

https://dspace.org/

The role of alcohol dehydrogenase genes in the development of fetal alcohol syndrome in two South African Coloured communities

Naidoo, Dhamari

2008-02-21T11:08:48Z
http://hdl.handle.net/10539/4483

Downloaded from a DSpace repository. https://dspace.org/

The Role of Alcohol Dehydrogenase Genes in the Development of Fetal Alcohol Syndrome in Two South African Coloured Communities

Dhamari Naidoo
Division of Human Genetics
National Health Laboratory Service and,
School of Pathology, University of Witwatersrand
Johannesburg
South African

A dissertation submitted in the Faculty of Health Sciences, University of Witwatersrand, in fulfilment of the requirements for the degree of Masters of Science in Medicine.

Declaration

I, Dhamari Naidoo, hereby declare that this is my own unaided work, unless otherwise stated. The statistical analyses were either checked by a statistician or, in the case of the logistic regression analysis, haplotype inference and linkage disequilibrium calculations, performed by a statistician with detailed consultation. It is being submitted for the degree of Masters in Medical Science in Human Genetics at the University of Witwatersrand. It has not been submitted for any other degree at any other university.

…….. ……..

Dhamari Naidoo Date

Bsc(Hons)

Acknowledgements

Prof. Michele Ramsay for all the help, guidance and most importantly your patience throughout my project

Dr Lize van der Merwe for all the much appreciated help with the statistical analysis

Lillian Ouko, Zane Lombard, Shelley Macauley, Phillip Haycock and Desmond Schnugh, thank you for your friendship and for making my time spent on this project an enjoyable one

Thank you to all my colleagues from the Human Genetics Department for the technical assistance and friendship

University of Witwatersrand and National Research Foundation for the financial assistance provided throughout my project

University of Witwatersrand and the Paul and Stella Lowenstein Charitable and Educational Trust Fund for the opportunity of attending the European Society of Human Genetics conference (Prague 2005) to present my work

To David

Abstract

Fetal alcohol syndrome (FAS) is a common cause of mental retardation and is attributable to the teratogenic effects of alcohol exposure in utero in individuals with genetic susceptibility. The Coloured communities from the Western and Northern Cape regions have some of the highest recorded incidence rates (~70 affected children per 1000 live births) in the world.

The candidate genes selected for this study belong to the family of alcohol dehydrogenase genes that code for enzymes which metabolise alcohol. The ADH1B and ADH1C genes have previously been examined in the Western Cape Coloured community and the enzyme encoded by the allele ADH1B*2 was significantly associated with protection against the development of FAS. ADH4, a new candidate gene, was selected due to its role in both the alcohol and retinol metabolic pathways.

A case-control genetic association study was performed to examine the potential roles of the ADH1B, ADH1C and ADH4 genes in the etiology of FAS in two Coloured populations from the Northern and Western Cape. Single nucleotide polymorphisms found within the candidate genes were typed by PCR-based methods in samples from the FAS children, their mothers and controls.

Significant associations were observed in the Western Cape cohort but were not replicated in the Northern Cape. Allelic association tests revealed that ADH1B*2 may be a protective marker as it occurred more commonly in the controls than the mothers ($p= 0.038$). The alleles of the polymorphic variant, ADH4.8, have been shown to influence the promoter activity of ADH4 (the 'A' allele has been shown to increase the activity of the promoter when compared to the 'C' allele as the same position). The alleles of this polymorphic marker were significantly associated with the risk for FAS. The 'A' allele was shown to occur more commonly in the mothers and FAS-affected children ($p= 0.002$ and 0.035 respectively) when compared to the controls, suggesting a role in disease susceptibility while the 'C' allele was shown to occur more commonly in the controls. It was also observed that ADH1B and ADH4.8 when examined together in a haplotype demonstrated an association

with susceptibility to the disease. While the 2-C haplotype (ADH1B-ADH4.8) was shown to be associated with protection against the development of FAS, the 1-A haplotype was associated with increased susceptibility. The results suggest that mothers with the common ADH1B*1 allele and presumably a normal ADH1B function but an increased level of ADH4 (allele 'A') as a result of the promoter mutation, will, when the blood alcohol concentration is high, have an increased risk of having a child with FAS. Conversely when the mothers have a faster alcohol metabolising rate due to the allele ADH1B*2 and normal levels of ADH4 protein (allele 'C'), the circulating alcohol in the blood is removed efficiently resulting in maternal protection against developing the disease.

This study has also highlighted the genetic diversity within individuals of the South African Coloured population. Haplotype analysis and logistic regression revealed that the Western and Northern Cape Coloured communities are genetically different and as a result, the samples could not be pooled for analysis. Although the two groups of controls were genetically diverse, haplotype analysis revealed that the sample of mothers and FAS-affected children were not statistically different between the provinces thus possibly suggesting a similar genetic etiology for the disease. The results from this study suggest that the ADH genes do play a role in the pathogenesis of FAS.

Abbreviations

ADH	alcohol dehydrogenase
ALDH	aldehyde dehydrogenase
APS	ammonium persulphate
ARBD	alcohol-related birth defects
ARND	alcohol-related neurodevelopmental disorder
ARMS	amplified refractory mutation detection system
bp	base pair
CGF	continuous gene flow
CNS	central nervous system
dATP	deoxyadenine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanine triphosphate
dTTP	deoxythymine triphosphate
dNTPs	deoxynucleotide triphosphates
dH_2O	distilled water
ddH_2O	deionised distilled water
DNA	deoxyribonucleic acid
EDTA	ethylenediamine tetra-acetic acid
EM	expectation-maximum algorithm
FAE	fetal alcohol effects
FAS	fetal alcohol syndrome
FASD	fetal alcohol spectrum disorders
g	gram
GABA	gamma-aminobutyric acid
HWE	Hardy-Weinberg Equilibrium
IBD	Identical by descent
IOM	Institute of Medicine
kb	kilobases
kD	kiloDalton

K_m	Michealis Menton constant
l	litre
LD	linkage disequilibrium
M	molar
NAD	nicotinamide adenine dinucleotide
NaCl	sodium chloride
NEB	New England Biolabs
PCR	polymerase chain reaction
pmol	picomole
RFLP	restriction fragment length polymorphism
ROS	reactive oxygen species
SASA	simultaneous allele specific amplification
SNP	single nucleotide polymorphism
STR	short tandem repeat
TBE	trisborate-EDTA
TDT	Transmission Disequilibrium Test
TE	tris-EDTA
TEMED	N,N,N1,N1-tetramethylethylnediamine
WT	wild type
µl	microlitre
ml	millilitre

Table of contents

Declaration ..... **ii**

Acknowledgements ..... **iii**

Abstract ..... **v**

Abbreviations ..... **vii**

List of Figures ..... **xii**

List of Tables ..... **xiii**

1 Introduction **2**

1.1 Historical perspective on FAS 2

1.2 Pathogenesis of FAS 3

1.2.1 Criteria for diagnosis 3

1.3 Epidemiology 5

1.3.1 Sociobehavioural risk factors 5

1.3.2 Maternal Risk factors for women in South Africa 6

1.3.3 World wide prevalence of FAS 7

1.3.4 FAS in South Africa 8

1.4 Teratogenic effects of alcohol exposure 8

1.4.1 Stages of fetal development 8

1.4.2 Mechanisms of alcohol's effects on the fetus 9

1.4.2.1 Free radicals, oxidative distress and apoptosis 9

1.4.2.2 Cell adhesion and migration 11

1.4.2.3 Nutrition 11

1.4.2.4 Hormonal factors 11

1.4.2.5 Acetaldehyde 12

1.5 The role of genetics in FAS 12

1.5.1 Twin studies 12

1.5.2 Animal models 13

1.5.2.1 Rodent models 13

1.5.2.2 Nonmammalian models 14

1.5.2.3 Large animal models 14

1.6 Identification of genetic factors involved in complex diseases 15

1.6.1 Linkage analysis 15
1.6.2 Case-control association studies 15
1.7 Alcohol metabolism 17
1.7.1 Alcohol dehydrogenases 18
1.7.2 Tissue expression 18
1.7.3 ADH1B 20
1.7.4 ADH1C 21
1.7.5 ADH4 22
1.7.6 Aldehyde dehydrogenases (ALDH) 24
1.7.6.1 ALDH2 24
1.8 Retinol metabolism 24
1.8.1 Inhibition of retinoic acid synthesis by ethanol 25
1.8.1.1 ADH4 26
1.8.1.2 ALDH1A2 28
1.9 Study objectives 29
2 Subjects and Methods 31
2.1 Subjects 31
2.1.1 Blood storage 32
2.2 Methods 32
2.2.1 DNA extraction and storage 32
2.2.2 Polymerase chain reaction (PCR) 32
2.2.3 Selection of markers 33
2.2.3.1 ADH1B 33
2.2.3.2 ADH1C 36
2.2.3.3 ADH4 36
2.2.3.4 ALDH1A2 37
2.2.4 Primer design 37
2.2.5 Genotyping 38
2.2.5.1 ARMS 38
2.2.5.2 SASA 39
2.2.5.3 RFLP 40

2.2.5.4 Agarose gel electrophoresis 43
2.2.5.5 DNA sequencing 43
2.2.5.6 Statistical Analysis 45
3 Results 50
3.1 Genotype and allele frequencies 50
3.2 Hardy-Weinberg equilibrium (HWE) 56
3.3 Tests for associations with susceptibility for FAS 57
3.3.1 Allele and genotype associations 57
3.3.2 Haplotype analysis 58
3.3.2.1 WC vs NC 59
3.3.3 Logistic regression 64
3.4 Linkage disequilibrium (LD) 66
4 Discussion 70
4.1 The ADH1B and ADH1C genes and their role in the susceptibility to FAS 70
4.1.1 ADH1B 70
4.1.2 ADH1C 71
4.2 The role of ADH4 in the pathogenesis of FAS 72
4.3 The impact of the maternal and fetal genotypes on disease outcome 77
4.4 Genetic diversity of the Western and Northern Coloured populations 78
4.5 ADH and brain development 81
4.6 Limitations of the study 82
4.7 Future directions 84
4.8 Conclusion 85
5 References Cited 87
5.1 Electronic references 97
6 Appendix A: Solutions 98
7 Appendix B: Molecular weight markers 101
8 Appendix C: Ethics approval for the Western Cape study 102
9 Appendix D: Ethics approval for the Northern Cape study 103
10 Appendix E: Complete genotype dataset for the Western and Northern Cape 104

List of Figures

Figure 1.1: A schematic representation of the facial characteristics of a child with FAS (Larkby and Day, 1997) 4
Figure 1.2: Schematic representation (not to scale) of the ADH cluster spanning ~380kb and the relative position of the ADH4 gene. Adapted from Osier et al.,(2002). 19
Figure 1.3: Schematic representation of how ethanol inhibits retinol metabolism via ADH pathway (Duester, 1991) 27
Figure 2.1: An agarose gel image showing a restriction enzyme digest with *Msl*I at exon 3 of the ADH1B gene. 34
Figure 2.2: An agarose gel image showing a restriction enzyme digest with *Alw N*I at exon 9 of the ADH1B gene. 34
Figure 2.3: An illustration of the scoring of alleles in ADH1B and their combined haplotypes which are referred to as ADH1B alleles in all publications.. 35
Figure 2.4: An agarose gel image showing a restriction enzyme digest with *Ssp*I at the ADH1C locus.. 36
Figure 2.5: An agarose gel image demonstration a restriction enzyme digest with *Mbo*I for ADH4.4.. 37
Figure 2.6: Illustration of the ARMS PCR technique.. 38
Figure 2.7: An agarose gel image showing an ARMS PCR profile for ADH4.8.. 39
Figure 2.8: An illustration of the SASA method. 40
Figure 2.9: Electropherograms depicting the allelic variation at position 183bp at the ADH4.4 locus. 44
Figure 2.10: Electropherograms depicting the allelic variation at position 99bp for the ADH4.8 locus. 44
Figure 3.1: LD estimates for all the controls. 67
Figure 3.2: LD estimates for the controls from the Western Cape 67
Figure 3.3: LD estimates for controls from the Northern Cape 68

List of Tables

Table 1.1: Three broad categories used for the diagnosis of FAS 3

Table 1.2: New nomenclature for ADH genes and enzymes 19

Table 2.1: Description of restriction enzyme digestions to detect RFLPs 41

Table 2.2: Candidate genes with associated DNA polymorphisms and the PCR conditions used for genotyping 42

Table 2.3: The equations for the measures of LD for two alleles, A and B, at different loci 47

Table 3.1: Allele and genotype frequencies for ADH1B in the study groups from the Western and Northern Cape provinces. 52

Table 3.2: Allele and genotype frequencies for ADH1C in the study groups from the Western and Northern Cape provinces.. 53

Table 3.3: Allele and genotype frequencies for rs1126670 (ADH4.4) in the study groups from the Western and Northern Cape provinces.. 54

Table 3.4: Allele and genotype frequencies for ADH4.8 (rs1800759) in the study groups from the Western and Northern Cape. 55

Table 3.5: Hardy-Weinberg equilibrium analysis p-values shown for all the polymorphisms tested in the Western and Northern Cape for each group (FAS children, Mothers and Controls) 56

Table 3.6: Case-control associations with individual polymorphic loci associated with the ADH1B, ADH1C and ADH4 genes and the p-values for the allelic and genotypic association tests. 60

Table 3.7: Haplotype analysis showing the estimated haplotype frequencies (*f*) in the controls and cases (FAS children or mothers) and the likelihood of association (p) in the Western Cape 61

Table 3.8: Haplotype analysis showing the estimated haplotype frequencies (*f*) in the controls and cases (FAS children or mothers) and the likelihood of association (p) in the Northern Cape 62

Table 3.9: Haplotypes analysis between the provinces for the FAS-affected children, mothers and controls 63

Table 3.10: Logistic regression analysis on genotype probabilities in FAS-affected children vs Controls, mothers of children with FAS vs Contols and NC vs WC (Controls only) 65
Table 4.1: Published allele and genotype frequencies from various studies for three polymorphic variants in the ADH4 genes compared to the observed frequencies in the two Coloured samples. 76

Chapter 1

Introduction

1 Introduction

It is widely accepted that alcohol exposure is the leading cause of preventable mental retardation in developed countries (Abel and Sokol, 1986). Since the first reports of fetal alcohol syndrome (FAS) by Jones and Smith in 1973, three decades of research has lead to a better understanding of the disease etiology. However the incidence of FAS remains unchanged in many communities.

1.1 Historical perspective on FAS

The first reports of the effects of maternal drinking on the developing fetus were published in 1892 by Templeman (as cited in Burd et al., 2003). In Dundee, Scotland, Templeman studied the deaths of 258 infants born to alcoholic mothers. He noted that 46% of the deaths were a result of excessive drinking by the mothers during weekends. In 1899 Sullivan reported on the role of maternal alcoholism as the cause of infant deaths. He studied 100 incarcerated women and showed that the infant mortality rate was double in the alcoholic mothers. He concluded that 'maternal intoxication' was the main source of damage to the fetus (Burd et al., 2003).

In 1968 Lemoine and colleagues published a report in a French medical journal describing the physical and behavioural characteristics of children born to alcoholic women.
They had observed similar facial characteristics, growth deficiency and psychomotor disturbances in the 127 offspring. A diagnosis of maternal alcoholism was made in children based on their resemblances to each other (Streissguth et al., 1980).

Later, in 1973 and 1974, Jones and Smith published detailed descriptions of characteristics of children born to alcoholic women. In the first article published in 1973, they examined eight children who were born to alcoholic mothers and had alerted physicians to the altered facial and behavioural characteristics. They presented a detailed case report on each of three Native American, three black and two white children. They described the shared anomalies among the children including developmental delay, short palpebral fissures, epicanthic folds, small jaws, flattened midface, joint anomalies, altered palmar crease patterns and cardiac abnormalities. They concluded that the defective fetal development could have been attributed to maternal alcoholism. The term FAS was introduced in the second article written

by Jones and Smith which was published in *The Lancet* five months later. They described an additional three cases of Native American children and characterised more of the shared anomalies (as cited in Armstrong, 1998).

1.2 Pathogenesis of FAS

1.2.1 Criteria for diagnosis

The adverse effects of fetal exposure to alcohol result in a spectrum of disorders that has been termed fetal alcohol spectrum disorders (FASD). FAS is at the severe end of this spectrum and is characterised by a confirmed history of maternal drinking during pregnancy, characteristic facial features, growth deficiency and neurocognitive deficiencies (Hoyme et al., 2005).

As the physical characteristics can vary in severity between individuals, researchers have narrowed down the key features for diagnosis to three broad categories (Abel,1998; Weinberg, 1997) (Table 1.1).

Table 1-1: Three broad categories used for the diagnosis of FAS

Growth retardation	
	decreased birth weight for gestational age
	failure to thrive
	disproportionate ratio of weight to height
Characteristic facial features	
	short palpebral fissures
	flat upper lip
	flattened philtrum
	flat midface
Central nervous system abnormalities	
	small head size
	structural abnormalities e.g. small brain
	impairment of fine motor skills
	neurosensory hearing loss
	impaired hand-eye coordination

The diagnosis of FAS is usually made in infants and young children as the features become more subtle with age. Studies have shown that the facial characteristics become less distinctive as the child enters adolescence and as the weight approaches the mean for the general population. Other characteristics often associated with FAS, such as neurodevelopmental and behavioural problems remain unchanged with increasing age (Streissguth et al., 1991; Michaelis EK and Michaelis ML, 1994). Although the facial features of FAS change during puberty, the symptoms that persist include microcephaly, short palpebral fissures and thin upper lip with indistinct philtrum and epicanthal folds (Figure 1.1).

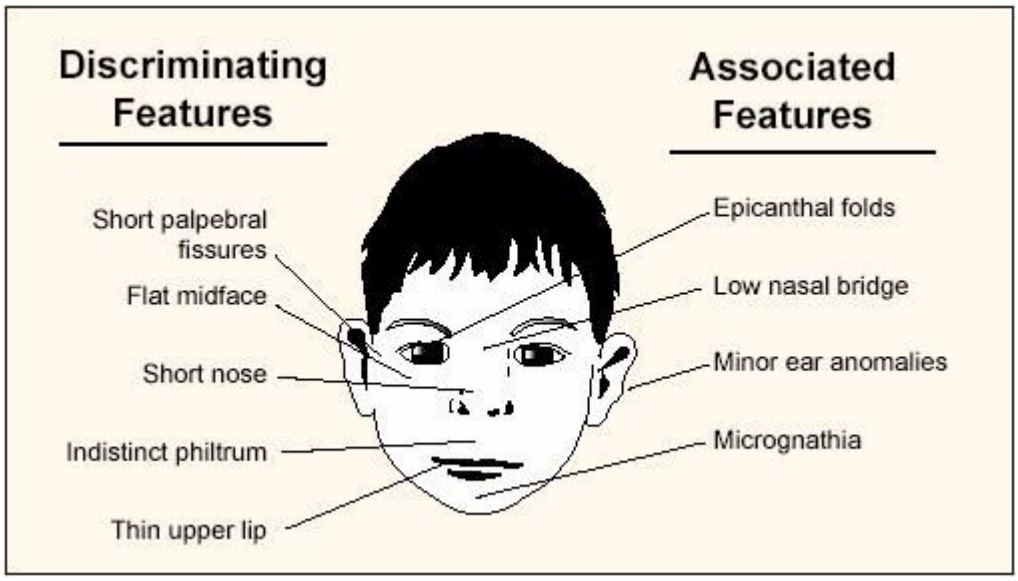


Figure 1.1: A schematic representation of the facial characteristics of a child with FAS (Larkby and Day, 1997)

Two sets of published guidelines are now being used that make the characterisation and evaluation of children who have been prenatally exposed to alcohol, much easier. The first set of criteria published by the Institute of Medicine (IOM) of the National Academy of Sciences, developed five diagnostic categories for FAS and alcohol related effects. These were: FAS with confirmed maternal alcohol exposure, FAS without confirmed maternal alcohol exposure, partial FAS with confirmed maternal alcohol exposure, alcohol related birth defects (ARBD) and alcohol-related neurodevelopmental disorder (ARND) (Hoyme et al., 2005).

Astley and Clarren published an objective set of guidelines in 2000 termed the Washington criteria (Astley and Clarren, 2000). These criteria retain the four key features of FAS: growth deficiency, facial phenotype, CNS damage and alcohol exposure in utero. Each factor is scored on a 4-point Linkert scale with 1 representing the absence of the feature and 4 the classical presentation. This system has been shown to be accurate in diagnosing a child but it is impractical for routine use in a clinical setting (Hoyme et al., 2005).

1.3 Epidemiology

The search for unique maternal characteristics, risk factors and protective factors has been a goal for researchers since the identification of FAS as a disease. Identifying these characteristics is important for facilitating the effective prevention of the disease and understanding its complex etiology.

Epidemiological studies have identified advanced maternal age, smoking, use of harmful drugs, low socioeconomic status and ethnicity as some of the traits associated with FAS (Sokol et al., 1980). A few of these characteristics will be discussed below.

1.3.1 Sociobehavioural risk factors

Alcohol intake pattern. The common pattern of alcohol intake among the mothers of FAS children is 'binge drinking'. This pattern of drinking results in brief exposures to high blood alcohol levels and is thought to cause more cellular damage than prolonged exposure to lower levels of alcohol (Streissguth et al., 1980).

Socioeconomic status. All epidemiological studies that have been reported in the United States have one common factor; mothers at greatest risk for having a child with FAS were from low socioeconomic backgrounds. These studies were conducted in inner city hospitals where the community was predominantly African American and characterised by poverty compared to cases from middle class Caucasian mothers (Streissguth et al., 1980). In the original case reports by Jones and Smith in which 11 children were characterised with FAS, they all had mothers who were living on welfare even though they were from different ethnic groups (Jones and Smith, 1973). It is believed that low socioeconomic status rather than biological factors relating to ethnicity is a major factor for FAS.

Smoking. Smoking has adverse effects on pregnancy outcome and can result in decreased birth weight. Smoking has been correlated with alcohol intake, as studies have shown that groups of heavy drinkers have the same number of heavy smokers. Smoking can increase the susceptibility to alcohol's toxic effect by mechanisms such as hypoxia (Streissguth et al., 1980).

1.3.2 Maternal Risk factors for women in South Africa

An epidemiological study examining the risk factors for women from the Western Cape showed that risk factors did not differ from those found for women from around the world (Viljoen et al., 2002). The mothers of the FAS children had a less formal education and were from low socioeconomic backgrounds. The mothers were from families with extensive drinking backgrounds and their partners were usually heavy drinkers. These women also smoked during their pregnancies but did not abuse other substances other than alcohol and tobacco (Viljoen et al., 2002).

Binge drinking over weekends is a common occurrence in some communities throughout South Africa and usually occurs at home or in shebeens which are equivalent to informal bars. In the Western Cape, alcohol had been consumed daily on many farms due to the 'Dop' system which has now been outlawed. This system was a form of partial payment to farm workers and has resulted in heavy and episodic drinking (Crome and Glass, 2000). An epidemiological study in the Western Cape revealed that drinking during pregnancy was common in parts of the Western Cape. Among pregnant mothers, 34.4% of those who drank during pregnancy were from metropolitan areas and 46.1% to 50.8% were from rural areas (Croxford and Viljoen, 1999;Viljoen et al., 2002).

Although the results from this study were specific to the Coloured community in the Western Cape, it may hold true for other low socioeconomic communities in other parts of South Africa.

1.3.3 World wide prevalence of FAS

Estimates of the prevalence of FAS vary between 0.2 and 1 per 1000 live births world wide. It must be noted that factors such as ethnicity, cultural backgrounds and varying diagnostic criteria could greatly influence the estimates for the prevalence of FAS around the world. Two reviews, the first by Abel (1995) and the other by Sampson et al., (1997) have been published that criticise population-based studies that published the prevalence of FAS.

In 1995 Abel reported the US incidence of FAS for the period between 1973-1992, to be 1.95/1000 births; 0.26/1000 when analysing middle/upper class Caucasians and 2.29/1000 for African/Native American groups. Another important observation from this report was that low socioeconomic circumstances rather than ethnic backgrounds played a greater role in the determination of prevalence rates (Abel, 1995).

In 1997 Sampson reported the incidence of FAS for different time periods at two US sites and one in France. The mothers were from different ethnic groups in the two samples; the US mothers were from minority groups and the mothers from France were Caucasian. The mothers from both groups suffered from alcoholism and came from low socioeconomic backgrounds. The researchers in France used different diagnostic criteria to the Americans. The diagnostic criteria used divided FAS into three categories; moderate FAS, full FAS and severe FAS. The incidences varied according the diagnostic criteria used; 1.3/1000 for severe FAS to 4.8/1000 for all FAS types (moderate, full or severe). When the US diagnostic criteria were applied to the French data, the incidence of FAS was 2.3/1000 live births. Estimates reported for the US were 2.8/1000 live births (Sampson et al., 1997). This report highlighted the influence different diagnostic criteria can have on the incidence rates observed in different communities. Both these reports provided evidence for a strong relationship between alcohol intake and low socioeconomic circumstances.

Indigenous communities have also been identified as having higher incidences of FAS compared to the general population for a country. May reported a study in 1991 looking at three Native American groups; Navajo, Peublo and SouthWest plains. The study was conducted over two time periods, 1969-1977 and 1978-1982. Results showed variation in the

prevalence of FAS between cultural groups (1.0, 1.3 and 17.5/1000 births respectively) and also indicated an increase in prevalence over time. The results were indicative of increasing drinking behaviour over time (May et al., 2000). The prevalence of FAS was also studied in the Alaskan indigenous populations from 1977 to 1992. Strict criteria were used to diagnose FAS and Native Alaskan children comprised 83% of the cases. The prevalence reported was 3.0/1000 among the Alaskan and 0.2/1000 for the non-native population (Egeland et al., 1998).

1.3.4 FAS in South Africa

The Coloured community of the Western Cape has the highest recorded prevalence of FAS in the world with the latest incidence levels at approximately 70 per 1000 children (May et al., 2007). An initial study conducted by May et al., (2000) assessed children of school going age from 11 predominantly Coloured and Black schools and one predominantly white school in the Western Cape. The prevalence of FAS was shown to be much higher than expected, with 61% of the cases occurring in rural areas. No cases of fetal alcohol syndrome were observed in the white school and the rate in the Black/Coloured schools was 49.3/1000. The minimum prevalence rates were calculated as 39.2-42.0 as the age-specific rate was 39.2 per 1000 for children of school going age. This paper provided insights into the social influences on fetal alcohol syndrome and has confirmed that low socioeconomic circumstance does increase the risk for the disease. Current studies are being conducted to ascertain whether this holds true for other low socioeconomic communities throughout South Africa (Viljoen, personal communication).

1.4 Teratogenic effects of alcohol exposure

As a teratogen, alcohol is able to induce malformations in a developing embryo. The extent of the damage is dependant on the time, dose and pattern of alcohol exposure.

1.4.1 Stages of fetal development

To understand the teratogenic effects of alcohol we need to understand the different stages of fetal development. There are two major periods of fetal development; the embryonic period which is up to eight weeks of gestation and the fetal period, from eight weeks till birth. The first stage of events of embryo development is cell division and proliferation followed by the

second stage of cellular growth and differentiation by which time cells become specialized in structure and function. The final stage is migration of maturing cells to their ultimate locations in the developing embryo where they will remain. It is during the embryonic period of development that malformations are formed due to the exposure of teratogens directly in the mother's bloodstream or through the mother's diet (Michaelis EK and Michaelis ML,1994). During the early stages of pregnancy alcohol exposure results in morphological malformations while growth is most affected during the later stages of development. Damage to the CNS as a result of alcohol exposure occurs throughout the gestational period (Larkby and Day, 1997).

All of the above stages are influenced and directed by nutritional, hormonal and cellular factors and alcohol can affect many of these factors, thereby influencing organ formation and growth (Michaelis EK and Michaelis ML,1994).

1.4.2 Mechanisms of alcohol's effects on the fetus

Numerous mechanisms have been postulated to explain how alcohol exerts its effects on fetal development but the molecular pathway(s) leading to FAS remain a mystery. Due to its small molecular size, alcohol is freely able to move through the placenta resulting in nearly equal concentrations in the blood of the mother and the fetus (Michaelis and Michaelis,1994). The resulting toxic effects on the fetus may be due to the direct exposure to alcohol or via secondary effects of the maternal/placental functions. Some of the mechanisms postulated include: a) inappropriate induction of apoptosis and the generation of free radicals, b) cell adhesion defects, c) growth factors and d) the inhibition of retinoic acid synthesis.

1.4.2.1 Free radicals, oxidative distress and apoptosis

Many actions of alcohol on the developing organism result in cell death via a process known as apoptosis. Apoptosis affects only individual cells and leaves the adjacent cells intact. A balance exists between certain protein factors which can activate or block apoptosis and disruption of this balance might be involved in alcohol-induced apoptosis (Olney, 2004).

One such factor is gamma-aminobutyric acid (GABA). This is a major inhibitory substance in the brain. Studies have shown that prenatal alcohol exposure increases the activity of

GABAnergic neurons resulting in a lesser degree of firing of the neurons compared with controls (Chaudhuri, 2000). An important neurotransmitter, glutamate, found in the brain and spinal cord is also affected by prenatal alcohol exposure. A decrease in GABAnergic activity or increased levels of glutamate result in hyper stimulation of neurons and this may result in mitochondrial damage. Growth factors also affected by alcohol include the epidermal growth factors and insulin-like growth factors.

Another key factor that can induce apoptosis is oxidative stress. Oxidative stress occurs as a result of having excess levels of free radicals in the cells. Free radicals are highly reactive molecules that are formed during biochemical reactions in the cell. Many of the free radicals also contain oxygen and are known as reactive oxygen species (ROS). The levels of free radicals and ROS are controlled by antioxidants which are scavenger molecules. The normal antioxidant levels in the cell can be reduced by the presence of alcohol and consequently, oxidative stress can occur. This stress causes damage to cellular membranes, DNA and protein (Cohen-Kerem and Koren, 2003).

Studies in cell lines and animal models have demonstrated a direct effect of oxidative stress by the formation of free radicals on apoptosis. The cellular damage caused by free radicals is a consequence of the peroxidation of lipids and alteration of enzymatic activity. Results from oxidative stress can be manifested as chromosomal abnormalities, enzymatic malfunction and disruption of cellular membranes. Formation of ROS can be induced in the mitochondria in hepatocytes exposed to ethanol. This is achieved by the reduction in mitochondria-derived components of electron transport. Oxidative stress is induced indirectly by reducing the intracellular antioxidant capacity for example, by reducing the levels of glutathione peroxidase (Cohen-Kerem and Koren, 2003).

Oxidative stress can also lead to fetal hypoxia. It has been shown that maternal intake of alcohol restricts the maternal blood flow resulting in reduced oxygen supply to the fetus. One of the proposed mechanisms for fetal hypoxia is the overproduction of prostaglandins. Their activity has been shown to be elevated during alcohol exposure and results in the constriction of the blood vessels in the uterus, placenta and fetus (Schenker et al., 1990).

Diminished blood supply to fetal tissues can result in apoptosis and may also induce the formation of free radicals in some embryonic tissues. An organ that is very sensitive to hypoxia is the brain and this can result in reduced brain size and weight which is one of the characteristics of FAS.

1.4.2.2 Cell adhesion and migration

Glial cells which are non-neuronal cells are needed for normal brain development. Various types of glial cells exist with specialized functions. One of their functions is to assist in the migration of newly formed neurons to their final locations. This is mediated by the presence of radial glia, which are able to direct the neurons to their correct destinations. Once at the final destination, they change into star-shaped astrocytes. However, with prenatal alcohol exposure, radial cells may become astrocytes prematurely, and as a result, neurons stop migrating and result in abnormal positions (Goodlett and Horn, 2001).

1.4.2.3 Nutrition

One of the characteristics of FAS is low birth weight where children generally remain small for their age. For normal growth and development to occur during gestation the constant transfer of amino acids and glucose from the mother to the fetus across the placental barrier is required. Several studies have indicated that alcohol directly inhibits the transport of amino acids and glucose therefore depriving fetal tissues of the energy sources and materials needed for cell proliferation, growth and differentiation (Michaelis EK and Michaelis ML,1994). In low socioeconomic status groups, maternal nutrition is thought to play an important role.

1.4.2.4 Hormonal factors

The release of hormones from maternal glands, fetal glands and the placenta influence the formation and development of tissues. Experimental studies on animals exposed to alcohol *in utero* show a reduction in corticosteroid hormones in the brain and blood. These hormones regulate aspects of metabolism and can influence the organism's response to stress. A deficiency in this hormone could affect the fetus's response to stress. Thyroid hormone deficiencies have been identified in babies born to alcoholic mothers and these have a deleterious effect on the development of some tissues such as the brain (Michaelis EK and Michaelis ML,1994).

1.4.2.5 Acetaldehyde

Ethanol metabolism, either through the alcohol dehydrogenase (ADH) or microsomal pathways, results in acetaldehyde. This metabolite has been shown to induce FAS-related malformations *in vivo* (Menegola et al., 2001). The human placenta is capable of transferring acetaldehyde from the mother to the fetus and of converting ethanol to acetaldehyde. Experimental evidence has shown that high acetaldehyde levels were reversible upon alcohol withdrawal but were increased by pregnancy and lactation (Lieber, 2000). There is also evidence for the carcinogenicity of acetaldehyde, when it is inhaled it can cause the formation of tumors of the respiratory tract in nasal mucosa of rats. It also has the ability to induce chromosomal aberrations, micronuclei and sister chromatid exchanges and it can also interact with DNA to form DNA adducts which is the initiating step of chemical carcinogenesis (Yokoyama and Omori, 2003).

1.5 The role of genetics in FAS

1.5.1 Twin studies

In 1974 Palmer et al., (1974) and colleagues (as cited in Streissguth and Dehaene, 1993), described monozygotic twins who were exposed to alcohol *in utero*. The twins were described as being identical at birth and all follow through examinations. They were both mildly retarded and their motor development was also restricted. Four pairs of dizygotic twins were also described in detail. In three of these, one twin was seen to be more severely affected than the other. Only one pair of twins were discordant in diagnosis, one twin had FAS and the other had FAE (fetal alcohol effects). A later report by Crain et al., (1983) described a dizygotic twin pair in which the twin with the smaller head circumference and most growth deficiency also had more severe mental retardation.

In 1993, Streissguth and Dehaene described 16 pairs of twins which included five monozygotic and 11 dizygotic twin pairs. All the monozygotic twins were concordant for diagnosis e.g., in two pairs of twins, both twins had FAS, in one pair both had FAE and the other pair there was no diagnosis. In the dizygotic twins seven of the pairs were concordant for diagnosis and the remaining four discordant. Among the twins, there were two pairs of

twins where one twin had FAS and the other had FAE, and another two pairs were one twin had no diagnosis and the other twin had FAE.

The above studies demonstrated that monozygotic twins show a higher rate of concordant alcohol related diagnosis when compared with dizygotic twins. These studies also examined IQ and there again the monozygotic twins were concordant more often with IQ when compared to the dizygotic twins. These studies confirm that genetic factors can influence the developmental effects of in utero alcohol exposure. Maternal response to alcohol and alcohol metabolism and the fetal genotype all play a vital role in the outcome of the disease.

1.5.2 Animal models

Animal models play a significant role in developing strategies to understand and prevent the damage caused by prenatal alcohol exposure. The damage caused results from multiple mechanisms that are dependant on dose, pattern and the timing of exposure. By using animal models, researchers have been able to replicate human physical characteristics in the rat and therefore provide further evidence that alcohol does have a direct toxic effect on the body.

They have also advanced the knowledge on areas such as effects of prenatal alcohol exposure on the immune system, hormonal systems, and the CNS.

1.5.2.1 Rodent models

The rat is an animal model that has been widely used by researchers. It has been used to study the effect of alcohol, particularly on the cerebellum. In one study, alcohol was administered postnatally during the period of rapid development of the cerebellum because the structures that develop *in utero* in humans only develop postnatally in the rat (Cudd, 2005). A review by Ponnappa and Rubin described studies that demonstrated that due to alcohol exposure, there was a reduction in the number of synaptic contacts and in the number of Purkinje cells which is thought to contribute to impaired motor control (Ponnappa and Rubin, 2000).

Mice have also been used as an animal model and have advantages such as short gestation periods and a greater potential for genetic manipulation compared to the rat. Mouse embryos

were exposed to alcohol at specific times during gestation and it was demonstrated that *in utero* alcohol exposure in mice can produce facial features that are characteristic to FAS in humans (Sulik et al., 1981). Studies on mouse embryos have revealed that cell death specifically occurs in neural crest cells (Smith, 1997).

The guinea pig, an animal model with a longer gestational period (68 days) allows researchers to target a particular developmental period. They also have the feature of prenatal brain growth, so researchers can study prenatal alcohol exposure during the third trimester of the guinea pig development. Studies have shown that prenatal alcohol exposure has resulted in low birth weight and reduced fetal brain weight (Kimura et al., 1999). Deficits in the cerebral cortex, hippocampus and cerebellum have also been observed.

1.5.2.2 Nonmammalian models

The chick model has been useful in studying the development of the face following prenatal alcohol exposure. The zebra fish, round worm and fruit fly (Drosophila) have also been used as models and they have the advantages of simple nervous systems and short generation times. These systems allow researchers to study questions that involve genetics and development. But a concern relating to these systems is the high dose of alcohol needed to induce malformations, suggesting that the mechanisms of damage might be different to those observed in lower dose alcohol concentration for larger animals (Cudd, 2005).

1.5.2.3 Large animal models

The use of large animal models to study prenatal alcohol exposure requires the insertion of catheters into the fetus. The intubation results in high fetal loss and is considered a great liability. Although nonhuman primates exhibit many of the same behaviour patterns as humans, these studies are complicated by the need for anesthetics and the need to restrain animals after instrumentation. Investigators have also used the pig as a model to study the behavioural effects of prenatal alcohol exposure. There are disadvantages such as the expense of maintaining the animals, the large litters and the fetuses that are difficult to intubate. Sheep have also been used as a model and studies include investigating the effect of alcohol on brain metabolism and the actions of alcohol on brain activity and fetal brain neurotransmitter activity (Cudd, 2005).

1.6 Identification of genetic factors involved in complex diseases

Complex diseases result from the interaction of multiple genes and environmental factors, with each gene contributing towards the overall phenotype. The identification of causative genes involved in complex diseases has posed great challenges for researchers. Two strategies are generally used by researchers to identify and characterise genes; linkage analysis and case-control association studies.

1.6.1 Linkage analysis

This method attempts to identify a region of the genome that is transmitted within families along with the disease phenotype (Silverman and Palmer, 2000). The families are studied for a higher-than expected number of shared alleles among affected individuals within a family (Carlson et al., 2004). One approach involves the use of affected sib-pairs. If the affected sib-pairs share markers that are identical by descent (IBD) the Mendelian expectations of sharing alleles IBD 0,1,2 with frequencies of 0.25, 0.5 and 0.25 will be expected. If there is linkage between the disease predisposing locus and a marker then deviations from Mendelian expectations will be observed. The transmission disequilibrium test is also used to identify markers that are transmitted from parents to the affected offspring (Daly and Day, 2001). A limitation to this approach is the need to collect DNA samples from the parents, the affected individuals and the siblings, and it may become more difficult to perform when studying late onset diseases.

The use of linkage analysis to study complex diseases has been less successful than its use in elucidating single genes in Mendelian inherited diseases. This is due to multiple gene interactions and the influence of the environment on disease outcome, heterogeneity due to ethnic differences, small sample sizes and inappropriate selection of controls (Risch, 2000). It has, however, been used with limited success in identifying regions of the genome that contain susceptibility genes for complex diseases.

1.6.2 Case-control association studies

The aim of this strategy is to test the association between a genetic polymorphism and a phentoype. Alleles are thought to play a role in a disease or act as a marker for a disease when the frequencies differ significantly between the cases and controls. It can also be

inferred that these alleles may be in linkage disequilibrium with another allele which may play a causal role in the disease, at a nearby locus (Schork, 1997). Significant differences in allele and genotype frequencies between cases and controls are regarded as support for the involvement of an allele in disease susceptibility (Silverman and Palmer, 2000), however further studies would need to confirm a causative effect. This strategy is thought to be more powerful than linkage analysis for the detection of the more common disease alleles that confer modest disease risk. Another advantage of an association analysis is the ability to more easily collect larger numbers of unrelated affected individuals than to collect large numbers of families (Carlson et al., 2004).

There are a few considerations to take into account when performing a case control study and these include: choice of candidate gene and polymorphism for study, patient and control recruitment methods, matching of controls and the number of subjects to be studied.

Choice of candidate gene and polymorphism. It is important to choose a candidate gene that has relevance to the pathogenesis of the disease. Consideration of the functional effects of the polymorphism is also important because studying functionally relevant polymorphisms instead of random neutral polymorphisms might lead to the detection of the causal disease genes. A gene is selected as candidates depending on a number of variables such as it's expression in the tissues that are affected by the disease, it's proposed functions, the biological pathway in which it occurs or it's homology to other genes. A disadvantage to this approach is the large number of genes that can be considered as candidates for any given disease (Schork, 1997).

Statistical aspects of study design. Consideration should be given to correcting for multiple testing. Although many methods have been proposed for correcting for multiple testing, the appropriate test to use in a given situation remains unclear. Many researchers have suggested using the Bonferroni method of correction but this method results in a decrease in power to detect any gene affects and may be too severe a correction for multiple testing when considering candidate genes that have not been randomly chosen (Daly and Day, 2001).

Recruitment of subjects and controls. The careful selection of cases and controls is very important as one must ensure that cases do not include a heterogenous collection of phenotypes which can reduce the power of the study (Daly and Day, 2001). Only cases diagnosed with FAS were included in this study and those classified with 'partial' FAS were excluded. Spurious associations may arise due to population stratification or population substructure. This arises as a result of recent population admixture or can result from different ethnicities between cases and controls (Daly and Day, 2001). Our study focused on two of the Coloured communities of South Africa, which have resulted from population admixture. We carefully selected our controls so that they were geographically and ethnically matched to the cases.

For this study we have chose candidate gene case-control association study design. We selected our candidate genes according to their proposed role in alcohol metabolism, since this disease results from the abuse of alcohol during pregnancy. The polymorphisms associated with each candidate gene were selected with respect to the published frequency data and functional relevance.

1.7 Alcohol metabolism

Molecular genetic research on alcoholism has drawn attention to the importance of alcohol and acetaldehyde metabolising enzymes. Functional polymorphisms exist for various genes involved in alcohol metabolism and they have the ability to alter the rate of metabolism. The molecular mechanisms that control the elimination and metabolism of alcohol are important in understanding the chemical basis of alcohol toxicity.

Ninety two to 95% of alcohol metabolism occurs in the hepatocytes which are located in the liver. Alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH) are two enzymes involved in the oxidation of alcohol (ethanol) to carbon dioxide and water. Alcohol is metabolised first to acetaldehyde by Class I ADH, with NAD^+ as the coenzyme in a reversible reaction. The acetaldehyde then becomes oxidised to acetate by aldehyde dehydrogenases, with NAD^+ also acting as the coenzyme in a non-reversible reaction. The end products of this pathway are carbon dioxide and water, after the oxidation of acetate in

the Kreb's cycle (Norberg et al., 2003). Due to the equilibrium constant favouring the direction of a reduction of acetaldehyde to ethanol, acetaldehyde is usually found at lower concentrations than ethanol or acetate (Warren and Li, 2005). Secondary ethanol metabolism occurs in the microsomal fraction of the liver and this pathway is known as microsomal ethanol oxidising system or MEOS. This is the major non-ADH pathway and its activity increases after chronic alcohol consumption. The cytochrome P450 family of enzymes, namely CYP2E1, is involved in this pathway (Agarwal, 2001).

1.7.1 Alcohol dehydrogenases

The alcohol dehydrogenase gene family encodes enzymes that metabolise a wide variety of substrates including ethanol and retinol. Human ADH is a dimeric protein consisting of two subunits with a molecular weight of 40kD each. Seven different genetic loci code for human ADH and the ADH enzymes are derived from the association of different types of subunits. The seven genes have been organised into five classes (Table 1.2) based on amino acid sequence alignments and similarities, catalytic properties and patterns of tissue-specific expression (Ramchandani et al., 2001). Originally the genes were named according to the sequence of discovery but a new nomenclature has been adopted and will be used throughout this dissertation. The ADH gene family is clustered on chromosome 4q22 spanning 367kb (Figure 1.2). Class I genes (ADH1A, ADH1B, ADH1C) are the low K_m forms for ethanol oxidation and are considered to play a major role in ethanol metabolism while Class II and Class III are the high K_m forms. Class III which is a glutathione dependant formaldehyde dehydrogenase does not favour ethanol as a substrate and the kinetic features of Class V have not yet been defined.

1.7.2 Tissue expression

The human ADH genes are differentially expressed in different tissues which is a very important feature for the physiological consequences of alcohol metabolism in specific cells and tissues. A study conducted by Estonius et al., (1996) examined tissue distributions for the different ADH enzymes using 23 adult and four fetal tissues. Class I transcripts showed a wide range of expression levels among different tissues. It occurred with highest levels in the adult liver and varying levels in the stomach, small intestine, colon, adrenal cortex, heart and lung. Class I mRNA was mainly expressed in the liver. It was also observed in the small

intestine, pancreas and stomach but at low levels. Class III showed a different distribution pattern with transcripts occurring in all tissues examined. Class II ADH activity was detected mainly in stomach mucosa but also in oesophagus, gingiva and tongue tissues. These results indicate that ADH is found in most tissues in the body but with highest levels in the liver.

Table 1-2: New nomenclature for ADH genes and enzymes (www.gene.ucl.ac.uk/nomenclature)

ADH class	new gene nomenclature	Enzyme subunit
I	ADH1A	α
I	ADH1B	β
I	ADH1C	γ
II	ADH4	π
III	ADH5	χ
IV	ADH7	σ
V	ADH6	not identified

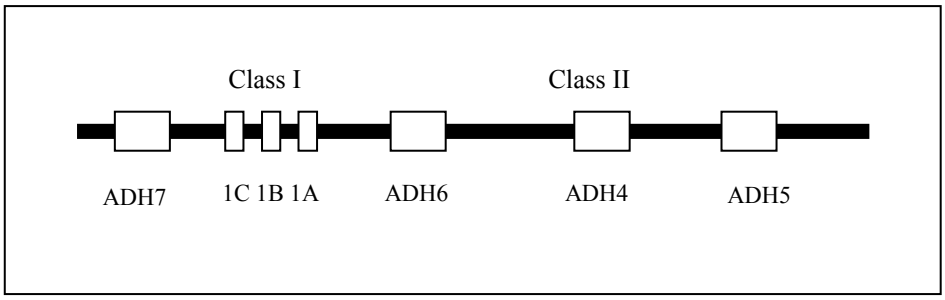


Figure 1.2: Schematic representation (not to scale) of the ADH cluster spanning ~380kb and the relative position of the ADH4 gene. Adapted from Osier et al.,(2002).

1.7.3 ADH1B

This enzyme is largely responsible for most of the liver ADH activity and remains active in the tissue throughout life. Three allelic variants have been identified; ADH1B*1, ADH1B*2 and ADH1B*3 which encode the subunits β1, β2 and β3 respectively (Chen et al., 1999). The isoenzyme encoded by ADH1B*2 has a higher turnover rate for ethanol metabolism than ADH1B*1. The product of thisallele differs from the ADH1B*1, the common allele, by a single amino acid substitution resulting an arginine to histidine exchange at position 47 (Chen et al., 1999). For individuals with the ADH1B*3 allele, the rate of ethanol oxidation is also greater than for individuals with the ADH1B*1 allele (Warren and Li, 2005). It differs from ADH1B*1 by an arginine to cysteine substitution at position 369 (Chen et al., 1999). It has been suggested that ADH1B*2 and ADH1B*3 protect against alcohol induced teratogenicity due to faster ethanol metabolism at higher ethanol concentrations but not all studies support this assertion (Warren and LI, 2005).

ADH1B*1 is the most common allele amongst Caucasians (frequency of 0.85 to 1) and African-Americans (frequency of 0.66 to 0.85), ADH1B*2 is the most frequent allele in the East Asian populations occurring at frequencies between 0.65 to 0.85. ADH1B*3 has been identified in African-American populations at frequencies between 0.15 to 0.33 (Burd et al., 2003).

In the literature, there are many reports regarding the ADH1B polymorphism and alcoholism. It has been suggested that ADH1B*2 might protect against alcoholism in Asian persons (Muramatsu et al., 1995) and Caucasian populations (Borras et al., 2000), though the mechanism remains unclear. It has been suggested that the unpleasant effects such as facial flushing following alcohol intake directly affect the amount of alcohol ingested. This was supported by studies where this allele was found to occur at a lower frequency in alcoholics. However there have been contradicting studies such as the one conducted by Russian researchers, where a negative association between ADH1B*2 and alcohol misuse was reported (Ogurtsov et al., 2001).

The effect of the ADH1B polymorphism on alcohol-exposed fetuses has been studied in the United States. McCarver et al., (1997) noted the presence of maternal ADH1B*3 was

associated with a decreased risk for reduced birth weight and birth length. In contrast, another study Stoler et al., (2002) reported on 404 women of whom 108 were African-American. They found that 60% of the African-American infants whom they classified as affected with FAS had an ADH1B*3 allele compared with only 29% in the unaffected infants. These results differed from those of McCarver et al., (1997) who indicated that the absence of ADH1B*3 was protective against adverse fetal outcome whereas other studies found that the presence of ADH1B*3 was protective.

The role of ADH1B in FAS etiology has also been examined in other cohorts. Eriksson et al., (2001) suggested that ADH1B*2 may be protective against developing the disease by preventing the mother from consuming alcohol due to the adverse effects experienced. The ADH1B*2 allele was also found to confer protection or to be a marker against the development of FAS in the South African Coloured community from the Western Cape. In this study the ADH1B locus was examined in 56 FAS children, their mothers and a group of controls. The ADH1B*2 allele occurred more commonly in the controls than in the FAS-affected children or their mothers however no significant associations were observed for ADH1B*3 (Viljoen et al., 2002).

1.7.4 ADH1C

Two allelic variants have been identified; ADH1C*1 and ADH1C*2 which encode the subunits γ1 and γ2 respectively. ADH1C*1 andADH1C*2 differ by an amino acid substitution of isoluecine by valine at position 349 in exon 8 (Osier et al., 2002). ADH1C*2 which is the lower activity allele, is the common allele in most populations with ADH1C*1 being the predominant allele amongst East Asians and Africans occurring in 90% of samples. The two alleles, ADH1C*1 and ADH1C*2, are equally distributed with in Caucasian populations (Chen et al., 1999). In 1999 Osier et al., demonstrated that ADH1B*2 and ADH1C*1, which are 21kb apart, were in strong linkage disequilibrium in a sample of Taiwanese Chinese alcoholic individuals. Through the use of haplotype analysis, the researchers found that the association with alcoholism was caused by ADH1B*2 allele and the decreased frequency of ADH1C*1 allele in the Taiwanese Chinese alcoholic individuals compared to the controls was due to the strong linkage disequilibrium with ADH1B*2 (Osier et al., 1999).

1.7.5 ADH4

ADH4 is the sole member of Class II. It was first characterised from stomach mucosa and was designated σ-ADH or μ-ADH. It is predominantly expressed in the epithelial tissue of the upper digestive tract i.e. esophagus, gingival and tongue (Baraona et al., 1991).

This enzyme has two important functions that makes it an interesting candidate gene for this study. Firstly it affects the amount of ingested alcohol that reaches the bloodstream as a result of first-pass metabolism. It is the first metabolic barrier as a result of its tissue distribution that ingested alcohol comes into contact with. Secondly it functions as a rate-limiting step in retinoic acid (RA) synthesis. Alcohol is able to competitively inhibit retinol from binding to ADH4 which functions as an efficient retinol dehydrogenase. This leads to the reduced levels of retinoic acid which is important during embryogenesis (Yin et al., 1999).

A study by Edenberg et al., (1999) characterised three polymorphisms in the promoter region of ADH4 that affected its expression; -192bp (T/A), -159bp (G/A) and -75 (A/C). They found that promoter activity increased if the 'A' allele was present instead of the 'C' allele at position -75bp. They hypothesised that the 'A' allele may have a protective effect by modulating alcohol metabolism thereby reducing the risk of developing alcoholism. A study conducted in Brazil examined the association of the promoter polymorphisms in alcoholic patients (Guindalini et al., 2005). Their results were in agreement with Edenberg et al., (1999) as the 'C' allele at -75bp was associated with alcohol dependence (odds ratio of 1.6).

Recently Luo et al., (2005) published their results on the association of polymorphic variants in the ADH4 gene with alcohol and drug dependence. They studied seven polymorphic variants in the ADH4 gene in European Americans and African Americans and four of the seven SNPs they examined were chosen for the present study; rs1800759, rs1126670, rs1126671, rs1042363. They found that within the European American controls, the genotype distribution frequencies were in HWE but in the alcohol and drug-dependant groups, they were not. Two SNPs, rs1042363 and rs1800759, showed the greatest degree of Hardy-Weinberg disequilibrium in both patient groups. Significant associations were found at the genotype level but not with alleles or haplotype frequency distributions. They concluded that

ADH4 genotypes were associated with alcohol and drug dependence with rs1042363 and rs1800759 being the markers closest to the risk loci.

In 2006 Edenberg and colleagues characterised 110 SNPs encompassed within and around the ADH gene cluster on chromosome 4q in an attempt to determine significant associations between SNPs in the region and alcohol dependence. The families studied were recruited from the COGA study and included individuals of European American and African American ethnicity who had been diagnosed with alcoholism. In their study they observed strong linkage disequilibrium within the ADH genes but low linkage disequilibrium between the ADH genes. Surprisingly they had observed a strong association between alcohol dependence and SNPs in the ADH4 gene, low levels of association with the ADH1A and ADH1B genes and no significant associations with the ADH1C gene.

The strongest region of association in the ADH4 gene was from intron 1 to 19.5kb downstream of exon 9 in which 12 SNPs were identified. Although previous studies have demonstrated the association between rs1800759 and alcohol dependence (Iida et al., 2002;Giundalani et al.,2005), the study by Edenberg et al., (2006) found no significant associations. However this gene is located 860bp away from the SNP with the highest significant association, rs4148886, and is more than likely in strong linkage disequilibrium with it.

No significant associations were found with the coding SNPs of ADH1B in the European American families as the alleles occurred at low frequencies. However, significant associations were observed with three adjacent SNPs in the ADH1B gene that occurred between intron 1 and the promoter region through to 1.5kb upstream of the initiation codon. A strong association was observed for ADH1B*3 and alcoholism in the African American sample which provided support for a role of ADH1B*3 in alcoholism susceptibility. The study by Edenberg et al., (2006) has provided strong evidence for a role of ADH4 in alcoholism and a weaker role for ADH1A and ADH1B in alcoholism susceptibility.

This study represents the first molecular genetic analysis of the ADH4 gene in the development of FAS.

1.7.6 Aldehyde dehydrogenases (ALDH)

Acetaldehyde is the metabolic product of ethanol metabolism in the liver. The major oxidation of acetaldehyde in the liver and other organs is carried out by aldehyde dehydrogenase. A number of isoenzymes of ALDH encoded by different gene loci have been detected in human organs and tissues which differ in their electrophoretic mobility, kinetic properties, and cellular and tissue distribution (Agarwal, 2001).

1.7.6.1 ALDH2

This gene encodes a mitochondrial enzyme that is expressed in various tissues but with the highest expression levels occurring in the liver. It contributes to acetaldehyde oxidation due to its low K_m values, this facilitates the rapid clearance of acetaldehyde. Studies have shown that individuals who are heterozygous or homozygous for the ALDH2*2 allele, exhibit the 'alcohol flushing syndrome' due to elevated levels of blood acetaldehyde. This is the result of a single base mutation in exon 12, resulting in a glutamate to lysine substitution which renders the catalytic properties of the enzyme inactive (Ramchandani et al., 2001). ALDH2*2 homozygosity alone, regardless of ADH1B and ADH1C protects against alcoholism due to the accumulation of acetaldehyde in the blood (Osier et al., 1999). This gene also exhibits ethnic differences between Orientals, Caucasians and Africans. The ALDH2*2 allele is frequently found among Asian individuals and there is evidence to suggest that it protects against alcoholism.

1.8 Retinol metabolism

Vitamin A is essential throughout life and its influence is most critical during pregnancy and childhood when cells proliferate and differentiate in response to different levels of vitamin A or RA. It is obtained from the diet and stored as retinyl esters, they are then hydrolysed to retinol and released into the bloodstream and transported bound to retinol-binding protein. Cells which require retinoic acid (RA) take up retinol and convert it to RA by the action of two types of enzymes. The first, retinol dehydrogenase, converts retinol to retinaldehyde and the second enzyme, retinal dehydrogenase, converts retinal to retinoic acid or vitamin A (Maden, 2000). Vitamin A exerts its function through the oxidised metabolites of retinol.

Late in the 1930s it was recognised that maternal deficiency of vitamin A during pregnancy resulted in fetal death and congenital malformations. Wilson and co workers later

documented the congenital abnormalities with the major tissues affected being the heart, ocular tissues, urogenital and respiratory systems (Wilson et al., 1953). Embryonic malformations also result from excess vitamin A. In 1954 Cohlan identified malformations in rats when they were exposed to excess vitamin A. Major target tissues included the heart, skull, skeleton, CNS, brain, eyes and cranofacial structures (as cited in Zile, 1998). The overlap of tissues being affected by vitamin A deficiency or excess has indicated a critical role for vitamin A in the development of many organs (Zile, 1998).

1.8.1 Inhibition of retinoic acid synthesis by ethanol

Although alcohol dehydrogenase has traditionally been associated with ethanol metabolism, it does however also function in retinol metabolism. Both ADH and ALDH are able to oxidise retinol to retinaldehyde to RA. RA regulates gene expression in several biological processes by binding to nuclear receptors. ADH acts as a competitive inhibitor of ADH-catalysed retinol oxidation because it can use either ethanol or retinol as a substrate (Duester, 2000) (Figure 1.3). During a state of intoxication, class I ADH becomes saturated with ethanol. The excess ethanol may start binding to other ADHs such as ADH4. This will prevent the retinol from binding to the ADH4 thus blocking retinol oxidation and resulting in reduced RA levels during embryogenesis. Studies have also revealed that ADH1C may be transcriptionally regulated by RA (Duester, 1991). It is therefore suggested that ethanol consumption may reduce RA levels in the fetus, resulting in birth defects (Keir, 1991).

Yelin et al., (2005) tested the hypothesis that the reduction of RA levels as a result of increased concentration of ethanol, affected several regulatory factors. They had exposed *Xenopus* embryos to ethanol and RA and examined the expression of known targets of RA signalling. These targets included the *Hox, chordin, gsc, Cyp26* and *Otx2* genes. The data revealed that while ethanol induced the up-regulation of *gsc*, RA had an opposite effect and down-regulated the expression of the gene. It was also shown that RA down-regulated the *Otx2* and *chordin* genes while ethanol up-regulated these genes. The opposite effects of ethanol and RA were also observed in the *Hox* genes. These observations were consistent with the model that ethanol competitively inhibits RA synthesis, resulting in the reduction of RA levels in tissues (Yelin et al., 2005).

1.8.1.1 *ADH4*

ADH4 is the best suited to the role of retinol dehydrogenase due to its high catalytic efficiency for retinol oxidation. Molecular modelling studies have shown that the ADH4 active site is quite large and can easily accommodate retinol, indicating that it is not only adapted to accommodate ethanol (Yin et al., 1999). Gene expression studies in the mouse have shown that ADH4 is less efficient at ethanol metabolism than it is at retinol metabolism. ADH4 mRNA was undetectable at developmental stage E6.5 but was detectable at low levels at E7.5 in the primitive streak mesoderm and at higher levels at E8.5. The simultaneous expression of ADH4, ALDH1 and RALDH2(ALDH1A2) at stage E7.5 when RA is first detected, provides a compelling argument that these enzymes are designed to function in RA synthesis. The ability of ethanol to reduce the level of RA in mouse embryos at stage E7.5 suggests that ethanol prevents the initiation of retinoid signalling needed for neural crest survival. This may be the result of ethanol-inhibition of RA synthesis catalysed by ADH4, therefore making cells that express ADH4 targets for the destructive effects of ethanol intoxication during early embryonic development (Duester, 2001). Along with ADH4, Class 1 ADH was also found to function *in vitro* as a retinol dehydrogenase. It is less effective than ADH4 because its catalytic efficiency is 10-fold lower and docking studies indicate that retinol has a better fit with ADH4 than with ADH1A, ADH1B and ADH1C (Duester, 2001).

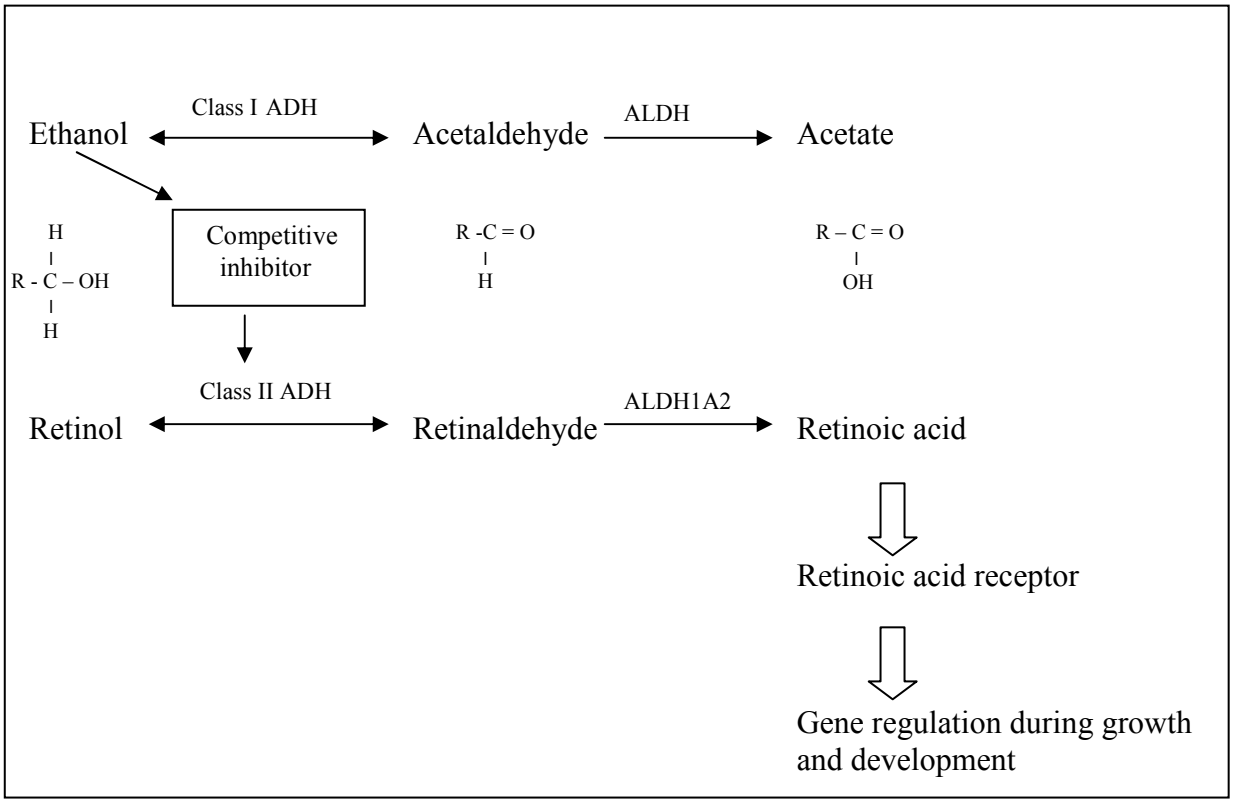


Figure 1.3: Schematic representation of how ethanol inhibits retinol metabolism via ADH pathway (Duester, 1991)

1.8.1.2 ALDH1A2

ALDHs are important in detoxification pathways because their isozymes are usually associated with altered drug metabolism and disease phenotype (Vasiliou et al., 2000). Members of the ALDH family catalyse the irreversible oxidation of retinaldehyde to RA. Two ALDH enzymes, cytosolic ALDH1A1 and mitochondrial ALDH1A2 have been studied in detail for their involvement in retinol oxidation.

ALDH1A2. This gene is also known as RALDH2, and is the second enzyme that forms retinoic acid which is bound to cellular retinol-binding protein (CRBP), as well as unbound retinal. cDNAs for ALDH1A2 have been isolated from mouse, rat, human and chicken. ALDH1A2 has been well conserved throughout evolution as shown by high amino acid sequence similarity. Studies from mouse and rat reveal that ALDH1A2 is strongly expressed in early embryonic trunk, forebrain and testis. It has the ability to oxidise retinaldehyde with high specificity. Both ALDH1A1 and ALDH1A2 contribute equally to retinoic acid synthesis but there is a difference in their affinity for the substrate retinol. RA affects mRNA levels for each differently in different tissues suggesting different physiological roles for each in generating RA. Knockout mice for ALDA1A2 have demonstrated the importance of this gene in RA synthesis as results showed the disruption of early embryonic development (Vasiliou et al., 2000).

Thus for both the ADH and ALDH families, the forms most efficient for retinol metabolism differ from those that are most efficient for ethanol metabolism. ALDH1A2 was also specifically chosen as a candidate gene for this study due to its role in retinol metabolism.

1.9 Study objectives

The aim of this study was to examine the role of the alcohol dehydrogenase genes in the development of FAS. Samples were collected from FAS-affected children, their mothers and healthy controls from two Coloured populations located in the Western and Northern Cape provinces.

Objectives:

To select polymorphisms within three ADH genes; ADH1B, ADH1C and ADH4 and within the ALDH gene; ALDH1A2.

To design and optimise various PCR-based methods to genotype the individuals.

To asses possible associations by comparing the observed allele, genotype and haplotype frequencies in the FAS-affected children, their mothers and controls for each region.

To perform logistic regression analysis to examine the relationships between the variables disease status and genetic variation.

To do linkage disequilibrium analysis to examine the extent of linkage disequilibrium between the markers studied.

To use the haplotype analysis, linkage disequilibrium analysis and logistic regression analysis to examine the genetic diversity between the two populations from the provinces sampled.

Chapter 2

Subjects and Methods

2 Subjects and Methods

2.1 Subjects

This study was performed as part of a larger study initiated by Prof Denis Viljoen. Through this study I have had access to samples collected from the Western Cape and Northern Cape. Following informed consent, blood samples were collected from subjects from the two regions; Western Cape (Wellington) and the Northern Cape (De Aar). Blood samples from FAS-affected children, their mothers and in some cases fathers and siblings were collected. A total of 64 FAS-affected children, 63 mothers and 190 controls were genotyped from Wellington. A total of 45 FAS-affected children, 36 mothers and 112 controls were genotyped from the Northern Cape. All samples collected were ethnically and geographically matched.

Samples were collected from the Coloured community as alcohol abuse has been identified as a serious health issue. The Coloured population is unique due to their genetic contributions from several populations including groups from Europe, Asia, India, Malay Archipelago, Madagascar, West Africa and also from indigenous South African groups (Botha, 1972). Individuals who participated as control subjects were asked about their ethnicity to make sure that the cases and controls were appropriately matched for ethnicity and geographic origins. Samples from individuals who were not from the communities of interest were excluded from the analysis.

Children were diagnosed with FAS or 'partial' FAS by trained clinicians from the Department of Human Genetics at the NHLS. The initial screening involved measurements for head circumference, body height and weight. If measurements were below the 10^{th} percentile for growth, they were examined by two independent physicians. When both physicians were in agreement with the diagnosis, a maternal interview was conducted. If a maternal history of alcohol abuse during pregnancy had been identified, a neurodevelopmental examination of the child was performed. Finally a case conference was held to determine the most accurate diagnosis (Viljoen et al., 2001). All samples were collected following a protocol approval by Ethics committees for research on human subjects

either from University of Cape Town or University of Witwatersrand (Appendix C and Appendix D)

2.1.1 Blood storage

A venous blood sample of 5-10ml was collected after informed consent was provided. The blood was collected in purple top vacutainer tubes which contain the anti coagulant agent called ethylenediamine tetra-acetic acid (EDTA). Samples were stored at -20°C after which the blood was transferred into 50ml NUNC tubes (Sterilin) until DNA extractions were performed.

2.2 Methods

2.2.1 DNA extraction and storage

Samples from the Western Cape had previously been extracted using the 'salting out method' (Miller et al., 1988) and were available for this study. All the DNA samples were resuspended in 1X TE buffer (Appendix A) and stored at 4°C. Samples from the Northern Cape were extracted using a commercially available kit, Flexigene (Qiagen). These extractions were performed by the FAS research group and I participated in these extractions. All DNA samples were resuspended in the buffer provided and stored at 4°C. All working stock solutions were diluted to 100ng/µl after being quantified using the Nanodrop® spectrophotometer.

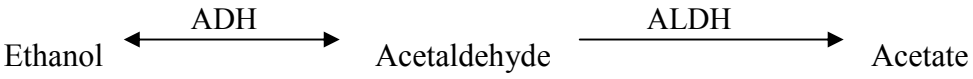
2.2.2 Polymerase chain reaction (PCR)

PCR allows for the exponential amplification of target specific DNA products. A set of primers are used that are complementary to the target sequence and flank the segment of DNA that is to be amplified. The template DNA is denatured so that the primers are allowed to anneal to the single-stranded target sequence. The annealed primers are then extended with a heat stable DNA polymerase such as DNA Taq polymerase. Repeated cycles of denaturation, annealing and extension result in the exponential growth of the target DNA sequence. Positive controls (DNA templates with known genotypes) and negative controls (all reagents but no DNA template) were used as control reactions for all PCR amplification experiments. All PCR primers used for this study were synthesised by Inqaba

Biotechnologies (Pretoria, South Africa). The general PCR protocol used was as follows; an initial denaturation step for 1 minute at 95°C, this was followed by 25 to 30 cycles of: denaturation at 95°C for 30 seconds, annealing at a primer set appropriate temperature that varied from 30 seconds to 1 minute and an elongation step at 72°C that varied between 30 seconds and 1 minute. The final step was another elongation period at 72°C for 5 minutes.

2.2.3 Selection of markers

The candidate genes selected for this study include ADH1B, ADH1C, ADH4 and ALDH1A2. The first three genes function in the first step of the metabolism of alcohol and ADLH1A2 functions in the non-reversible second step.



Polymorphisms associated with the ADH1B and ADH1C loci had previously been genotyped in the samples from the Western Cape and for the purpose of this study will be typed in the new cohort of samples collected from the Northern Cape. ADH4 and ALDH1A2 associated polymorphisms have not been studied in either the Western Cape or the Northern Cape and will be genotyped in both groups. The entire dataset will be analysed together.

2.2.3.1 ADH1B

The ADH1B alleles were previously typed in the Western Cape samples as reported in 2001 by (Viljoen et al., 2001) and in the MSc thesis of (Tshabalala, 2003). Results from the previous studies indicated that the ADH1B*2 allele was found to be significantly more commonly in the controls than the cases, suggesting that it may confer protection against the development of FAS or, be a marker for such protection.

ADH1B is a polymorphic locus characterised by three alleles, these are: ADH1B*1, ADH1B*2 and ADH1B*3 which are in fact three ‘haplotypes’. Since there is a large body of data referring to the haplotypes as alleles, this terminology will be used in this dissertation. The two polymorphisms that characterise the haplotype occur in exons 3 and 9. In exon 3 at position 47, a arginine/histidine substitution results in alleles ADH1B*1 (the common allele)

and ADH1B*2. In exon 9 at position 369, cysteine is substituted by arginine and results in the common allele ADH1B*1 and ADH1B*3. Restriction enzyme digestions were used to determine the alleles that were present at each exon. Positive (digested PCR product) and negative controls (undigested PCR product) were run on each gel as control reactions for the digestions. The alleles were combined into the final 'haplotype' which was used to ascribe a genotype at the ADH1B locus for each individual (Figure 2.3).

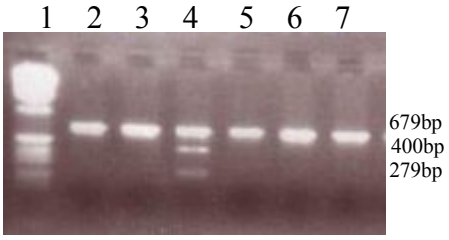


Figure 2.1: An agarose gel image showing a restriction enzyme digest with *Msl*I at exon 3 of the ADH1B gene. Lane 1 represents the molecular weight ladder, lane 2, 3, 5, 6, 7 are scored as homozygous 1/1 (ADH1B*1/ ADH1B*1) and lane 4 is scored as heterozygous 1/2 (ADH1B*1/ ADH1B*2)

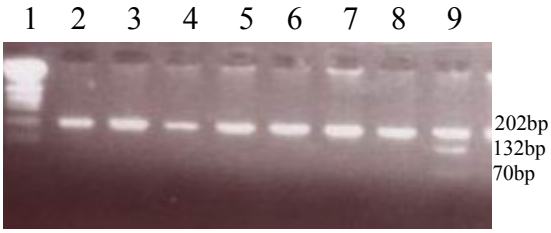


Figure 2.2: An agarose gel image showing a restriction enzyme digest with *Alw N*I at exon 9 of the ADH1B gene. Lane 1 is the molecular weight ladder, lanes 2 to 8 are scored as homozygous 1/1 (ADH1B*1/ ADH1B*1) and lane 9 is scored as a heterozygous 1/3 (ADH1B*1/ ADH1B*3).

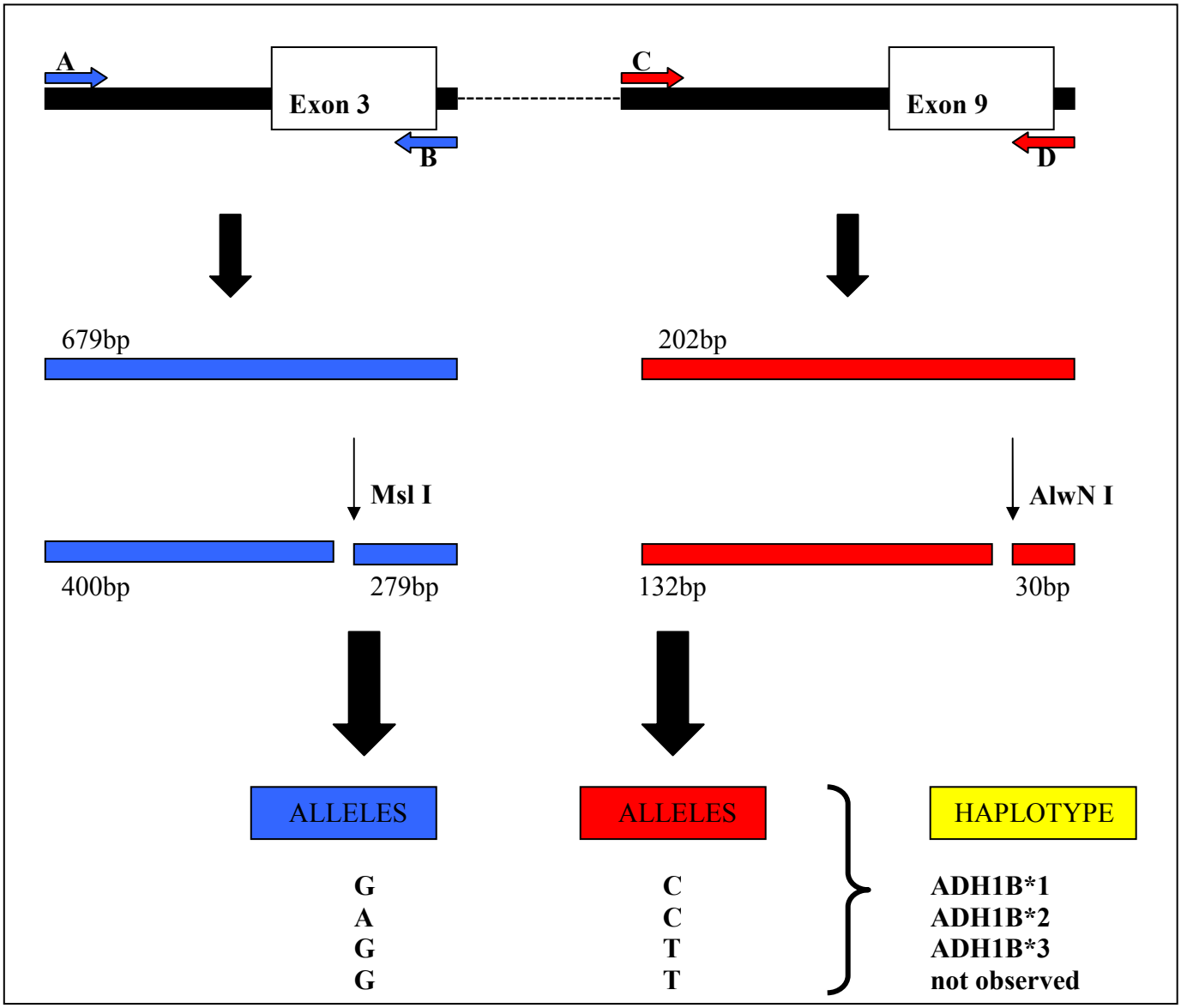


Figure 2.3: An illustration of the scoring of alleles in ADH1B and their combined haplotypes which are referred to as ADH1B alleles in all publications. The alleles present in exons 3 and 9 are scored using PCR with the primers denoted A, B, C, D respectively and the information is combined to determine the final genotype for ADH1B.

2.2.3.2 ADH1C

This polymorphic locus has two alleles, which result in the presence of either isoleucine or valine at position 349 in exon 8. The common allele is ADH1C*1 (isoleucine), and the rarer allele is ADH1C*2 (valine). They are distinguished using a restriction enzyme digestion with *Ssp*I as the restriction enzyme (NEB) (Figure 2.4).

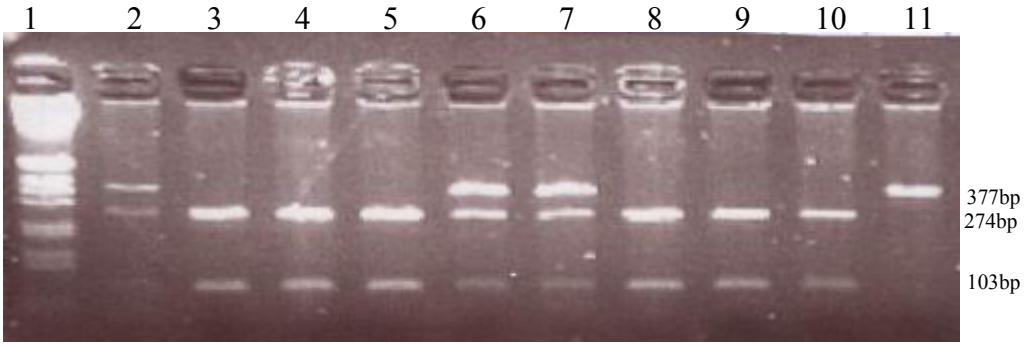


Figure 2.4: An agarose gel image showing a restriction enzyme digest with *Ssp*I at the ADH1C locus.. Lanes 2, 6, 7 are scored as heterozygous 1/2 (ADH1C*1/ ADH1C*2). Lanes 3, 4, 5, 8, 9, 10 are scored as homozygous 2/2 (ADH1C*2/ ADH1C*2) and lane 11 is scored as homozygous 1/1 (ADH1C*1/ ADH1C*1)

2.2.3.3 *ADH4*

This gene was selected as the main candidate in the study due to its role in both the ethanol and retinol metabolism pathways. The SNPs located in the ADH4 gene were selected from dbSNP according to frequency data and their potential functional impact or potential effect on the promoter activity (Edenberg et al., 1999; Iida et al., 2002). The SNPs selected from dbSNP include rs1042364 which results in an arginine/glysine substitution, rs1126671 which results in valine/isoluecine substitution and rs1126670, a synonymous amino acid change of proline/proline and two that affect promoter activity, rs1800759 and ADH4-2 (Edenberg et al., 1999; Iida et al., 2002). For the purpose of this study we will refer to rs1800759 as ADH4.8 (Figure 2.7) and rs1126670 as ADH4.4 (Figure 2.5).

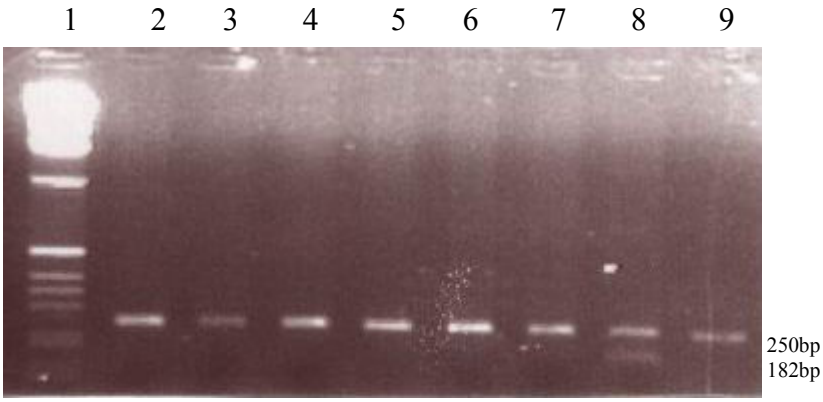


Figure 2.5: An agarose gel image demonstration a restriction enzyme digest with *Mbo*I for ADH4.4. Lane 1 is the molecular weight ladder, lanes 2 to 7 and 9 are individuals who are homozygous T/T. Lane 8 is heterozygous G/T.

2.2.3.4 ALDH1A2

This gene was chosen as a candidate because it functions in the second part of the ethanol metabolism pathway. It is involved in the conversion of acetaldehyde to acetate. Two polymorphic markers (namely rs1061278 and rs1837855) were selected according to frequency data and potential functional impact.

2.2.4 Primer design

Published primer sequences were used for two polymorphisms in the ADH4 gene, ADH4.8 and ADH4.2 (Iida et al., 2002; Edenberg et al., 1999). For the remainder of the polymorphisms, primers were designed using a program called Primer Select (DNA Star). The resulting primers were subjected to a BLAST search against the entire genome (http://genome.ucsc.edu) to make sure that they were 100% specific to the region of interest. Primer lengths varied from 25 to 33 bases in length (Table 2.2).

2.2.5 Genotyping

Different PCR-based methods were used to determine the SNP genotypes. All PCR products were resolved on agarose gels by gel electrophoresis (Table 2.1).

2.2.5.1 ARMS

When performing the amplification refractory mutation system (ARMS) a total of three primers are used for each reaction to detect a single base change. Two primers, one each corresponding to the 'normal' and 'mutant' alleles and a third common primer are used in a single reaction. Normal and mutant DNA can be distinguished by demonstrating differential annealing of either the 'normal' or 'mutant' primers to the normal or mutant DNA sequence respectively. These primers are designed in such a way that a single base change occurs at the 3' end of the primer. As a PCR control measure, a second set of primers that anneal elsewhere on the genome (arrows C in figure 2.6) are used as an internal control to ensure that the absence of a band is not due to PCR failure. The normal allele and mutant allele are amplified in separate PCR reactions (Figure 2.6 and Figure 2.7).

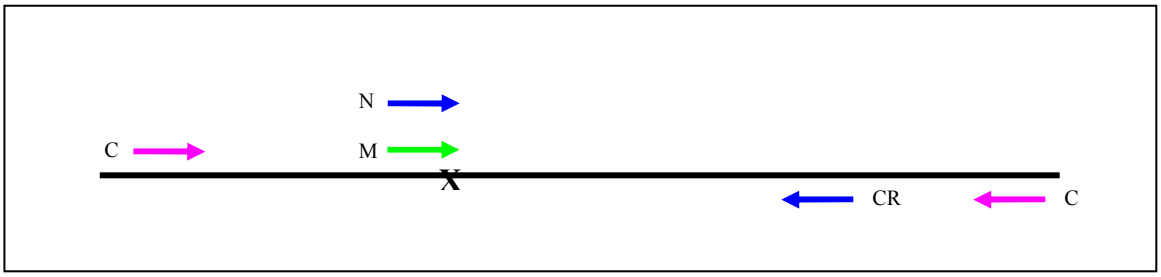


Figure 2.6: Illustration of the ARMS PCR technique. C represents the primers that will form the internal control band. N is the primer for the normal allele, M is the primer for the mutant allele and CR presents the common reverse primer. X marks the position of the polymorphism being detected (which is part of the primer). Amplification generates two PCR products, depending on the allele present. C/C = control band; N/CR product = band with normal allele; M/CR product= band with mutant allele.

When this system was initially designed, it resulted in non-specific binding of both the normal and mutant primers to the template DNA. The primers were subsequently redesigned by shortening the primer lengths and adding a second mismatch at the

penultimate base at the 3' end of the primer. This resulted in higher specificity for the PCR reaction and genotypes (in a subset of the samples confirmed by DNA sequencing).

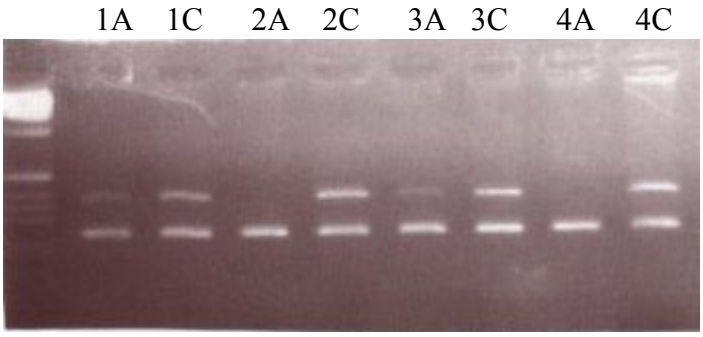


Figure 2.7: An agarose gel image showing an ARMS PCR profile for ADH4.8. Amplification is carried out separately for the normal and mutant alleles for each individual. The two PCR products are run on a gel and analysed together to form a genotype for one individual. Amplification of both PCR products indicates a heterozygous genotype while amplification of either the mutant or normal alleles indicates a homozygous genotype.

2.2.5.2 SASA

The Simultaneous Allele Specific Amplification (SASA) (DelRio-LaFreniere and McGlennen, 2001) is an adaptation of the ARMS method. This method uses four primers in one PCR reaction thereby allowing for quicker genotyping (Figure 2.8). One set of primers amplifies the locus specific product which serves as the internal control band. The third and fourth primers are the allele-specific primers, one for the 'normal' allele and one for the 'mutant' allele. A second mismatch was inserted at the penultimate base at the 3'end to ensure reduction of non-specific primer binding. The 'mutant' primer together with the reverse primer of the first set of primers will amplify the locus specific product. The wild type or 'normal' primer together with the forward primer of the first set of locus specific primers will form a PCR product (Figure 2.8). This system is designed in such a way that when the primers anneal, different size fragments are generated for each of the alleles. All PCR products were resolved on agarose gels using gel electrophoresis to determine the amplicon size and corresponding genotype. Although a great deal of optimisation was performed, results were not obtained for this system as the PCR conditions could not be optimised.

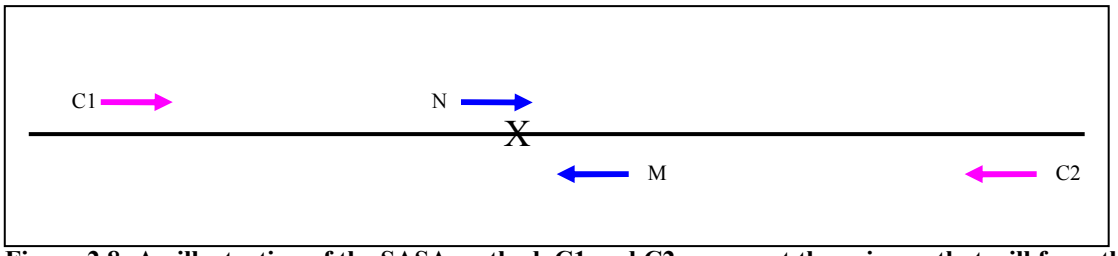


Figure 2.8: An illustration of the SASA method. C1 and C2 represent the primers that will form the locus specific product. N is the forward primer for the normal allele and M is the reverse primer for the mutant allele. X marks the position of the polymorphism being detected. Amplification is performed generating two different sized PCR products allowing for simultaneous detection of the normal and mutant alleles. C1/C2 product= locus specific product; N/C2 product= band with normal allele; M/C1 product=band with mutant allele.

2.2.5.3 RFLP

Restriction enzyme digestion was used when a single base pair change created or abolished a restriction enzyme recognition site. The PCR product was cut by the restriction enzyme resulting in different lengths of products depending on the presence or absence of the restriction site. The restriction enzyme recognition sites were identified using a program called MapDraw (DNA Star) or NEBcutter, v2 (www.neb.com). The digestion procedures were performed according to the manufacturer's recommendations and 5 units of enzyme were used per reaction for all digestions of PCR products. All digestions were carried out at the recommended temperatures for 3 hours to ensure complete digestion. All digested PCR products were resolved on agarose gels to determine the genotype (Table 2.1).

Table 2-1: Description of restriction enzyme digestions to detect RFLPs

Marker	Restriction enzyme (units)	Allele sizes	Reference
ADH1B*2	*Msl*I (5U)	Allele 1: 679bp Allele 2: 400bp + 279bp	Osier et al.2002
ADH1B*3	*AlwN*I (5U)	Allele 1: 202bp Allele 2: 132bp + 70bp	Xu et al.1988
ADH1C*1	*Ssp*I (5U)	Allele 1: 377bp Allele 2: 274bp + 103bp	Osier et al.2002
rs1126671	*HpyCH4*III (5U)	Allele A: 363bp Allele G: 241bp + 122bp	Present study
rs1126670	*Mbo*I (5U)	Allele T: 250bp Allele G: 182bp + 68bp	Present study
rs1061278	*Bfa*I (5U)	Allele A: 366bp Allele G: 166bp + 200bp	Present study
rs1837855	*Hpy188*I (5U)	Allele C: 440bp Allele T: 215bp + 225bp	Present study

Key: bp= base pairs

Table 2-2: Candidate genes with associated DNA polymorphisms and the PCR conditions used for genotyping

Gene	I.D	Polymorphism	Method of detection	Primer Sequences(5' to 3')	Tm
ADH1B	ADH1B*2	G→A	RFLP	F: attctaaattgtttaattcaag R: actaacacagaattacgg	55°C
	ADH1B*3	C→T	RFLP	F: tggactctcacaacaagcat R: ttgataacatctctgaagag	59°C
ADH1C	ADH1C*1	A→T	RFLP	F: ttgtttatctgtgatttttttgt R: cgttactgtagaatacaaag	51°C
ADH4	ADH4-2	A→C	ARMS	F: ggtgcttatccttattaataatcagatcaata R(N): tagtggttacatacstscaaattcttttcatgtaactt R(M): tagtggttacatacstscaaattcttttcatgtaactc	59.8°C
	ADH4.8	A→C	ARMS	F(N): cagcaacaaaggagaaaagaa F(M): cagcaacaaaggagaaaagtc R: attaatattaactgaacacccttct	54.5°C
	rs1042364	A→G	SASA	F: tccctccttggtacattctaa R: tatagactttgtacctgttata F(N): aagatgccaggagcaattca R(M): ggaatactatctgattgaatgtgaac	57.4°C
	rs1126671	A→G	RFLP	F: ggctcaatctacaatgat R: gaaattcttagagtgaag	50.5°C
	ADH4.4	T→G	RFLP	F: gtcacccctggttcgactt R: atccacctgcacagtcaag	59.8°C
ALDH1A2	rs1061278	A→G	RFLP	F: gcttgcagggagtaaagcaggc R: gggctgctgaatgcactgtcg	66.2°C
	rs1837855	T→C	RFLP	F: ctccatggttcatccagcataac R: tccaggaagccttaatactagg	58.6°C

Key: F = forward, R= reverse, F(N)= forward normal, F(M)= forward mutant, Tm= PCR annealing temperature

2.2.5.4 Agarose gel electrophoresis

Electrophoresis through agarose gels using 1X TBE buffer, is the standard method used to separate, identify and purify DNA products. Between 10–15 µl of PCR product or digested PCR product was mixed with Ficoll dye (Appendix A) and then loaded into the wells of the agarose gels. The appropriate percentage of agarose for the gel was chosen depending on the size of the fragments that needed to be separated, e.g. for fragments that are small in size, a high percentage gel is used as the pore sizes are smaller due to the higher concentration of agarose used and vice versa for a low percentage gel. A constant voltage was applied and a molecular weight ladder was included in every gel, either a 1Kb or 1Kb+ DNA ladder (NEB, Appendix B). Positive and negative controls were also run on every gel for quality control purposes. Once adequate separation occurred, a photograph of the gel was taken as a permanent record, using the Vacutec Gel Documentation system.

2.2.5.5 DNA sequencing

For each new method that was designed, sequencing was performed to confirm (i) the correct scoring of genotypes and (ii) that each system was optimised correctly for the specific detection of the alleles. Before sequencing was performed on the ABI Prism™ 377 DNA Sequencer, a number of steps were performed.

The selected PCR product was cleaned using the Nucleospin Extraction II kit. This column clean up procedure was performed to remove any unbound primers and dNTPs. Products were subsequently resolved on an agarose gel to ensure that sufficient PCR products remained after the column clean up procedure.

When it was confirmed that there was sufficient PCR product, cycle sequencing was performed to label the PCR products using the BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems). Individual dideoxy nucleotides are labelled with different terminating fluorophores for each nucleotide. Different sized fragments are produced and when run through the sequencer gel or capillary, it provides sequence data for the entire length of the PCR product.

After cycle sequencing, a final column clean was performed using the DyeEx™2.0 Spin (Qiagen) protocol. This removes any unbound components from the cycle sequencing kit.

Prior to loading into the wells of the vertical polyacrylamide gel, samples were denatured at 95°C for 2 minutes. The PCR products were resolved in 1X TBE buffer under constant voltage and data were analysed using the ABI Prism™ Sequencing Analysis software. Sequencing data were compared to a reference sequence to ensure the fidelity of the fragment.

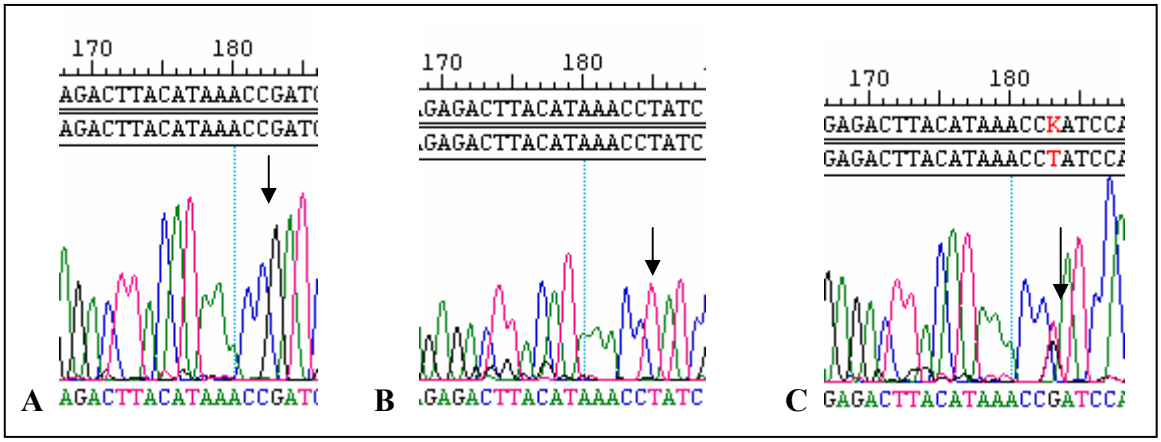


Figure 2.9: Electropherograms depicting the allelic variation at position 183bp at the ADH4.4 locus. A: G/G genotype, B: T/T genotype (position 185bp) and C: G/T genotype

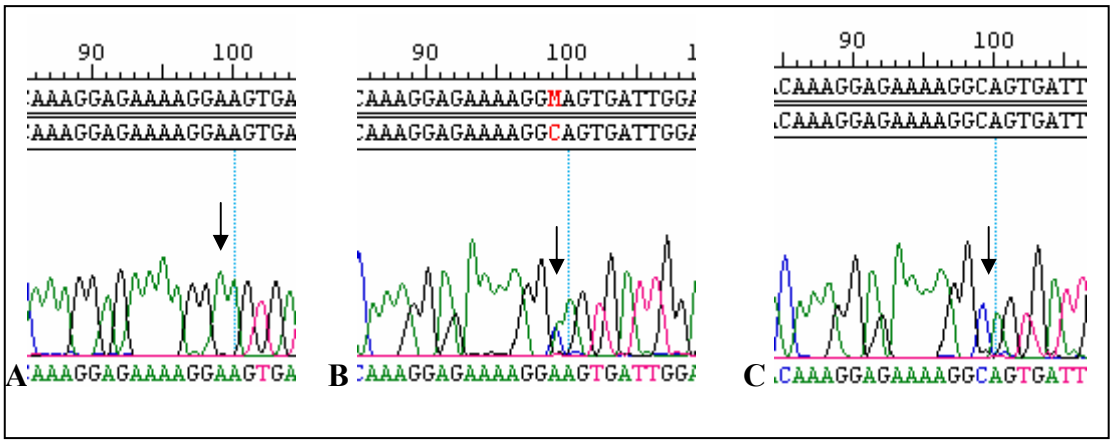


Figure 2.10: Electropherograms depicting the allelic variation at position 99bp for the ADH4.8 locus. A: A/A genotype, B: A/C genotype and C: C/C genotype

2.2.5.6 Statistical Analysis

The following analyses were performed by the author using a web-based statistical program (http://www.home.clara.net/sisa/); calculation of allele and genotype frequencies, HWE and associations using contingency tables. Statistically significant deviations from Hardy-Weinberg equilibrium were indicated by χ^2 p-values less than 0.05. Contingency tables (using the Fishers exact test) comparing the observed allele and genotype frequencies were used to detect possible associations between disease susceptibility and the marker of interest. All of the above were checked by a statistician.

Haplotype inference, association studies, LD calculations and regression analysis were done by the statistician, Dr Lize van der Merwe, following consultation. The results were discussed with the statistician to ensure that they were correctly interpreted. All tests were calculated using the program 'R' with the 'Genetics' package (R development core team, 2005; Warnes and Leisch, 2005). The haplotypes were calculated using the package 'Haplostats' (Sinwell and Schaid, 2005).

The groups were examined for compliance to Hardy-Weinberg expectation to determine whether the groups sampled for this study were in equilibrium, providing evidence for a representative sample of a large randomly mating population. When two alleles, 'A' and 'a' occur (with frequencies 'p' and 'q' respectively), $p + q = 1$. HWE predicts that the genotype frequencies of AA, Aa and aa are p^2, 2pq and q^2 .This test assesses the potential sampling errors in the groups used for this study as well as possible genotyping errors. The exact test for HWE, using genotype distribution data for each polymorphic marker that was tested within each province and separately for the FAS children, mothers and controls was used. This method is preferred to the standard χ^2 goodness-to-fit model as it controls for type I error (Wigginton et al., 2005).

Associations between the polymorphic variants and the disease phenotype were assessed by the use of four statistical tests; allelic and genotypic associations, inferred haplotype analysis and logistic regression. For the allelic and genotypic tests, the Fisher's exact test was used. The allelic and genotypic frequencies were compared between the FAS-

affected children and the controls and between the mothers and controls for each province and a p-value was calculated. All p-values less than 0.05 were considered to be statistically significant.

Four polymorphic markers were analysed together to calculate haplotypes with the markers in the following order as they were arranged in the genome: ADH1B-ADH1C-ADH4.4-ADH4.8. The expectation-maximum (EM) algorithm as described by Schaid et al., (2002) was used to estimate the haplotype frequencies. Global tests for associations and haplotype-specific tests were calculated. The haplotype-specific scores allow for the evaluation of individual haplotypes when the global score indicates a statistically significant finding.

Haplotype distribution and frequencies were examined in both provinces. Each haplotype was given a score according to the difference in frequencies between the cases (either FAS-affected children or mothers) and the controls. A global simulation p-value (Global sim p) represents the p-value for the group comparison and 'p' represents the p-value for each individual haplotype. For each province, the FAS-affected children were compared to the controls and the mothers were compared to the controls. In addition, an analysis was performed to examine the genetic makeup of the population. Firstly all the samples were grouped together (FAS-effected children, mothers and controls) and compared between provinces then the FAS-affected children were compared to each other between the provinces (FAS-affected children in WC vs FAS-affected children in NC) and the same for the mothers and controls.

Logistic regression analysis was used to determine the odds of being a case or a control according to the observed genotype. Two models were derived to compare the FAS-affected children to the controls and secondly to compare the mothers to the controls. A third model was derived to examine genetic make-up by combining the controls from both provinces. For each model all the terms or covariates such as 'province' and 'genotype' were added, and in a stepwise fashion terms were dropped. The Akaike criterion, as implemented in the R package, was used to remove the covariates until the

best fit for each model was found (the Akaike criterion is at a minimum) (Venables, 2002). A global p-value was calculated which represented how best the model fitted the data. All p-values less than 0.05 were considered as statistically significant. Individual p-values were also calculated for each covariate in the model. Results from these analyses were used to confirm any associations found in the allelic, genotypic and haplotype analyses.

Markers that are in close proximity to disease-causing variants on a chromosome are co-inherited more often than expected and this is demonstrated by linkage disequilibrium. It was therefore important to examine the extent of LD between the markers used in this case-control association study. In association studies the polymorphic variants that are tested may not necessarily be the causative factors for susceptibility and therefore association studies rely on linkage disequilibrium to help identify disease-causing markers (Zondervan and Cardon, 2004). LD is influenced by a number of factors such as recombination, time, natural selection, mutation, genetic drift, population substructure (e.g. admixture), population size and mating patterns. The commonly used measurement for LD is the disequilibrium coefficient *D;* $D = P_{11} - p_1q_1$ where P_{11} is the observed frequency of 1/1 haplotype, p_1 is the frequency of the '1' allele at locus 1 and q_1 is the population frequency of '1' at locus 2 (Haines and Pericak-Vancem, 1998). The disequilibrium coefficient can be difficult to interpret because its maximum value is dependant on allele frequencies. Two measures, correlation coefficient (r^2) and Lewontin's *D*' have been standardised to lie between -1 and 1 are often used instead of *D* (Haines and Pericak-Vance, 1998).

Table 2.3: The equations for the measures of LD for two alleles, A and B, at different loci

Correlation coefficient	$r^2 = D/\{(Pr(A)[1-Pr(A)]\ Pr(B)[1-Pr(B)]\}$
Lewontin's *D*	$D' = \begin{cases} D/\min\{Pr(A)[1-Pr(B), Pr(B)[1-Pr(A)]\} \text{ if } D > 0 \\ D/\min\{-Pr(A)\ Pr(B), -[1-Pr(A)][1-Pr(B)]\} \text{ if } D < 0 \end{cases}$

Key: Pr(A) = frequency of A allele; Pr(B) = frequency of B allele

If both r^2 and *D*' are zero then no allelic association exists, and they fall between -1 and 1, thus their maximum values are not dependant on allele frequencies.

Linkage disequilibrium was used to study the genetic composition of the two Coloured populations that were studied by examining the patterns of LD between the two control groups in the Western and Northern Cape

Post-test power based on allelic frequencies was to provide insight into the likelihood that the observed results would be replicated in another similar sized study.

When studying complex diseases, the chance of detecting false positives exists and increases with the number of polymorphisms examined. Therefore consideration should be given to correcting for multiple testing. However, using tests such as the Bonferroni test which is considered to be very conservative runs the risk of loosing any significant associations found. Chance associations may also be detected if the correction methods are not stringent enough or with the use of tests that assume independence. It must be noted that different polymorphisms within a gene or different genes may not be independent. A possible solution to these problems is to use the false discovery rate and sequential analysis (Romero et al., 2002). In this study we chose not to correct for multiple testing as the SNPs were non-randomly selected.

Chapter 3

Results

3 Results

To detect associations numerous statistical analyses were performed, these included allele and genotype association tests, calculating haplotypes and studying their distribution amongst the groups and between populations. Logistic regression analysis was performed to examine the relationship between variables such as disease status, genetic variation, geographic origin and linkage disequilibrium. The results are presented and analysed in the following sections.

3.1 Genotype and allele frequencies

Genotypes determined at polymorphic loci in the ADH1B, ADH1C and ADH4 genes were generated as described in the methods section. ADH1B, ADH1C and ADH4 genotyping was performed in the samples from FAS-affected children referred to as 'FAS children', in the mothers of the FAS-affected children, referred to as 'Mothers' and finally in random ethnically and geographically matched controls referred to as 'Controls'. The allele frequencies were determined by simple gene counting. A brief comment on the observations of each locus follows. Genotype and allele counts and frequencies are shown in Tables 3.1 -3.4.

ADH1B (Table 3.1): The common allele in all groups from both regions was ADH1B*1 occurring at frequencies of over 0.90, with the exception of the control group in the Western Cape where it occurred at a frequency of 0.85. Alleles ADH1B*2 and ADH1B*3 were relatively rare (11% in the WC controls) but showed some differences i.e. (geographic origins and study groups). In the Northern Cape, ADH1B*2 was not observed in the FAS-affected children and was only observed once in the mothers. In the Western Cape the highest frequency was observed in the controls. The genotypes ADH1B(2/2), (2/3) and (3/3) were not observed in any of the groups examined in the Northern Cape, however they were observed in the controls from the Western Cape, although there were very few individuals with those genotypes.

ADH1C (Table 3.2): The common allele at this locus was ADH1C*1, occurring at high frequencies in all groups, ranging from 0.77 in the Northern Cape controls to 0.86 in the

mothers in the Northern Cape. The most common genotype was ADH1C(1/1) with the least frequent being ADH1C(2/2). The latter was not observed in the mothers from the Northern Cape and was only observed once in the Western Cape.

ADH4 (Table 3.3): Genotyping was performed with two polymorphic variants associated within the ADH4 locus; ADH4.4 (rs1126670) and ADH4.8 (rs1800759). The third marker, rs1226710, proved to be uninformative with all the individuals homozygous for the common allele. The most common genotype at ADH4.4 was T/T, occurring at high frequencies in all three groups examined from both the provinces. The rarest genotype in all three groups was G/G. The two alleles of the ADH4.8 polymorphism occurred at frequencies of ~0.50 and as a result, the common genotype was A/C in all groups from both regions. In the Western Cape, the A/A genotype was more common in the FAS-affected children and their mothers than the controls while in the Northern Cape, the A/A genotype was more common in the controls than the FAS-affected children or mothers.

ALDH1A2: Both polymorphic markers tested proved to be uninformative and were not included in the analyses.

Table 3-1: Allele and genotype frequencies for ADH1B in the study groups from the Western and Northern Cape provinces. The frequency of genotypes and alleles are represented by *f* and the observed number of individuals is shown by *n*.

Province		FAS children		Mothers		Controls	
Western Cape	**Genotype**	***f***	***n***	***f***	***n***	***f***	***n***
	1/1	0.83	48	0.84	49	0.74	132
	1/2	0.09	5	0.07	4	0.18	32
	1/3	0.09	5	0.09	5	0.05	9
	2/2	0		0		0.01	2
	2/3	0		0		0.02	3
	3/3	0		0		0.01	1
	Total	**58**		**58**		**179**	
	Allele						
	1	0.91	106	0.92	107	0.85	305
	2	0.04	5	0.03	4	0.11	39
	3	0.04	5	0.04	5	0.04	14
	Total	**116**		**116**		**358**	
Northern Cape	**Genotype**						
	1/1	0.91	41	0.89	32	0.92	102
	1/2	0		0.03	1	0.05	5
	1/3	0.09	4	0.08	3	0.04	4
	2/2	0		0		0	
	2/3	0		0		0	
	3/3	0		0		0	
	Total	**45**		**36**		**111**	
	Allele						
	1	0.96	86	0.94	68	0.96	213
	2	0		0.01	1	0.02	5
	3	0.04	4	0.04	3	0.02	4
	Total	**90**		**72**		**222**	

Table 3-2: Allele and genotype frequencies for ADH1C in the study groups from the Western and Northern Cape provinces. The frequency of genotypes and alleles are represented by *f* and the observed number of individuals is shown by *n*.

Province		FAS children		Mothers		Controls	
Western Cape	**Genotype**	***f***	***n***	***f***	***n***	***f***	***n***
	1/1	0.67	39	0.60	34	0.57	102
	1/2	0.29	17	0.39	22	0.41	73
	2/2	0.03	2	0.01	1	0.02	3
	Total		**58**		**57**		**178**
	Allele						
	1	0.82	95	0.79	90	0.78	277
	2	0.18	21	0.21	24	0.22	79
	Total		**116**		**114**		**356**
Northern Cape	**Genotype**	***f***	***n***	***f***	***n***	***f***	***n***
	1/1	0.70	31	0.72	26	0.59	66
	1/2	0.25	11	0.28	10	0.37	41
	2/2	0.05	2	0	0	0.04	5
	Total		**44**		**36**		**112**
	Allele						
	1	0.83	73	0.86	62	0.77	173
	2	0.17	15	0.14	10	0.23	51
	Total		**88**		**72**		**224**

Table 3-3: Allele and genotype frequencies for rs1126670 (ADH4.4) in the study groups from the Western and Northern Cape provinces. The frequency of genotypes and alleles are represented by *f* and the observed number of individuals is shown by *n*.

Province		**FAS children**		**Mothers**		**Controls**	
Western Cape	**Genotype**	***f***	***n***	***f***	***n***	***f***	***n***
	G/G	0.02	1	0.04	2	0.01	1
	T/G	0.27	14	0.31	15	0.25	41
	T/T	0.71	37	0.65	32	0.75	124
	Total		**52**		**49**		**166**
	Allele						
	G	0.15	16	0.19	19	0.13	43
	T	0.85	88	0.81	79	0.87	289
	Total		**104**		**98**		**332**
Northern Cape	**Genotype**	***f***	***n***	***f***	***n***	***f***	***n***
	G/G	0.02	1	0.03	1	0.02	2
	T/G	0.19	8	0.21	7	0.26	28
	T/T	0.79	33	0.76	25	0.72	77
	Total		**42**		**33**		**107**
	Allele						
	G	0.12	10	0.14	9	0.15	32
	T	0.88	74	0.86	57	0.85	182
	Total		**84**		**66**		**214**

Table 3-4: Allele and genotype frequencies for ADH4.8 (rs1800759) in the study groups from the Western and Northern Cape. The frequency of genotypes and alleles are represented by *f* and the observed number of individuals is shown by *n*.

Province		FAS children		Mothers		Controls	
Western Cape	**Genotype**	***f***	***n***	***f***	***n***	***f***	***n***
	A/A	0.26	8	0.38	12	0.16	25
	A/C	0.58	18	0.47	15	0.48	74
	C/C	0.16	5	0.16	5	0.36	56
	Total		**31**		**32**		**155**
	Allele						
	A	0.55	34	0.61	39	0.40	124
	C	0.45	28	0.39	25	0.60	186
	Total		**62**		**64**		**310**
Northern Cape	**Genotype**	***f***	***n***	***f***	***n***	***f***	***n***
	A/A	0.16	7	0.14	5	0.26	29
	A/C	0.64	28	0.65	22	0.50	56
	C/C	0.20	9	0.21	7	0.24	27
	Total		**44**		**34**		**112**
	Allele						
	A	0.48	42	0.47	32	0.51	114
	C	0.52	46	0.53	36	0.49	110
	Total		**88**		**68**		**224**

3.2 Hardy-Weinberg equilibrium (HWE)

At loci ADH1B, ADH4.4 and ADH4.8, all genotype frequencies were found to be in HWE. The genotypes at ADH1C in the Western Cape controls (p= 0.0153) were found to deviate from HWE. No p-value could be generated using the 'R' program for the FAS affected children samples from the Northern Cape group since only two of the possible six genotypes were observed.

If a sample is not in HWE, it may suggest one of the following: (a) genotyping error, (b) a sample which poorly represents the populations or (c) a chance finding. To address the first issue, all data were re-examined for genotyping errors (and were subsequently shown to be correct when analysed independently by another researcher). Since only genotypes at the ADH1C locus deviated from HWE, it does not suggest that the samples collected is not representative of the population in the Western Cape.

Table 3-5: Hardy-Weinberg equilibrium analysis p-values shown for all the polymorphisms tested in the Western and Northern Cape for each group (FAS children, Mothers and Controls)

Marker	FAS children	Mothers	Controls
Western Cape			
ADH1B	1	1	0.4367
ADH1C	1	0.4243	**0.0153**
ADH4.4	1	1	0.3156
ADH4.8	0.4764	1	1
Northern Cape			
ADH1B	N/A	1	1
ADH1C	0.5845	1	0.7926
ADH4.4	0.4543	0.4665	1
ADH4.8	0.1285	0.1655	1

Key: **bold** = $p < 0.05$

3.3 Tests for associations with susceptibility for FAS

Allele and genotype frequencies of the FAS-affected children and their mothers were used to test for association between FAS and the alcohol metabolism gene loci (Table 3.6).

3.3.1 Allele and genotype associations

The polymorphisms for ADH1B and ADH1C had previously been studied in the Western Cape region and these results were published in 2001 (Viljoen et al., 2001). An additional fourteen FAS-affected children and 13 mothers were added to the Western Cape cohort and the data were re-analysed.

The allele frequencies observed at ADH1B confirmed a significant association (p=0.038) in the mothers of the Western Cape but did not reach significance (p=0.089) when the FAS-affected children were compared to the controls. The post-test power for detecting a significant difference, given the sample size, was only 57%. Allele 1 (ADH1B*1) occurred in 92% of the mothers compared to 85% in the controls. Allele 2 (ADH1B*2) occurred in 4% of the mothers and in 11% of the controls and allele 3 (ADH1B*3) occurred at the same frequency (4%) in both groups. It was inferred that allele 2 was likely to be a marker for protection as it was significantly more common in the control than the mothers. The allele comparison at ADH1B between the FAS-affected children and the controls in the Western Cape almost reached significance (p= 0.089). The lack of significance, compared to the published results, may be attributed to the single additional allele 2 observed in the FAS children.

ADH1B genotype comparisons did not reveal significant associations. The ADH1B(1/1) genotype was the most common genotype amongst all the groups in both provinces with the 1/2, 2/2 and 2/3 genotypes being less common (Table 3.2).

No significant associations were found for ADH1C when analysing the genotypes or the alleles. The common genotype for ADH1C was ADH1C(1/1) occurring at high

frequencies in the FAS-affected children, their mothers and controls in both the Western and Northern Cape regions. The least common genotype was ADH1C(2/2) in all three groups, however it was not observed in the mothers from the Northern Cape region.

Two polymorphic markers (ADH4.4 and ADH4.8) were genotyped at the ADH4 locus. Allelic association studies were significant for the second locus, ADH4.8, in the Western Cape for both the FAS children (p= 0.035) and their mothers (p= 0.002) when compared to the controls. The allele responsible for the association was the 'A' allele as it was significantly more common in the FAS-affected children and their mothers than the controls (55%, 61% and 40%, respectively). Interestingly, when genotypes were compared a significant association was only found in the mothers (p=0.010 for the genotype C/C), although the FAS children did show borderline significance (p= 0.061). It is worth noting that in the Western Cape the A/A genotype was more common in the FAS children and mothers compared to the controls (26%, 38% and 16%, respectively) while the C/C genotype was more common in the controls than in the FAS children and mothers (36%, 16% and 16% respectively). These trends were not observed in the Northern Cape. The post-test power for detecting associations with the ADH4.8 locus was higher in the Western Cape (44% and 75%) when compared to the Northern Cape (6% and 6%).

3.3.2 Haplotype analysis

When comparing the mothers to the controls in the Western Cape (Table 3.7), the global simulation p-value almost reached significance (p= 0.065) indicating that there might be a significant difference between the mothers and the controls. The most common haplotype was **11TC** occurring in 39% of the controls and 30% in the mothers. Individuals with a **21TC** haplotype were more likely to be controls (p = 0.010) and individuals with a **11TA** or **11GA** haplotype were more likely to be cases (p = 0.043 and p = 0.035, respectively). In the Northern Cape, the mothers were not significantly different from the controls (as indicated by the non-significant global simulation p-value, p=0.788)). No haplotypes showed significant differences in this group. When the mothers were compared to each other between the two provinces the global simulation p-value was 0.398 indicating that the mothers' genotypes were similar in each of the provinces (Table 3.9).

The results for the comparison of FAS-affected children and controls within each province showed no significant difference in either the Northern or Western Cape regions ($p = 0.806$ and $p = 0.450$) (Table 3.8). The most common haplotype was **11TC** occurring with a frequency of 38% in the FAS-affected children in Northern Cape and 33% in the Western Cape FAS-affected children. In the Western Cape individuals with **21TC** haplotype were more likely to be controls ($p = 0.033$) and individuals with **11TA** were more likely to be a FAS-affected individual ($p= 0.015$) occurring with a frequency of 31% compared to 19% in the controls.

3.3.2.1 WC vs NC

Interesting differences were observed when comparing the three study groups within the provinces (Table 3.9). The results show that the controls were significantly different to each other ($p= 0.0019$). The FAS-affected children and the mothers in the different provinces were not significantly different to each other ($p= 0.339$ and $p = 0.398$ respectively). The comparison of controls between the provinces resulted in a highly significant result ($p= 0.0019$) suggesting that there are significant genetic differences between the Coloured populations in the two provinces. To elucidate the genetic makeup of the two populations sampled, all three sample groups were combined and compared between the provinces. The results show that the two Coloured populations seem to be genetically different ($p=0.0067$). These results guided us not to pool the samples of the two provinces as they appeared to have significantly different histories and structures. The results of the haplotype analysis however, showed that even though the Coloured populations in the provinces appeared to differ from each other, the genotypes of the FAS-affected children and mothers were similar in the Western and Northern Cape regions.

Table 3-6: Case-control associations with individual polymorphic loci associated with the ADH1B, ADH1C and ADH4 genes and the p-values for the allelic and genotypic association tests. Bold represents all significant results.

Marker	FAS-affected children vs Controls		Mothers vs Controls	
	Allele	Genotype	Allele	Genotype
Western Cape				
ADH1B	0.089	0.346	**0.038**	0.071
ADH1C	0.432	0.201	0.897	0.947
ADH4.4	0.515	0.477	0.140	0.098
ADH4.8	**0.035**	0.061	**0.002**	**0.010**
Northern Cape				
ADH1B	0.171	0.211	0.584	0.577
ADH1C	0.286	0.374	0.132	0.270
ADH4.4	0.581	0.563	1.000	0.698
ADH4.8	0.706	0.299	0.678	0.278

Key: **bold** = $p<0.05$

Table 3-7: Haplotype analysis showing the estimated haplotype frequencies (*f*) in the controls and cases (FAS children or mothers) and the likelihood of association (p) in the Western Cape Coloured population

Haplotype	Global sim p	Hap score	control *f*	case *f*	p
FAS vs Controls	0.450				
21TC		-2.135	0.088	0.033	**0.033**
12TC		-1.314	0.111	0.040	0.098
11TC		-0.117	0.376	0.332	0.899
12TA		0.095	0.045	0.084	0.919
12GA		0.098	0.052	0.007	0.929
21TA		0.135	0.010	0.000	0.908
31TA		0.417	0.033	0.044	0.687
11GC		0.573	0.009	NA	0.632
11GA		0.905	0.065	0.093	0.376
11TA		2.302	0.193	0.311	**0.015**
Mothers vs Controls	0.065				
21TC		-2.586	0.088	0.000	**0.010**
11TC		-1.073	0.376	0.298	0.303
12TC		-0.929	0.111	0.086	0.354
12GA		0.296	0.052	0.048	0.785
31TA		0.537	0.033	0.045	0.617
12TA		0.612	0.045	0.063	0.558
11TA		2.153	0.193	0.284	**0.043**
11GA		2.185	0.065	0.140	**0.035**

Key: Hap score = haplotype score, *f* = frequency, NA = not available, **bold** = $p<0.05$

Table 3-8: Haplotype analysis showing the estimated haplotype frequencies (*f*) in the controls and cases (FAS children or mothers) and the likelihood of association (p) in the Northern Cape Coloured population.

Haplotype	Global sim p	Hap score	control *f*	case *f*	p
FAS vs Controls	0.806				
12GA		-1.428	0.040	0.003	0.141
11GA		-0.667	0.085	0.063	0.491
12TA		-0.245	0.069	0.079	0.825
12TC		-0.243	0.096	0.089	0.816
11TA		0.085	0.297	0.287	0.930
11TC		0.969	0.352	0.389	0.323
31TA		1.097	0.016	0.038	0.223
Mothers vs Controls	0.788				
12TC		-1.060	0.096	0.038	0.288
12TA		0.894	0.069	0.100	0.392
12GA		-0.830	0.040	NA	0.376
11TA		-0.601	0.297	0.196	0.559
21TC		-0.446	0.022	NA	0.931
11GA		0.346	0.085	0.111	0.726
31TA		1.186	0.016	0.041	0.279
11TC		1.374	0.352	0.479	0.152

Key: Hap score = haplotype score, *f* = frequency, NA = not available, **bold** = $p<0.05$

Table 3-9: Haplotypes analysis between the provinces for the FAS-affected children, mothers and controls

Haplotype	Global sim p	Hap score	WC *f*	NC *f*	p
All samples	**0.0067**				
21TC		-3.768	0.062	0.016	**0.000**
21TA		-2.087	0.011	0.000	**0.033**
12GA		-0.844	0.051	0.031	0.399
31TA		-0.829	0.036	0.025	0.412
12TC		-0.518	0.101	0.086	0.602
11GA		-0.085	0.081	0.079	0.933
11GC		0.550	0.013	0.016	0.420
12GC		0.855	NA	0.011	0.420
11TC		0.923	0.357	0.378	0.357
12TA		0.954	0.047	0.070	0.346
11TA		2.170	0.226	0.285	**0.032**
Controls	**0.0019**				
21TC		-3.431	0.088	0.022	**0.000**
31TA		-1.046	0.033	0.016	0.284
12TC		-0.649	0.111	0.096	0.516
11TC		-0.316	0.376	0.352	0.756
12GA		0.087	0.052	0.040	0.933
11GA		0.799	0.065	0.085	0.420
12TA		1.236	0.045	0.069	0.218
12GC		1.689		0.020	**0.056**
11TA		2.772	0.193	0.297	**0.005**
FAS	0.339				
12GA		-1.545	0.007	0.003	0.131
11GA		-0.372	0.093	0.093	0.063
11TA		-0.150	0.311	0.287	0.894
31TA		-0.015	0.044	0.038	1.000
12TA		0.273	0.084	0.079	0.789
12TC		0.654	0.040	0.089	0.529
11TC		0.876	0.332	0.389	0.392
11GC		0.890	NA	0.046	0.375
Mothers	0.398				
12GA		-1.073	0.048	NA	0.301
12TC		-0.791	0.086	0.038	0.446
11GA		-0.756	0.140	0.111	0.434
12TA		-0.502	0.063	0.100	0.621
11TA		-0.425	0.284	0.196	0.677
31TA		-0.105	0.045	0.041	0.840
11TC		2.305	0.298	0.470	**0.022**

Key: Hap score = haplotype score, *f* = frequency, NA = not available, WC = Western Cape, NC = Northern Cape, **bold** = $p<0.05$

3.3.3 Logistic regression

Three logistic regression models were generated; two to assess significant associations between the polymorphisms studied, geographic region and disease outcome and a third model to examine the genetic differences between the samples from the two geographic regions (Table 3.10).

In the first model, the genotype data for the provinces were pooled and the FAS-affected children were compared to the controls. This model contained the covariates of province and the marker ADH4.8. The baseline genotype for this model was ADH4.8(A/A) and the baseline province Northern Cape and all the individual p-values were based on the baselines. This model demonstrates that an individual from the Western Cape has reduced probability of being a case ($p= 0.018$). The results also indicated that an individual with either genotype A/C or C/C has an increased probability of being a control or a reduced chance of being affected, compared with the A/A genotype.

The second model compared the mothers to the controls and highlighted the difference in patterns observed in the Western and Northern Cape regions. This model also highlighted the interaction of ADH4.8 in each province. Individuals in the Northern Cape with a ADH4.8(A/C) or ADH4.8(C/C) genotype have an increased probability of being a mother of a FAS child. However in the Western Cape individuals with the A/C and C/C genotypes have a reduced chance of being a case and those with the A/A genotype have an increased probability of being a case.

To assess whether the two populations groups were similar to each other, in order to pool the data to increase the power, a third model was generated using only the controls from both provinces. When comparing the controls between the two regions, the model revealed that individuals with ADH1B(2/2), ADH1B(2/3) and ADH1B(3/3) genotypes were more likely to occur in the Western Cape. This model suggested a significant genetic difference between the Coloured populations of the Northern and Western Cape regions. This was also observed in the haplotype analysis and as a result of the genetic

diversity between the two regions, the samples could not be combined to create a larger sample size.

Table 3-10: Logistic regression analysis on genotype probabilities in FAS-affected children vs Controls, mothers of children with FAS vs Contols and NC vs WC (Controls only)

Model	Covariate	estimate	std error	p value
FAS-affected children vs Controls				
Global sim p				**0.011**
	Intercept	-1.005	0.312	0.0010
	WC	-0.636	0.268	0.0180
	ADH4.8(A/C)	-0.636	0.342	0.0180
	ADH4.8(C/C)	-0.409	0.415	0.3250
Mothers vs Controls				
Global sim p				**0.016**
	Intercept	-1.758	0.484	0.0000
	ADH4.8(A/C)	0.824	0.546	0.1310
	ADH4.8(C/C)	0.408	0.644	0.5260
	ADH4.8(A/A)WC	1.024	0.598	0.0870
	ADH4.8(A/C)WC	-0.662	0.379	0.0810
	ADH4.8(C/C)WC	-1.066	0.631	0.0910
NC vs WC(controls)				
Global sim p				**0.001**
	Intercept	0.2580	0.132	0.0505
	ADH1B(1/2)	1.5980	0.499	0.0013
	ADH1B(1/3)	0.5530	0.615	0.3680
	ADH1B(2/2)	15.308	1029.121	0.9880
	ADH1B(2/3)	15.308	840.274	0.9850
	ADH1B(3/3)	15.308	1455.398	0.9910

Key: **bold**= $p<0.05$

3.4 Linkage disequilibrium (LD)

In this study, LD was examined in the controls independently for each province and then combined. This resulted in the observation of interesting patterns of LD. When referring to the figures, all red blocks indicate high linkage disequilibrium (as indicated by the corresponding p-value), orange indicates moderate linkage disequilibrium and light brown indicates low or no linkage disequilibrium. The measures of LD (as calculated in the R package); D and D' and r are shown in each block respectively.

When all the controls were combined, ADH1B had the strongest LD with ADH4.4 and a high LD with ADH1C (Figure 3.1). ADH4.4 and ADH4.8 were also in very strong LD with each other. In the Western Cape cohort, LD patterns were similar to those observed when all the control samples were combined (Figure 3.2). It was shown that ADH1B was in strong LD with ADH4.4, and in moderate to high LD with ADH1C, and ADH4.4 and ADH4.8 were in strong LD. However, the patterns of LD observed in the Northern Cape cohort differed slightly from that in the Western Cape cohort (Figure 3.3). ADH1B was in high LD with ADH1C but ADH4.4 and ADH4.8 were in strong LD with each other.

Generally, LD between ADH1B and ADH1C and between ADH4.4 and ADH4.8 is higher than the LD between ADH1C and ADH4.4. This was expected since a greater physical distance occurs between ADH1C and ADH4.4.The observed differences in patterns of LD between the provinces suggest that there might be differences in population histories for the two groups. These differences contribute to the LD patterns observed when the two groups are combined This is supported by haplotype and logistic regression analysis.

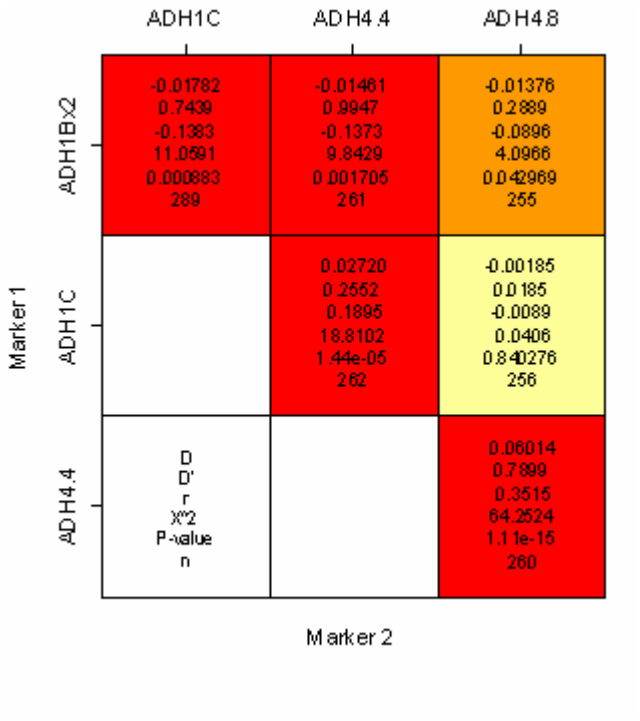


Figure 3.1: LD estimates for all the controls.

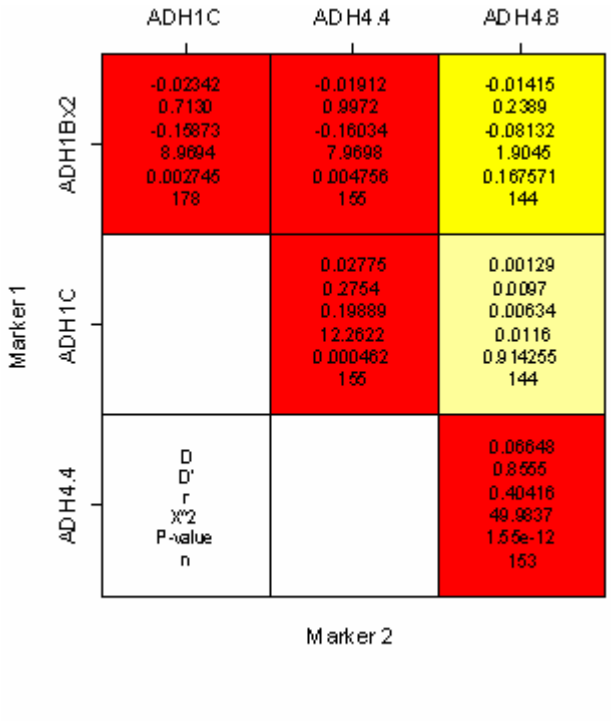


Figure 3.2: LD estimates for the controls from the Western Cape

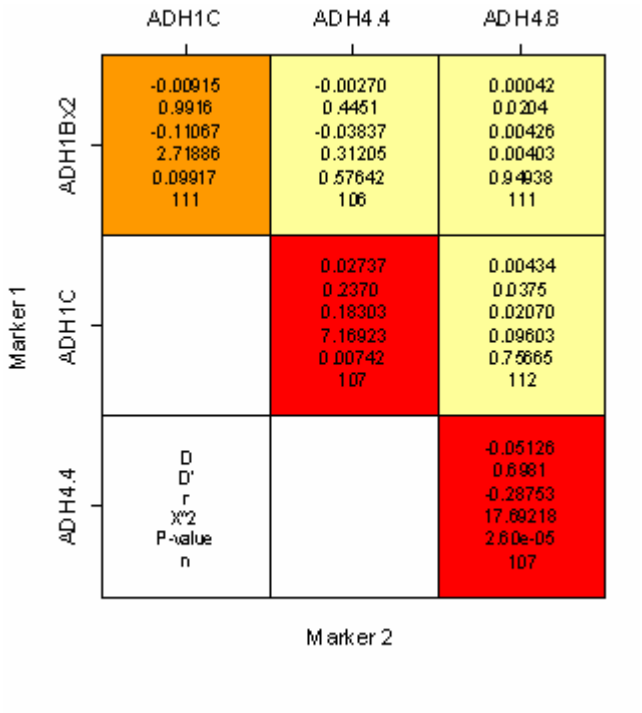


Figure 3.3: LD estimates for controls from the Northern Cape

Chapter 4

Discussion

4 Discussion

In this study we adopted the candidate gene case-control study design approach and focused our attention on the alcohol dehydrogenase genes since they play a vital role in first pass metabolism of alcohol. Several polymorphic variants of the ADH1B, ADH1C and ADH4 genes have been examined in two Coloured populations from the Western and Northern Cape regions. These two population groups were chosen as the incidence of FAS has been found to be the highest in the world.

It is known that FAS is a result of fetal exposure to alcohol. Although no individual mechanism has been identified as the main cause for the resulting damage, studies have shown that alcohol is able to affect the fetus by three possible routes. It can affect the fetus directly or through the exposure to metabolites such as acetaldehyde or alcohol can inhibit retinoic acid (RA) synthesis. By examining the ADH genes in the two Coloured communities, this project sought out to determine whether they play an important role in the disease pathogenesis.

4.1 The ADH1B and ADH1C genes and their role in the susceptibility to FAS

4.1.1 ADH1B

In the South African Western Cape Coloured population the ADH1B*2 allele has been associated with a protective effect against the development of FAS i.e. it occurred at a significantly higher frequency in the control population than the sample group of mothers (Viljoen et al., 2001). However, in the present study (representing an extension thereof), no significant associations were observed in the FAS-affected children samples from the Western Cape, perhaps due to the inclusion of the additional FAS-affected children. This observation could be due to the small sample size and low power. The analysis of the ADH1B genotypes (in the present study) resulted in a weakly significant association. The group of mothers from the Western Cape had a significant association between ADH1B*2 and the protection against the development of FAS, which was consistent with the observations from the previous study by Viljoen et al., (2001). Interestingly, no

significant associations were observed for allelic or genotypic association analysis in the Northern Cape Coloured sample.

Haplotype analysis performed to examine the interaction of the polymorphisms described in this study revealed that the alleles of ADH1B and ADH4 played a role in disease susceptibility. Individuals with both the ADH1B*2 allele and the 'C' allele of ADH4.8 have a greater probability of being in the control group in the Western Cape region (Table 3.7). This suggests that the presence of these two alleles in a haplotype confers some protection against the development of FAS. The presence of the ADH1B*1 allele and the 'A' allele of ADH4.8 in a haplotype showed an increased risk of being a case as it had a high haplotype score of 2.772 (p= 0.005). This is an interesting finding as it suggests that the interaction between these two loci or variants in strong LD with them, may increase the risk of being affected with the disease or being a mother of a FAS-affected child.

4.1.2 ADH1C

The ADH1C gene has been found to be associated with an effect that is protective against alcoholism only because of its strong disequilibrium with ADH1B*2. This was demonstrated in a haplotype based analysis between Taiwanese Chinese alcoholic and non-alcoholic individuals. The association with alcoholism was due to an allelic variation at the ADH1B*2 site. The observation that the ADH1C*1 allele was decreased in the alcoholic Taiwanese Chinese individuals relative to the control individuals was thought to be a result of the strong linkage disequilibrium with the ADH1B*2 allele(Osier et al., 1999).

The ADH1C*1 allele which encodes an enzyme with a high V_{max} results in efficient alcohol metabolism. An individual with an efficient alcohol metabolising enzyme may be likely to consume more alcohol and accumulate more metabolites thereby increasing the risk of FAS. It was therefore surprising not to find any significant associations with this locus and FAS in our study cohort. LD analysis revealed that there was stronger LD between ADH4.4 and ADH1C than ADH1B and ADH1C in the Western and Northern Cape regions.

The only significant finding was that the control frequencies did not meet with HWE expectation in the Western Cape region. This could be due to population stratification, selection bias, skewed or non-representative sampling and chance. All the results were independently scored for genotype assignment and duplicate sample testing suggests that genotyping errors are an unlikely explanation. It is probably a chance observation given that with a 5% significance level, one would expect deviation from HW proportions in one out of every 20 markers tested. From these studies it seems unlikely that ADH1C plays a significant role in FAS susceptibility in the Western and Northern Cape provinces.

4.2 The role of ADH4 in the pathogenesis of FAS

ADH4 plays a role in two important physiological processes; alcohol metabolism and RA synthesis. A model proposed by Duester (1991) hypothesises that alcohol interferes with RA synthesis via the ADH pathway. The resulting RA deficiency may contribute to the development of FAS since the symptoms of RA deficiency overlap with those of FAS. Supporting evidence for this comes from kinetic studies which show that when concentrations of alcohol in the bloodstream are high, alcohol competitively inhibits retinol from binding to ADH4 and reduces the levels of local retinoic acid. RA is critical for embryogenesis and organogenesis (Yin et al., 2003).

The allelic association tests identified the 'A' allele at ADH4.8 (rs1800759) as a susceptibility factor as it occurred more commonly in the FAS-affected children samples and the mother's samples than in the control samples ($p= 0.035$ and $p= 0.002$ respectively). At the genotypic level, the A/A genotype was more common in the mothers and the C/C genotype was more common in the controls ($p = 0.010$). This genotype difference was only marginally significant in the FAS-affected children ($p= 0.061$). The implications of these results in the mothers compared to the FAS-affected children will be discussed in section 4.3.

Logistic regression analysis revealed that ADH4.8 was a possible susceptibility marker but that its effect was population specific. The C/C genotype was more common in the mothers and the FAS-affected children (Table 3.4) from the Western Cape only. However this was not a statistically significant finding. Further support for ADH4.8 as a possible susceptibility marker was provided by haplotype analysis (section 4.11). The haplotype analysis suggested that the 'C' allele may confer a possible protective effect while the 'A' allele was more common in the FAS children and their mothers, possibly playing a role in increasing the risk of developing FAS.

Previous studies examining the alleles of ADH4.8 have shown that the frequencies vary between different populations. The frequency distribution observed in our study of the Western Cape population was most similar to that described by (Luo et al., 2005) in the European American population they sampled. The allele frequencies in the African American population were quite different to those in the European Americans with the 'A' allele being the most common. The study by (Guindalini et al., 2005) showed allele distributions similar to those observed in the African Americans with the most common allele being the 'A' allele (Table 4.1). Both studies by (Guindalini et al., 2005) and (Luo et al., 2005) found that the 'A' allele was associated with an increased risk of developing alcoholism.

Previous studies have identified the allelic variants of ADH4.8 as being responsible for the change in the ADH4 promoter activity. The presence of the 'C' allele has been shown to decrease the promoter activity while the presence of the 'A' allele increases the promoter activity (Edenberg et al., 1999). An interesting observation can be made from the haplotype analysis. The ADH1B*1 allele together with the 'A' allele from ADH4.8 was found to increase the probability of being a FAS-affected child or a mother of a FAS-affected child while the ADH1B*2 allele together with the 'C' allele was shown to confer protection against developing the disease.

A hypothesis can be proposed from these observations that may provide support for the model proposed by Duester (1999). This model proposes that a diversion of fetal ADH

activity away from retinol metabolism may lead to the lowering of retinoic acid levels in tissues. This is due to the ADH enzymes being involved in alcohol metabolism during alcohol intoxication because of their high k_m for alcohol rather than retinol. As a result these enzymes will not be available for retinol metabolism. Thus a high maternal dose of alcohol in an individual who has a relatively low rate of alcohol metabolism may lead to prolonged inhibitory effect on RA synthesised by the fetus for normal development

The ADH1B*1 allele which results in a slower alcohol clearance rate (Takeshita et al., 1996) and therefore results in an increased amount of alcohol in the bloodstream, was found to be associated with the 'A' allele of ADH4.8 which causes increased promoter activity. Due to the increased promoter activity, the level of ADH4 protein increases and may bind to some of the excess alcohol circulating in the bloodstream. This may competitively inhibit retinol from binding to the ADH4 protein as proposed by the Duester (1991). If the excess alcohol levels are of such a magnitude that even in the presence of increased levels of ADH4 they block retinol binding, the local retinoic acid levels may decrease and this may increase susceptibility to FAS as retinoic acid is required for normal embryogenesis and it is thought that decreased levels of RA may contribute to the development of FAS. This may explain why these two alleles were associated with an increased risk of being a FAS-affected child or being a mother of a FAS-affected child.

The ADH1B*2 allele has a faster alcohol clearance rate (Takeshita et al., 1996) and has been associated with the 'C' allele of ADH4.8. The 'C' allele results in decreased promoter activity. Due to the faster alcohol metabolism rate at ADH1B, the excess alcohol is metabolised rapidly thus allowing retinol synthesis to occur optimally as the excess alcohol will not inhibit it from binding to ADH4. This is suggestive of a maternal protective effect.

The result of fast alcohol metabolism is the rapid accumulation of acetaldehyde. This metabolite is able to pass through the placenta into the fetus and may have a toxic effect on the fetus (Chaudhuri, 2000). The build up of acetaldehyde also results in adverse

symptoms that should prevent further drinking but its effects go unnoticed due to the extreme binge drinking patterns occurring in these communities. The results from this study indicate that acetaldehyde may not play a significant role as one would expect. The data suggested that having a faster alcohol clearance rate due to the interaction between ADH1B and ADH4 may have a protective effect against FAS.

Table 4-1: Published allele and genotype frequencies from various studies for three polymorphic variants in the ADH4 genes compared to the observed frequencies in the two Coloured samples.

Marker	European Americans[a]		African Americans[a]		Caucasian[b]		African Brazilian[b]		Western Cape[c]		Northern Cape[c]	
	n	*f*	n	*f*	n	*f*	n	*f*	n	*f*	n	*f*
ADH4.8 (rs1800759)												
A/A	44	0.145	33	0.688	14	0.200	7	0.318	25	0.160	29	0.260
A/C	154	0.507	15	0.313	53	0.760	14	0.636	74	0.480	56	0.500
C/C	106	0.349	0	0.000	3	0.04	1	0.045	56	0.360	27	0.240
A	242	0.398	81	0.844	81	0.578	28	0.636	124	0.400	114	0.510
C	366	0.602	15	0.156	59	0.421	16	0.363	186	0.600	110	0.490
ADH4.4 (rs1126670)												
C/C	24	0.077	1	0.021					1	0.010	2	0.020
A/C	145	0.468	14	0.298					41	0.250	28	0.260
A/A	141	0.455	32	0.681					124	0.750	77	0.720
C	193	0.311	16	0.170					43	0.130	32	0.150
A	427	0.689	78	0.830					289	0.870	182	0.850
rs1126671												
A/A	24	0.078	1	0.021					0	0	0	0
A/G	147	0.476	15	0.313					2	0.014	1	0.009
G/G	138	0.447	32	0.667					132	0.985	103	0.990
A	195	0.316	17	0.177					2	0.007	1	0.004
G	423	0.684	79	0.823					266	0.992	207	0.995

Key: a = Luo et al.2005, b = Guindalani et al.2005, c = current study

4.3 The impact of the maternal and fetal genotypes on disease outcome

Prenatal alcohol exposure results in a spectrum of disorders of which, FAS is at the extreme end. Studies have shown that not all children born to mothers who drink during their pregnancies are born with FAS. Some children are unharmed while others may have partial characteristics of FAS. This raises the question as to whether the maternal or fetal genotypes or a combination of both, play a role in the disease outcome. In this study we examined the FAS-affected children and their mothers to determine whether their genotypes do play a role in disease pathogenesis.

The marker ADH4.8(rs1800759) best demonstrated the impact of the maternal and fetal genotypes on the disease outcome. The mothers had a higher frequency of the 'A' allele compared to the FAS-affected children (61% vs 55%). The 'A' allele has been shown to increase the activity of the ADH4 promoter activity. The FAS-affected children had a higher frequency of the 'C' allele compared to the mothers (45% vs 39%) and this allele has a lower promoter activity. Genotypic association tests revealed that the A/A genotype was significantly different between the mothers and the controls ($p= 0.010$). This suggests that mothers with a faster alcohol metabolism rate are at an increased risk of having a child with FAS. This could be due to factors such as the exposure to acetaldehyde as a result of the fast alcohol metabolism or the competitive inhibition between alcohol and retinol, preventing retinol from binding to ADH4 and resulting in reduced retinoic acid synthesis. Evidence to support this hypothesis can be drawn from a study conducted by (Bhalla et al., 2005). ADH activity was measured in the liver and intestines in rats that were prenatally exposed to alcohol. The study found that the ADH activities were low in the pups which accounted for the high concentration of alcohol in the tissues. Since no ADH activity has been detected in the placenta and ADH activity seems to be low in the fetus, it suggests that the maternal metabolic pathway is largely responsible for the removal of circulating alcohol from the blood (Bhalla et al., 2005).

However, when examining the ADH1B locus, we found that ADH1B*2 was protective in the mothers from the Western Cape. This allele is known to metabolise alcohol faster

(Takeshita et al., 1996) and this faster clearance rate would prevent the fetus from being exposed to too much alcohol.

This finding contradicts the observations made for ADH4.8 which showed that faster alcohol metabolism increased the risk of having a child with FAS. Although these results seem to contradict each other, the effect of ADH4.8 on disease outcome appears to be greater than ADH1B due to the power at which the associations were detected. The post-test power revealed that the results for ADH4.8 had a 75% chance of being replicated in another study indicating that these findings were more than likely to be real findings and not a spurious association.

4.4 Genetic diversity of the Western and Northern Coloured populations

One of the biggest confounding factors in studying complex diseases is population stratification. Association studies are usually performed in samples of unrelated individuals to identify susceptibility loci. Spurious associations may arise due to the differences in allele frequencies if the cases and controls are not appropriately matched by ethnicity (Risch, 2000). To overcome this problem in the current study, the cases and controls were ethnically and geographically matched. It is thought by some researchers that the use of admixed populations would allow for the easier detection of disease-causing variants. However, in the present study this is complicated by the fact that admixture is an ongoing process. The South African Coloured population falls within the continuous gene flow (CGF) model of admixture (Halder and Shriver, 2003). This model states that admixture is occurring in each generation at a steady rate. As a result linkage disequilibrium is increased between markers on the same chromosome and it is thought that the linkage disequilibrium may be helpful in identifying possible disease markers.

To examine the genetic make-up in the two regions, haplotype analysis, logistic regression and linkage disequilibrium analysis were performed. A very important and perhaps not entirely surprising finding from this study was the genetic diversity observed between the two South African Coloured populations resident in the Northern and

Western Cape regions. Due to the genetic diversity observed between the two regions, each dataset generated was analysed separately.

Through haplotype analysis, we found that the Western and Northern Cape controls were significantly different to each other ($p = 0.0019$). Further support was provided by logistic regression which revealed that the presence or absence of three ADH1B genotypes; ADH1B(2/2), ADH1B(2/3) and ADH1B(3/3) distinguished the two province cohorts from each other (these genotypes were only observed in the Western Cape controls). Finally linkage disequilibrium testing revealed different patterns between the two populations. Published studies on alcoholic and non-alcoholic individuals from three Taiwanese populations ($n = 61$ and $n = 87$ for alcoholic samples and controls respectively) and Taiwanese Chinese populations ($n = 128$ and $n = 135$ for alcoholic samples and controls respectively) have shown that ADH1B is in strong LD with ADH1C however there was low LD between ADH4 and the ADH1 gene cluster (Osier et al., 1999). In 1987, Murray et al. showed that LD existed between two polymorphic sites (*StuI* and *XbaI*) on the ADH1C gene in a sample of European Americans. LD has also been examined between ADH1B, ADH1C, ADH4 and ADH5 in samples of mixed European Americans and Swedes (Edman and Maret, 1992). The researchers found strong LD between ADH1B and ADH1C and between ADH4 and ADH5. Low LD was observed between the Class I genes and either the ADH4 or ADH5 genes. More recently Osier et al., (1999) examined LD in ADH1B, ADH1C and ADH7 genes in 40 populations. Six polymorphic sites were examined in the Class I gene family and one in the ADH7 gene. In all the populations studied, significant LD occurred within the Class I genes sites. Pairwise LD also occurred between ADH7 and some sites in the ADH1 gene cluster. However in the South African Coloured population it was found that ADH1B was in stronger LD with ADH4.4 than ADH1C. This finding shows that the South African Coloured community is genetically different to other populations throughout the world. In the Western Cape, LD showed that ADH1B had highest LD with ADH4.4 and in the Northern Cape, ADH1C was in high LD with ADH4.4. This demonstrates the contrasting results obtained for the two provinces.

The finding that the Western Cape and Northern Cape Coloured people are different to each other was not surprising as a previous study examining Y chromosome variation also showed significant genetic diversity between groups of South African Coloureds (Motlidile,2004). Individuals from three different Coloured communities within South Africa were studied; Cape Malay, Cape Coloured from the Western Cape and a Coloured community from Johannesburg. The study showed that the Cape Coloured community was most similar to the Johannesburg Coloured community and both these differed from the Cape Malay group. The Cape Malay community was shown to have a higher Y chromosome contribution from Asia (46%) compared to the higher African paternal contributions in the Cape Coloured and Johannesburg Coloured groups (Motlidile,2004). These findings support historical data on the movement of people into the Western Cape.

Settlement in the Western Cape dates as far back as 1652 when some of the first European settlers arrived to form a supply base for travellers. As a consequence of this colonisation and the need for cheap labour, slaves were brought in from Asia, India, East and West Africa (Botha, 1972). The Coloured populations each have a unique genetic background due to differing genetic contributions from the Europeans, Asians, Africans and local indigenous groups that settled in the Western Cape.

The Coloured communities were chosen for this study because they have the highest rates in FAS in the world. We had hoped to combine the two datasets generated from the two provinces sampled to create a larger sample size. Due to the diverse genetic contributions from the various parental populations, the datasets could not be pooled. An important criterion in evaluating a case-control study is the ability to replicate the results. Although we failed to replicate some significant results in the Northern Cape province, all the appropriate measures were taken to attempt to control for population stratification. Even though the populations were found to have different genetic backgrounds, when the mothers and the FAS-affected children were compared to each other between the two provinces, haplotype analysis revealed that there were no significant differences between these two groups. This in itself is a very important finding as it suggests a similar genetic

background within these two groups that may reflect a common role in disease susceptibility.

4.5 ADH and brain development

The most severe symptom of FAS is poor neurological development. Alcohol can affect the developing brain either directly or indirectly, causing regions of the brain that deal with cognitive skills, motor skills, memory, learning and decision making to be severely affected.

Direct alcohol exposure causes abnormalities in the serotonin/glutamate neurotransmitter system, the premature conversion of glial cells into astrocytes and cell death via apoptosis and necrosis. Free radicals, which are produced as a result of alcohol metabolism cause oxidative stress. The occurrence of excess free radicals increases the toxicity levels in the cell which can induce apoptosis (Goodlett and Horn, 2001). The CNS is also vulnerable to damage by alcohol because the cells located there have a very low threshold for alcohol (Welch-Carre, 2005).

Could the ADH genes regulate the amount of alcohol exposure to the brain? The distribution of ADH has been studied in the brain and neither ADH1 nor ADH4, the enzymes largely responsible for alcohol metabolism, have been found to be expressed in human brain tissue. However ADH3 has been found in the regions of the brain that are severely affected by alcohol exposure such as the hippocampus, cerebellum and corpus callosum. It has been proposed that ADH3 might play a role in alcohol metabolism in these tissues (Galter et al., 2003). In recent years, an alternative pathway has been proposed for brain ethanol-oxidising properties that involves the enzyme catalase. It is thought that brain catalase oxidises ethanol through its peroxidatic activity (Quertemont et al., 2005).

The ADH genes involved in retinoid signalling via the ADH pathway have not been detected in the brain implying that retinoic acid synthesis occurs via other secondary pathways. There is increasing evidence that retinoid signalling plays an important role in

the function of the adult brain. Detection of retinoid components in the brain suggest that retinoic acid can by synthesised in certain regions of the brain (Lane and Bailey, 2005). Animal models have also been used to demonstrate that the retinoic acid present in brain tissue can activate gene expression in these same areas.

As none of the ADH loci we studied appear to be present in the brain it would appear that any susceptible or protective effect in brain development would be a result of lowering levels of blood alcohol that reach the brain during development.

4.6 Limitations of the study

A major limitation of this study is the small sample size. This study demonstrated that the samples from the Western and Northern Cape regions could not be pooled and therefore larger numbers of samples need to be collected from each particular region. Numerous factors contribute to the difficulties in collecting samples, these include; access to the rural areas, children left in the care of guardians due to early mortality of mothers, lack of telephones, high mobility within and out of the regions in search of employment. It is also difficult to collect information and blood samples or cheek swabs due to the sensitive nature of the disease. There is a risk of families being ostracised from their communities because of the stigma associated with FAS.

The controls were randomly selected individuals collected from the two Coloured communities and their unknown FAS status could be considered a confounding factor in this study.

Many of the statistically significant associations found in the Western Cape were not detected in the Northern Cape. The sample size needs to be increased in each region to determine if these differences are due to low power in detecting associations or are real differences.

The extent of admixture needs to be studied in great detail to detect to which extent different parental populations have contributed to the populations. A possible future

direction for this project is to use family-based association study designs such as the transmission disequilibrium test (TDT). These strategies have become common in the study of complex diseases as they avoid spurious associations related to population stratification. Another strategy being used to assess population stratification is the use of polymorphic markers that are not linked to the candidate genes associated with the disease. An absence of an association will provide evidence that any association found in the cases and controls is not spurious (Silverman and Palmer, 2000). Family studies pose certain disadvantages such as the need to genotype more individuals (affected sibling and two parents), locate all the members of the family and in particular to locate the fathers. In the instances where fathers are absent, the siblings will be used. This study design will be used in future studies to support the results from the current association study. Great effort is being made to collect the large numbers of samples that are required from these communities.

In this study, we did not correct for multiple testing. The chances of finding spurious associations when studying complex diseases is always high when examining large numbers of polymorphisms in a complex disease. To overcome this problem stringent tests such as the Bonferroni correction have been used. However, the use of such corrections often means the loss of genuine associations. It is thought by some researchers that finding spurious associations when studying a polymorphism that has functional significance to the disease in question, is greatly reduced (Daly and Day, 2001). Our most significant finding in this study involved the ADH4.8 (rs1800759) polymorphism which has been shown to affect promoter activity.

4.7 Future directions

The ADH4 gene locus has been identified as a possible susceptibility locus involved in the pathogenesis of FAS. With the completion of Phase I of the HapMap project (Altshuler et al., 2005) it is now possible to select other informative SNPs in the ADH4 gene that might be associated with the development of FAS. HapMap aims to create a database containing common human sequence variation. The database will contain information on common SNPs and information on LD throughout the genome. One advantage of this is the ability to select tag SNPs for association studies. The use of such informative markers reduces redundancy and minimises the loss of information. However it needs to be established whether the populations we study exhibit similar LD and haplotype blocks to those used to generate the data for HapMap.

The SASA method developed to detect SNP rs1042363 from the ADH4 gene, needs to be optimised to test whether it plays a role in the disease as it was shown previously to be associated with alcohol and drug dependence. Due to the lack of robustness of technique in our hands, this method could not be optimised. It has been shown to be very specific method for detecting genotypes and will minimise the time taken generating results as compared to the ARMS method.

As mentioned previously, family based association studies may help to overcome the effect of population stratification. Samples are currently being collected for this part of the study.

Significant associations have been found with ADH4 and other genes involved in the retinal pathway that may play a role in FAS etiology. One such gene is ALDH1A2 (also known as RALDH2), needs to be genotyped in the current cohort of samples.

4.8 Conclusion

South Africa has the highest incidence of FAS in the world (40 per 1000 affected children in the Coloured community of the Western (Viljoen et al.,2001) and it poses a major health risk. The identification of the genetic factors that place a fetus at risk of the harmful effects of prenatal alcohol exposure and the elucidation of the mechanism by which alcohol causes its toxic effects may give researchers a better understanding of the disease.

The data from this study has led to the proposal of a model involving ADH1B and ADH4 in the development of FAS. When the ADH1B*2 allele occurs with the 'C' allele of ADH4.8, it results in a protective effect or acts as a marker of decreased risk. This suggests that individuals with a faster alcohol metabolism via the ADH1B gene may be protected from developing FAS. When ADH1B*1 occurred with the 'A' allele of ADH4.8, this increased the risk of being affected with FAS. In this instance, the excess alcohol due to the slower alcohol clearance rate may competitively inhibit retinol from binding to ADH4, thus resulting in reduced levels of retinoic acid synthesis. These results provide support for a model proposed by Duester (1991) suggesting that the etiology of FAS is due to the competition of alcohol and retinol for ADH4 leading to increased or decreased levels of RA.

Our results indicate that mothers with a faster alcohol metabolism were at an increased risk of having a child with FAS. In the fetus, it was shown that a slower metabolism protected the fetus from developing FAS. There is some contradiction between the roles of ADH1B and ADH4, where one functions differently the other is quantitatively increased, it is difficult to tell which gene plays a more important role. However one can assume that the interaction between ADH1B and ADH4 plays an important role in disease etiology.

An interesting outcome of this study was the observed diverse genetic backgrounds of the two Coloured groups sampled. Significant associations were not always replicated

between the groups possibly as a result of this diversity, indicating that susceptibility or protective factors may vary in these genetically distinct populations.

This study has provided evidence that ADH4 is likely to play an important role in the pathogenesis of FAS due to the highly significant associations that were found in the Coloured group of samples from the Western Cape province.

5 References Cited

Abel, E.L. (1995) An update on incidence of FAS: FAS is not an equal opportunity birth defect. *Neurotoxicol Teratol* **17:** 437-43.

Abel, E.L. (1998) Fetal alcohol Abuse Syndrome. Plenum Press, New York.

Abel, E.L. and Sokol, R.J. (1986) Fetal alcohol syndrome is now leading cause of mental retardation. *Lancet* **2:** 1222.

Agarwal, D.P. (2001) Genetic polymorphisms of alcohol metabolizing enzymes. *Pathol Biol (Paris)* **49:** 703-9.

Altshuler, D., Brooks, L.D., Chakravarti, A., Collins, F.S., Daly, M.J., and Donnelly, P. (2005) A haplotype map of the human genome. *Nature* **437:** 1299-320.
Notes: CORPORATE NAME: International HapMap Consortium.

Armstrong, E.M. (1998) Diagnosing moral disorder: the discovery and evolution of fetal alcohol syndrome. *Soc Sci Med* **47:** 2025-42.

Astley, S.J. and Clarren, S.K. (2000) Diagnosing the full spectrum of fetal alcohol-exposed individuals: introducing the 4-digit diagnostic code. *Alcohol Alcohol* **35:** 400-10.

Baraona, E., Yokoyama, A., Ishii, H., Hernandez-Munoz, R., Takagi, T., Tsuchiya, M., and Lieber, C.S. (1991) Lack of alcohol dehydrogenase isoenzyme activities in the stomach of Japanese subjects. *Life Sci* **49:** 1929-34.

Bhalla, S., Kaur, K., Mahmood, A., and Mahmood, S. (2005) Postnatal development of alcohol dehydrogenase in liver and intestine of rats exposed to ethanol in utero. *Indian J Med Res* **121:** 39-45.

Borras, E., Coutelle, C., Rosell, A., Fernandez-Muixi, F., Broch, M., Crosas, B. et al. (2000) Genetic polymorphism of alcohol dehydrogenase in europeans: the ADH2*2 allele decreases the risk for alcoholism and is associated with ADH3*1. *Hepatology* **31:** 984-9.

Botha, M.C. (1972) Blood group gene frequencies. An indication of the genetic constitution of population samples in Cape Town. *Am J Roentgenol Radium Ther Nucl Med* **115:** Suppl:1-27.

Burd, L., Cotsonas-Hassler, T.M., Martsolf, J.T., and Kerbeshian, J. (2003) Recognition and management of fetal alcohol syndrome. *Neurotoxicol Teratol* **25:** 681-8.

Carlson, C.S., Eberle, M.A., Kruglyak, L., and Nickerson, D.A. (2004) Mapping complex disease loci in whole-genome association studies. *Nature* **429:** 446-52.

Chaudhuri, J.D. (2000) Alcohol and the developing fetus--a review. *Med Sci Monit* **6:** 1031-41.

Chen, C.C., Lu, R.B., Chen, Y.C., Wang, M.F., Chang, Y.C., Li, T.K., and Yin, S.J. (1999) Interaction between the functional polymorphisms of the alcohol-metabolism genes in protection against alcoholism. *Am J Hum Genet* **65:** 795-807.

Cohen-Kerem, R. and Koren, G. (2003) Antioxidants and fetal protection against ethanol teratogenicity. I. Review of the experimental data and implications to humans. *Neurotoxicol Teratol* **25:** 1-9.

Crain, L.S., Fitzmaurice, N.E., and Mondry, C. (1983) Nail dysplasia and fetal alcohol syndrome. Case report of a heteropaternal sibship. *Am J Dis Child* **137:** 1069-72.

Crome, I.B. and Glass, Y. (2000) The DOP system: a manifestation of social exclusion. A personal commentary on 'Alcohol consumption amongst South African farm workers: a post-apartheid challenge, by L. London 1999'. *Drug Alcohol Depend* **59:** 207-8.

Croxford, J. and Viljoen, D. (1999) Alcohol consumption by pregnant women in the Western Cape. *S Afr Med J* **89:** 962-5.

Cudd, T.A. (2005) Animal model systems for the study of alcohol teratology. *Exp Biol Med (Maywood)* **230:** 389-93.

Daly, A.K. and Day, C.P. (2001) Candidate gene case-control association studies:

advantages and potential pitfalls. *Br J Clin Pharmacol* **52:** 489-99.

DelRio-LaFreniere, S.A. and McGlennen, R.C. (2001) Simultaneous allele-specific amplification: a strategy using modified primer-template mismatches for SNP detection--application to prothrombin 20210A (factor II) and factor V Leiden (1691A) gene mutations. *Mol Diagn* **6:** 201-9.

Duester, G. (1991) A hypothetical mechanism for fetal alcohol syndrome involving ethanol inhibition of retinoic acid synthesis at the alcohol dehydrogenase step. *Alcohol Clin Exp Res* **15:** 568-72.

Duester, G. (1999) Function of alcohol dehydrogenase and aldehyde dehydrogenase gene families in retinoid signaling. *Adv Exp Med Biol* **463:** 311-9.

Duester, G. (2000) Families of retinoid dehydrogenases regulating vitamin A function: production of visual pigment and retinoic acid. *Eur J Biochem* **267:** 4315-24.

Duester, G. (2001) Genetic dissection of retinoid dehydrogenases. *Chem Biol Interact* **130-132:** 469-80.

Edenberg, H.J., Jerome, R.E., and Li, M. (1999) Polymorphism of the human alcohol dehydrogenase 4 (ADH4) promoter affects gene expression. *Pharmacogenetics* **9:** 25-30.

Edenberg, H.J., Xuei, X., Chen, H.J., Tian, H., Wetherill, L.F., Dick, D.M. et al. (2006) Association of alcohol dehydrogenase genes with alcohol dependence: a comprehensive analysis. *Hum Mol Genet* **15:** 1539-49.

Edman, K. and Maret, W. (1992) Alcohol dehydrogenase genes: restriction fragment length polymorphisms for ADH4 (pi-ADH) and ADH5 (chi-ADH) and construction of haplotypes among different ADH classes. *Hum Genet* **90:** 395-401.

Egeland, G.M., Perham-Hester, K.A., Gessner, B.D., Ingle, D., Berner, J.E., and Middaugh, J.P. (1998) Fetal alcohol syndrome in Alaska, 1977 through 1992: an administrative prevalence derived from multiple data sources. *Am J Public Health* **88:** 781-6.

Eriksson, C.J., Fukunaga, T., Sarkola, T., Chen, W.J., Chen, C.C., Ju, J.M. et al. (2001) Functional relevance of human adh polymorphism. *Alcohol Clin Exp Res* **25:** 157S-163S.

Estonius, M., Svensson, S., and Hoog, J.O. (1996) Alcohol dehydrogenase in human tissues: localisation of transcripts coding for five classes of the enzyme. *FEBS Lett* **397:** 338-42.

Galter, D., Carmine, A., Buervenich, S., Duester, G., and Olson, L. (2003) Distribution of class I, III and IV alcohol dehydrogenase mRNAs in the adult rat, mouse and human brain. *Eur J Biochem* **270:** 1316-26.

Goodlett, C.R. and Horn, K.H. (2001) Mechanisms of alcohol-induced damage to the developing nervous system. *Alcohol Res Health* **25:** 175-84.

Guindalini, C., Scivoletto, S., Ferreira, R.G., Breen, G., Zilberman, M., Peluso, M.A., and Zatz, M. (2005) Association of genetic variants in alcohol dehydrogenase 4 with alcohol dependence in Brazilian patients. *Am J Psychiatry* **162:** 1005-7.

Halder, I. and Shriver, M.D. (2003) Measuring and using admixture to study the genetics of complex diseases. *Hum Genomics* **1:** 52-62.

Haines, J.L., Pericak-Vance, M.A. (1998) Approaches to gene mapping in complex human diseases. Wiley-Liss, New York

Hoyme, H.E., May, P.A., Kalberg, W.O., Kodituwakku, P., Gossage, J.P., Trujillo, P.M. et al. (2005) A practical clinical approach to diagnosis of fetal alcohol spectrum disorders: clarification of the 1996 institute of medicine criteria. *Pediatrics* **115:** 39-47.

Iida, A., Saito, S., Sekine, A., Kondo, K., Mishima, C., Kitamura, Y. et al. (2002) Thirteen single-nucleotide polymorphisms (SNPs) in the alcohol dehydrogenase 4 (ADH4) gene locus. *J Hum Genet* **47:** 74-6.

Jones, K.L. and Smith, D.W. (1973) Recognition of the fetal alcohol syndrome in early infancy. *Lancet* **2:** 999-1001.

Keir, W.J. (1991) Inhibition of retinoic acid synthesis and its implications in fetal alcohol syndrome. *Alcohol Clin Exp Res* **15:** 560-4.

Kimura, K.A., Chiu, J., Reynolds, J.N., and Brien, J.F. (1999) Effect of chronic prenatal ethanol exposure on nitric oxide synthase I and III proteins in the hippocampus of the near-term fetal guinea pig. *Neurotoxicol Teratol* **21:** 251-9.

Lane, M.A. and Bailey, S.J. (2005) Role of retinoid signalling in the adult brain. *Prog Neurobiol* **75:** 275-93.

Larkby, C. and Day, N. (1997) The effects of prenatal alcohol exposure. *Alcohol Health Res World* **21:** 192-8.

Lieber, C.S. (2000) Alcohol and the liver: metabolism of alcohol and its role in hepatic and extrahepatic diseases. *Mt Sinai J Med* **67:** 84-94.

Luo, X., Kranzler, H.R., Zuo, L., Lappalainen, J., Yang, B.Z., and Gelernter, J. (2005) ADH4 Gene Variation is Associated with Alcohol Dependence and Drug Dependence in European Americans: Results from HWD Tests and Case-Control Association Studies. *Neuropsychopharmacology*.

Maden, M. (2000) The role of retinoic acid in embryonic and post-embryonic development. *Proc Nutr Soc* **59:** 65-73.

May, P.A., Brooke, L., Gossage, J.P., Croxford, J., Adnams, C., Jones, K.L. et al. (2000) Epidemiology of fetal alcohol syndrome in a South African community in the Western Cape Province. *Am J Public Health* **90:** 1905-12.

May, P.A., Gossage, J.P., Marais, A.S., Adnams, C.M., Hoyme, H.E., Jones, K.L. et al. (2007) The epidemiology of fetal alcohol syndrome and partial FAS in a South African community. *Drug Alcohol Depend* **88:** 259-71.

McCarver, D.G., Thomasson, H.R., Martier, S.S., Sokol, R.J., and Li, T. (1997) Alcohol dehydrogenase-2*3 allele protects against alcohol-related birth defects among African Americans. *J Pharmacol Exp Ther* **283:** 1095-101.

Menegola, E., Broccia, M.L., Di Renzo, F., and Giavini, E. (2001) Acetaldehyde in vitro exposure and apoptosis: a possible mechanism of teratogenesis. *Alcohol* **23:** 35-9.

Michaelis EK and Michaelis ML Cellular and molecular bases of alcohol's teratogenic effects. *Alcohol Health and Research World* **18:** 17-22.

Miller, S.A., Dykes, D.D., and Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* **16:** 1215.

Motlidile, T.(2004) Y chromosome variation in the Coloured Community of South Africa. Masters Thesis, University of Witwatersrand.

Muramatsu, T., Wang, Z.C., Fang, Y.R., Hu, K.B., Yan, H., Yamada, K. et al. (1995) Alcohol and aldehyde dehydrogenase genotypes and drinking behavior of Chinese living in Shanghai. *Hum Genet* **96:** 151-4.

Norberg, A., Jones, A.W., Hahn, R.G., and Gabrielsson, J.L. (2003) Role of variability in explaining ethanol pharmacokinetics: research and forensic applications. *Clin Pharmacokinet* **42:** 1-31.

Ogurtsov, P.P., Garmash, I.V., Miandina, G.I., Guschin, A.E., Itkes, A.V., and Moiseev, V.S. (2001) Alcohol dehydrogenase ADH2-1 and ADH2-2 allelic isoforms in the Russian population correlate with type of alcoholic disease. *Addict Biol* **6:** 377-383.

Olney, J.W. (2004) Fetal alcohol syndrome at the cellular level. *Addict Biol* **9:** 137-49; discussion 151.

Osier, M., Pakstis, A.J., Kidd, J.R., Lee, J.F., Yin, S.J., Ko, H.C. et al. (1999) Linkage disequilibrium at the ADH2 and ADH3 loci and risk of alcoholism. *Am J Hum Genet* **64:** 1147-57.

Osier, M.V., Pakstis, A.J., Soodyall, H., Comas, D., Goldman, D., Odunsi, A. et al. (2002) A global perspective on genetic variation at the ADH genes reveals unusual patterns of linkage disequilibrium and diversity. *Am J Hum Genet* **71:** 84-99.

Palmer, R.H., Oullette, E.M., Warner, L., and Leichtman, S.R. (1974) Congenital malformations in offspring of a chronic alcoholic mother. *Pediatrics* **53:** 490-4.

Ponnappa, B.C. and Rubin, E. (2000) Modeling alcohol's effects on organs in animal models. *Alcohol Res Health* **24:** 93-104.

Quertemont, E., Tambour, S., and Tirelli, E. (2005) The role of acetaldehyde in the neurobehavioral effects of ethanol: a comprehensive review of animal studies. *Prog Neurobiol* **75:** 247-74.

Ramchandani, V.A., Bosron, W.F., and Li, T.K. (2001) Research advances in ethanol metabolism. *Pathol Biol (Paris)* **49:** 676-82.

Risch, N.J. (2000) Searching for genetic determinants in the new millennium. *Nature* **405:** 847-56.

Romero, R., Kuivaniemi, H., Tromp, G., and Olson, J. (2002) The design, execution, and interpretation of genetic association studies to decipher complex diseases. *Am J Obstet Gynecol* **187:** 1299-312.

Sampson, P.D., Streissguth, A.P., Bookstein, F.L., Little, R.E., Clarren, S.K., Dehaene, P. et al. (1997) Incidence of fetal alcohol syndrome and prevalence of alcohol-related neurodevelopmental disorder. *Teratology* **56:** 317-26.

Schenker, S., Becker, H.C., Randall, C.L., Phillips, D.K., Baskin, G.S., and Henderson, G.I. (1990) Fetal alcohol syndrome: current status of pathogenesis. *Alcohol Clin Exp Res* **14:** 635-47.

Schork, N.J. (1997) Genetics of complex disease: approaches, problems, and solutions. *Am J Respir Crit Care Med* **156:** S103-9.

Silverman, E.K. and Palmer, L.J. (2000) Case-control association studies for the genetics of complex respiratory diseases. *Am J Respir Cell Mol Biol* **22:** 645-8.

Smith, S.M. (1997) Alcohol-induced cell death in the embryo. *Alcohol Health Res World*

21: 287-97.

Sokol, R.J., Miller, S.I., and Reed, G. (1980) Alcohol abuse during pregnancy: an epidemiologic study. *Alcohol Clin Exp Res* **4:** 135-45.

Stoler, J.M., Ryan, L.M., and Holmes, L.B. (2002) Alcohol dehydrogenase 2 genotypes, maternal alcohol use, and infant outcome. *J Pediatr* **141:** 780-5.

Streissguth, A.P., Aase, J.M., Clarren, S.K., Randels, S.P., LaDue, R.A., and Smith, D.F. (1991) Fetal alcohol syndrome in adolescents and adults. *JAMA* **265:** 1961-7.

Streissguth, A.P. and Dehaene, P. (1993) Fetal alcohol syndrome in twins of alcoholic mothers: concordance of diagnosis and IQ. *Am J Med Genet* **47:** 857-61.

Streissguth, A.P., Landesman-Dwyer, S., Martin, J.C., and Smith, D.W. (1980) Teratogenic effects of alcohol in humans and laboratory animals. *Science* **209:** 353-61.

Sulik, K.K., Johnston, M.C., and Webb, M.A. (1981) Fetal alcohol syndrome: embryogenesis in a mouse model. *Science* **214:** 936-8.

Takeshita, T., Mao, X.Q., and Morimoto, K. (1996) The contribution of polymorphism in the alcohol dehydrogenase beta subunit to alcohol sensitivity in a Japanese population. *Hum Genet* **97:** 409-13.

Tshabalala, M (2003) Assesment of candidate genes for Fetal Alcohol Syndrome. Masters Thesis, University of Witwatersrand.

Vasiliou, V., Pappa, A., and Petersen, D.R. (2000) Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chem Biol Interact* **129:** 1-19.

Venables, W.N. and Ripley, B. D. Modern Applied Statistics with S, 4th edition. 2002. Springer.

Viljoen, D., Croxford, J., Gossage, J.P., Kodituwakku, P.W., and May, P.A. (2002) Characteristics of mothers of children with fetal alcohol syndrome in the Western Cape Province of South Africa: a case control study. *J Stud Alcohol* **63:** 6-17.

Viljoen, D.L., Carr, L.G., Foroud, T.M., Brooke, L., Ramsay, M., and Li, T.K. (2001) Alcohol dehydrogenase-2*2 allele is associated with decreased prevalence of fetal alcohol syndrome in the mixed-ancestry population of the Western Cape Province, South Africa. *Alcohol Clin Exp Res* **25:** 1719-22.

Warren, K.R. and Li, T.K. (2005) Genetic polymorphisms: impact on the risk of fetal alcohol spectrum disorders. *Birth Defects Res A Clin Mol Teratol* **73:** 195-203.

Weinberg, N.Z. (1997) Cognitive and behavioural deficits associated with parental alcohol use. *J Am Acad Child Adolesc Psychiatry* **36:** 1177-86.

Welch-Carre, E. (2005) The neurodevelopmental consequences of prenatal alcohol exposure. *Adv Neonatal Care* **5:** 217-29.

Wigginton, J.E., Cutler, D.J., and Abecasis, G.R. (2005) A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* **76:** 887-93.

Wilson, J.G., Roth, C.B., and Warkany, J. (1953) An analysis of the syndrome of malformations induced by maternal vitamin A deficiency. Effects of restoration of vitamin A at various times during gestation. *Am J Anat* **92:** 189-217.

Yelin, R., Schyr, R.B., Kot, H., Zins, S., Frumkin, A., Pillemer, G., and Fainsod, A. (2005) Ethanol exposure affects gene expression in the embryonic organizer and reduces retinoic acid levels. *Dev Biol* **279:** 193-204.

Yin, S.J., Chou, C.F., Lai, C.L., Lee, S.L., and Han, C.L. (2003) Human class IV alcohol dehydrogenase: kinetic mechanism, functional roles and medical relevance. *Chem Biol Interact* **143-144:** 219-27.

Yin, S.J., Han, C.L., Lee, A.I., and Wu, C.W. (1999) Human alcohol dehydrogenase family. Functional classification, ethanol/retinol metabolism, and medical implications. *Adv Exp Med Biol* **463:** 265-74.

Yokoyama, A. and Omori, T. (2003) Genetic polymorphisms of alcohol and aldehyde dehydrogenases and risk for esophageal and head and neck cancers. *Jpn J Clin Oncol* **33:**

111-21.

Zile, M.H. (1998) Vitamin A and embryonic development: an overview. *J Nutr* **128:** 455S-458S.

Zondervan, K.T. and Cardon, L.R. (2004) The complex interplay among factors that influence allelic association. *Nat Rev Genet* **5:** 89-100.

5.1 Electronic references

Gregory Warnes and Friedrich Leisch (2005). Genetics: Population Genetics. R package version 1.1.4.
http://www.R-project.org

Jason P. Sinnwell and Daniel J. Schaid (2005). Haplostats: Statistical Analysis of Haplotypes with Traits and Covariates when Linkage Phase is Ambiguous. R package version 1.2.2
http://mayoresearch.mayo.edu/mayo/research/biostat/schaid.cfm

R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

www.neb.com was used to determine the restriction enzymes that were used for the RFLP analysis

http://genome.ucsc.edu was used to perform a BLAST for the primer sequences against the genome

www.gene.ucl.ac.uk/nomenclature was used to determine the correct ADH gene nomenclature

http://www.home.clara.net/sisa/ was used to calculate allele and genotype frequencies, HWE and associations between markers of interest and disease status

6 Appendix A: Solutions

1% Alconox solution

1g Alconox per 100ml dH_2O

0.5M EDTA (pH 8.0)

93.6g EDTA into 400ml dH_2O

pH to 8.0 with NaOH pellets

Make to 500ml with dH_2O

Autoclave

10X dNTPs

125µl dATP(10mM stock)

125µl dCTP(10mM stock)

125µl dGTP(10mM stock)

125µl dTTP(10mM stock)

500µl ddH_2O

Ficoll loading dye

50g sucrose

10ml 0.5M EDTA

0.1g 0.1% bromophenol blue

10g 10% Ficoll

make up to a final volume of 100ml with dH_20

Saturated NaCl

100ml autoclaved dH_2O

Slowly add 40g NaCl until saturated.

Let NaCl precipitate out before using by gently agitating

Sucrose-Triton X lysing buffer

20ml 1M Tris HCL

10ml 1m Mgcl2

20ml Triton-X 100

Make up to 2l with dH_2O

Autoclave

Add 109.5g sucrose per litre just before use and keep chilled

0.1M Spermidine

0.5g Spermidine

20ml dH_2O

Aliquot into 1ml Eppendorf tubes

Store at -20°C

T20E5

10ml 1M Tris-HCL (pH 8.0)

5ml 0.5M EDTA (pH 8.0)

Make up to final volume of 500ml with dH_2O

Autoclave

1M Tris-HCL(pH 8.0)

60.55g Tris in 400ml dH_2O

pH using concentrated HCL

Make up to final volume of 500ml with dH_2O

Molecular weight marker

10.9µl 1kb ladder

5µl ficoll dye

84µl 1X TE buffer

Polyacrylamide gel master mix for sequencing

36g	Urea
10ml	10X TBE buffer
10.6ml	40% BIS-acrylamide
50ml	ddH_2O

Filter and store in the dark at 4°C

Polyacrylamide gel mix

30ml	Gel master mix
150µl	10% APS solution
18µl	TEMED

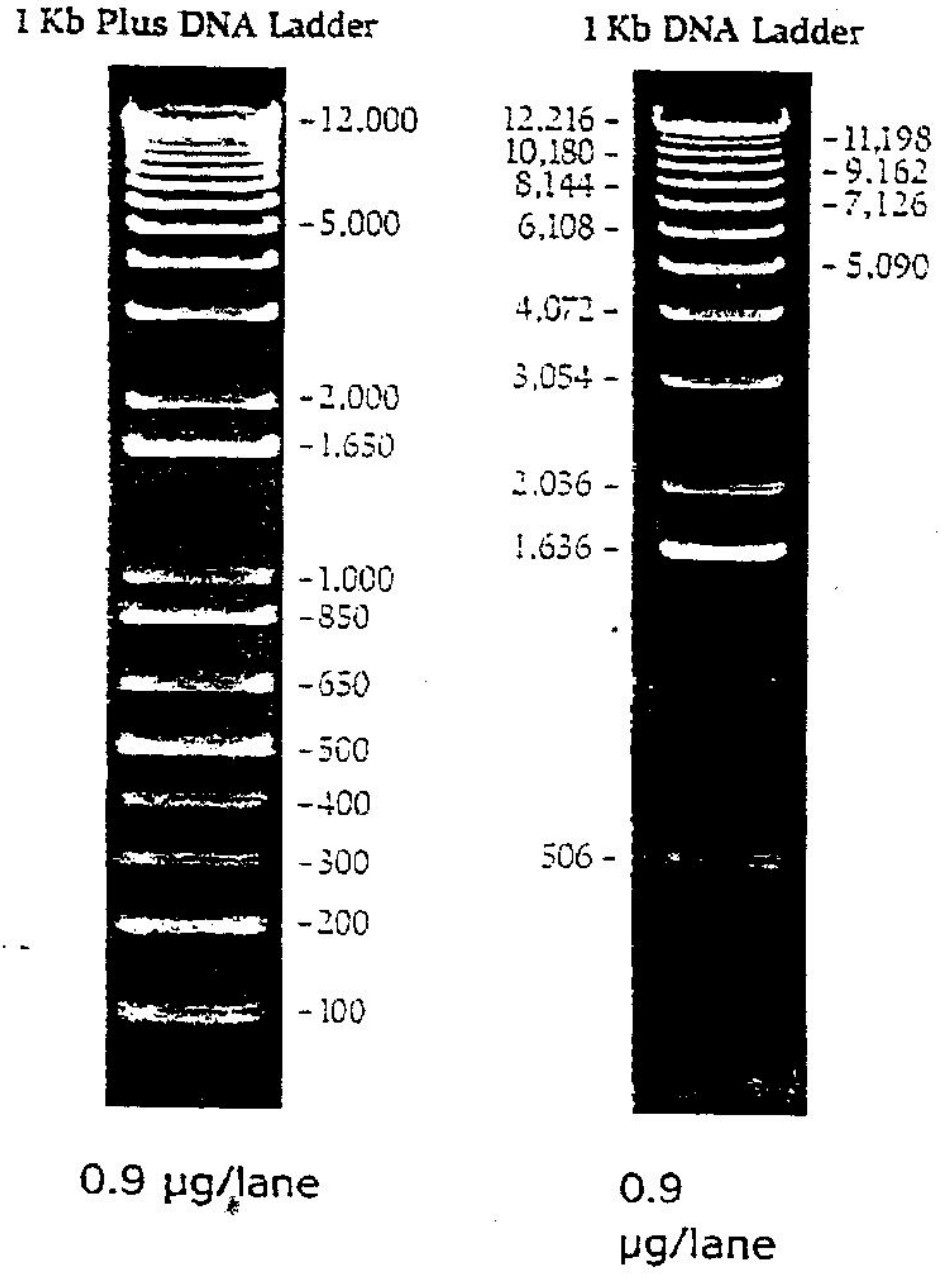
10X TBE buffer

108g	Tris
55g	Boric Acid
7.44g	EDTA

Make up to final volume of 1l with dH_2O

Autoclave

7 Appendix B: Molecular weight markers



8 Appendix C: Ethics approval for the Western Cape study

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Division of the Deputy Registrar (Research)

COMMITTEE FOR RESEARCH ON HUMAN SUBJECTS (MEDICAL)
Ref: R14/49 Tshabalala

CLEARANCE CERTIFICATE **PROTOCOL NUMBER** M00/08/26

PROJECT Assessment of Candidate Genes For Fetal Alcohol Syndrome

INVESTIGATORS Mr ME Tshabalala

DEPARTMENT Dept Human Genetics, SAIMR

DATE CONSIDERED 00/08/25

DECISION OF THE COMMITTEE *

Approved unconditionally

DATE 00/08/28 **CHAIRMAN** .....(Professor P E Cleaton-Jones)

* Guidelines for written "informed consent" attached where applicable.

c c Supervisor: Prof M Ramsay
Dept of Dept of Human Genetics, SAIMR

Works2\tsm0015\HumEth97.wdb\M 00/08/26

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10001, 10th Floor, Senate House, University.

I/we fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee.

DATE 7/9/00 SIGNATURE M Ramsay

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

9 Appendix D: Ethics approval for the Northern Cape study

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Division of the Deputy Registrar (Research)

COMMITTEE FOR RESEARCH ON HUMAN SUBJECTS (MEDICAL)
Ref: R14/49 Naidoo

CLEARANCE CERTIFICATE **PROTOCOL NUMBER** M03-10-28

PROJECT The Roe of Alcohol Dehydrogenase 2 (ADH2) and Alcohol Dehydrogenase 4 (ADH4) Genes in the Development of Fetal Alcohol Syndrome (FAS) in the South African Coloured Population

INVESTIGATORS D Naidoo

DEPARTMENT School of Pathology, NHLS

DATE CONSIDERED 03-10-31

DECISION OF THE COMMITTEE Approved unconditionally

Unless otherwise specified the ethical clearance is valid for 5 years but may be renewed upon application
This ethical clearance will expire on 1 January 2008.

DATE 03-11-28 CHAIRMAN.....(Professor P E Cleaton-Jones)

* Guidelines for written "informed consent" attached where applicable.

c c Supervisor: Prof M Ramsay
Dept of School of Pathology, NHLS
Works2\lain0015\HumEth97.wdb\M 03-10-28

==========

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and ONE COPY returned to the Secretary at Room 10001, 10th Floor, Senate House, University.

I/we fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. I agree to a completion of a yearly progress form. I/we agree to inform the Committee once the study is completed.

DATE 27/11/2003 SIGNATURE Naidoo

PLEASE QUOTE THE PROTOCOL NO IN ALL QUERIES .: M 03-10-28
PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

10 Appendix E: Complete genotype dataset for the Western and Northern Cape

ID	Status	ADH1B	ADH1C	ADH4.4	ADH4.8	prov
NPF 001	FAS	1/1	1/2	T/T	A/C	NC
NPF 002	M	1/1	1/2	T/T	C/C	NC
NPF 009	M	1/1	1/1	T/T	C/C	NC
NPF 010	FAS	1/1	1/1	T/T	A/C	NC
NPF 012	M	1/1	1/1	T/T		NC
NPF 013	FAS	1/1				NC
NPF 015	M	1/1	1/2	T/T	A/C	NC
NPF 017	FAS	1/1	1/1	T/T	A/C	NC
NPF 018	FAS	1/1	1/2	T/T	A/C	NC
NPF 019	M	1/1	1/1	T/T	C/C	NC
NPF 020	M	1/1	1/2	G/T	A/A	NC
NPF 022	FAS	1/1	1/1	G/T	A/C	NC
NPF 024	FAS	1/1	1/1	T/T	A/C	NC
NPF 025	M	1/1	1/1	T/T	A/C	NC
NPF 026	FAS	1/1	1/1	T/T	A/C	NC
NPF 027	M	1/1	1/1	T/T	A/A	NC
NPF 028	FAS	1/3	1/1	T/T	A/C	NC
NPF 029	M	1/3	1/2		A/A	NC
NPF 031	FAS	1/1	1/1	T/T	A/C	NC
NPF 032	M	1/1	1/1	T/T	A/C	NC
NPF 033	M	1/1	1/1	T/T	A/C	NC
NPF 034	FAS	1/1	1/1	T/T	A/C	NC
NPF 035	M	1/1	1/1	G/T	A/C	NC
NPF 036	FAS	1/1	1/1	G/T	C/C	NC
NPF 038	FAS	1/1	1/1	G/T	A/C	NC
NPF 039	M	1/1	1/1	T/T	A/A	NC
NPF 040	M	1/1	1/1	T/T	C/C	NC
NPF 042	FAS	1/1	1/1	T/T	C/C	NC
NPF 043	M	1/1	1/1			NC
NPF 044	FAS	1/1	1/1	G/T	A/A	NC
NPF 045	FAS	1/1	2/2		C/C	NC
NPF 046	M	1/1	1/2		A/C	NC
NPF 047	FAS	1/1	1/2	T/T	C/C	NC
NPF 048	M	1/1	1/1	T/T	C/C	NC
NPF 049	FAS	1/3	1/1	T/T	A/C	NC
NPF 050	M	1/3	1/1	T/T	A/C	NC
NPF 052	M	1/1	1/1	T/T	A/C	NC
NPF 053	FAS	1/1	1/1	T/T	A/A	NC
NPF 054	FAS	1/3	1/1	G/T	A/C	NC
NPF 055	M	1/1	1/1	G/G	A/C	NC
NPF 057	M	1/1	1/1	G/T	A/C	NC
NPF 059	FAS	1/1	1/1	G/G	A/C	NC
NPF 060	FAS	1/3	1/1	T/T	A/C	NC
NPF 061	M	1/3	1/1	T/T	A/C	NC
NPF 062	FAS	1/1	1/1	T/T	C/C	NC

NPF 063	M	1/1	1/1	G/T	A/C	NC
NPF 064	M	1/1	1/1	G/T	A/C	NC
NPF 065	FAS	1/1	2/2	T/T	C/C	NC
NPF 066	FAS	1/1	1/1	T/T	A/C	NC
NPF 067	M	1/1	1/2	T/T	A/C	NC
NPF 069	FAS	1/1	1/1	T/T	A/C	NC
NPF 070	M	1/1	1/1	T/T	A/C	NC
NPF 071	M	1/1	1/1	T/T	C/C	NC
NPF 072	FAS	1/1	1/1	T/T	A/C	NC
NPF 074	M	1/1	1/1	T/T	A/C	NC
NPF 078	FAS	1/1	1/1	T/T	A/C	NC
NPF 080	M	1/2	1/1	T/T	A/C	NC
NPF 082	FAS	1/1	1/1	T/T	A/C	NC
NPF 083	FAS	1/1	1/1	T/T	C/C	NC
NPF 085	FAS	1/1	1/2	G/T	A/C	NC
NPF 086	FAS	1/1	1/1	T/T	A/A	NC
NPF 089	FAS	1/1	1/1	T/T	A/C	NC
NPF 090	M	1/1	1/2	T/T	A/C	NC
NPF 091	FAS	1/1	1/2	T/T	C/C	NC
NPF 092	M	1/1	1/2	T/T	C/C	NC
NPF 093	FAS	1/1	1/1	T/T	C/C	NC
NPF 094	FAS	1/1	1/1	G/T	A/C	NC
NPF 095	FAS	1/1	2/2	T/T	A/A	NC
NPF 096	M	1/1	1/2	T/T	A/C	NC
NPF 098	FAS	1/1	1/2	T/T	A/C	NC
NPF 099	FAS	1/1	1/2	T/T	A/C	NC
NPF 102	FAS	1/1	1/1	T/T	A/A	NC
NPF 104	FAS	1/1	1/1		A/C	NC
NPF 105	M	1/1	1/1	T/T	A/C	NC
NPF 106	FAS	1/1	1/2	T/T	A/A	NC
NPF 107	M	1/1	1/2	T/T	A/C	NC
NPF 108	FAS	1/1	1/2	T/T	A/C	NC
NPF 109	M	1/1	1/2	G/T	A/A	NC
NPF 110	FAS	1/1	1/2	G/T	A/A	NC
NPF 111	M	1/1	1/1	G/T	A/C	NC
NPF 112	FAS	1/1	1/1	T/T	A/C	NC
NPC 001	control	1/1	1/2	T/T	C/C	NC
NPC 003	control	1/1	2/2	G/T	A/A	NC
NPC 004	control	1/2	1/1	G/T	A/C	NC
NPC 005	control	1/1	1/1		C/C	NC
NPC 006	control	1/1	1/2	G/T	A/C	NC
NPC 007	control	1/1	1/1	G/T	A/C	NC
NPC 008	control	1/1	1/1	T/T	C/C	NC
NPC 009	control	1/1	1/2	T/T	A/C	NC
NPC 010	control	1/1	1/2	T/T	C/C	NC
NPC 011	control	1/1	1/2		A/C	NC
NPC 012	control	1/3	1/1		A/C	NC
NPC 013	control	1/1	1/2	T/T	A/C	NC

NPC 014	control	1/1	1/2	T/T	C/C	NC
NPC 015	control	1/1	1/2	T/T	A/A	NC
NPC 017	control	1/2	1/1	T/T	C/C	NC
NPC 018	control	1/1	1/1	T/T	A/A	NC
NPC 019	control	1/1	1/2	T/T	A/C	NC
NPC 021	control	1/1	1/1	T/T	C/C	NC
NPC 022	control	1/1	1/1	G/T	A/A	NC
NPC 023	control	1/1	1/2	T/T	C/C	NC
NPC 024	control	1/1	1/1	G/T	A/A	NC
NPC 025	control	1/1	1/2	G/T	A/C	NC
NPC 026	control	1/1	1/1	T/T	C/C	NC
NPC 027	control	1/1	1/1	T/T	A/C	NC
NPC 028	control	1/1	1/1	T/T	A/C	NC
NPC 029	control	1/1	1/2	T/T	C/C	NC
NPC 030	control	1/1	1/2	T/T	C/C	NC
NPC 031	control	1/2	1/1	T/T	A/C	NC
NPC 032	control	1/1	1/1	T/T	A/C	NC
NPC 033	control	1/1	1/1	T/T	A/C	NC
NPC 034	control	1/1	1/1	T/T	C/C	NC
NPC 035	control	1/1	1/1	T/T	C/C	NC
NPC 036	control	1/1	1/2	T/T	A/C	NC
NPC 037	control	1/1	2/2	G/G	A/C	NC
NPC 038	control	1/1	1/1	T/T	A/C	NC
NPC 039	control	1/1	1/2	T/T	A/C	NC
NPC 040	control	1/1	1/1	T/T	A/C	NC
NPC 041	control	1/2	1/2	T/T	C/C	NC
NPC 042	control	1/1	1/2	T/T	A/C	NC
NPC 043	control		1/2	T/T	A/A	NC
NPC 044	control	1/1	1/2	G/T	A/C	NC
NPC 045	control	1/1	2/2	T/T	A/A	NC
NPC 046	control	1/1	1/2	T/T	A/A	NC
NPC 048	control	1/1	1/1	T/T	C/C	NC
NPC 049	control	1/1	1/1	T/T	A/A	NC
NPC 050	control	1/1	1/2	T/T	A/C	NC
NPC 051	control	1/1	1/1	G/T	A/A	NC
NPC 052	control	1/1	1/1	T/T	A/C	NC
NPC 053	control	1/1	1/1	T/T	A/A	NC
NPC 054	control	1/1	1/2	G/T	A/C	NC
NPC 055	control	1/1	1/2	T/T	A/C	NC
NPC 056	control	1/1	1/2	T/T	A/C	NC
NPC 057	control	1/1	1/1	T/T	C/C	NC
NPC 058	control	1/1	1/2	T/T	C/C	NC
NPC 059	control	1/1	1/2	G/T	A/C	NC
NPC 060	control	1/1	1/1	T/T	A/C	NC
NPC 061	control	1/1	1/1	G/T	A/C	NC
NPC 062	control	1/1	2/2	G/T	C/C	NC
NPC 063	control	1/2	1/1	T/T	A/C	NC
NPC 064	control	1/1	1/1	T/T	C/C	NC

NPC 065	control	1/1	1/1	T/T	C/C	NC
NPC 066	control	1/3	1/1	T/T	A/A	NC
NPC 067	control	1/1	1/1	G/T	A/C	NC
NPC 068	control	1/1	1/2		A/C	NC
NPC 069	control	1/1	1/2	T/T	C/C	NC
NPC 070	control	1/1	1/1	T/T	A/C	NC
NPC 071	control	1/1	1/2	T/T	A/C	NC
NPC 072	control	1/1	1/1	T/T	A/A	NC
NPC 073	control	1/1	1/1	T/T	A/C	NC
NPC 074	control	1/1	1/2		A/A	NC
NPC 075	control	1/1	1/2	T/T	C/C	NC
NPC 077	control	1/1	1/1	T/T	A/A	NC
NPC 079	control	1/1	1/1	G/T	A/A	NC
NPC 080	control	1/1	1/1	T/T	A/C	NC
NPC 081	control	1/3	1/2	G/T	A/A	NC
NPC 082	control	1/1	1/1	T/T	C/C	NC
NPC 083	control	1/1	1/1	T/T	C/C	NC
NPC 084	control	1/1	1/1		A/C	NC
NPC 085	control	1/1	1/1	T/T	A/C	NC
NPC 086	control	1/1	1/2	G/T	A/C	NC
NPC 087	control	1/3	1/1	T/T	A/C	NC
NPC 088	control	1/1	1/2	G/T	A/A	NC
NPC 089	control	1/1	1/1	T/T	A/C	NC
NPC 090	control	1/1	1/1	T/T	A/C	NC
NPC 091	control	1/1	1/2	G/T	A/A	NC
NPC 092	control	1/1	1/1	T/T	A/C	NC
NPC 093	control	1/1	1/1	T/T	A/C	NC
NPC 094	control	1/1	1/2	G/T	C/C	NC
NPC 095	control	1/1	1/1	G/T	A/A	NC
NPC 096	control	1/1	1/2	T/T	A/C	NC
NPC 097	control	1/1	1/1	T/T	A/C	NC
NPC 098	control	1/1	1/1	T/T	C/C	NC
NPC 099	control	1/1	1/1	G/T	A/C	NC
NPC 100	control	1/1	1/1	T/T	A/A	NC
NPC 101	control	1/1	1/1	G/T	A/A	NC
NPC 102	control	1/1	1/1	G/T	A/C	NC
NPC 103	control	1/1	1/1	G/T	A/C	NC
NPC 104	control	1/1	1/1	T/T	A/C	NC
NPC 105	control	1/1	1/2	T/T	A/C	NC
NPC 106	control	1/1	1/1	T/T	A/A	NC
NPC 107	control	1/1	1/1	T/T	A/C	NC
NPC 108	control	1/1	2/2	G/T	A/C	NC
NPC 110	control	1/1	1/2	G/G	A/A	NC
NPC 111	control	1/1	1/2		A/C	NC
NPC 112	control	1/1	1/1	T/T	A/C	NC
NPC 113	control	1/1	1/1	T/T	C/C	NC
NPC 114	control	1/1	1/1	T/T	A/A	NC
NPC 115	control	1/1	1/1	G/T	A/A	NC

NPC 116	control	1/1	2/2	G/T	A/A	NC
NPC 117	control	1/1	1/1	T/T	A/C	NC
NPC 118	control	1/1	1/1	G/T	A/A	NC
NPC 119	control	1/1	1/2	T/T	A/A	NC
1.1	M	1/1	1/2			WC
1.2	FAS	1/3	1/1			WC
2.1	M	1/1	1/2			WC
2.2	FAS	1/2	1/1			WC
3.1	M	1/1	1/1			WC
3.2	FAS	1/1	1/1			WC
4.1	M	1/1	1/1			WC
4.2	FAS	1/1	1/1			WC
5.1	M	1/1	1/1			WC
5.2	FAS	1/1	1/2			WC
6.1	M	1/1	1/1			WC
6.2	FAS	1/1	1/1			WC
7M	M	1/1	1/2		A/A	WC
7.2C	FAS	1/1	1/1	G/T	A/A	WC
8.1M	M	1/1	1/1		C/C	WC
8C	FAS	1/3	1/1		A/C	WC
9.1	M	1/1	1/2			WC
9.2	FAS	1/1	1/2			WC
10M	M	1/3	1/1	T/T	A/A	WC
1OC	FAS	1/3	1/1			WC
12.1	M	1/1	1/1			WC
12.2	FAS	1/1	1/1			WC
13.1M	M	1/1	1/1	G/T	A/C	WC
13.2C	FAS			T/T		WC
14.1M	M	1/1	1/1	G/T	A/A	WC
14C	FAS	1/1	1/1	T/T		WC
15M	M	1/1	1/2	T/T	A/A	WC
15.2C	FAS	1/1	1/1		A/C	WC
16.1M	M	1/1	1/1	G/T		WC
16.2C	FAS	1/1	1/2	T/T	A/A	WC
17M	M	1/1	1/1	T/T		WC
17C	FAS	1/1	1/1	T/T		WC
18M	M			T/T	A/A	WC
18C	FAS	1/3	1/2	T/T	A/A	WC
19M	M	1/1	1/2		C/C	WC
19C	FAS	1/1	1/2	T/T	A/C	WC
20M	M	1/1	1/1	T/T	A/C	WC
20C	FAS	1/1	1/1	T/T		WC
23M	M	1/1	1/2	G/T		WC
23C	FAS	1/1	2/2	G/T	A/C	WC
24M	M	1/1	1/1	G/T		WC
24C	FAS	1/1	1/2	G/T	A/A	WC
25M	M	1/1	1/2		A/A	WC
25C	FAS	1/1	1/1	G/T	A/A	WC

26M	M	1/1	1/1	T/T		WC
26C	FAS	1/1	1/1	T/T		WC
31M	M	1/1	1/2	G/G		WC
31.2C	FAS	1/1	1/2	G/G	A/C	WC
32.1M	M			T/T		WC
32.2	FAS	1/2	1/2	T/T	A/C	WC
34M	M	1/1	1/1	T/T	A/C	WC
34C	FAS	1/1	1/1	T/T	A/C	WC
35M	M	1/1	1/1	T/T	A/C	WC
35C	FAS	1/1	1/1	T/T	C/C	WC
36M	M	1/1	1/2	T/T	A/C	WC
36C	FAS	1/1	1/2	G/T	A/C	WC
37M	M	1/1	1/2	T/T		WC
37C	FAS	1/1	1/1	T/T	A/A	WC
38M	M	1/1	1/1		A/A	WC
38C	FAS	1/1	1/1			WC
39.1	M	1/1	1/1	T/T		WC
39C	FAS	1/1	1/2	G/T		WC
40M	M	1/1	1/1	T/T	A/A	WC
40C	FAS	1/1	1/1			WC
41M	M	1/2	1/2			WC
41C	FAS	1/1	1/2			WC
42M	M	1/1	1/1	T/T	A/C	WC
42C	FAS	1/1	1/1	T/T	A/C	WC
43M	M	1/1	1/1	T/T		WC
43C	FAS	1/1	1/1	G/T		WC
44.1	M	1/1	1/2		C/C	WC
44.2c	FAS	1/1	1/2	T/T	A/C	WC
45.1	M	1/1	1/1			WC
45C	FAS	1/1	1/1			WC
46M	M	1/1	1/2			WC
46C	FAS	1/1	1/1	T/T		WC
47M	M	1/1	1/1	T/T		WC
47C	FAS	1/3	1/1	T/T		WC
48.1	M	1/2	1/2	T/T		WC
48C	FAS	1/1	1/1	G/T	A/A	WC
49.1M	M	1/2	1/1	T/T		WC
49.2	FAS	1/2	1/1	G/T	A/A	WC
50M	M	1/1	1/2	T/T	C/C	WC
50C	FAS	1/1	1/1	T/T	A/C	WC
51M	M	1/3	1/1		A/C	WC
51C	FAS	1/1	1/1	T/T	A/C	WC
53.1	M					WC
53.2	FAS			T/T	A/C	WC
55.1	M	1/1		T/T		WC
55.4	FAS			T/T	A/C	WC
57.1	M			T/T		WC
57.3	FAS			T/T		WC

58.1	M			T/T	A/C	WC
58.3	FAS			T/T		WC
59.1	M	1/2	1/1	T/T	A/C	WC
59.4	FAS	1/2	1/1	T/T	A/C	WC
60.1	M	1/1	1/2	T/T	A/C	WC
60.3	FAS	1/2	1/1	T/T	C/C	WC
61.1	M	1/1	1/2		C/C	WC
61.2	FAS	1/1	1/2	T/T	C/C	WC
63.3	M	1/1	1/1			WC
63.4	FAS	1/1	1/1	T/T	A/C	WC
64.1	M	1/1	1/1	G/T		WC
64.4	FAS	1/1	1/1	G/T		WC
66.1	M	1/1	1/2	T/T		WC
66.2	FAS	1/1	1/2	G/T		WC
67.1	M	1/1	1/2	G/T		WC
67.4	FAS	1/1	1/1	G/T		WC
68.1	M	1/3	1/1	T/T	A/C	WC
68.2	FAS	1/1	1/1	T/T	A/C	WC
69.1	M	1/3	1/1	G/T		WC
69.3	FAS	1/1	1/1	G/T	A/C	WC
70.1	M	1/1	1/1	T/T	A/A	WC
70.3	FAS	1/1	1/1		A/C	WC
71.1	M	1/1	1/1	G/T		WC
71.2	FAS	1/1	1/1			WC
72.1	M	1/3	1/1	T/T		WC
72.3	FAS	1/1	1/2		C/C	WC
73.1	M	1/1	1/2	G/T		WC
73.2	FAS	1/1	2/2			WC
74.1	M	1/1	1/2	T/T	A/A	WC
74.3	FAS	1/1	1/1	T/T		WC
75.1	M	1/1	2/2	T/T	A/C	WC
75.2	FAS	1/1	1/2	T/T		WC
76.1	M	1/1	1/1	T/T	A/C	WC
76.2	FAS	1/1	1/2	G/T		WC
76.3	FAS			T/T		WC
100	control	1/2	1/2	T/T	C/C	WC
101	control	1/1	1/2	G/T		WC
102	control	1/2	1/1			WC
103	control	1/1	1/2			WC
104	control	1/1	1/2			WC
105	control	1/1	1/1			WC
106	control	1/1	1/1			WC
107	control	1/1	1/1			WC
108	control	1/1	1/1			WC
109	control	1/1	1/1			WC
110	control	1/2	1/2			WC
111	control	1/1	1/2	T/T	C/C	WC
112	control	1/2	1/1	T/T		WC

113	control	1/1				WC
114	control	1/1	1/1			WC
115	control	1/1	1/1			WC
116	control	1/3	1/1	T/T		WC
117	control	1/1	1/2	G/T		WC
118	control	1/1	1/1	G/T		WC
119	control	1/1	1/1			WC
120	control	1/2	1/2	T/T		WC
121	control	1/1	1/1	T/T	A/C	WC
122	control	1/2	1/2	T/T		WC
123	control	1/1	1/1	G/T		WC
124	control	1/1	1/1	T/T	A/A	WC
125	control	1/1	1/1	T/T	A/C	WC
126	control	1/1	1/1			WC
127	control	1/1	2/2	T/T	A/C	WC
128	control	1/1	1/2	T/T	C/C	WC
129	control	1/1	1/1	T/T		WC
130	control	1/2	1/1	T/T	A/C	WC
131	control	1/1	1/1	T/T	A/A	WC
132	control	1/1	1/2	G/T	A/A	WC
133	control	1/1	1/1	T/T		WC
134	control	1/1	1/1	T/T	A/C	WC
135	control	1/1	1/1	G/T	C/C	WC
136	control	1/1	1/2	G/T	A/C	WC
137	control	1/1	1/2	T/T	A/C	WC
138	control	1/2	1/1	T/T	A/C	WC
139	control	1/1	1/2	T/T	C/C	WC
140	control	1/1	1/1	T/T	A/C	WC
141	control	1/1	1/1	G/T	A/C	WC
142	control	1/3	1/1	T/T	A/C	WC
143	control	1/1	1/1	T/T	A/C	WC
144	control	1/1	1/1	T/T	A/A	WC
145	control	1/1	1/1	T/T	C/C	WC
146	control	1/1	1/2	T/T	C/C	WC
147	control	1/1	1/2	G/T	A/C	WC
148	control	1/1	1/2	T/T	A/A	WC
149	control	1/1	1/1	G/T	A/C	WC
150	control	1/3	1/1	T/T	A/A	WC
151	control	1/1	1/2	T/T		WC
152	control	1/1	1/1	T/T	A/C	WC
153	control	1/1	1/1	T/T	A/A	WC
154	control	1/2	1/1	T/T	C/C	WC
155	control	1/1	1/1	T/T	C/C	WC
156	control	1/1	1/1	T/T	A/C	WC
157	control	1/1	1/2	T/T	A/A	WC
158	control	1/2	1/1	G/T	A/C	WC
159	control	1/1	1/2	G/T	A/C	WC
160	control	1/1	1/2	G/T	A/C	WC

161	control	1/1	1/2	G/T	A/C	WC
162	control	1/1	1/2	T/T	A/C	WC
163	control	1/1	1/1	T/T	C/C	WC
164	control	1/1	1/1	T/T	C/C	WC
165	control			G/T	A/C	WC
167	control			T/T	A/C	WC
168	control			T/T	A/C	WC
169	control			T/T	C/C	WC
200	control	1/2	1/1			WC
201	control	1/1	1/1			WC
202	control	1/1	1/1			WC
203	control	1/1	1/2	T/T	C/C	WC
204	control	1/2	1/2		C/C	WC
205	control	1/1	2/2	T/T	C/C	WC
206	control	1/1	1/1	T/T	C/C	WC
207	control	1/1	1/1	T/T	A/A	WC
208	control	1/2	1/1	T/T	C/C	WC
209	control	1/1	1/2	T/T	A/C	WC
210	control	1/2	1/2	T/T	A/C	WC
211	control	1/1	1/2	G/T	A/A	WC
212	control	1/1	1/2	T/T	C/C	WC
213	control	1/1	1/2	T/T	A/C	WC
214	control	1/1	1/2	T/T	A/C	WC
215	control	1/1	1/2	G/T	A/C	WC
216	control	1/1	1/1	T/T	C/C	WC
217	control	1/1	1/2	T/T	A/C	WC
218	control	1/1	1/1			WC
219	control	1/3	1/1			WC
220	control	1/1	1/2	G/T	A/C	WC
221	control	1/1	1/1	T/T	C/C	WC
222	control	1/2	1/1	T/T	A/C	WC
223	control	1/1	1/1	G/T	A/C	WC
224	control	1/1	1/2	T/T	A/A	WC
225	control	1/2	1/2	T/T	C/C	WC
226	control	1/1	1/1	T/T	C/C	WC
227	control	1/2	1/2	T/T	C/C	WC
228	control	2/3	1/1	T/T	A/C	WC
229	control	1/3	1/1	T/T	A/C	WC
230	control	1/1	1/1	T/T	A/C	WC
231	control	1/3	1/1	T/T	A/C	WC
232	control	1/1	1/1			WC
233	control	1/1	1/2	T/T	C/C	WC
234	control	3/3	1/1	T/T	A/A	WC
235	control	1/1	1/2	G/T	A/C	WC
236	control	1/1	1/2			WC
237	control	1/2	1/1	G/T	C/C	WC
238	control	1/1	1/1	T/T	C/C	WC
239	control	1/1	1/2	G/T	A/C	WC

240	control	1/1	1/2	G/T	A/C	WC
241	control	1/1	1/2	G/T	A/C	WC
242	control	1/1	1/1	G/T	A/A	WC
243	control	2/2	1/1	T/T	C/C	WC
244	control	1/1	1/1	T/T	C/C	WC
245	control	1/1	1/1	G/T	A/C	WC
246	control	1/1	1/1	G/T	A/C	WC
247	control	1/1	1/1	T/T	C/C	WC
248	control	1/1	1/2	T/T	C/C	WC
249	control	1/2	1/2	T/T	A/C	WC
250	control	1/1	1/1	T/T	A/C	WC
251	control	1/1	1/1	T/T	C/C	WC
252	control	1/1	1/2	T/T	C/C	WC
253	control	1/1	1/2	G/T	A/C	WC
254	control	1/1	1/2	T/T	C/C	WC
255	control	1/1	1/2	T/T	A/C	WC
256	control	1/1	1/2	T/T	C/C	WC
257	control	1/1	1/2	T/T	C/C	WC
258	control	1/2	1/2			WC
259	control	1/1	1/2	T/T	C/C	WC
260	control	1/1	1/1	T/T	C/C	WC
261	control	1/1	1/1	T/T	C/C	WC
262	control	1/1	1/2	G/T	A/C	WC
263	control	1/2	1/2	T/T	C/C	WC
264	control	2/2	1/1	T/T	C/C	WC
265	control	1/3	1/1	T/T	A/C	WC
266	control	1/1	1/1	T/T	A/C	WC
267	control	1/2	1/1	T/T		WC
268	control	1/2	1/1	T/T	C/C	WC
269	control	1/1	1/1	G/T		WC
270	control	2/3	1/2	T/T	C/C	WC
271	control	1/1	1/2	T/T	C/C	WC
272	control	1/1	1/1	G/T	A/C	WC
273	control	1/1	1/1	G/T	A/C	WC
274	control	1/2	1/2	T/T	A/A	WC
275	control	1/1	1/1	T/T	A/C	WC
276	control	1/1	1/1	T/T	A/C	WC
277	control	1/3	1/1	T/T	A/A	WC
278	control	1/1	1/1	T/T	C/C	WC
279	control	1/2	1/1	T/T	A/C	WC
280	control	1/1	1/2	T/T	C/C	WC
281	control	1/1	1/2	G/T	A/C	WC
282	control	1/1	1/1	T/T	A/A	WC
283	control	1/1	1/2	T/T	A/C	WC
284	control	1/1	1/1		A/C	WC
285	control	1/1	1/1	T/T	C/C	WC
286	control	1/2	1/1	T/T	A/C	WC
287	control	1/1	1/1	T/T	C/C	WC

288	control	1/1	1/1	T/T	A/C	WC
289	control	1/2	1/1	T/T	C/C	WC
290	control	1/1	1/1	G/T	A/C	WC
291	control	1/1	1/2	G/T	A/A	WC
292	control	1/1	1/2	T/T	A/C	WC
293	control	1/2	1/1	T/T	A/C	WC
294	control	1/1	1/1	G/T	A/A	WC
295	control	1/1	1/2	T/T	A/A	WC
296	control	1/1	1/2	G/T	A/A	WC
297	control	1/1	1/2	G/T	A/C	WC
298	control	1/1	1/2	T/T	A/A	WC
299	control	1/1	2/2	T/T	C/C	WC
300	control	1/2	1/2	T/T	C/C	WC
301	control	1/1	1/2	G/G	A/A	WC
302	control	1/1	1/2	G/T	A/C	WC
303	control	1/2	1/1	T/T	C/C	WC
304	control	1/1	1/1	T/T	A/C	WC
305	control	1/2	1/1	T/T	C/C	WC
306	control	1/2	1/1	T/T	A/A	WC
307	control	1/3	1/2	T/T	A/A	WC
308	control	2/3	1/1	T/T	C/C	WC
309	control	1/1	1/2	T/T	A/C	WC
310	control	1/1	1/1	T/T	C/C	WC
311	control	1/1	1/1	T/T	A/C	WC
312	control	1/1	1/2	G/T	A/C	WC
313	control	1/1	1/1	T/T	C/C	WC
315	control			T/T	A/C	WC
316	control			T/T	C/C	WC
317	control			G/T	A/C	WC
318	control			T/T	A/C	WC
319	control			T/T	A/A	WC
320	control			T/T	A/C	WC
321	control			T/T	A/C	WC

Key : C = child; M = mother; WC =Western Cape; NC = Northern Cape