The Association of HLA Class II Genetic and Expression Level Variation with Response to the Hepatitis B Vaccine in South African Laboratory Workers

Hadassa Goldfein

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science

Johannesburg, 2017

The financial assistance of the National Research Foundation (DAAD-NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the DAAD-NRF.

DECLARATION

I, **Hadassa Goldfein** (**561370**), am a student registered for the degree of Master of Science in the academic year 2017.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged.
- I have not submitted this work before for any other degree or examination at this or any other University.
- The information used in the Dissertation HAS NOT been obtained by me while employed by, or working under the aegis of, any person or organisation other than the University.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signed  02 day of AUG 2017

ABSTRACT

The hepatitis B virus (HBV) vaccine has contributed greatly to decreasing the HBV epidemic. However, it remains unclear why 5-10% of individuals do not mount an adequate antibody response. Previous studies have shown that genetic variation influences HBV vaccine response. Since such studies are lacking in South African individuals, we examined the associations between HBV vaccine response and genetic variation in *HLA-DPB1*, additional candidate genes and *HLA-DPB1* expression levels in a South African cohort. *HLA-DPA1* and *-DPB1* allele typing was performed using Luminex technology, twenty-four candidate SNPs were typed by MassArray Analysis and *HLA-DPB1* mRNA expression levels were measured by qPCR. *HLA-DPB1*01:01*, **04:01:01G* and **09:01* and SNPs and haplotypes in *IL1B, IL4, IL12B, IFNG* and the *HLA* region were significantly associated with HBV vaccine response. A trend of lower *HLA-DPB1* expression associating with better anti-HBs response was observed, although this was not significant. Response to the HBV vaccine is multi-genic but *HLA-DP* plays an important role.

This dissertation is dedicated to my parents

ACKNOWLEDGEMENTS

I am humbled by the number of people and the institutions who have assisted me over the past two years and who have made this research possible. I would like to express my sincerest gratitude and appreciation here:

- My supervisor, Dr Debbie de Assis Rosa, for her dedicated supervision, guidance and support throughout. For her patience, advice and encouragement – not only when things did not go as planned but also to push myself beyond my 'comfort zone'. I am privileged to have had such a wonderful supervisor.
- My co-supervisor, Dr Melinda Suchard, for her advice, valuable recommendations and thought provoking immunology tutorials and discussions. For training me on the Luminex and Flow Cytometer and for allowing me access to her labs, instruments and skilled staff members.
- The laboratory and department heads who gave us time and access to their staff for recruitment purposes.
- The participants of the study who gave both of their time and their physical selves. Without them, this research would have been impossible.
- The principle study nurse, Sr Nkole Mashilo who went above and beyond her duties in assisting with enrolling the participants into the study.
- Sr Maleshwane Magasa for assisting with the booster vaccinations, blood draws and for allowing us to make use of her rooms at the NICD.
- Dr Villyen Motaze for helping with blood draws when needed.
- Jackie Lubbe for her assistance at the very start of the project, for her organisational skills, problem solving capabilities, for her assistance with the sample processing. For her constant willingness to help and her friendship and company which made those very long days enjoyable.
- Lillian Makhathini for performing the serology testing, and for her patience in explaining the technique and instrument as well as for her general assistance in the lab.
- Professor Robert Veale and Dr Heather Hong for providing control cells lines/DNA for the *HLA* typing experiments.
- Dr Dana Savulescu and Clement Adu-Gyamfi for their assistance with the Luminex instrument.
- Dr Aaron Abera who kindly provided electropherogram images of my MassArray results as well as a detailed MassArray protocol.
- Dr Heather Hong who assisted with cell stimulations, and for her interest, dedication and willingness to help wherever possible.
- Dr Diana Schramm for helping out in times of need - with blood collection tubes, Luminex instrument troubleshooting and locating kits.

- Dr Maria Paximadis for helping me troubleshoot the RNA extractions and generously allowing me to make use of some of her Agilent RNA Nano kit. Also thanks to Gemma Koor for helping me get the RNA extractions optimised, and for the Agilent assay training and assistance.
- Lerato Seakamela, Raffaella Williams and Wayne Howard for training and assisting me with the Applied Bioystems Real-Time PCR system.
- The CVI department as a whole, both staff and students, for their willingness to assist with anything that I had trouble with in the labs.
- The students of GH700 for the interesting discussions and pleasant environment whenever I was there.
- My parents and siblings and other family and friends for their love and support (and distraction) during my studies, and for making the effort, again and again, to try to understand what I do.
- Wits PGMA, DAAD-NRF and PRF for funding.

PRESENTATIONS ARISING FROM THIS STUDY

Poster presentation. Goldfein H, Lubbe J, Makhathini L, Suchard M and de Assis Rosa D. Immunity to hepatitis B virus in a cross-sectional cohort of vaccinated laboratory staff from Johannesburg, South Africa. 5th Conference of the South African Immunology Society. 6- 9 March 2016, Johannesburg.

Oral presentation. Goldfein H, Lubbe J, Makhathini L, Suchard M and de Assis Rosa D. Genetic associations with antibody response to the hepatitis B virus vaccine in South African laboratory workers. Wits molecular biosciences research thrust (MBRT) Postgraduate research day 5 Dec 2016, Johannesburg, (second best talk).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

A	Adenine
A230/A260/A280	Absorbance at 230 nm/ 260 nm/ 280 nm
ACD	Acid Citrate Dextrose
AFND	Allele Frequency Net Database
AIDS	Acquired Immunodeficiency Syndrome
Anti-HBc	Antibodies against HBV core antigen
Anti-HBs	Antibodies against HBV surface antigen
APC	Antigen presenting cell
ASHI	American Society for Histocompatibility and Immunogenetics
BCR	B cell receptor
bp	Base pairs
C	Cytosine
CAT	Clot Activator Tube
CD	Cluster of differentiation
cDNA	Complementary DNA
CEU	Utah residents with Northern and Western European Ancestry
CMJAH	Charlotte Maxeke Johannesburg Academic Hospital
CO_2	Carbon dioxide
Cq	Quantification cycle
CTL	Cytotoxic T lymphocyte
CTLA	Cytotoxic T lymphocyte associated protein
CVI	Centre for Vaccines and Immunology
CXCR	Chemokine (C-X-C motif) receptor
Da	Dalton
DC	Dendritic cell
dCq	Delta quantification cycle
ddCq	Delta delta quantification cycle
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
ddNTP	Dideoxynucleotide triphosphate
dUTP	Deoxyuridine triphosphate
EDTA	Ethylenediaminetetraacetic acid
EPI	Expanded program of immunization
FcR	Fc receptor

FCS	Foetal calf serum
FDC	Follicular dendritic cell
G	Guanine
g	Relative centrifugal force
GWAS	Genome-wide association study
HBcAg	Hepatitis B core antigen
HBeAg	Hepatitis B e antigen
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus
HCW	Health care worker
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
HWE	Hardy-Weinberg equilibrium
ICAM	Intercellular adhesion molecule
ICOS	Inducible costimulatory
IFNG	Interferon gamma gene
IFNα	Interferon-alpha
IFNɣ	Interferon gamma
Ig	Immunoglobulin
IL	Interleukin
IPO8	Importin-8
ITGAL	Integrin alpha-L
ITU	Indian Telugu in the United Kingdom
IU	International units
kg	Kilogram
l or L	Litre
LD	Linkage disequilibrium
LFA	Lymphocyte function-associated antigen
LWK	Luhya in Webuye, Kenya
M	Molar
MAF	Minor allele frequency
MAPK	Mitogen-activated protein kinase
MGB	Minor groove binder
MHC	Major histocompatibility complex
ml	Millilitre
mRNA	Messenger RNA

NFQ	Nonfluorescent quencher
ng	Nanogram
NHLS	National Health Laboratory Service
NICD	National Institute for Communicable Diseases
nm	Nanometre
NRTI	Nucleoside reverse transcriptase inhibitors
OR	Odds ratio
ORF	Open-reading frame
PBMC	Peripheral Blood Mononuclear Cells
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PCV7	Pneumococcal Conjugate Vaccine (heptavalent)
PD	Programmed death
PHA	Phytohemagglutinin
qPCR	Quantitative polymerase chain reaction
rHBsAg	Recombinant hepatitis B surface antigen
RIN	RNA integrity number
RLU	Relative light units
RNA	Ribonucleic acid
SA	South Africa
SAP	Shrimp alkaline phosphatase
SAPE	Streptavidin-conjugated with phycoerythrin
S/CO	Signal to cut-off ratio
SD	Standard deviation
SNP	Single nucleotide polymorphism
SSO	Sequence-specific oligonucleotide
T	Thymine
TBP	TATA-binding protein
TCR	T cell receptor
TFh	Follicular helper T
TGF	Transforming growth factor
Th	Helper T
TNF	Tumour necrosis factor
Treg	Regulatory T
UK	United Kingdom
UNG	Uracil-DNA glycosylase

USA	United States of America
UTR	Untranslated region
UV	Ultraviolet
WHO	World Health Organisation
YRI	Yoruba in Ibadan, Nigeria
°C	Degrees Celsius
µl	Microlitre

CHAPTER 1

Introduction

1.1 Hepatitis B Virus

The hepatitis B virus (HBV) is an enveloped DNA virus of the *Hepadnaviridae* family (Shepard et al., 2006), identified in 1970 (Dane et al., 1970). The virus is 42 nm in diameter, and is comprised of an envelope that surrounds an inner capsid, or core, which contains the viral genome and a polymerase with reverse transcriptase activity (Nassal and Schaller, 1993; Seeger and Mason, 2000; Shepard et al., 2006).

The HBV genome is a partially double-stranded DNA molecule (Locarnini, 2004; Nassal and Schaller, 1993; Seeger and Mason, 2000; Shepard et al., 2006). It is only 3 200 base pairs (bp) in size (varying slightly with genotype) but has a compact organisation: the open-reading frames (ORF) overlap, as do all regulatory and coding regions (Kramvis et al., 2005; Nassal and Schaller, 1993; Shepard et al., 2006). There are four ORFs: one ORF (*C*) encodes the hepatitis B core antigen (HBcAg) and the hepatitis B e antigen (HBeAg), another ORF (*P*) encodes the polymerase, a third ORF (*S*) encodes the envelope proteins and the last ORF (*X*) encodes a transcriptions trans-activator protein (Kramvis et al., 2005; Locarnini, 2004).

HBcAg assembles into the icosahedral capsid particles which is surrounded by the envelope (Seeger and Mason, 2000). The envelope is made up of small (S), middle (M) and large (L) surface proteins (Nassal and Schaller, 1993; Seeger and Mason, 2000). The S surface protein is known as hepatitis B surface antigen (HBsAg) and can also be found as self-assembling, non-infectious envelope particles (Seeger and Mason, 2000; Shepard et al., 2006), which may serve to trap antibodies against the virus, thereby weakening the immune response against HBV (Nassal and Schaller, 1993). The viral polymerase is important for viral replication (Seeger and Mason, 2000). The functions of the X protein and HBeAg are unknown; however, the X protein is thought to function in the establishment of infection *in vivo* and HBeAg possibly influences the level of viral replication and immune responses (Nassal and Schaller, 1993; Seeger and Mason, 2000).

There are ten genotypes of HBV (A-J), whose genomes differ from each other by more than 8% (Kramvis et al., 2005; Lin and Kao, 2013; Locarnini, 2004; Sunbul, 2014). Additionally, some of the genotypes can be further classified into sub-genotypes (Custer et al., 2004; Locarnini, 2004), and based on the variation of HBsAg, serological subtypes can be identified (Kramvis et al., 2005). The major subtypes include *adw*, *adr*, *ayw* and *ayr*, but the identification of additional sub-determinants has increased the number of serological subtypes (Kramvis et al., 2005).

1.2 Hepatitis B Virus Epidemiology

According to the World Health Organisation (WHO), more than two billion people worldwide are, or have been, infected with HBV (WHO, 2009). Geographic regions are classified into three categories of HBV endemicity according to the prevalence of chronic HBV: high (≥8%), intermediate (2-8%), and low (<2%) (Romano et al., 2011). The life risk of exposure to HBV (prior to vaccine introduction) is 60% or greater in high endemic areas, 20-60% in intermediate endemic regions and below 20% in low endemic areas (Romano et al., 2011). HBV is highly endemic to Asia and Africa, especially sub-Saharan Africa (Kew, 1996; Kiire, 1996; Romano et al., 2011).

The HBV genotypes have distinct geographical distributions, and are associated with differences in disease progression, which may contribute to regional differences in endemicity (Lin and Kao, 2013; Sunbul, 2014). Genotype A is common in Northern Europe and sub-Saharan Africa; genotypes B and C are found in Asia and C in Australia; genotype D occurs across Africa, Europe, India and the Mediterranean region; E is restricted to West Africa; G is observed in the United States, France and Germany and F and H is found in Central and South America (Custer et al., 2004; Lin and Kao, 2013; Locarnini, 2004; Sunbul, 2014). Genotypes I and J are newly defined and have been observed in Vietnam and in Japan, respectively (Sunbul, 2014).

Transmission of HBV occurs by percutaneous or mucosal exposure to infected bodily fluids, which can occur either through horizontal or vertical transmission routes (Shepard et al., 2006). The route of transmission may also play a role in the distribution of genotypes, for example genotypes B and C are predominant in Asia where vertical transmission is the major route, whereas other genotypes are found in locations where horizontal transmission is the dominant route (Lin and Kao, 2013). In Africa, HBV is most commonly transferred in early life via child-to-child transmission and in later life by sexual contact (Burnett et al., 2012; Hennig and Hall, 2012; Kew, 1996). Health care workers (HCWs) are also at risk from occupation-related blood or body fluid exposures, particularly in countries endemic for HBV (Burnett et al., 2012). Transmission may also occur indirectly via contaminated items since HBV can remain viable on surfaces for more than seven days (Shepard et al., 2006).

HBV replicates in the liver cells and can result in asymptomatic infection or acute or chronic hepatitis (Shepard et al., 2006; WHO, 2009). Acute HBV infection is detected by a combination of the presence of the HBsAg as well as antibodies against the HBV core and surface antigens,

that is, anti-HBc and anti-HBs, respectively (Shepard et al., 2006). HBsAg is cleared from the blood in individuals who recover from HBV infection (and these individuals will have anti-HBs as well), but HBsAg remains present in chronic carriers as they do not develop anti-HBs (Hennig and Hall, 2012; Shepard et al., 2006). Immunity against HBV after natural infection is marked by the presence of anti-HBs, at a titre of more than 10 IU/L (Shepard et al., 2006; WHO, 2009). Acute HBV infection does not usually require treatment. Chronic HBV infection can be treated with antivirals, such as pegylated interferon-alpha (IFNα), or nucleoside reverse transcriptase inhibitors (NRTIs) such as, entecavir, lamivudine and tenofovir (de Clercq et al., 2010; WHO, 2015). Treatment guidelines for chronic HBV infection were issued by WHO in 2015. HBV infection can, however, be prevented by vaccination (Hennig and Hall, 2012).

It is estimated that 360 million infected individuals are chronic carriers, at high risk of suffering from the sequelae of cirrhosis and hepatocellular carcinoma (WHO, 2009). The rate of chronic carriage (uncleared chronic infection) of HBV worldwide is about 10% in individuals older than five years of age, with a higher prevalence in younger children (Frodsham, 2005; Shepard et al., 2006). In Africa (prior to the introduction of the HBV vaccine), the prevalence of chronic carriage was approximately 15%, but differed between gender and regions, even within a country (Kiire, 1996). Additionally, the carrier rates in rural areas are higher than those in urban areas, and HBV infection is more prevalent in black South Africans than in the Caucasian, Indian and Coloured population groups in South Africa (SA) (Kew, 1996, 2008). In black South Africans, the carrier rate was 9.6% and 76% had been exposed to HBV (Kiire, 1996), whereas in Coloured, Indian and Caucasian South African populations, the prevalence of chronic carriage ranges between 0.2-0.4% and only 3-5% had previous exposure, prior to the introduction of vaccination (Kew, 1996). HBV genotype may contribute to the higher chronic carriage rate observed in Africa compared to the global carriage rate, as genotype A (commonly found in sub-Saharan Africa) is associated with increased rates of chronic HBV infection (Kew, 2008; Sunbul, 2014). The observed differences in carrier rates may also be due to variations in host genetics, as described below.

1.3 Factors Influencing HBV Infection Outcome

Several factors have been found to influence the outcome of HBV infection and clearance or persistence of the virus (Frodsham, 2005). These factors include age at infection, viral load, gender and host genetic factors (Frodsham, 2005). Most published studies on host genetic factors and HBV persistence have associated the major histocompatibility complex (*MHC*) - known as human leukocyte antigen (*HLA*) in humans - particularly class II molecules, with HBV infection outcome (Frodsham, 2005) and details of these findings are summarised below.

A study in a Gambian population found that *HLA-DRB1*13:02* was protective against the development of chronic hepatitis B in adults and children (Thursz et al., 1995), and this was later replicated in a Caucasian (American) and Asian populations (Mbarek et al., 2011; Thio et al., 2003). Yan et al. (2012) conducted a meta-analysis of 2609 cases and 606 controls including various ethnicities, and showed that *HLA-DR*01*, **04* and **13* alleles were significantly associated with HBV clearance in clearance, while *DR*03* and **07* alleles had a significantly increased risk of chronic HBV persistence. *HLA-DQA1*05:01* and *HLA-DQB1*03:01* were associated with chronic infection in African-Americans (Thio et al., 1999). A meta-analysis of HBV infection outcome with *HLA-DQB1* alleles in 815 cases and 731 controls found *HLA-DQB1*03:03* and **06:04* significantly associated with decreased risk of chronic HBV infection, whereas *HLA-DQB1*02:01*, **03:01* and **05:02* associated with increased risk of chronic infection (Huang et al., 2016). A genome-wide association study (GWAS) identified the HLA class II *DP* locus as being strongly associated with HBV infection outcome in Asian populations (Kamatani et al., 2009). The haplotypes including *HLA-DPA1*01:03* and either *HLA-DPB1*04:01* or **04:02* were associated with protection against the development of chronic HBV, while the haplotypes including *HLA-DPA1*02:02* and either *HLA-DPB1*03:01* or *05:01* were associated with increased risk of chronic HBV (Kamatani et al., 2009). In an African- and European-American cohort, *HLA-DPB1*04:01* was also associated with HBV clearance and *HLA-DPB1*01:01* was associated with HBV persistence (Thomas et al., 2012).

A role for *HLA-DP* variants in HBV infection risk, HBV clearance and hepatocellular carcinoma was replicated in several subsequent studies in Asian populations (Guo et al., 2011; Hu et al., 2012; Mbarek et al., 2011; Nishida et al., 2012; Thomas et al., 2012). A meta-analysis of 29 case-control studies involving a total of 62 050 subjects was performed (Yu et al., 2015) and found that the A alleles of *HLA-DPA1* rs3077 and *HLA-DPB1* rs9277535 significantly decreased the risk of HBV infection and increased the possibility of HBV clearance. In African-

and European-Americans, *HLA-DPB1* rs9277534 showed a stronger association with HBV infection outcome than the rs9277535 variant (Thomas et al., 2012). Additionally, Thomas et al. (2012) showed that variants in the *HLA-DP* region that were associated with *HLA-DP* expression levels predicted recovery from HBV infection. More specifically, the rs9277534 GG genotype was associated with HBV persistence and with a higher *HLA-DPB1* expression level (Thomas et al., 2012).

HLA class I alleles have also been associated with HBV infection outcome in Caucasian Americans: *HLA-A*03:01* and *-B*08* have been associated with viral clearance and persistence, respectively (Thio et al., 2003). Various other non-HLA genes have been associated with HBV infection outcome, such as variations in the vitamin D receptor and tumour necrosis factor-alpha (*TNFα*) genes, both of which play a role in immune response (reviewed in Frodsham, 2005).

There are limited studies of host genetic associations with HBV infection outcome in populations from the African continent (Hennig and Hall, 2012). Several have been performed in Gambian cohorts (Bellamy et al., 1999,1998; Frodsham et al., 2006; Thursz et al., 1996, 1995) and a few in the west African countries including Senegal (Dieye et al., 1999; Obami-Itou et al., 2000) and Togo and Benin (Bronowicki et al., 2008). However, no such studies have been performed in South Africans.

1.4 Hepatitis B Vaccine

The HBV vaccine was first licensed in the United States in 1981 and due to its success, it is now part of the routine vaccination schedules for infants in many countries (Shepard et al., 2006). The introduction of immunization at birth has greatly reduced the occurrence of HBV transmission in many countries (WHO, 2009). The HBV vaccine has been included in the South African Expanded Program of Immunisation (EPI) schedule since 1995 (Kew, 2008; Young et al., 2013). Since the introduction of the HBV vaccine in SA, less than 0.5% of vaccinated children from both rural and urban environments, were found to be HBsAg-positive and HBV DNA-positive due to natural infection (Hino et al., 2001; Schoub et al., 2002; Tsebe et al., 2001). Additionally, protective levels of anti-HBs were found in 84 to 87% of the vaccinated cohorts (Hino et al., 2001; Schoub et al., 2002; Tsebe et al., 2001).

The vaccine contains the HBsAg which is the viral protein that elicits an immune response in natural infections (Burnett et al., 2012; Young et al., 2013). The first HBV vaccine consisted of purified HBsAg derived from plasma of chronic carriers of HBV (Szmuness et al., 1981). This vaccine was soon replaced by HBsAg produced in yeast using recombinant technologies (McAleer et al., 1984; Zuckerman, 1996). The HBV vaccine is usually given in a three-dose series (Shepard et al., 2006; WHO, 2009), although as of 2016, in addition to the HBV vaccine doses given at 6, 10 and 14 weeks of age, a booster dose at 18 months was introduced into the South African EPI (Motaze and Suchard, 2016).

1.4.1 Serological markers of HBV immune response

An anti-HBs response of more than 10 IU/L after three vaccine doses is the accepted correlate of protection against HBV infection (Young et al., 2013), with some individuals mounting responses of several thousand IU/L (Tripathy et al., 2011). While an anti-HBs titre > 10 IU/L indicates immunity to HBV, this immunity may have arisen either from natural infection or from vaccination (Krajden et al., 2005; Shepard et al., 2006). The source of immunity to HBV can be determined by the presence or absence of anti-HBc (Krajden et al., 2005) since the core protein of the virus is not included in the vaccine preparation (Szmuness et al., 1981). An anti-HBc negative result indicates no previous infection with HBV and therefore immunity because of vaccination, whereas an anti-HBc positive result indicates infection with HBV which may be resolved infection (an absence of HBsAg) or ongoing infection (a presence of HBsAg) (Krajden et al., 2005).

1.4.2 HBV vaccine responses

A decline in anti-HBs level over time following infection or immunisation occurs in some cases (Shepard et al., 2006; Zuckerman, 1996) and is associated with an increasing risk of infection (Jack et al., 1999). However, previously vaccinated individuals whose antibody titres decrease – even to undetectable levels – years after vaccination are generally still protected from infections upon subsequent exposures (Shepard et al., 2006; WHO, 2009) due to immunological memory (Zuckerman, 1996). Immune memory to HBsAg is achieved by both memory T and B cells, however, their detection is difficult as a result of their low frequency, especially in the peripheral blood (Bauer and Jilg, 2006). These cells represent a second correlate of protection against HBV. In a small study of successful vaccinees whose antibody levels had declined to undetectable levels over 4-8 years, the number of HBsAg-specific memory T and B cells was found to be approximately 1-2 and 2-4 cells per one million of peripheral blood lymphocytes respectively (Bauer and Jilg, 2006). Despite the low frequencies of T and B memory cells, re-vaccination of the subjects resulted in good anti-HBs responses with titres ranging between 110 and 13500 IU/L after one month (Bauer and Jilg, 2006).

1.5 Vaccine Efficacy

Vaccines in general have saved millions of lives in the past century and are among the most cost-effective health interventions available (Ehreth, 2003). Vaccination has resulted in the eradication of smallpox, the near eradication of polio and the control of other diseases such as measles and tetanus (Ehreth, 2003). However, the problems encountered in designing successful vaccines against certain diseases, such as human immunodeficiency virus (HIV), malaria and hepatitis C, as well as the less than 100% efficacy of most vaccines (Grimm and Ackerman, 2013) highlight the gaps in our knowledge of how vaccines work. The immunological mechanisms underlying vaccine response must be fully elucidated for the rational design and development of effective novel and improved vaccines.

There are several terms, including immunogenicity, efficacy and effectiveness, used to differentiate measures of how well a vaccine works. Immunogenicity is defined as the ability of the vaccine to induce an immune response in a vaccinated individual (Banaszkiewicz and Radzikowski, 2013). Efficacy of a vaccine is the percentage reduction of disease in a vaccinated group of people compared to an unvaccinated group, measured in a randomized, placebo-controlled study (Banaszkiewicz and Radzikowski, 2013) and since it is measured under controlled conditions, it proves the best possible scenario (Weinberg and Szilagyi, 2010). Effectiveness is measured from observational studies of protection against disease after vaccine use in the general population (Banaszkiewicz and Radzikowski, 2013), thus providing a more realistic measure of how well the vaccine reduces the disease in the population (Weinberg and Szilagyi, 2010).

Additionally, protection against infection and/or disease is ideally associated with a ‘correlate of immunity’, a quantifiable indication of immunity that can be directly measured (Plotkin, 2008). For many vaccines, antibody quantity is the predominant correlate of immunity, although in some cases correlates of protection may be T cell responses or quantity of memory T or B cells (Plotkin, 2010, 2008). Correlates of immunity may be absolute quantities, where a particular level of response is highly correlated with protection (Plotkin, 2008). However, the case is often more complex as many correlates of immunity are relative, meaning that while most infections are prevented at a certain level of response, some will occur above that level (Plotkin, 2008). Furthermore, there may be no identifiable or easily measurable correlate of immunity in which case a surrogate measure, which is not in itself protective and is only indirectly related to the true correlate of protection, must be used as a substitute (Plotkin, 2008).

There can also be more than one correlate of immunity since vaccine-induced protection may be achieved by multiple mechanisms (Plotkin, 2010). These factors, among others, illustrate that certain mechanisms of action of vaccines may be unknown, while other aspects may be well described for the same vaccine and can also make it difficult to assess the success of a vaccine.

1.6 HBV Vaccine Efficacy

Despite the success of the HBV vaccine, studies have shown that a percentage of healthy subjects do not produce a sufficient antibody response (Zuckerman, 1996). HBV vaccine non-response is characterized by an antibody titre of less than 10 IU/L (Zuckerman, 1996). Non-responders continue to be susceptible to HBV infection (Zuckerman, 1996). Globally, approximately 5 to 10% of healthy individuals fail to mount an antibody response to the HBV vaccine (Zuckerman, 1996). In SA, although only a few studies have been performed in this population, a somewhat higher percentage of individuals do not respond to the HBV vaccine. The percentage of non-responders in SA ranges from 7% (Mphahlele et al., 2002) to 20% (Young et al., 2013), with an average of 14% (calculated from data from Hino et al., 2001; Mphahlele et al., 2002; Schoub et al., 2002; Simani et al., 2009; Tsebe et al., 2001; Young et al., 2013). Most of these studies were performed in children under three years of age, and only Tsebe et al. (2001) noted an association of younger age and more recent vaccination with higher anti-HBs levels. The relatively high non-response rate of 20% was observed in women between the ages of 18 and 49 years (Young et al., 2013), and may be as a result of the older age of the subjects, as older age is a clinical factor of non-response to the HBV vaccine (Zuckerman, 2006); however, the high non-response rate is likely due to the anti-HBs titres being measured after only two vaccine doses (Young et al., 2013).

There is a direct correlation between HBV vaccine efficacy and anti-HBs titre (Jack et al., 1999) and hence, antibody titre is the standardly used and easily implemented correlate of protection for HBV vaccine studies (Plotkin, 2008). T cell correlates of protection for HBV vaccine have also been identified (Carollo et al., 2013) but are less easy to implement as a standardised test (Suchard, 2012). Due to these well-established correlations of protection, responders and non-responders to HBV vaccination can be clearly defined. Therefore, HBV vaccination is an excellent model by which to study the mechanisms resulting in vaccine response versus non-response. This, in turn, could offer insight into mechanisms of immune response which may assist in the rational development of effective vaccines, not only against HBV, but also against other infectious diseases for which no vaccines are yet available.

In order to understand mechanisms of vaccine response, a thorough understanding of the immunology behind the vaccine-induced immune response must be obtained. In the case of HBV, the immune response towards HBsAg provided in vaccines is thought to be very similar

to the immune response to natural HBV infection. The following sections provide a brief review of the immune response to HBV vaccine.

1.7 Adaptive Immune Response To HBsAg

HBsAg (whether from natural infection or from vaccine), like many other protein antigens induces both a T cell and a B cell response (Cupps et al., 1984; Milich and Leroux-Roels, 2003).

1.7.1 T cell response

Antigen presenting cells (APCs), such as macrophages, dendritic cells (DCs), and B lymphocytes, internalize pathogens or antigens (such as HBV or HBsAg) and process them into peptides that can be attached to HLA molecules and presented on the cell surface (Milich and Leroux-Roels, 2003). There are two classes of HLA molecules, class I and II. Class I molecules generally present intracellular antigens (such as viral peptides) and activate $CD8^+$ cytotoxic T lymphocytes (CTLs) to kill the infected cell, and class II molecules usually present extracellular antigens and activate $CD4^+$ helper T (Th) cells which produce cytokines that induce cellular and humoral responses (Bonilla and Oettgen, 2010).

The class I and class II HLA molecules are the most polymorphic human proteins (Janeway et al., 2001; The MHC sequencing consortium, 1999). This highly polymorphic nature of the HLA molecules is thought to be driven by evolutionary pressure for the immune system to identify constantly evolving pathogens (The MHC sequencing consortium, 1999). The *HLA* genes are located on chromosome 6 in one of the most gene-dense regions of the human genome (Janeway et al., 2001; The MHC sequencing consortium, 1999). The *HLA* class I and class II genes are also polygenic with three class I *HLA* genes (*HLA-A*, *HLA-B* and *HLA-C*) encoding the α chain of HLA class I molecules and three pairs of *HLA* class II genes (*HLA-DR*, *HLA-DQ* and *HLA-DP*) encoding the α and β chains of class II molecules (Janeway et al., 2001). The presence of several genes encoding HLA molecules allows an individual to present a broad variety of antigenic peptides (Janeway et al., 2001).

There are several subsets of T cells, the largest subsets include $CD4^+$ Th cells, $CD8^+$ CTLs which are activated via the classical HLA class I and II molecules (Bonilla and Oettgen, 2010). A small T cell subset known as gamma delta T cells, as well as the natural killer T (NKT) cell subset do not recognise the the HLA class I and II molecules but rather recognise the nonclassical CD1 family of molecules (Bonilla and Oettgen, 2010). T cells can be furhter defined by a variety of cell surface markers indicative of exposure to antigen, state of mutation,

state of activation, memory, homing preference or other functional properties (Broere et al., 2011).

$CD4^+$ and $CD8^+$ naïve T cells are activated by two major signals which are received via their interactions with peptides presented via HLA on the surface of APCs. The first signal is achieved through the binding of the T cell receptor (TCR) to the HLA-peptide complex present on the APC cell surface (Bromley et al., 2001). In $CD8^+$ T cells, this signal involves HLA class I molecules which are expressed on the surface of any nucleated cell, and the interaction is stabilized by CD8 (Bonilla and Oettgen, 2010; Broere et al., 2011; Bromley et al., 2001). In the case of $CD4^+$ T cells, the interaction involves HLA class II molecules, expressed exclusively on APCs, and is stabilized by CD4 (Bonilla and Oettgen, 2010; Broere et al., 2011; Bromley et al., 2001).

The second signal for T cell activation is a costimulatory signal (Bromley et al., 2001). There are many costimulatory signalling/receptor pairs involved at the T cell/APC immune synapse as seen in **Figure 1.1**. An example of costimulatory molecules involved in CTL activation is Fas ligand on the $CD8^+$ T cell binding to Fas on the surface of the target cell (Bonilla and Oettgen, 2010; Broere et al., 2011). The classical example involved in Th cell activation is CD28 on the $CD4^+$ T cell interacting with B7 (CD80/CD86) on the APC (Bromley et al., 2001).

Upon activation, T cells proliferate rapidly in a clonal fashion with each clone having the same TCR and antigen specificity (Milich and Leroux-Roels, 2003). Some of these develop into memory T cells . Clonal CTLs kill infected cells either by the cytolytic granules granzymes and perforin, or through triggering Fas/Fas ligand-mediated apoptosis and also produce various cytolines (Bonilla and Oettgen, 2010; Broere et al., 2011). Th cells produce cytokines that assist in enhancing humoral and cellular immune responses (Bonilla and Oettgen, 2010).

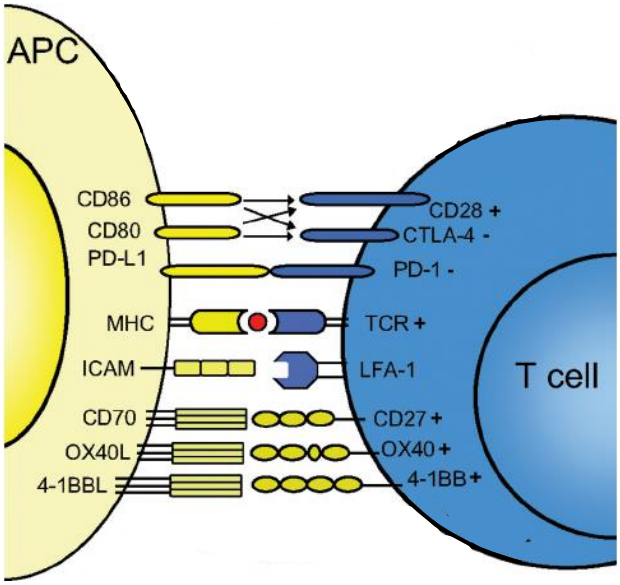


Figure 1.1. Interactions made at the immunological synapse between the T cell and the interacting APC. CD, cluster of differentiation; CTLA, cytotoxic T lymphocyte associated protein; ICAM, intercellular adhesion molecule; LFA, lymphocyte function-associated antigen; MHC, major histocompatibility complex; PD(-L), programmed death (-ligand); TCR, T cell receptor. Adapted from den Haan et al., 2014.

Activated $CD4^+$ T cells can be further differentiated into Th1, Th2, Th17, follicular helper T (TFh) and regulatory T (Treg) cell subsets according to the cytokines they produce (Broere et al., 2011). Th1 cell subset characteristically produces interleukin (IL)2, interferon gamma (IFNɣ) and TNFα, while Th2 cells produce IL4, IL5 and IL13 (Abbas et al., 1996; Bonilla and Oettgen, 2010; Broere et al., 2011). Cytokines which assist development of Th1 cells include IFNɣ and IL12 and those that support Th2 development include IL4 (Abbas et al., 1996; Broere et al., 2011). Th1 cells are predominantly involved in enhancing cellular immune responses whereas Th2 cells are necessary for antibody production by B cells (Broere et al., 2011), although this is an oversimplified categorisation since Th-type cytokines can be involved in humoral immune responses and *vice versa* (Abbas et al., 1996). For example, IFNɣ is involved in macrophage activation but also in the production of antibodies involved in opsonization and phagocytosis (Abbas et al., 1996). The differentiation of Th cells into these subsets are influenced by the antigen and by exposure to particular cytokines (Broere et al., 2011) and may also be influenced by HLA allele (Jafarzadeh and Shokri, 2012; Mangalam et al., 2013) and/or its expression levels (Thomas et al., 2012).

The Th17 cell subset is a more recently described group of Th cells that produce IL17 and IL22 (Bonilla and Oettgen, 2010; Broere et al., 2011). It is the first subset produced upon infection

and plays a role in pro-inflammatory actions (Broere et al., 2011). TFh cells are important in B cell differentiation and the in the generation of memory B cells (Crotty, 2011). Conversely, Treg cells are involved in down-regulation of immune responses (Broere et al., 2011). Treg cells produce anti-inflammatory cytokines such as IL10 and Transforming growth factor beta (TGFβ) (Bonilla and Oettgen, 2010; Broere et al., 2011). Treg also form a small set of $CD8^+$ T cells (Broere et al., 2011).

1.7.2 B cell response (T-dependent)

In a manner analogous to T cell activation, B cells are activated after receiving two major signals (Bonilla and Oettgen, 2010). The first signal involves the cross-linking of B cell receptors (BCRs). BCR cross-linking is caused by the interaction of BCRs with whole native antigen in its 3-dimensional conformation (Bonilla and Oettgen, 2010). Cross-linking of BCRs results in the activation of intracellular signalling pathways (Bonilla and Oettgen, 2010). The first signal results in B cells being able to interact with activated T cells which will provide the second signal (Bonilla and Oettgen, 2010).

B cells themselves are APCs which present peptides on HLA class II molecules after internalizing and processing antigen recognised by the BCR (Bonilla and Oettgen, 2010). The antigen presenting B cell can then interact with a Th cell specific for the HLA class II-peptide complex (Bonilla and Oettgen, 2010). The HLA class II-peptide complex expressed on the surface of the B cell is recognized by the TCR of the interacting T cell, which was previously activated by interaction with an APC (Milich and Leroux-Roels, 2003). This interaction between TCR and HLA-peptide complex is stabilized by the CD4 molecule on the T cell surface (Broere et al., 2011). Additionally, costimulatory signals are generated through the interaction of signal/receptor pairs on the B and T cell, such as the CD40-CD40 ligand interaction (**Figure 1.2**) (Milich and Leroux-Roels, 2003). Both cells involved in this B cell/Th cell immunological synapse are activated: the Th cell secretes Th2-type cytokines (such as IL4, IL5 and IL13) and these cytokines, together with the signals described above, result in the differentiation of the B cell into antibody-secreting plasma cells or memory cells (Bonilla and Oettgen, 2010; Broere et al., 2011).

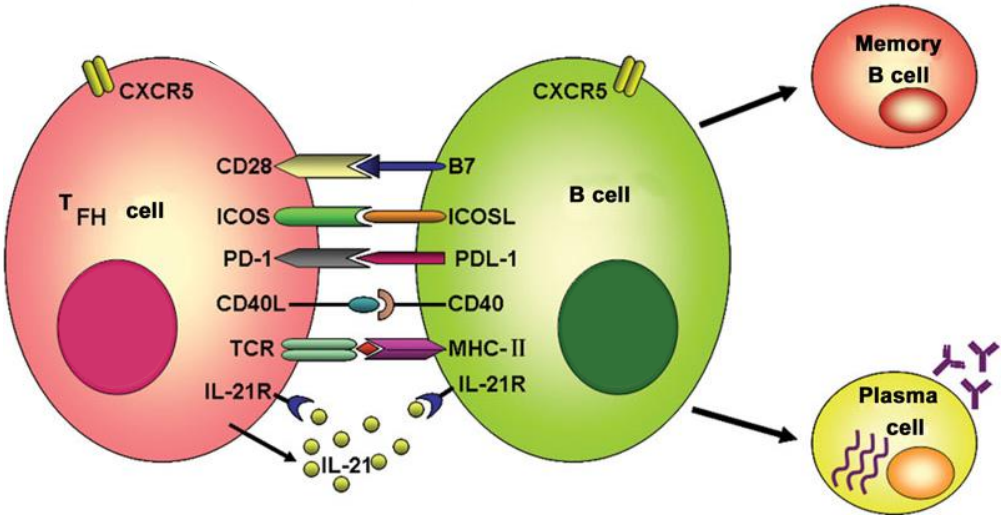


Figure 1.2. Costimulatory pairs in B cell/T cell interactions. A follicular T helper (T_{FH}) cell interacting with a B cell to induce the production of memory B cells or antibody-producing plasma cells. CD, cluster of differentiation; CD40L, CD40 ligand; CXCR5, chemokine C-X-C motif receptor 5; ICOS(L), inducible costimulatory (ligand); IL-21(R), interleukin-21 (receptor); MHC, major histocompatibility complex; PD(L)-1, programmed death (ligand)-1; TCR, T cell receptor. Adapted from Shekhar and Yang, 2012.

Additionally, the humoral response is assisted by certain proteins of the complement system (Milich and Leroux-Roels, 2003). The generation of memory B cells which occurs in germinal centres is reliant on contact with antigens held by follicular dendritic cells (FDCs) (Milich and Leroux-Roels, 2003). FDCs retain antigens that are coated with the complement components C3d and C4d (Milich and Leroux-Roels, 2003). However, FDCs do not process and present antigens via MHC; they present intact antigen-antibody complexes on their cell surface using Fc receptors (FcR) or complement receptors CD21 or CD35 (Milich, 2002).

In responders to HBV vaccine, both memory T and B cells are produced which provide protection against HBV infection upon subsequent exposure to the virus (Hennig et al., 2008).

While immune response is vital in protection against disease, its regulation is equally important. B and T lymphocytes can be kept in check by the inhibition of costimulatory signals, in which case the cells become anergic (unable to respond to antigen) (Wang and Zheng, 2013). However, more commonly described is the mechanism of regulation through the production of regulatory cytokines by Treg cells, including both $CD4^+$ and $CD8^+$ Treg cells (Broere et al., 2011; Wang and Zheng, 2013). Treg cells are mainly involved in the regulation of T cells, but they have been found to regulate other cell types as well, including B cells (Wang and Zheng, 2013). For example, the regulatory cytokine IL10 is capable of suppressing both cellular (via

Th1 cells) and humoral (via Th2 cells) immune responses by reducing production of IL2, TNFα and IL5 (Broere et al., 2011). IL10 is also able to reduce T cell activation indirectly by down-regulating APC expression of HLA and costimulatory molecules (Broere et al., 2011). In addition to suppressing B cells by cytokine production, Treg cells can release granzyme B and perforin to kill B cells (Wang and Zheng, 2013). Moreover, activated B cells can also produce cytokines, including IL4, IL 6, IL 10, IL23 and TNFα, which add to the regulation of immune responses (Wang and Zheng, 2013).

As can be seen from the above simplified explanation of the immune responses to an antigen (such as HBsAg), there are many interacting pathways involving numerous molecules. Natural genetic variation in the genes encoding these molecules can cause changes in the immune pathways described. This, in turn, may result in non-responsiveness to HBV vaccination. From this we can understand the hypotheses put forward to explain the mechanisms underlying non-response.

1.8 Factors Causing Non-Response To HBV Vaccine

1.8.1 Clinical factors

Several clinical factors have been associated with poor antibody response to HBV vaccination. These factors include the site and route of injection, gender, age, body mass, cigarette smoking and immunosuppression or immunodeficiency (Averhoff et al., 1998; Chathuranga et al., 2013; Cleveland et al., 1994; Fisman et al., 2002; Ingardia et al., 1999; Milner and Beck, 2012; Roome et al., 1993; Shaw et al., 1989; Tohme et al., 2011; Young et al., 2013; Zuckerman, 1996). In a Gambian twin study, a relatively high heritability was observed for anti-HBs response (Newport et al., 2004), suggesting a genetic mechanism for non-response to HBV vaccine. However, genetic and/or cellular mechanisms underlying these associations with lack of response to vaccination remain unclear, as discussed below.

1.8.2 Immunogenetic hypotheses investigated to explain non-response

While numerous studies have investigated immunogenetic causes of non-response to the HBV vaccine in Caucasian and Asian populations (such as Höhler et al., 2002; Kardar et al., 2002; Li et al., 2013), few have been performed in African populations (Hennig et al., 2008; Newport et al., 2005; Ryckman et al., 2010), and none in South African individuals.

Defects in the genes encoding any molecule involved in any step along the immune pathways described above could potentially cause a lack of antibody production in response to HBV vaccination. Appropriate genes can be studied using a candidate-specific approach or using a genome-wide approach. The candidate gene approach involves analysing a gene/mutation that may play a role in the phenotype of interest, either due to its function or due to other evidence that may implicate it (Hirschhorn and Daly, 2005). Genome-wide association studies (GWAS) involve surveying the whole genome without any prior hypothesis as to the causal candidates in order to identify genetic variants commonly associated with a phenotype (Hirschhorn and Daly, 2005). To date the candidate gene approach has been used more often than GWAS for identifying gene variants associated with non-responsiveness to the HBV vaccine, possible due to the expense and large sample sizes required by GWAS.

The literature review below summarises findings from GWAS, as well as from candidate approaches which studied associations with, and mechanisms for, non-response to HBV vaccination.

1.8.2.1 Identifying genetic associations with HBV vaccine response

There is value in examining genetic associations in multiple populations or ethnic groups. Rates of disease and of other phenotypes, such as adverse response to treatment, vary between populations due to variation in both genetic and non-genetic risk factors (Risch et al., 2002; Rosenberg et al., 2002). Specific alleles or allele frequencies can differ across populations and a variant might only occur, or be common, in certain populations (Adeyemo and Rotimi, 2010; Keita et al., 2004; Myles et al., 2008; Rosenberg et al., 2010). Furthermore, the effect sizes of risk variants may differ across populations, so that the underlying determinants of the same disease may vary between populations (Rosenberg et al., 2010; Tang, 2006). Additionally, differences in linkage disequilibrium (LD) patterns across populations means that finely localizing the risk variant(s) will be more probable in populations with lower LD (Rosenberg et al., 2010; Teo et al., 2010). As considerable genetic variation may be specific to regions or populations, there is an interest in the potential role of the geographic distribution of genetic variation in contributing to phenotypic differences between groups (Keita et al., 2004; Mountain and Risch, 2004).

As mentioned, genome-wide approaches aim to identify genetic variants associated with the phenotype of interest without any *a priori* assumptions or candidate genes in mind. While several GWAS have been performed looking for genetic associations with HBV infection outcomes (Al-Qahtani et al., 2013; Hu et al., 2013; Kim et al., 2013b; Mbarek et al., 2011; Nishida et al., 2012), relatively few GWAS have been performed looking for genetic associations with HBV vaccine response. One GWAS performed in an Indonesian population, with an initial scan of over 400 000 single nucleotide polymorphisms (SNPs), found multiple significant genetic determinants within the *HLA* region (Png et al., 2011), including rs3135363 in *HLA-DR*, rs9277535 in *HLA-DP* and rs9267665 in the *HLA* class III region. Pan et al. (2013) performed a GWAS in a Han Chinese cohort and found that the strongest associations HBV vaccine non-response were 21 variants in the *HLA* region on chromosome 6. These 21 SNPs could be tagged by rs477515, rs28366298, rs3763316 and rs13204672, but the strongest association was rs477515 (upstream of *HLA-DRB1*) which reached genome wide significance even after Bonferroni correction (Pan et al., 2013). A recent GWAS performed in Taiwan found many significant associations in the *HLA-DP* locus, with rs7770370, an *HLA-DPB1* intronic variant, reaching genome-wide significance (Wu et al., 2015).

Another study investigated 715 SNPs in 133 candidate genes that specifically excluded the MHC region and found that variations in mitogen-activated protein kinase 8 (*MAPK8*), as well

as in the cytokine genes interferon gamma (*IFNG*) and *IL10RA*, affected peak anti-HBs levels in the studied Gambian population (Hennig et al., 2008). While MAPK, a kinase, is not an obvious player in immune response pathways, it is a key player in the signal transduction pathways that result in the consequent response after receiving a signal (Hennig et al., 2008). This highlights that molecules involved in these transduction pathways cannot be ignored when attempting to identify causative molecules of HBsAg non-responsiveness.

1.8.2.2 Defects related to antigen presentation in association with HBV vaccine response

One major hypothesis suggested to explain non-responsiveness to the HBV vaccine is defective antigen presentation which results in either an inefficient T cell response or an inefficient B cell response. Defective antigen presentation could involve HLA-mediated antigen presentation (during APC/T or T/B cell interactions), or HLA independent antigen presentation (during B/FDC interaction) as discussed in section 1.7.2 above. Several aspects of antigen presentation have been examined in the literature in association with HBV vaccine response, including *HLA* variation, binding of HBV antigen to *HLA* alleles, variation in secondary signalling molecules, variation in complement influencing antigen presentation and TCR variation. Results of these studies are reviewed below.

HLA genes are the most polymorphic genes involved in T cell/APC interactions and influence the outcomes of many viral and bacterial infections and vaccines (reviewed in Blackwell et al., 2009; Yucesoy et al., 2013) therefore, they are obvious candidate genes to be examined. Many significant associations between particular *HLA* alleles and non-response to HBV vaccination have been detected (Hennig and Hall, 2012). *HLA* class II alleles and their association with antibody response to the HBV vaccine have been studied extensively in Caucasian and Asian subjects (Amirzargar et al., 2008; Caillat-Zucman et al., 1998; Craven et al., 1986; Desombere et al., 1998; Höhler et al., 2002b; Kimman et al., 2007; Langö-Warensjö et al., 1998; Martinetti et al., 2000; Milich and Leroux-Roels, 2003; Wang et al., 2004; Wu et al., 2014, 2013).

In order to improve the statistical power of these studies and to summarise published data, Li et al. (2013) performed a meta-analysis involving fifteen studies and a total of 2 308 Caucasian and Asian subjects. The meta-analysis showed the following significant associations between *HLA* variants and HBV vaccine response: *HLA-DRB1*01*, *-DRB1*13:01*, *-DRB1*15*, *HLA-DQB1*05* (**05:01*) and *-DQB1*06* (**06:02*) were associated with a good (>10 IU/L) antibody response. *HLA* class II alleles associated with non-response (<10 IU/L) to the HBV vaccine

were *HLA-DRB1*03* (**03:01*), *-DRB1*04*, *-DRB1*07* and *-DRB1*13:02* and *-DQB1*02* (Li et al., 2013). Relatively few studies have studied *HLA-DP* variation, but *HLA-DPA1*01:03* and *HLA-DPB1*02:02*, **04:01* and **13:01* have been associated with good response, although **04:01* is the most consistent association (Desombere et al., 1998; Martinetti et al., 1995; Mineta et al., 1996; Wu et al., 2015, 2014, 2013) and HLA-*DPB1*05:01*, **09:01* and **11:01* have been associated with non-responsiveness in Asian and Caucasian populations (Desombere et al., 1998; Wataya et al., 2001).

Most of these risk and protective alleles associated with HBV vaccine response are similar to those identified in association with HBV infection outcome. Only *HLA-DRB1*13:02* showed contrary effects – a risk allele for HBV vaccine non-response, but protective against chronic HBV infection. Additionally, there were several alleles associated with HBV vaccine response that were not reported as being associated with HBV infection outcome and *vice versa*.

HLA class I associations have not been studied as extensively as class II in the context of HBV vaccine response because exogenous protein antigen, such as the HBsAg, is expected to be presented by HLA class II and not by HLA class I. *HLA-A*26:02* (Mineta et al., 1996) and *HLA-Cw*03* (Wang et al., 2004) are associated with low antibody production. Barnaba et al. (1990) found that if antigen cross presentation occurs in B cells (that is, antigen is displayed via B cell HLA class I instead of on HLA class II as usually expected), then the B cell may be targeted for antigen-specific CTL-mediated lysis and lower antibody production may occur. Factors driving HLA cross presentation are unclear.

The above describes associations found between single *HLA* alleles and responsiveness, however, studies involving associations with the genes of the *MHC* region are confounded by the high levels of linkage disequilibrium (LD) between the *HLA* loci (Paximadis et al., 2012). As a result of this LD, it is often difficult to identify the precise causative variation involved, and therefore, haplotypes are described in association with responsiveness to HBsAg. For example, of the *HLA* class II alleles associated with non-response described above, *HLA* allele *HLA-DQB1*06:02* is in LD with *HLA-DRB1*15:01* and *HLA-DQB1*02:01*, *-DRB1*07:01* and *-DRB1*03:01* are in strong LD with each other (Wang et al., 2004; Zuckerman, 1996). Other haplotypes associated with non-response to the HBV vaccine are haplotypes *HLA-DRB1*07-DQB1*02*, *HLA-DRB1*07-DQB1*03* (Wang et al., 2004) and *HLA-B18-DRB1*03:01-DQB1*02:01* (Milich and Leroux-Roels, 2003). Haplotypes associated with good response to HBV vaccination include *HLA-B*15-Cw*03*, *HLA-DRB1*15-DQB1*06*

(Wang et al., 2004), *HLA-DQB1*06:02-DQA1*01:02-DR15* and *HLA-DQB1*06:03-DQA1*01:03-DR13:01* (Milich and Leroux-Roels, 2003). Haplotypes including the *HLA-DPB1* locus have not been examined.

HLA allele frequency distributions differ between different ethnicities (Paximadis et al., 2012; Sidney et al., 2010) and different *HLA* alleles could potentially mediate different immune responses in different populations. Few studies have been conducted in African populations, however, the association observed between *HLA-DRB1*07:01* and *HLA-DQB1*02* and non-responsiveness and between *HLA-DRB1*15* and *HLA-DQB1*06* and good response to the HBV vaccine were confirmed in one study in African-American subjects (Wang et al., 2004). Most of the class I and class II *HLA-DR* and *-DQ* alleles associated with response to HBV vaccination are present at reasonable frequencies in South African black populations (Lombard et al., 2006; Paximadis et al., 2012; Pirie et al., 2001), as shown in **Table 1.1**. Only *HLA-A*26:02* and *-B*18* are not common in a black South African population as they are found at a frequency of less than 5%. However, of the *HLA-DP* alleles associated with vaccine response, only *HLA-DPB1*04:01* is found at a frequency greater than 5% in black South African populations (Lombard et al., 2006).

Table 1.1. *HLA* alleles associated with HBV vaccine response and their frequency in black South African populations

HLA allele	Association with HBV vaccine response*	Reference for association with HBV vaccine response	Frequency (%)	Reference for frequency
*A*26:02*	Low	Mineta et al., 1996	0	Paximadis et al., 2012
*B*15*	High	Wang et al., 2004	12.9	Paximadis et al., 2012
*B*18*	Low	Milich and Leroux-Roels, 2003	5.1	Paximadis et al., 2012
*Cw*03*	Low	Wang et al., 2004	6.8	Paximadis et al., 2012
*DRB1*01*	High	Caillat-Zucman et al., 1998; Desombere et al., 1998; Höhler et al., 2002b	7.2	Pirie et al., 2001; Lombard et al., 2006; Paximadis et al., 2012
*DRB1*03 (03:01)*	Low	Amirzargar et al., 2008; Caillat-Zucman et al., 1998; Craven et al., 1986; Martinetti et al., 2000	32.4 (13.7)	Pirie et al., 2001; Lombard et al., 2006; Paximadis et al., 2012
*DRB1*04*	Low	Amirzargar et al., 2008	10.5	Pirie et al., 2001; Lombard et al., 2006; Paximadis et al., 2012
*DRB1*07*	Low	Amirzargar et al., 2008; Craven et al., 1986; Desombere et al., 1998; Martinetti et al., 2000; Wang et al., 2004	12.6	Pirie et al., 2001; Lombard et al., 2006; Paximadis et al., 2012
*DRB1*13:01*	High	Langö-Warensjö et al., 1998	10.4	Lombard et al., 2006; Paximadis et al., 2012
*DRB1*13:02*	Low	Langö-Warensjö et al., 1998	9.2	Lombard et al., 2006; Paximadis et al., 2012
*DRB1*15*	High	Caillat-Zucman et al., 1998; Höhler et al., 2002b; Langö-Warensjö et al., 1998; Wang et al., 2004	16.2	Pirie et al., 2001; Lombard et al., 2006; Paximadis et al., 2012
*DQA1*01:02*	High	Langö-Warensjö et al., 1998	41.1	Pirie et al., 2001; Lombard et al., 2006
*DQA1*01:03*	High	Langö-Warensjö et al., 1998	12.45	Pirie et al., 2001; Lombard et al., 2006
*DQB1*02 (02:01)*	Low	Amirzargar et al., 2008; Desombere et al., 1998; Martinetti et al., 2000; Wang et al., 2004	31.3 (16.1)	Pirie et al., 2001; Lombard et al., 2006
*DQB1*03*	Low	Martinetti et al., 2000	30.3	Pirie et al., 2001; Lombard et al., 2006
*DQB1*05 (05:01)*	High	Desombere et al., 1998	24.3 (20.5)	Pirie et al., 2001; Lombard et al., 2006
*DQB1*06 (06:02)*	High	Langö-Warensjö et al., 1998; Wang et al., 2004	58.5 (28.9)	Pirie et al., 2001; Lombard et al., 2006
*DQB1*06:03*	High	Langö-Warensjö et al., 1998; Wang et al., 2004	8.5	Pirie et al., 2001; Lombard et al., 2006
*DPB1*02:02*	High	Mineta et al., 1996; Wu et al., 2013	1.3	Lombard et al., 2006
*DPB1*04:01*	High	Desombere et al., 1998; Mineta et al., 1996; Wu et al., 2015, 2013	15.5	Lombard et al., 2006
*DPB1*05:01*	Low	Wu et al., 2015, 2014, 2013	0.6	Lombard et al., 2006
*DPB1*09:01*	Low	Wu et al., 2013	0.0	Lombard et al., 2006
*DPB1*11:01*	Low	Desombere et al., 1998	2.8	Lombard et al., 2006
*DPB1*13:01*	High	Martinetti et al., 1995; Mineta et al., 1996	2.3	Lombard et al., 2006

*High antibody response characterised by antibody titre > 10 IU/L
Low antibody response characterised by antibody titre < 10 IU/L

The common association observed between *HLA* alleles and responsiveness to vaccines would seem to imply that lack of response is due to defective antigen presentation, that is, that certain HLA alleles do not bind HBsAg peptides for T cell recognition. However, Godkin et al. (2005) studied the affinity of HBsAg-derived peptide binding to selected HLA alleles and found that *HLA* alleles previously associated with non-responsiveness often show normal binding capabilities. In addition, it has been shown that APCs with HLA class II molecules associated with poor/non-response used in *in vitro* studies, were able to bind and present HBsAg peptides to T cells and stimulate appropriate T cell responses such as proliferation (Desombere et al., 1995).

Activation of the T and B cells may be affected by variation in secondary signalling molecules such as those shown in **Figures 1.1** and **1.2**. Therefore, response to HBV vaccination may also be influenced by variation in secondary signalling molecules, such as genetic mutations found in the candidate genes *CD58* (Ryckman et al., 2010) and Integrin alpha-L (*ITGAL*) which were identified in a GWAS (Hennig et al., 2008). CD58, or Lymphocyte function-associated antigen (LFA)-3, is a molecule present on APCs that has previously been associated with T cell proliferation and/or differentiation in response to a poxvirus vaccine in mice, and it may be responsible for such a response after stimulation by other vaccines including the HBV vaccine (Ryckman et al., 2010). As a result, variations in the *CD58* gene may result in a defective T cell response; this would explain the association observed between variant alleles of *CD58* and lower antibody production (Ryckman et al., 2010). ITGAL, or LFA-1, is a protein with a central role in immune cell-cell interaction; thus, variation in the gene encoding ITGAL may detrimentally affect the T and B cell interaction, thereby leading to a decreased production of antibodies (Hennig et al., 2008). A study of genetic variation in some other secondary signalling molecules, specifically B7, suggested no role for variation at this locus in HBV vaccine response (Desombere et al., 2005). Reduced expression of the CD40 ligand was observed on Th cells in non-responders to the HBV vaccine (Goncalves et al., 2004). The CD40 ligand-CD40 interaction is important in the activation of B cells and the switching of immunoglobulin (Ig) classes (Goncalves et al., 2004). Molecules present on T cells such as CD25 and CD69 that are involved in T cell activation, have also been found to be expressed at lower levels in HBV vaccine non-responders compared to responders (Goncalves et al., 2004). However, there has not been a focused study of genetic variation in secondary signal interactions between APC and T cells, or between T and B cells.

Genetic variation along the B cell/FDC antigen presentation pathway (which is not HLA-dependent) may also influence response to HBV vaccination. For example, C4 and C3 molecules of the complement system play a role in FDC antigen presentation to B cells in germinal centres, as explained previously (Höhler et al., 2002a; Milich and Leroux-Roels, 2003). Deficiencies in the complement components C4 are common in the Caucasian population where *C4A* gene deletions and the non-expressed *C4AQ0* allele have been associated with non-response to the HBsAg (Höhler et al., 2002a; Milich and Leroux-Roels, 2003). While this association has not been studied in Africans, *C4A* deficiencies are present in South African black populations, with differences between ethnic groups (Creemers and du Toit, 1996; Steuer et al., 1994), and therefore, may play a role in non-responsiveness in African populations. The influence of genetic variation in other molecules along this pathway, such as FcγR, CD21 or CD35, has not been examined.

Lastly, the possibility of a 'hole' in the TCR repertoire has been considered as a possible cause of poor response to the HBV vaccine, where the T cells are unable to recognise particular HLA-peptide complexes due to changes in their TCR (Desombere et al., 2000; Roukens and Visser, 2011). However, findings regarding the TCR repertoire in HBV vaccine responders and non-responders do not support this theory (Soroosh et al., 2003).

1.8.2.3 T cell response/cytokine production pathway in association with HBV vaccine response

If defective antigen presentation is not the immunological cause of non-response to HBV vaccination, then defects in immune events that occur after antigen presentation should be considered, such as defects in T cell proliferation, cytokine production or memory cell formation.

A poor T cell proliferative response has been demonstrated in non-responders to the HBV vaccine (Goncalves et al., 2004). Cytokine studies have shown that non-responders do indeed exhibit different cytokine levels compared to those observed in responders (Garner-Spitzer et al., 2013; Goncalves et al., 2004; Jafarzadeh and Shokri, 2003; Jafarzadeh et al., 2004; Jarrosson et al., 2004; Kardar et al., 2002; Larsen et al., 2000; Velu et al., 2008; Wataya et al., 2001). Cytokine profiles of responders were found to differ from those of non-reponders, indicating a T cell dysfunction involved in non-responsiveness to the HBV vaccine (Jafarzadeh and Shokri, 2003; Kardar et al., 2002; Velu et al., 2008). In particular, decreased production of

various Th1 and Th2 cytokines including IL2, IL4, IL10, IL13 TNFα, TNFβ and IFNɣ, was found in non-responders compared to responders to the HBV vaccine (Goncalves et al., 2004; Jafarzadeh and Shokri, 2003; Kardar et al., 2002; Larsen et al., 2000; Velu et al., 2008). This is in contrast to the control of phytohaemagluttinin (PHA) induced cytokine production, which induces similar quantities of cytokines in responders and non-responders (Jafarzadeh and Shokri, 2003; Kardar et al., 2002; Velu et al., 2008).

Other cytokines may also be of interest in vaccine-induced immune response. For example, IL28B has been associated with decreased Th2 cytokine production, decreased B lymphocyte proliferation and antibody production, and with lower seroconversion rates to the Influenza virus vaccine (Egli et al., 2014). The effect of *IL28B* variation on HBV infection is still a matter of debate. For example, Kim et al. (2013a) and Shi et al. (2015) found that genetic variants of *IL28B* were associated with outcomes of HBV infection and Lampertico et al. (2013) found that genetic variants of *IL28B* were associated with treatment outcomes of chronic HBV infection, while Martin et al. (2010) found that genetic variants of *IL28B* were not associated with outcomes of chronic HBV infection. Thus, whether IL28B associates with HBV vaccine non-response would be an interesting aspect to explore. Such an association, or lack thereof, may provide insight into whether non-responsiveness is a general or antigen-specific defect.

Garner-Spitzer et al., (2013) examined if non-responders to HBV vaccination were also non-responders to other antigens, that is, whether the cytokine defects were antigen specific or not. In response to a non-HBV antigen, HBV vaccine non-responders produced lower levels of IL2 and IFNɣ and showed an increased production of IL10 and TNFβ (Garner-Spitzer et al., 2013) which may suggest a non-antigen specific mechanism for non-response. The high level of IL10 production may be responsible for the reduced T cell proliferation to antigens observed in these individuals (Garner-Spitzer et al., 2013). Additionally, non-responders were found to have a somewhat lower mean percentage of naïve T cells and higher percentages of T effector memory cells, which may possibly be a compensation for lack of antibody (Garner-Spitzer et al., 2013). However, another study that identified differences in the cytokine profiles of non-responders and responders after HBsAg stimulations, noted that these differences were absent after stimulation with tetanus toxoid (Larsen et al., 2000), which suggests an antigen-specific non-response.

Consistent with the idea that defects in cytokine response in some non-responders may be due to a non-antigen specific mechanism, genetic polymorphisms in the genes encoding cytokines

and/or their receptors may contribute to poor antibody response. For instance, allele 2 of the *IL12B* promoter (which contains a 4-bp deletion) has been associated with HBsAg non-responsiveness, while the *IL4* TTC haplotype has been associated with response to the HBV vaccine in both European- and African-Americans (Wang et al., 2004).

In summary, there is evidence supporting several mechanisms explaining defective T cell proliferation, abnormal cytokine production or defective T cell differentiation in HBV vaccine non-responders. In addition, cytokine production profiles have been associated with the HLA alleles (Jafarzadeh and Shokri, 2012), which still supports the hypothesis of a specific antigen-presentation defect driving aberrant cytokine production. However, how this may occur has not yet been deciphered.

1.8.2.4 B cell response/antibody production pathway in association with HBV vaccine response

While the average percentages of naïve and specific $CD19^+$ B cells were comparable between responders and non-responders prior to vaccination, HBsAg-specific B cells have been found to be significantly lower in non-responders than in responders following vaccination (Shokrgozar et al., 2001). Additionally, anti-HBs titres and activated B cells were found to be positively correlated (Shokrgozar et al., 2001). This would indicate that the decreased antibody production characteristic of non-response may be due to defects in the B cell response rather than in the antibody production pathway (Shokrgozar et al., 2001). This may be due to defects in the pathways generating the primary B cell repertoire or a defect in the B cell proliferation phase (Shokrgozar et al., 2001). This, in turn, can partially be caused by defects in the Th cell pathways as described above.

Moreover, a higher percentage of transitional (immature) B cells has been observed in non-responders compared to responders (Garner-Spitzer et al., 2013). These B cells possibly have regulatory function through the production of IL10 (Garner-Spitzer et al., 2013). This regulation may affect Th cells as well, which would then contribute to defects in B cell proliferation and differentiation, as described previously.

A defect in the generation of memory B cells could contribute to the low antibody levels observed in non-responders. Variation in the *C4* and *C3* genes of the complement system as explained previously may affect B cell activation and B cell memory development (Höhler et al., 2002a; Milich and Leroux-Roels, 2003). However, the frequency of B memory cells in non-responders has not been compared to that in responders.

To conclude, there are numerous genetic and functional immune characteristics which associate with antibody response to HBV vaccine. Variation in the *HLA* class II alleles and changes in T cell cytokine production seem to be the most common characteristics associated with HBsAg responsiveness. However, it is unclear why such strong associations with *HLA* exist since antigen presentation seems to be functional in non-responders (Desombere et al., 1995). Few studies have examined both *HLA* variation and cytokine production in the same individual (except for Jafarzadeh and Shokri, 2012). It is still unclear whether variation in the *HLA* class II alleles and defective T cell function (the most common associations) are independent or relate to each other in some way. Thus, despite the findings described above, the mechanism underlying non-responsiveness to HBV vaccination is yet to be explained.

1.8.3 HLA expression and HBV vaccine response

There are still many immunogenetic mechanisms contributing to vaccine response that have not yet been examined in the context of the HBV vaccine. Furthermore, although defective antigen presentation, so far, is not thought to be a likely cause of non-response, it cannot be ruled out as several aspects of antigen presentation have not been fully examined. For example, the role of variation in HLA class II expression levels in vaccine response has not yet been studied, and this aspect is discussed further below.

HLA class I molecules are expressed on all nucleated cells, with HLA-C being expressed at a level much lower than (approximately 10% of) that of HLA-A and -B (Zemmour and Parham, 1992). Expression of HLA class II molecules (on B cells, macrophages and DCs) is tissue-specific, varies with the developmental/maturation stage of the cell as well as in response to antigen and cytokine exposure (Glimcher and Kara, 1992). The frequency of cells expressing HLA-DP is generally lower than the frequency of cells expressing HLA-DR, but higher than that expressing HLA-DQ (Edwards et al., 1986, 1985; Gansbacher and Zier, 1988).

Allele-specific HLA-C expression levels have been shown to affect the progression of certain diseases, including HIV and Crohn's disease (Apps et al., 2013; Kulkarni et al., 2011). Expression levels of HLA-DP alleles have been associated with recovery from HBV infection (Thomas et al., 2012) and with graft-versus-host disease during tissue transplants (Petersdorf et al., 2015). In particular, Thomas et al. (2012) showed that HLA-DPB1*01:01 had a higher expression level compared to the *04:01 allele, and HLA-DPB1*01:01 was associated with chronic HBV infection, while *04:01 was associated with HBV clearance.

Additionally, it has been suggested that the concentration of HLA-peptide complexes available on the surface of APCs may affect Th cell differentiation into Th1 and Th2 cells; such an association could provide a link between HLA allele and T cell function (Jafarzadeh and Shokri, 2012). In particular, increased HLA class II expression is predicted to sway Th response towards Th1 differentiation (Baumgart et al., 1998; Cavalli et al., 2016). However, the effect of HLA expression on response to the HBV vaccine is yet to be investigated.

This research project therefore aims to examine the role of *HLA-DP* genetic variation and expression levels in response to HBV vaccination in a vaccinated South African cohort.

1.9 Aims and Objectives

This research study aimed to firstly establish a South African cohort in which response to HBV vaccination, and immunogenetic mechanisms for this response, could be explored. This was the focus of chapter 2.

Our second aim was to explore *HLA-DP* and other *HLA* class II and cytokine SNP variation in South African individuals, and to investigate the association between these genetic variants and anti-HBs titres. This work is discussed in chapter 3.

Thirdly, we aimed to explore the relationship between *HLA-DPB1* expression levels and HBV vaccine response in South Africans. We hypothesized that high expressed *HLA-DP* alleles may associate with poor anti-HBs responses, whereas low expressed *HLA-DP* alleles may associate more with good HBV vaccine response. We also examined whether *HLA* SNPs from chapter 3 were associated with *HLA-DPB1* expression levels. This work is discussed in chapter 4.

A final summary is given in chapter 5.

Study objectives included:

1. Recruit 150 South African laboratory workers previously vaccinated against HBV, collect blood samples, test HBV antibody titres and classify participants as responders, poor-responders or non-responders to the HBV vaccine.

2. Extract DNA from blood samples and genotype DNA for *HLA-DP* alleles, selected SNPs in the *HLA* class II region and selected SNPs in cytokine genes.

3. Analyse the genetic variation data by comparing the variation observed in different ethnic groups in the South African cohort, and by comparing the genetic variation to reference populations; and finally determine whether genetic variation associates with response to the HBV vaccine.

4. Extract RNA from blood samples and use this in qPCR to measure *HLA-DPB1* expression levels.

5. Analyse whether *HLA-DPB1* allele-specific mRNA expression levels occur, and if expression levels associate with response to vaccine. More specifically:
    a) Test if *HLA-DPB1*04:01* homozygosity or heterozygosity associates with low expression and *HLA-DPB1*01:01* homozygosity associates with high expression.
    b) Test if other SNPs in the *HLA-DPB1* region (rs9277534, rs7770370 and rs931) associated with expression level.
    c) And compare *HLA-DPB1* expression levels with anti-HBs titres to test if low-responders had high or low *HLA-DPB1* expression levels.

CHAPTER 2

Establishment and Characterisation of a South African Cohort for Study of HBV Vaccination

2.1 Introduction

As described in Chapter 1, HBV has affected more than two billion people worldwide and approximately 360 million infected individuals are chronic carriers, at high risk of suffering from the sequelae of cirrhosis and hepatocellular carcinoma (WHO, 2009). HBV is highly endemic to Asia and Africa, especially sub-Saharan Africa (Kew, 1996; Kiire, 1996; Romano et al., 2011). In Africa, the carrier rate is approximately 15%, but differs between gender and regions, even within a country (Kiire, 1996), compared to the global carrier rate of 10% (Frodsham, 2005; Shepard et al., 2006).

However, HBV infection can be prevented by vaccination. Since its inclusion in the South African EPI since 1995 (Kew, 2008; Young et al., 2013), it has greatly reduced the HBV infection rates (Hino et al., 2001; Tsebe et al., 2001; Schoub et al., 2002). Despite the success of the HBV vaccine, studies have shown that 5-10% of healthy individuals, globally, do not produce sufficient antibody titres (<10 IU/L) following vaccination (Averhoff et al., 1998; Cleveland et al., 1994; Coates et al., 2001; Dienstag et al., 1984; Roome et al., 1993; Szmuness et al., 1982, 1981, 1980). The few studies that have been performed in South Africans, observed a higher rate of non-response, with an average of 14% (calculated from data from Hino *et al.*, 2001; Tsebe *et al*., 2001; Mphahlele *et al.*, 2002; Schoub *et al*., 2002; Simani *et al*., 2009; Young *et al*., 2013). Many European and Asian studies have investigated the factors contributing to HBV vaccine non-response and several demographic and genetic factors have been associated, yet the immunogenetic mechanisms of non-response are yet to be explained.

There is a lack of South African data on the genetics of HBV vaccination, so in order to explore immunogenetic mechanisms of HBV vaccine response in South Africans, a cohort was needed for this study.

The aim of this chapter was to describe the establishment of a South African cohort in which response to HBV vaccination and the immunogenetic mechanisms for response, could be investigated. The objectives were to recruit 150 South African laboratory workers, collect blood samples and test for HBV antibody titres in order to classify particpants as responders, poor responders or non-responders to the HBV vaccine.

2.2 Materials and Methods

2.2.1 Recruitment and enrolment

South African health care and laboratory workers from the National Health Laboratory Service (NHLS) at Charlotte Maxeke Johannesburg Academic Hospital (CMJAH), the NHLS in Braamfontein, the National Institute for Communicable Diseases (NICD) and the School of Pathology at Wits University were recruited to participate in this study. Many adults in SA remain unvaccinated against HBV since the vaccine was only introduced into the South African EPI in 1995, given at 6, 10 and 14 weeks (and recently 18 months) of age. Adults from these workplaces were targeted since immunity to HBV is a routine requirement of employment. Inclusion criteria were:

- NHLS / School of Pathology health care and laboratory workers with the exception of staff at NICD Centre for Vaccines and Immunology (CVI) (due to study collaboration with CVI)
- previously vaccinated against HBV
- HIV-negative.
- South African, any ethnicity
- over 18 years old

The study was approved by the Human Research Ethics Committee (Medical) of the University of the Witwatersrand, South Africa (clearance certificate no. M140954; see Appendix A) and written, informed consent was obtained from the participants. One hundred and seventy South African, HIV-negative health care and laboratory workers were recruited into this study over three months August-October 2015.

HIV testing, using a rapid HIV test (Alere Determine^{TM} HIV – 1/2), was performed by an independent phlebotomist after counselling. Researchers were not made aware of the identities of any participant excluded due to positive HIV results. Following confirmation of negative HIV status, a venous blood sample of 33 ml was taken. Serum, plasma and cells were isolated and stored on the same day as collection as described below.

Information regarding demographics (age, sex, weight, ethnicity), hepatitis infection history, HBV vaccination history and other factors that may affect response to the HBV vaccine (smoker-status, history of cancer, auto-immune disease) was requested through participant self-recall in the form of a brief questionnaire (Appendix B). Where possible, additional data with

regards to HBV vaccination history and antibody titres were obtained from NHLS records, with permission granted from the Head of Occupational Health, NHLS, and from the participants.

2.2.2 Preparation of blood samples for storage

The 33 ml of blood taken per participant were collected into seven blood collection tubes, as described in **Table 2.1**. In the case that the total blood volume required was not obtained, whatever blood was taken was used as resourcefully as possible.

Table 2.1. Blood collected and purpose thereof

Tube type	Number of tubes	Volume per tube* (ml)	Purpose of storage
CAT	1	2	Serum for antibody tests
EDTA	1	2	Full blood count**
EDTA	2	4	Future DNA and RNA extraction
ACD	3	7	Plasma and PBMCs for functional assays
Total	7	33	

*Volumes are approximate
**Full blood count was performed by Ampath Laboratories
Abbreviations: CAT, Clot Activator Tube; EDTA, Ethylenediaminetetraacetic Acid; ACD, Acid Citrate Dextrose; PBMC, Peripheral Blood Mononuclear Cells

Serum was obtained by centrifugation of whole blood in Clot Activator (CAT) tubes at 1 250 g for 10 minutes and stored at -75°C. Plasma was removed from three Acid Citrate Dextrose (ACD) tubes after centrifugation at 750 g for 10 minutes. The plasma was then centrifuged at 750 g for 15 minutes and the purified plasma was stored at -75°C.

Peripheral Blood Mononuclear Cells (PBMC) were isolated from Ethylenediaminetetraacetic Acid (EDTA) and ACD blood tubes. PBMCs were isolated by density gradient centrifugation over Ficoll-Paque PLUS (GE Healthcare Bio-Sciences AB, Sweden). After removal of the plasma, the blood was diluted approximately 1:1 with Phosphate buffered saline (PBS) (Lonza, BioWhittaker, USA). Fifteen millilitres of blood/PBS mixture were gently layered onto 15 ml of Ficoll and centrifuged at 500 g (Hermle Z400-K) for 30 minutes at 23°C, brake off. This resulted in the formation of four distinct layers based on density: plasma/PBS, PBMCs, Ficoll and erythrocytes/granulocytes. The plasma/PBS layer was first discarded before gently transferring the PBMC layer to a new tube. The PBMCs were washed twice with 30 ml PBS

and pelleted by centrifugation at 250 g for 15 minutes. After discarding the supernatants, the pellets derived from the ACD and EDTA tubes were resuspended in 10 ml and 5 ml RPMI-1640 medium (Lonza, BioWhittaker, USA) respectively.

The PBMCs were counted either manually or using an automated cell counter. For manual cells counts, 10 µl of the cell suspension was mixed with an equal volume of 0.1% Trypan Blue solution (Sigma, UK) and applied to a disposable Neubauer improved haemocytometer (DigitalBio, Korea). Live and dead cells were counted at 400 × magnification using a light microscope. Similarly, for automated cell counts (using the Countess™ Automated Cell Counter, Invitrogen), 10 µl of the cell suspension were mixed with an equal volume of 0.4% Trypan Blue (Life Technologies, USA) and applied to a Countess™ cell counting chamber slide. In both cases, cell viability ($\frac{live\ cell\ count}{total\ cell\ count} \times 100$), total number of live cells obtained and number of live cells per ml of whole blood were calculated. Following the cell counts, approximately 4-5 × 10^6 cells (where possible) were removed from the cells derived from blood from the EDTA tubes and stored directly at -20°C for the purpose of subsequent DNA extractions. The erythrocyte/granulocyte layer under the Ficoll layer originating from the blood of the EDTA tubes was also saved and stored for DNA extraction. The remaining cells derived from blood from the EDTA tubes were mixed with Buffer RLT Plus/β-mercaptoethanol solution (Qiagen, USA) and stored at -75°C for subsequent RNA extractions.

The cells derived from blood from the ACD tubes were stored for functional assays. Cells were diluted to 2 million cells/ml (unless >60 million cells were obtained, then the concentration was increased to 4 or 6 million cells/ml) with freeze media consisting of inactivated FBS (HyClone, USA) with 10% Dimethyl Sulfoxide (DMSO) (Sigma Life Science, USA). Aliquots of 1 ml were placed in a CoolCell/Mr Frosty at -80°C for approximately 24 hours for rate-controlled freezing (-1°C/minute), after which the cells were removed for long term storage at <-80°C.

2.2.3 Serology tests and classification of participants

Serum samples stored at -75°C were batched and used for antibody/antigen testing. If serum was not available, plasma was used instead. Three different antibody/antigen tests were used in order to classify study participants as having experienced current or prior infection, or as being immune due to vaccination or potential non-responders to vaccination. These tests

included tests of hepatitis B surface and core antibody (anti-HBs and anti-HBc) and test for hepatitis B surface antigen (HBsAg).

Hepatitis B surface and core antibody (anti-HBs and anti-HBc respectively) testing was performed by staff of CVI, NICD using the ARCHITECT i 1000SR Immunoassay Analyzer using standard operating procedures. The ARCHITECT Anti-HBs assay (Abbott Ireland, Ireland) was used to measure anti-HBs, while the ARCHITECT Anti-HBc II assay (Abbott, Germany) was used detect anti-HBc. These assays are two-step chemiluminescent immunoassays described below.

For anti-HBs detection, the sample was combined with recombinant HBsAg (rHBsAg) coated microparticles. Anti-HBs present in the sample bound to the antigen coating the microparticles. Following a wash step, acridinium-labelled rHBsAg conjugate was added. After washing, Trigger solutions were added to the mixture resulting in a chemiluminescent reaction, which was measured in relative light units (RLU). There was a direct relationship between RLUs and the amount of anti-HBs in the sample. The analyser determines the concentration of anti-HBs in the specimen using a calibration curve. A concentration greater or equal to 10 IU/L indicated the presence of protective levels of anti-HBs due to either vaccination or infection.

The detection of anti-HBc used a similar principle, but using rHBcAg coated microparticles, and an acridinium-labelled anti-human conjugate. The presence or absence of anti-HBc in the specimen was determined by comparing its chemiluminescent signal to the cut-off signal determined from the calibration. A value greater than or equal to the cut-off signal (1.0) was indicative of the presence of anti-HBc in the sample, suggesting previous or current infection.

HBsAg was tested in any samples positive for anti-HBc in order to differentiate between current and resolved HBV infections, and in samples with anti-HBs titre <10 IU/L in order to rule out a current HBV infection. Any positive HBsAg results were confirmed by repeating the assay. The ARCHITECT HBsAg Qualitative assay is a one-step chemiluminescent immunoassay, where the sample is combined with anti-HBs coated microparticles and acridinium-labelled anti-HBs conjugate. After washing, the Trigger solutions are added and the resulting chemiluminescent reaction is measured as RLUs. A value greater than or equal to the cut-off value of 1.0 indicates the presence of HBsAg in the sample.

Together, the results of the antibody and antigen testing were interpreted as follows in **Table 2.2**.

Table 2.2. Classification of participants according to serology results

Anti-HBs (IU/L)	Anti-HBc (S/CO)	HBsAg (S/CO)	Classification
<10	<1	<1	Potential non-responder to vaccine
<10	>1	>1	Current infection
>10	<1	Not tested	Responder to vaccine
>10	>1	<1	Previous infection

Abbreviations: Anti-HBs, Antibodies against hepatitis B surface antigen; Anti-HBc, Antibodies against hepatitis B core antigen; HBsAg, Hepatitis B surface antigen; IU/L, International Units per Litre; S/CO, Signal to cut-off ratio

In addition, for responders where complete vaccine history and previous antibody testing records were available, study participants were classified as follows in **Table 2.3**.

Table 2.3. Classification of participants based on vaccine history and anti-HBs titre records

Anti-HBs (IU/L) (measured within 6 months after 3 doses of vaccine)	Classification
<10	Non-responder
<200	Low-responder
>200	Normal-responder

Abbreviations: Anti-HBs, Antibodies against hepatitis B surface antigen; IU/L, International Units per Litre

However, due to many instances of incomplete vaccination records, and because our HBV serology testing was not performed within six months of vaccination in most participants, it could not be ascertained if individuals with anti-HBs <10 IU/L and anti-HBc <1 ("Potential non-responder to vaccine") were true non-responders to the HBV vaccine, or if their antibodies had just waned over time, or if they had never been vaccinated nor exposed to HBV. Therefore, a single dose HBV vaccine booster (Heberbio HBV, The Biovac Institute, SA) was offered to participants with Anti-HBs titre <20 IU/L, and negative anti-HBc and HBsAg results. Anti-HBs titres were retested 4-8 weeks after booster. After this round of testing, study participants were classified as follows in **Table 2.4**.

Table 2.4. Final classification of participants based on pre- and post-booster serology results

Anti-HBs pre-boost in 2015 (IU/L)	Anti-HBs post-boost in 2015 (IU/L)	Anti-HBc (S/CO)	HBsAg (S/CO)	Classification
<10	<10	<1	<1	Non-responder
<10	<200	<1	<1	Low-responder
<10	>200	<1	<1	Normal-responder

Abbreviations: Anti-HBs, Antibodies against hepatitis B surface antigen; Anti-HBc, Antibodies against hepatitis B core antigen; HBsAg, Hepatitis B surface antigen; IU/L; International Units per Litre; S/CO; Signal to cut-off ratio

Thus, the low-responders included participants with historical records of anti-HBs titres <200 IU/L within six months of receiving the third dose of HBV vaccine, as well as participants who remained with anti-HBs titres <200 IU/L post-boost in 2015.

Analysis of associations between demographic data and responses to HBV vaccine (normal-responders (>200 IU/L) versus low-responders (<200 IU/l)) were assessed using SPSS v24.0. Fisher's exact test or chi-squared tests were used to assess associations between response and discrete variables (for example, gender, ethnicity), whereas Mann-Whitney U test was used to assess associations between response and continuous variables (age and weight).

2.3 Results

2.3.1 HBV infection and other exclusion criteria

One hundred and seventy South African, HIV-negative laboratory workers were recruited into this study over the period August-October 2015.

Four participants were excluded from this study due to discrepant self-reports of HBV vaccination (signed consent form which stated they had been vaccinated but on questionnaire stated as not vaccinated). Serum samples from the remaining 166 participants were tested for levels of anti-HBs and anti-HBc. We found that fourteen individuals tested positive for anti-HBc, suggesting prior or current infection (14/166 = 8%) (**Table 2.5**). Samples from these individuals were therefore tested for HBsAg. Of these fourteen anti-HBc positive individuals, four tested negative for anti-HBs but positive for HBsAg (indicating current infection), while ten tested positive for anti-HBs and negative for HBsAg (indicating a resolved previous infection). These participants were immune to HBV infection due to natural infection and not due to vaccination, and were thus excluded from this study (**Table 2.5**). Participants with current or prior infection were advised of their status and advised to consult a medical doctor. Additionally, three participants lacked any records of HBV vaccination and tested negative for anti-HBs, anti-HBc and HBsAg. These test results could arise in true non-responders to the vaccine or in unvaccinated individuals. Without proof of prior vaccination, these individuals were also excluded from the study. In total, 21 individuals were excluded from the study, leaving a final sample size of 149 for further investigation.

One of the questions on the questionnaire asked if participants were aware of ever being infected with Hepatitis B. Responses to this question indicated that 156 (94%) of the cohort identified as never having been infected with HBV, while the remaining 6% did not know if they had previously had an HBV infection or not. No participant reported knowledge of previous or current HBV infection. However, from our testing, 8% of the cohort had either had a previous infection or were currently infected; thirteen of these fourteen had self-reported that they had never been infected with HBV while only one was not sure.

Table 2.5 Cohort classified according to serology results before vaccine booster

Anti-HBs in 2015 (IU/L)	Anti-HBc in 2015 (S/CO)	HBsAg (S/CO)	Classification	N (%)
<10	<1	<1	Potential non-responder to vaccine	13 (8%)*
<10	>1	>1	Current infection	4 (2%)
>10	<1	Not tested	Responder to vaccine	139 (84%)
>10	>1	<1	Previous infection	10 (6%)
				Total = 166

*Three of these 'potential non-responders' were ultimately excluded from the study due to lack of vaccination records together with the lack of detectable anti-HBs titres

Abbreviations: Anti-HBs – Antibodies against hepatitis B surface antigen; Anti-HBc – Antibodies against hepatitis B core antigen; HBsAg – Hepatitis B surface antigen; IU/L – International Units per Litre; S/CO – Signal to cut-off ratio

2.3.2 Characterisation of cohort regarding response to HBV vaccine

Thirteen of the 166 participants tested negative for anti-HBs (<10 IU/L), anti-HBc and HBsAg (< 1 (S/CO)) in this first round of antibody testing, and were therefore classified preliminarily as 'potential non-responders to HBV vaccine'. It was unclear if these thirteen were true non-responders or if their antibody levels had just waned over time since vaccination. Three of these 'potential non-responders' were ultimately excluded from the study due to lack of vaccination records together with the lack of detectable anti-HBs titres since they were likely to have never been vaccinated. A booster vaccination was offered to all participants with anti-HBs titres <20 IU/L. Anti-HBs measurement 4 to 8 weeks after booster vaccination showed that all individuals responded to the booster with anti-HBs titres at least >80 IU/L, suggesting that there were no true non-responders to the HBV vaccine in this cohort with titres <10 IU/L.

As seen in **Table 2.6**, participants were classified as 'low-responders' if they had an anti-HBs titre <200 IU/L post three-dose vaccine (based on records) or post booster vaccination. Eight participants (5%) were identified as low responders, while the remaining participants (n=141; 95%) were classified as normal-responders. All participants were advised of their antibody titres pre- and post-booster vaccination.

Participants' self-reported HBV vaccine response status differed somewhat to our test results. Of the 149 participants immune to HBV due to vaccination, four participants (2.7%) self-reported as being non-responders and three participants (2%) self-reported as being low-responders, however none of these participants had records of anti-HBs titre within six months

of vaccination and none were classified as 'low-responders' according to our classification method. Eighty-six participants (57.7%) self-reported as being responders and 56 participants (37.6%) were unsure of their response status. Of the eight participants classified as low-responders in this study, three self-reported as being responders to the HBV vaccine, while the remaining five did not know their response status.

Table 2.6. Classification of participants as low-responders or normal-responders to the HBV vaccine

Anti-HBs in 2015 (IU/L)	Anti-HBs post boost in 2016 (IU/L)	Anti-HBc in 2015 (S/CO)	HBsAg (S/CO)	Classification	N (%)
<10	<10	<1	<1	Non-responder	0 (0%)
<10	<200	<1	<1	Low-responder	8 (5.4%)*
>10	>200	<1	<1	Normal-responder	141 (94.6%)**
					Total=149

*Includes those with post-booster anti-HBs titres >200 IU/L but with records of anti-HBs <200 IU/L 4-8 weeks after three-dose vaccination series

**Includes those with pre-booster anti-HBs titres <10 IU/L but who responded to the boost and had no records of low-response following HBV vaccination

Abbreviations: Anti-HBs, Antibodies against hepatitis B surface antigen; Anti-HBc, Antibodies against hepatitis B core antigen; HBsAg, Hepatitis B surface antigen; IU/L, International Units per Litre; S/CO, Signal to cut-off ratio

2.3.3 Cohort demographics and association with anti-HBs titre

The total cohort (n=149) was largely female (79%) with a mean age of 39 years (ranging from 21 to 74 years) and consisted of Caucasian (50%), black African (33%), Asian (15%) and Coloured (2%) South African individuals (**Table 2.7**). In the Asian category, four participants self-identified as Asian or Chinese, while eighteen self-identified as Indian.

The responder group and low-responder group did not differ significantly with regards to sex, age (at recruitment (n=149) or vaccination (n=114)) or ethnicity. However, the two groups did differ significantly with regards to weight (n=141, p=0.012), with the mean weight of responders being 72 kg (range 40-130 kg) and that of the low-responders being 86 kg (range 70-106 kg).

Table 2.7. Demographics of the responder versus low-responder groups

	Responders	Low-responders	p-value	Total, N (%)
Mean age, years (SD)*	39.1 (11.7)	37.9 (9.0)	0.853	
Mean age at vaccination, years (SD)*	28.1 (10.4)	31.3 (7.9)	0.121	
Mean weight, kg (SD)*	71.9 (18.0)	86.1 (13.5)	**0.012**	
African, N (%)**	47 (33.3)	2 (25.0)	1.000	49 (32.9)
Caucasian, N (%)**	71 (50.4)	4 (50.0)	1.000	75 (50.3)
Asian, N (%)**	21 (14.9)	1 (12.5)	1.000	22 (14.8)
Coloured, N (%)**	2 (1.4)	1 (12.5)	0.154	3 (2.0)
Female, N (%)**	114 (80.9)	4 (50.0)	0.059	118 (79.2)

*Mann-Whitney U test used to obtain p-value
**Fisher's exact test used to obtain p-value
Abbreviation: SD, Standard deviation

Additional factors that may affect immune response were investigated. These factors investigated were: smoker at the time of recruitment, smoker during the vaccination period, presence of a malignancy (previous or current), immunodeficiency, auto-immune disease and liver disease. The only factor that significantly differed between the responders and low-responders was smoking during the vaccination period, where this was more frequent than expected in the low-responders (**Table 2.8**).

Table 2.8. Comparison of factors that may affect immune response between responders and low-responders

	Responders, N (%)	Low-responders, N (%)	p-value*	Total, N (%)
Smoker in 2015	10 (7.1)	2 (25)	0.128	12 (8.1)
Smoker during vaccination period	6 (4.3)	3 (37.5)	**0.008**	9 (6.1)
Have had a malignancy	4 (2.9)	0	1.000	4 (2.7)
Immunodeficient	0	0	-	0
Have an auto-immune disease	22 (15.7)	3 (37.5)	0.134	25 (16.9)
Have a liver disease	1 (0.7)	0	1.000	1 (0.7)

*assessed by Fisher's exact test

Thus, in this cohort, weight and smoker status at the time of vaccination were significantly different between normal-responders and low-responders.

2.4 Discussion

2.4.1 Knowledge of HBV infection and vaccination status

We noted that none of the currently or previously infected individuals were aware of their infection status. Such lack of knowledge of HBV infection status has been seen previously, where 56% of South African health care workers (HCW) who had experienced occupational exposure to HBV, did not know their HBV infection status (Khan and Ross, 2013). This lack of knowledge may arise due to the fact that HBV, which is endemic to sub-Saharan Africa, may be acquired in early childhood in rural areas or later in life in urban areas and is often asymptomatic or only mildly symptomatic (Kew, 2008). Furthermore, although easily identifiable, most cases in SA are not notified (Kew, 2008). Lack of knowledge of HBV infection status among laboratory workers may lead to unnecessary prophylactic treatment upon accidental exposure and/or to unnecessary HBV vaccination; however, this is a cost-effectiveness concern, rather than a health issue (Immunization Action Coalition, 2005). It may also lead to transmission of the virus to unvaccinated contacts, who could be vaccinated if they were aware of the infection risk.

There were some differences seen in the self-reports of response status and our serology results, however, it is difficult to comment on the significance thereof due to the lack of vaccination and anti-HBs titre records with which to corroborate self-reports. Nevertheless, self-knowledge of HBV vaccine response status was low, with 37.6% of our cohort unsure of their status. This was unexpected as laboratory workers' anti-HBs levels are measured after vaccination to assess response. This may suggest lack of communication between the laboratory workers and the health care professionals performing these tests. Due to the risk of occupational exposure to HBV, it is important for laboratory workers to know the risks, to be vaccinated against HBV and to know their response status.

2.4.2 HBV vaccine response in the cohort

From published reports of HCWs, 4-10% do not respond (<10 IU/L) to three doses of the HBV vaccine (Averhoff et al., 1998; Cleveland et al., 1994; Coates et al., 2001; Dienstag et al., 1984; Roome et al., 1993; Szmuness et al., 1982). These rates are similar in the normal adult population (Coates et al., 2001; Rendi-wagner et al., 2001; Szmuness et al., 1981, 1980). However, we observed zero non-responders in this cohort.

A difference in response rates between these studies and the current one may be the demographic characteristics of the cohorts. Factors that increase the risk of non-response include male gender and increasing age (Cleveland et al., 1994; Averhoff et al., 1998; Fisman et al., 2002; Tohme et al., 2011), as discussed below in section 2.4.3. Our cohort was mostly female (80.9%) and 90% of the cohort was younger than 40 years at the time of vaccination. This may have contributed to the lack of non-responders in the cohort.

Additionally, this difference in response rates may be due to the fact that these aforementioned studies are reporting seroconversion rates following a three-dose immunization schedule whereas individuals who are at a risk of occupational exposure to HBV (including laboratory workers) are recommended to receive additional doses if they do not respond to the initial vaccination series (Immunization Action Coalition, 2005). Additional HBV doses increase the chance of seroconversion (Clemens et al., 1997; Das et al., 2003; Desombere et al., 1998; Heininger et al., 2010; Struve et al., 1994). The reason as to why some individuals respond only after additional doses is yet to be explained (Heininger et al., 2010). It has been speculated that vaccination may result in cell-mediated immune responses and the humoral response only develops after booster vaccination (Chiaramonte et al., 1995); indeed, it has been shown that some non-responders do develop a cellular immune response *in vitro* (Jarrosson et al., 2004). Furthermore, as individuals respond to different numbers of booster doses (and non-responders show variable levels of response *in vitro*), there seems to be different grades of poor response which suggests that there is a qualitative difference between poor responders that is yet to be explained (Desombere et al., 1998; Jarrosson et al., 2004; Zuckerman, 1996).

Furthermore, a factor which hinders the classification of responders/non-responders in a cross-sectional study, so long after vaccination, is the waning of anti-HBs levels (Shepard et al., 2006; Zuckerman, 1996). From available records, it is clear that anti-HBs levels do wane over time, as expected. Boosting responder individuals whose antibodies wane to even <10 IU/L is possibly unnecessary (Banatvala and Van Damme, 2003) as previous responders remain protected due to memory B and T cells (Bauer and Jilg, 2006).

2.4.3 Clinical factors affecting HBV vaccine response

We found that increased weight and smoking during the vaccination period was associated with increased risk of low anti-HBs response. Such associations have been seen previously (Averhoff et al., 1998; Roome et al., 1993; Yang et al., 2016). We note that BMI (kg/m^2) would

be a better measure than weight alone, however, height was not included in the questionnaire and thus, BMI could not be calculated.

In attempt to explain the observed association between weight and HBV vaccine response, it has been suggested that the vaccine may be inadvertently deposited into fat rather than into muscle in overweight individuals (Shaw et al., 1989). This may result in delayed absorption of the antigen thereby allowing time for its enzymatic denaturation (Ingardia et al., 1999; Milner and Beck, 2012; Shaw et al., 1989). Deposition of the vaccine into fat may also result in reduced immune response due to the scarcity of phagocytes and/or APCs in deep fat (Shaw et al., 1989). Additionally, overweight individuals have altered levels of hormones and nutrients such as leptin, adiponectin and proinflammatory cytokines (Milner and Beck, 2012). Leptin is a growth factor that can increase T cell proliferation and survival; this and the increase in proinflammatory cytokines may suppress the Th2 or antibody response leading to the increased risk of non-response associated with increasing weight. In contrast, leptin deficiency noted in obese individuals due to T cells becoming resistant to leptin signalling may explain the reduced vaccine-induced immunity observed in obese persons (Milner and Beck, 2012). This lack of leptin signalling may also enhance Treg cell proliferation which could result in the suppression of the immune response (Milner and Beck, 2012).

With regards to smoking, higher leukocyte counts have been observed in smokers compared to non-smokers (Holt, 1987; Jensen et al., 1998; Noble and Penny, 1975). It has also been observed that nicotine suppresses leukocyte chemotaxis, decreases phagocytic capabilities and inhibits T cell function (Noble and Penny, 1975). As HBV vaccine response is T cell dependent (Cupps et al., 1984), such suppression of T cell function would adversely affect the HBV vaccine-induced immune response (Ingardia et al., 1999).

Together with weight and smoking, increasing age and male sex are often associated with increased risk of non-responsiveness to the HBV vaccine (Averhoff et al., 1998; Cleveland et al., 1994; Fisman et al., 2002; Tohme et al., 2011). However, we did not find age or sex associated with response to the HBV vaccine.

Although we did not find an association, there is a biologically plausible reason for the association with age – the phenomenon of immunosenescence noted in older individuals, where both cellular and humoral immunity wane with age (Effros, 2007; Ginaldi et al., 1999; Hodes, 1997). Moreover, a meta-analysis performed in 2002 (Fisman et al.) determined that there was a significantly increased risk of non-response to the HBV vaccine with increasing age even

though they reviewed a similar number of studies that found such an association and those that did not. They suggested that insufficient statistical power is the reason behind studies not observing an association with age and HBV vaccine response (Fisman et al., 2002). This may well apply to this study due to the small sample size of low-responders and the lack of any true non-responders. Additionally, as 90% (134/149) of participants were younger than 40 years of age at the time of vaccination (which is plausible due to NHLS laboratory workers being vaccinated against HBV at the start of their career which would usually translate to vaccination at younger age), it is possible that age did not affect the anti-HBs titres significantly in this study.

Although we did not find an association between sex and HBV vaccine response, such an association has been observed in previous studies on HBV vaccine response (Averhoff et al., 1998; Chathuranga et al., 2013; Cleveland et al., 1994). Poorer immune response in males has been noted with other vaccines as well (Klein et al., 2015) and also has a biologically plausible explanation: oestrogen has the capability of activating signalling pathways and inducing transcription of various genes which act on multiple immune pathways, such as B cell proliferation and survival, T cell responses and cytokine regulation (Klein et al., 2015; Lang, 2004). Oestrogen has also been shown to increase immunoglobulin (Ig) production (Kanda and Tamaki, 1999), whereas testosterone inhibits the production of IgG and IgM (Kanda et al., 1996). Furthermore, there are several immune-related genes located on the X chromosome and as X inactivation in females may favour the inactivation of mutant genes, females would have an advantage over males in the case of deleterious mutations in these genes (Klein et al., 2015). Furthermore, the low numbers of males in this study may have affected the findings and the lack of true non-responders may have masked any significant differences between the anti-HBs responses of the sexes.

Chronic disease has been associated with HBV vaccine response in older vaccinees (>40 years) and not in those <40 years (Averhoff et al., 1998). This may explain why we did not find an association between HBV vaccine response and the other factors that may affect immune response (presence of malignancy, immunodeficiency, autoimmune disease, liver disease) in our cohort that was relatively young (mean age of 29.7 years) at time of vaccination.

CHAPTER 3

Associations Between Genetic Variation and HBV Vaccine Response in a South African Cohort

3.1 Introduction

HLA-DP is a member of the HLA class II molecules. Each molecule is composed of an alpha and beta chain which are encoded by the *HLA-DPA1* and *–DPB1* genes, respectively. APCs, such as macrophages, dendritic cells and B cells, express HLA class II molecules which are involved in the presentation of extracellular antigenic peptides to $CD4^+$ T cells. The interaction between the HLA-peptide complex and the T cell initiate the immune response. Thus, *HLA* associations with various diseases and vaccine responses have a biological basis.

In general, studies investigating *HLA-DP* genes and their impact on human disease are somewhat under-represented compared to *HLA-DR* and *–DQ*. This may be due to *HLA-DPA1* and *–DPB1* being less polymorphic than *HLA-DR* and *–DQ* genes and also because HLA-DP cell surface expression levels are likely to be lower than that of HLA-DR (Edwards et al., 1986; Thomas et al., 2012). Therefore, this study aimed to investigate the role of *HLA-DP* variation in HBV vaccine response. GWAS have consistently associated the *HLA* region with both HBV infection outcome and HBV vaccine response (Hu et al., 2013; Kamatani et al., 2009; Mbarek et al., 2011; Nishida et al., 2012; Pan et al., 2013; Png et al., 2011; Wu et al., 2015). Moreover, the most significant associations in GWAS of HBV infection outcomes in Asian populations are variants located within the *HLA-DP* region (Guo et al., 2011; Kamatani et al., 2009; Nishida et al., 2012). Additionally the *HLA-DP* locus has also been strongly associated with HBV vaccine response in GWAS in Asian populations (Wu et al., 2015).

Cytokines are key players in the immune response, functioning to promote particular immune pathways, such as CTL-mediated cell lysis or antibody production (Broere et al., 2011). Furthermore, cytokine profiles differ between HBV vaccine responders and non-responders (Larsen et al., 2000; Kardar et al., 2002; Jafarzadeh and Shokri, 2003; Goncalves et al., 2004; Velu et al., 2008). Thus, genetic variation in these cytokine genes may affect their functioning and is therefore, important to investigate.

As there is no South African data on the genetics of HBV vaccination, this study aimed to investigate the role of *HLA-DP* and cytokine variation in HBV vaccine response in a South African cohort. We aimed to genotype both the *HLA-DPA1* and *-DPB1* genes, as well as SNPs in the *HLA* class II region (with a focus on the the *HLA-DP* region) and in cytokine genes

3.2 Materials and Methods

3.2.1 DNA extraction

Human DNA was extracted from the stored erythrocyte/granulocyte layer using the DNA Isolation Kit (Invitrogen, Life Technologies), according to the manufacturers protocol. The erythrocyte/granulocyte mixture was thawed at room temperature and 600 µl was combined with 900 µl Red Cell Lysis Buffer. After vortexing, the granulocytes were pelleted by centrifugation at 12 000 g for 1 minute, and this wash step was repeated a second time. Protease in Nuclear Lysis Buffer (125 µl) was added to the pellet after which the tube was incubated at 65°C in a heat block for 10 minutes with a brief vortex every 2-3 minutes. The 275 µl of Protein Clearing Solution was then added to the sample, followed by a 10 minute incubation on ice. The cell debris was pelleted by centrifugation for 5 minutes at 13 500 g. The supernatant (containing the DNA) was carefully transferred to a clean tube. In order to precipitate the DNA, ice cold 90% ethanol was slowly added to the solution and the tube was inverted several times. In the case that no DNA was visible, the tube was incubated in a -20°C freezer for 10 minutes or overnight. The DNA was pelleted by centrifugation at 13 500 g for 5 minutes. The supernatant was discarded while being careful not to disrupt the DNA pellet. Ethanol (70%) was added to the pellet before centrifuging at 13 500 g for 2 minutes. The ethanol was discarded and the pellet air-dried. The DNA was resuspended in DNA Suspension Buffer by incubation at 65°C for 15 minutes.

The concentration and quality of the DNA was assessed by using the NanoDrop 1000 spectrophotometer using absorbance at 260 nm (A260) for DNA concentration and the ratio of absorbance at 260 nm to that at 280 nm (A280) and to the absorbance at 230 nm (A230) as quality assessments. The A260/A280 ratio of pure DNA should be approximately 1.8, while a lower ratio indicated the presence of protein or other contaminants that absorb at 280 nm. The A260/A230 ratio should ideally be higher than the A260/A280 ratio (usually 2.0 – 2.2); a lower ratio indicated the presence of contaminants such as EDTA and phenol. The DNA was then diluted to a working concentration of 20 ng/µl with distilled water and stored at -20°C. To ensure that the DNA was of appropriate quality for downstream applications, several samples were selected to be used in a conventional polymerase chain reaction (PCR) run. As all samples were successfully amplified, the DNA was considered of sufficient quality to be used in downstream applications.

3.2.2 *HLA-DPA1* and -*DPB1*

3.2.2.1 *HLA-DPA1* and -*DPB1* genotyping

HLA-DPA1 and -*DPB1* genotyping of the cohort DNA samples was performed using the LIFECODES HLA-DPA1/DPB1 SSO Typing Kit (Immucor) on a Luminex instrument at the NICD according to the manufacturer's instructions. The Luminex technology allowed for the simultaneous detection of multiple analytes in a single reaction tube or well by the use of sequence-specific oligonucleotide (SSO) probes bound to colour-coded microbeads (Heinemann, 2009).

This procedure for typing *HLA* class II alleles began with the amplification (by PCR) of exon 2 (and exon 3 for *HLA-DPB1*), which is the most polymorphic region of the *HLA* class II gene (Heinemann, 2009). Using biotinylated primers (Heinemann, 2009). The components and cycling conditions of the PCR amplification procedure are listed in **Tables 3.1** and **3.2** respectively. A no-template control and a positive control were included. The positive control was either DNA extracted from MCF7 cell line (kindly provided by Professor R Veale, Wits University) whose *HLA-DPB1* alleles have previously been determined (Adams et al., 2005) or International Histocompatibility Working Group reference DNA sample IHW09021 (kindly provided by Dr H Hong, NICD).

Table 3.1. Reaction components for HLA-DP amplification

Component	Amount per PCR reaction
LIFECODES Master Mix	6 µl
Lifecodes Taq Polymerase	0.2 µl
Genomic DNA	Total of ~80 ng
Nuclease-free water	To 20 µl final volume

Table 3.2. Thermal cycler conditions for amplification

Step	Temperature	Incubation Time	Number of Cycles
1	95°C	3 minutes	1
2	95°C	15 seconds	12
	60°C	30 seconds	
	72°C	30 seconds	
3	95°C	10 seconds	28
	63°C	30 seconds	
	72°C	30 seconds	
4	72°C	2 minutes	1
5	4°C	∞	1

The amplified DNA was denatured and hybridized to complementary allele-specific SSO probes attached to the colour-coded microbeads (Heinemann, 2009) which are provided with the kit. The subsequent PCR products were then hybridized with the microbeads according to the conditions in **Table 3.3**.

Table 3.3. Thermal cycler conditions for hybridization

Temperature	Incubation Time
97°C	2 minutes
47°C	10 minutes
56°C	8 minutes
56°C	HOLD

Streptavidin-conjugated with phycoerythrin (SAPE), a fluorescent molecule was then added to each well. The streptavidin is capable of binding to the biotin on the PCR products. A solution of 170 µl SAPE was added to each well at the 56°C hold.

The plate was then transferred to the Luminex instrument for reading. Lastly, the beads were passed in a single stream through two lasers. One laser excited the colour in the microbead (this identified which microbead was being read), while the other excited the colour on the bead surface, the SAPE (Heinemann, 2009).

The Luminex analyser detected the signal and the software assigned the HLA alleles (Heinemann, 2009). In order to exclude non-specific background signals and to normalize the data, each microbead mixture contains positive and negative control probes (Heinemann, 2009). MatchIT! DNA software from Immucor was used to analyse the data

(http://www.immucor.com/) along with the allele assignment protocol recommended by the American Society for Histocompatibility and Immunogenetics (ASHI). The assignment of HLA alleles was based on the sequences listed in the IMGT/HLA database (http://www.ebi.ac.uk/ipd/imgt/hla/) (Heinemann, 2009).

Luminex *HLA* genotyping typically produces results with 2-digit, 4-digit nomenclature or G group nomenclature, but does not usually produce results with 6 or 8-digit nomenclature. These nomenclature types refer to the following (http://hla.alleles.org/nomenclature/naming.html):

- 6-8-digit nomenclature e.g. *HLA-DPA*01:03:01:01*. The 5th and 6th digits refer to synonymous DNA changes in coding regions, while the 7th and 8th digits refer to differences in a non-coding region.
- 4-digit nomenclature e.g. *HLA-DPA*01:03*. This is a unique HLA protein.
- 2-digit nomenclature: e.g. *HLA-DPA*01*. Each 2-digit group either encodes a particular serological antigen, or includes a group of alleles with high sequence homology.
- G group: group of alleles with the same exon 2 sequence (for class II alleles) or same exon 2+3 sequences (for class I alleles) e.g. *HLA-DPA1*02:02:02G* includes alleles *02:02:02* and *02:06*.

However often the 4-digit Luminex genotyping results are ambiguous (that is, where different allele combinations produce identical sequences). Therefore, Luminex results are often reported as either 2-digit genotypes, ambiguous genotypes (that cannot be further defined without further laboratory work) or genotypes containing G group alleles.

3.2.2.2 Statistical analysis of *HLA-DPA1* and -*DPB1* alleles

HLA-DPA1 and –*DPB1* allele frequencies were compared between the different South African ethnicities using chi-squared tests or Fisher's exact test in SPSS v24.0. The South African Coloured population were not included in these analyses due to the small sample size (n=3).

HLA-DPA1 and –*DPB1* allele frequencies within each ethnicity were compared to data from the Allele Frequency Net Database (AFND, http://www.allelefrequencies.net/) using Fisher's exact tests in SPSS v24.0. Due to the limited *HLA-DP* data available from South African populations, the South African black population in this study was compared to sub-Saharan black populations, the South African Caucasian population was compared to European Caucasoid populations (in particular to data from France, Norway, Sweden and UK/England)

and the South African Asian population was compared to data from India due to majority (18/22) of the 'Asian' group self-identifying as Indian. Although individuals of East Asian and South Asian descent often have different allele frequencies, participants from these population groups were combined and collectively classified as Asian in order to increase the sample size in this group. The South African Coloured population were not included in these analyses due to the small sample size (n=3).

The *HLA-DPA1* and *–DPB1* allele frequencies in the total cohort were compared between responders and low-responders using Fisher's exact tests in SPSS v24.0.

3.2.3 SNP genotyping

3.2.3.1 SNP selection

To investigate the associations between SNPs in candidate genes (such as *HLA-DP* regulatory regions, cytokine genes or costimulatory signalling genes) and response to HBV vaccination, we chose 31 SNPs of interest to be genotyped from such gene regions. SNPs were selected by surveying the literature for SNPs associated with HBV vaccine response, HBV infection outcome, other vaccine response and/or HLA-DP expression levels.

SNPs were included if the minor allele frequency (MAF) in the Luhya in Webuye, Kenya (LWK) was >5%. The LWK population was used as a proxy reference population as there is limited whole genome or *HLA-DP* data available for South Africans. MAF >5% in LWK was confirmed, by using ENGINES (http://spsmart.cesga.es/engines.php) and 1000 Genomes data (http://browser.1000genomes.org/index.html). We note that our cohort was of multiple South African ethnicities, but we have tried to select SNPs in this study that would be applicable to further genetic studies of black South Africans. Linkage disequilibrium (LD) between the 31 selected SNPs in the LWK population was determined using the LD TAG SNP Selection tool at the National Institute of Environmental Health Sciences website (https://snpinfo.niehs.nih.gov/snpinfo/snptag.htm). In the case where there was significant LD between SNPs, a tagSNP was selected to represent the relevant SNPs. The SNPs (of the 31 selected SNPs) captured by these tagSNPs are shown in Appendix C.

While 31 SNPs of interest were identified (**Table 3.4**), only 24 SNPs were compatible with single multiplex genotyping using the Mass Array system.

Table 3.4. Details of SNPs selected for genotyping

SNP ID	Gene/Gene Region	Chromosomal location	Location relative to Gene	Association	Allele Association	Study Population	References
rs1800872*	*IL10*	1:206773062	Upstream	HBV vaccine response	CC ↑ response	Taiwanese	Lin et al., 2011
					(in haplotype C ↑ response)	Germans	Höhler et al., 2005
				Measles vaccine response	T ↑ response	European-American;	Dhiman et al., 2007
rs1800871*	*IL10*	1:206773289	Upstream	HBV vaccine response	CC ↑ response	Taiwanese	Lin et al., 2011
					(in haplotype C ↑ response)	Germans	Höhler et al., 2005
				PCV7 vaccine response	A ↓ response	European-American	Yucesoy et al., 2009
				Measles vaccine response	A ↑ response		Dhiman et al., 2007
rs1800896	*IL10*	1:206773552	Upstream	HBV vaccine response	(in haplotype A ↑ response)	Taiwanese	Lin et al., 2011
						Germans	Höhler et al., 2005
				Diphtheria vaccine response	A ↑ response;	European-American	Yucesoy et al., 2009
rs1143634	*IL1beta*	2:112832813	Exon, synonymous	HBV vaccine response	A ↑ response	European-American	Yucesoy et al., 2002
rs1143633	*IL1beta*	2:112832890	Intron	HBV vaccine response	(haplotype AG ↓ response)	Chinese	Chen et al., 2011
rs1143627	*IL1beta*	2:112836810	5′ UTR	HBV vaccine response		Chinese	Chen et al., 2011
rs2069763	*IL2*	4:122456327	Exon, synonymous	Measles vaccine response	G ↑ response	European-American	Dhiman et al., 2007
rs2243248	*IL4*	5:132672952	Upstream	HBV vaccine response	(in haplotype T ↑ response)	Various ethnicities (USA)	Wang et al., 2004
					C ↓ response	Chinese	Chen et al., 2011
rs2243250	*IL4*	5:132673462	Upstream	HBV vaccine response	(in haplotype T ↑ response)	Various ethnicities (USA)	Wang et al., 2004
					T(T) ↑ response	Taiwanese	Lin et al., 2012
						Asian	Cui et al., 2013
				HBV infection outcome	T ↑ susceptibility	Caucasian	Zheng et al., 2013
					T ↓ susceptibility	Chinese males	Lu et al., 2014

rs2070874	*IL4*	5:132674018	5′ UTR	HBV vaccine response	(in haplotype C ↑ response)	Various ethnicities (USA)	Wang et al., 2004
					T ↑ response	Asian	Cui et al., 2013
				HBV infection outcome	T ↓ susceptibility	Chinese	Lu et al., 2014
rs3212227	*IL12B*	5:159315942	3′ UTR	HBV vaccine response	AA ↑ response	European American	Yucesoy et al., 2009
				Measles vaccine response	AA ↓ response		Dhiman et al., 2007
rs3213093	*IL12B*	5:159323971	Intron	Measles vaccine response	GG ↓ response	European-American	Dhiman et al., 2007
rs4711998*	*IL17A*	6:52185555	Upstream	HBV vaccine response	GG ↑ response	Iranian	Borzooy et al., 2016
rs2069727	*IFNG*	12:68154443	Downstream	HBV vaccine response	C ↑ response	Gambian	Hennig et al., 2008
rs1805010*	*IL4R*	16:27344882	Exon, missense	HBV vaccine response	AA ↑ response	Taiwanese	Lin et al., 2012
				Tetanus		European-American	Yucesoy et al., 2009
				PCV7			
				Measles vaccine response			Dhiman et al., 2007
rs1805015	*IL4R*	16:27362859	Exon, missense	HBV vaccine response	C ↑ response	Chinese	Chen et al., 2011
rs2230433	*ITGAL*	16:30506720	Exon, missense	HBV vaccine response	C ↑ response	Gambian	Hennig et al., 2008
rs8099917	*IL28B*	19:39252525	Upstream	HBV viral load	T(T) ↓ viral load	Chinese	Li et al., 2011
				IL28B expression levels;	TT ↑ expression		
				Influenza vaccine response	TT ↓ response	Canadian	Egli et al., 2014
rs3135363	*HLA-DRA*	6:32421871	Upstream	HBV vaccine response	C ↓ response	Indonesian	Png et al., 2011
						Chinese	Pan et al., 2013
rs7453920	*HLA-DQB2*	6:32762235	Intron	HBV infection outcome	A ↓ susceptibility	Japanese	Kamatani et al., 2009
						Thai	
						Chinese	Mbarek et al., 2011
							Hu et al., 2012
							Hu et al., 2013
rs2395309	*HLA-DPA1*	6:33058469	Upstream	HBV infection outcome	G ↑ susceptibility	Japanese Thai	Kamatani et al., 2009
						Chinese	Guo et al., 2011
rs2071349	*HLA-DPA1/DPB1*	6:33075743	Intron (DPA1), upstream (DPB1)	-	-	-	-

rs7770370	*HLA-DPB1*	6:33081144	Intron	HBV vaccine response	A ↑ response	Taiwan	Wu et al., 2015
						Korean	Roh et al., 2016
rs3749985*	*HLA-DPB1*	6:33086656	3′ UTR	-	-	-	-
rs931	*HLA-DPB1*	6:33086773	3′ UTR	-	-	-	-
rs9277533*	*HLA-DPB1*	6:33086944	3′ UTR	HBV infection outcome	GG ↑ susceptibility	European- and African-American	Thomas et al., 2012
				DPB1 expression levels	GG ↑ expression		
rs9277534	*HLA-DPB1*	6:33087030	3′ UTR	HBV infection outcome	GG ↑ susceptibility	European- and African American	Thomas et al., 2012
				DPB1 expression levels	GG ↑ expression		
				Rubella virus vaccine response	GG ↓ response		Lambert et al., 2014
rs9277535	*HLA-DPB1*	6:33087084	3′ UTR	HBV infection outcome	A ↓ susceptibility	Thai	Kamatani et al., 2009
						Japanese	Mbarek et al., 2011
						Chinese	Hu et al., 2012
							Hu et al., 2013
							Guo et al., 2011
				HBV vaccine response	A ↑ response	Indonesian	Png et al., 2011
						Korean	Roh et al., 2016
rs9277536*	*HLA-DPB1*	6:33087113	3′ UTR	HBV infection outcome	GG ↑ susceptibility	European- and African-American	Thomas et al., 2012
				DPB1 expression levels	GG ↑ expression		
rs3130186	*HLA-DPB1*	6:33088430	3′ UTR	-	-	-	-
rs2064479	*HLA-DPB1*	6:33104463	Downstream	Rubella virus vaccine response	C ↑ response	European-American	Lambert et al., 2014

*SNPs not genotyped

Abbreviations: HLA, Human Leukocyte Antigen; IFNG, Interferon-gamma; IL, Interleukin; HBV, Hepatitis B Virus; PCV7, Pneumococcal Conjugate Vaccine (heptavalent); USA, United States of America; UTR, untranslated region

3.2.3.2 MassArray Analysis

The selected SNPs were genotyped by iPLEX® / MassArray® System by Inqaba Biotec (Pretoria, SA). SNP rs numbers were submitted to Inqaba Biotec for evaluation by the Assay Design Suite software which was used to design the PCR and extension primers. Of the 31 SNPs submitted, it was possible to multiplex 24 of the SNPs in a single well.

A region of DNA that contained each SNP was amplified by PCR using the components listed in **Table 3.5**, according to the cycling conditions in **Table 3.6**. The unincorporated deoxynucleotide triphosphates (dNTPs) were dephosphorylated by shrimp alkaline phosphatase (SAP) at 37°C for 40 minutes, followed by a 5 minute incubation period at 85°C to inactivate the enzyme. The samples were then submitted to the single base extension reaction using the extension primers and mass-modified dideoxynucleotides (ddNTPs) (see **Tables 3.7** and **3.8** for reaction components and cycling conditions). The single base extension reaction involved using an extension primer that bound directly adjacent to the SNP site. ddNTPs were utilized for this extension so that the primer was extended by only one nucleotide that is complementary to the nucleotide at the SNP site. The ddNTPs were also mass-modified so that DNA molecules produced differ by mass based on the allele (Gabriel et al., 2009; Agena Bioscience, iPLEX Reagents User Guide).

Table 3.5. Reaction components for PCR of target regions

Component	Amount per PCR reaction (µl)
Nanopure Water	0.8
10× PCR Buffer	0.5
25 mM $MgCl_2$	0.4
25 mM dNTP mix	0.1
0.5 µM Primer Pool	1.0
Polymerase, 5 U/µl	0.2
DNA template (25 ng/µl)	2.0

Table 3.6. Thermal cycler conditions for amplification of target regions

Step	Temperature	Incubation Time	Number of Cycles
1	94°C	2 minutes	1
2	94°C	30 seconds	
	56°C	30 seconds	46
	72°C	60 seconds	
3	72°C	5 minutes	1
4	4°C	∞	1

Table 3.7. Reaction components for single base extension reaction

Component	Amount per PCR reaction (µl)
NanoPure Water	0.6182
10× iPLEX Buffer Plus	0.2
iPLEX dNTP Termination mix, 45×	0.2
Probe Pool (11 µM each)	0.9409
iPLEX Polymerase, 220×	0.0409
DNA template (PCR+SAP)	7.0

Table 3.8. Thermal cycler conditions for single base extension reaction

Step	Temperature	Incubation Time	Number of Cycles
1	94°C	30 seconds	1
2	95°C	5 seconds	60
	52°C	3 seconds	5
	80°C	3 seconds	
3	72°C	3 minutes	1
4	15°C	HOLD	1

The products from the single base extension reaction were diluted with approximately 20 µl NanoPure water after which the resin was added to the wells in order to desalt the extension products. The samples were transferred to the chip using a Nanodispenser where they crystalized with the MALDI matrix on the chip, and analysed by the MassARRAY Analyzer.

These products were then applied to a microchip and underwent Matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectrometry whereby a laser is used to generate ionized products. The ions accelerate into a tube and towards the detector. The time of flight of these ionized molecules differ depending on the mass of the molecule, which was measured by the mass spectrometer. The MassARRAY Analyzer can distinguish between molecules that differ by 16 Da, and therefore mass of PCR products determined which alleles were present. The MassARRAY Typer software automatically generated reports that identify the SNP alleles (homozygous and heterozygous) in each sample and SNP genotypes were finally exported into an Excel worksheet.

3.2.3.3 Statistical analysis of SNP data

gPLINK v1.07 (Purcell et al., 2007) was used to analyse the SNP results. To remove individuals or SNPs with a high amount of missing data, both the individual and SNP missingness rates were set thresholds of $p<0.1$. To remove SNPs with a low MAF, a MAF threshold of $p<0.05$ was applied, and to remove SNPs that fail Hardy-Weinberg equilibrium (HWE) checks, a HWE threshold of $p<0.001$ was used.

Allele, genotype and haplotype frequencies in the South African cohort

Allele and genotype frequencies for each SNP were counted in the whole South African cohort and within each ethnicity in our cohort. Allele frequencies for each SNP were compared among ethnicities using chi-square tests in SPSS v24.0.

Haplotypes were generated where >1 SNP per gene or chromosome were tested, including ten SNPs in the *HLA* class II region, three SNPs in *IL1B*, three SNPs in *IL4*, two SNPs in *IL12B* and two SNPs on chromosome 16 (genes *IL4R* and *ITGAL*). For the cytokine/signalling molecule genes (*IL1B*, *IL4*, *IL12B* and *IL4R* and *ITGAL*), two-, three- four- and five- SNP haplotypes were generated using the sliding window approach in gPLINK (Purcell et al., 2007). Linkage patterns (D′) between SNPs were analysed using Haploview (Barrett et al., 2005). See Appendix D for r^2 values. The LD blocks were generated using the Confidence Intervals algorithm (Gabriel et al., 2002).

For the *HLA* region, we used the Arlequin v3.5 software (Excoffier and Lischer, 2010) to phase the genotypic data into haplotypes which included the 10 SNPs within the *HLA* region. We performed haplotype phasing using the Arlequin pseudo-Bayesian approach (ELB algorithm)

with 2000 iterations (Excoffier and Lischer, 2010) in the total cohort (n=149). This algorithm estimated haplotype frequencies in the total cohort, and also produced the estimated gametic phase of each individual. Because our cohort was of mixed ethnicity, we also repeated this phasing process in separate ethnic groups within our cohort to see if predicted haplotypes changed.

Comparisons of allele frequencies to other populations

Allele frequencies in our South African cohort from the 24 SNPs were compared to data from the 1000 Genomes Project (The 1000 Genomes Project Consortium, 2015), which is summarised in Ensembl (release 87) (Yates et al., 2016), using Fisher's exact tests in SPSS v24.0. The following reference populations were used: Luhya in Webuye, Kenya (LWK) and Yoruba in Ibadan, Nigeria (YRI) populations for the South African black population; Utah residents with Northern and Western European ancestry (CEU) for the South African Caucasian population and the Indian Telugu in the United Kingdom (ITU) for the South African Asian population. These reference populations were selected merely as representative of populations who share ancestry with the South African ethnic groups used in this study.

Association of alleles, genotypes and haplotypes with response to HBV vaccination

We compared SNP allele, genotype and haplotype frequencies in normal-responders to HBV vaccine versus low-responders to HBV vaccine. Both univariate and multivariate analyses were performed as described below.

For univariate analyses, SNP allele frequencies were compared in responders versus low-responders using Fisher's exact test in gPLINK v1.07 (Purcell et al., 2007). SNP genotype frequencies were also compared in responders versus low-responders using chi-squared tests and three different genotypic models in gPLINK: genotypic, dominant and recessive. The genotypic model assessed the association between HBV vaccine response and the three possible genotypes (aa vs Aa vs AA, where 'a' is the minor allele and 'A' is the major allele). The dominant model assessed the association between HBV vaccine response and having at least one minor allele versus not having any minor allele (aa, Aa vs AA), and the recessive model assessed the association between HBV vaccine response and having two minor alleles versus having at least one major allele (aa vs AA, Aa). We also compared haplotype frequencies in the cases and controls using gPLINK. For the HLA region, haplotypes involving

SNPs with D′ value >0.8 (as determined by Haploview), were tested for association with vaccine response using gPLINK, and haplotypes of the HLA region generated by Arlequin were tested for association with HBV vaccine response using Fisher's exact test in SPSS v24.0.

Those SNPs that were significant in the univariate analysis were carried forward to multivariate analyses taking significant demographic variables (weight and smoker status at time of vaccination) into account as covariates. Multivariate analysis was conducted using logistic regression in gPLINK taking significant demographic variables (weight and smoker status at time of vaccination) into account as covariates.

The empirical p-values were considered in the univariate and multivariate analyses. The empirical p-values (P_{EMP1}) were generated using a Monte Carlo-based method and ten thousand simulations in gPLINK. The empirical p-values corrected for multiple testing (P_{EMP2}) were calculated using the proportion of permutations in which any of the test statistics exceeded that particular observed statistic. P-values <0.05 were considered significant.

The effects (odds ratio (OR)) presented in the results is with respect to the minor allele for individual SNP association analysis. An OR of greater than one indicates that the minor allele increases the risk of low-response to the HBV vaccine relative to the major allele and *vice versa*. For haplotype association analysis, the OR is with respect to the alleles indicated.

3.3 *HLA-DPA1* and *–DPB1* Genotyping Results

The *HLA-DPA1* and *–DPB1* alleles were typed using DNA from 149 participants.

3.3.1 *HLA-DPA1*

The *HLA-DPA1* typing was successful for all samples (n=149). *HLA-DPA1* Luminex results were grouped as shown in **Table 3.9** for analysis purposes. We observed five alleles at 4-digit resolution, while the *HLA-DPA1*02:02:01* and the ambiguous **02:02/02:04/02:05* alleles were grouped together as a low resolution '**02:other*' category (meaning **02* excluding **02:01* alleles). The frequencies of these alleles in the total South African cohort are shown in **Table 3.9** and the frequencies of the *HLA-DPA1* alleles in the different ethnicities are shown in **Table 3.10**. Allele *HLA-DPA1*01:03* was the most frequent allele in all ethnicities (58% of total cohort and 31-77% in different ethnicities), and *HLA-DPA1*02:01* alleles were also common (20% of total cohort and 17-24% in different ethnicities).

Table 3.9. *HLA-DPA1* allele groupings and frequencies in the cohort, n=149

Luminex Result	Allele	Count in Total Cohort (2n)	% in Total Cohort
01:03:03 *01:03:01G*	*01:03*	172	57.7
01:04	*01:04*	1	0.3
02:01 *02:01:01G* *02:01:02*	*02:01*	59	19.8
02:02:01 *02:02/02:04/02:05*	*02:other* (not *02:01*)	43	14.4
03:01	*03:01*	21	7
04:01	*04:01*	2	0.7
Total		**298**	**100**

Several *HLA-DPA1* allele frequencies differed significantly by ethnicity (**Table 3.10**). *HLA-DPA1*01:03* and *-DPA1*02:other* allele frequencies differed significantly between the South African ethnic groups ($p<0.0001$). Due to low frequencies of *HLA-DPA1*03:01* allele, the ethnicities were compared in a pairwise fashion using Fisher's exact test: frequencies differed significantly between African and Caucasian ($p<0.0001$), between African and Asian ($p=0.007$) and between Caucasian and Asian ($p=0.0003$) populations. These results remained significant after Bonferroni adjustment **(Table 3.10)**.

Table 3.10. *HLA-DPA1* allele frequencies according to ethnicity

Allele	African (count (%))	Caucasian (count (%))	Asian (count (%))	Coloured (count (%))	Chi-squared test p-value	Fisher's exact test p-value	Bonferroni corrected p-value
n	49	75	22	3			
2n	98	150	44	6			
01:03	30 (30.6)	115 (76.7)	26 (59.1)	1 (16.7)	<0.0001	NA	<0.0001
01:04	1 (1.0)	0	0	0	NA	*0.395 **1.000	1.000 1.000
02:01	23 (23.5)	26 (17.3)	9 (20.5)	1 (16.7)	0.493	NA	1.000
02:other **(not *02:01*)**	25 (25.5)	6 (4)	8 (18.2)	4 (66.7)	<0.0001	NA	<0.0001
03:01	18 (18.4)	2 (1.3)	1 (2.3)	0 (0)	NA	*<0.0001 **0.007 ***0.0003	<0.0001 0.042 0.002
04:01	1 (1)	1 (0.7)	0	0	NA	*1.000 **1.000 ***1.000	1.000 1.000 1.000

P-values assessed by chi-squared test (excluding Coloured) or Fisher's exact test
*African vs Caucasian
**African vs Asian
***Caucasian vs Asian
Abbreviation: NA, not applicable

A comparison of the *HLA-DPA1* allele frequencies observed in this study to the allele frequencies from reference populations is shown in **Table 3.11**. Allele frequencies from several reference populations were grouped into 'superpopulations' (sub-Saharan black, Caucasian and Indian) to provide a broad view of genetic affinities of South African ethnic groups. Specific reference populations (for example, only Kenya) could alternatively have been used here to confirm which global populations are most similar to South African population groups. The reference population frequencies were obtained from the AFND as described previously in section 3.2.2.2. There were no significant differences in the allele frequencies in this cohort compared to the reference populations, as assessed by Fisher's exact test and therefore, Bonferroni corrections for multiple comparisons were not performed for these p-values.

Table 3.11. *HLA-DPA1* allele frequencies in this study compared to reference populations

DPA1 Alleles	SA black[a]	Sub-Saharan, black[b]	p-value	SA white[a]	European, Caucasoid[c]	p-value	SA Asian/ Indian[a]	Indian[d]	p-value
01:03	0.306	0.290	0.877	0.767	0.842	0.284	0.591	0.593	1.000
01:04	0.010	0.001	1.000	0.000	0.004	1.000	0.000	0.025	0.246
02:01	0.235	0.338	0.121	0.173	0.127	0.553	0.205	0.212	1.000
02:other	0.255	0.177	0.233	0.040	0.070	0.537	0.182	0.161	0.851
03:01	0.184	0.174	1.000	0.013	0.002	1.000	0.023	0.000	0.497
04:01	0.010	0.022	1.000	0.007	0.001	1.000	0.000	0.008	1.000
Total	1.000	1.002		1.000	1.046		1.001	0.999	

Reference allele frequencies labelled in blue
[a]Allele frequencies observed in this study
[b]Average allele frequencies across sub-Saharan black African populations (Cameroon, Gabon, Gambia, Kenya, Liberia, Nigeria, Uganda) in AFND
[c]Average allele frequencies across Caucasian populations from France, Norway, Sweden, UK (AFND)
[d]Allele frequencies from a population in Bombay, India (AFND)
Abbreviations: SA, South African

There were no significant associations between *HLA-DPA1* alleles and response to the HBV vaccine as determined by Fisher's exact test ($p>0.05$), as seen in **Table 3.12**. As there were no significant associations, Bonferroni corrections for multiple comparisons were not performed for these p-values.

Table 3.12. Frequency of *HLA-DPA1* alleles in responders versus low-responders

HLA-DPA1 Allele	Count in Responders (2n=282)	Count in Low-responders (2n=16)	p-value
01:03	165	7	0.30
01:04	1	0	1.00
02:01	53	6	0.10
02:other	40	3	0.71
03:01	21	0	0.61
04:01	2	0	1.00
Total	**282**	**16**	

3.3.2 *HLA-DPB1*

HLA-DPB1 allele assignment failed for 1 sample (genotyped sample size was therefore n=148). Additionally, there were two cases where an unambiguous *HLA-DPB1* allele could not be identified from the ambiguous string of alleles (*HLA-DPB1*131:01/241:01/343:01/361:01/407:01/447:01*), even as a G group. This result was confirmed by repeating the typing.

We observed 25 different *HLA-DPB1* alleles, the *HLA-DPB1* allele frequencies in the different ethnicities are shown in **Table 3.13**. Allele *HLA-DPB1*04:01:01G* was the most common allele in the Caucasian (41.3%) and Asian (45.5%) populations. However, *HLA-DPB1*01:01* was the most frequent allele in the African (27.1%) and Coloured (50%) populations. This was followed closely by *HLA-DPB1*105:01* in the African population (26%).

There were statistically significant differences in *HLA-DPB1* allele frequencies across the different ethnicities. *HLA-DPB1*01:01*, **04:01:01G* and **105:01* allele frequencies in the South African black population were significantly different to the other populations ($p<0.0001$), and these remained significant after Bonferroni adjustment for multiple comparisons ($p<0.0001$). The allele frequencies of *HLA-DPB1*04:02:1G* and **13:01:01G* significantly differed between the black African and Caucasian ($p<0.0001$ and $p=0.035$ respectively) and between Caucasian and Asian ($p=0.009$ and $p=0.037$ respectively) populations. *HLA-DPB1*05:01:01G* and **18:01* allele frequencies only differed significantly between black and Caucasian populations ($p=0.045$ and $p=0.003$ respectively). While *HLA-DPB1*09:01* allele frequencies significantly differed between black and Asian ($p=0.009$), Caucasian and Asian ($p=0.002$). However, only the difference in *HLA-DPB1*04:02:01G* allele frequencies between the black African and Caucasian populations remained significant after Bonferroni adjustment for multiple comparisons ($p=0.001$) (**Table 3.13**).

Table 3.13. *HLA-DPB1* frequencies according to SA ethnicities

DPB1 Allele	African (n (%))	Caucasian (n (%))	Asian (n (%))	Coloured (n (%))	Total cohort (n (%))	Chi-squared test p-value	Fisher's exact test p-value	Bonferroni corrected p-value
n	48	75	22	3	148			
2n	96	150	44	6	296			
01:01	26 (27.1)	7 (4.7)	1 (2.3)	3 (50)	37 (12.5)	<0.0001	NA	<0.0001
02:01:02G	10 (10.4)	24 (16.0)	7 (15.9)	0 (0)	41 (13.9)	0.369	NA	1.000
02:02	0 (0)	0 (0)	1 (2.3)	0 (0)	1 (0.3)	NA	**0.314 ***0.227	1.000 1.000
03:01:01G	6 (6.3)	9 (6.0)	2 (4.5)	0 (0)	17 (5.7)	0.933	NA	1.000
04:01:01G	4 (4.2)	62 (41.3)	20 (45.5)	1 (16.7)	87 (29.4)	<0.0001	NA	<0.0001
04:02:01G	0 (0)	20 (13.3)	0 (0)	0 (0)	20 (6.8)	NA	*<0.0001 ***0.009	0.001 0.225
05:01:01G	0 (0)	7 (4.7)	2 (4.5)	0 (0)	9 (3.0)	NA	*0.045 **0.097 ***1.000	1.000 1.000 1.000
06:01	0 (0)	2 (1.3)	0 (0)	0 (0)	2 (0.7)	NA	*0.522 ***1.000	1.000 1.000
09:01	0 (0)	0 (0)	4 (9.1)	1 (16.7)	5 (1.7)	NA	**0.009 ***0.002	0.225 0.050
10:01	0 (0)	5 (3.3)	0 (0)	0 (0)	5 (1.7)	NA	*0.160 ***0.590	1.000 1.000
11:01:01	4 (4.2)	4 (2.7)	0 (0)	0 (0)	8 (2.7)	NA	*0.715 **0.308 ***0.576	1.000 1.000 1.000
13:01:01G	5 (5.2)	1 (0.7)	3 (6.8)	0 (0)	9 (3.0)	NA	*0.035 **0.706 ***0.037	0.875 1.000 0.925
14:01	1 (1.0)	1 (0.7)	0 (0)	0 (0)	2 (0.7)	NA	*1.000 **1.000 ***1.000	1.000 1.000 1.000
15:01	1 (1.0)	0 (0)	0 (0)	0 (0)	1 (0.3)	NA	*0.390 **1.000	1.000 1.000
16:01	0 (0)	1 (0.7)	0 (0)	0 (0)	1 (0.3)	NA	*1.000 ***1.000	1.000 1.000
17:01:01G	1 (1.0)	3 (2.0)	1 (2.3)	0 (0)	5 (1.7)	NA	*1.000 **0.531 ***1.000	1.000 1.000 1.000
18:01	6 (6.3)	0 (0)	0 (0)	0 (0)	6 (2.0)	NA	*0.003 **0.177	0.075 1.000
19:01:01G	1 (1.0)	0 (0)	0 (0)	0 (0)	1 (0.3)	NA	*0.390 **1.000	1.000 1.000
20:01:01	0 (0)	1 (0.7)	0 (0)	0 (0)	1 (0.3)	NA	*1.000 ***1.000	1.000 1.000
26:01:02	0 (0)	1 (0.7)	2 (4.5)	0 (0)	3 (1.0)	NA	*1.000 **0.097 *0.129	1.000 1.000 1.000
34:01	1 (1.0)	0 (0)	0 (0)	0 (0)	1 (0.3)	NA	*0.390 **1.000	1.000 1.000
58:01	1 (1.0)	0 (0)	0 (0)	0 (0)	1 (0.3)	NA	*0.390 **1.000	1.000 1.000
105:01	25 (26.0)	2 (1.3)	1 (1.3)	1 (16.7)	29 (9.8)	<0.0001	NA	<0.0001
106:01	2 (2.1)	0 (0)	0 (0)	0 (0)	2 (0.7)	NA	*0.151 **1.000	1.000 1.000
131:01	2 (2.1)	0 (0)	0 (0)	0 (0)	2 (0.7)	NA	*0.151 **1.000	1.000 1.000

P-values assessed by chi-squared test (excluding Coloured) or Fisher's exact test
*African vs Caucasian
**African vs Asian
***Caucasian vs Asian
Abbreviations: SA, South African; NA, not applicable

The *HLA-DPB1* allele frequencies observed in this study were compared to those in reference populations, shown in **Table 3.14**. There were significant differences between the *HLA-DPB1*04:02:01G* and the **17:01:01G* allele frequencies in the South African black population of this study compared to the sub-Saharan black reference populations, but only the difference between the *HLA-DPB1*04:02:01G* allele frequencies remained significant after Bonferroni correction for multiple comparisons ($p<0.0001$). The difference observed with regards to the **04:02* allele frequencies (0.0% in our cohort versus 18.6% in other cohorts) was due to the non-reporting of the *HLA-DPB1*105* allele in the reference black populations. There is a single base difference in exon 3 between alleles **04:02* and **105* which is not detected in standard sequence-based typing of *HLA-DPB1* which uses only exon 2 data, and was however, detected in our study using Luminex SSO genotyping. The *HLA-DPB1*04:02:01G* allele frequencies also significantly differed between the South African Asian population and the reference Indian population, however this did not remain significant after Bonferroni correction for multiple comparisons ($p=0.396$).

Table 3.14. Comparison between *HLA-DPB1* allele frequencies in this study cohort and reference populations

DPB1 Alleles	SA black[a]	Sub-Saharan, black[b]	p-value	SA white[a]	European, Caucasoid[c]	p-value	SA Asian/ Indian[a]	Indian[d]	p-value
01:01	0.271	0.311	0.640	0.047	0.049	1.000	0.023	0.041	0.683
02:01:02G	0.104	0.112	1.000	0.160	0.126	0.689	0.159	0.195	0.582
02:02	0.000	0.002	1.000	0.000	0.006	1.000	0.023	0.000	0.497
03:01:01G	0.063	0.051	1.000	0.060	0.110	0.311	0.045	0.045	1.000
04:01:01G	0.042	0.046	1.000	0.413	0.429	0.886	0.455	0.359	0.197
04:02:01G	0.000	0.186	<0.0001	0.133	0.121	1.000	0.000	0.066	0.014
05:01:01G	0.000	0.000	-	0.047	0.026	0.721	0.045	0.009	0.212
06:01	0.000	0.003	1.000	0.013	0.018	1.000	0.000	0.000	-
09:01	0.000	0.001	1.000	0.000	0.011	1.000	0.091	0.032	0.134
10:01	0.000	0.011	1.000	0.033	0.015	0.683	0.000	0.011	1.000
11:01:01	0.042	0.012	0.369	0.027	0.021	1.000	0.000	0.002	1.000
13:01:01G	0.052	0.040	1.000	0.007	0.016	1.000	0.068	0.075	1.000
14:01	0.010	0.005	1.000	0.007	0.015	1.000	0.000	0.029	0.246
15:01	0.010	0.012	1.000	0.000	0.009	1.000	0.000	0.011	1.000
16:01	0.000	0.007	1.000	0.007	0.005	1.000	0.000	0.000	-
17:01:01G	0.010	0.091	0.018	0.020	0.009	1.000	0.023	0.024	1.000
18:01	0.063	0.049	1.000	0.000	0.001	1.000	0.000	0.005	1.000
19:01:01G	0.010	0.016	1.000	0.000	0.006	1.000	0.000	0.000	-
20:01:01	0.000	0.003	1.000	0.007	0.005	1.000	0.000	0.000	-
26:01:02	0.000	0.006	1.000	0.007	0	1.000	0.045	0.053	1.000
34:01	0.010	0.010	1.000	0.000	0	-	0.000	0.000	-
58:01	0.010	0.000	1.000	0.000	0	-	0.000	0.000	-
105:01	0.26	NR	-	0.013	NR	-	0.013	NR	-
106:01	0.021	NR	-	0.000	NR	-	0.000	NR	-
131:01 **(ambiguous)**	0.021	NR	-	0.000	NR	-	0.000	NR	-
Total	0.999	0.974		1.001	0.998		0.990	0.957	

Reference allele frequencies labelled in blue
[a]Allele frequencies observed in this study
[b]Average allele frequencies across sub-Saharan black African populations (Cameroon, Gabon, Gambia, Kenya, Liberia, Nigeria, Uganda, Zimbabwe) in AFND
[c]Average allele frequencies across Caucasian populations from France, Norway, Sweden, UK (AFND)
[d]Allele frequencies from a population in Bombay, India (AFND)
Abbreviations: SA, South African; NR, not reported

We found that allele *HLA-DPB1*04:01:01G* was significantly more frequent in the normal responders compared to the low-responders (p=0.004), whereas *HLA-DPB1*01:01* and **09:01* were significantly more frequent in the low-responders compared to the normal responders (p=0.036 and 0.025 respectively, **Table 3.15**). After Bonferroni correction for multiple comparisons, the significance of these associations was lost.

Table 3.15. *HLA-DPB1* allele frequencies in responders and low-responders

Allele	Count in Responders (%)	Count in Low-responders (%)	p-value	Bonferroni corrected p-value
n	140	8		
2n	280	16		
01:01 (01:01:01G and 01:01:02G)	32 (11.4)	5 (31.3)	**0.036**	0.9
02:01:02G	38 (13.6)	3 (18.8)	0.472	1.0
02:02	1 (0.4)	0 (0)	1.000	1.0
03:01:01G	17 (6.1)	0 (0)	0.610	1.0
04:01:01G	87 (31.1)	0 (0)	**0.004**	0.1
04:02:01G	17 (6.1)	3 (18.8)	0.084	1.0
05:01:01G	9 (3.2)	0 (0)	1.000	1.0
06:01	2 (0.7)	0 (0)	1.000	1.0
09:01	3 (1.1)	2 (12.5)	**0.025**	0.625
10:01	5 (1.8)	0 (0)	1.000	1.0
11:01:01	7 (2.5)	1 (6.3)	0.362	1.0
13:01:01G	9 (3.2)	0 (0)	1.000	1.0
14:01	2 (0.7)	0 (0)	1.000	1.0
15:01	1 (0.4)	0 (0)	1.000	1.0
16:01	1 (0.4)	0 (0)	1.000	1.0
17:01:01G	5 (1.8)	0 (0)	1.000	1.0
18:01	5 (1.8)	1 (6.3)	0.286	1.0
19:01:01G	1 (0.4)	0 (0)	1.000	1.0
20:01:01	1 (0.4)	0 (0)	1.000	1.0
26:01:02	3 (1.1)	0 (0)	1.000	1.0
34:01	1 (0.4)	0 (0)	1.000	1.0
58:01	1 (0.4)	0 (0)	1.000	1.0
105:01	29 (10.4)	0 (0)	0.382	1.0
106:01	1 (0.4)	1 (6.3)	0.105	1.0
131:01	2 (0.7)	0 (0)	1.000	1.0
Total	**280 (100.5)**	**16 (100.3)**		

The frequencies of the three alleles that were significantly associated with HBV vaccine response in this cohort (*HLA-DPB1*01:01, *04:01:01G* and **09:01*) in the different ethnicities are shown graphically in **Figure 3.1**.

*HLA-DPB1*01:01* which associated with susceptibility to chronic HBV in the literature (Thomas et al., 2012), and here associated with poor vaccine response, was significantly more frequent in black African individuals compared to Asian and Caucasian individuals ($p<0.0001$), as seen in **Figure 3.1** (also refer to **Table 3.13**).. These results remained significant after Bonferroni correction for multiple comparisons. *HLA-DPB1*04:01*, associated with clearance of HBV infection (Kamatani et al., 2009; Thomas et al., 2012), as well as with detectable anti-HBs levels post booster vaccination (Wu et al., 2013) in the literature and here associated with good vaccine response, was significantly more frequent in Asian and Caucasian individuals compared to African individuals ($p<0.0001$), even after Bonferroni adjustment for multiple comparisons ($p<0.0001$), refer also to **Table 3.13**. *HLA-DPB1*09:01*, previously associated with undetectable anti-HBs following booster vaccination (Wu et al., 2013) and here associated with poor vaccine response, was found only in Asian and Coloured individuals in our cohort. *HLA-DPB1*09:01* was significantly more frequent in Asian individuals compared to African and Caucasian individuals ($p=0.009$ and 0.002, respectively) (**Table 3.13)**; although only the difference between Asian and Caucasian individuals remained significant after adjustment for multiple comparisons ($p=0.012$).

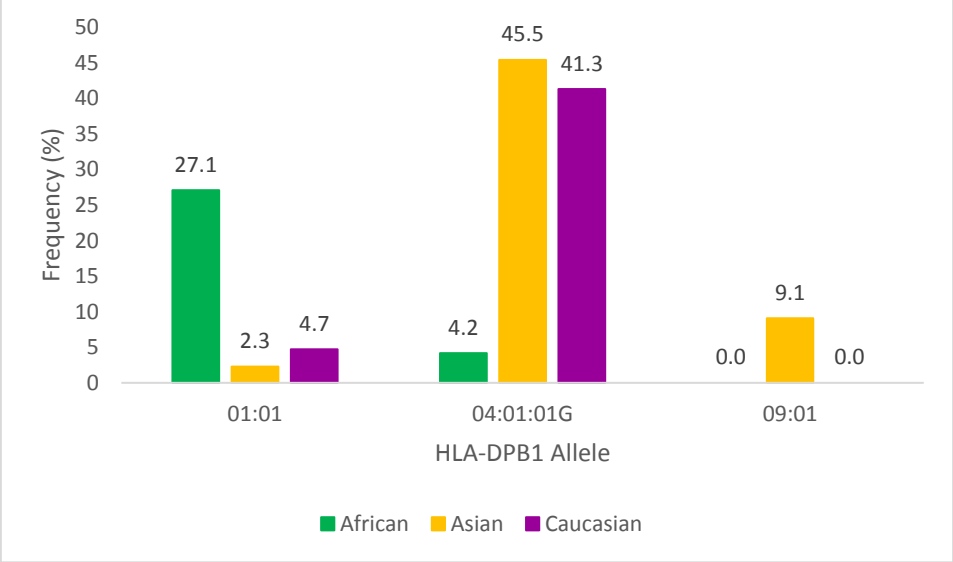


Figure 3.1. Frequencies of *HLA-DPB1*01:01, *04:01* and **09:01* according to ethnicity

3.4 Results of SNP Genotyping

3.4.1 MassArray Analysis

All 24 SNPs that passed the design stage were successfully genotyped in the 149 participants. Examples of the electropherograms produced by MassArray Analysis from which the genotypes are determined are shown in **Figure 3.2**.

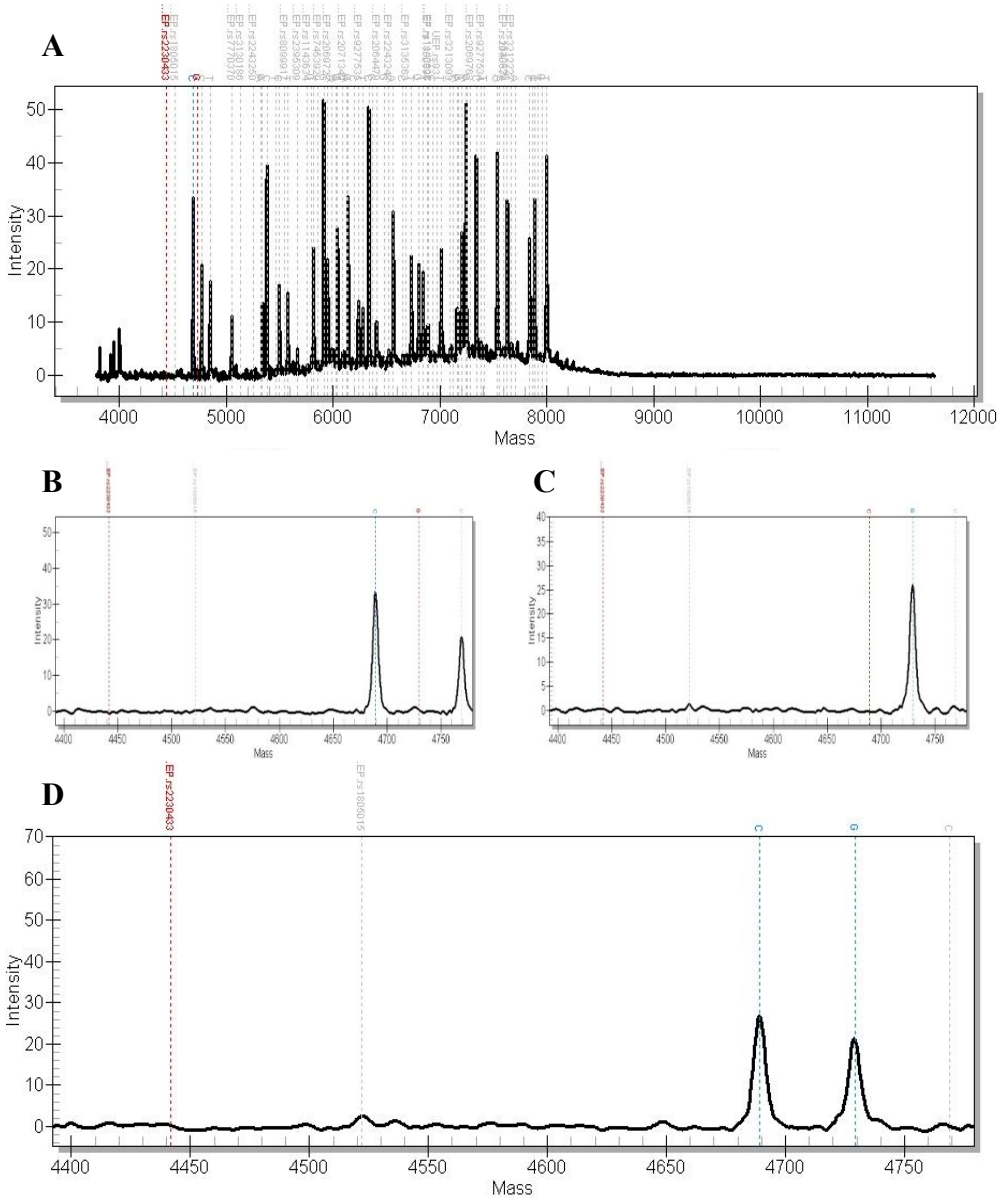


Figure 3.2. **Example electropherograms produced by MassArray Analysis.** (**A**) All SNP calls in a multiplex, SNP rs2230433 indicated in red. (**B**) Homozygous C call for rs2230433. (**C**) Homozygous G call for rs2230433. (**D**) Heterozygous GC call for rs2230433.

All SNPs passed the MAF QC threshold of 0.05 (see **Table 3.16**) and were in HWE (p>0.001) (Appendix E), thus all 24 SNPs were included in the analyses.

3.4.2 Allele frequencies within South African ethnicities

The MAFs of the SNPs in the total cohort and within the different ethnicities are shown in **Table 3.16**. Due to frequency differences in different populations, there were seven cases where the allele deemed to be the minor allele in the whole cohort was actually the major allele in a particular ethnicity.

Most of the SNP allele frequencies (18/24) differed between the ethnicities, as determined by chi-square test, and 14 of these remained significant after Bonferroni correction for multiple comparisons (**Table 3.16**). The allele frequencies of the *IL4R* and *ITGAL* SNPs significantly differed between all three ethnicities. The allele frequencies of the *IL2* and *IL4* SNPs as well as three of the *HLA-DPB1* SNPs (rs7770370, rs931, rs9277534) significantly differed between the African and Caucasian and between the African and Asian populations. The *IL10* SNP, rs1143627 of *IL1B*, the *HLA-DRA* SNP and rs2395309 of *HLA-DPA1* differed significantly between the Caucasian and African populations and between the Caucasian and Asian populations. The allele frequencies of rs1143633 of *IL1B*, the *HLA-DQB2* SNP, rs2064479 of *HLA-DPB1*, and the *IFNG* and *IL28B* SNPs differed significantly between the African and Caucasian populations.

Table 3.16. MAF of 24 SNPs in the cohort

SNP ID	Gene	Chromosome	Position (bp)	Minor allele[a]	Major allele[a]	MAF				p-value	Bonferroni corrected p-value
						Total Cohort, n=149	African, n=49	Caucasian, n=75	Asian, n=22		
rs1800896	*IL10*	1	206773552	C	T	0.40	0.32	0.49	0.23	**0.0005**	**0.011**
rs1143634	*IL1B*	2	112832813	A	G	0.18	0.12	0.22	0.16	0.161	1.000
rs1143633	*IL1B*	2	112832890	T	C	0.27	0.15	0.35	0.23	**0.004**	0.096
rs1143627	*IL1B*	2	112836810	A	G	0.48	0.28	0.64*	0.41	**<0.0001**	**<0.0001**
rs2069763	*IL2*	4	122456327	A	C	0.23	0.04	0.30	0.43	**<0.0001**	**<0.0001**
rs2243248	*IL4*	5	132672952	G	T	0.20	0.40	0.11	0.05	**<0.0001**	**<0.0001**
rs2243250	*IL4*	5	132673462	T	C	0.36	0.67*	0.17	0.27	**<0.0001**	**<0.0001**
rs2070874	*IL4*	5	132674018	T	C	0.32	0.55*	0.17	0.27	**<0.0001**	**<0.0001**
rs3212227	*IL12B*	5	159315942	G	T	0.26	0.26	0.23	0.34	0.199	1.000
rs3213093	*IL12B*	5	159323971	T	C	0.26	0.28	0.23	0.34	0.225	1.000
rs3135363	*HLA-DRA*	6	32421871	C	T	0.22	0.12	0.32	0.09	**<0.0001**	**0.001**
rs7453920	*HLA-DQB2*	6	32762235	A	G	0.29	0.17	0.38	0.25	**0.003**	0.072
rs2395309	*HLA-DPA1*	6	33058469	G	A	0.35	0.50	0.22	0.39	**0.0002**	**0.005**
rs2071349	*HLA-DPA1*	6	33075743	G	C	0.13	0.11	0.13	0.16	0.578	1.000
rs7770370	*HLA-DPB1*	6	33081144	G	A	0.46	0.69*	0.33	0.34	**<0.0001**	**<0.0001**
rs931	*HLA-DPB1*	6	33086773	A	G	0.36	0.51*	0.25	0.32	**0.0003**	**0.009**
rs9277534	*HLA-DPB1*	6	33087030	G	A	0.37	0.54*	0.27	0.32	**0.0001**	**0.004**
rs9277535	*HLA-DPB1*	6	33087084	G	A	0.22	0.22	0.21	0.25	0.781	1.000
rs3130186	*HLA-DPB1*	6	33088430	T	C	0.22	0.22	0.21	0.25	0.781	1.000
rs2064479	*HLA-DPB1*	6	33104463	T	C	0.33	0.46	0.24	0.30	**0.003**	0.072
rs2069727	*IFNG*	12	68154443	C	T	0.32	0.22	0.39	0.32	**0.033**	0.792
rs1805015	*IL4R*	16	27362859	C	T	0.24	0.39	0.19	0.07	**<0.0001**	**<0.0001**
rs2230433	*ITGAL*	16	30506720	C	G	0.42	0.49	0.32	0.68*	**<0.0001**	**<0.0001**
rs8099917	*IL28B*	19	39252525	G	T	0.13	0.04	0.20	0.09	**0.001**	**0.024**

[a]Minor and major alleles determined for the entire cohort

*as seen from these frequencies, the so-called minor allele is actually the major allele in this ethnicity

Abbreviations: MAF, minor allele frequency

3.4.3 Comparison of allele frequencies to other populations

Allele frequencies of the 24 SNPs in our South African ethnicities were compared to data from Ensembl (release 87) (Yates et al., 2016). The South African black population was compared to the average allele frequencies of the LWK and YRI populations, the South African Caucasian population was compared to the CEU population and the South African Asian population was compared to the ITU population, as seen in **Table 3.17**. As determined by Fisher's exact test, the allele frequencies of rs2243248 and rs2243250, both of the *IL4* gene, differed significantly between our South African black population and the reference populations (p=0.001 and p=0.023 respectively). The *HLA-DPB1* intron variant, rs7770370, differed significantly between the South African Caucasian population and the CEU population (p=0.014). Three SNPs - rs2069763 in *IL2*, rs3135363 upstream of *HLA-DRA* and rs7770370 in *HLA-DPB1* - significantly differed in allele frequencies between the South African Asian population and the ITU population (p= 0.038, p=0.043 and p=0.015, respectively). However, only the difference in allele frequency of rs2243248 between the South African black and reference populations remained significant after Bonferroni correction for multiple comparisons.

Table 3.17. Comparison of allele frequencies in SA populations to reference populations

SNP ID	Allele	Allele Frequency		p-value	Allele Frequency		p-value	Allele Frequency		p-value
		SA black[a] (n=49)	LWK/YRI		SA white[a] (n=75)	CEU		SA Asian/Indian[a] (n=22)	ITU	
rs1800896	C	0.32	0.31	1.000	0.49	0.52	0.777	0.23	0.24	1.000
rs1143634	A	0.12	0.11	1.000	0.22	0.23	1.000	0.16	0.13	0.689
rs1143633	T	0.15	0.17	0.847	0.35	0.37	0.883	0.23	0.24	1.000
rs1143627	A	0.28	0.32	0.644	0.64	0.65	1.000	0.41	0.38	0.772
rs2069763	A	0.04	0.10	0.164	0.30	0.40	0.182	0.43	0.28	0.038
rs2243248	G	0.40	0.17	0.001	0.11	0.09	0.814	0.05	0.13	0.081
rs2243250	T	0.67	0.82	0.023	0.17	0.13	0.553	0.27	0.18	0.175
rs2070874	T	0.55	0.53	0.887	0.17	0.13	0.553	0.27	0.18	0.175
rs3212227	G	0.26	0.38	0.095	0.23	0.19	0.603	0.34	0.42	0.308
rs3213093	T	0.28	0.38	0.176	0.23	0.19	0.603	0.34	0.42	0.308
rs3135363	C	0.12	0.19	0.241	0.32	0.19	0.051	0.09	0.20	0.043
rs7453920	A	0.17	0.28	0.090	0.38	0.46	0.316	0.25	0.24	1.000
rs2395309	G	0.50	0.59	0.256	0.22	0.18	0.596	0.39	0.36	0.770
rs2071349	G	0.11	0.12	1.000	0.13	0.12	1.000	0.16	0.28	0.060
rs7770370	G	0.69	0.63	0.456	0.33	0.17	0.014	0.34	0.18	0.015
rs931	A	0.51	0.57	0.478	0.25	0.27	0.872	0.32	0.37	0.552
rs9277534	G	0.54	0.62	0.316	0.27	0.28	1.000	0.32	0.38	0.459
rs9277535	G	0.22	0.17	0.476	0.21	0.24	0.735	0.25	0.30	0.527
rs3130186	T	0.22	0.17	0.476	0.21	0.24	0.735	0.25	0.30	0.527
rs2064479	T	0.46	0.53	0.396	0.24	0.21	0.735	0.30	0.33	0.761
rs2069727	C	0.22	0.14	0.197	0.39	0.43	0.666	0.32	0.41	0.240
rs1805015	C	0.39	0.48	0.254	0.19	0.17	0.854	0.07	0.07	1.000
rs2230433	C	0.49	0.43	0.478	0.32	0.29	0.759	0.68	0.77	0.205
rs8099917	G	0.04	0.06	0.748	0.20	0.15	0.457	0.09	0.18	0.097

Reference alleles labelled in blue

[a]Allele frequencies observed in this study

Abbreviations: LWK, Luhya in Webuye, Kenya; SA, South African; YRI, Yoruba in Ibadan, Nigeria; CEU, Utah residents with Northern and Western European ancestry; ITU, Indian Teluga in the UK

3.4.4 Genotype frequencies

The genotype frequencies of the 24 SNPs in the total cohort and within the South African ethnicities can be seen in Appendix F.

3.4.5 Haplotype frequencies

3.4.5.1 Haplotype frequencies in cytokine genes

Haplotype analysis was performed using the sliding window approach for the chromosomes for which two or more SNPs were genotyped, that is, chromosome 2 (*IL1B*), chromosome 5 (*IL4* and *IL12B*) and chromosome 16 (*IL4R* and *ITGAL*). The frequencies of these haplotypes in the different South African ethnicities are shown in **Table 3.18**.

Table 3.18. Chromosome 2, 5 and 16 haplotype frequencies

Haplotype					Frequency (%)			
					Total cohort	SA black	SA Caucasian	SA Asian
Chromosome 2								
		rs1143634	**rs1143633**	**rs1143627**				
		G	T	A	23.40	6.58	34.30	22.70
		A	C	A	13.10	1.66	20.10	13.60
		G	C	A	11.80	19.30	9.61	4.61
		G	T	G	2.64	8.22	1.07	0.00
		A	C	G	4.56	10.10	1.88	2.34
		G	C	G	44.00	53.60	33.10	56.80
Chromosome 5								
rs2243248	**rs2243250**	**rs2070874**	**rs3212227**	**rs3213093**				
G	T	T	G	T	1.00	2.00	0.00	1.27
T	T	T	G	T	5.80	11.00	0.00	10.60
T	T	C	G	T	1.20	3.80	0.00	0.00
G	C	C	G	T	3.60	3.70	3.60	3.28
T	C	C	G	T	13.80	5.00	19.70	19.00
G	T	T	T	C	3.40	9.60	0.00	0.00
T	T	T	T	C	19.60	25.40	16.70	15.40
G	C	T	T	C	1.90	6.20	0.00	0.00
G	T	C	T	C	3.30	7.00	0.00	0.00
G	C	C	T	C	6.20	10.40	7.70	0.00
T	C	C	T	C	38.60	7.40	52.30	50.50
T	T	C	T	T	0.00	1.00	0.00	0.00
G	T	T	T	T	0.00	1.00	0.00	0.00
T	T	T	T	C	0.00	6.60	0.00	0.00
Chromosome 16								
			rs1805015	**rs2230433**				
			C	C	9.00	16.50	6.02	4.18
			T	C	33.60	32.50	26.00	64.00
			C	G	15.20	22.30	12.60	2.64
			T	G	42.20	28.70	55.40	29.20

Linkage maps for the SNPs in the cytokine genes, based on the whole cohort, were generated using Haploview, shown in **Figure 3.3**. In the *IL1B* gene, rs1143634 was in complete LD with rs1143633, and rs1143633 was partially in LD with rs114627. In the *IL4* gene, rs2243250 was in near complete LD with rs2070874, while rs3212227 was in complete LD with rs3213093 in *IL12B*. There was no LD between the rs1805015 and rs2230433 SNPs on chromosome 16.

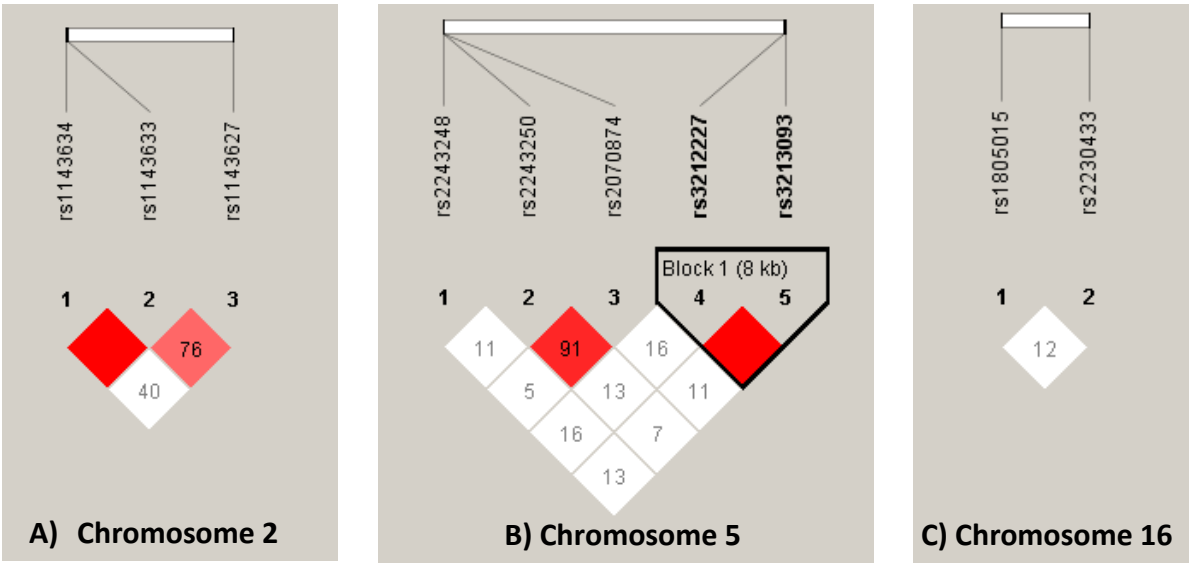


Figure 3.3. LD maps generated by Haploview for SNPs in chromosomes 2, 5, and 16. Numbers within the blocks represent D′ values. A) Chromosome 2, all SNPs located within IL1B gene region. B) Chromosome 5, SNPs located within IL4 and IL12B gene regions. C) Chromosome 16, SNPs located within the IL4R and ITGAL genes.

3.4.5.2 Haplotypes frequencies of *HLA* class II region

Haplotypes of the *HLA* region, including ten SNPs (one in *HLA-DR*, one in *–DQ* and eight in *–DPA1* or *–DPB1* genes), and the *HLA-DPA1* and *-DPB1* alleles were generated using Arlequin v3.5. The ten most frequent haplotypes of the *HLA* class II region SNPs and the *HLA-DP* alleles, in the total cohort are shown in **Table 3.19**. The complete set of these haplotypes and their frequencies are shown in Appendix G.

The ten most frequent haplotypes account for 45.5% of the total cohort. Five of these ten most frequent haplotypes differ significantly between the different South African ethnicities and four remained significant following Bonferroni correction for multiple comparisons (**Table 3.19**).

Table 3.19. Chromosome 6 SNP and *HLA-DP* allele haplotype frequencies

Haplotype												Frequency (%)					
rs3135363	**rs7453920**	**rs2395309**	**rs2071349**	**rs7770370**	**rs931**	**rs9277534**	**rs9277535**	**rs3130186**	**rs2064479**	***HLA-DPA1***	***HLA-DPB1***	**Total cohort**	**SA black**	**SA Caucasian**	**SA Indian**	**p-value**	**Bonferroni corrected p-value**
T	G	A	C	A	G	A	A	C	C	*01:03*	*04:01*	10.7	0.0	10.7	36.1	**<0.0001**	**<0.0001**
T	A	A	C	A	G	A	A	C	C	*01:03*	*04:01*	7.7	0.0	15.3	0.0	**<0.0001**	**<0.0001**
T	G	G	C	G	A	G	A	C	T	*02:other*	*01:01*	6.0	15.2	0.0	0.0	**<0.0001**	**<0.0001**
T	G	A	C	A	A	G	G	T	T	*01:03*	*03:01*	4.0	6.1	3.3	2.8	0.458	1.000
T	G	G	C	G	A	G	A	C	T	*02:01*	*01:01*	3.7	5.1	2.0	2.8	0.485	1.000
C	G	A	C	A	G	A	A	C	C	*01:03*	*04:01*	3.0	0.0	6.0	0.0	**0.002**	**0.020**
T	G	A	G	A	G	A	A	C	C	*01:03*	*02:01*	2.7	3.0	6.7	5.6	0.424	1.000
C	A	A	C	A	G	A	A	C	C	*01:03*	*04:01*	2.7	0.0	1.3	0.0	0.367	1.000
C	A	A	C	G	G	A	A	C	C	*01:03*	*04:01*	2.7	0.0	4.7	5.6	0.054	0.540
T	A	A	G	A	G	A	A	C	C	*01:03*	*02:01*	2.3	4.0	0.0	0.0	**0.017**	0.170

The linkage map for the *HLA* region SNPs, based on the whole cohort, as generated by Haploview, is shown in **Figure 3.4**. The *HLA-DPA1* SNPs rs2395309 and rs2071349 were in complete LD. In the *HLA-DPB1* 3′ untranslated region (UTR), rs931 was in complete LD with rs9277534, and rs9277534, rs9277535 and rs3130186 were in complete LD with one another. There were also several SNPs in partial LD with one another (**Figure 3.4**).

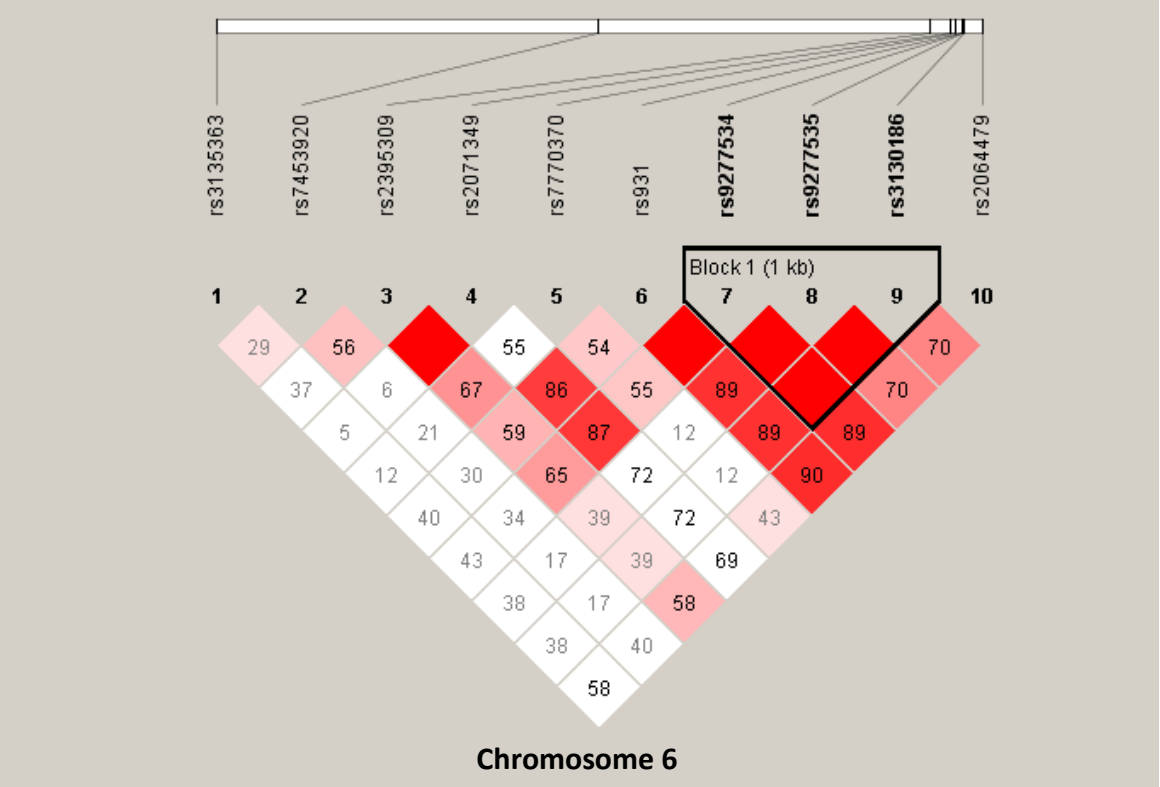


Figure 3.4. LD map of chromosome 6 SNPs generated by Haploview. Numbers within the blocks represent D′ values.

3.4.6 Association of alleles, genotypes and haplotypes with antibody response to HBV vaccination

Only significant results are shown in this section (the complete results set is shown in Appendix H). SNPs that were significant at the univariate level were then analysed in multivariate analysis taking weight and 'smoker during vaccination period' status into account using logistic regression. Results from both univariate and multivariate analyses are shown in a **Table 3.20**.

Significant SNP associations using the four models (allelic, genotypic, dominant and recessive) are shown in **Table 3.20**.

Table 3.20. SNP associations with HBV vaccine response

Model	Gene	SNP ID	UNIVARIATE			MULTIVARIATE			Risk of LR
			P_{EMP1}	P_{EMP2}	OR	P_{EMP1}	P_{EMP2}	OR	
Allelic	*IL1B*	rs1143634	**0.038**	0.560	2.93	0.135	0.560	2.62	↑
	HLA-DPB1	rs931	**0.048**	0.360	3.23	**0.013**	0.071	2.16	↑
Genotypic	*IL1B*	rs1143634	**0.048**	0.599	-	0.052	0.407	-	
	IL4	rs2070874	**0.044**	0.512	-	0.111	0.350	-	
	HLA-DPB1	rs7770370	**0.010**	0.199	-	**0.027**	0.180	-	
	IFNG	rs2069727	**0.029**	0.421	-	0.209	0.059	-	
Dominant	*IL4*	rs2070874	**0.028**	0.329	-	**0.004**	**0.029**	12.08	↑
	IFNG	rs2069727	**0.019**	0.320	-	**0.030**	0.124	0.12	↓
Recessive	*HLA-DPB1*	rs7770370	**0.012**	0.117	-	**0.011**	**0.050**	8.52	↑

Abbreviations: P_{EMP1}, pointwise empirical p-value; P_{EMP2}, p-value corrected for multiple comparisons; OR, odds ratio; LR; low-response

3.4.6.1 Associations between *HLA* SNPs and vaccine response

SNP rs931 (located in the 3′ UTR of *HLA-DPB1*) was significantly associated with HBV vaccine response in our South African cohort during univariate allelic analysis. The A variant (minor allele) was associated with an increased risk of low-response to the HBV vaccine relative to the major allele (OR>1.0).

Genotypes of only one SNP in the *HLA* region were significantly associated with HBV response in cohort. SNP rs7770370, located in an *HLA-DPB1* intron, was significantly associated with HBV vaccine response in genotypic and recessive models of inheritance. These results indicate that the rs7770370 GG genotype was associated with an increased risk of low anti-HBs response. The GG genotype of rs7770370 was found at rather low frequencies <10% in the different ethnicities in our study cohort.

The *HLA-DPB1* SNPs rs931 and rs7770370 were still significant after multivariate analysis and rs7770370 remained significant even after correction for multiple comparisons in the recessive model of inheritance. This indicates a strong association of these two variants with HBV vaccine response.

From the *HLA* class II region haplotypes (**Table 3.19**), it appears that the rs931 A allele and the rs7770370 G allele are inherited with *HLA-DPB1*01:01*, all of which were associated with poor anti-HBs response in this South African cohort; whereas the rs931 G allele and rs7770370 A allele are inherited with *HLA-DPB1*04:01*, which were all associated with good anti-HBs response in this study.

3.4.6.2 Associations between *HLA* SNP haplotypes and vaccine response

Haplotypes including SNPs of the *HLA* region with an LD D′ value greater than 0.8 (as determined by Haploview) were tested for association with vaccine response. Six haplotypes were significantly associated after univariate analysis, and all of these haplotypes remained significant after multivariate analysis (**Table 3.21**). However, after adjustment for multiple comparisons, the significance of these haplotypes was lost (see Appendix H for the complete set of results).

Table 3.21. *HLA* region haplotypes significantly associated with vaccine response

Haplotype								Univariate			Multivariate			
rs2395309	**rs2071349**	**rs931**	**rs9277534**	**rs9277535**	**rs3130186**	**rs2064479**	**Haplotype Frequency**	P_{EMP1}	P_{EMP2}	OR	P_{EMP1}	P_{EMP2}	OR	**Risk of LR**
A	C						0.523	**0.042**	0.251	0.30	**0.045**	0.308	0.30	↓
	C	G					0.523	**0.032**	0.260	0.31	**0.029**	0.255	0.25	↓
	C		A				0.506	**0.031**	0.296	0.32	**0.028**	0.268	0.26	↓
		A	G				0.356	**0.013**	0.257	2.79	**0.006**	0.191	3.47	↑
		G	A	A	C		0.628	**0.021**	0.296	0.38	**0.011**	0.219	0.30	↓
		G	A	A	C	C	0.606	**0.012**	0.187	0.30	**0.005**	0.121	0.21	↓

Abbreviations: P_{EMP1}, pointwise empirical p-value; P_{EMP2}, p-value corrected for multiple comparisons; OR, odds ratio; LR, low-response

Four of the six significant haplotypes included SNPs rs931 and rs9277534, located within the 3′ UTR of *HLA-DPB1*. The haplotype involving the A allele of rs931 and the G allele of rs9277534 is the only haplotype associated with an increased risk of low anti-HBs response. rs931, which was also significant at an allelic level of analysis, has not previously been associated with HBV vaccine response. As mentioned previously, rs931 alleles seem to be linked to the *HLA-DPB1*01:01* and **04:01* alleles, this also seems to be the case with rs2395309, rs9277534 and rs2064479.

The frequencies of the haplotypes composed of the ten *HLA* region SNPs and the *HLA-DPA1* and *HLA-DPB1* alleles were compared between the low-responders and responders. Only two haplotypes were significantly associated with HBV vaccine response, and these are shown in **Table 3.22**.

Table 3.22. *HLA* class II region haplotypes associated with HBV vaccine response

Haplotype												Frequency (n, %)		p-value
rs3135363	**rs7453920**	**rs2395309**	**rs2071349**	**rs7770370**	**rs931**	**rs9277534**	**rs9277535**	**rs3130186**	**rs2064479**	***HLA-DPA1***	***HLA-DPB1***	**Responders**	**Low-responders**	
T	G	G	C	A	A	G	G	T	T	*02:01*	*09:01*	3 (1.1)	2 (12.5)	0.025
T	A	A	C	G	G	A	A	C	C	*01:03*	*04:02*	1 (0.4)	2 (12.5)	0.008

Both of the haplotypes are observed at significantly higher frequencies in low-responders, however their significance is lost after Bonferroni correction for multiple comparisons (p=1.000 and p=0.760, respectively). These haplotypes are only similar at SNPs rs3135363 and rs2071349; and only *HLA-DPB1*09:01*, present in the first haplotype, was previously associated with HBV vaccine response in this study. However, due to the large number of possible haplotypes and the small sample sizes in each, significance testing may not be very reliable.

3.4.6.3 Associations between cytokine SNPs and vaccine response

***IL1B*:** SNP rs1143634 (a synonymous variant in the *IL1B* gene) was significantly associated with HBV vaccine response in our South African cohort during univariate allelic analysis. The A variant (minor allele) was associated with an increased risk of low-response to the HBV vaccine relative to the major allele (OR>1.0).

IL4: rs2070874, located in the 5′ UTR of *IL4*, was significantly associated with HBV vaccine response in the genotypic and dominant models of inheritance. These results suggested that in our cohort having at least one T allele at rs2070874 was associated with low anti-HBs response. In our cohort, the T allele was much more frequent in black Africans (55%) than in other ethnicities (<30%; **Table 3.16**). The associations observed with the *IL4* SNP rs2070874 remained significant after multivariate analysis in the dominant model of inheritance. Moreover, *IL4* SNP rs2070874 stayed significant after adjustment for multiple comparisons, indicating a very significant result.

IFNG: the rs2069727 SNP which is located downstream of the *IFNG* gene was significantly associated with HBV vaccine response in the genotypic and dominant models of inheritance. Having at least one C allele at rs2069727 in our cohort was associated with an increased anti-HBs response. The associations observed with the *IFNG* rs2069727 remained significant after

multivariate analysis in the dominant model of inheritance. These results re-enforce the association described for the univariate analysis.

3.4.6.4 Associations between haplotypes in cytokine genes and vaccine response

Haplotype frequencies for *IL1B* on chromosome 2, *IL4* and *IL12B* on chromosome 5, and *IL4R* and *ITGAL* on chromosome 16 were compared between low-responders and normal-responders. Only significant results are shown here (see Appendix H for the full results set). Results from both univariate and multivariate analyses are shown in a single table, **Table 3.23**.

Table 3.23. Significant haplotype associations with vaccine response

Haplotype				Haplotype Frequency	Univariate			Multivariate			Risk of LR
					P_{EMP1}	P_{EMP2}	OR	P_{EMP1}	P_{EMP2}	OR	
Chromosome 2 *IL1B*											
rs1143634	**rs1143633**	**rs1143627**									
A	C			0.181	**0.039**	0.457	3.48	0.151	0.857	2.62	↑
A	C	G		0.046	**0.009**	0.071	10.5	**0.006**	**0.047**	19.4	↑
Chromosome 5 *IL4*			*IL12B*								
rs2243250	**rs2070874**	**rs3212227**	**rs3213093**								
T	T	T		0.231	**0.029**	0.564	2.79	**0.010**	0.187	4.35	↑
T	T	T	C	0.227	**0.029**	0.526	2.84	**0.008**	0.174	4.38	↑
	T	T		0.251	**0.050**	0.736	2.45	**0.017**	0.259	3.74	↑
	T	T	C	0.244	**0.043**	0.692	2.49	**0.013**	0.252	3.74	↑

Abbreviations: P_{EMP1}, pointwise empirical p-value; P_{EMP2}, p-value corrected for multiple comparisons; OR, odds ratio; LR, low-response

Two haplotypes involving the SNPs of *IL1B* surpassed the univariate threshold of $P_{EMP1}<0.05$, but only the three-SNP haplotype, rs1143634-rs1143633-rs1143627 ACG, was significant after multivariate analysis (**Table 3.21)**. Furthermore, this three-SNP haplotype was significant after adjusting for multiple comparisons ($P_{EMP2}<0.05$). Both these haplotypes were associated with increased risk of low antibody response to the HBV vaccine (OR>1.0). This haplotype contained rs1143634 A, which was also significantly associated with low vaccine response in allelic and genotypic models.

Four haplotypes in chromosome 5 encompassing genes *IL4* and *IL12B* (all various combinations of rs2243250-rs2070874-rs3212227-rs3213094 TTTC), were significantly associated with increased risk of low anti-HBs response, and the associations remained significant following multivariate analysis. All haplotypes included the *IL4* SNP (rs2070874 T) that was also associated with poor HBV vaccine response in genotype analysis.

In summary, the *HLA* and cytokine SNPs that remained significantly associated with HBV vaccine response in our cohort in multivariate analysis were rs931 (*HLA-DPB1*), rs7770370 (*HLA-DPB1*), rs2070874 (*IL4*) and rs2069727 (*IFNG*); and haplotypes that remained significantly associated with HBV vaccine response in our cohort in multivariate analysis involved the *HLA* class II region, *IL1B* and *IL4/IL12B* genes (*HLA* class II region haplotypes rs2395309-rs2071349 AC, rs207349-rs931 CG, rs2071349-rs9277534 CA, rs931-rs9277534-rs9277535-rs3130186 GAAC and rs931-rs9277534-rs9277535-rs3130186-rs2064479 GAACC, and cytokine gene *IL1B* haplotype rs1143634-rs1143633-rs1143627 ACG and *IL4/IL12B* haplotypes rs2243250-rs2070874-rs3212227 TTT, rs2243250-rs2070874-rs3212227-rs3213093 TTTC, rs2070874-rs3212227 TT and rs2070874-rs3212227-rs3213093 TTC). Those that remained significant in multivariate analysis and after correction for multiple testing included rs2070874 (*IL4*) and rs7770370 (*HLA-DPB1*), and the *IL1B* haplotype rs1143634-rs1143633-rs1143627 ACG.

3.5 Discussion

We investigated *HLA* class II variation and cytokine SNPs in a South African cohort, and compared genetic variation to response to HBV vaccination.

3.5.1 *HLA* class II variation in the South African cohort

We reported *HLA-DPA1* and *-DPB1* allele frequencies and *HLA* class II region SNP allele and genotype frequencies in the whole cohort, as well as per ethnic group in the cohort.

There is limited *HLA-DP* region data available for South African cohorts. No data on *HLA-DPA1* alleles in South African cohorts was found, and very few papers in the literature have reported on *HLA-DPB1* data in South Africans (Lombard et al., 2006; Tikly et al., 2004; Tshabalala et al., 2015). Additionally, *HLA-DPB1*105:01* frequencies have not been reported previously, likely due to the inability to differentiate between **105:01* and **04:02* using standard sequence-based typing techniques. The *HLA* class II region SNP data have also not been previously reported in South African populations.

We found that most of the *HLA-DPA1/DPB1* allele frequencies and SNP frequencies differed significantly among different ethnic groups in our cohort. This was similar to reports for other *HLA* loci in South African populations (Paximadis et al., 2012). In addition, significant differences among the allele frequencies of the three ethnicities were found for 7/10 SNPs.

In contrast, *HLA* and SNP frequencies were similar to those observed in reference (shared ancestry) populations for all South African ethnicities examined. The few cases where differences were observed between the ethnic groups of this cohort and the reference populations may be due to the small sample size of each group. In addition, the reference populations used may not be ideal reference populations due to the broad ancestral grouping used. For example, there were no data in AFND for South African black populations so several sub-Saharan black populations were used as the reference instead; yet African populations remain largely subdivided (Campbell and Tishkoff, 2008).

3.5.2 Cytokine genetic variation in the South African cohort

We reported allele and genotype frequencies of fourteen SNPs in cytokine genes in a South African cohort, and compared these frequencies between different South African ethnicities.

The investigated SNPs have been reported on previously in South African cohorts (Govan et al., 2003; Meenagh et al., 2002; Moller et al., 2010; Naicker et al., 2009; Wadley et al., 2013), and differences among ethnicities have been previously reported (Govan et al., 2003; Meenagh et al., 2002). We observed SNP allele frequency differences between the South African ethnic groups in majority (11/14) of the cytokine SNPs.

There were only a few significant differences when comparing the South African ethnic groups to reference populations. Reasons for these differences were discussed above.

3.5.3 Genetic variation and HBV vaccine response

In addition to the clinical factors influencing anti-HBs titres, genetic factors also play a role. The heritability of anti-HBs titres after vaccination has been estimated to be more than 70% (Newport et al., 2004).

3.5.3.1 *HLA-DPA1* variation

HLA-DPA1 alleles

We observed five *HLA-DPA1* alleles (at 4-digit resolution). This relatively low number of alleles was expected as, globally, *HLA-DPA1* has the lowest number of alleles compared to the other class II *HLA* genes (Sidney et al., 2010; Tshabalala et al., 2015). Four *HLA-DPA1* alleles account for more than 90% of all *HLA-DPA1* genes and -*DPA1*01:0*3 representing approximately 60% of them (Sidney et al., 2010). These frequencies are seen in this study as well

Some previous studies have associated *HLA-DPA1*01:03* and *02:02* with HBV infection outcome and vaccine response (Kamatani et al., 2009; Mineta et al., 1996). However, another report found no association between *HLA-DPA1* alleles and HBV infection outcome (Thomas et al., 2012). In concordance with Thomas et al., (2012), our study did not find any significant associations between *HLA-DPA1* alleles and anti-HBs titres. Additionally, the small sample size may not have allowed the detection of small associations.

HLA-DPA1 region SNPs

The two *HLA-DPA1* SNPs investigated, rs2395309 located upstream and rs2071349 located in an intron, were not associated with HBV vaccine response in the allelic or genotypic analyses.

However, they are in complete LD with one another (**Figure 3.4**) and the rs2395309-rs2071349 haplotype AC was associated with decreased risk of low anti-HBs response.

The rs2395309 C allele was previously associated with increased susceptibility of chronic HBV infection (Guo et al., 2011; Kamatani et al., 2009), so our finding that the A allele (in the haplotype) contributes positively to the response suggests that the same genetic variation contributes to both poor response to HBV infection and poor response to HBV vaccination. While we did not study *HLA-DPA1* rs3077 in this study, this SNP has been associated with HBV vaccine response (Pan et al., 2013) as well as with HBV infection outcome in Asian populations (Guo et al., 2011; Hu et al., 2012; Kamatani et al., 2009; Mbarek et al., 2011; Nishida et al., 2012). The *HLA-DPA1* SNPs rs2395309 and rs3077 are in high LD in Europeans and Asians (O'Brien et al., 2011), as well as in Africans (as determined from the LWK population which is representative of the South African black population). The haplotype containing rs2395309 that associated with poor vaccine response in the current study should also be examined for linkage to rs3077.

The SNP rs2071349 has not been associated with HBV vaccine response previously.

Although not formally analysed, it appears that rs2395309 A allele is linked to *HLA-DPA1*01:03*, while the rs2395309 G allele is linked to *HLA-DPA1*02* alleles. *HLA-DPA1*01:03* has previously been associated with HBV infection clearance and with good HBV vaccine response (Kamatani et al., 2009; Mineta et al., 1996), and the rs2395309 A allele (in haplotype) was associated with good HBV vaccine response in this study, which supports the observation of the linkage that possibly exists between them.

3.5.3.2 *HLA-DPB1* variation

HLA-DPB1 alleles

In this study, *HLA-DPB1*01:01* and *09:01* alleles were significantly associated with low anti-HBs titre after vaccination, and *HLA-DPB1*04:01:01G* allele was significantly associated with good anti-HBs response after vaccination.

*HLA-DPB1*0101* has previously been reported in association with persistent HBV infection (Kamatani et al., 2009; Thomas et al., 2012), but it has not been previously associated with HBV vaccine response. This could be due to the fact that many of the vaccine response association studies were performed in Asian populations which have low *HLA-DPB1*01:01*

allele frequencies, and none were performed in African populations which have high *HLA-DPB1*01:01* frequencies (Thomas et al., 2012; **Table 3.14**).

*HLA-DPB1*09:01* has previously been reported in association with HBV vaccine non- or low-response (Desombere et al., 1998; Wu et al., 2013). The previous association of *HLA-DPB1*09:01* with HBV vaccine response was identified in an Asian population; and although the association was observed in our total cohort, the *09:01* allele is significantly more common in our South African Asian population compared to the South African black or Caucasian populations.

In this study, we found the *HLA-DPB1*04:01:01G* allele to be significantly associated with good anti-HBs response. Both **04:01* and **04:02* have previously been reported as associated with HBV infection clearance (Kamatani et al., 2009; Thomas et al., 2012). *HLA-DPB1*04:01* has also previously been associated with good HBV vaccine response (Desombere et al., 1998; Martinetti et al., 1995; Wu et al., 2015, 2013).

While ethnicity was not associated with HBV vaccine response in our analyses of demographic data, we did note that certain risk alleles were increased/decreased in particular South African ethnicities. The *HLA-DPB1*01:01* (risk) allele was significantly more frequent in the South African black population, while the *-DPB1*04:01* (protective) allele was significantly less frequent in the South African black population compared to the other populations. If the same risk alleles act in both HBV infection and HBV vaccine response then this might indicate a genetic reason as to why chronic HBV infection is increased in the South African black population compared to other population groups (Kew, 2008, 1996).

HLA-DPB1 region SNPs

We examined six *HLA-DPB1* region SNPs (rs7770370, rs931, rs9277534, rs9277535, rs3130186, rs2064479) and two (rs7770370 and rs931) were associated with antibody response to HBV vaccination. In addition, six haplotypes were associated with antibody response to HBV vaccination in multivariate analysis.

The rs7770370 intronic variant was the one of two *HLA-DPB1* SNP to be associated with HBV vaccine response in the allelic and genotypic analyses in this study, but the only one to remain significant after adjustment for multiple comparisons. The rs7770370 GG genotype was associated with an increased risk of low anti-HBs response. Similar results have been reported previously in Asian populations (Roh et al., 2016; Wu et al., 2015).

The other SNP associated with HBV vaccine response in the allelic model was rs931, located in the 3′ UTR. It was also in complete or partial LD with other SNPs of located in the 3′ UTR and downstream of the *HLA-DPB1* gene. rs931 has not been associated with HBV vaccine response previously, so the protective effect of the G allele cannot be confirmed. The rs931 A allele was more frequent in the black African population compared to the others in this South African cohort. These frequencies were similar to those of reference populations (**Table 3.17**).

In our study, the rs9277534, rs9277535 and rs313086 *HLA-DPB1* 3′ UTR variants were in complete LD with one another (**Figure 3.4**). The AAC haplotype of these SNPs was associated with a decreased risk of low-response. In a population of European- and African-Americans, the SNP rs9277534 showed the most significant association with HBV infection outcome, even stronger than the SNPs identified in Asian studies (Thomas et al., 2012). The rs9277534 GG genotype was previously associated with chronic HBV infection (Thomas et al., 2012) and with a decrease in Rubella virus antibodies (Lambert et al., 2014). Thus, the effect of rs9277534 is consistent in HBV infection outcome (Thomas et al., 2012), Rubella virus vaccine response (Lambert et al., 2014) and with HBV vaccine response (in this study). While rs9277535 has previously been associated with HBV vaccine response in an Indonesian and Korean populations (Png et al., 2011; Roh et al., 2016), this association was not observed in a Chinese population (Pan et al., 2013). However, rs9277535 has been repeatedly associated with HBV infection outcome in Asian populations in both GWAS (Hu et al., 2013; Kamatani et al., 2009; Mbarek et al., 2011) and candidate-gene studies (Guo et al., 2011; Hu et al., 2012; Wang, 2011). The rs9277535 A allele has a protective effect in both HBV vaccine response and infection outcome. The SNP rs3130186 has not been associated with HBV previously, likely due to its LD with adjacent SNPs; it is in complete LD with rs9277535 in the LWK population and in this South African population (determined by Haploview; **Figure 3.4**).

The variant rs2064479, located downstream of *HLA-DPB1*, was associated with HBV vaccine response in our study. The C allele was the protective allele in this study, which has also been found to be the case in Rubella virus vaccine response (Lambert et al., 2014).

These SNPs have previously been found to be in LD with *HLA-DP* alleles. For example, *HLA-DPB1*01:01* and **05:01* was in perfect LD with the G allele rs9277534, but **04:01* was in LD with the A allele in Asian and European- and African-American populations (Thomas et al., 2012); while the rs7770370 A allele was correlated with the *HLA-DPB1*02:01*, **02:02*, **03:01*, **04:01* and **14:01*, and negatively correlated with *HLA-DPB1*05:01* allele in an

Asian population (Wu et al., 2015); and *HLA-DPB1*05:01* is also in LD with rs9277535 G in an Asian population (Kamatani et al., 2009). Consequently, it is difficult to ascertain the precise causative variant, or combination thereof, although Thomas et al. (2012) did find that the rs9277534 variant had a stronger effect than any particular *HLA* allele.

Of the SNPs included in the haplotype analysis, the allele frequencies of rs2395309, rs931, rs9277534 and rs2064479 significantly differ between the South African ethnicities, with increased frequency of the minor allele in the South African black population (**Table 3.16**). In fact, the minor alleles of rs931 and rs9277534 of the total cohort are the major alleles in the South African black population. These allele frequencies are similar to those observed in the reference populations (**Table 3.17**). Of these four SNPs, the alleles of rs2395309, rs931, and rs9277534 that are associated with an increased risk of low-response (that is, G, A and G alleles respectively) are found at higher frequencies in the South African black population. Together with the *HLA-DPB1* allele findings, this supports the genetic reason behind an increased rate of chronic HBV infection and HBV vaccine non-response (average of 14%, as discussed in section 1.6) in the South African black population (Hino et al., 2001; Kew, 2008, 1996; Mphahlele et al., 2002; Schoub et al., 2002; Simani et al., 2009; Tsebe et al., 2001; Young et al., 2013). These findings also suggest the importance of the *HLA-DP* region in the HBV immune response.

3.5.3.3 Other SNPs of the *HLA* region

Neither the SNP upstream of *HLA-DRA* (rs3135363) nor the *HLA-DQB2* intronic SNP (rs7453920) were significantly associated with HBV vaccine response. Although these SNPs represent a very small portion of the variation of the *HLA-DR* or *-DQ* genes that may play a role in the response to the HBV vaccine, the *HLA-DP* region is the focus of association in this study.

In summary, rs931 which associated with HBV vaccine response on its own and in a haplotype, and rs2071347 and rs3130186 which associated with HBV vaccine response in haplotype, were novel associations. Additionally, SNPs that were previously associated with HBV infection outcome and/or other vaccine response, and were associated (in haplotype) with HBV vaccine response in the current study, included rs2395309, rs9277534 and rs2064479. The risk alleles of these SNPs were consistent in this, and previous studies. *HLA-DPB1*01:01* was previously associated with HBV chronic infection, and in this study, with poor HBV vaccine response.

The findings relating to SNPs rs7770370 and rs9277535, and to *HLA-DPB1*04:01* and **09:01*, alleles were consistent with previous findings of HBV vaccine response (and HBV infection outcome).

3.5.3.4 Cytokine gene variation

It has been noted that variation within the *HLA* region only accounts for part of the HBV vaccine response variability, suggesting that other factors contribute to this response (Mineta et al., 1996). Indeed, non-*HLA* genes may account for as much as 60% of the genetic factors influencing HBV vaccine response (Höhler et al., 2002b). Additionally, non-*HLA* genetic variation may be linked with the *HLA* alleles (Yucesoy et al., 2009). Furthermore, the HBV vaccine response ranges from no anti-HBs production to low anti-HBs levels and to very high titres which suggests that multiple gene interactions may underlie the mechanism of vaccine-induced immune response (Desombere et al., 1998). Thus, variation in cytokine genes was also investigated in this study. The following genes contained SNPs that were significantly associated with HBV vaccine response: *IL1B*, *IL4*, *IL12B* and *IFNG*.

***IL1B*:** IL1β is produced by monocytes, macrophages and dendritic cells (Garlanda et al., 2013). It plays a role in inflammation and host defence and is involved in Th17 cell differentiation (Garlanda et al., 2013).

In this study, the A allele of rs1143634 (a synonymous SNP in the *IL1B* gene) was associated with an increased risk of low-response in our South African cohort at the univariate level. In contrast, the A allele has previously been associated with a good response to the HBV vaccine in European-Americans adults (Yucesoy et al., 2002), although this association was not seen in infants (Yucesoy et al., 2009).

Although a synonymous SNP does not change the amino acid encoded by the codon, it can interrupt a splicing donor site, which could result in an inappropriate stop codon or exon skipping and thereby generating an incomplete mRNA (Thi Tran et al., 2005). The shorter mRNA would result in a truncated protein that may be non-functional or may be degraded rapidly (Bellone et al., 2006). It has been reported that the rs1143634 AA genotype is associated with increased IL1β levels (Pociot et al., 1992). As the A allele was associated with low response in our study, increased IL1β levels may lead to decreased anti-HBs response.

Additionally, the *IL1B* haplotype ACG (involving SNPS rs1143634, rs1143633 and rs1143627) was associated with an increased risk of low-response, and remained significant

after multivariate analysis and after adjusting for multiple comparisons (P_{EMP2}<0.05). Previously, the AG haplotype defined by SNPs rs1143633 and rs1143627 was found to correlate with non-response to HBV vaccine in a Chinese population (Chen et al., 2011). Therefore, the risk allele of rs1143633 also differs from previous studies but, the risk-associated G allele of rs1143627 is consistent between the two studies.

IL4: rs2070874 located in the 5′ UTR of *IL4* was significantly associated with HBV vaccine response in the genotypic and dominant models of inheritance. These results suggested that in our cohort having at least one T allele at rs2070874 was associated with low anti-HBs response. Contradictory associations have been reported in the literature regarding the role of the rs2070874 allele in HBV vaccine response: the T allele has been associated with increased vaccine response in Asian populations but not in Caucasian populations (Cui et al., 2013), and the C allele (in a haplotype) has been associated with increased response in various ethnicities in the USA (Wang et al., 2004). Our study supported the findings by Wang et al. (2004).

While the *IL12B* gene variants were not associated with HBV vaccine response in allelic and genotypic analyses, haplotypes involving the *IL4* and *IL12B* gene variants were significantly associated. The TTTC haplotype defined by rs2243250, rs2070874, rs3212227 and rs3213093 was associated with an increased risk of low-response. This haplotype included the *IL4* SNP (rs2070874 T) that was also associated with poor HBV vaccine response in genotype analysis.

Previously, the TTC haplotype of rs2243248, rs2243250 and rs2070874 was associated with increased response in an African- and European-American population (Wang et al., 2004). While rs2243248 was not associated with HBV vaccine response in this study, the effect of the T allele of rs2243250 was contradictory and yet the rs2070874 risk allele was consistent. Moreover, the T allele of rs2243250 was associated with chronic HBV infection in Caucasians (Zheng et al., 2013), but HBV clearance in a Chinese male population (Lu et al., 2014). The contradictory findings regarding rs2243250 may be due to ethnicity of the cohorts. Lu et al. (2014) noted that susceptibility to HBV infection was linked to the rs2243250 C allele, but that a meta-analysis found the risk allele to be a T at this position in Caucasian populations (Zheng et al., 2013). Differences in allele frequencies between populations may explain the observation that the same allele may have different effects in different ethnic groups (Lu et al., 2014). Indeed, the rs2243250 T allele is the minor allele in our South African Caucasian and Asian populations, but the major allele in the South African black (**Table 3.16**). Thus, a minor, risk allele in one population may be the major, non-risk allele in another (Lu et al., 2014). However,

the same differences in the allele frequencies between the South African ethnic groups were seen for both rs224350 and rs2070874, which showed consistent effects in the studies. Alternatively, these results may suggest the rs2070874 variant is more predictive of the immune response than the other *IL4* SNPs investigated.

With regards to the *IL12B* SNPs, rs3212227 TT genotype has been associated with increased antibody response against the HBV vaccine but decreased antibody response against the measles virus vaccine in European-American populations (Dhiman et al., 2007; Yucesoy et al., 2009). In this study, the *IL12B* rs3212227 T allele was only associated with low anti-HBs response in haplotype with *IL4* SNPs. Possibly, the effect of rs3212227 is influenced by the effect of the other SNPs in the haplotype. The *IL12B* rs3213093 CC genotype has previously been associated with decreased antibody response to the measles vaccine in a European-American population (Dhiman et al., 2007), and here the C allele (in the IL4 linked haplotype) was associated with low-response to the HBV vaccine.

IFNG: the rs2069727 SNP which is located downstream of the *IFNG* gene, was significantly associated with HBV vaccine response in the genotypic and dominant models of inheritance. Having at least one C allele at rs2069727 in our cohort was significantly associated with an increased anti-HBs response. In line with the results of our study, the C allele of rs2069727 has previously been associated with a good anti-HBs response in Gambians, although this was not significant in a replication study (Hennig et al., 2008)

Having at least one T allele at rs2070874, located in the 5′ UTR of the *IL4* gene, was significantly associated with the low-responder phenotype this study. *IL4* gene variants have previously been associated with both HBV vaccine response and HBV infection outcome, although the results are often seemingly contradictory. This risk-associated T allele has been observed an American population of various ethnicities (Wang et al., 2004) and negatively influences long-term anti-HBs duration (Wang et al., 2012). Furthermore, a small meta-analysis investigating *IL4* polymorphisms with response to HBV vaccination and HBV infection outcome found that the T allele of rs2070874 was associated with high response to the vaccine in Asian populations but this association was not seen in Caucasian populations (Cui et al., 2013).

While the *IL12B* gene variants were not associated with HBV vaccine response in allelic and genotypic analyses, haplotypes involving the *IL4* and *IL12B* gene variants were significantly

associated. The TTTC haplotype defined by rs2243250, rs2070874, rs3212227 and rs3213093 was associated with an increased risk of low-response.

Previously, the TTC haplotype of rs2243248, rs2243250 and rs2070874 in an African- and European-American population was associated with increased response. While rs2243248 was not associated with HBV vaccine response in this study, the effect of the T allele of rs2243250 was contradictory in the two studies and yet the rs2070874 risk allele was consistent. Moreover, the T allele of rs2243250 was associated with chronic HBV infection in Caucasians (Zheng et al., 2013), but HBV clearance in a Chinese male population (Lu et al., 2014). The ethnicity of the cohorts may play a role in the difference seen in the effect of rs2243250. Lu et al. (2014) noted that susceptibility to HBV infection was linked to the rs2243250 C allele, but that a meta-analysis found the risk allele to be a T at this position in Caucasian populations (Zheng et al., 2013). This difference in roles of a particular allele in different populations may be due to the differences in allele frequencies between populations, based on the genetic disparities between ethnicities (Lu et al., 2014). Indeed, the rs2243250 T allele is the minor allele in our South African Caucasian and Asian populations, but the major allele in the South African black (**Table 3.16**). Thus, a minor, risk allele in one population may be the major, non-risk allele in another (Lu et al., 2014). However, the same differences in the allele frequencies between the South African ethnic groups were seen for both rs224350 and rs2070874. Alternatively, these results may suggest the rs2070874 variant is more predictive of the immune response than the other *IL4* SNPs investigated.

With regards to the *IL12B* SNPs, rs3212227 TT genotype has been associated with increased antibody response against the HBV vaccine but decreased antibody response against the measles virus vaccine in European-American populations (Dhiman et al., 2007; Yucesoy et al., 2009). In this study, the T allele was associated with low anti-HBs response in a haplotype. The allele frequencies of rs3212227 do not differ significantly between the South African ethnic groups or between the South African ethnicities and the reference populations so ethnic differences are unlikely to explain the contradictions. Possibly, the effect of rs3212227 is influenced by the effect of the other SNPs in the haplotype. The rs3213093 CC genotype has previously been associated with decreased antibody response to the Measles vaccine in a European-American population (Dhiman et al., 2007), and here the C allele (in the haplotype) is associated with low-response to the HBV vaccine.

The rs2069767 variant, downstream of the *IFNG* gene, was one of the most significant associations detected with HBV vaccine response in a Gambian population, although this was not significant in a replication study (Hennig et al., 2008). The protective effect of the C allele was consistent in both studies.

In summary, cytokine SNPs associated with HBV vaccine response (on their own or in haplotype) that were consistent with previous studies on HBV vaccine response include rs1143627, rs2070874, rs2069727. The SNP rs32130186 has not previously been associated with HBV vaccine response, but had a consistent effect to response to measles vaccination. Several SNPs, including rs1143634, rs1143633, rs2243250 and rs3212227, had contradictory effects to those reported in the literature (and often these contradictory effects were noted in the literature as well).

From these results, and noted previously (Hu et al., 2013; Lu et al., 2014, 2010), there are often contradictory results where an association may be identified in one study but not in another or the same polymorphism is associated with different roles in different studies. These differences may arise due to various factors: genetic heterogeneity is apparent in different ethnic populations and HBV infection and vaccine response may be described as complex traits; therefore, the population being studied may affect the study outcomes due to genetic differences that exist (such as allele frequencies) and the numerous factors that will contribute to the phenotype. Additionally, population structure may play a role – where LD patterns differ and as such, an identified variant may not be the causal variant but is linked to the causal variant in some populations but not in others. Furthermore, the study populations' attributes often differ between the studies, such as age, weight, gender proportions, time since onset/vaccination and case/control definitions.

Nonetheless, these results support the significance of genetic variation in interacting cytokine genes in vaccine-induced immune responses. It is possible that the effect of an individual SNP may be limited, but contributes to the non-responder phenotype along with other variations.

3.5.4 Sample size and statistical power

An effective sample size is required to achieve an adequate statistical power that can reliably detect the causal variants of a phenotype (Hong and Parks, 2012). A sample size that is too small to detect a true association increases the rate of false negative results and thus, reduces

the reliability of the study (Hong and Parks, 2012). Power is also affected by MAF, case:control ratio, the frequency of phenotype of interest, LD, effect size and number of SNPs studied (Hong and Parks, 2012). The required sample size decreases with increasing MAF, increasing LD, increasing strength of effect size and decreasing number of SNPs studied (Hong and Parks, 2012).

The required case sample sizes for the different genetic models, under the assumptions of 5% phenotype prevalence, 5% MAF, complete LD, equal cases and controls and a 5% significance level, are shown in **Table 3.24**. For even a dominant genetic model where there is a very strong genetic association with the phenotype (that is, heterozygotes have an OR of 2 and homozygotes have an OR of 3), the case sample size should be at least 90 to achieve 80% power (**Table 3.24**). Thus, in this study, a sample size of at least 90 non-responders or low-responders should have been used to detect effects of dominant genotypes of OR>2 at 80% power, and should be much greater (in the region of 1000s) to detect effects of recessive genotypes of OR>2 at 80% power (**Table 3.24**). For SNPs with weak genetic association with the phenotype (for example, OR~1.3), the power to detect these would require even greater sample sizes (**Table 3.24**). The sample size of this study (n=8 cases) is clearly insufficient to achieve sufficient power to detect any effect size with confidence. Therefore, significant results in this section should be viewed with caution.

Table 3.24. Number of cases necessary to achieve 80% power for the different genetic models in a case-control study (adapted from Hong and Park, 2012)

Genetic Model	$OR_{heterozygotes}/OR_{homozygotes}$			
	1.3/1.6	**1.5/2**	**2/3**	**2.5/4**
Allelic	1 974	789	248	134
Dominant	606	258	90	53
Co-dominant	2 418	964	301	161
Recessive	20 294	8 390	2 776	1 536

Assumptions: 5% phenotype prevalence, 5% MAF, complete LD (D′=1), 1:1 case:control ratio and 5% significance level for single marker analyses
Abbreviations: OR, odds ratios

CHAPTER 4

Associations Between *HLA-DPB1* Alleles, *HLA-DPB1* Expression and HBV Vaccine Response in a South African Cohort

4.1 Introduction

The complexity of the HLA effects in immune response extends beyond the genetic variation to include HLA expression levels. *HLA-DPB1*01:01*, **02:01*, **04:01*, **04:02* and **05:01* are common alleles in most populations (Sidney et al., 2010) and represent alleles associated with both HBV recovery and persistence (Kamatani et al., 2009; Thomas et al., 2012) and with both low and high anti-HBs response (Desombere et al., 1998; Martinetti et al., 1995; Mineta et al., 1996; Wu et al., 2014, 2013). Yet, these alleles share largely overlapping peptide binding repertoires (Sidney et al., 2010). This data supports the findings that the mechanism of HBV vaccine response is not (or not entirely) due to a defect in antigen binding capabilities (Desombere et al., 2005, 1995; Godkin et al., 2005). Thus, Thomas et al., (2012) postulated that the effect of *HLA-DP* on HBV infection outcome may be due to differences in the expression levels rather than to differences in peptide presentation.

HLA expression levels have been shown to affect disease progression of a variety of diseases. HLA-C expression levels have been implicated in HIV disease progression (Apps et al., 2013; Thomas et al., 2009) and Crohn's disease (Apps et al., 2013). Increased HLA-C expression levels associated with slower progression to Acquired Immunodeficiency Syndrome (AIDS), attributed, in part, to an enhanced CTL-mediated response; whereas higher HLA-C levels increase the risk of Crohn's disease (Apps et al., 2013). Both variants upstream of the *HLA-C* gene and in its 3′ UTR have been associated with the differential expression levels (Kulkarni et al., 2011; Thomas et al., 2009). Additionally, seroconversion to the influenza vaccine has been associated with higher HLA-DR expression (Egli et al., 2014). HLA-DPB1 expression levels have been associated with HBV infection outcome, with lower expression levels associating with HBV clearance (Thomas et al., 2012). The major determinant of the HLA-DP expression levels was the rs9277534 variant, located in the *HLA-DPB1* 3′ UTR (Thomas et al., 2012). Individuals with GG genotype (which was linked to the *HLA*-DPB1*01:01 allele) had higher levels of HLA-DP expression than individuals with AA/AG genotypes. These findings are in contrast to another study which found that the rs9277535 G allele associated with both lower levels of HLA-DP expression and with HBV persistence (O'Brien et al., 2011).

Based on the findings of Thomas et al. (2012), our study hypothesis was that individuals with *HLA-DPB1*04:01* versus *01:01* homozygous genotypes would have low and high *HLA-DPB1* expression levels respectively, which might influence their Th1/Th2 cytokine production. Poor responders were also hypothesised to also have high *HLA-DPB1* expression levels, similar to

the *HLA-DPB1*01:01* homozygotes. We also wanted to test how *HLA-DPB1* expression levels changed after HBsAg stimulation.

Therefore, we aimed to investigate whether *HLA-DPB1* expression levels were associated with HBV vaccine response. As both *HLA-DPB1* SNPs and alleles have been associated with variability in vaccine response we aimed to investigate whether *HLA-DPB1* expression levels differed between groups defined by both *HLA-DPB1* alleles and *HLA-DPB1* SNP genotypes; and whether *HLA-DPB1* expression levels directly affected anti-HBs titre.

4.2 Materials and Methods

4.2.1 Selection of samples for expression analysis

Twenty-seven participants were selected for analysis of *HLA-DPB1* expression analysis based on their HBV vaccine response status and on their *HLA-DPB1* alleles. We selected all poor responders in our cohort (n=8), as well as all participants in our cohort who were *04:01* homozygous (n=8), or *01:01* homozygous (n=4) for further analysis of *HLA-DPB1* expression levels, as shown in **Table 4.1**. We had space on our plates for extra samples and therefore also included *04:01* heterozygotes (n=7) in these analyses.

4.2.2 Stimulation of selected samples with HBsAg

We stimulated cells with whole HBsAg subtype *adw* (Novus Biologicals, UK). This subtype is used in several common vaccines such as Engerix-B (Hernández-Bernal et al., 2011) as well as in the vaccine used for the booster vaccinations in this study (Heberbio HBV, The Biovac Institute, SA). Moreover, the '*a*' determinant is common to all types of HBV and anti-HBs mainly target this determinant (Kramvis et al., 2005; Michel and Tiollais, 2010); thus we were confident of the stimulating capabilities of this HBsAg subtype on PBMCs from HBV vaccinated individuals. We used PHA (Sigma, Germany) as a positive stimulation control because it is a commonly used mitogen of T lymphocytes, and commonly used in stimulation assays (Jafarzadeh and Shokri, 2012, 2003; Jarrosson et al., 2004; Kardar et al., 2002; Velu et al., 2008; Wataya et al., 2001).

4.2.3 Thawing of PBMCs

PBMCs were thawed in a 37°C water bath and then transferred drop-wise to warmed R20 media (RPMI-1640 with 20% foetal calf serum (FCS)). The cells were pelleted by centrifugation at 250 g for 10 minutes at 23°C. The supernatant was discarded and the cells resuspended in the remaining media. This wash was repeated once, after which the volume was increased to 2 ml using R20 media. Ten microliters of the cell suspension were removed for a cell count using Trypan Blue and the TC20 Automated Cell counter. (If necessary, additional cells were thawed to ensure sufficient cells were recovered for a consistent count across all stimulations). The cells were then incubated overnight at 37°C in a humidified CO_2 (5%) incubator. The cell count was repeated the following day.

Table 4.1. Samples selected for *HLA-DPB1* expression assays

Sample ID	Category	Ab titre 1 (IU/L; after 3 doses, from records)	Ab titre 2 (IU/L; cross sectional, this study)	Ab titre 3 (IU/L; after boost, this study)	*HLA-DPB1* Genotype
HEP035	Poor responder	107 (2 mo after 3rd dose)	106	NA	*01:01/18:01*
HEP057	Poor responder	45 (1 mo after 3rd dose)	0	146 (1 mo after boost)	*09:01/09:01*
HEP059	Poor responder	122 (3 mo after 3rd dose)	1	83 (1 mo after boost)	*02:01/02:01*
HEP071	Poor responder	17 (1 mo after 3rd dose)	1	>1000 (1 mo after boost)	*01:01/04:02*
HEP086	Poor responder	198 (5 mo after 3rd dose)	13	96 (1 mo after boost)	*04:02/11:01*
HEP102	Poor responder	196 (1 mo after 3rd dose)	4	468 (1 mo after boost)	*01:01/106:01*
HEP111	Poor responder	74 (5 mo after 3rd dose)	12	>1000 (2 mo after boost)	*01:01/01:01*
HEP134	Poor responder	198 (2 mo after 3rd dose)	6	151 (1 mo after boost)	*02:01/04:02*
HEP011	*0401* homozygote, High responder	?	>1000	NA	*04:01/04:01*
HEP017	*0401* homozygote, High responder	?	>1000	NA	*04:01/04:01*
HEP037	*0401* homozygote, High responder	>1000	43	NA	*04:01/04:01*
HEP065	*0401* homozygote, High responder	>1000	407	NA	*04:01/04:01*
HEP068	*0401* homozygote, High responder	>1000	90	NA	*04:01/04:01*
HEP085	*0401* homozygote, High responder	?	>1000	NA	*04:01/04:01*
HEP091	*0401* homozygote, High responder	?	>1000	NA	*04:01/04:01*
HEP157	*0401* homozygote, High responder	?	>1000	NA	*04:01/04:01*
HEP001	*0101* homozygote	742 (2 mo after 3rd dose)	3	>1000 (1 mo after boost)	*01:01/01:01*
HEP003	*0101* homozygote	?	7	675 (1 mo after boost)	*01:01/01:01*
HEP049	*0101* homozygote	?	64	NA	*01:01/01:01*
HEP108	*0101* homozygote	236 (5 mo after 3rd dose)	23	NA	*01:01/01:01*
HEP005	*0401* heterozygote (High responder)	?	>1000	NA	*04:01/02:01*
HEP015	*0401* heterozygote (High responder)	>1000 (2 mo after 3rd dose)	>1000	NA	*04:01/05:01*
HEP019	*0401* heterozygote (High responder)	?	>1000	NA	*04:01/09:01*
HEP060	*0401* heterozygote (High responder)	?	>1000	NA	*04:01/02:01*
HEP090	*0401* heterozygote (High responder)	>1000	>1000	NA	*04:01/01:01*
HEP092	*0401* heterozygote (High responder)	?	>1000	NA	*04:01/03:01*
HEP096	*0401* heterozygote (High responder)	?	>1000	NA	*04:01/09:01*

? – no records

Abbreviations: Mo, months; NA, not applicable as participant did not receive a boost

4.2.4 *In vitro* stimulation of PBMCs

PBMC stimulation in cohort samples was conducted as follows. Two million cell/ml were stimulated with 10 µg/ml PHA (as a control), or 5 µg/ml purified HBsAg *adw* protein in a final volume of 0.5 ml and were incubated for 5 days at 37°C in a humidified CO_2 (5%) incubator. Thereafter, the cell suspensions were centrifuged at 250 g for 10 minutes at 23°C. The cell pellets were combined with (350 µl) RLT/β-mercaptoethanol solution and stored at -70°C for future RNA extraction.

4.2.5 *HLA-DPB1* mRNA expression level analysis

HLA-DPB1 mRNA was measured under two conditions in selected samples: firstly, at the time of sample collection without any specific antigen stimulation, and secondly after cell stimulation with HBsAg. Samples used for the first measurement were PBMC stored in RLT/β-mercaptoethanol solution on the day of blood collection. Samples used for the antigen specific measurements were the remaining cells after 5 days of stimulation with HBsAg, stored in RLT/β-mercaptoethanol solution.

4.2.5.1 Extraction of RNA

Total RNA was extracted from the lysate of PBMCs stored in RLT/β-mercaptoethanol solution using the QIAamp RNA Blood Mini Kit (Qiagen), a column based method. The lysate was thawed at room temperature and transferred to a QIAshredder spin column which was then centrifuged at maximum speed (16 100 g) for 2 minutes. One volume of 70% ethanol was added to the homogenised lysate in the collection tube and mixed by pipetting. This lysate was then transferred to a QIAamp spin column and centrifuged for 15 seconds at 10 200 g. If the sample volume exceeded the maximum loading volume of 700 µl, the sample was loaded in successive aliquots and the centrifugation step repeated. After transferring the spin column to a new collection tube, 700 µl Buffer RW1 was applied to it and the column centrifuged for 15 seconds at 10 200 g. Five hundred microliters of Buffer RPE was added to the spin column in a new collection tube and centrifuged again at 10 200 g for 15 seconds. After discarding the flow-through, an additional 500 µl Buffer RPE was added to the spin column which was then centrifuged at 16 200 g for 3 minutes. The spin column was then placed in a new collection tube and centrifuged at 16 100 g for 1 minute in order to minimize the possibility of Buffer RPE carryover. Finally, the spin column was transferred to a 1.5 ml microcentrifuge tube, 30

µl RNase-free water (warmed to 95°C) was applied to the column membrane and centrifuged at 10 200 g for 1 minute to elute the RNA.

Five microliters were removed for RNA quality and quantity analysis by NanoDrop 1000 and Agilent 2100 Bioanalyzer using the Agilent RNA 6000 Nano kit. The extracted RNA samples were stored at -70°C until use.

4.2.5.2 RNA quantity and quality assessment

The Agilent technique is an automated procedure for assessing RNA quality by calculating an RNA integrity number (RIN). RIN values are based on several features that contribute information about RNA integrity, not only on the traditional measure of 18S to 28S rRNA ratio (Schroeder et al., 2006). RIN values range from 10 (intact RNA) to 1 (totally degraded RNA). The technique utilizes microcapillary electrophoresis whereby molecules are separated according to size in gel filled capillaries, and detection is by laser-induced fluorescence of an intercalating dye. Each chip can analyse 12 samples and requires only 1 µl of sample (Schroeder et al., 2006).

The gel was prepared by filtering the gel matrix through the provided spin filter and then adding 1 µl of the dye concentrate to a 65 µl aliquot of the filtered gel matrix. The gel-dye mix was vortexed thoroughly and then centrifuged for 10 minutes at 13 000 g. The tube was protected from light and stored at room temperature until use. The ladder and samples to be analysed were heat-denatured for 2 minutes at 70°C and then cooled immediately on ice.

The chip was then prepared according to the manufacturer's instructions, adding the gel, ladder, marker and samples to specified wells. The chip was vortexed prior to inserting it into the Bioanalyzer for analysis. The RIN for good quality RNA to be used in downstream applications is debatable. Intact RNA is generally considered for samples with RIN >8, but samples with RIN >6 are considered good enough for downstream quantitative PCR (qPCR) (de Cremoux et al., 2011; Fleige and Pfaffl, 2006). Furthermore, it has been suggested that the amplification of small PCR products (<250 bp) is less dependent on good RNA quality than larger amplicon sizes (Fleige and Pfaffl, 2006).

Additionally, the A260/A280 and A260/A230 quality ratios were determined for each sample by NanoDrop.

4.2.5.3 cDNA synthesis

The High Capacity RNA-to-cDNA Kit (Applied Biosystems) was used to synthesize cDNA. The reaction mix included 1× RT Buffer (containing dNTPs, random octamers and oligo dT-16), 1× RT Enzyme mix (MultiScribe™ MuLV reverse transcriptase and an RNase inhibitor) and 234 ng RNA template in a final volume of 20 µl. cDNA was synthesized at 37°C for 60 min, followed by 95°C for 5 minutes and then 4°C for infinity. The samples were stored at -20°C until use. The cDNA was then evaluated by UV absorbance using the NanoDrop 1000.

4.2.5.4 qPCR

Relative gene expression level of *HLA-DPB1* was determined by qPCR and the Comparative quantification cycle (Cq) method, also known as the delta delta Cq (ddCq) method (Livak and Schmittgen, 2001). The mRNA levels of *HLA-DPB1* were normalized against two reference genes, Importin-8 (*IPO8*) and TATA-binding protein (*TBP*). The reference genes were selected based on previous reports of their stability in PBMCs and stimulated PBMCs (Bibova et al., 2012; Ledderose et al., 2011; Piehler et al., 2010; Radke et al., 2014; Wang et al., 2014). Predesigned TaqMan® Gene Expression Assays (Applied Biosystems) were used. The assay IDs for *HLA-DPB1*, *IPO8* and *TBP* were Hs03045104_m1, Hs00183533_m1 and Hs00427621_m1 respectively. These assays rely on three sequence-specific binding events involving two primers and a probe which are designed to ensure high target specificity and to guarantee amplification at least ten Cq values earlier than the gene with the closest sequence homology (Applied Biosystems, 2010).

To determine whether the efficiencies of the three genes were similar, as well as to determine the minimum quantity of cDNA required, a standard curve was constructed for each gene, from the average results of three replicate wells, using 150 ng, 15 ng, 1.5 ng, 0.15 ng and 0.015 ng cDNA.

To analyse cohort samples, a qPCR reaction mix of 20 µl was used, consisting of 1× TaqMan® Gene Expression Master Mix (Applied Biosystems), 1× TaqMan Gene Expression Assay (Applied Biosystems), and 17.6 ng of cDNA. The TaqMan® Gene Expression Master Mix contains AmpliTaq Gold® DNA Polymerase, Uracil-DNA glycosylase, dNTPs (with dUTP) and ROX™ Passive Reference in an optimised buffer. The TaqMan® Gene Expression Assays consist of a PCR primer pair and a hydrolysis probe with FAM™ dye label on the 5′ end and minor groove binder (MGB) and nonfluorescent quencher (NFQ) on the 3′ end. Amplification

was achieved using the Applied Biosystems 7500 Real-Time PCR System under the cycling conditions listed in **Table 4.2**. Each gene was quantified on its own (single plex), each sample was measured in triplicate wells and a no-template control was included in each run.

Since the ddCq method requires the efficiencies of the genes to be similar, a validation experiment was performed using the data generated by the standard curves. A semi-regression line was produced by plotting the delta Cq (dCq) value versus log input amount. If the absolute value of the slope is <0.1, the efficiencies of the genes are considered sufficiently similar for the ddCq method to be used (Applied Biosystems, 2004; Livak and Schmittgen, 2001).

Table 4.2: Cycling conditions for qPCR

Step	Temperature	Incubation Time	Number of Cycles
1*	50°C	2 minutes	1
2	95°C	10 minutes	1
3	95°C	15 seconds	
	60°C	1 minute	45

*Required for Uracil-DNA glycosylase (UNG) activity.

The dCq was calculated by subtracting the geometric mean Cq of the reference genes from the Cq of the target gene. The samples were grouped based on *HLA-DPB1* alleles (Low-responders, *04:01* homozygotes, *04:01* heterozygotes, *01:01* homozygotes), as well as on genotypes of the SNP rs9277534 and the SNPs significantly associated with HBV vaccine response in the study. The ddCq value for each group was determined by subtracting the mean dCq of an assigned calibrator group from the mean dCq of that group.

The quality of the qPCR data was controlled by considering the standard deviation (SD) of the replicates as well as the Cq values. We removed any outliers of samples with replicates SD >0.3, and removed any samples with an average Cq >35 from the analysis. Inter-plate variation was controlled for by running the same calibrator sample in each plate and adjusting the Cq values accordingly.

Using SPSS v24.0, ANOVA (or the non-parametric Kruskal-Wallis test) were performed to compare expression levels between ≥3 groups:

1. *HLA-DPB1* genotypes, in particular, low-responders versus *HLA-DPB1*04:01* homozygotes versus **01:01* homozygotes versus **04:01* heterozygotes
2. Genotypes of the SNP previously associated with *HLA-DPB1* expression levels (rs9277534) (Thomas et al., 2012), and SNPs in the *HLA* class II region that were significantly associate with HBV vaccine response (rs931, rs7770370)

The Mann-Whitney U test was performed to analyse differences in expression levels between two groups, that is, when genotypes were combined, for example, rs7770370 GG vs AG+AA.

The expression levels were compared to anti-HBs titre using a Spearman's rank-order correlation in SPSS v24.0.

4.3 Results

4.3.1 RNA quality

The RNA extracted from the PBMCs stored on the day of blood collection was of good quality: the samples (n=27) had an average RIN of 8.7 (range 4.5 - 9.6) and a high average concentration of 157 ng/µl (range 26-414 ng/µl). An example of Agilent digital gel image and electropherograms is shown in **Figure 4.1**. The average A260/A280 quality ratio determined by NanoDrop was also very good at 1.96 (range 1.3-2.1). The average A260/A230 ratio was 1.39 (range 0.3-2.3), suggesting that other contaminants, such as ethanol, were present in the samples. We used all 27 RNA samples from the collection time-point for cDNA synthesis and qPCR.

Next, we tested the RNA extracted from HBsAg-stimulated PBMC samples after five-day stimulation. Due to limited stimulation reagents, only 24 samples were stimulated with HBsAg. Agilent quality indicators for the 24 RNA samples extracted from cells remaining from the five-day stimulation with HBsAg included an average concentration of 70 ng/µl. The Agilent assay could only determine a RIN for seven of the 24 samples tested, and showed, an average RIN of 5.2 (for seven samples) which is relatively poor quality. The average A260/A280 ratio following RNA extraction was 1.89 (range 1.8-1.9). However, the A260/A230 ratio at this stage was less than ideal (lower than 2), with an average of 1.67, suggesting the presence of contaminants. Overall, the RNA extracted from these post-stimulation samples was of poor quality/degraded, as determined by the Agilent assay. RNA quality in two samples stimulated with PHA was also of poor quality as measured by RIN, suggesting that the long stimulation period or the culture conditions played a role in RNA degradation. Unfortunately, we therefore did not continue with qPCR for samples post-antigen stimulation. Consequently, pre- and post-stimulation expression levels could not be compared and HBsAg-specific *HLA-DPB1* expression levels could not be compared between low-responders and responders.

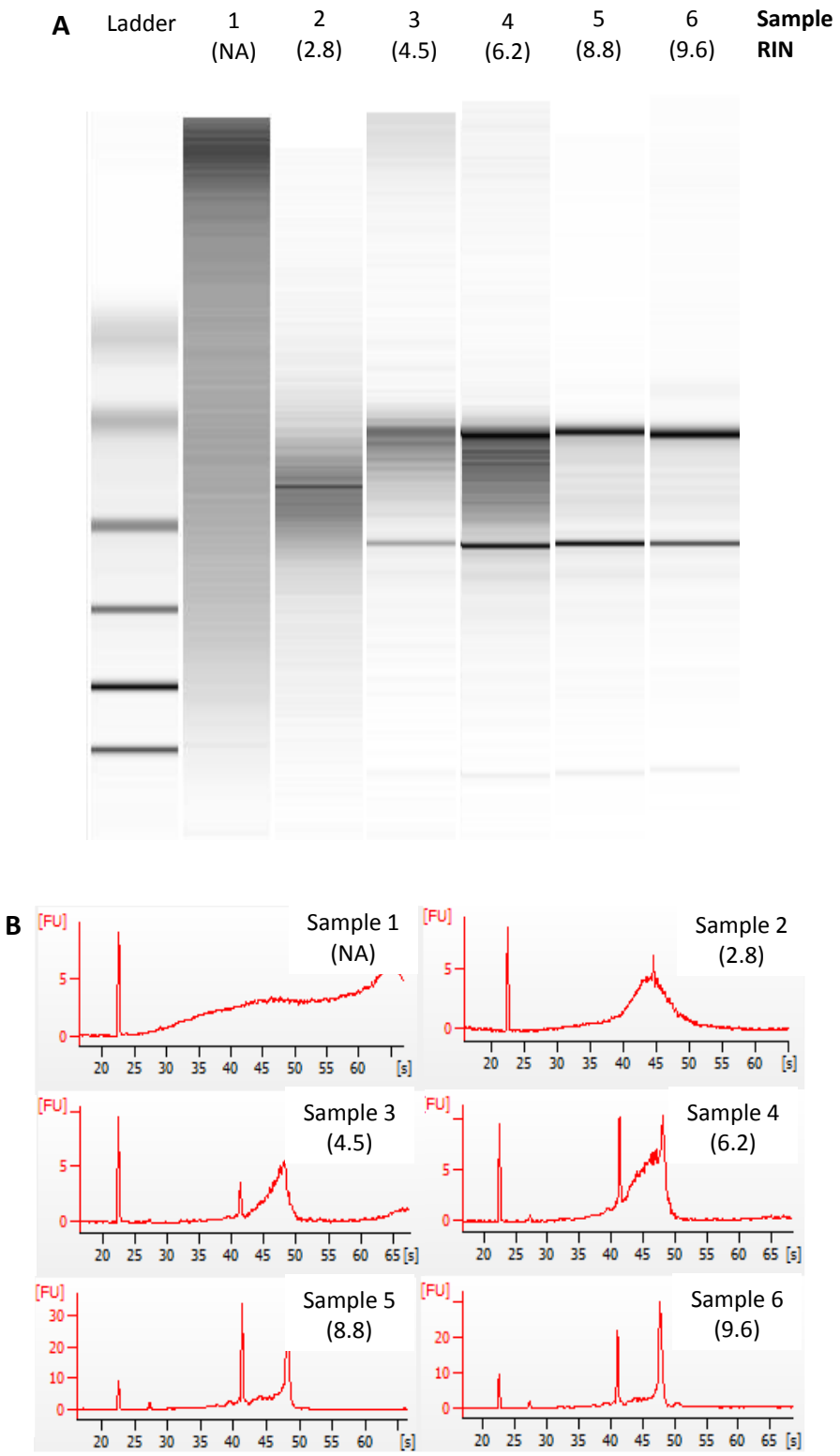


Figure 4.1. An example of Agilent assay gel images and electropherograms for total RNA extracted from six samples. RIN values are shown in brackets below sample number. Samples 1-6 are in order of increasingly good quality (RIN values ranging from 'not any' – 9.6)

A) Digital gel images for Ladder and 6 samples.

B) The corresponding digital electropherograms. The y-axis represents fluorescence units [FU], and the x-axis represents runtime [s].

4.3.2. qPCR validation

The efficiency of the qPCR reactions for each gene was determined from the standard curves (**Figure 4.2**). *HLA-DPB1* had a percentage efficiency of 83%, whereas *IPO8* and *TBP* had efficiencies of 81% and 82%, respectively.

The standard curves also showed that 15 ng input cDNA resulted in Cq values of approximately 25 and 30 for the target and reference genes, respectively. As the sample with the lowest RNA concentration would result in a cDNA input of 17.6 ng, it was decided that this would be a satisfactory input amount as it would result in Cq values <30 for all genes.

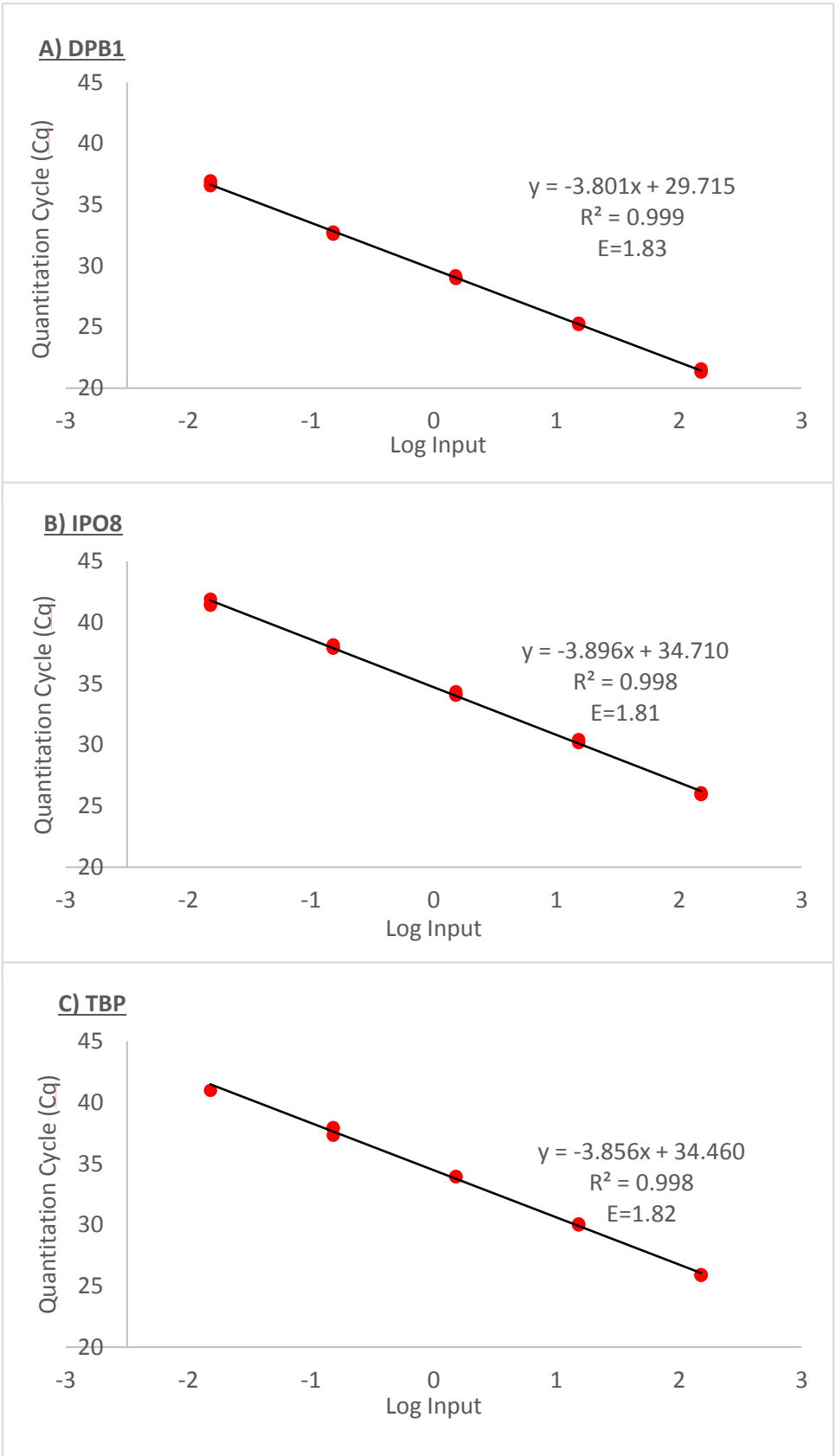


Figure 4.2. Standard curves of Cq versus log cDNA input for *HLA-DPB1* (A), *IPO8* (B) and *TBP* (C). The equation of the line and the R^2 and E values for each graph are shown alongside the line of best fit.

The semi-log regression line of the dCq versus log input amount (**Figure 4.3)** was drawn to check if the efficiencies of the PCR reactions were similar enough to use the Livak method. The slope of this line was 0.0693. As this value was <0.1, the efficiencies of the genes were similar and therefore, the Comparative Cq method could be used (Applied Biosystems, 2004; Livak and Schmittgen, 2001). The E value of 1.82 was therefore used for the ddCq calculations below.

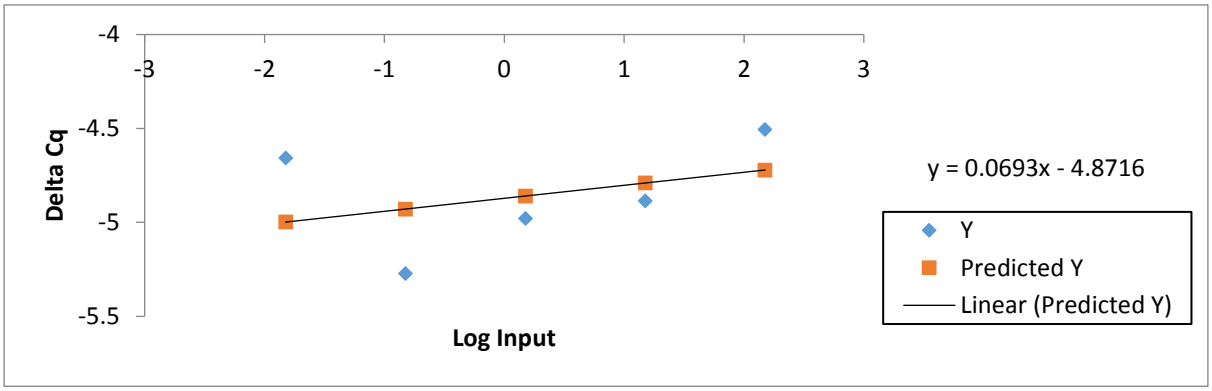


Figure 4.3. Regression line of dCq ($Cq_{target} - Cq_{geometric\ mean\ of\ references}$) versus log input amount. The key and the formula of the line are shown to the right of the graph.

4.3.3 qPCR assays in cohort samples

4.3.3.1 cDNA quality

Using cDNA generated from 27 samples pre-stimulation, *HLA-DPB1* Cq values (averaged from three replicate wells) ranged from 23.79 to 27.05 and had a mean value of 25.04. Mean Cq values for *IPO8* and *TBP* for the same 27 samples were 28.67 (range 27.45 - 29.80) and 29.15 (range 27.60 - 30.3). The average SD for the three replicate samples was 0.07 for all genes and the SD was never more than 0.2. The SD range for *HLA-DPB1* was 0.01 to 0.18, for *IPO8*: 0.02 to 0.19, and for *TBP*: 0 to 0.19. These values suggested that the replicates were reliable and the results were of good quality; thus, no replicates were excluded based on the above results.

4.3.3.2 *HLA-DPB1* relative expression (pre-stimulation), compared to genetic data

We compared relative *HLA-DPB1* expression levels in the 27 samples used for qPCR to:

1. *HLA-DPB1* genotypes, in particular, low-responders versus *HLA-DPB1*04:01* homozygotes versus **01:01* homozygotes versus **04:01* heterozygotes

2. Genotypes of the SNP previously associated with *HLA-DPB1* expression levels (rs9277534) (Thomas et al., 2012), and SNPs in the *HLA* class II region that were significantly associate with HBV vaccine response (rs931, rs7770370)

With regards to *HLA-DPB1* genotype: the 27 samples (pre-stimulation) used in qPCR were grouped into low-responders (n=8), *HLA-DPB1*04:01* homozygotes (n=8), *HLA-DPB1*01:01* homozygotes (n=4) and *HLA-DPB1*04:01* heterozygotes (n=7) (**Table 4.1)**. Initial statistics of the qPCR dCq data identified a significant outlier (as assessed by inspection of box-plot of dCq values). Closer inspection of this sample showed that it had a RIN of 4.5 and this was considered a valid reason to exclude the sample from the analysis. Thus, n=26, with only three individuals in the *HLA-DPB1*01:01* homozygote group. Following removal of this sample, the data were determined to be normally distributed (Shapiro-Wilk test, $p>0.05$ for all groups), there were no further outliers in each group (assessed by inspection of a box-plot) and there was homogeneity of variance (Levene's test, $p>0.05$), thus an ANOVA was performed to compare *HLA-DPB1* expression among low-responders (n=8), *HLA-DPB1*04:01* homozygotes (n=8), *HLA-DPB1*01:01* homozygotes (n=3) and *HLA-DPB1*04:01* heterozygotes (n=7). Data is presented as mean dCq ± standard deviation, but for graphical representation, fold change is used (**Figure 4.4**).

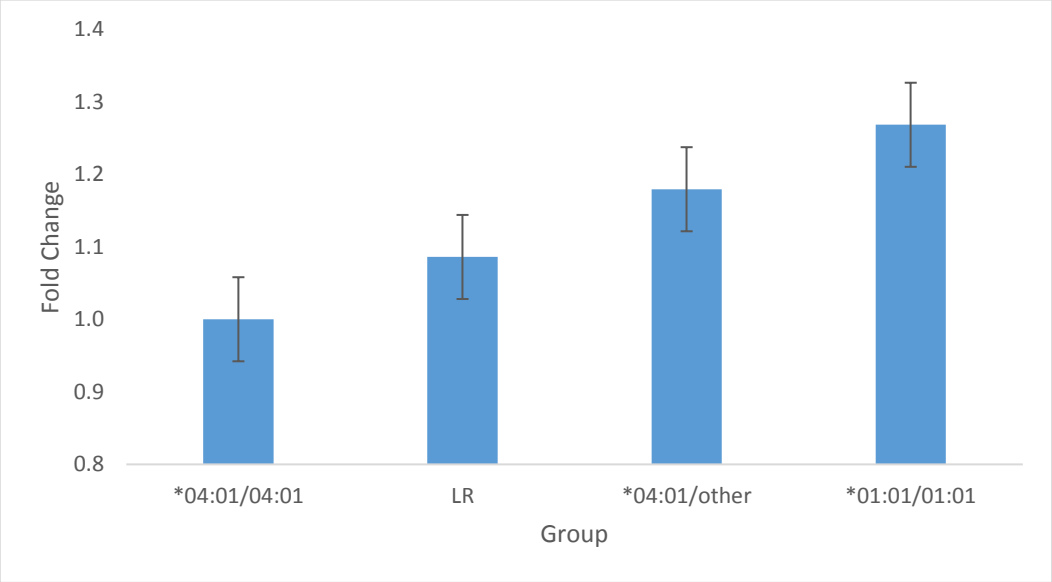


Figure 4.4. Comparison of *HLA-DPB1* gene expression levels between *04:01* homozygotes, low-responders, *04:01* heterozygotes and *01:01* homozygotes. Fold change (1.82^{-ddCq}) relative to *04:01* homozygotes is shown on the y-axis. P=0.277.

The *HLA-DPB1* gene expression levels increased slightly from the *04:01* homozygotes (-3.73±0.44) to low-responders (-3.86±0.33) to *04:01* heterozygotes (-4.00±0.20) to *01:01* homozygotes (-4.12±0.19) in that order (see **Figure 4.4**). These results confirmed the literature description that the *HLA-DPB1*04:01* allele is associated with lower expression and the *01:01* with higher *HLA-DPB1* expression (Thomas et al., 2012). However, the differences between these groups were not significant ($p>0.05$); indeed, the fold changes (1.82^{-ddCq}) relative to the *04:01* homozygote group were <2. The expression level of the poor responders was intermediate between *04:01* homozygotes and *01:01* homozygotes but was also not significantly different to these groups. Due to the very small sample sizes, significance testing may not have been very reliable.

Next, the *HLA-DPB1* expression levels were compared to an *HLA* class II SNP which has previously been associated with *HLA-DPB1* expression levels and HBV infection outcome (rs9277534, *HLA-DPB1* 3′ UTR) (Thomas et al., 2012), and to *HLA* class II SNPs which were significantly associated with antibody response to HBV vaccine, including rs7770370 in *HLA*-DPB1 intron and rs931 in *HLA-DPB1* 3′ UTR (**Figure 4.5**).

Since rs9277534 3'UTR GG genotype was associated with HBV infection persistence, and with higher *HLA-DPB1* expression in the literature (Thomas et al., 2012), we aimed to confirm whether this SNP would be associated with *HLA-DPB1* expression levels in this cohort. Once again, as can be seen in **Figure 4.5 (A)**, although rs9277534 GG had the highest expression (-4.10±0.16), followed by GA (-3.94±0.28) and then AA (-3.73±0.40), the differences were not significant ($p=0.09$). Additionally, there was no significant difference between the GG (-4.10±0.16) and the AX (-3.83±0.36) genotypes ($p=0.062$), as determined by a Mann-Whitney U test. These results were expected due to the fact that rs9277534 A allele is in LD with *HLA-DPB1*04:01* and the G allele with *HLA-DPB1*01:01* in African- and European-Americans (Thomas et al., 2012).

The *HLA-DPB1* intron variant rs7770370 was associated with response to the HBV in this cohort (in the genotypic and recessive models). Additionally, having an rs7770370 A allele has been associated with good HBV vaccine response (Roh et al., 2016; Wu et al., 2015). Thus, it was expected that the A allele would be associated with lower *HLA-DPB1* mRNA expression levels and therefore the association of this SNP with *HLA-DPB1* mRNA expression levels was tested. Having an A allele did result in lower expression levels than the GG genotype (**Figure**

4.5 (B)). However, the differences in fold expression among genotypes was not significant (Kruskal Wallis H test, p=0.287).

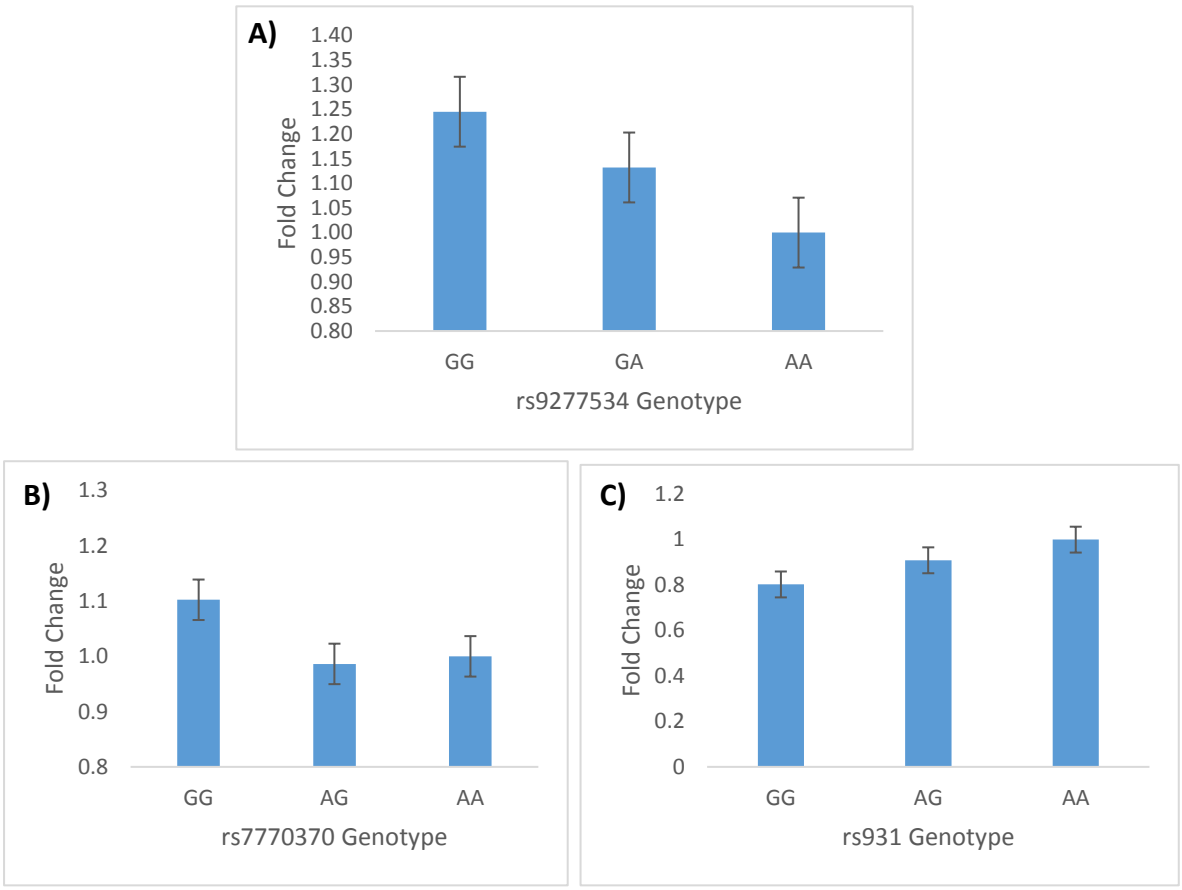


Figure 4.5. *HLA-DPB1* expression levels according to:

A) rs9277534 genotypes, p=0.090
B) rs7770370 genotypes, p=0.287
C) rs931 genotypes, p=0.090
Fold change (1.82^{-ddCq}) is shown on the y-axis.

rs931 is located in the 3′ UTR of *HLA-DPB1* and was associated with HBV vaccine response at allelic level with the A allele associating with an increased risk of low-response. Therefore, the A allele was expected to be associated with higher *HLA-DPB1* expression levels. This was indeed the case (**Figure 4.5 (C)**), however the differences were not significant (p=0.09), as assessed by ANOVA.

4.3.3.3 *HLA-DPB1* relative expression (pre-stimulation), compared to anti-HBs

In order to assess whether *HLA-DPB1* expression directly affected the anti-HBs titre, a Spearman's rank-order correlation was performed, comparing dCq values for the 26 samples

to the corresponding anti-HBs titre at time of blood collection. There was no significant correlation between *HLA-DPB1* expression (pre-stimulation) and anti-HBs titre (p=0.843).

4.4 Discussion

We found a similar trend to that noted by Thomas et al. (2012) – higher *HLA-DPB1* expression levels in low-responders and *HLA-DPB1*01:01* homozygotes compared to *HLA-DPB1*04:01* homozygotes, and higher expression levels in individuals with the rs9277534 GG genotype. However, these associations did not reach significance at the 5% level. The mRNA expression level study by Thomas et al. (2012) was performed in healthy European-American subjects, comparing ten AX samples to five GG samples; whereas this study was performed in mixed-ethnicity group involving twenty AX samples and six GG samples. The ethnicity may have influenced the results as level of gene expression has been noted to differ between individuals, and particularly, between population groups; although this is often due to differences in allele frequencies between the populations (Spielman et al., 2007). Additionally, HBsAg-specific expression may be more important to investigate in vaccine response.

We also found that when comparing expression levels between the genotypes of *HLA-DPB1* SNPs rs7770370 and rs931, the alleles associated with low anti-HBs response were also linked to higher *HLA-DPB1* expression levels, although not statistically significant.

We hypothesize that genetic variation of the *HLA-DP* region may affect *HLA-DPB1* expression, which in turn, may influence the Th1/Th2 cytokine production, thereby affecting anti-HBs production. Several studies have associated MHC expression levels with the cytokine profile in other phenotypes (Baumgart et al., 1998; Cavalli et al., 2016), or cytokine with MHC expression level (de Waal Malefyt et al., 1991). This theory could provide a link between the two characteristics commonly associated with HBV vaccine response, namely, *HLA* allele and cytokine profile.

In summary, these results show a trend of lower *HLA-DPB1* expression associating with better anti-HBs response (by proxy of other allele association with HBV vaccine response), as has been seen with HBV infection outcome (Thomas et al., 2012), although these associations were not significant. It is possible that a clearer (and significant) difference would be seen following HBsAg-specific stimulation and/or between high-responders and true non-responders to the HBV vaccine.

CHAPTER 5

Summary and Conclusions

Characterising the mechanism underlying HBV vaccine response is important in order to identify targets for more effective HBV vaccines that will provide protection to all vaccinees. In the broader picture, understanding the mechanisms behind HBV vaccine non-response may shed light on the pathways of normal immune response to vaccines, and thereby enable us to improve the current vaccines and to design novel vaccines against diseases for which no vaccine is yet available.

This research aimed to, firstly, establish a South African cohort in which response to HBV vaccination, and immunogenetic mechanisms for this response, could be explored. Secondly, we aimed to explore *HLA-DP* and other *HLA* class II and cytokine SNP variation in South African individuals, and to investigate the association between these genetic variants and anti-HBs titres. Thirdly, we aimed to explore the relationship between *HLA-DPB1* expression levels and HBV vaccine response in South Africans.

Despite not finding any true non-responders in our cohort, we found several significant associations when comparing low anti-HBs responders to normal responders. *HLA-DPB1*01:01* and *09:01* alleles were associated with low anti-HBs response, while **04:01* was associated with good HBV vaccine response. We identified four SNPs that significantly associated with HBV vaccine response in our cohort in multivariate analysis: rs931 (*HLA-DPB1*) which has not previously been associated with HBV vaccine response, rs7770370 (*HLA-DPB1*), rs2070874 (*IL4*) and rs2069727 (*IFNG*). We also identified several significant haplotypes involving SNPs in the *HLA-DP* region and in the cytokine genes *IL1B* and *IL4/IL12B*. Those that remained significant in multivariate analysis and after correction for multiple testing included rs2070874 (*IL4*) and rs7770370 (*HLA-DPB1*), and the *IL1B* haplotype rs1143634-rs1143633-rs1143627 ACG. Thus, it can be concluded that *HLA* class II, particularly *HLA-DP* genetic variation and cytokine genetic variation associates with antibody response to the HBV vaccination in South African individuals.

The *HLA-DPB1* expression levels were not significantly associated with HBV vaccine response, but a trend of lower *HLA-DPB1* expression associating with better anti-HBs response was observed. Thus, *HLA-DPB1* expression levels may influence immune responses to HBV vaccination however these findings need to be confirmed.

It seems that a complex interplay of genes and pathways is involved in HBV vaccine response, and the combination of these factors will determine the overall effect and the immune response. Nevertheless, such association studies enable us to better understand the complexity and the

diversity of genetic factors that influence the efficacy of the HBV vaccine. These results may assist in the identification of specific genetic targets to be used in the development of more effective and novel vaccines.

There are several similarities between the genetic associations observed in HBV infection outcome and in HBV vaccine response, including *HLA* alleles and *HLA* and cytokine SNPs. These findings suggest that similar mechanisms cause persistent HBV infection and HBV vaccine non-response. Thus, these two phenotypes can act as models for each other and assist in understanding the mechanism underlying both these phenotypes. Typically, one would assume that immune response to a viral infection (that is, cellular-mediated) should differ to that of a protein antigen, as is found in the HBV vaccine (that is, humoral-mediated). This highlights the overlap between pathways of immune response.

Inconsistencies between this study's results and previous studies' results were noted for some of the associations, for which there could be several explanations. Firstly, associations may be specific to particular populations or ethnic groups. We observed significant differences between the different South African ethnic groups, and these ethnic groups were similar to similar ancestral reference groups. Differences in associations between populations may be due to the population-specific genetic structure, including variations in allele frequencies, different LD structures and the consequent fact that different SNPs may be in LD with the causal SNP in different ethnic groups. Secondly, differences in other study cohort characteristics, such as age, or study definitions, for example, studies may compare non-responders (<10 IU/L) or low-responders (typically 10-100 IU/L) to responders (>10 IU/L) or to high-responders (>1000 IU/L), may exist between studies which could affect the associations observed. Thirdly, study sample sizes may differ, and small associations may not be detected in small sample sizes.

5.1 Limitations

There were several limitations of this study. Importantly, the sample size is small (n=149) and the proportion classified as low-responders is very small (n=8). Such small sample sizes may not have sufficient power to detect small differences or effects of a particular factor, which could lead to important associations being missed. In addition, with such a small sample size, the study was underpowered which consequently reduced the reliability of the results. Another limitation that may have masked important associations is the classification we used due to the lack of true non-responders in our cohort. This resulted in us comparing low-responders (though responders nonetheless) to normal-responders, and thus associations with HBV vaccine response that may be detected when comparing more distinct groups such as non-responders to responders, or even high-responders, may have been missed. Additionally, the cohort was of mixed ethnicities and differences in allele frequencies and LD patterns between ethnic groups may have masked associations that are population-specific. Furthermore, the presence of population stratification is indicated by the observed differences in allele frequencies between the South African populations. The cohort largely consisted of discrete populations which are remotely related, yet, unfortunately, the sample size was too small to analyse ethnicities separately in order to identify population-specific associations. Methods are available to correct for population stratification, however, these generally involve the typing of numerous SNPs which are unrelated to the SNPs under investigation (Mingyao et al., 2010). These methods are more suited to GWAS or large gene association studies, and were not possible to utilise in this study. Including ethnicity as a covariate in the analyses could have at least accounted for some of the population stratification.

With regards to the *HLA-DPB1* allele associations: the *HLA-DP* genotyping technique reported many alleles in G groups. If the association with HBV vaccine response is not based on antigen binding capabilities, this may have masked some associations. This could have also affected the differences in *HLA-DPB1* expression levels since the hypothesis was based on *HLA-DPB1*04:01* alleles, whereas we used **04:01:01G* group alleles in the comparison. We also did not use antigen-specific stimulated samples for the expression work. It is possible that stimulation would have resulted in increased expression of the *HLA-DPB1* gene, at which point differences between the low-responders and responders may have been significant. Additionally, we only measured expression level in a portion of the cohort. A larger sample size may have allowed for the detection of small differences.

Furthermore, several limitations arose from the cross-sectional study design, including differences in vaccine type used, vaccination site and schedule and time since vaccination. These differences were exacerbated by the lack of vaccination history and antibody titre records which consequently made it difficult to account for them in analysis. Additionally, the participant historical data were obtained by way of self-reported questionnaires which is subject to recall bias. There was also a fair amount of missing data due to inability to recall certain facts. Both of these may have distorted the data. A prospective or retrospective study design, with full vaccination history and antibody titre would have provided a more complete and homogenous sample from which more clear outcomes could have been concluded.

Another limitation of the study is that the demographics of the cohort is not a true representation of the general South African population. According to Statistics South Africa's mid-year population estimates report (2016), the South African population is majority black African (80.7%), followed by Coloured and Caucasian (8.8% and 8.1%, respectively) and lastly by Indian/Asian (2.5%). Whereas, in our cohort, Caucasian was the majority ethnicity (50.3%) followed by black African (32.9%), Asian (14.8%) and Coloured (2.0%). Additionally, approximately 51% of the South African population is estimated to be female (Statistics South Africa, 2016), but our cohort was 79.2% female. Lastly, 30.1% of the South population is estimated to be younger than 15 years and 8.0% is 60 years or older (Statistics South Africa, 2016), whereas the age of our cohort ranged from 21 to 74 years. Thus, results obtained from this study may not apply to the general South African population.

5.2 Strengths

This research has gone some way to addressing the paucity of *HLA-DP* and cytokine SNP data in SA. To the best of our knowledge, this is the first study to investigate the genetics of HBV vaccine response in SA.

HLA-DP is less often studied compared to the other classical *HLA* loci, so this data contributes to a somewhat lacking area.

The low-responders were confirmed with booster vaccination and re-testing of antibody titres thereafter, and we did not only rely on anti-HBs, HBsAg testing was performed to confirm that these individuals were not currently infected with HBV. Results of these tests were shared to the participants.

The cohort was established by myself, with the help of many people who are acknowledged in the Acknowledgements section. This included recruitment, sample processing and storage, assay optimization (including for PBMC isolation, DNA extraction, HLA-DP allele genotyping, RNA extraction, cDNA synthesis and qPCR), data analysis and feedback to participants regarding their results.

5.3 Future Work

Future work in this area of research could include establishing a larger cohort in order for the study to have adequate power. The inclusion of true non-responders could also assist in identifying important factors in the HBV vaccine-induced immune response. Expanding the number genes investigated would be beneficial in attempting to understand the genetics underlying HBV vaccine response. In particular, including the *HLA-DQ* and *-DR* alleles in the analysis could provide essential information on the role of haplotypes of the HLA class II genes in HBV vaccine response. Furthermore, it is possible that multiple variants could contribute small effects which influence the antibody response to the HBV vaccine in a cumulative manner. The only way to test this would be to perform GWAS. As noted in the introduction, GWAS related to Hepatitis B vaccination outcomes have been performed in Asian populations (Png et al., 2011; Pan et al., 2014; Wu et al., 2015), and identified SNPs in the *HLA* class II region as important. GWAS of HBV vaccine response have not yet been performed in a South African population. The recent introduction of an African-specific genotyping array (Mulder et al., 2016) would make such a study more relevant. Another important aspect would be to use HBsAg-stimulated samples in the expression level experiments, and to perform these experiments in a larger sample. It would also be of interest to quantify cytokine levels in non- and/or low-responders versus responders in order to assess the T cell-mediated responses involved in HBV vaccine response. Additionally, determining the association between cytokine SNP genotypes and their corresponding cytokine levels would provide information on the functional importance of cytokine gene variation. Cytokine quantification could assist in determining whether HLA-DP antigen density contributes to the fate of the Th cell, and the subsequent cytokines produced.

REFERENCES

Abbas, A.K., Murphy, K.M., and Sher, A. (1996). Functional diversity of helper T lymphocytes. Nature *383*, 787–793.

Adams, S., Robbins, F., Chen, D., Wagage, D., Holbeck, S.L., Iii, H.C.M., Stroncek, D., and Marincola, F.M. (2005). HLA class I and II genotype of the NCI-60 cell lines. J. Transl. Med. *3*, doi:10.1186/1479-5876-3-11.

Adeyemo, A., and Rotimi, C. (2010). Genetic variants associated with complex human diseases show wide variation across multiple populations. Public Health Genomics. *13*, 72-79.

Agena Bioscience, iPLEX Reagents User Guide, USG-CUS-071.

Al-Qahtani, A., Khalak, H.G., Alkuraya, F.S., Al-Hamoudy, W., Alswat, K., Al Balwi, M. a, Al Abdulkareem, I., Sanai, F.M., and Abdo, A. a (2013). Genome-wide association study of chronic hepatitis B virus infection reveals a novel candidate risk allele on 11q22.3. J. Med. Genet. *1*, doi:10.1136/jmedgenet-2013-101724.

Amirzargar, A.A., Mohseni, N., Shokrgozar, M.A., Arjang, Z., Ahmadi, N., Behzadi, M.Y., Amanzadeh, A., and Shokri, F. (2008). HLA-DRB1, DQA1 and DQB1 Alleles and Haplotypes Frequencies in Iranian Healthy Adult Responders and Non-Responders to Recombinant Hepatitis B Vaccine. Iran. J. Immunol. *5*, 92–99.

Applied Biosystems (2004). Guide to Performing Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR, 117GU17-01. 1–60.

Applied Biosystems White Paper (2010). Gene Expression Assay Performance Guaranteed With the TaqMan ® Assays QPCR Guarantee Program. http://www.gene-quantification.de/lifetech-cms-088.

Apps, R., Qi, Y., Carleson, J.M., Chen, H., Gao, X., Thomas, R., Yuki, Y., Del Prete, G.Q., Goulder, P., Brumme, C.J., et al. (2013). Influence of HLA-C Expression Level on HIV Control. Science (80-.). *340*, 87–91.

Averhoff, F., Mahoney, F., Coleman, P., Schatz, G., Hurwitz, E., and Margolis, H. (1998). Immunogenicity of hepatitis B vaccines: Implications for persons at occupational risk of hepatitis B virus infection. Am. J. Prev. Med. *15*, 1–8.

Banaszkiewicz, A., and Radzikowski, A. (2013). Efficacy , effectiveness , immunogenicity - are not the same in vaccinology. World J. Gastroenterol. *19*, 7217–7218.

Banatvala, J.E., and Van Damme, P. (2003). Hepatitis B vaccine – do we need boosters ? J. Viral Hepat. *10*, 1–6.

Barnaba, V., Franco, a, Alberti, a, Benvenuto, R., and Balsano, F. (1990). Selective killing of hepatitis B envelope antigen-specific B cells by class I-restricted, exogenous antigen-specific T lymphocytes. Nature *345*, 258–260.

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

Bauer, T., and Jilg, W. (2006). Hepatitis B surface antigen-specific T and B cell memory in individuals who had lost protective antibodies after hepatitis B vaccination. Vaccine *24*, 572–577.

Baumgart, M., Moos, V., Schuhbauer, D., and Muller, B. (1998). Differential expression of major histocompatibility complex class II genes on murine macrophages associated with T cell cytokine profile and protective suppressive effects. Proc. Natl. Acad. Sci. U. S. A. *95*, 6936–6940.

Bellamy R., Ruwende C., McAdam K.P., Thursz M., Sumiya M., Summerfield J., et al. (1998). Mannose binding protein deficiency is not associated with malaria, hepatitis B carriage nor tuberculosis in Africans. QJM *91*, 13–8.

Bellamy R., Ruwende C., Corrah T., McAdam K.P., Thursz M., Whittle H.C., et al. (1999). Tuberculosis and chronic hepatitis B virus infection in Africans and variation in the vitamin D receptor gene. J. Infect. Dis. *179*, 721–4.

Bellone, E., Balestra, P., Ribizzi, G., Schenone, A., Zocchi, G., Di Maria, E., Ajmar, F., and Mandich, P. (2006). An abnormal mRNA produced by a novel PMP22 splice site mutation associated with HNPP. J. Neurol. Neurosurg. Psychiatry *77*, 538–540.

Bibova, I., Linhartova, I., Stanek, O., Rusnakova, V., Kubista, M., Suchanek, M., Vasakova, M., and Sebo, P. (2012). Detection of immune cell response to M . tuberculosis – specific antigens by quantitative polymerase chain reaction. Diagn. Microbiol. Infect. Dis. *72*, 68–78.

Blackwell, J.M., Jamieson, S.E., and Burgner, D. (2009). HLA and infectious diseases. Clin. Microbiol. Rev. *22*, 370–385.

Bonilla, F.A., and Oettgen, H.C. (2010). Adaptive immunity. J. Allergy Clin. Immunol. *125*, S33–S40.

Borzooy, Z., Streinu-cercel, A., Mirshafiey, A., Khamseh, A., Mahmoudie, M.K., Navabi, S.S., Nosrati, M., Najafi, Z., Hosseini, M., and Jazayeri, S.M. (2016). IL-17 and IL-22 genetic polymorphisms in HBV vaccine non- and low- responders among healthcare workers. GERMS *6*, 14–20.

Broere, F., Apasov, S.G., Sitkovsky, M. V, Eden, W. Van, and van Eden, W. (2011). T cell subsets and T cell-mediated immunity.

Bromley, S.K., Burack, W.R., Johnson, K.G., Somersalo, L., Sims, T.N., Sumen, C., Davis, M.M., Shaw, A.S., Allen, P.M., and Dustin, M.L. (2001). The immunological synapse. Annu. Rev. Immunol. *19*, 375–396.

Bronowicki J.P., Abdelmouttaleb I., Peyrin-Biroulet L., Venard V., Khiri H., Chabi N., et al. (2008). Methylenetetrahydrofolate reductase 677 T allele protects against persistent HBV infection in West Africa. J. Hepatol. *48*, 532–9.

Burnett, R.J., Kramvis, A., Dochez, C., and Meheus, A. (2012). An update after 16 years of hepatitis B vaccination in South Africa. Vaccine *30*, C45–C51.

Caillat-Zucman, S., Gimenez, J.J., Wambergue, F., Albouze, G., Lebkiri, B., Naret, C., Moynot, a, Jungers, P., and Bach, J.F. (1998). Distinct HLA class II alleles determine antibody response to vaccination with hepatitis B surface antigen. Kidney Int. *53*, 1626–1630.

Carollo, M., Palazzo, R., Bianco, M., Pandolfi, E., Chionne, P., Fedele, G., Tozzi, A.E., Carsetti, R., Romanò, L., Ausiello, C.M., et al. (2013). Hepatitis B specific T cell immunity induced by primary vaccination persists independently of the protective serum antibody level. Vaccine *31*, 506–513.

Cavalli, G., Hayashi, M., Jin, Y., Yorgov, D., Santorico, S.A., Holcomb, C., Spritz, R.A., and Dinarello, C.A. (2016). MHC class II super-enhancer increases surface expression of HLA-DR and HLA-DQ and affects cytokine production in autoimmune vitiligo. PNAS *113*, 1363–1368.

Chathuranga, L.S., Noordeen, F., and Abeykoon, A.M.S.B. (2013). Immune response to hepatitis B vaccine in a group of health care workers in Sri Lanka. Int. J. Infect. Dis. *17*, 1078–1079.

Chen, J., Liang, Z., Lu, F., Fang, X., Liu, S., Zeng, Y., Zhu, F., Chen, X., Shen, T., Li, J., et al. (2011). Toll-like receptors and cytokines/cytokine receptors polymorphisms associate with non-response to hepatitis B vaccine. Vaccine *29*, 706–711.

Chiaramonte, M., Ngatchu, T., Majori, S., Baldo, V., Moschen, M.E., Renzulli, G., and Trivello, R. (1995). Response to an Extra Dose of Hepatitis B Vaccine and Specific Antibody Persistence in Non-Responders to Primary Immunization. Scand. J. Gastroenterol. *30*, 601–603.

Clemens, R., Stinger, R., Kruppenbacherj-, J., Hiibel, W., Stanbury, W., Bock, H.L., and Jilg, W. (1997). Booster immunization of low- and non-responders after a standard three dose hepatitis B vaccine schedule-results of a post-marketing surveillance. Vaccine *15*, 349–352.

de Clercq, E., Férir, G., Kaptein, S., and Neyts, J. (2010). Antiviral treatment of chronic hepatitis B virus (HBV) infections. Viruses *2*, 1279–1305.

Cleveland, J.L., Siew, C., Lockwod, S.A., Gruninger, S.E., Chang, S.-B., Neidle, E.A., and Russel, C.M. (1994). Factors associated with hepatitis B vaccine response among dentists. J. Dent. Res. *73*, 1029–1035.

Coates, T., Wilson, R., Patrick, G., André, F., and Watson, V. (2001). Hepatitis B vaccines: assessment of the seroprotective efficacy of two recombinant DNA vaccines. Clin. Ther. *23*, 392–403.

Craven, D.E., Awdeh, Z.L., Kunches, L.M., Yunis, E.J., Dienstag, J.L., Wener, B.G., Polk, F., Snydman, D.R., Platt, R., Crumpacker, C.S., et al. (1986). Nonresponsiveness to Hepatitis B Vaccine in Health Care Workers. Ann. Intern. Med. *105*, 356–360.

Creemers, P.C., and du Toit, E.D. (1996). C4 polymorphism and extended HLA haplotypes in Namibian San and Khoi and in South African Xhosa. Tissue Antigens *47*, 111–116.

de Cremoux, P. De, Valet, F., Gentien, D., Lehmann-che, J., Scott, V., Tran-perennou, C., Barbaroux, C., Servant, N., Vacher, S., Sigal-zafrani, B., et al. (2011). Importance of pre-analytical steps for transcriptome and Importance of pre-analytical steps for transcriptome and RT-qPCR analyses in the context of the phase II randomised multicentre trial REMAGUS02 of neoadjuvant chemotherapy in breast cancer patients. BMC Cancer *11*, 215–225.

Crotty, S. (2011). Follicular helper CD4 T cells (Tfh). Annu. Rev. Immunol. *29*, 621–663.

Cui, W., Sun, C.-M., Deng, B.-C., and Liu, P. (2013). Association of polymorphisms in the interleukin-4 gene with response to hepatitis B vaccine and susceptibility to hepatitis B virus infection: a meta-analysis. Gene *525*, 35–40.

Cupps, T.R., Gerin, J.L., Purcell, R.H., Goldsmith, P.K., and Fauci, A.S. (1984). In vitro antigen-induced antibody responses to hepatitis B surface antigen in man. Kinetic and cellular requirements. J. Clin. Invest. *74*, 1204–1213.

Custer, B., Sullivan, S., and Hazlet, T. (2004). Global epidemiology of hepatitis B virus. J. Clin. Gastroenterol. *38*, S158–S168.

Dane, D.S., Cameron, C.H., and Briggs, M. (1970). Virus-like particles in serum of patients with Australia-antigen- associated hepatitis. La *295*, 695–698.

Das, K., Gupta, R.K., Kumar, V., and Kar, P. (2003). Immunogenicity and reactogenicity of a recombinant hepatitis B vaccine in subjects over age of forty years and response of a booster dose among nonresponders. World J. Gastroenterol. *9*, 1132–1134.

Desombere, I., Hauser, P., Rossau, R., Paradijs, J., and Leroux-roels, C. (1995). Nonresponders to Hepatitis B vaccine can present envelope particles to T lymphocytes. J. Immunol. *154*, 520–529.

Desombere, I., Willems, A., and Leroux-Roels, G. (1998). Response to hepatitis B vaccine: multiple HLA genes are involved. Tissue Antigens *51*, 593–604.

Desombere, I., Gijbels, Y., Verwulgen, a., and Leroux-Roels, G. (2000). Characterization of the T cell recognition of hepatitis B surface antigen (HBsAg) by good and poor responders to hepatitis B vaccines. Clin. Exp. Immunol. *122*, 390–399.

Desombere, I., Cao, T., Gijbels, Y., and Leroux-Roels, G. (2005). Non-responsiveness to hepatitis B surface antigen vaccines is not caused by defective antigen presentation or a lack of B7 co-stimulation. Clin. Exp. Immunol. *140*, 126–137.

Dhiman, N., Ovsyannikova, I.G., Cunningham, J.M., Vierkant, R.A., Kennedy, R.B., Pankratz, V.S., Poland, G.A., and Jacobson, R.M. (2007). Associations between Measles Vaccine Immunity and Single-Nucleotide Polymorphisms in Cytokine and Cytokine Receptor Genes. J. Infect. Dis. *195*, 29.

Dienstag, J.L., Werner, B.G., Polk, B.F., Snydman, D.R., Craven, D.E., Platt, R., Crumpacker, C.S., Ouellet-Hellstrom, R. and Grady, G.F. (1984). Hepatitis B vaccine in health care personnel: safety, immunogenicity, and indicators of efficacy. Annals of Int. Med. *101*, 34-40.

Dieye A., Obami-Itou V., Barry M.F., Raphenon G., Thiam A., Ndiaye R., et al. (1999). Association between Class I HLA alleles and HBs antigen carrier status among blood donors in Senegal. Dakar Méd *44*, 166–70.

Edwards, J.A., Jones, D.B., Evans, P.R., and Smith, J.L. (1985). Differential expression of HLA class II antigens on human fetal and Adult Lymphocytes and Macrophages. Immunology *55*, 489–500.

Edwards, J.A., Durant, B.M., Jones, D.B., Evans, P.R., and Smith, J.L. (1986). Differential expression of HLA class II antigens in fetal human spleen: relationship of HLA-DP, DQ, and DR to immunoglobulin expression. J. Immunol. *137*, 490–497.

Effros, R.B. (2007). Role of T lymphocyte replicative senescence in vaccine efficacy. Vaccine *25*, 599–604.

Egli, A., Santer, D.M., Shea, D.O., Barakat, K., Syedbasha, M., Vollmer, M., Baluch, A., Bhat, R., Groenendyk, J., Joyce, M.A., et al. (2014). IL-28B is a Key Regulator of B- and T-Cell Vaccine Responses against Influenza. PLoS Pathog. *10*, e1004556.

Ehreth, J. (2003). The globla value of vaccination. Vaccine *21*, 596–600.

Excoffier, L., and Lischer, H.E.L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. *10*, 564–567.

Fisman, D.N., Agrawal, D., and Leder, K. (2002). The effect of age on immunologic response to recombinant hepatitis B vaccine: a meta-analysis. Clin. Infect. Dis. *35*, 1368–1375.

Fleige, S., and Pfaffl, M.W. (2006). RNA integrity and the effect on the real-time qRT-PCR performance. Mol. Aspects Med. *27*, 126–139.

Frodsham, A.J. (2005). Host genetics and the outcome of hepatitis B viral infection. Transpl. Immunol. *14*, 183–186.

Frodsham A.J., Zhang L., Dumpis U., Taib N.A., Best S., Durham A., et al. (2006). Class II cytokine receptor gene cluster is a major locus for hepatitis B persistence. Proc. Natl. Acad. Sci. USA *103*, 9148–53.

Gabriel., S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., et al. (2002). The structure of haplotype blocks in the human genome. Science *296*, 2225-2229.

Gabriel, S., Ziaugra, L., and Tabbaa, D. (2009). SNP Genotyping Using the Sequenom MassARRAY iPLEX Platform. 1–18.

Gansbacher, B., and Zier, K.S. (1988). Regulation of HLA-DR , DP , and DQ Expression in Activated T Cells. Cell. Immunol. *117*, 22–34.

Garlanda, C., Dinarello, C.A., and Mantovani, A. (2013). The Interleukin-1 Family : Back to the Future. Immunity *39*, 1003–1018.

Garner-Spitzer, E., Wagner, A., Paulke-Korinek, M., Kollaritsch, H., Heinz, F.X., Redlberger-Fritz, M., Stiasny, K., Fischer, G.F., Kundi, M., and Wiedermann, U. (2013). Tick-borne encephalitis (TBE) and hepatitis B nonresponders feature different immunologic mechanisms in response to TBE and influenza vaccination with involvement of regulatory T and B cells and IL-10. J. Immunol. *191*, 2426–2436.

Ginaldi, L., De Martinis, M., D'Ostilio, A., Marini, L., Loreto, M.F., Corsi, M.P., and Quaglino, D. (1999). The Immune System in the Elderly: I. Specific Humoral Immunity. Immunol. Res. *20*, 101–108.

Glimcher, L.H., and Kara, C.J. (1992). SEQUENCES AND FACTORS : A guide to MHC class-II transcription. Annu Rev Immunol *10*, 13–49.

Godkin, A., Davenport, M., and Hill, A.V.S. (2005). Molecular analysis of HLA class II associations with hepatitis B virus clearance and vaccine nonresponsiveness. Hepatology *41*, 1383–1390.

Goncalves, L., Albarran, B., Salmen, S., Borges, L., Fields, H., Montes, H., Soyano, A., Diaz, Y., and Berrueta, L. (2004). The nonresponse to hepatitis B vaccination is associated with impaired lymphocyte activation. Virology *326*, 20–28.

Grimm, S.K., and Ackerman, M.E. (2013). Vaccine design: Emerging concepts and renewed optimism. Curr. Opin. Biotechnol. *24*, 1078–1088.

Guo, X., Zhang, Y., Li, J., Ma, J., Wei, Z., Tan, W., and O'Brien, S.J. (2011). Strong influence of human leukocyte antigen (HLA)-DP gene variants on development of persistent chronic hepatitis B virus carriers in the Han Chinese population. Hepatology *53*, 422–428.

den Haan, J.M., Arens, R., and van Zelm, M.C. (2014). The activation of the adaptive immune system : Cross-talk between antigen-presenting cells, T cells and B cells. Immunol. Lett. *162*, 103–112.

Heinemann, F.M. (2009). HLA Genotyping and Antibody Characterization Using the Luminex ™ Multiplex Technology. Transfus. Med. Hemotherapy *36*, 273–278.

Heininger, U., Gambon, M., Gruber, V., and Margelli, D. (2010). Successful hepatitis B immunization in non- and low responding health care workers. Hum. Vaccin. *6*, 725–728.

Hennig, B.J., and Hall, A.J. (2012). Host genetic factors in hepatitis B infection, liver cancer and vaccination response: A review with a focus on Africa. Sci. Total Environ. *423*, 202–209.

Hennig, B.J., Fielding, K., Broxholme, J., Diatta, M., Mendy, M., Moore, C., Pollard, A.J., Rayco-Solon, P., Sirugo, G., Sande, M.A. Van Der, et al. (2008). Host genetic factors and vaccine-induced immunity to hepatitis B virus infection. PLoS One *3*, e1898.

Hernández-Bernal, F., Aguilar-Betancourt, A., Aljovin, V., Arias, G., Valenzuela, C., Alejo, K.P. De, Hernández, K., Oquendo, O., Figueredo, N., Figueroa, N., et al. (2011). Comparison of four recombinant hepatitis B vaccines applied on an accelerated schedule in healthy adults. Hum. Vaccin. *7*, 1026–1036.

Hino, K., Katoh, Y., Vardas, E., Sim, J., Okita, K., and Carman, W.F. (2001). The effect of introduction of universal childhood hepatitis B immunization in South Africa on the prevalence of serologically negative hepatitis B virus infection and the selection of immune escape variants. Vaccine *19*, 3912–3918.

Hirschhorn, J.N., and Daly, M.J. (2005). Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. *6*, 95–108.

Hodes, R.J. (1997). Aging and the immune system. Immunol. Rev. *160*, 5–8.

Höhler, T., Stradmann-Bellinghausen, B., Starke, R., Sänger, R., Victor, A., Rittner, C., and Schneider, P.M. (2002a). C4A deficiency and nonresponse to hepatitis B vaccination. J. Hepatol. *37*, 387–392.

Höhler, T., Reuss, E., Evers, N., Dietrich, E., Rittner, C., Freitag, C.M., Vollmar, J., Schneider, P.M., and Fimmers, R. (2002b). Differential genetic determination of immune responsiveness to hepatitis B surface antigen and to hepatitis A virus : a vaccination study in twins. Lancet *360*, 991–995.

Höhler, T., Reuss, E., Freitag, C.M., and Schneider, P.M. (2005). A functional polymorphism in the IL-10 promoter influences the response after vaccination with HBsAg and hepatitis A. Hepatology *42*, 72–

76.

Holt, P.G. (1987). Immune and inflammatory function in cigarette smokers. Thorax *42*, 241–249.

Hong, E.P. and Park, J. W. (2012). Sample size and statistical power calculation in genetic association studies. G&I *10*, 117-122.

Hu, L., Zhai, X., Liu, J., Chu, M., Pan, S., Jiang, Y., Wang, H., and J, C. (2012). Genetic variants in human leukocyte antigen/DP--DQ influence both hepatitis B virus clearance and hepatocellular carcinoma development. Hepatology *55*, 1426–1431.

Hu, Z., Liu, Y., Zhai, X., Dai, J., Jin, G., Wang, L., Zhu, L., Yang, Y., Liu, J.J., Chu, M., et al. (2013). New loci associated with chronic hepatitis B virus infection in Han Chinese. Nat. Genet. *45*, 1499–1503.

Huang, J., Xiong, L., Wang, J.I.N., Liu, Y., Zhu, Q., Lei, J.U.N., and Zhou, Z. (2016). Association between the HLA-DQB1 polymorphisms and the susceptibility of chronic hepatitis B : A comprehensive meta-analysis. Biomed. Reports *4*, 557–566.

Immunization Action Coalition (2005). Hepatitis B and the health care worker: CDC answers frequently asked questions about how to protect health care workers. Item #P2109. Available from: www.immunize.org/catg.d/2109hcw.pdf. [Accessed: 02/02/2017].

Ingardia, C.J., Kelley, L., Steinfeld, J.D., and Wax, J.R. (1999). Hepatitis B vaccination in pregnancy: factors influencing efficacy. Obstet. Gynacology *93*, 983–986.

Jack, A.D., Hall, A.J., Maine, N., Mendy, M., and Whittle, H.C. (1999). What level of hepatitis B antibody is protective? J. Infect. Dis. *179*, 489–492.

Jafarzadeh, A., and Shokri, F. (2003). The antibody response to HBs antigen is regulated by coordinated Th1 and Th2 cytokine production in healthy neonates. Clin. Exp. Immunol. *131*, 451–456.

Jafarzadeh, A., and Shokri, F. (2012). TH1 and TH2 responses are influenced by HLA antigens in healthy neonates vaccinated with recombinant hepatitis B vaccine. Iran. J. Allergy, Asthma Immunol. *11*, 308–315.

Jafarzadeh, A., Kardar, G.A., Khoshnoodi, J., and Shokri, F. (2004). Downregulation of IL-12 Production in Healthy Non-Responder Neonates to Recombinant Hepatitis B Vaccine. Iran. Biomed. J. *8*, 41–45.

Janeway, C.A. Jr., Travers, P., Walport, M. and Shlomchik, M.J. (2001). Immunobiology: the immune system in health and disease. 5th ed. Garland Science, New York. The major histocompatibility complex and its functions. Available from: http://www.ncbi.nlm.nih.gov/books/NBK27156/ [First accessed: 1/3/2015].

Jarrosson, L., Kolopp-sarda, M.N., Aguilar, P., Béné, M.C., Lepori, M.L., Vignaud, M.C., Faure, G.C., and Kohler, C. (2004). Most humoral non-responders to hepatitis B vaccines develop HBV-specific cellular immune responses. Vaccine *22*, 3789–3796.

Jensen, E.J., Pedersen, B., Frederiksen, R., and Dahl, R. (1998). Prospective study on the effect of smoking and nicotine substitution on leucocyte blood counts and relation between blood leucocytes and lung function. Thorax *53*, 784–789.

Kamatani, Y., Wattanapokayakit, S., Ochi, H., Kawaguchi, T., Takahashi, A., Hosono, N., Kubo, M., Tsunoda, T., Kamatani, N., Kumada, H., et al. (2009). A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. Nat. Genet. *41*, 591–595.

Kanda, N., and Tamaki, K. (1999). Estrogen enhances immunoglobulin production by human PBMCs. J. Allergy Clin. Immunol. *103*, 282–288.

Kanda, N., Tsuchida, T., and Tamaki, K. (1996). Testosterone inhibits immunoglobulin production by

human peripheral blood mononuclear cells. Clin. Exp. Immunol. *106*, 410–415.

Kardar, G. a., Jeddi-Tehrani, M., and Shokri, F. (2002). Diminished Th1 and Th2 cytokine production in healthy adult nonresponders to recombinant hepatitis B vaccine. Scand. J. Immunol. *55*, 311–314.

Keita, S.O.Y., Kittles, R.A., Royal, C.D.M., Bonney, G.E., Furbert-Harris, P., Dunston, G.M., and Rotimi, C.N. (2004). Conceptualizing human variation. Nat. Genet. *36*, S17-S20.

Kew, M.C. (1996). Progress towards the comprehensive control of hepatitis B in Africa: a view from South Africa. Gut *38*, S31–S36.

Kew, M.C. (2008). Hepatitis B virus infection: the burden of disease in South Africa. South. African J. Epidemiol. Infect. *23*, 4–8.

Khan, F.Y., and Ross, A.J. (2013). Hepatitis B Immunisation amongst doctors and laboratory personnel in KwaZulu-Natal, South Africa. African J. Prim. Heal. Care Fam. Med. *5*, 1–6.

Kiire, C.F. (1996). The epidemiology and prophylaxis of hepatitis B in sub-Saharan Africa: a view from tropical and subtropical Africa. Gut *38*, S5–S12.

Kim, S.U., Song, K.J., Chang, H.Y., Shin, E.-C., Park, J.Y., Han, K.-H., Chon, C.Y., and Ahn, S.H. (2013a). Association between IL28B Polymorphisms and Spontaneous Clearance of Hepatitis B Virus Infection. PLoS One *8*, e69166.doi: 10.137/journal.pone.0069166.

Kim, Y.J., Kim, H.Y., Lee, J.-H., Yu, S.J., Yoon, J.-H., Lee, H.-S., Kim, C.Y., Cheong, J.Y., Cho, S.W., Park, N.H., et al. (2013b). A genome-wide association study identified new variants associated with the risk of chronic hepatitis B. Hum. Mol. Genet. *22*, 4233–4238.

Kimman, T.G., Vandebriel, R.J., and Hoebee, B. (2007). Genetic Variation in the Response to Vaccination. Community Genet. *10*, 201–217.

Klein, S.L., Marriott, I., and Fish, E.N. (2015). Sex-based differences in immune function and responses to vaccination. Trans. R. Soc. Trop. Med. Hyg. *109*, 9–15.

Krajden, M., McNabb, G., and Petric, M. (2005). The laboratory diagnosis of hepatitis B virus. Can. J. Infect. Dis. Med. Microbiol. *16*, 65–72.

Kramvis, A., Kew, M., and Francois, G. (2005). Hepatitis B virus genotypes. Vaccine *23*, 2409–2423.

Kulkarni, S., Savan, R., Qi, Y., Gao, X., Yuki, Y., Bass, S.E., Martin, M.P., Hunt, P., Deeks, S.G., Telenti, A., et al. (2011). Differential microRNA regulation of HLA-C expression and its association with HIV control. Nature *472*, 495–498.

Lambert, N.D., Haralambieva, I.H., Kennedy, R.B., Ovsyannikova, I.G., Pankratz, V.S., and Poland, G.A. (2014). Polymorphisms in HLA-DPB1 Are Associated With Differences in Rubella Virus – Specific Humoral Immunity After Vaccination. J. Infect. Dis. *211*, 1–8.

Lampertico, P., Vigan, M., Cheroni, C., Facchetti, F., Invernizzi, F., Valveri, V., Soffredini, R., Abrignani, S., Francesco, R. De, and Colombo, M. (2013). IL28B Polymorphisms Predict Interferon-Related Hepatitis B Surface Antigen Seroclearance in Genotype D Hepatitis B e Antigen–Negative Patients With Chronic Hepatitis B. Hepatology *57*, 890–896.

Lang, T.J. (2004). Estrogen as an immunomodulator. Clin. Immunol. *113*, 224–230.

Langö-Warensjö, A., Cardell, K., and Lindblom, B. (1998). Haplotypes comprising subtypes of the DQB1*06 allele direct the antibody response after immunisation with hepatitis B surface antigen. Tissue Antigens *52*, 374–380.

Larsen, C.E., Xu, J., Lee, S., Dubey, D.P., Uko, G., Yunis, E.J., and Alper, C. a (2000). Complex cytokine responses to hepatitis B surface antigen and tetanus toxoid in responders, nonresponders and subjects naive to hepatitis B surface antigen. Vaccine *18*, 3021–3030.

Ledderose, C., Heyn, J., Limbeck, E., and Kreth, S. (2011). Selection of reliable reference genes for quantitative real-time PCR in human T cells and neutrophils. BMC Res. Notes *4*, 427.

Li, W., Jiang, Y., Jin, Q., Shi, X., Jin, J., Gao, Y., Pan, Y., Zhang, H., Jiang, J., and Niu, J. (2011). Expression and gene polymorphisms of interleukin 28B and hepatitis B virus infection in a Chinese Han population. Liver Int. *31*, 1118–1126.

Li, Z.-K., Nie, J.-J., Li, J., and Zhuang, H. (2013). The effect of HLA on immunological response to hepatitis B vaccine in healthy people: A meta-analysis. Vaccine *31*, 4355–4361.

Lin, C., and Kao, J. (2013). Hepatitis B Virus Genotypes : Clinical Relevance and Therapeutic Implications. Curr. Hepat. Rep. *12*, 124–132.

Lin, Y.J., Lan, Y.C., Wan, L., Lin, T.H., Chen, D.Y., Tsai, C.H., Liu, C.S., Hsueh, K.C., and Tsai, F.J. (2011). Serological surveillance and IL-10 genetic variants on anti-HBs titers: Hepatitis B vaccination 20years after neonatal immunization in Taiwan. Clin. Chim. Acta *412*, 766–773.

Lin, Y.J., Lan, Y.C., Huang, Y.C., Lin, T.H., Huang, S.M., Lai, C.C., Liu, C.S., Lin, C.W., Chen, S.Y., and Tsai, F.J. (2012). Effects of cytokine and cytokine receptor gene variation on high anti-HB titers: Following up on Taiwan's neonatal hepatitis B immunization program. Clin. Chim. Acta *413*, 1194–1198.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402–408.

Locarnini, S. (2004). Molecular virology of hepatitis B virus. Semin. Liver Dis. *24 Suppl 1*, 3–10.

Lombard, Z., Brune, A.E., Hoal, E.G., Babb, C., Van Helden, P.D., Epplen, J.T., and Bornman, L. (2006). HLA class II disease associations in southern Africa. Tissue Antigens *67*, 97–110.

Lu, Y., Wu, X., Huang, H., and Dai, L. (2010). Allele polymorphisms of interleukin-10 and hepatitis B, C virus infection. Chin. Med. J. (Engl). *123*, 1338–1344.

Lu, Y., Wu, Z., Peng, Q., Ma, L., Zhang, X., Zhao, J., Qin, X., and Li, S. (2014). Role of IL-4 Gene Polymorphisms in HBV-Related Hepatocellular Carcinoma in a Chinese Population. PLoS One *9*, e110061.

Mangalam, A.K., Taneja, V., and David, C.S. (2013). HLA class II molecules influence susceptibility versus protection in inflammatory diseases by determining the cytokine profile. J Immunol *190*, 513–518.

Martin, M.P., Qi, Y., Goedert, J.J., Hussain, S.K., Kirk, G.D., Hoots, W.K., Buchbinder, S., Carrington, M., and Thio, C.L. (2010). IL28B Polymorphism Does Not Determine Outcomes of Hepatitis B Virus or HIV Infection. J. Infect. Dis. *202*, 1749–1753.

Martinetti, M., Cuccia, M., Daielli, C., Ambroselli, F., Gatti, C., Pizzochero, C., Belloni, C., Orsolini, P., and Salvaneschi, L. (1995). Anti-HBV neonatal immunization with recombinant vaccine. Part II. Molecular basis of the impaired alloreactivity. Vaccine *13*, 555–560.

Martinetti, M., De Silvestri, a, Belloni, C., Pasi, a, Tinelli, C., Pistorio, a, Salvaneschi, L., Rondini, G., Avanzini, M. a, and Cuccia, M. (2000). Humoral response to recombinant hepatitis B virus vaccine at birth: role of HLA and beyond. Clin. Immunol. *97*, 234–240.

Mbarek, H., Ochi, H., Urabe, Y., Kumar, V., Kubo, M., Hosono, N., Takahashi, A., Kamatani, Y., Miki, D., Abe, H., et al. (2011). A genome-wide association study of chronic hepatitis B identified novel risk locus in a Japanese population. Hum. Mol. Genet. *20*, 3884–3892.

Michel, M., and Tiollais, P. (2010). Hepatitis B vaccines : Protective efficacy and therapeutic potential. Pathol. Biol. *58*, 288–295.

Milich, D.R. (2002). Influence of C4A deficiency on nonresponse to HBsAg vaccination: A new

immune response gene. J. Hepatol. *37*, 396–399.

Milich, D.R., and Leroux-Roels, G.G. (2003). Immunogenetics of the response to HBsAg vaccination. Autoimmun. Rev. *2*, 248–257.

Milner, J.J., and Beck, M.A. (2012). The impact of obesity on the immune response to infection. Proc. Nutr. Soc. *71*, 298–306.

Mineta, M., Tanimura, M., Tana, T., Yssel, H., Kashiwagi, S., and Sasazuki, T. (1996). Contribution of HLA class I and class II alleles to the regulation of antibody production to hepatitis B surface antigen in humans. Int. Immunol. *8*, 525–531.

Mingyao, L., Reilly, M.P., Rader, D.J., and Wang, L.S. (2010). Correcting population stratification in genetic association studies using a phylogenetic approach. Bioinformatics. *26*, 798-806.

Mountain, J.L., and Risch, N. (2004). Assessing genetic contributions to phenotypic differences among 'racial' and 'ethnic' groups. Nat. Genet. *36*. S48-S53.

Mphahlele, M.J., Tshatsinde, E.A., Burnett, R.J., and Aspinall, S. (2002). Protective efficacy and antibody follow-up of hepatitis B vaccine within the South African expanded programme on immunisation. S. Afr. Med. J. *92*, 612–613.

Mulder, N.J., Adebiyi, E., Alami, R., Benkahla, A., Brandful, J., Doumbia, S., Everett, D., Fadlelmola, F.M., Gaboun, F., Gaseitsiwe, S., et al. (2016). 3ABioNet, a sustainable pan-African bioinformatics network for human heredity and health in Africa. Genome Res. *26*, 271-277.

Myles, S., Davison, D., Barrett, J., Stoneking, M., and Timpson, N. (2008). Worldwide population differentiation at disease-associated SNPs. BMC Med. Gemomics. *1*, 22.

Nassal, M., and Schaller, H. (1993). Hepatitis B virus replication. Trends Microbiol. *1*, 221–228.

Newport, M.J., Goetghebuer, T., Weiss, H. a, Whittle, H., Siegrist, C., and Marchant, a (2004). Genetic regulation of immune responses to vaccines in early life. Genes Immun. *5*, 122–129.

Newport, M.J., Goetghebuer, T., and Marchant, A. (2005). Hunting for immune response regulatory genes: vaccination studies in infant twins. Expert Rev. Vaccines *4*, 739–746.

Nishida, N., Sawai, H., Matsuura, K., Sugiyama, M., Ahn, S.H., Park, J.Y., Hige, S., Kang, J.H., Suzuki, K., Kurosaki, M., et al. (2012). Genome-wide association study confirming association of HLA-DP with protection against chronic hepatitis B and viral clearance in Japanese and Korean. PLoS One *7*, 1–8.

Noble, R.C., and Penny, B.B. (1975). Comparison of leukocyte count and function in smoking and nonsmoking Comparison of Leukocyte Count and Function in Smoking and Nonsmoking Young Men. Infect. Immun. *12*, 550–555.

Obami-Itou V., Barry M.F., Raphenon G., Thiam A., Ndiaye R., Ndiaye M., et al. (2000). Serological HLA class I alleles in Senegalese blood donors detected HBs Ag positive. Immunol. Lett. *74*, 229–32.

O'Brien, T.R.O., Kohaar, I., Pfeiffer, R.M., Maeder, D., Yeager, M., Schadt, E.E., and Prokunina-Olsson (2011). Risk alleles for chronic hepatitis B are associated with decreased mRNA expression of HLA-DPA1 and HLA-DPB1 in normal human liver. Genes Immun. *12*, 428–433.

Pan, L., Zhang, L., Zhang, W., Wu, X., Li, Y., Yan, B., and Zhu, X. (2013). A genome-wide association study identifies polymorphisms in the HLA-DR region associated with non-response to hepatitis B vaccination in Chinese Han populations. Hum. Mol. Genet. 1–10.

Paximadis, M., Mathebula, T.Y., Gentle, N.L., Vardas, E., Colvin, M., Gray, C.M., Tiemessen, C.T., and Puren, A. (2012). Human leukocyte antigen class I (A, B, C) and II (DRB1) diversity in the black and Caucasian South African population. Hum. Immunol. *73*, 80–92.

Petersdorf, E.W., Malkki, M., O'hUigin, C., Carrington, M., Gooley, T., Haagenson, M.D., Horowitz, M.M., Spellman, S.R., Wang, T., Ph, D., et al. (2015). High HLA-DP Expression and Graft-versus-Host Disease. N. Engl. J. Med. *373*, 599–609.

Piehler, A.P., Grimholt, R.M., Øvstebø, R., and Berg, J.P. (2010). Gene expression results in lipopolysaccharide-stimulated monocytes depend significantly on the choice of reference genes. BMC Immunol. *11*, 21.

Pirie, F.J., Hammond, M.G., Motala, A.A., and Omar, M.A. (2001). HLA class II antigens in South African blacks with type I diabetes. Tissue Antigens *57*, 348–352.

Plotkin, S.A. (2010). Correlates of Protection Induced by Vaccination. Clin. Vaccine Immunol. *17*, 1055–1065.

Plotkin, S. a (2008). Vaccines: correlates of vaccine-induced immunity. Clin. Infect. Dis. *47*, 401–409.

Png, E., Thalamuthu, A., Ong, R.T.H., Snippe, H., Boland, G.J., and Seielstad, M. (2011). A genome-wide association study of hepatitis B vaccine response in an Indonesian population reveals multiple independent risk variants in the HLA region. Hum. Mol. Genet. *20*, 3893–3898.

Pociot, F., Mølvig, J., Wogensen, L., Worsaae, H., and Nerup, J. (1992). A TaqI polymorphism in the human interleukin-1 beta (IL-1 beta) gene correlates with IL-1 beta secretion in vitro. Eur. J. Clin. Invest. *22*, 396–402.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.

Radke, L., Giese, C., Lubitz, A., Hinderlich, S., Sandig, G., Hummel, M., and Frohme, M. (2014). Reference gene stability in peripheral blood mononuclear cells determined by qPCR and NanoString. Microchim. Acta *181*, 1733–1742.

Rendi-wagner, P., Kundi, M., Stemberger, H., Wiedermann, G., Holzmann, H., Hofer, M., Wiesinger, K., and Kollaritsch, H. (2001). Antibody-response to three recombinant hepatitis B vaccines : comparative evaluation of multicenter travel-clinic based experience. Vaccine *19*, 2055–2060.

Risch, N., Burchard, E., Ziv, E., and Tang, H. (2002). Categorization of humans in biomedical research: genes, race and disease. Genome Biol. *3*, comment 2007.

Roh, E.Y., Youn, J.H., In, J.W., Lee, N., Shin, S., and Song, E.Y. (2016). Association of HLA-DP variants with the responsiveness to Hepatitis B virus vaccination in Korean Infants. Vaccine *34*, 2602–2607.

Romano, L., Paladini, S., Van Damme, P., and Zanetti, A.R. (2011). The worldwide impact of vaccination on the control and protection of viral hepatitis B. Dig. Liver Dis. *43*, S2–S7.

Roome, A.J., Walsh, S.J., Cartter, M.L., and Hadler, J.L. (1993). Hepatitis B vaccine responsiveness in Connecticut public safety personnel. Jama *270*, 2931–2934.

Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A., and Feldman, M.W. (2002), Genetic structure of human populations. Science *298*, 2381-2385.

Rosenberg, N.A., Huang, L., Jewett, E.M., Szpiech, Z.A., Jankovic, I., and Boehnke, M. (2010). Genome-wide association studies in diverse populations. Nat. Rev. Genet. *11*, 356-366.

Roukens, A.H., and Visser, L.G. (2011). Hepatitis B vaccination strategy in vaccine low and non-responders: A matter of quantity of quality? Hum. Vaccin. *7*, 654–657.

Ryckman, K.K., Fielding, K., Hill, A. V., Mendy, M., Rayco-Solon, P., Sirugo, G., van der Sande, M. a., Waight, P., Whittle, H.C., Hall, A.J., et al. (2010). Host genetic factors and vaccine-induced immunity to HBV infection: Haplotype analysis. PLoS One *5*, e12273.

Schoub, B.D., Matai, U., Singh, B., Blackburn, N.K., and Levin, J.B. (2002). Universal immunization of infants with low doses of a low-cost , plasma-derived hepatitis B vaccine in South Africa. Bull. World Health Organ. *80*, 277–281.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., and Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol. Biol. *7*, 3.

Seeger, C., and Mason, W.S. (2000). Hepatitis B Virus Biology. Microbiol. Mol. Biol. Rev. *64*, 51–68.

Shaw, F.E., Guess, H.A., Roets, J.M., Mohr, F.E., Coleman, P.J., Mandel, E.J., Roehm, R.R., Talley, W.S., and Hadler, S.C. (1989). Effect of anatomic injection site, age and smoking on the immune response to hepatitis B vaccination. Vaccine *7*, 425–430.

Shekhar, S., and Yang, X. (2012). The darker side of follicular helper T cells : from autoimmunity to immunodeficiency. Cell. Mol. Immunol. *9*, 380–385.

Shepard, C.W., Simard, E.P., Finelli, L., Fiore, A.E., and Bell, B.P. (2006). Hepatitis B virus infection: Epidemiology and vaccination. Epidemiol. Rev. *28*, 112–125.

Shi, X., Chi, X., Pan, Y., Gao, Y., Li, W., Yang, C., Zhong, J., Xu, D., Zhang, M., Minuk, G., et al. (2015). IL28B Is Associated with Outcomes of Chronic HBV Infection. Yonsei Med. J. *56*, 625–633.

Shokrgozar, M. a, Shokri, F., and Cell, N. (2001). Enumeration of hepatitis B surface antigen-specific B lymphocytes in responder and non-responder normal individuals vaccinated with recombinant hepatitis B surface antigen. Immunology *104*, 75–79.

Sidney, J., Steen, A., Moore, C., Ngo, S., Sidney, J., Steen, A., Moore, C., Ngo, S., Chung, J., and Peters, B. (2010). Five HLA-DP Molecules Frequently Expressed in the Worldwide Human Population Share a Common HLA Supertypic Binding Specificity. J. Immunol. *184*, 2492–2503.

Simani, O.E., Leroux-roels, G., Franc, G., Burnett, R.J., Meheus, A., and Mphahlele, M.J. (2009). Reduced detection and levels of protective antibodies to hepatitis B vaccine in under 2-year-old HIV positive South African children at a paediatric outpatient clinic. Vaccine *27*, 146–151.

Soroosh, P., Shokri, F., Azizi, M., and Jeddi-Tehrani, M. (2003). Analysis of T-cell receptor beta chain variable gene segment usage in healthy adult responders and nonresponders to recombinant hepatitis B vaccine. Scand. J. Immunol. *57*, 423–431.

Spielman, R.S., Bastone, L.A., Burdick, J.T., Morley, M., Ewens, W.J., and Cheung, V.G. (2007). Common genetic variants account for differences in gene expression among ethnic groups. Nat. Genet. *39*, 226–231.

Steuer, M.K., Oudshoorn, M., Brenden, M., Fimmers, R., Neugebauer, M., Toit, E., Baur, M.P., and Mauff, G. (1994). An estimate on the frequency of duplicated haplotypes and silent alleles of human C4 protein polymorphism. Tissue Antigens *43*, 88–94.

Struve, J., Aronsson, B., Frenning, B., Forsgren, M., and Weiland, O. (1994). Seroconversion after Additional Vaccine Doses to Non-responders to Three Doses of Intradermally or Intramuscularly Administered Recombinant Hepatitis B Vaccine. Scand. J. Infect. Dis. *26*, 468–470.

Suchard, M.S. (2012). Missing: A diagnostic technique to enumerate antigen-specific T cells. Crit. Rev. Oncol. Hematol. *83*, 276–282.

Sunbul, M. (2014). Hepatitis B virus genotypes : Global distribution and clinical importance. World J. Gastroenterol. *20*, 5427–5434.

Statistics South Africa (2016). Statistical release P0302. Mid-year population estimates 2016. Available at: http://www.statssa.gov.za/publications/P0302/P03022016.pdf (Accessed: 29/06/2017).

Szmuness, W., Stevens, C.E., Harley, E.J., Zang, E.A., Oleszko, W.R., William, D.C., Sadovsky, R.,

Morrison, J.M., and Kellner, A. (1980). Demonstration of efficacy in a controlled clinical trial in a high-risk population in the United States. N. Engl. J. Med. *303*, 833–841.

Szmuness, W., Stevens, C.E., Zang, E.A., Harley, J., and Kellner, A. (1981). A Controlled Clinical Trial of the Efficacy of the Hepatitis B Vaccine (Heptavax B): A Final Report. Hepatology *1*, 377–385.

Szmuness, W., Stevens, C.E., Harley, E.J., Zang, E.A., Alter, H.J., Taylor, P.E., DeVera, A., Chen, G.T.S., Kellner, A., and Group, T.D.V.T.S. (1982). Hepatitis B Vaccine in medical staff of hemodialysis units - efficacy and subtype cross-protection. N. Engl. J. Med. *307*, 1481–1486.

Tang, H. (2006). Confronting ethnicity-specific disease risk. Nat. Genet. *38*, 13-15.

Teo, Y.Y., Small, K.S., and Kwiatkowski, D.P. (2010). Methodological challenges of genome-wide association analysis in Africa. Nat. Rev. Genet. *11*, 149-160.

The 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. Nature *526*, 68-74. doi:10.1038/nature15393.

The MHC sequencing consortium (1999). Complete sequence and gene map of a human major histocompatibility complex. Nature *401*, 921–923.

Thi Tran, H.T., Takeshima, Y., Surono, A., Yagi, M., Wada, H., and Matsuo, M. (2005). A G-to-A transition at the fifth position of intron-32 of the dystrophin gene inactivates a splice-donor site both in vivo and in vitro. Mol. Genet. Metab. *85*, 213–219.

Thio, C.L., Carrington, M., Marti, D., O'Brien, S.J., Vlahov, D., Nelson, K.E., Astemborski, J., and Thomas, D.L. (1999). Class II HLA alleles and hepatitis B virus persistence in African Americans. J. Infect. Dis. *179*, 1004–1006.

Thio, C.L., Thomas, D.L., Karacki, P., Gao, X., Marti, D., Kaslow, R. a, Goedert, J.J., Hilgartner, M., Strathdee, S. a, Duggal, P., et al. (2003). Comprehensive Analysis of Class I and Class II HLA Antigens and Chronic Hepatitis B Virus Infection. J. Virol. *77*, 12083–12087.

Thomas, R., Apps, R., Qi, Y., Gao, X., Male, V., O'hUigin, C., O'Connor, G., Ge, D., Fellay, J., Martin, J.N., et al. (2009). HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. Nat. Genet. *41*, 1290–1294.

Thomas, R., Thio, C.L., Apps, R., Qi, Y., Gao, X., Marti, D., Stein, J.L., Soderberg, K. a., Moody, M. a., Goedert, J.J., et al. (2012). A Novel Variant Marking HLA-DP Expression Levels Predicts Recovery from Hepatitis B Virus Infection. J. Virol. *86*, 6979–6985.

Thursz, M.R., Kwiatkowski, D., Allsopp, C.E., Greenwood, B.M., Thomas, H.C., and Hill, A. V (1995). Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. N. Engl. J. Med. *332*, 1065–1069.

Thursz M., Kwiatkowski D., McGuire W., Hill A.V., Thomas H.C. (1996). TNF promoter polymorphism and the outcome of hepatitis B virus infection. Gut *39*, A19.

Tikly, M., Rands, A., Mchugh, N., Wordsworth, P., and Welsh, K. (2004). Human leukocyte antigen class II associations with systemic sclerosis in South Africans. Tissue Antigens *63*, 487–491.

Tohme, R.A., Awosika-Olumo, D., Nielsen, C., Khuwaja, S., Scott, J., Xing, J., Drobeniuc, J., Hu, D.J., Turner, C., Wafeeg, T., et al. (2011). Evaluation of hepatitis B vaccine immunogenicity among older adults during an oubreak response in assisted living facilities. Vaccine *29*, 9316–9320.

Tripathy, S., Sati, H.C., Saha, S., Shankar, R., and Singh, V.K. (2011). Study of immune response after hepatitis B vaccination in medical students and health care workers. Indian J. Prev. Soc. Med. *42*, 314–321.

Tsebe, K. V., Burnett, R.J., Hlungwani, N.P., Sibara, M.M., Venter, P. a., and Mphahlele, M.J. (2001). The first five years of universal hepatitis B vaccination in South Africa: Evidence for elimination of

HBsAg carriage in under 5-year-olds. Vaccine *19*, 3919–3926.

Tshabalala, M., Mellet, J., and Pepper, M.S. (2015). Human Leukocyte Antigen Diversity : A Southern African Perspective. J. Immunol. Res. doi:10.1155/2015/746151.

Ed. V. Motaze and M.S. Suchard (2016). Vaccine information for mothers and caregivers. First edition. Published by Ideas Wise and Wonderful for the National Institute for Communicable Diseases. ISBN 978-0-620733-15-1.

Velu, V., Saravanan, S., Nandakumar, S., Shankar, E.M., Vengatesan, A., Jadhav, S.S., Kulkarni, P.S., and Thyagarajan, S.P. (2008). Relationship between T-lymphocyte cytokine levels and sero-response to hepatitis B vaccines. World J. Gastroenterol. *14*, 3534–3540.

Wang, L. (2011). Evaluation of Genetic Susceptibility Loci for Chronic Hepatitis B in Chinese : Two Independent Case-Control Studies. 11–16.

Wang, P., and Zheng, S.G. (2013). Regulatory T cells and B cells: implication on autoimmune diseases. Int. J. Clin. Exp. Pathol. *6*, 2668–2674.

Wang, C., Tang, J., Song, W., Lobashevsky, E., Wilson, C.M., and Kaslow, R. a. (2004). HLA and Cytokine Gene Polymorphisms Are independently Associated with Responses to Hepatitis B Vaccination. Hepatology *39*, 978–988.

Wang, J., Wang, Y., Wang, H., Hao, X., Wu, Y., and Guo, J. (2014). Selection of Reference Genes for Gene Expression Studies in Porcine Whole Blood and Peripheral Blood Mononuclear Cells under Polyinosinic : Polycytidylic Acid Stimulation. Asian Australas. J. Anim. Sci. *27*, 471–478.

Wang, Y., Xu, P., Zhu, D., Zhang, S., Bi, Y., Hu, Y., and Zhou, Y.H. (2012). Association of polymorphisms of cytokine and TLR-2 genes with long-term immunity to hepatitis B in children vaccinated early in life. Vaccine *30*, 5708–5713.

Wataya, M., Sano, T., Kamikawaji, N., Tana, T., Yamamoto, K., and Sasazuki, T. (2001). Comparative analysis of HLA restriction and cytokine production in hepatitis B surface antigen-specific T cells from low- and high-antibody responders in vaccinated humans. J. Hum. Genet. *46*, 197–206.

Weinberg, G.A., and Szilagyi, P.G. (2010). Vaccine Epidemiology : Efficacy , Effectiveness , and the Translational Research Roadmap. J. Infect. Dis. *201*, 1607–1610.

World Health Organization (2009). Hepatitis B vaccines: WHO position paper. Wkly. Epidemiol. Rec. *84*, 405–420.

Wu, T., Chu, C., Liao, H.C., Lin, S., Ho, T., Lin, M., Lin, H.H., and Wang, L. (2014). HLA-DPB1 and anti-HBs titer kinetics in hepatitis B booster recipients who completed primary hepatitis B vaccination during infancy. Genes Immun. *15*, 47–53.

Wu, T., Chen, C., Lai, S., Lin, H.H., Chu, C., and Wang, L. (2015). SNP rs7770370 in HLA - DPB1 loci as a major genetic determinant of response to booster hepatitis B vaccination : Results of a genome-wide association study. J. Gastroenterol. Hepatol. *30*, 891–899.

Wu, T.W., Chu, C.C., Ho, T.Y., Chang Liao, H.W., Lin, S.K., Lin, M., Lin, H.H., and Wang, L.Y. (2013). Responses to booster hepatitis B vaccination are significantly correlated with genotypes of human leukocyte antigen (HLA)-DPB1 in neonatally vaccinated adolescents. Hum. Genet. *132*, 1131–1139.

Yan, Z.H., Fan, Y., Wang, X.H., Mao, Q., Deng, G.H., and Wang, Y.M. (2012). Relationship between HLA-DR gene polymorphisms and outcomes of hepatitis B viral infections: A meta-analysis. World J. Gastroenterol. *18*, 3119–3128.

Yang, S., Tian, G., Cui, Y., Ding, C., Deng, M., Yu, C., Xu, K., Jingjing, R., Yao, J., Li, Y., et al. (2016). Factors influencing immunologic response to hepatitis B vaccine in adults. Sci. Rep. *6*, doi: 10.1038/srep27251.

Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L., et al. (2016). Ensembl 2016. Nucleic Acids Res. *44*, D710–D716.

Young, K.M., Gray, C.M., and Bekker, L.G. (2013). Is obesity a risk factor for vaccine non-responsiveness? PLoS One *8*, 1–6.

Yu, L., Cheng, Y., Cheng, M., Yao, Y., Zhang, Q., Zhao, X., Liu, H., Hu, Y., Mu, M., Wang, B., et al. (2015). Quantitative assessment of common genetic variations in HLA-DP with hepatitis B virus infection , clearance and hepatocellular carcinoma development. Sci. Rep. *5*, DOI: 10.1038/srep14933.

Yucesoy, B., Sleijffers, A., Kashon, M., Garssen, J., de Gruijl, F.R., Boland, G.J., van Hattum, J., Simeonova, P.P., Luster, M.I., and van Loveren, H. (2002). IL-1beta gene polymorphisms influence hepatitis B vaccination. Vaccine *20*, 3193–3196.

Yucesoy, B., Johnson, V.J., Fluharty, K., Kashon, M.L., Slaven, J.E., Wilson, N.W., Weissman, D.N., Biagini, R.E., Germolec, D.R., and Luster, M.I. (2009). Influence of cytokine gene variations on immunization to childhood vaccines. Vaccine *27*, 6991–6997.

Yucesoy, B., Talzhanov, Y., Johnson, V.J., Wilson, N.W., Biagini, R.E., Wang, W., Frye, B., Weissman, D.N., Germolec, D.R., Luster, M.I., et al. (2013). Genetic variants within the MHC region are associated with immune responsiveness to childhood vaccinations. Vaccine *31*, 5381–5391.

Zemmour, J., and Parham, P. (1992). Distinctive polymorphism at the HLA-C locus: implications for the expression of HLA-C. J. Exp. Med. *176*, 937–950.

Zheng, Z., Li, X., Li, Z., and Ma, X.-C. (2013). IL-4 - 590C/T Polymorphism and Susceptibility to Liver Disease: A Meta-Analysis and Meta-Regression. DNA Cell Biol. *32*, 443–450.

Zuckerman, J.N. (1996). Nonresponse to hepatitis B vaccines and the kinetics of anti-HBs production. J. Med. Virol. *50*, 283–288.

APPENDIX A
Ethical Clearance

R14/49 Dr Debra de Assis Rosa et al

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M140954

NAME: (Principal Investigator)	Dr Debra de Assis Rosa et al
DEPARTMENT:	School of Molecular and Cell Biology National Institute of Communicable Diseases
PROJECT TITLE:	Immuogenetic Mechanisms of Non-Response to Hepatitis B Vaccince in South African Health-Workers
DATE CONSIDERED:	03/10/2014
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	

APPROVED BY: Professor Cleaton-Jones, Chairperson, HREC (Medical)

DATE OF APPROVAL: 21/11/2014

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Secretary in Room 10004, 10th floor, Senate House, University.
I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**.

Principal Investigator Signature Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

APPENDIX B
Participant Questionnaire

Participant questionnaire.
Please fill in, and circle the correct answer where required. Please hand this completed questionnaire to the nurse upon blood draw and enrolment into the study.

Date of questionnaire completion by participant: _____(YYYY/MMM/DD)

Participant details:

Name: _____

Date of birth: _____/____/____ (YYYY/MMM/DD)

Gender _____

Ethnicity: _____

Weight (kg): _____

Study number: _____ (to be filled in by investigator)

Infection history

Have you ever been infected with hepatitis? yes / no / don't know

If yes, when? _____ / current / don't know

If yes, what type of hepatitis? A / B / C / D / E / don't know

If yes, do you know if you are still a chronic carrier? yes / no / don't know

Are you currently taking medication for hepatitis infection? _____

Vaccination history

Have you been previously vaccinated for Hep B? yes / no / don't know

Dates previously vaccinated? _____

Place previously vaccinated? _____

Number of doses received? 1 / 2 / 3 / more than 3 / don't remember

Can you provide any proof of vaccination e.g. immunisation card if necessary? yes / no

Do you know if you are a responder / non-responder to HBV vaccination? Yes / no

If known non-responder: please explain how you came to know this? e.g. Who diagnosed this

_____

If known non-responder: please explain how this was managed? e.g. further booster.

_____

_____

Do you know if you are a responder / non-responder to other vaccines? yes / no / don't know. If yes, to which other vaccines are you a non-responder?

_____

Participant name: _____

Study number: _____ (to be filled in by investigator)

Factors affecting immune responses

Do you smoke currently? Yes / no. If yes, number of cigarettes per day: _____

Did you smoke at during the period of vaccination: yes / no

If yes, number of cigarettes per day at that time: _____

Any previous or current malignancy? Yes / no

If yes: further details (type and date) _____

If yes: being treated? yes / no / type of treatment: _____

_____

Do you have any primary immunodeficiency? yes / no / don't know

If yes, type? _____

If yes: are you currently taking medication for this disease? Yes / no / details: _____

Do you suffer any auto-immune diseases? yes / no / don't know

If yes, what type? Eczema / asthma / inflammatory bowel disease / arthritis / celiac disease / other: _____

If yes: are you currently taking medication for this disease? Yes / no / details:

_____

Current infections / illness? _____

Do you have any liver disease? _____

Current medication _____

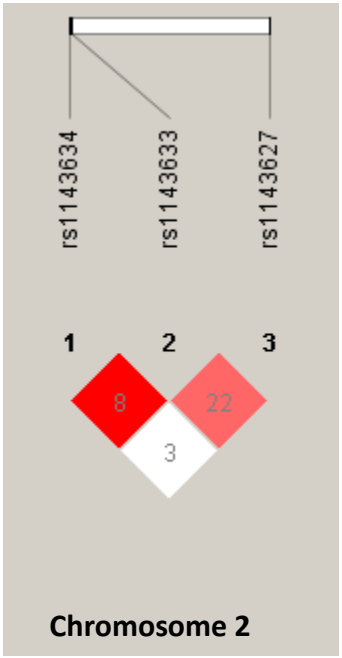
APPENDIX C

SNPs (of the 31 Selected SNPs) Captured by TagSNPs

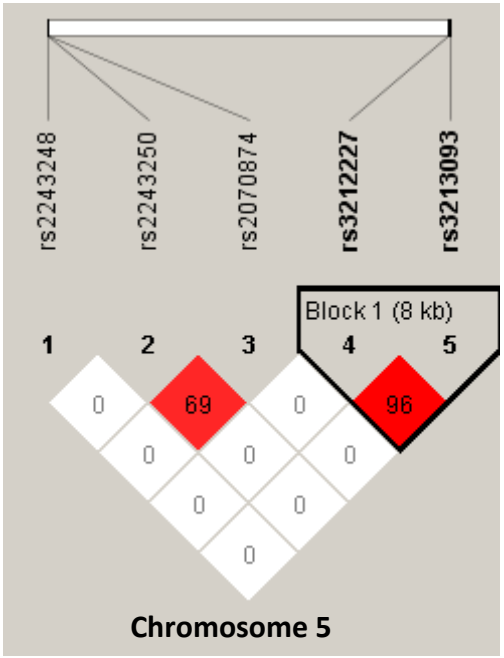
Chromosome	TagSNP	Captured SNPs
6	rs2395309	rs3077, rs2301220
6	rs3130186	rs9277535, rs3117229

APPENDIX D

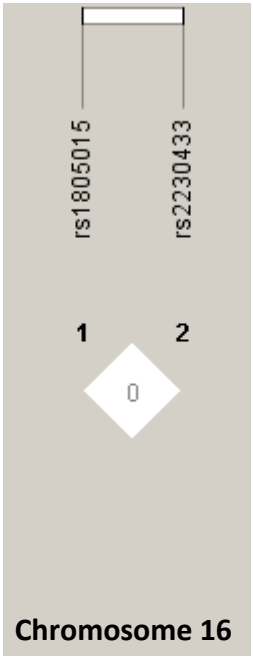
LD Maps Generated by Haploview, r^2 values



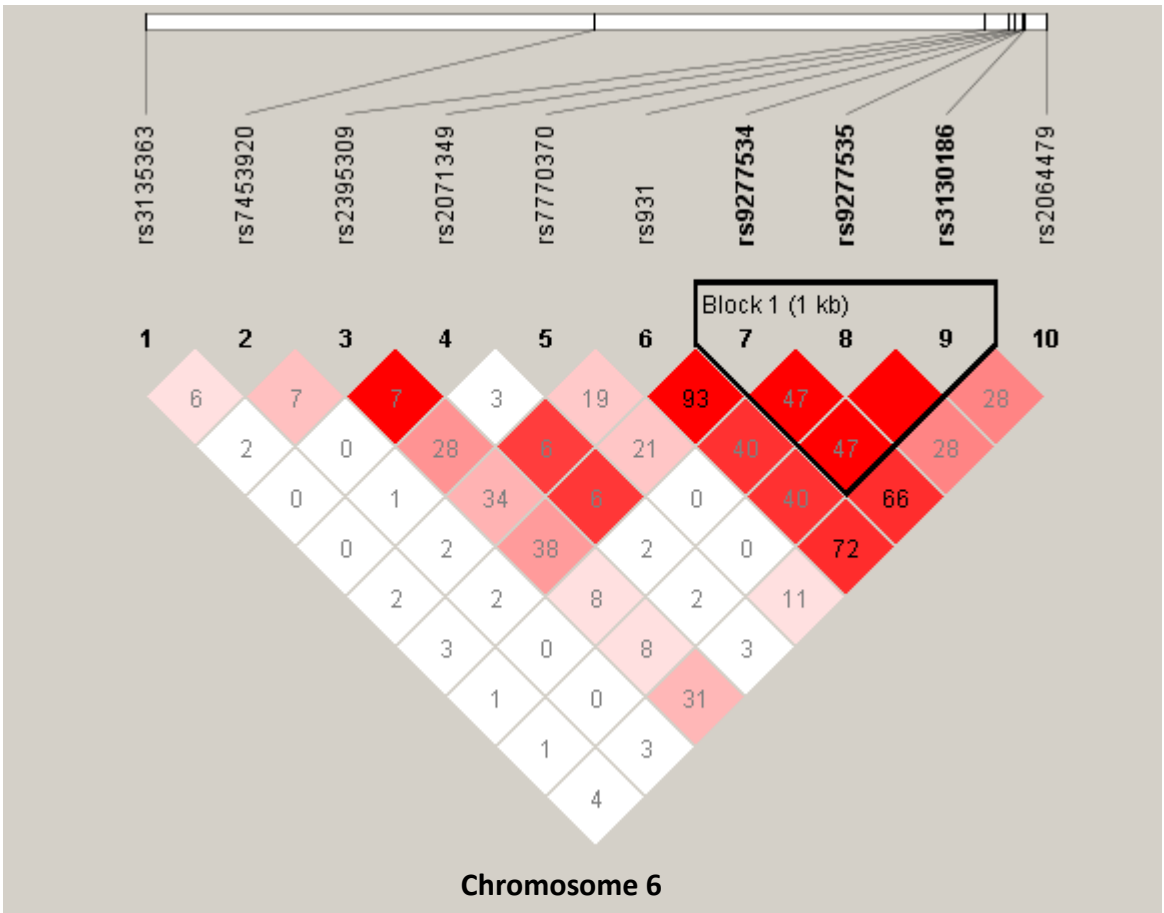
Chromosome 2



Chromosome 5



Chromosome 16



Chromosome 6

APPENDIX E

SNP Hardy-Weinberg Equilibrium P-Values

SNP ID	Total cohort		African		Caucasian		Asian		Coloured	
	HWE (Unaff)	HWE (All)	HWE (Unaff)	HWE (All)	HWE (Unaff)	HWE (All)	HWE (Unaff)	HWE (All)	HWE (Unaff)	HWE (All)
rs1800896	1.00	1.00	0.75	0.75	1.00	1.00	1.00	1.00	1.00	1.00
rs1143634	0.37	0.41	1.00	1.00	0.50	0.75	1.00	1.00	1.00	1.00
rs1143633	0.13	0.06	1.00	1.00	0.31	0.20	0.54	0.27	1.00	1.00
rs1143627	0.04	0.02	1.00	1.00	0.80	0.81	0.07	0.07	1.00	1.00
rs2069763	0.47	0.64	0.06	0.10	0.78	0.60	0.66	0.66	1.00	1.00
rs2243248	0.17	0.12	1.00	0.77	0.56	0.22	1.00	1.00	1.00	1.00
rs2243250	0.01	0.02	0.52	0.74	0.66	1.00	0.58	0.62	1.00	1.00
rs2070874	0.003	0.01	0.16	0.17	0.66	1.00	0.58	0.62	1.00	1.00
rs3212227	0.38	0.39	0.71	0.71	0.20	0.20	0.66	1.00	1.00	1.00
rs3213093	0.28	0.29	0.47	0.49	0.20	0.20	0.66	1.00	1.00	1.00
rs3135363	0.81	1.00	0.54	0.53	0.28	0.43	1.00	1.00	1.00	1.00
rs7453920	0.15	0.23	1.00	1.00	0.05	0.09	1.00	1.00	1.00	1.00
rs2395309	0.13	0.10	0.57	0.58	0.73	0.74	1.00	0.65	1.00	1.00
rs2071349	1.00	0.71	0.48	0.46	1.00	0.60	1.00	1.00	1.00	1.00
rs7770370	0.87	0.74	0.75	1.00	0.05	0.19	0.66	1.00	0.33	0.20
rs931	0.09	0.07	0.16	0.16	0.75	1.00	1.00	0.62	1.00	1.00
rs9277534	0.15	0.16	0.26	0.27	1.00	1.00	1.00	0.62	1.00	1.00
rs9277535	0.62	0.64	0.42	0.42	1.00	1.00	1.00	0.57	1.00	1.00
rs3130186	0.62	0.64	0.42	0.42	1.00	1.00	1.00	0.57	1.00	1.00
rs2064479	0.70	0.46	0.57	0.58	1.00	1.00	1.00	1.00	1.00	1.00
rs2069727	0.57	0.26	0.23	0.21	0.63	1.00	0.62	0.62	1.00	1.00
rs1805015	0.82	0.51	1.00	1.00	0.68	1.00	1.00	1.00	1.00	1.00
rs2230433	0.17	0.32	0.56	0.57	0.42	0.60	0.33	0.62	1.00	1.00
rs8099917	0.28	0.28	1.00	1.00	0.73	0.49	1.00	1.00	1.00	1.00

Unaff – Responders

All – Responders and low-responders

APPENDIX F
SNP Genotype Frequencies

SNP	Genotype	Frequency (%)			
		Total cohort	SA black	SA Caucasian	SA Asian
rs1800896	CC	16.1	8.2	24.0	4.5
	CT	47.7	46.9	50.7	36.4
	TT	36.2	44.9	25.3	59.1
rs1143634	AA	2.0	0.0	4.0	0.0
	AG	32.2	24.5	36.0	31.8
	GG	65.8	75.5	60.0	68.2
rs1143633	TT	10.1	2.0	16.0	9.1
	TC	32.9	26.5	38.7	27.3
	CC	57.0	71.4	45.3	63.6
rs1143627	AA	28.2	8.2	13.3	27.3
	AG	40.3	38.8	45.3	27.3
	GG	31.5	53.1	41.3	45.5
rs2069763	AA	6.0	2.0	6.7	13.6
	AC	33.6	4.1	46.7	59.1
	CC	60.4	93.9	46.7	27.3
rs2243248	GG	6.0	14.3	2.7	0.0
	GT	27.5	51.0	17.3	9.1
	TT	66.4	34.7	80.0	90.9
rs2243250	TT	17.4	8.2	2.7	9.1
	TC	36.9	49.0	28.0	36.4
	CC	45.6	42.9	69.3	54.5
rs2070874	TT	14.8	24.5	2.7	9.1
	TC	34.2	40.8	28.0	36.4
	CC	51.0	34.7	69.3	54.5
rs3212227	GG	8.1	8.2	8.0	9.1
	GT	34.9	34.7	30.7	50.0
	TT	57.0	57.1	61.3	40.9
rs3213093	TT	8.7	10.2	8.0	9.1
	TC	34.9	34.7	30.7	50.0
	CC	56.4	55.1	61.3	40.9
rs3135363	CC	4.7	2.0	8.0	0.0
	CT	34.2	20.4	48.0	18.2
	TT	61.1	77.6	44.0	81.8

	AA	6.0	2.0	9.3	4.5
rs7453920	AG	45.6	30.6	57.3	40.9
	GG	48.3	67.3	33.3	54.5
	GG	15.4	26.5	5.3	18.2
rs2395309	GA	38.9	46.9	33.3	40.9
	AA	45.6	26.5	61.3	40.9
	GG	2.0	2.0	2.7	0.0
rs2071349	GC	21.5	18.4	21.3	31.8
	CC	76.5	79.6	76.0	68.2
	GG	21.5	8.2	6.7	9.1
rs7770370	GA	48.3	44.9	52.0	50.0
	AA	30.2	46.9	41.3	40.9
	AA	16.1	28.6	6.7	13.6
rs931	AG	38.9	40.8	37.3	36.4
	GG	45.0	30.6	56.0	50.0
	GG	16.8	24.5	6.7	13.6
rs9277534	GA	40.9	42.9	40.0	36.4
	AA	42.3	32.7	53.3	50.0
	GG	4.0	2.0	4.0	9.1
rs9277535	GA	36.2	40.8	34.7	31.8
	AA	59.7	57.1	61.3	59.1
	TT	4.0	2.0	4.0	9.1
rs3130186	TC	36.2	40.8	34.7	31.8
	CC	59.7	57.1	61.3	59.1
	TT	12.1	22.4	5.3	9.1
rs2064479	TC	41.6	46.9	37.3	40.9
	CC	46.3	30.6	57.3	50.0
	CC	12.1	8.2	14.7	13.6
rs2069727	CT	39.6	28.6	48.0	36.4
	TT	48.3	63.3	37.3	50.0
	CC	6.7	14.3	2.7	0.0
rs1805015	CT	34.9	49.0	32.0	13.6
	TT	58.4	36.7	65.3	86.4
	CC	20.1	20.4	12.0	13.6
rs2230433	CG	45.0	57.1	40.0	36.4
	GG	34.9	22.4	48.0	50.0
	GG	2.7	0.0	5.3	0.0
rs8099917	GT	20.8	8.2	29.3	18.2
	TT	76.5	91.8	65.3	81.8

APPENDIX G

Chromosome 6 SNP and HLA-DP Haplotype Frequencies

Haplotype												Frequency (%)			
rs3135363	rs7453920	rs2395309	rs2071349	rs7770370	rs931	rs9277534	rs9277535	rs3130186	rs2064479	HLA-DPA1	HLA-DPB1	Total cohort	SA black	SA Caucasian	SA Indian
T	G	A	C	A	G	A	A	C	C	01:03	04:01	10.7	0.0	10.7	36.1
T	A	A	C	A	G	A	A	C	C	01:03	04:01	7.7	0.0	15.3	0.0
T	G	G	C	G	A	G	A	C	T	02:other	01:01	6.0	15.2	0.0	0.0
T	G	A	C	A	A	G	G	T	T	01:03	03:01	4.0	6.1	3.3	2.8
T	G	G	C	G	A	G	A	C	T	02:01	01:01	3.7	5.1	2.0	2.8
C	G	A	C	A	G	A	A	C	C	01:03	04:01	3.0	0.0	6.0	0.0
T	G	A	G	A	G	A	A	C	C	01:03	02:01	2.7	3.0	6.7	5.6
C	A	A	C	A	G	A	A	C	C	01:03	04:01	2.7	0.0	1.3	0.0
C	A	A	C	G	G	A	A	C	C	01:03	04:01	2.7	0.0	4.7	5.6
T	A	A	G	A	G	A	A	C	C	01:03	02:01	2.3	4.0	0.0	0.0
T	G	A	C	A	G	A	A	C	C	01:03	04:02	2.3	0.0	4.7	0.0
T	G	A	C	G	G	A	A	C	C	03:01	105:01	2.3	7.1	0.0	0.0
T	G	A	G	G	G	A	A	C	C	01:03	02:01	2.3	0.0	0.0	8.3
C	A	A	G	A	G	A	A	C	C	01:03	02:01	2.0	1.0	1.3	0.0
T	G	G	C	A	A	G	G	T	T	02:01	09:01	1.7	0.0	0.0	8.3
T	G	G	C	G	A	G	G	T	T	02:other	05:01	1.7	0.0	2.7	0.0
T	G	A	C	A	G	A	A	C	C	03:01	105:01	1.7	5.1	0.7	0.0
C	G	A	C	A	G	A	A	C	C	03:01	105:01	1.7	4.0	0.0	0.0
C	A	A	C	G	G	A	A	C	C	01:03	04:02	1.7	0.0	4.0	0.0
T	G	G	C	A	A	G	G	T	T	02:01	10:01	1.7	0.0	2.7	0.0
T	G	A	C	G	A	G	G	T	C	01:03	18:01	1.3	5.1	0.0	0.0
C	G	A	G	A	G	A	A	C	C	01:03	02:01	1.3	0.0	2.7	0.0
T	G	A	C	A	G	A	A	C	C	01:03	02:01	1.0	0.0	1.3	0.0
T	A	A	C	G	G	A	A	C	C	01:03	04:02	1.0	0.0	2.7	0.0
C	A	G	C	G	A	G	A	C	T	02:01	01:01	1.0	0.0	2.7	0.0
T	G	G	C	G	A	G	G	T	T	02:01	11:01	1.0	2.0	0.0	0.0
T	G	G	C	G	G	G	G	T	C	02:01	11:01	1.0	0.0	1.3	0.0
T	G	G	C	A	G	A	A	C	C	02:other	04:01	0.7	0.0	0.0	5.6
T	G	G	C	A	A	G	G	T	T	02:01	14:01	0.7	1.0	0.7	0.0
T	A	A	C	G	G	A	A	C	C	01:03	04:01	0.7	1.0	2.7	2.8
C	G	G	C	G	A	G	A	C	T	02:other	01:01	0.7	2.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	02:01	105:01	0.7	2.0	0.0	0.0
T	G	G	C	G	A	G	G	T	C	02:01	13:01	0.7	2.0	0.0	0.0
C	G	G	C	G	A	G	G	T	C	02:01	13:01	0.7	0.0	0.0	0.0
T	G	G	C	G	A	G	G	T	T	02:01	05:01	0.7	0.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	02:01	17:01	0.7	0.0	0.0	0.0
T	G	G	C	G	A	G	G	T	C	02:01	11:01	0.7	0.0	1.3	0.0
T	A	A	C	G	A	G	G	T	T	01:03	06:01	0.7	0.0	1.3	0.0
T	G	G	C	G	A	G	G	T	C	02:other	18:01	0.7	0.0	0.0	0.0
T	G	A	C	A	G	A	A	C	C	01:03	105:01	0.7	0.0	0.0	0.0
T	A	G	C	G	A	G	A	C	T	02:01	01:01	0.7	3.0	0.0	0.0
T	G	G	C	G	A	G	A	C	T	02:01	26:01	0.3	0.0	0.0	5.6

T	A	G	C	G	A	G	A	C	T	02:other	01:01	0.3	1.0	0.0	0.0
T	G	A	C	A	G	A	A	C	T	03:01	105:01	0.3	0.0	0.0	2.8
C	A	A	C	A	G	A	A	C	C	01:03	105:01	0.3	1.0	0.0	0.0
T	G	G	C	G	G	A	A	C	T	02:other	34:01	0.3	1.0	0.0	0.0
T	G	A	G	A	A	G	G	T	T	01:03	16:01	0.3	0.0	0.7	0.0
T	G	G	C	G	A	G	A	C	T	02:other	17:01	0.3	1.0	0.0	0.0
T	G	A	C	G	G	A	A	C	C	01:03	105:01	0.3	1.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	02:01	13:01	0.3	1.0	0.0	0.0
T	G	G	C	G	G	A	A	C	T	02:01	19:01	0.3	1.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	02:01	04:01	0.3	0.0	0.0	0.0
T	G	A	G	G	G	A	A	C	C	01:03	04:02	0.3	0.0	0.0	0.0
C	G	A	C	G	A	G	A	C	T	01:03	26:01	0.3	0.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	02:other	105:01	0.3	0.0	0.0	0.0
T	A	A	C	G	A	G	G	T	T	01:04	15:01	0.3	0.0	0.0	0.0
T	G	G	C	A	G	A	A	C	T	02:01	02:01	0.3	0.0	0.0	0.0
T	G	A	G	A	A	G	G	T	T	01:03	02:01	0.3	0.0	0.7	0.0
C	A	A	C	A	A	G	G	T	T	01:03	03:01	0.3	0.0	0.7	0.0
C	A	G	C	G	A	G	A	C	T	02:01	26:01	0.3	0.0	0.0	0.0
T	G	G	C	G	A	G	G	T	C	02:other	13:01	0.3	0.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	04:01	17:01	0.3	0.0	0.0	0.0
T	A	G	C	G	A	G	G	T	T	02:other	05:01	0.3	0.0	0.0	0.0
T	G	G	C	A	A	G	G	T	T	02:01	03:01	0.3	0.0	1.3	0.0
C	G	G	C	G	A	G	G	T	T	02:other	05:01	0.3	0.0	0.7	0.0
T	A	G	C	G	A	G	G	T	T	02:other	13:01	0.3	0.0	0.0	0.0
T	G	A	C	A	G	G	A	C	C	01:03	04:01	0.3	0.0	0.0	0.0
T	G	G	C	A	G	A	A	C	C	02:01	02:01	0.3	0.0	0.7	0.0
T	G	G	C	G	G	A	A	C	T	02:other	106:01	0.3	0.0	0.0	0.0
C	A	A	G	G	G	A	A	C	C	01:03	02:01	0.3	1.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	02:other	106:01	0.3	1.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	02:other	04:01	0.3	0.0	0.0	0.0
T	A	A	C	A	G	A	A	C	C	01:03	04:02	0.3	0.0	0.0	0.0
T	A	A	C	A	G	A	A	C	C	01:03	03:01	0.3	0.0	0.0	0.0
T	G	G	C	G	A	G	G	T	T	02:other	02:02	0.3	0.0	0.0	0.0
C	A	A	C	A	G	A	A	C	C	01:03	03:01	0.3	0.0	0.0	0.0
T	G	G	C	G	A	G	G	T	T	02:01	17:01	0.3	0.0	0.0	0.0
T	A	A	G	G	G	A	A	C	C	01:03	02:01	0.3	1.0	0.0	0.0
C	G	A	C	A	A	G	G	T	T	01:03	03:01	0.3	0.0	0.7	0.0
C	A	G	C	G	A	G	A	C	T	02:01	131	0.3	0.0	0.0	0.0
T	G	A	C	G	A	G	A	C	C	01:03	13:01	0.3	0.0	0.0	0.0
T	A	A	C	G	A	G	G	T	T	01:03	131	0.3	0.0	0.0	0.0
T	G	G	C	A	G	A	A	C	C	02:other	105:01	0.3	3.0	0.0	0.0
T	G	A	C	G	G	A	A	C	C	01:03	04:02	0.3	0.0	0.0	0.0
C	A	A	C	A	G	A	A	C	C	01:03	04:02	0.3	0.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	02:01	02:01	0.3	0.0	0.0	0.0
C	G	A	C	A	G	A	A	C	C	01:03	04:02	0.3	0.0	0.0	0.0
T	G	A	C	A	A	G	G	T	T	01:03	20:01	0.3	0.0	0.7	0.0
T	G	A	G	G	G	A	A	C	T	03:01	58:01	0.3	1.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	04:01	105:01	0.3	1.0	0.0	0.0
C	A	A	C	G	G	A	A	C	C	03:01	105:01	0.3	0.0	0.0	0.0
T	A	G	C	G	A	G	A	C	T	02:other	?	0.3	1.0	0.0	0.0
T	G	G	C	G	G	A	A	C	C	02:01	?	0.3	1.0	0.0	0.0
C	G	A	C	G	G	A	A	C	C	03:01	105:01	0.3	0.0	0.0	0.0

T	G	G	C	G	G	G	G	T	C	02:01	13:01	0.3	1.0	0.0	0.0
T	G	A	C	G	A	G	G	T	T	01:04	15:01	0.0	1.0	0.0	0.0
T	A	G	C	G	A	G	A	C	T	02:01	131	0.0	1.0	0.0	0.0
C	G	A	C	G	A	G	G	T	C	01:03	18:01	0.0	1.0	0.0	0.0
T	G	A	C	G	A	G	G	T	C	01:03	13:01	0.0	1.0	0.0	0.0
T	A	A	C	G	A	G	A	C	T	01:03	131	0.0	1.0	0.0	0.0
C	A	A	C	A	G	A	A	C	C	03:01	105:01	0.0	1.0	0.0	0.0
C	G	G	C	G	G	G	G	T	C	02:01	11:01	0.0	1.0	0.0	0.0
T	G	A	C	G	G	A	A	C	C	01:03	04:01	0.0	2.0	0.0	0.0
T	G	A	C	A	G	G	G	T	C	01:03	04:01	0.0	1.0	0.0	0.0
T	G	G	C	G	A	G	A	C	T	02:01	11:01	0.0	2.0	0.0	0.0
T	G	G	C	G	G	A	A	C	T	02:01	106:01	0.0	1.0	0.0	0.0
T	A	G	C	A	A	G	G	T	T	02:01	09:01	0.0	0.0	0.0	2.8
C	A	G	C	G	A	G	G	T	C	02:01	13:01	0.0	0.0	0.0	2.8
C	A	G	C	G	A	G	G	T	C	02:other	13:01	0.0	0.0	0.0	2.8
T	A	G	C	A	A	G	G	T	T	02:other	13:01	0.0	0.0	0.0	2.8
T	G	G	C	A	G	A	A	C	C	02:01	17:01	0.0	0.0	0.0	2.8
T	A	A	C	A	G	A	A	C	C	03:01	105:01	0.0	0.0	0.7	0.0
T	G	G	C	A	A	G	G	T	T	02:other	10:01	0.0	0.0	0.7	0.0
C	G	G	C	G	G	A	A	C	C	02:01	04:01	0.0	0.0	0.7	0.0
C	G	A	C	G	G	A	A	C	C	01:03	04:02	0.0	0.0	2.0	0.0
T	G	A	G	G	A	G	A	C	T	01:03	26:01	0.0	0.0	0.7	0.0
C	G	A	C	A	G	A	A	C	C	01:03	02:01	0.0	0.0	0.7	0.0
C	G	G	C	A	G	A	A	C	T	02:01	02:01	0.0	0.0	0.7	0.0
T	G	G	C	G	A	G	G	T	C	04:01	13:01	0.0	0.0	0.7	0.0
C	G	G	C	G	G	A	A	C	C	02:01	17:01	0.0	0.0	1.3	0.0
T	G	G	C	A	A	G	G	T	T	02:01	05:01	0.0	0.0	1.3	0.0
C	G	G	C	A	G	A	A	C	C	02:01	02:01	0.0	0.0	0.7	0.0
C	A	A	C	G	G	A	A	C	C	01:03	17:01	0.0	0.0	0.7	0.0
C	G	A	G	G	G	A	A	C	C	01:03	02:01	0.0	0.0	0.7	0.0

APPENDIX H
Complete Results Set (SNP Data)

Univariate single SNP analysis

SNP	Minor Allele	ALLELIC MODEL			GENOTYPIC MODEL		DOMINANT MODEL		RECESSIVE MODEL	
		P_{EMP1}	P_{EMP2}	OR	P_{EMP1}	P_{EMP2}	P_{EMP1}	P_{EMP2}	P_{EMP1}	P_{EMP2}
rs1800896	C	0.450	1.000	1.54	0.602	1.000	0.673	1.000	0.626	1.000
rs1143634	A	**0.038**	0.560	2.93	**0.048**	0.599	0.117	0.762	0.164	0.809
rs1143633	T	0.602	1.000	0.63	0.447	1.000	0.471	1.000	0.580	1.000
rs1143627	A	0.495	1.000	0.63	0.574	1.000	0.257	0.965	1.000	1.000
rs2069763	A	0.755	1.000	1.14	0.680	1.000	0.720	1.000	1.000	1.000
rs2243248	G	0.376	1.000	1.92	0.259	0.995	0.455	0.999	0.414	1.000
rs2243250	T	0.167	0.830	2.41	0.120	0.874	0.069	0.614	0.632	1.000
rs2070874	T	0.211	0.937	2.24	**0.044**	0.512	**0.028**	0.329	1.000	1.000
rs3212227	G	0.371	1.000	0.40	0.852	1.000	0.480	1.000	1.000	1.000
rs3213093	T	0.251	0.995	0.39	0.762	1.000	0.479	1.000	1.000	1.000
rs3135363	C	1.000	1.000	0.82	0.192	0.972	0.484	1.000	0.312	0.974
rs7453920	A	1.000	1.000	0.81	0.292	0.998	0.507	1.000	0.413	1.000
rs2395309	G	0.116	0.802	2.53	0.178	0.933	0.279	0.997	0.110	0.620
rs2071349	G	0.435	1.000	1.63	0.201	0.980	1.000	1.000	0.161	0.809
rs7770370	G	0.065	0.694	2.76	**0.010**	0.199	1.000	1.000	**0.012**	0.117
rs931	A	**0.048**	0.360	3.23	0.067	0.723	0.082	0.621	0.135	0.682
rs9277534	G	0.086	0.624	2.99	0.065	0.739	0.159	0.801	0.142	0.724
rs9277535	G	0.360	1.000	1.65	0.399	1.000	0.736	1.000	0.289	0.961
rs3130186	T	0.360	1.000	1.65	0.399	1.000	0.736	1.000	0.289	0.961
rs2064479	T	0.062	0.600	2.79	0.075	0.800	0.275	0.996	0.063	0.379
rs2069727	C	0.116	0.821	0.29	**0.029**	0.421	**0.019**	0.320	1.000	1.000
rs1805015	C	0.776	1.000	0.71	0.240	0.978	0.486	1.000	0.444	1.000
rs2230433	C	0.822	1.000	0.80	0.165	0.967	0.711	1.000	0.376	0.996
rs8099917	G	0.709	1.000	0.43	1.000	1.000	0.687	1.000	1.000	1.000

Haplotype analysis - chromosome 2

Haplotype			Frequency	Univariate		
				P_{EMP1}	P_{EMP2}	OR
rs1143634 (IL1B)	**rs1143633** (IL1B)	**rs1143627** (IL1B)				
G	T		0.265	0.559	1.000	0.66
G	T	A	0.234	0.364	0.998	0.49
G	T	G	0.026	0.224	0.999	2.33
A	C		0.181	**0.039**	0.457	3.48
A	C	A	0.131	0.526	1.000	1.52
A	C	G	0.046	**0.009**	0.071	10.50
G	C		0.554	0.421	0.999	0.65
G	C	A	0.118	0.455	1.000	0.46
G	C	G	0.440	0.574	1.000	0.74
	T	A	0.233	0.378	0.998	0.50
	C	A	0.250	0.985	1.000	1.00
	T	G	0.032	0.286	1.000	1.98
	C	G	0.485	0.513	1.000	1.33

Haplotype analysis - chromosome 5

Haplotype					Frequency	Univariate		
						P_{EMP1}	P_{EMP2}	OR
rs2243248 (IL4)	**rs2243250** (IL4)	**rs2070874** (IL4)	**rs3212227** (IL12B)	**rs3213093** (IL12B)				
G	C				0.112	0.471	1.000	1.49
G	C	C			0.096	0.281	1.000	1.82
G	C	C	G		0.035	0.853	1.000	0.00
G	C	C	G	T	0.036	0.975	1.000	0.00
G	C	C	T		0.062	0.058	0.881	2.90
G	C	C	T	C	0.062	0.057	0.876	2.92
G	C	T			0.018	0.553	1.000	0.00
G	C	T	T		0.019	0.741	1.000	0.00
G	C	T	T	C	0.019	0.614	1.000	0.00
G	T				0.086	0.097	0.981	2.52
G	T	C			0.035	0.503	1.000	1.86
G	T	C	T		0.032	0.508	1.000	1.36
G	T	C	T	C	0.033	0.494	1.000	1.45
G	T	T			0.049	0.130	0.998	2.40
G	T	T	G		0.014	0.263	1.000	0.08
G	T	T	G	T	0.010	0.226	1.000	0.02
G	T	T	T		0.034	0.070	0.907	3.55
G	T	T	T	C	0.034	0.071	0.891	3.60
T	C				0.529	0.062	0.742	0.37
T	C	C			0.527	0.064	0.742	0.38
T	C	C	G		0.139	0.221	0.993	0.16
T	C	C	G	T	0.138	0.219	0.993	0.15
T	C	C	T		0.386	0.234	0.991	0.50
T	C	C	T	C	0.386	0.240	0.992	0.50
T	T				0.273	0.253	0.995	1.69
T	T	C			0.024	0.527	1.000	0.26
T	T	C	G		0.011	0.227	1.000	0.75
T	T	C	G	T	0.012	0.425	1.000	0.71
T	T	C	T		0.014	0.433	1.000	0.16
T	T	T			0.252	0.171	0.958	1.96
T	T	T	G		0.055	0.581	1.000	1.35
T	T	T	G	T	0.058	0.610	1.000	1.36
T	T	T	T		0.197	0.181	0.958	2.14
T	T	T	T	C	0.196	0.185	0.958	2.13
	C	C			0.623	0.141	0.936	0.53
	C	C	G		0.173	0.109	0.962	0.15
	C	C	G	T	0.173	0.123	0.962	0.14
	C	C	T		0.450	0.730	1.000	0.82
	C	C	T	C	0.450	0.745	1.000	0.82
	C	T			0.018	0.440	1.000	0.00
	C	T	T		0.018	0.440	1.000	0.00
	C	T	T	C	0.018	0.440	1.000	0.00
	T	C			0.058	0.365	1.000	1.10
	T	C	G		0.011	0.185	1.000	3.60
	T	C	G	T	0.011	0.250	1.000	3.69
	T	C	T		0.047	0.590	1.000	0.72

T	C	T	C	0.044	0.537	1.000	0.74
T	T			0.301	0.107	0.845	2.18
T	T	G		0.070	0.647	1.000	0.85
T	T	G	T	0.070	0.646	1.000	0.84
T	T	T		0.231	**0.029**	0.564	2.79
T	T	T	C	0.227	**0.029**	0.527	2.84
	C	G		0.187	0.233	0.992	0.32
	C	G	T	0.187	0.237	0.992	0.31
	C	T		0.494	0.612	1.000	0.76
	C	T	C	0.495	0.616	1.000	0.76
	T	G		0.068	0.593	1.000	0.70
	T	G	T	0.069	0.625	1.000	0.71
	T	T		0.251	**0.050**	0.736	2.45
	T	T	C	0.244	**0.043**	0.692	2.49
		G	T	0.255	0.196	0.990	0.42
		T	C	0.738	0.182	0.982	2.42

Haplotype analysis - chromosome 6

Haplotype							Frequency	Univariate		
								P_{EMP1}	P_{EMP2}	OR
rs2395309 (HLA-DPA1)	**rs2071349** (HLA-DPA1)	**rs931** (HLA-DPB1)	**rs9277534** (HLA-DPB1)	**rs9277535** (HLA-DPB1)	**rs3130186** (HLA-DPB1)	**rs2064479** (HLA-DPB1)				
A	C						0.523	**0.042**	0.251	0.30
A	G						0.128	0.430	0.981	1.60
G	C						0.349	0.123	0.449	2.26
	C	A					0.349	0.092	0.454	2.22
	C	G					0.523	**0.032**	0.260	0.31
	G	G					0.121	0.724	1.000	1.04
	C		A				0.506	**0.031**	0.296	0.32
	C		G				0.367	0.083	0.531	2.11
	G		A				0.122	0.678	1.000	1.03
		A	G				0.356	**0.013**	0.257	2.79
		A	G	A			0.151	0.056	0.414	2.40
		A	G	A	C		0.151	0.056	0.414	2.40
		A	G	A	C	T	0.151	0.056	0.414	2.40
		A	G	G			0.205	0.214	0.893	1.85
		A	G	G	T		0.205	0.214	0.893	1.85
		A	G	G	T	C	0.048	0.127	0.608	3.58
		A	G	G	T	T	0.157	0.763	1.000	1.22
		G	A				0.628	**0.021**	0.296	0.38
		G	A	A			0.628	**0.021**	0.296	0.38
		G	A	A	C		0.628	**0.021**	0.296	0.38
		G	A	A	C	C	0.606	**0.012**	0.187	0.30
		G	A	A	C	T	0.021	0.213	0.857	3.87
		G	G				0.017	0.837	1.000	0.00
		G	G	G			0.017	0.837	1.000	0.00
		G	G	G	T		0.017	0.837	1.000	0.00
		G	G	G	T	C	0.017	0.837	1.000	0.00
			A	A			0.628	**0.021**	0.296	0.38
			A	A	C		0.628	**0.021**	0.296	0.38
			A	A	C	C	0.606	**0.011**	0.187	0.30
			A	A	C	T	0.022	0.164	0.857	3.86
			G	A			0.151	0.056	0.414	2.40
			G	A	C		0.151	0.056	0.414	2.40
			G	A	C	T	0.151	0.056	0.414	2.40
			G	G			0.221	0.384	0.957	1.70
			G	G	T		0.221	0.384	0.957	1.70
			G	G	T	C	0.065	0.205	0.919	2.43
			G	G	T	T	0.156	0.726	1.000	1.23
					A	C	0.779	0.359	0.957	0.59
					G	T	0.221	0.384	0.957	1.70

Haplotype analysis - chromosome 16

Haplotype		Frequency	Univariate		
			P_{EMP1}	P_{EMP2}	OR
rs1805015 (*IL4R*)	**rs2230433** (*ITGAL*)				
C	C	0.090	0.211	1.000	0.14
C	G	0.152	0.888	1.000	1.08
T	C	0.336	0.925	1.000	1.06
T	G	0.422	0.723	1.000	1.20