table below.

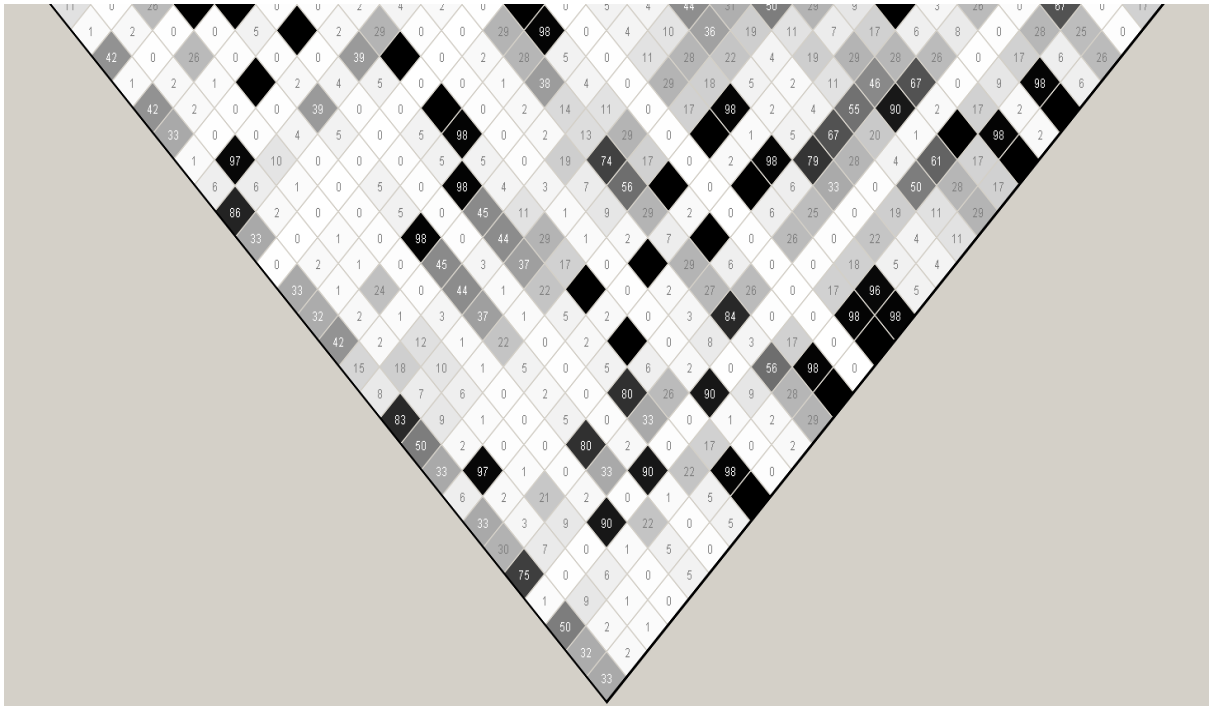


Figure 3: LD plot for SNP within 10kpb of rs11045819

Table 14: Summary of R^2 value for SNPs within 10kbp of rs6759892

SNP	R^2
rs11045818	1
rs17329885	1
rs11045820	0.98
rs11045821	1
rs12812279	1
rs11045823	0.98
rs11045524	1
rs4381410	1
rs2169969	0.94

Results were corrected using logistic regression in SPSS for gender, ethnicity, type of CNI at baseline, type of MPA at baseline, use of induction agent at baseline, type of transplant donor and HLA-MM.

Table 15-16: Results of logistic regression for these confounders.

Table 15: Logistic regression for entire cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
10	ABCC2	exm848464	rs2273697	Alle F	0.17	1	0.6	0.28-1.25
10	ABCC2	exm848464	rs2273697	Dom	0.10	1	0.55	0.27-1.3
12	SLCO1B1	exm988942	rs4149056	Alle F	0.07	1	0.44	0.18-1.07
12	SLCO1B1	exm988942	rs4149056	Dom	0.04	1	0.39	0.16-0.96

Table 16: Logistic regression for Caucasian cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
12	SLCO1B1	exm989046	rs34671512	Alle F	0.008	0.26	3.53	1.38-9.01
12	SLCO1B1	exm989046	rs34671512	Dom	0.008	0.26	3.53	1.38-9.01
12	SLCO1B1	exm-rs4149032	rs4149032	Alle F	0.003	0.1	4.88	1.74-13.7
12	SLCO1B1	exm-rs4149032	rs4149032	Dom	0.045	1	2.14	1.02-4.5
12	SLCO1B1	exm-rs4149032	rs4149032	Rec	0.005	0.17	3.77	1.5-9.46
2	UGT1A5	exm277212	rs3755321	Alle F	0.03	0.99	2.39	1.1-5.2
2	UGT1A5	exm277212	rs3755321	Dom	0.02	0.66	2.51	1.17-5.39
10	ABCC2	exm848464	rs2273697	Alle F	0.045	1	0.44	0.2-0.98
10	ABCC2	exm848464	rs2273697	Dom	0.045	1	0.44	0.2-0.98
12	SLCO1B1	exm988936	rs11045819	Dom	0.05	1	2.03	0.99-4.14
2	UGT1A1	exm-rs6742078	rs6742078	Dom	0.01	0.33	2.65	1.25-5.6
2	UGT1A1	exm-rs887829	rs887829	Dom	LD			
2	UGT1A1	exm-rs4148325	rs4148325	Dom	LD			

Time to event analysis

SNP's with significant results were then analysed for time to event analysis using Kaplan Meier (KM) survival curves with Log Rank test. Significant KM results were then analysed using cox regression analysis to adjust for the confounding variables gender, ethnicity, type of CNI at baseline, type of MPA at baseline, use of induction agent at baseline, type of transplant received and HLA-MM.

Table 17: Time to BPAR all (days) entire cohort

Chr	Gene	EXM	rs number	Model	KM log rank	KM curve	Cox reg p value	RR	95% CI
10	ABCC2	exm848464	rs2273697	Allele F	0.172	na	na	na	na
10	ABCC2	exm848464	rs2273697	Dom	0.1	na	na	na	na
12	SLCO1B1	exm988942	rs4149056	Allele F	0.07	na	na	na	na
12	SLCO1B1	exm988942	rs4149056	Dom	0.03	6	0.03	0.384	0.16-0.91

KM Curve 6 (p=0.03)

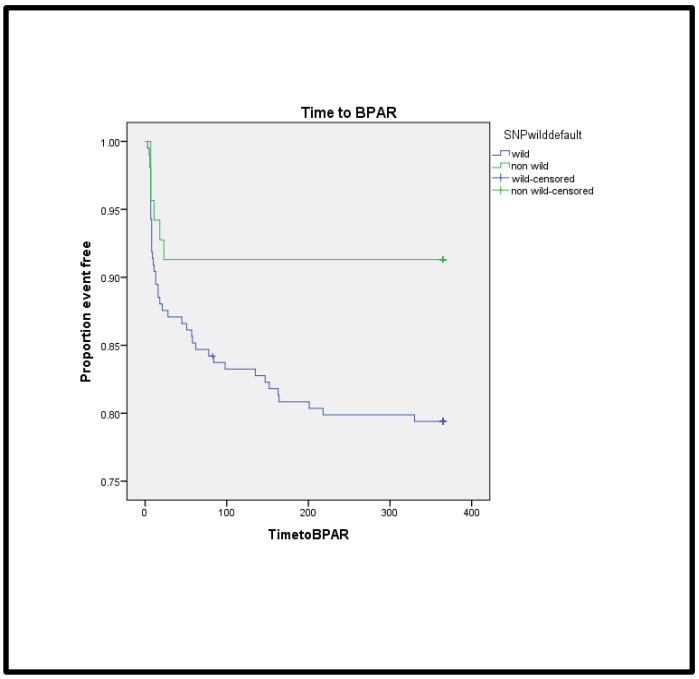
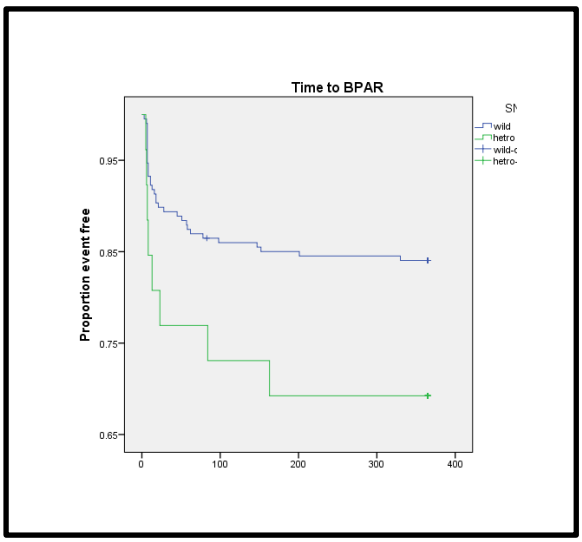


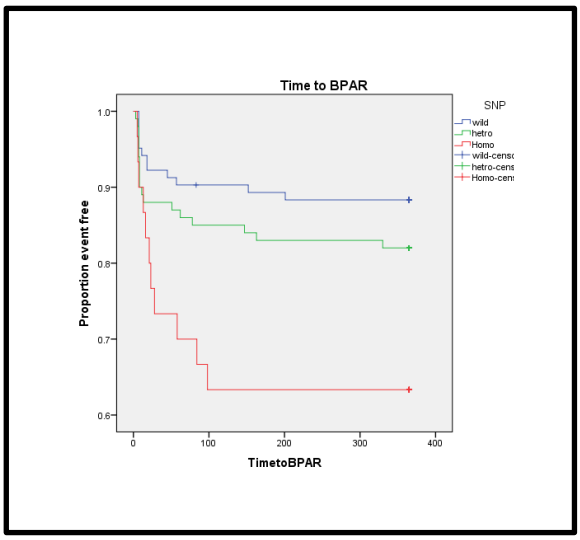
Table 17: Time to BPAR all (days) Caucasian cohort

Chr	Gene	EXM	rs number	Model	KM log rank	KM curve	Cox reg p value	RR	95% CI
12	SLCO1B1	exm989046	rs34671512	Dom	0.048	7	0.2	1.7	0.8-3.7
12	SLCO1B1	exm-rs4149032	rs4149032	All SNP	0.005	8	0.001	4.3	1.8-10.3
12	SLCO1B1	exm-rs4149032	rs4149032	Dom	0.03	9	0.02	2.3	1.1-4.6
12	SLCO1B1	exm-rs4149032	rs4149032	Rec	0.002	10	0.002	3.2	1.5-6.6
2	UGT1A5	exm277212	rs3755321	All SNP	0.16	na	na	Na	Na
2	UGT1A5	exm277212	rs3755321	Dom	0.09	na	na	Na	Na
10	ABCC2	exm848464	rs2273697	All SNP	0.6	na	na	Na	Na
10	ABCC2	exm848464	rs2273697	Dom	0.07	na	na	Na	Na
12	SLCO1B1	exm988936	rs11045819	All SNP	0.002	11	<0.0001	8.3	2.6-26.2
12	SLCO1B1	exm988936	rs11045819	Dom	0.007	12	0.01	2.3	1.2-4.2
12	SLCO1B1	exm988936	rs11045819	Rec	0.003	13	0.001	6.8	2.2-20.8
2	UGT1A1	exm-rs6742078	rs6742078	All SNP	0.04	14	0.1	1.7	0.9-3.4
2	UGT1A1	exm-rs887829	rs887829	Dom	0.04	14	0.1	1.7	0.9-3.4
2	UGT1A1	exm-rs4148325	rs4148325	Dom	0.04	14	0.1	1.7	0.9-3.4

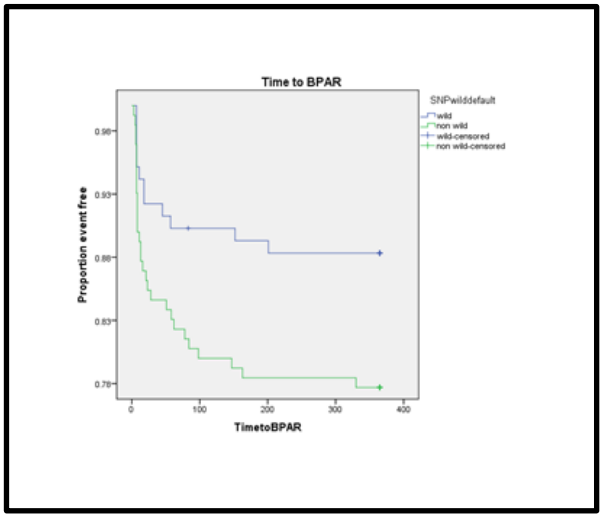
KM Curve 7 (p=0.048)



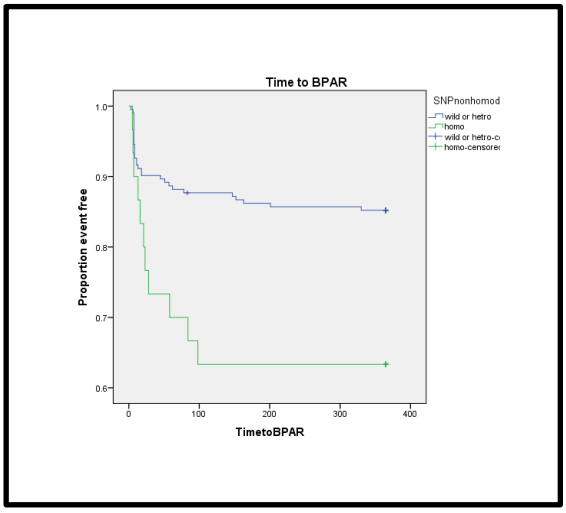
KM Curve 8 (p=0.005)



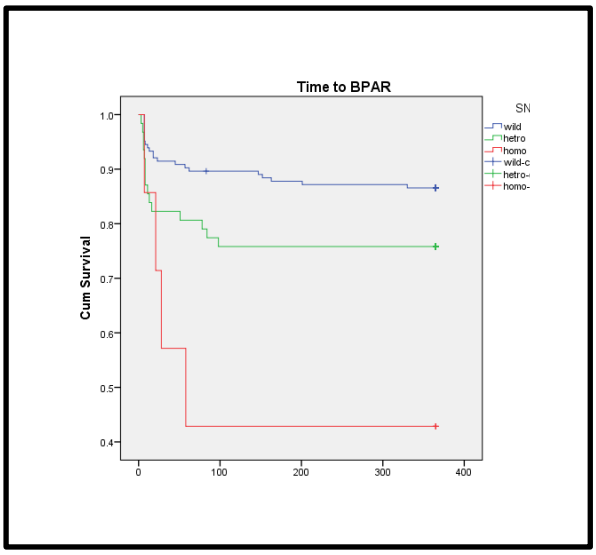
KM Curve 9 (p=0.03)



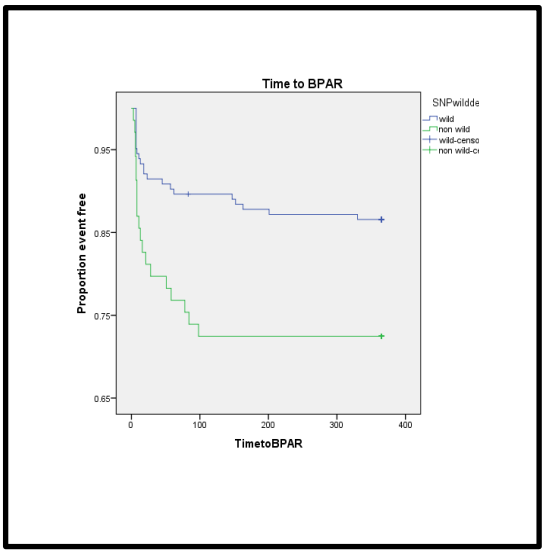
KM Curve 10 (p=0.002)



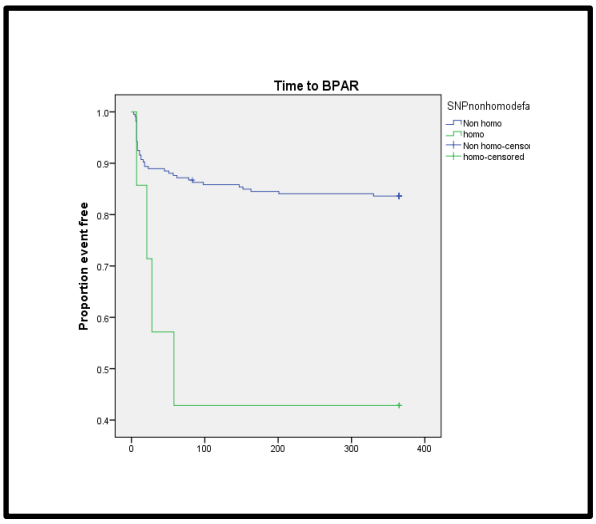
KM Curve 11 (p=0.002)



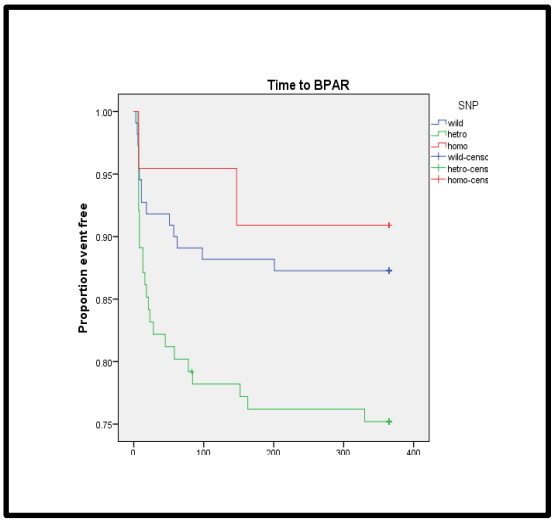
KM Curve 12 (p=0.007)



KM Curve 13 (p=0.003)



KM Curve 14 (p=0.04)



Biopsy Proven Acute Rejection Vascular or Cellular

Table 19: Results prior to QC of the data

Chr	gene	SNP	rs number	Model	P value	P value after Log reg	Exp (β)	95% CI for Exp (β)
2	UGT1A6	exm277163	rs6759892	Alle F	0.02	0.07	0.46	0.2-1.05
12	SLCO1B1	exm-rs4149032	rs4149032	Rec	0.015	0.003	4.1	1.7-10.4
12	SLCO1B1	exm988936	rs11045819	Alle F	0.03	0.09	2.0	0.9-4.7
12	SLCO1B1	exm988936	rs11045819	Dom	0.04	0.05	2.2	1.0-4.97
12	SLCO1B1	exm988942	rs4149056	Alle F	0.004	0.02	0.17	0.04-0.8
12	SLCO1B1	exm988942	rs4149056	Dom	0.009	0.015	0.16	0.04-0.7
12	SLCO1B1	exm-rs4363657	rs4363657	Alle F	0.049	0.02	0.2	0.06-0.8
12	SLCO1B1	exm-rs4363657	rs4363657	Dom	0.025	0.03	0.28	0.09-0.9
12	SLCO1B1	exm989046	rs34671512	Alle F	0.04	0.05	2.7	1.0-7.0
12	SLCO1B1	exm989046	rs34671512	Dom	0.04	0.05	2.7	1.0-7.0

Table 20-22: Results of Fisher's exacts tests for BPAR v or c following QC of the data

Table 20 : Allele frequency model

CHR	Gene	SNP	rs number	Model	P value	Bonf
12	SLCO1B1	exm988942	rs4149056	Alle f	0.006	0.2
12	SLCO1B1	exm-rs4149032	rs4149032	Alle f	0.018	0.6
12	SLCO1B1	exm988936	rs11045819	Alle f	0.025	0.8
12	SLCO1B1	exm989046	rs34671512	Alle f	0.042	1
2	UGT1A5	exm277212	rs3755321	Alle f	0.05	1

Table 21: Dominant model

CHR	Gene	SNP	rs number	Model	P value	Bonf
12	SLCO1B1	exm988942	rs4149056	Dom	0.008	0.26
12	SLCO1B1	exm988936	rs11045819	Dom	0.035	1
12	SLCO1B1	exm989046	rs34671512	Dom	0.036	1
12	SLCO1B1	exm-rs4363657	rs4363657	Dom	0.037	1

Table 22: Recessive model

CHR	Gene	SNP	rs number	Model	P value	Bonf
12	SLCO1B1	exm-rs4149032	rs4149032	Rec	0.013	0.43

Table 23: Results following logistic regression for PC of entire cohort (N=278)

CHR	gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
10	CYP2C8	exm844097	rs10509681	Alle F	2.2	0.37	1.1-4.5	0.03	0.99
10	ABCC2	exm848464	rs2273697	Alle F	0.45	0.4	0.2-1.0	0.056	1
10	ABCC2	exm848464	rs2273697	Dom	0.37	0.47	0.1-0.9	0.03	0.99
12	SLCO1B1	exm988936	rs11045819	Dom	2.29	0.39	1.1-4.9	0.03	0.99
10	CYP2C8	exm844097	rs10509681	Dom	2.34	0.4	1.1-5.2	0.04	1

Table 24: Results following logistic regression for PC of Caucasian cohort (N=233)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
12	SLCO1B1	exm989046	rs34671512	Alle f	4.4	0.5	1.7-11.7	0.003	0.1
12	SLCO1B1	exm-rs4149032	rs4149032	Alle f	2.1	0.3	1.2-3.8	0.01	0.33
12	SLCO1B1	exm988936	rs11045819	Alle f	2.3	0.4	1.2-4.6	0.016	0.5
2	UGT1A5	exm277212	rs3755321	Alle f	2.5	0.4	1.1-5.4	0.025	0.8
12	SLCO1B1	exm988942	rs4149056	Alle f	0.2	0.7	0.05-0.9	0.029	0.96
12	SLCO1B1	exm989046	rs34671512	Dom	4.4	0.5	1.7-11.7	0.003	0.1
12	SLCO1B1	exm988942	rs4149056	Dom	0.2	0.8	0.04-0.8	0.027	0.9
12	SLCO1B1	exm988936	rs11045819	Dom	2.5	0.4	1.1-5.7	0.029	0.96
12	SLCO1B1	exm-rs4363657	rs4363657	Dom	0.3	0.6	0.08-0.9	0.038	1
2	UGT1A5	exm277212	rs3755321	Dom	2.5	0.4	1.0-5.8	0.041	1
12	SLCO1B1	exm-rs4149032	rs4149032	Rec	3.8	0.5	1.4-9.8	0.007	0.23

Significant SNPs were checked for Linkage disequilibrium, none of the SNP's were found to be in LD with each other

Results were corrected using logistic regression in SPSS for HLAMM, Gender, Ethnicity, Type of CNI at baseline, Type of MPA at baseline, Use of induction agent at baseline, Type of transplant donor and HLAMM.

Table 25-26: Results of logistic regression for these confounders.

Table 25: Logistic regression for entire cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
10	CYP2C8	exm844097	rs10509681	All F	0.6	1	1.9	0.2-23
10	CYP2C8	exm844097	rs10509681	Dom	0.74	1	1.2	0.4-3.2
10	ABCC2	exm848464	rs2273697	All F	0.27	1	0.6	0.2-1.5
10	ABCC2	exm848464	rs2273697	Dom	0.25	1	0.6	0.3-1.4
12	SLCO1B1	exm988936	rs11045819	Dom	0.05	0.3	2.3	1.0-5.1

Table 26: Logistic regression for Caucasian cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
12	SLCO1B1	exm989046	rs34671512	All F	0.01	0.33	3.7	1.3-10.5
12	SLCO1B1	exm989046	rs34671512	Dom	0.01	0.33	3.7	1.3-10.5
12	SLCO1B1	exm-rs4149032	rs4149032	All F	0.006	0.2	5.4	1.6-18.1
12	SLCO1B1	exm-rs4149032	rs4149032	Rec	0.004	0.13	4.7	1.6-13.7
12	SLCO1B1	exm988936	rs11045819	All F	0.04	1	7.1	1.1-46.3
12	SLCO1B1	exm988936	rs11045819	Dom	0.06	1	2.3	1-5.5
2	UGT1A5	exm277212	rs3755321	All F	0.05	1	2.6	1-6.64
2	UGT1A5	exm277212	rs3755321	Dom	0.03	0.99	2.8	1.1-7.0
12	SLCO1B1	exm988942	rs4149056	All F	0.03	0.99	0.2	0.04-0.9
12	SLCO1B1	exm988942	rs4149056	Dom	0.02	0.66	0.2	0.04-0.8
12	SLCO1B1	exm-rs4363657	rs4363657	Dom	0.04	1	0.2	0.07-0.9

Time to event analysis

SNP's with significant results were then analysed for time to event analysis using KM survival curves with Log Rank test. Significant KM results were then analysed using cox regression analysis to adjust for the confounding variables Gender, Ethnicity, Type of CNI at baseline, Type of MPA at baseline, Use of induction agent at baseline, Type of transplant received and HLAMM.

Table 27: Time to BPAR v or C (days) entire cohort

Chr	Gene	SNP	rs number	Model	KM log rank	KM curve	Cox reg	RR	95% CI
12	SLCO1B1	exm988936	rs11045819	Dom	0.048	15	0.06	2.03	0.98-4.2

KM Curve 15 (p=0.048)

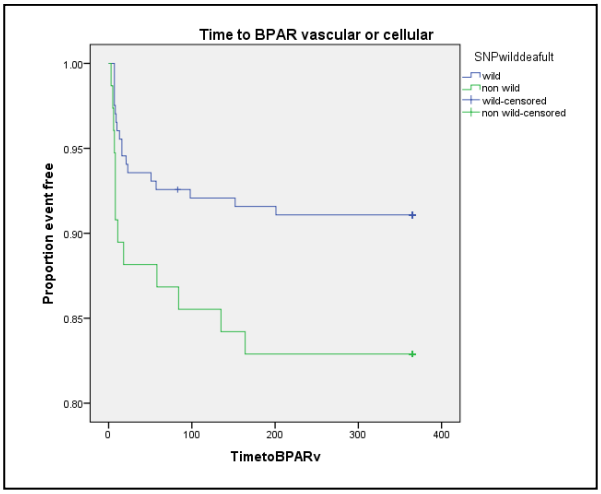
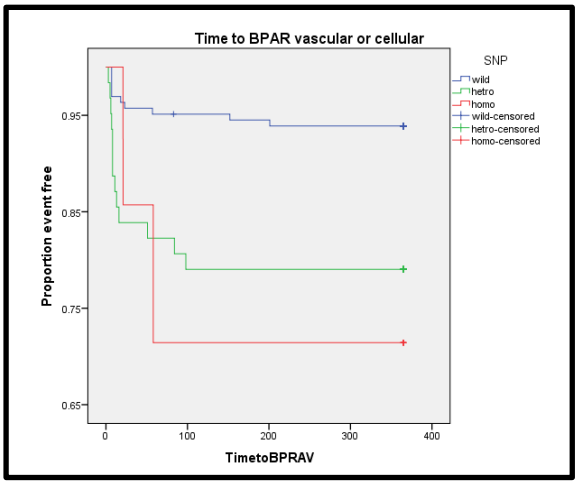


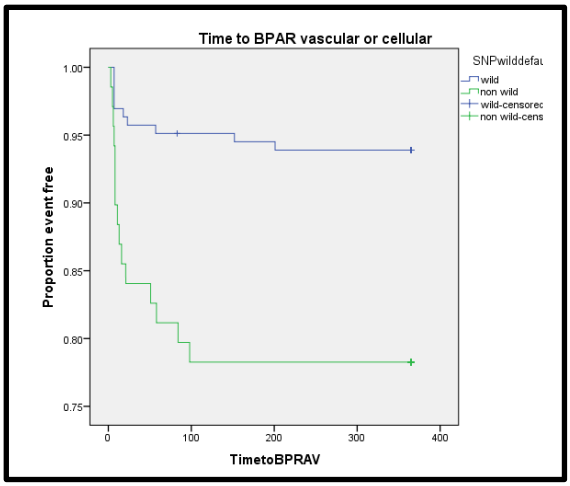
Table 28: Time to BPAR V or C (days) Caucasian cohort

Chr	Gene	EXM	rs number	Model	KM log rank	KM curve	Cox reg	RR	95% CI
12	SLCO1B1	exm989046	rs34671512	All F	0.001	16	0.004	3.4	1.5-7.8
12	SLCO1B1	exm989046	rs34671512	Dom	<0.0001	17	0.001	3.7	1.7-1.9
12	SLCO1B1	exm-rs4149032	rs4149032	All F	0.001	18	<0.0001	7.2	2.5-21.0
12	SLCO1B1	exm-rs4149032	rs4149032	Rec	<0.0001	19	<0.0001	5.5	2.3-13.2
12	SLCO1B1	exm988936	rs11045819	All F	0.001	20	0.004	3.4	1.5-7.8
12	SLCO1B1	exm988936	rs11045819	Dom	<0.0001	21	0.001	3.7	1.7-1.7
2	UGT1A5	exm277212	rs3755321	All F	0.02	22	0.02	2.7	1.2-6.3
2	UGT1A5	exm277212	rs3755321	Dom	0.02	23	0.01	2.9	1.3-6.5
12	SLCO1B1	exm988942	rs4149056	All F	0.37	na	na	na	Na
12	SLCO1B1	exm988942	rs4149056	Dom	0.2	na	na	na	Na
12	SLCO1B1	exm-rs4363657	rs4363657	Dom	0.28	na	na	na	Na

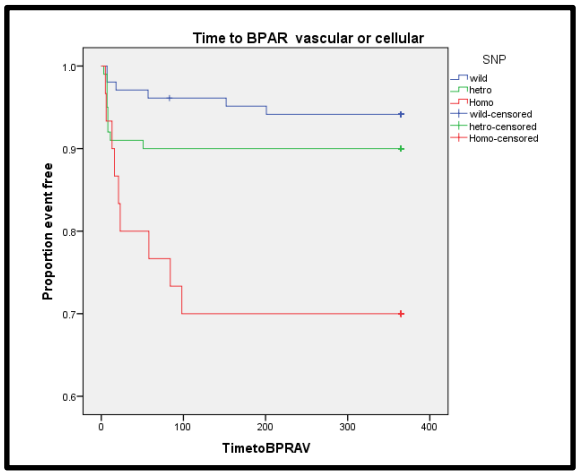
KM Curve 16 (p=0.004)



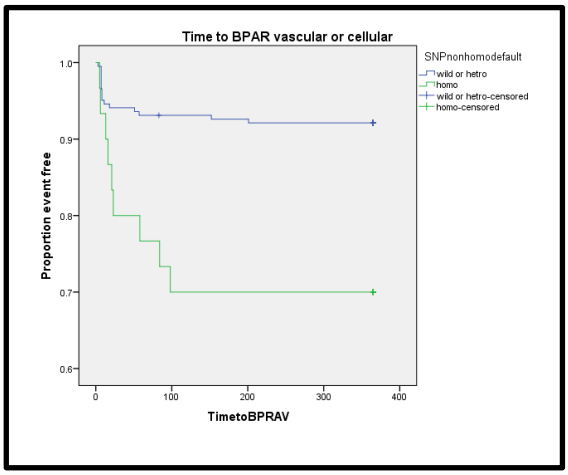
KM Curve 17 (p=0.001)



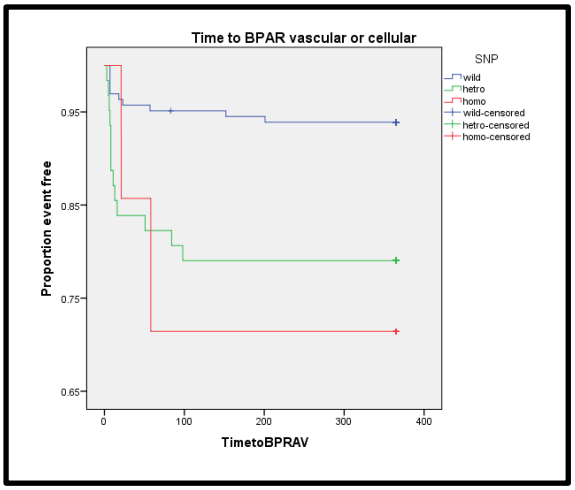
KM Curve 18 (p<0.0001)



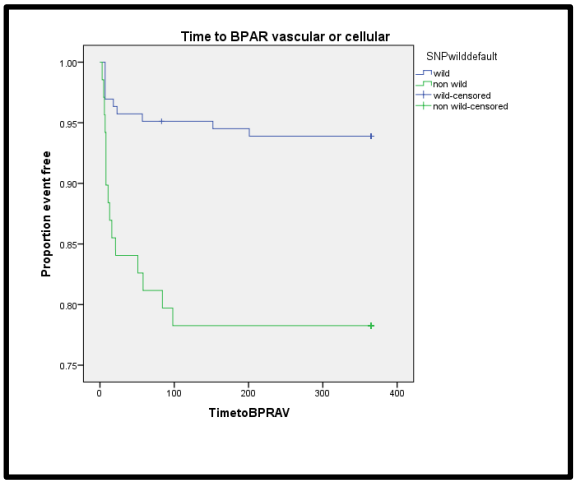
KM Curve 19 (p<0.0001)



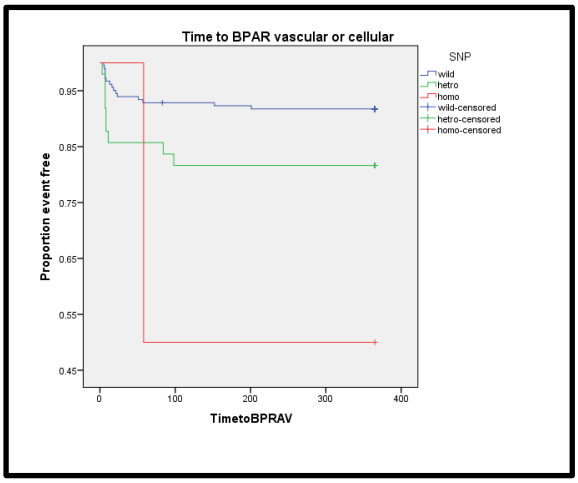
KM Curve 20 (P=0.004)



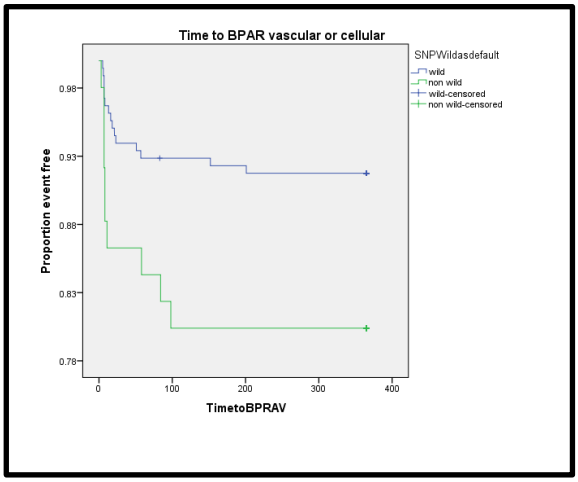
KM Curve 21 (p=0.001)



KM Curve 22 (P=0.02)



KM Curve 23 (p=0.01)



Gastrointestinal side effects

Table 29: Results prior to QC of the data

Chr	Gene	SNP	Rs number	Model	P value	P value after log reg	Exp β	95% CI exp β
2	UGT1A9	exm-rs2602381	rs2602381	Dom	0.03	0.025	0.5	0.3-0.9
2	UGT1A6	exm277163	rs6759892	Dom	0.04	0.03	1.8	1.1-3.0

Table 30: Results of fishers exact test for GISE following QC of the data

CHR	Gene	SNP	rs number	Model	P value	Bonf
2	UGT1A9	exm-rs2602381	rs2602381	Dom	0.006	0.2
2	UGT1A6	exm277163	rs6759892	Dom	0.02	0.66
2	UGT1A6	exm277188	rs1105879	Dom	0.05	1

Table 31: Results following logistic regression for PC of entire cohort (N=278)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
4	ABCG2	exm412774	rs34783571	All F	9.1	1.2	0.9-89.4	0.059	1
4	ABCG2	exm412774	rs34783571	Dom	9.1	1.2	0.9-89.4	0.059	1

Table 32: Results following logistic regression for PC of Caucasian cohort (N=233)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
2	UGT1A9	exm-rs2602381	rs2602381	Dom	2.2	0.4	1.1-4.3	0.03	0.99

Results were corrected using logistic regression in SPSS for Gender, Ethnicity, Type of CNI at baseline, Type of MPA at baseline, Use of induction agent at baseline, valganciclovir prophylaxis Type of transplant received and HLAMM. As before

Tables 33-34: Results of logistic regression for these confounders

Table 33: Logistic regression for entire cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
4	ABCG2	exm412774	rs34783571	All F	0.07	1	3.1	0.9-10.4
4	ABCG2	exm412774	rs34783571	Dom	0.07	1	3.1	0.9-10.4

Table 34: Logistic regression for Caucasian cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
2	UGT1A9	exm-rs2602381	rs2602381	Dom	0.035	1	2.2	1.1-4.5

No SNPs on the chip were found to be in LD with these SNPs. The LD plot for rs2602381 is shown in figure 5 It is not in complete LD with any other SNPs but has some degree of LD with 5 other SNPs within the same gene.

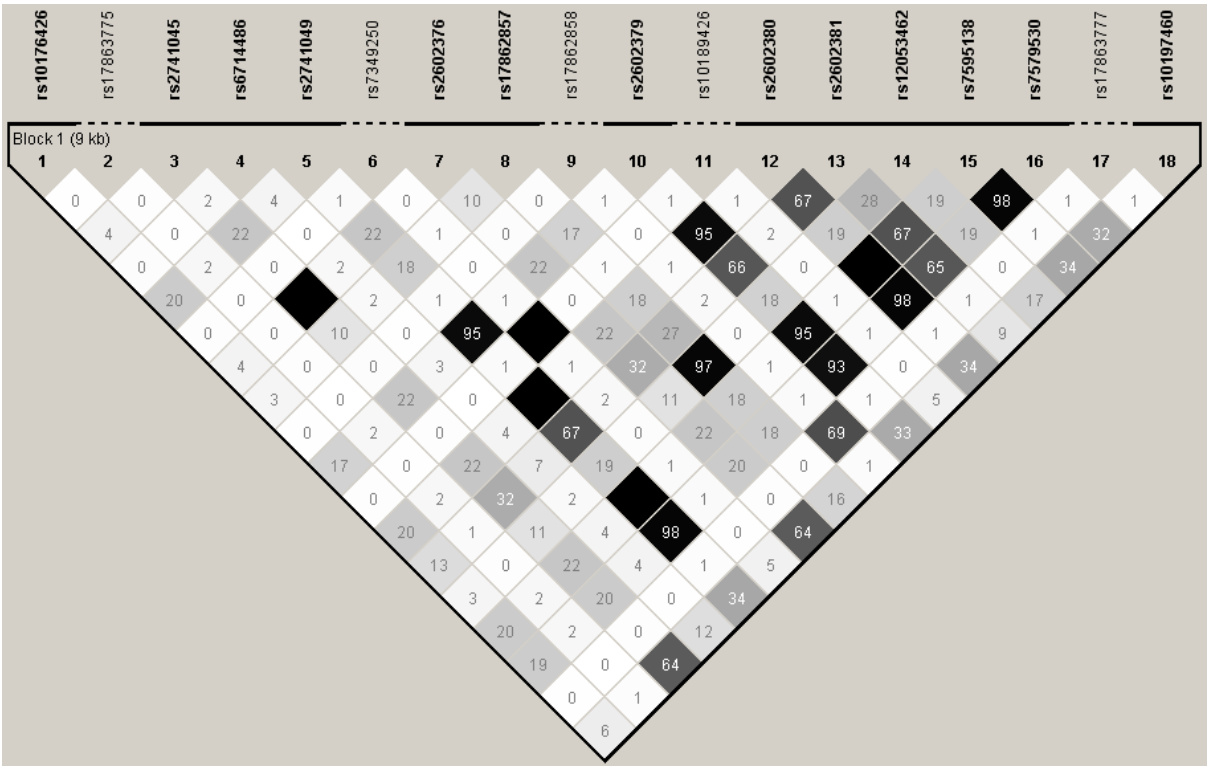


Figure 5: LD plot for SNP within 10kpb of rs2602381, no SNPs were found to be in complete LD with rs2602381

Upper Gastrointestinal (UGISE) side effects

Table 36: Results prior to QC of the data

Chr	gene	SNP	rs number	Model	P value	P value after Log reg	Exp β	95% CI exp (β)
2	UGT1A9	exm-rs2602381	rs2602381	Dom	0.056	0.03	0.4	0.2-0.9
2	UGT1A6	exm277163	rs6759892	Dom	0.04	0.04	2.3	1.0-5.1

Table 37-38: Results of Fisher's exact test for UGISE following QC of the data

Table 37: Allele frequency

CHR	Gene	SNP	rs number	Model	P value	Bonf
7	ABCB1	exm631843	rs2229109	All F	0.02	0.66

Table 38: Dominant model

CHR	Gene	SNP	rs number	Model	P value	Bonf
7	ABCB1	exm631843	rs2229109	Dom	0.22	1
2	UGT1A6	Exm277163	rs6759892	Dom	0.23	1

After logistic regression for PC of the entire cohort no results remained significant

Table 39: Results following logistic regression for PC of Caucasian cohort (N=233)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
7	ABCB1	exm631843	rs2229109	All F	3.79	0.65	1.05-13.6	0.04	1
7	ABCB1	exm631843	rs2229109	Dom	3.79	0.65	1.05-13.6	0.04	1
2	UGT1A6	exm277163	rs6759892	Dom	0.38	0.43	0.16-0.89	0.03	0.99

Results were corrected using logistic regression in SPSS for Gender, Ethnicity, Type of CNI at baseline, Type of MPA at baseline, use of induction agent at baseline, valganciclovir prophylaxis Type of transplant received and HLAMM.

Table 40: Results of logistic regression for the Caucasian only cohort, for these confounders

Table 40: Logistic regression for Caucasian cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
7	ABCB1	exm631843	rs2229109	All F	0.02	0.66	4.65	1.26-17.17
7	ABCB1	exm631843	rs2229109	Dom	0.02	0.66	4.65	1.26-17.17
2	UGT1A6	exm277163	rs6759892	Dom	0.018	0.6	0.33	0.13-0.83

Lower Gastrointestinal Side effects

No results were significant for LGISE prior to QC of the data

Table 41-42 Results of fishers exact test for LGISE following QC of the data.

Table 41: Allele frequency model

CHR	Gene	SNP	rs number	Model	P value	Bonf
2	UGT1A9	exm-rs2602381	rs2602381	All F	0.03	0.99
7	ABCB1	exm2266441	Rs3789243	All F	0.03	0.99

Table 42: Dominant model

CHR	Gene	SNP	rs number	Model	P value	Bonf
2	UGT1A9	exm-rs2602381	rs2602381	Dom	0.01	0.33

Table 43: Results following logistic regression for PC of entire cohort (N=278)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
7	ABCB1	exm631879	rs9282564	All F	0.41	0.4	0.2-0.9	0.03	0.99
12	SLCO1B1	exm989046	rs34671512	All F	0.27	0.65	0.08-0.98	0.046	1
7	ABCB1	exm631879	rs9282564	Dom	0.41	0.41	0.2-0.92	0.03	0.99
12	SLCO1B1	exm989046	rs34671512	Dom	0.27	0.65	0.08-0.98	0.046	1
12	SLCO1B1	exm988942	rs4149056	Dom	0.41	0.39	0.2-0.9	0.02	0.66
12	SLCO1B1	exm988933	rs2306283	Dom	0.54	0.29	0.3-0.96	0.037	1

Table 44: Results following logistic regression for PC of Caucasian cohort (N=233)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
2	UGT1A9	exm-rs2602381	rs2602381	Dom	2.4	0.4	1.1-5.19	0.03	0.99

No significant SNPs were found to be in LD.

Results were corrected using logistic regression in SPSS for Gender, Ethnicity, Type of CNI at baseline, Type of MPA at baseline, Use of induction agent at baseline, Type of transplant received and HLAMM.

Table 45-46: Results of logistic regression for these confounders. No results remained statistically significant

Table 45: Logistic regression of entire cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
7	ABCB1	exm631879	rs9282564	All F	0.57	1	0.81	0.38-1.7
7	ABCB1	exm631879	rs9282564	Dom	0.48	1	0.77	0.36-1.61
12	SLCO1B1	exm989046	rs34671512	All F	0.76	1	0.86	0.32-2.28
12	SLCO1B1	exm989046	rs34671512	Dom	0.76	1	0.86	0.32-2.28
12	SLCO1B1	exm988942	rs4149056	Dom	0.26	1	0.65	0.31-1.36
12	SLCO1B1	exm988933	rs2306283	Dom	0.16	0.96	0.65	0.35-1.18

Table 46: Logistic regression of Caucasian cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
2	UGT1A9	exm-rs2602381	rs2602381	Dom	0.062	1	2.15	0.96-4.8

MPA cessation

No results were significant for cessation of MPA prior to QC of the data

Table 47: Results of Fisher's exact results for MPA cessation following QC of the data

Table 47: Allele frequency model

Chr	Gene	SNP	rs number	Model	P value	Bonf
2	UGT1A6	exm277188	rs1105879	All F	0.045	1

Table 48: Results following logistic regression for PC on entire cohort (N=278)

Chr	gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
10	CYP2C8	exm844097	rs10509681	Alle F	9.6	1.1	1.3-72.7	0.03	0.99
12	SLCO1B1	exm988933	rs2306283	Alle F	2.4	0.4	1.1-5.2	0.03	0.99
4	ABCG2	exm412774	rs34783571	Alle F	0.5	0.3	0.3-0.95	0.035	1
12	SLCO1B1	exm988933	rs2306283	Dom	0.4	0.42	0.2-0.82	0.015	0.5
4	ABCG2	exm412774	rs34783571	Dom	9.6	1.03	1.3-72.7	0.029	0.96
10	CYP2C8	exm844097	rs10509681	Rec	19.7	1.26	1.7-233.5	0.02	0.66

After logistic regression for PC of Caucasian-only cohort no results were significant for MPA cessation.

Significant SNPs were checked for Linkage disequilibrium and no SNP's were in LD with each other.

Results were corrected using logistic regression in SPSS for Gender, Ethnicity, Type of CNI at baseline, Type of MPA at baseline, Use of induction agent at baseline, Valganciclovir prophylaxis, Type of transplant received and HLAMM.

Table 49: Logistic regression for entire cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
10	CYP2C8	exm844097	rs10509681	All F	0.27	1	4.3	0.3-58.6
10	CYP2C8	exm844097	rs10509681	Rec	0.3	1	4.0	0.3-53.5
12	SLCO1B1	exm988933	rs2306283	All F	0.34	1	0.5	0.14-2
12	SLCO1B1	exm988933	rs2306283	Dom	0.26	1	0.6	0.3-1.4
4	ABCG2	exm412774	rs34783571	All F	0.29	1	3.6	0.3-39.3
4	ABCG2	exm412774	rs34783571	Dom	0.29	1	3.6	0.3-39.3

MPA dose reduction

Table 50: Results pre QC of the data

Chr	gene	SNP	rs number	Model	P value	P value after Log reg	Exp β	95% CI exp (β)
2	UGT1A9	exm277410, exm277431	rs45449995 rs34622615	All F	0.039	0.024	0.25	0.08-0.83
2	UGT1A9	exm277410, exm277431	rs45449995 rs34622615	Dom	0.036	0.024	0.25	0.08-0.83
2	UGT1A6	exm277187	rs2070959	All F	0.003	0.28	0.75	0.47-1.27
2	UGT1A6	exm277187	rs2070959	Dom	0.004	0.21	0.73	0.45-1.2
2	UGT1A6	exm277187	rs2070959	Rec	<0.0001	0.47	0.74	0.33-1.7
16	CES1	exm1241669	Rs62028647	Dom	0.05	0.02	0.54	0.32-0.91

Table 51-52: Results of Fisher's exact test for MPA dose reduction following QC of the data

Table 51: Allele frequency model

CHR	Gene	SNP	rs number	Model	P value	Bonf
2	UGT1A6	exm277163	rs6759892	All F	0.036	1

Table 52: Dominant model

CHR	Gene	SNP	rs number	Model	P value	Bonf
2	UGT1A6	exm277163	rs679892	Dom	0.026	0.86
2	UGT1A6	exm277188	Rs1105879	Dom	0.05	1

Table 53: Results following logistic regression for PC of entire cohort (N=278)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
12	SLCO1B1	exm-rs4363657	rs4363657	Alle F	2.0	0.3	1.2-3.2	0.008	0.26
12	SLCO1B1	exm988942	rs4149056	Alle F	1.7	0.3	1.0-2.8	0.04	1
2	UGT1A9	exm277410	rs45449995	Alle F	0.3	0.1	0.1-1	0.05	1
2	UGT1A9	exm277431	rs34622615	Alle F	0.3	0.1	0.1-1	0.05	1
12	SLCO1B1	exm-rs4363657	rs4363657	Dom	2.0	0.3	1.1-3.5	0.017	0.56
2	UGT1A9	exm277410	rs45449995	Dom	0.3	0.6	0.1-1	0.05	1
2	UGT1A9	exm277431	rs34622615	Dom	0.3	0.6	0.1-1	0.05	1

Table 54: Results following logistic regression for PC of Caucasian population (N=233)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
2	UGT1A6	exm277163	rs6759892	Alle F	0.7	0.2	1-1.9	0.056	1
2	UGT1A6	exm277187	rs2070959	Dom	0.5	0.3	0.3-0.9	0.02	0.66
2	UGT1A6	exm277188	rs1105879	Dom	0.6	0.3	0.3-0.96	0.037	1
2	UGT1A6	exm277163	rs6759892	Dom	0.5	0.3	0.3-0.97	0.038	1

Significant SNPs were checked for Linkage disequilibrium and the following SNP's were found to be in 99.9% LD with each other.

exm277410
exm277431

Results were corrected using logistic regression in SPSS for Gender, Ethnicity, Type of CNI at baseline, Type of MPA at baseline, Use of induction agent at baseline, valganciclovir prophylaxis Type of transplant received and HLAMM.

Table 55-56 Results of logistic regression for these confounders.

Table 55: Logistic regression for entire cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
12	SLCO1B1	exm-rs4363657	rs4363657	Alle F	0.04	1	1.9	1-3.5
12	SLCO1B1	exm-rs4363657	rs4363657	Dom	0.058	1	1.7	0.98-3.1
12	SLCO1B1	exm988942	rs4149056	Alle F	0.2	1	1.5	0.8-2.8
2	UGT1A9	exm277410	rs45449995	Alle F	0.04	1	0.3	0.08-0.96
2	UGT1A9	exm277410	rs45449995	Dom	0.04	1	0.3	0.08-0.96
2	UGT1A9	exm277431	rs34622615	Alle F	0.04	1	0.3	0.08-0.96
2	UGT1A9	exm277431	rs34622615	Dom	0.04	1	0.3	0.08-0.96

Table 56: Logistic regression for Caucasian cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
2	UGT1A6	exm277163	rs6759892	Alle F	0.08	1	0.56	0.3-1.0
2	UGT1A6	exm277163	rs6759892	Dom	0.056	1	0.6	0.3-1.0
2	UGT1A6	exm277187	rs2070959	Dom	0.03	0.99	0.5	0.3-0.9
2	UGT1A6	exm277188	rs1105879	Dom	0.03	0.99	0.5	0.3-0.9

Infection

No results were significant for infection prior to QC of the data

Table 57-58: Results of fishers exact test for infection following QC of the data

Table 57: Allele frequency model

CHR	Gene	SNP	rs number	Model	P value	Bonf
10	ABCC2	exm848475	rs17222561	Alle F	0.046	1

Table 58: Dominant model

CHR	Gene	SNP	rs number	Model	P value	Bonf
10	ABCC2	exm848475	rs17222561	Dom	0.046	1

Table 59: Results following logistic regression for PC of entire cohort (N=278)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
10	ABCC2	exm848522	rs17222617	Alle F	3.4	0.58	1.1-10.7	0.035	1
10	ABCC2	exm848522	rs17222617	Dom	3.4	0.58	1.1-10.7	0.035	1
12	SLCO1B1	exm988933	rs2306283	Dom	0.54	0.3	0.3-0.97	0.04	1

After logistic regression for PC of the Caucasian- only cohort no results remained significant.

Should that be changed to Caucasian control cohort? If so that needs to be changed throughout and can be abbreviated to CCC.

Results were corrected using logistic regression in SPSS for Gender, Ethnicity, Type of CNI at baseline, Type of MPA at baseline, Use of induction agent at baseline, Use of valganciclovir prophylaxis, Type of transplant donor and HLAMM.

Table 60: Results of logistic regression for these confounders

Table 60: Logistic regression for entire cohort

Chr	Gene	EXM	rs number	Model	P value	Bonf	Exp (β)	95% CI Exp (β)
10	ABCC2	exm848522	rs17222617	Alle F	0.07	1	3.08	0.9-10.4
10	ABCC2	exm848522	rs17222617	Dom	0.07	1	3.08	0.9-10.4
12	SLCO1B1	exm988933	rs2306283	Dom	0.04	1	0.52	0.3-0.96

Unsupervised analysis results

Anaemia

Table61-63: Results for Fishers exact test for anaemia following QC of the data

Table 61: Allele frequency model

CHR	Gene	SNP	rs number	Model	P value	Bonf
6	Unspecified	exm-rs2621367	rs2621367	Alle F	0.00001	0.45
6	Unspecified	exm-rs2621366	rs2621366	Alle F	0.00003	0.82
6	Unspecified	exm-rs2621338	rs2621338	Alle F	0.00003	1
6	HLA-DOB	exm-rs2071474	rs2071474	Alle F	0.00005	1
6	HLA-DOB	exm-rs11244	rs11244	Alle F	0.00009	1

Table 62: Dominant model

CHR	Gene	SNP	rs number	Model	P value	Bonf
8	TMEM70	exm706302	rs1053079	Dom	0.00004	1
6	Unspecified	exm-rs2621367	rs2621367	Dom	0.00005	1
6	Unspecified	exm-rs2621338	rs2621338	Dom	0.00005	1
6	Unspecified	exm-rs2621366	rs2621366	Dom	0.00009	1
21	between KRTAP7-1 and KRTAP11-1	exm-rs7283316	rs7283316	Dom	0.0001	1

Table 63: Recessive model

CHR	Gene	SNP	rs number	Model	P value	Bonf
11	GRAMD1B	exm2249628	rs10893053	Rec	0.00005	1
18	Unspecified	exm2253444		Rec	0.0003	1
8	ASAP1	exm720843	rs966185	Rec	0.0004	1
3	AHSG	exm370881	rs4917	Rec	0.0008	1
22	CHEK2	exm-rs738722	rs738722	Rec	0.0008	1

Manhattan plot for unsupervised analysis of Anaemia

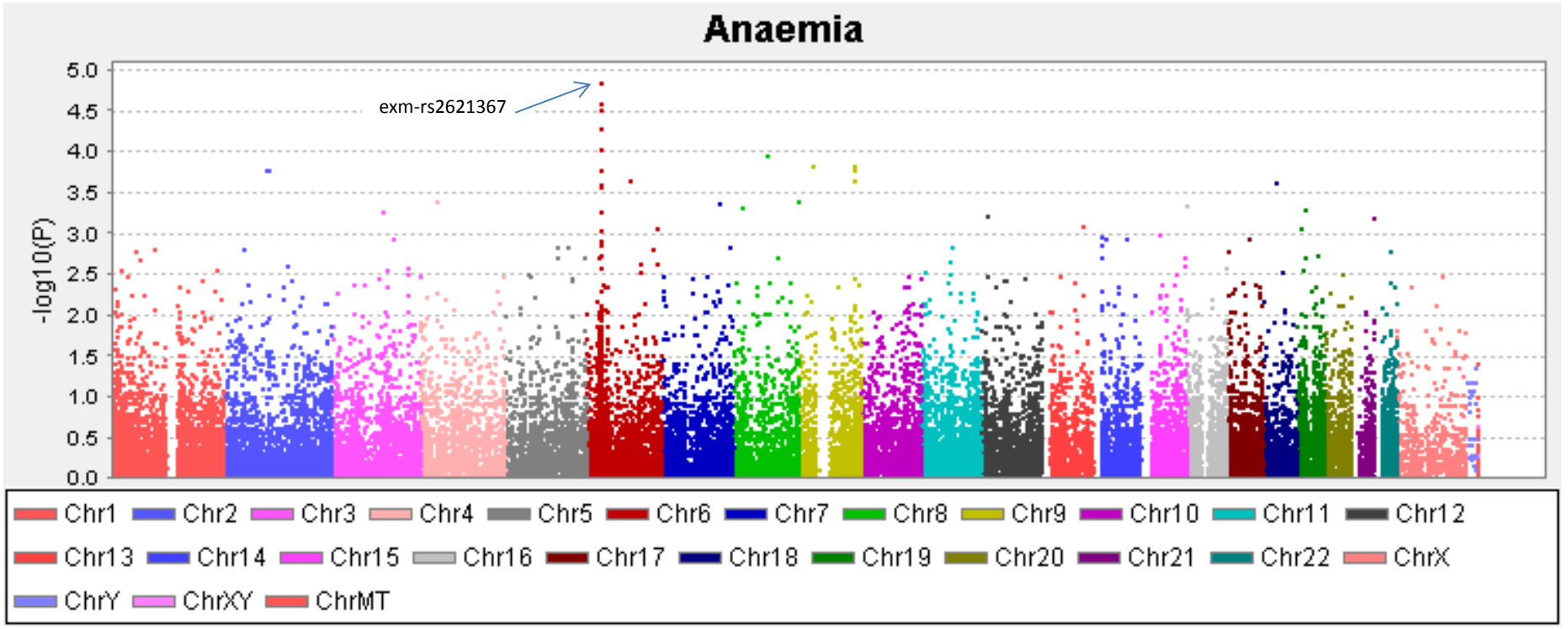


Table 64: Results following logistic regression for PC of entire population (N=278)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value	Bonf
8	TMEM70	exm706302	rs1053079	Alle F	3.0	0.3	1.8-5.1	0.00001	1
8	TMEM70	exm706306	rs1053077	Alle F	3.0	0.3	1.8-5.1	0.00001	1
6	Unspecified	exm-rs2621367	rs2621367	Alle F	0.4	0.2	0.3-0.7	0.00006	1
3	CLSTN2	exm-rs11708189	rs11708189	Alle F	1.9	0.2	1.3-2.9	0.001	1
2	ABCG8	exm190428	rs6544718	Alle F	0.5	0.3	0.3-0.7	0.001	1
8	TMEM70	exm706302	rs1053079	Dom	3.2	0.3	1.9-5.5	0.00003	1
8	TMEM70	exm706306	rs1053077	Dom	3.2	0.3	1.9-5.5	0.00003	1
6	Unspecified	exm-rs2621367	rs2621367	Dom	0.4	0.3	0.2-0.6	0.00006	1
6	Unspecified	exm-rs2621338	rs2621338	Dom	0.4	0.3	0.2-0.6	0.00006	1
21	between KRTAP7-1 and KRTAP11-1	exm-rs7283316	rs7283316	Dom	0.4	0.3	0.2-0.6	0.0001	1
8	ASAP1	exm720843	rs966185	Rec	3.3	0.3	1.7-6.3	0.0003	1
1	AGBL4	exm2264835	rs657452	Rec	3.6	0.4	1.8-7.5	0.0004	1
3	AHSG	exm370881	rs4917	Rec	3.5	0.4	1.7-7.1	0.0007	1
3	AHSG	exm370882	rs4918	Rec	3.5	0.4	1.7-7.1	0.0007	1
18	Unspecified	exm2253444		Rec	0.3	0.4	0.1-0.6	0.0008	1

Biopsy proven acute rejection (all)

Tables 65-67: Results of fishers exact test for BPAR all following QC of the data

Table 65: Allele frequency model

CHR	Gene	SNP	rs number	Model	P value	Bonf
12	VWF	exm976501	rs35335161	All F	0.00004	1
18	Between CDH2 and ARIH2P1	exm-rs11083271	rs11083271	All F	0.0002	1
4	CSN3	exm404085	rs3775739	All F	0.0002	1
11	OR10A2	exm887083	rs10839632	All F	0.0003	1
11	PDHX	exm900157	rs11539202	All F	0.0004	1

Table 66: Dominant model

CHR	Gene	SNP	rs number	Model	P value	Bonf
18	Between CDH2 and ARIH2P1	exm-rs11083271	rs11083271	Dom	0.00002	0.52
12	VWF	exm976501	rs35335161	Dom	0.00006	1
11	PDHX	exm900157	rs11539202	Dom	0.0002	1
7		exm2270569		Dom	0.0004	1
4	CSN3	exm404085	rs3775739	Dom	0.0005	1

Table 67: Recessive model

CHR	Gene	SNP	rs number	Model	P value	Bonf
2	Unspecified	exm-rs7584993	rs7584993	Rec	0.000067	1
21	LINC00478	exm2272957	rs239049	Rec	0.0001	1
1	CDC42BPA	exm154259	rs1929860	Rec	0.0002	1
10	VSTM4	exm823678	rs13088	Rec	0.0003	1
6	NUP153	exm518663	rs2228375	Rec	0.0005	1

Manhattan plot for unsupervised analysis of BPAR

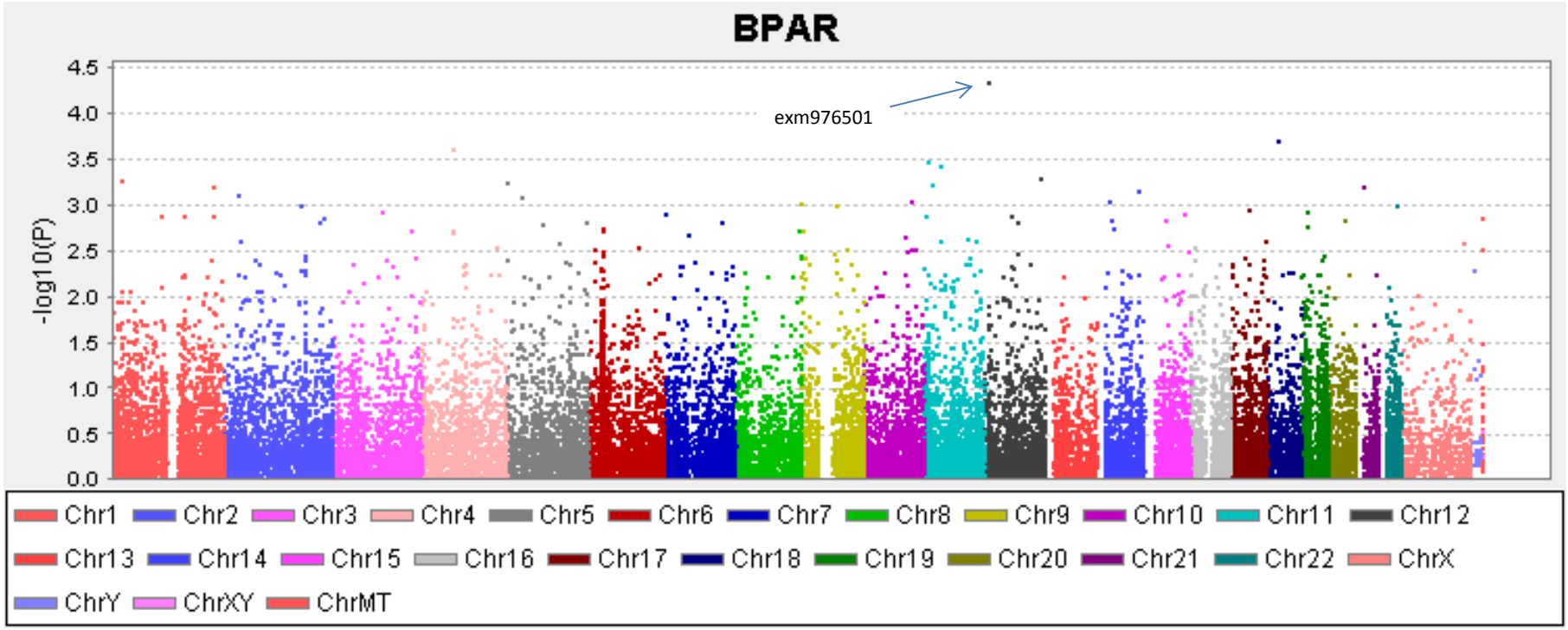


Table 68: Results following logistic regression for PC of entire cohort (N=278)

CHR	gene	SNP	rs number	Model	OR	SE	95% CI	P value
12	VWF	exm976501	rs35335161	Alle F	5.4	0.5	2.2-13.5	0.0003
18	between CDH2 and ARIH2P1	exm-rs11083271	rs11083271	Alle F	0.4	0.3	0.2-0.6	0.0004
11	OR10A2	exm887083	rs10839632	Alle F	2.3	0.2	1.4-3.6	0.0005
1	between IGSF21 and KLHDC7A	exm-rs3007729	rs3007729	Alle F	2.2	0.2	1.4-3.4	0.0007
7	EGFR	exm-rs11979158	rs11979158	Alle F	2.8	0.3	1.5-5	0.0007
18	between CDH2 and ARIH2P1	exm-rs11083271	rs11083271	DOM	0.2	0.34	0.1-0.5	0.00003
12	VWF	exm976501	rs35335161	DOM	7.4	0.5	2.8-19.9	0.00006
7	Unspecified	exm2270569		DOM	3.8	0.4	1.8-8.0	0.0005
11	PDHX	exm900157	rs11539202	DOM	3.3	0.3	1.7-6.5	0.0005
2	Unspecified	exm2254828		DOM	0.3	0.3	0.2-0.6	0.0006
2	Unspecified	exm-rs7584993	rs7584993	REC	5.8	0.4	2.5-13.3	0.00004
1	CDC42BPA	exm154259	rs1929860	REC	5.8	0.5	2.4-13.9	0.0001
6	NUP153	exm518663	rs2228375	REC	4.2	0.4	1.9-9.3	0.0005
1	Between SRRM1 and CLIC4	exm-rs4601530	rs4601530	REC	5.2	0.5	2-13.3	0.0006
6	NUP153	exm-rs12199222	rs12199222	REC	4.3	0.4	1.8-9.9	0.0007

Gastrointestinal side effects

Tables 69-71: Results of fishers exact test for GISE following QC of the data

Table 69: Allele frequency model

CHR	Gene	SNP	rs number	Model	P value	Bonf
18	Unsupervised	exm2268111		All F	0.0001	1
6	PSORS1C1	exm-rs1265100	rs1265100	All F	0.0001	1
18	DCC	exm-rs7506909	rs7506909	All F	0.0001	1
12	ZNF605	exm2271816	rs7778	All F	0.0001	1
4	BOD1L	exm390066	rs3733557	All F	0.0002	1

Table 70: Dominant model

CHR	Gene	SNP	rs number	Model	P value	Bonf
15	ARRDC4	exm1192081	rs2130882	Dom	0.0002	1
8	RBPMS	exm2266605	rs2979531	Dom	0.0003	1
6	Between WASF5P and HLA-B	exm-rs3873386	rs3873386	Dom	0.0003	1
4	BOD1L	exm390066	rs3733557	Dom	0.0004	1
4	EPHA5	exm2269928	rs12644356	Dom	0.0004	1

Table 71: Recessive model

CHR	Gene	SNP	rs number	Model	P value	Bonf
6	Between BDH2P1 and C6orf168	exm-rs2132683	rs2132683	Rec	0.000079	1
12	ZNF605	exm2271816	rs7778	Rec	0.0001	1
19	DKKL1	exm1490415	rs2288481	Rec	0.0004	1
19	CKM	exm2268199	rs377993	Rec	0.0004	1
17	Unspecified	exm-rs7217319	rs7217319	Rec	0.0004	1

Manhattan plot for unsupervised analysis of BPAR

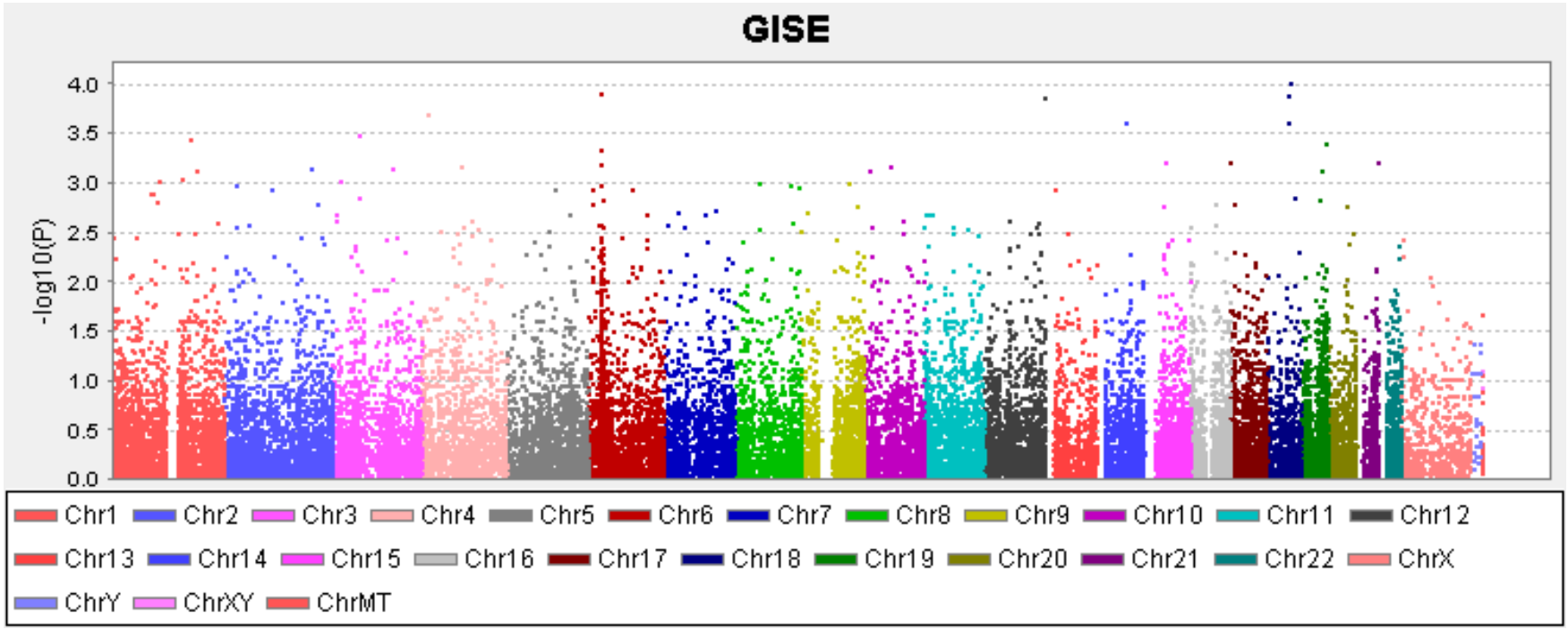


Table 72: Results following logistic regression for PC of entire cohort (N=278)

CHR	Gene	SNP	rs number	Model	OR	SE	95% CI	P value
6	PSORS1C1	exm-rs1265100	rs1265100	Alle F	2.6	0.2	1.6-4.2	0.00009
18	Unspecified	exm2268111		Alle F	0.4	0.2	0.3-0.7	0.0001
4	BOD1L	exm390066	rs3733557	Alle F	2.7	0.3	1.6-4.5	0.0001
19	NLRP2	exm1507217	rs1043673	Alle F	2.1	0.2	1.4-3.2	0.0002
10	MCM10	exm810209	rs2274110	Alle F	2.6	0.3	1.6-4.5	0.0003
6	MUC22	exm-rs2517554	rs2517554	Dom	0.3	0.3	0.2-0.6	0.0001
4	BOD1L	exm390066	rs3733557	Dom	3	0.3	1.7-5.3	0.0005
3	SETD5	exm287880	rs11542009	Dom	3.6	0.3	1.9-7.2	0.0002
6	PSORS1C1	exm-rs1265100	rs1265100	Dom	2.8	0.3	1.6-4.9	0.0002
15	ARRDC4	exm1192081	rs2130882	Dom	0.3	0.3	0.2-0.6	0.0002
19	DKKL1	exm1490415	rs2288481	Rec	7.7	0.6	2.5-23.2	0.0003
3	TMEM108	exm2255702	rs1197314	Rec	3	0.3	1.6-5.5	0.0004
17	ARHGEF15	exm1292049	rs3744647	Rec	3.2	0.3	1.7-6.2	0.0005
12	ZNF605	exm2271816	rs7778	Rec	5.6	0.5	2.1-14.8	0.0005
7	Between CNTNAP2 and MIR548T	exm2270711		Rec	2.8	0.3	1.6-4.9	0.0005

8 References

1. Illumina. Illumina Human Exome BeadChips Data sheet. 2011 [cited 2013 24th May 2013]; Available from: http://www.illumina.com/documents/products/datasheets/datasheet_humanexome_beadchips.pdf.

2. Dalal P, Grafals M, Chhabra D, Gallon L. Mycophenolate mofetil: safety and efficacy in the prophylaxis of acute kidney transplantation rejection. Therapeutics and clinical risk management. 2009;5(1):139-49. Epub 2009/05/14.

3. Golshayan D, Pascual M, Vogt B. Mycophenolic acid formulations in adult renal transplantation - update on efficacy and tolerability. Therapeutics and clinical risk management. 2009;5(4):341-51. Epub 2009/09/16.

4. Knight SR, Russell NK, Barcena L, Morris PJ. Mycophenolate mofetil decreases acute rejection and may improve graft survival in renal transplant recipients when compared with azathioprine: a systematic review. Transplantation. 2009;87(6):785-94. Epub 2009/03/21.

5. Moreso F, Hernandez D. Has the survival of the graft improved after renal transplantation in the era of modern immunosuppression? Nefrologia : publicacion oficial de la Sociedad Espanola Nefrologia. 2013;33(1):14-26. Epub 2013/02/01.

6. Bouamar R, Hesselink DA, van Schaik RH, Weimar W, van der Heiden IP, de Fijter JW, et al. Mycophenolic acid-related diarrhea is not associated with polymorphisms in SLCO1B nor with ABCB1 in renal transplant recipients. Pharmacogenetics and genomics. 2012;22(6):399-407. Epub 2011/09/01.

7. Sollinger HW, Sundberg AK, Leverson G, Voss BJ, Pirsch JD. Mycophenolate mofetil versus enteric-coated mycophenolate sodium: a large, single-center comparison of dose adjustments and outcomes in kidney transplant recipients. Transplantation. 2010;89(4):446-51. Epub 2010/02/24.

8. Michelon H, Konig J, Durrbach A, Quteineh L, Verstuyft C, Furlan V, et al. SLCO1B1 genetic polymorphism influences mycophenolic acid tolerance in renal transplant recipients. Pharmacogenomics. 2010;11(12):1703-13. Epub 2010/12/15.

9. Laftavi MR, Hai F, Laftavi H, Feng L, Said M, Patel S, et al. Mycophenolic acid dose reductions result in poor long-term renal allograft survival: comparison between mycophenolate sodium and mycophenolate mofetil. Transplantation proceedings. 2011;43(2):478-81. Epub 2011/03/29.

10. Clayton PA, McDonald SP, Chapman JR, Chadban SJ. Mycophenolate versus azathioprine for kidney transplantation: a 15-year follow-up of a randomized trial. Transplantation. 2012;94(2):152-8. Epub 2012/06/26.

11. Sprangers B, Kuypers DR, Vanrenterghem Y. Immunosuppression: Does one regimen fit all? Transplantation. 2011;92(3):251-61.

12. Thervet E, Anglicheau D, Legendre C, Beaune P. Role of pharmacogenetics of immunosuppressive drugs in organ transplantation. Therapeutic drug monitoring. 2008;30(2):143-50. Epub 2008/03/28.

13. Elens L, Hesselink DA, van Schaik RH, van Gelder T. Pharmacogenetics in kidney transplantation: recent updates and potential clinical applications. Molecular diagnosis & therapy. 2012;16(6):331-45. Epub 2012/11/30.

14. Gourishankar S, Houde I, Keown PA, Landsberg D, Cardella CJ, Barama AA, et al. The CLEAR study: a 5-day, 3-g loading dose of mycophenolate mofetil versus standard 2-g dosing in renal transplantation. Clinical journal of the American Society of Nephrology : CJASN. 2010;5(7):1282-9. Epub 2010/05/26.

15. Betonico GN, Abudd-Filho M, Goloni-Bertollo EM, Pavarino-Bertelli E. Pharmacogenetics of mycophenolate mofetil: a promising different approach to tailoring immunosuppression? Journal of nephrology. 2008;21(4):503-9. Epub 2008/07/25.

16. Sanchez-Fructuoso AI, Maestro ML, Calvo N, Viudarreta M, Perez-Flores I, Veganzone S, et al. The prevalence of uridine diphosphate-glucuronosyltransferase 1A9 (UGT1A9) gene promoter region single-nucleotide polymorphisms T-275A and C-2152T and its influence on mycophenolic acid pharmacokinetics in stable renal transplant patients. Transplantation proceedings. 2009;41(6):2313-6. Epub 2009/09/01.

17. MacPhee IAM, Fredericks S, Holt DW. Does pharmacogenetics have the potential to allow the individualisation of immunosuppressive drug dosing in organ transplantation? Expert Opinion on Pharmacotherapy. 2005;6(15):2593-605.

18. Ingelman-Sundberg M. Pharmacogenetics: an opportunity for a safer and more efficient pharmacotherapy. Journal of internal medicine. 2001;250(3):186-200. Epub 2001/09/14.

19. Evans WE, McLeod HL. Pharmacogenomics--drug disposition, drug targets, and side effects. The New England journal of medicine. 2003;348(6):538-49. Epub 2003/02/07.

20. Ma Q, Lu AYH. Pharmacogenetics, pharmacogenomics, and individualized medicine. Pharmacological Reviews. 2011;63(2):437-59.

21. Newman WG. Pharmacogenetics: transforming clinical medicine. The journal of the Royal College of Physicians of Edinburgh. 2012;42(3):244-7. Epub 2012/09/07.

22. Mutsatsa S, Currid TJ. Pharmacogenetics: a reality or misplaced optimism? Journal of psychiatric and mental health nursing. 2013;20(4):314-20. Epub 2012/04/20.

23. Li-Wan-Po A. Pharmacogenetics and personalized medicine. Journal of clinical pharmacy and therapeutics. 2012;37(6):617-9. Epub 2012/09/28.

24. Alvan G, Bertilsson L, Dahl ML, Ingelman-Sundberg M, Sjoqvist F. Moving toward genetic profiling in patient care: the scope and rationale of pharmacogenetic/ecogenetic investigation. Drug metabolism and disposition: the biological fate of chemicals. 2001;29(4 Pt 2):580-5. Epub 2001/03/22.

25. Kirchheiner J, Fuhr U, Brockmoller J. Pharmacogenetics-based therapeutic recommendations--ready for clinical practice? Nature reviews Drug discovery. 2005;4(8):639-47. Epub 2005/08/02.

26. Manolio TA, Collins FS. The HapMap and genome-wide association studies in diagnosis and therapy. Annual review of medicine. 2009;60:443-56. Epub 2009/07/28.

27. Geraci F. A comparison of several algorithms for the single individual SNP haplotyping reconstruction problem. Bioinformatics. 2010;26(18):2217-25. Epub 2010/07/14.

28. Zhang W, Dolan ME. Impact of the 1000 genomes project on the next wave of pharmacogenomic discovery. Pharmacogenomics. 2010;11(2):249-56. Epub 2010/02/09.

29. Howland RH. Future prospects for pharmacogenetics in the quest for personalized medicine. Journal of psychosocial nursing and mental health services. 2012;50(12):13-6. Epub 2013/03/05.

30. Johnson JA, Klein TE, Relling MV. Clinical implementation of pharmacogenetics: more than one gene at a time. Clinical pharmacology and therapeutics. 2013;93(5):384-5. Epub 2013/04/20.

31. Kato K. Impact of the next generation DNA sequencers. International journal of clinical and experimental medicine. 2009;2(2):193-202. Epub 2009/08/18.

32. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature. 2009;461(7261):272-6. Epub 2009/08/18.

33. Ovunc B, Otto EA, Vega-Warner V, Saisawat P, Ashraf S, Ramaswami G, et al. Exome sequencing reveals cubilin mutation as a single-gene cause of proteinuria. Journal of the American Society of Nephrology : JASN. 2011;22(10):1815-20. Epub 2011/09/10.

34. Sastre L. New DNA sequencing technologies open a promising era for cancer research and treatment. Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico. 2011;13(5):301-6. Epub 2011/05/21.

35. Amigo J, Salas A, Phillips C. ENGINES: exploring single nucleotide variation in entire human genomes. BMC bioinformatics. 2011;12:105. Epub 2011/04/21.

36. Picard N, Marquet P. The influence of pharmacogenetics and cofactors on clinical outcomes in kidney transplantation. Expert opinion on drug metabolism & toxicology.7(6):731-43.

37. Rosso Felipe C, de Sandes TV, Sampaio ELM, Park SI, Silva HT, Jr., Medina Pestana JO. Clinical impact of polymorphisms of transport proteins and enzymes involved in the metabolism of immunosuppressive drugs. Transplantation proceedings.41(5):1441-55.

38. MacPhee IAM. Use of pharmacogenetics to optimize immunosuppressive therapy. Therapeutic drug monitoring. 2010;32(3):261-4.

39. Cattaneo D, Baldelli S, Perico N. Pharmacogenetics of immunosuppressants: progress, pitfalls and promises. American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons. 2008;8(7):1374-83. Epub 2008/05/31.

40. Liang F, Holt DW, MacPhee IAM. Pharmacogenetics as a tool for optimising drug therapy in solid-organ transplantation. Expert Opinion on Pharmacotherapy. 2007;8(13):2045-58.

41. Yagil Y. Pharmacogenomic considerations for immunosuppressive therapy. Pharmacogenomics. 2003;4(3):309-19.

42. Sanghavi. K, Sangkuhl. K, Lamba. V. Mycophenolic acid Pathway, Pharmacokinetics/Pharmacodynamics. 2012; Available from: http://www.pharmgkb.org/pathway/PA165964832.

43. Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, et al. Pharmacogenomics knowledge for personalized medicine. Clinical pharmacology and therapeutics. 2012;92(4):414-7. Epub 2012/09/21.

44. Zhao W, Fakhoury M, Deschenes G, Roussey G, Brochard K, Niaudet P, et al. Population pharmacokinetics and pharmacogenetics of mycophenolic acid following administration of mycophenolate mofetil in de novo pediatric renal-transplant patients. Journal of clinical pharmacology. 2010;50(11):1280-91. Epub 2010/02/12.

45. Barraclough KA, Staatz CE, Isbel NM, Johnson DW. Therapeutic monitoring of mycophenolate in transplantation: is it justified? Current drug metabolism. 2009;10(2):179-87. Epub 2009/03/12.

46. Levesque E, Benoit-Biancamano MO, Delage R, Couture F, Guillemette C. Pharmacokinetics of mycophenolate mofetil and its glucuronide metabolites in healthy volunteers. Pharmacogenomics. 2008;9(7):869-79. Epub 2008/07/04.

47. Allison AC, Eugui EM. Mycophenolate mofetil and its mechanisms of action. Immunopharmacology. 2000;47(2-3):85-118. Epub 2000/07/06.

48. Wang J, Yang JW, Zeevi A, Webber SA, Girnita DM, Selby R, et al. IMPDH1 gene polymorphisms and association with acute rejection in renal transplant patients. Clinical pharmacology and therapeutics. 2008;83(5):711-7. Epub 2007/09/14.

49. Barraclough KA, Lee KJ, Staatz CE. Pharmacogenetic influences on mycophenolate therapy. Pharmacogenomics. 2010;11(3):369-90. Epub 2010/03/20.

50. Winnicki W, Weigel G, Sunder-Plassmann G, Bajari T, Winter B, Herkner H, et al. An inosine 5'-monophosphate dehydrogenase 2 single-nucleotide polymorphism impairs the effect of mycophenolic acid. The pharmacogenomics journal. 2010;10(1):70-6. Epub 2009/09/23.

51. Kuypers DR, Naesens M, Vermeire S, Vanrenterghem Y. The impact of uridine diphosphate-glucuronosyltransferase 1A9 (UGT1A9) gene promoter region single-nucleotide polymorphisms T-275A and C-2152T on early mycophenolic acid dose-interval exposure in de novo renal allograft recipients. Clinical pharmacology and therapeutics. 2005;78(4):351-61. Epub 2005/10/04.

52. Baldelli S, Merlini S, Perico N, Nicastri A, Cortinovis M, Gotti E, et al. C-440T/T-331C polymorphisms in the UGT1A9 gene affect the pharmacokinetics of mycophenolic acid in kidney transplantation. Pharmacogenomics. 2007;8(9):1127-41. Epub 2007/10/11.

53. Levesque E, Delage R, Benoit-Biancamano MO, Caron P, Bernard O, Couture F, et al. The impact of UGT1A8, UGT1A9, and UGT2B7 genetic polymorphisms on the pharmacokinetic profile of mycophenolic acid after a single oral dose in healthy volunteers. Clinical pharmacology and therapeutics. 2007;81(3):392-400. Epub 2007/03/07.

54. Honarbakhsh N, Rouini MR, Lesan-Pezeshki M, Javadi MR, Karimzadeh I, Mohebbi N, et al. Mycophenolic Acid pharmacokinetics early after kidney transplant. Experimental and clinical transplantation : official journal of the Middle East Society for Organ Transplantation. 2013;11(2):112-7. Epub 2012/11/28.

55. Kaplan B. Mycophenolic acid trough level monitoring in solid organ transplant recipients treated with mycophenolate mofetil: association with clinical outcome. Current medical research and opinion. 2006;22(12):2355-64. Epub 2007/01/30.

56. van Gelder T. Mycophenolate blood level monitoring: recent progress. American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons. 2009;9(7):1495-9. Epub 2009/06/13.

57. Le Meur Y, Buchler M, Thierry A, Caillard S, Villemain F, Lavaud S, et al. Individualized mycophenolate mofetil dosing based on drug exposure significantly improves patient outcomes after renal transplantation. American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons. 2007;7(11):2496-503. Epub 2007/10/03.

58. van Gelder T, Silva HT, de Fijter JW, Budde K, Kuypers D, Tyden G, et al. Comparing mycophenolate mofetil regimens for de novo renal transplant recipients: the fixed-dose concentration-controlled trial. Transplantation. 2008;86(8):1043-51. Epub 2008/10/24.

59. Tedesco-Silva H, Felipe CR, Park SI, Pinheiro-Machado PG, Garcia R, Slade A, et al. Randomized crossover study to assess the inter- and intrasubject variability of morning mycophenolic acid concentrations from enteric-coated mycophenolate sodium and mycophenolate mofetil in stable renal transplant recipients. Clinical transplantation. 2010;24(4):E116-23. Epub 2010/01/06.

60. Ji Y, Gu J, Makhov AM, Griffith JD, Mitchell BS. Regulation of the interaction of inosine monophosphate dehydrogenase with mycophenolic Acid by GTP. The Journal of biological chemistry. 2006;281(1):206-12. Epub 2005/10/26.

61. Sankatsing SU, Prins JM, Yong SL, Roelofsen J, van Kuilenburg AB, Kewn S, et al. Mycophenolate mofetil inhibits T-cell proliferation in kidney transplant recipients without lowering intracellular dGTP and GTP. Transplant international : official journal of the European Society for Organ Transplantation. 2008;21(11):1066-71. Epub 2008/08/14.

62. Rejman AS, Mansell MA, Grimes AJ, Joekes AM. Rapid correction of red-cell nucleotide abnormalities following successful renal transplantation. British journal of haematology. 1985;61(3):433-43. Epub 1985/11/01.

63. Goldsmith D, Carrey EA, Edbury S, Smolenski RT, Jagodzinski P, Simmonds HA. Mycophenolate mofetil, an inhibitor of inosine monophosphate dehydrogenase, causes a paradoxical elevation of GTP in erythrocytes of renal transplant patients. Clin Sci (Lond). 2004;107(1):63-8. Epub 2004/01/16.

64. Jagodzinski P, Lizakowski S, Smolenski RT, Slominska EM, Goldsmith D, Simmonds HA, et al. Mycophenolate mofetil treatment following renal transplantation decreases GTP concentrations in mononuclear leucocytes. Clin Sci (Lond). 2004;107(1):69-74. Epub 2004/01/16.

65. van Schaik RH, van Agteren M, de Fijter JW, Hartmann A, Schmidt J, Budde K, et al. UGT1A9 -275T>A/-2152C>T polymorphisms correlate with low MPA exposure and acute rejection in MMF/tacrolimus-treated kidney transplant patients. Clinical pharmacology and therapeutics. 2009;86(3):319-27. Epub 2009/06/06.

66. Johnson LA, Oetting WS, Basu S, Prausa S, Matas A, Jacobson PA. Pharmacogenetic effect of the UGT polymorphisms on mycophenolate is modified by calcineurin inhibitors. European journal of clinical pharmacology. 2008;64(11):1047-56. Epub 2008/06/24.

67. Prausa SE, Fukuda T, Maseck D, Curtsinger KL, Liu C, Zhang K, et al. UGT genotype may contribute to adverse events following medication with mycophenolate mofetil in pediatric kidney transplant recipients. Clinical pharmacology and therapeutics. 2009;85(5):495-500. Epub 2009/02/20.

68. Betonico GN, Abbud-Filho M, Goloni-Bertollo EM, Alvarenga MP, Guillemette C, Villeneuve L, et al. Influence of UDP-glucuronosyltransferase polymorphisms on mycophenolate mofetil-induced side effects in kidney transplant patients. Transplantation proceedings. 2008;40(3):708-10. Epub 2008/05/06.

69. Kagaya H, Inoue K, Miura M, Satoh S, Saito M, Tada H, et al. Influence of UGT1A8 and UGT2B7 genetic polymorphisms on mycophenolic acid pharmacokinetics in Japanese renal transplant recipients. European journal of clinical pharmacology. 2007;63(3):279-88. Epub 2007/01/11.

70. Yang JW, Lee PH, Hutchinson IV, Pravica V, Shah T, Min DI. Genetic polymorphisms of MRP2 and UGT2B7 and gastrointestinal symptoms in renal transplant recipients taking mycophenolic acid. Therapeutic drug monitoring. 2009;31(5):542-8. Epub 2009/09/05.

71. van Agteren M, Armstrong VW, van Schaik RH, de Fijter H, Hartmann A, Zeier M, et al. AcylMPAG plasma concentrations and mycophenolic acid-related side effects in patients undergoing renal transplantation are not related to the UGT2B7-840G>A gene polymorphism. Therapeutic drug monitoring. 2008;30(4):439-44. Epub 2008/07/22.

72. Fukuda T, Goebel J, Cox S, Maseck D, Zhang K, Sherbotie JR, et al. UGT1A9, UGT2B7, and MRP2 genotypes can predict mycophenolic acid pharmacokinetic variability in pediatric kidney transplant recipients. Therapeutic drug monitoring. 2012;34(6):671-9. Epub 2012/11/08.

73. Kagaya H, Miura M, Saito M, Habuchi T, Satoh S. Correlation of IMPDH1 gene polymorphisms with subclinical acute rejection and mycophenolic acid exposure parameters on day 28 after renal transplantation. Basic & clinical pharmacology & toxicology. 2010;107(2):631-6. Epub 2010/02/09.

74. Gensburger O, Van Schaik RH, Picard N, Le Meur Y, Rousseau A, Woillard JB, et al. Polymorphisms in type I and II inosine monophosphate dehydrogenase genes and association with clinical outcome in patients on mycophenolate mofetil. Pharmacogenetics and genomics. 2010;20(9):537-43. Epub 2010/08/04.

75. Sombogaard F, van Schaik RH, Mathot RA, Budde K, van der Werf M, Vulto AG, et al. Interpatient variability in IMPDH activity in MMF-treated renal transplant patients is correlated with IMPDH type II 3757T > C polymorphism. Pharmacogenetics and genomics. 2009;19(8):626-34. Epub 2009/07/21.

76. Fujiyama N, Miura M, Satoh S, Inoue K, Kagaya H, Saito M, et al. Influence of carboxylesterase 2 genetic polymorphisms on mycophenolic acid pharmacokinetics in Japanese renal transplant recipients. Xenobiotica; the fate of foreign compounds in biological systems. 2009;39(5):407-14. Epub 2009/03/11.

77. Naesens M, Kuypers DR, Verbeke K, Vanrenterghem Y. Multidrug resistance protein 2 genetic polymorphisms influence mycophenolic acid exposure in renal allograft recipients. Transplantation. 2006;82(8):1074-84. Epub 2006/10/25.

78. Miura M, Satoh S, Inoue K, Kagaya H, Saito M, Inoue T, et al. Influence of SLCO1B1, 1B3, 2B1 and ABCC2 genetic polymorphisms on mycophenolic acid pharmacokinetics in Japanese renal transplant recipients. European journal of clinical pharmacology. 2007;63(12):1161-9. Epub 2007/10/02.

79. Jacobson PA, Schladt D, Oetting WS, Leduc R, Guan W, Matas AJ, et al. Genetic determinants of mycophenolate-related anemia and leukopenia after transplantation. Transplantation. 2011;91(3):309-16. Epub 2010/11/26.

80. Bouamar R, Elens L, Shuker N, van Schaik RH, Weimar W, Hesselink DA, et al. Mycophenolic acid-related anemia and leucopenia in renal transplant recipients are related to genetic polymorphisms in CYP2C8. Transplantation. 2012;93(10):e39-40; author reply e1-2. Epub 2012/05/11.

81. Jacobson PAS, David; William S. Oetting; Robert Leduc; Weihau Guan; Arthur J. Matas; Vishal Lamba; Roslyn B. Mannon; Bruce A. Julian; Ajay Israni. Reply to "Genetic determinants of Mycophenolate-Related anaemia and leucopenia after transplantation". Transplantation. 2012;93(10):e41-e2.

82. Djebli N, Picard N, Rerolle JP, Le Meur Y, Marquet P. Influence of the UGT2B7 promoter region and exon 2 polymorphisms and comedications on Acyl-MPAG production in vitro and in adult renal transplant patients. Pharmacogenetics and genomics. 2007;17(5):321-30. Epub 2007/04/13.

83. Gilga J, Castledinea C, Fogarty D. UK Renal Registry 14th Annual Report:Chapter 1 UK RRT Incidence in 2010:national and centre-specific analyses. 2011.

84. MacPheea I, Webbb L, Casulab A, Udayaraj U. UK Renal Registry 14th Annual Report:Chapter 3 Demoraphic and BiochemistryProfile of Kidney Transplant Recipients in the UK in 2010: national and centrespecific analyses. 2011.

85. Castledine C, Casulaa A, Fogarty D. UK Renal Registry 14th Annual Report:Chapter 2 UK RRT Prevalence in 2010:national and centre-specific analyses. 2011.

86. Solez K, Colvin RB, Racusen LC, Haas M, Sis B, Mengel M, et al. Banff 07 classification of renal allograft pathology: updates and future directions. American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons. 2008;8(4):753-60. Epub 2008/02/26.

87. Peiffer DA, Gunderson KL. Design of tag SNP whole genome genotyping arrays. Methods Mol Biol. 2009;529:51-61. Epub 2009/04/22.

88. Huyghe JR, Jackson AU, Fogarty MP, Buchkovich ML, Stancakova A, Stringham HM, et al. Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. Nature genetics. 2013;45(2):197-201. Epub 2012/12/25.

89. Kiezun A, Garimella K, Do R, Stitziel NO, Neale BM, McLaren PJ, et al. Exome sequencing and the genetic basis of complex traits. Nature genetics. 2012;44(6):623-30. Epub 2012/05/30.

90. illumina. Illumina Beadchip genotyping. 2013; Available from: http://www.illumina.com/applications/genotyping.ilmn.

91. Shah TS, Liu JZ, Floyd JA, Morris JA, Wirth N, Barrett JC, et al. optiCall: a robust genotype-calling algorithm for rare, low-frequency and common variants. Bioinformatics. 2012;28(12):1598-603. Epub 2012/04/14.

92. Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. Basic statistical analysis in genetic case-control studies. Nature protocols. 2011;6(2):121-33. Epub 2011/02/05.

93. Fardo DW, Ionita-Laza I, Lange C. On quality control measures in genome-wide association studies: a test to assess the genotyping quality of individual probands in family-based association studies and an application to the HapMap data. PLoS genetics. 2009;5(7):e1000572. Epub 2009/07/25.

94. Laurie CC, Doheny KF, Mirel DB, Pugh EW, Bierut LJ, Bhangale T, et al. Quality control and quality assurance in genotypic data for genome-wide association studies. Genetic epidemiology. 2010;34(6):591-602. Epub 2010/08/19.

95. Southam L, Panoutsopoulou K, Rayner NW, Chapman K, Durrant C, Ferreira T, et al. The effect of genome-wide association scan quality control on imputation outcome for common variants. European journal of human genetics : EJHG. 2011;19(5):610-4. Epub 2011/01/27.

96. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics. 2007;81(3):559-75. Epub 2007/08/19.

97. Purcell S. PLINK V1.07. [31/08/2013]; Available from: http://pngu.mgh.harvard.edu/purcell/plink/.

98. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447(7145):661-78. Epub 2007/06/08.

99. Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, et al. Linkage disequilibrium in the human genome. Nature. 2001;411(6834):199-204. Epub 2001/05/11.

100. Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. Nature reviews Genetics. 2011;12(6):443-51. Epub 2011/05/19.

101. Goldstein JI, Crenshaw A, Carey J, Grant GB, Maguire J, Fromer M, et al. zCall: a rare variant caller for array-based genotyping: genetics and population analysis. Bioinformatics. 2012;28(19):2543-5. Epub 2012/07/31.

102. Do R, Kathiresan S, Abecasis GR. Exome sequencing and complex disease: practical aspects of rare variant association studies. Human molecular genetics. 2012;21(R1):R1-9. Epub 2012/09/18.

103. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nature genetics. 2006;38(8):904-9. Epub 2006/07/25.

104. Woillard JB, Rerolle JP, Picard N, Rousseau A, Drouet M, Munteanu E, et al. Risk of diarrhoea in a long-term cohort of renal transplant patients given mycophenolate mofetil: the significant role of the UGT1A8 2 variant allele. British journal of clinical pharmacology. 2010;69(6):675-83. Epub 2010/06/23.

105. Pruthia R, Casulaa A, I. M. UK Renal Registry 15th Annual Report: Chapter 3 Demographic and biochemistry profile of kidney transplant recipients in the UK in 2011: national and centre-specific analyses. 2013.

106. Ansell D, Feest T, Byrne C. The fifth annual report: The UK renal registry: Chapter 13: Transplantation. 2002.

107. Chen GD, Gu JL, Zhang XD, Qiu J, Wang CX, Chen LZ. Donor factors predictive for poor outcomes of living donor kidney transplantation. Transplantation proceedings. 2013;45(4):1445-8. Epub 2013/06/04.

108. Rao PS, Ojo A. The alphabet soup of kidney transplantation: SCD, DCD, ECD--fundamentals for the practicing nephrologist. Clinical journal of the American Society of Nephrology : CJASN. 2009;4(11):1827-31. Epub 2009/10/08.

109. Hellegering J, Visser J, Kloke HJ, D'Ancona FC, Hoitsma AJ, van der Vliet JA, et al. Poor early graft function impairs long-term outcome in living donor kidney transplantation. World journal of urology. 2012. Epub 2012/02/15.

110. Park HS, Hong YA, Kim HG, Choi SR, Sun IO, Chung BH, et al. Delayed graft function in living-donor renal transplantation: 10-year experience. Transplantation proceedings. 2012;44(1):43-6. Epub 2012/02/09.

111. Vu LT, Baxter-Lowe LA, Garcia J, McEnhill M, Summers P, Hirose R, et al. HLA-DR matching in organ allocation: balance between waiting time and rejection in pediatric kidney transplantation. Arch Surg. 2011;146(7):824-9. Epub 2011/07/20.

112. Transplant NBa. Activity in the UK, Activity Report 2011-2012. 2012.

113. Association JWPoTBTSaTR. United Kingdom Guidelines for living donor kidney transplantation. 2011.

114. Remuzzi G, Cravedi P, Costantini M, Lesti M, Ganeva M, Gherardi G, et al. Mycophenolate mofetil versus azathioprine for prevention of chronic allograft dysfunction in renal transplantation: the MYSS follow-up randomized, controlled clinical trial. Journal of the American Society of Nephrology : JASN. 2007;18(6):1973-85. Epub 2007/04/27.

115. Fishman JA, Emery V, Freeman R, Pascual M, Rostaing L, Schlitt HJ, et al. Cytomegalovirus in transplantation - challenging the status quo. Clinical transplantation. 2007;21(2):149-58. Epub 2007/04/12.

116. Tryphonopoulos P, Weppler D, Morris MI, Russo C, Nishida S, Levi DM, et al. The effect of donor-recipient cytomegalovirus serology on adult liver transplantation: a single center experience. Transplantation. 2011;92(9):1051-7. Epub 2011/08/31.

117. Tian YF, Liao CH, Chen MJ. Risk factors among donor characteristics which affect graft outcome in paired kidney transplantation. Transplantation proceedings. 2008;40(7):2281-4. Epub 2008/09/16.

118. Hodson EM, Ladhani M, Webster AC, Strippoli GF, Craig JC. Antiviral medications for preventing cytomegalovirus disease in solid organ transplant recipients. Cochrane Database Syst Rev. 2013;2:CD003774. Epub 2013/03/02.

119. Jacobson P, Schladt D, Oetting W, Leduc R, Guan W, Matas A, et al. Genetic determinants of mycophenolate related leukopenia and anemia from a 7-center consortium. American Journal of Transplantation. 2010;10:170.

120. Webster AC, Woodroffe RC, Taylor RS, Chapman JR, Craig JC. Tacrolimus versus ciclosporin as primary immunosuppression for kidney transplant recipients: meta-analysis and meta-regression of randomised trial data. BMJ. 2005;331(7520):810. Epub 2005/09/15.

121. Toledo AH, Hendrix L, Buchholz V, Fisher E, Newton K, Smith C, et al. Improvement of gastrointestinal symptoms after conversion from mycophenolate mofetil to enteric-coated mycophenolate sodium in liver transplant patients. Clinical transplantation.26(1):156-63.

122. Vanhove T, Kuypers D, Claes KJ, Evenepoel P, Meijers B, Naesens M, et al. Reasons for dose reduction of mycophenolate mofetil during the first year after renal transplantation and its impact on graft outcome. Transplant international : official journal of the European Society for Organ Transplantation. 2013. Epub 2013/06/12.

123. Einollahi B, Lessan-Pezeshki M, Rostami Z, Kalantar E, Afshar R, Beiraghdar F. Anemia after kidney transplantation in adult recipients: prevalence and risk factors. Transplantation proceedings. 2011;43(2):578-80. Epub 2011/03/29.

124. Malyszko J, Oberbauer R, Watschinger B. Anemia and Erythrocytosis in patients after kidney transplantation. Transplant international : official journal of the European Society for Organ Transplantation. 2012;25(10):1013-23. Epub 2012/06/22.

125. Malyszko J, Watschinger B, Przybylowski P, Durlik M. Anemia in solid organ transplantation. Annals of transplantation : quarterly of the Polish Transplantation Society. 2012;17(2):86-100. Epub 2012/06/30.

126. Cooper M, Salvadori M, Budde K, Oppenheimer F, Sollinger H, Zeier M. Enteric-coated mycophenolate sodium immunosuppression in renal transplant patients: efficacy and dosing. Transplant Rev (Orlando). 2012;26(4):233-40. Epub 2012/08/07.

127. Ortega F, Sanchez-Fructuoso A, Cruzado JM, Gomez-Alamillo JC, Alarcon A, Pallardo L, et al. Gastrointestinal quality of life improvement of renal transplant recipients converted from mycophenolate mofetil to enteric-coated mycophenolate sodium drugs or agents: mycophenolate mofetil and enteric-coated mycophenolate sodium. Transplantation. 2011;92(4):426-32. Epub 2011/07/16.

128. Fishman JA. Infection in solid-organ transplant recipients. The New England journal of medicine. 2007;357(25):2601-14. Epub 2007/12/21.

129. Snyder JJ, Israni AK, Peng Y, Zhang L, Simon TA, Kasiske BL. Rates of first infection following kidney transplant in the United States. Kidney international. 2009;75(3):317-26. Epub 2008/11/21.

130. Kosmadakis G, Daikos GL, Pavlopoulou ID, Gobou A, Kostakis A, Tzanatou-Exarchou H, et al. Infectious complications in the first year post renal transplantation. Transplantation proceedings. 2013;45(4):1579-83. Epub 2013/06/04.

131. Heemann U, Kliem V, Budde K, Hamza A, Jurgensen JS, Juarez F, et al. Mycophenolate mofetil maintenance therapy in renal transplant patients: long-term results of the TranCept STAY study. Clinical transplantation. 2012;26(6):919-26. Epub 2012/09/22.

132. Asberg A, Apeland T, Reisaeter AV, Foss A, Leivestad T, Heldal K, et al. Long-term outcomes after cyclosporine or mycophenolate withdrawal in kidney transplantation - results from an aborted trial. Clinical transplantation. 2013;27(2):E151-6. Epub 2013/01/29.

133. Lamba V, Sangkuhl K, Sanghavi K, Fish A, Altman RB, Klein TE. PharmGKB summary: mycophenolic acid pathway. Pharmacogenetics and genomics. 2014;24(1):73-9. Epub 2013/11/14.

134. Lunetta KL. Genetic association studies. Circulation. 2008;118(1):96-101. Epub 2008/07/02.

135. Forner K, Lamarine M, Guedj M, Dauvillier J, Wojcik J. Universal false discovery rate estimation methodology for genome-wide association studies. Human heredity. 2008;65(4):183-94. Epub 2007/12/13.

136. Maitland ML, Grimsley C, Kuttab-Boulos H, Witonsky D, Kasza KE, Yang L, et al. Comparative genomics analysis of human sequence variation in the UGT1A gene cluster. The pharmacogenomics journal. 2006;6(1):52-62. Epub 2005/11/30.

137. Gong QH, Cho JW, Huang T, Potter C, Gholami N, Basu NK, et al. Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. Pharmacogenetics. 2001;11(4):357-68. Epub 2001/07/04.

138. Krishnaswamy S, Hao Q, Al-Rohaimi A, Hesse LM, von Moltke LL, Greenblatt DJ, et al. UDP glucuronosyltransferase (UGT) 1A6 pharmacogenetics: II. Functional impact of the three most common nonsynonymous UGT1A6 polymorphisms (S7A, T181A, and R184S). The Journal of pharmacology and experimental therapeutics. 2005;313(3):1340-6. Epub 2005/03/12.

139. Pazik J, Oldak M, Dabrowski M, Lewandowski Z, Sitarek E, Podgorska M, et al. Association of UDP-glucuronosyltransferase 1A9 (UGT1A9) gene polymorphism with kidney allograft function. Annals of Transplantation.16(4):69-73.

140. Jordan de Luna C, Herrero Cervera MJ, Sanchez Lazaro I, Almenar Bonet L, Poveda Andres JL, Alino Pellicer SF. Pharmacogenetic study of ABCB1 and CYP3A5 genes during the first year following heart transplantation regarding tacrolimus or cyclosporine levels. Transplantation proceedings.43(6):2241-3.

141. Rausch F, Schicht M, Paulsen F, Ngueya I, Brauer L, Brandt W. "SP-G", a putative new surfactant protein--tissue localization and 3D structure. PloS one. 2012;7(10):e47789. Epub 2012/10/25.

142. Mittal RA, Hammel M, Schwarz J, Heschl KM, Bretschneider N, Flemmer AW, et al. SFTA2--a novel secretory peptide highly expressed in the lung--is modulated by lipopolysaccharide but not hyperoxia. PloS one. 2012;7(6):e40011. Epub 2012/07/07.

143. Low SK, Chung S, Takahashi A, Zembutsu H, Mushiroda T, Kubo M, et al. Genome-wide association study of chemotherapeutic agent-induced severe neutropenia/leucopenia for patients in Biobank Japan. Cancer science. 2013;104(8):1074-82. Epub 2013/05/08.

144. Lee JS, Bae JS, Kim JH, Kim JY, Park TJ, Pasaje CF, et al. Effect of diffuse panbronchiolitis critical region 1 polymorphisms on the risk of aspirin-exacerbated respiratory disease in Korean asthmatics. Respiratory care. 2012;57(5):758-63. Epub 2011/12/14.

145. Galvan CA, Elbarcha OC, Fernandez EJ, Beltramo DM, Soria NW. Rapid HCP5 single-nucleotide polymorphism genotyping: a simple allele-specific PCR method for prediction of hypersensitivity reaction to Abacavir. Clinica chimica acta; international journal of clinical chemistry. 2011;412(15-16):1382-4. Epub 2011/04/26.

146. Rodriguez-Novoa S, Cuenca L, Morello J, Cordoba M, Blanco F, Jimenez-Nacher I, et al. Use of the HCP5 single nucleotide polymorphism to predict hypersensitivity reactions to abacavir: correlation with HLA-B*5701. The Journal of antimicrobial chemotherapy. 2010;65(8):1567-9. Epub 2010/06/11.

147. Yogalingam G, Bonten EJ, van de Vlekkert D, Hu H, Moshiach S, Connell SA, et al. Neuraminidase 1 is a negative regulator of lysosomal exocytosis. Developmental cell. 2008;15(1):74-86. Epub 2008/07/09.

148. Stamatos NM, Liang F, Nan X, Landry K, Cross AS, Wang LX, et al. Differential expression of endogenous sialidases of human monocytes during cellular differentiation into macrophages. The FEBS journal. 2005;272(10):2545-56. Epub 2005/05/12.

149. Lillehoj EP, Hyun SW, Feng C, Zhang L, Liu A, Guang W, et al. NEU1 sialidase expressed in human airway epithelia regulates epidermal growth factor receptor (EGFR) and MUC1 protein signaling. The Journal of biological chemistry. 2012;287(11):8214-31. Epub 2012/01/17.

150. Barcellos LF, May SL, Ramsay PP, Quach HL, Lane JA, Nititham J, et al. High-density SNP screening of the major histocompatibility complex in systemic lupus erythematosus demonstrates strong evidence for independent susceptibility regions. PLoS genetics. 2009;5(10):e1000696. Epub 2009/10/24.

151. Gencheva M, Kato M, Newo AN, Lin RJ. Contribution of DEAH-box protein DHX16 in human pre-mRNA splicing. The Biochemical journal. 2010;429(1):25-32. Epub 2010/04/29.

152. Gencheva M, Lin TY, Wu X, Yang L, Richard C, Jones M, et al. Nuclear retention of unspliced pre-mRNAs by mutant DHX16/hPRP2, a spliceosomal DEAH-box protein. The Journal of biological chemistry. 2010;285(46):35624-32. Epub 2010/09/16.

153. Kottgen A, Pattaro C, Boger CA, Fuchsberger C, Olden M, Glazer NL, et al. New loci associated with kidney function and chronic kidney disease. Nature genetics. 2010;42(5):376-84. Epub 2010/04/13.

154. Torraco A, Verrigni D, Rizza T, Meschini MC, Vazquez-Memije ME, Martinelli D, et al. TMEM70: a mutational hot spot in nuclear ATP synthase deficiency with a pivotal role in complex V biogenesis. Neurogenetics. 2012;13(4):375-86. Epub 2012/09/19.

155. Spiegel R, Khayat M, Shalev SA, Horovitz Y, Mandel H, Hershkovitz E, et al. TMEM70 mutations are a common cause of nuclear encoded ATP synthase assembly defect: further delineation of a new syndrome. Journal of medical genetics. 2011;48(3):177-82. Epub 2010/12/15.

156. Brown MT, Andrade J, Radhakrishna H, Donaldson JG, Cooper JA, Randazzo PA. ASAP1, a phospholipid-dependent arf GTPase-activating protein that associates with and is phosphorylated by Src. Molecular and cellular biology. 1998;18(12):7038-51. Epub 1998/11/20.

157. Jian X, Brown P, Schuck P, Gruschus JM, Balbo A, Hinshaw JE, et al. Autoinhibition of Arf GTPase-activating protein activity by the BAR domain in ASAP1. The Journal of biological chemistry. 2009;284(3):1652-63. Epub 2008/11/20.

158. Goldsmith DJ, Carrey EA, Edbury SM, Marinaki AM, Simmonds HA. GTP concentrations are elevated in erythrocytes of renal transplant recipients when conventional immunosuppression is replaced by the inosine monophosphate dehydrogenase inhibitor mycophenolic acid mofetil (MMF). Nucleosides, nucleotides & nucleic acids. 2004;23(8-9):1407-9. Epub 2004/12/02.

159. Muendlein A, Stark N, Rein P, Saely CH, Geller-Rhomberg S, Geiger K, et al. Are AHSG polymorphisms directly associated with coronary atherosclerosis? Clinica chimica acta; international journal of clinical chemistry. 2012;413(1-2):287-90. Epub 2011/10/26.

160. Jung JY, Hwang YH, Lee H, Ro H, Chung W, Chae DW, et al. Association of AHSG gene polymorphisms and aortic stiffness in peritoneal dialysis patients. American journal of nephrology. 2010;31(6):510-7. Epub 2010/05/11.

161. Brown RM, Head RA, Morris AA, Raiman JA, Walter JH, Whitehouse WP, et al. Pyruvate dehydrogenase E3 binding protein (protein X) deficiency. Developmental medicine and child neurology. 2006;48(9):756-60. Epub 2006/08/15.

162. Sumi T, Matsumoto K, Shibuya A, Nakamura T. Activation of LIM kinases by myotonic dystrophy kinase-related Cdc42-binding kinase alpha. The Journal of biological chemistry. 2001;276(25):23092-6. Epub 2001/05/08.

163. Datta S, Chowdhury A, Ghosh M, Das K, Jha P, Colah R, et al. A genome-wide search for non-UGT1A1 markers associated with unconjugated bilirubin level reveals significant association with a polymorphic marker near a gene of the nucleoporin family. Annals of human genetics. 2012;76(1):33-41. Epub 2011/11/29.

164. Peddle L, Zipperlen K, Melay B, Hefferton D, Rahman P. Association of SEEK1 polymorphisms in Crohn's disease. Human immunology. 2004;65(7):706-9. Epub 2004/08/11.

165. Chang YT, Liu HN, Shiao YM, Lin MW, Lee DD, Liu MT, et al. A study of PSORS1C1 gene polymorphisms in Chinese patients with psoriasis. The British journal of dermatology. 2005;153(1):90-6. Epub 2005/07/21.

166. Slavotinek AM, Moshrefi A, Davis R, Leeth E, Schaeffer GB, Burchard GE, et al. Array comparative genomic hybridization in patients with congenital diaphragmatic hernia: mapping of four CDH-critical regions and sequencing of candidate genes at 15q26.1-15q26.2. European journal of human genetics : EJHG. 2006;14(9):999-1008. Epub 2006/06/01.

167. Yan Q, Wu X, Chen C, Diao R, Lai Y, Huang J, et al. Developmental expression and function of DKKL1/Dkkl1 in humans and mice. Reproductive biology and endocrinology : RB&E. 2012;10:51. Epub 2012/07/24.

168. Kusuhara S, Fukushima Y, Fukuhara S, Jakt LM, Okada M, Shimizu Y, et al. Arhgef15 promotes retinal angiogenesis by mediating VEGF-induced Cdc42 activation and potentiating RhoJ inactivation in endothelial cells. PloS one. 2012;7(9):e45858. Epub 2012/10/03.

169. Phillips EJ, Chung WH, Mockenhaupt M, Roujeau JC, Mallal SA. Drug hypersensitivity: pharmacogenetics and clinical syndromes. The Journal of allergy and clinical immunology. 2011;127(3 Suppl):S60-6. Epub 2011/03/05.

170. Pompeu YA, Stewart JD, Mallal S, Phillips E, Peters B, Ostrov DA. The structural basis of HLA-associated drug hypersensitivity syndromes. Immunological reviews. 2012;250(1):158-66. Epub 2012/10/11.

171. Medway C, Shi H, Bullock J, Black H, Brown K, Vafadar-Isfahani B, et al. Using In silico LD clumping and meta-analysis of genome-wide datasets as a complementary tool to investigate and validate new candidate biomarkers in Alzheimer's disease. International journal of molecular epidemiology and genetics. 2010;1(2):134-44. Epub 2010/01/01.

172. Bacanu SA, Whittaker JC, Nelson MR. How informative is a negative finding in a small pharmacogenetic study? The pharmacogenomics journal. 2012;12(2):93-5. Epub 2011/12/14.

173. Wilffert B, Swen J, Mulder H, Touw D, Maitland-Van der Zee AH, Deneer V. From evidence based medicine to mechanism based medicine. Reviewing the role of pharmacogenetics. International journal of clinical pharmacy. 2013;35(3):369-75. Epub 2010/11/05.

174. Wu TY, Peng Y, Pelleymounter LL, Moon I, Eckloff BW, Wieben ED, et al. Pharmacogenetics of the mycophenolic acid targets inosine monophosphate dehydrogenases IMPDH1 and IMPDH2: gene sequence variation and functional genomics. British journal of pharmacology. 2010;161(7):1584-98. Epub 2010/08/20.

175. Kang MJ, Kim HG, Kim JS, Oh do G, Um YJ, Seo CS, et al. The effect of gut microbiota on drug metabolism. Expert opinion on drug metabolism & toxicology. 2013;9(10):1295-308. Epub 2013/09/17.

176. Ramezani A, Raj DS. The gut microbiome, kidney disease, and targeted interventions. Journal of the American Society of Nephrology : JASN. 2014;25(4):657-70. Epub 2013/11/16.

177. Bakulski KM, Fallin MD. Epigenetic epidemiology: promises for public health research. Environmental and molecular mutagenesis. 2014;55(3):171-83. Epub 2014/01/23.

178. Mensaert K, Denil S, Trooskens G, Van Criekinge W, Thas O, De Meyer T. Next-generation technologies and data analytical approaches for epigenomics. Environmental and molecular mutagenesis. 2014;55(3):155-70. Epub 2013/12/12.

179. Zhong XB, Leeder JS. Epigenetic regulation of ADME-related genes: focus on drug metabolism and transport. Drug metabolism and disposition: the biological fate of chemicals. 2013;41(10):1721-4. Epub 2013/08/13.

180. Ingelman-Sundberg M, Gomez A. The past, present and future of pharmacoepigenomics. Pharmacogenomics. 2010;11(5):625-7. Epub 2010/04/27.

181. Gladding PA. Clinical applications of pharmacogenetics: present and near future. Cleveland Clinic journal of medicine. 2013;80(8):477-82. Epub 2013/08/03.