Allelic diversity of selected human neurotransmitter genes in South African ethnic groups

Clement Malan

Dissertation submitted in accordance with the requirements for the degree *Magister Scientiae* (Genetics) in the Faculty of Natural and Agricultural Sciences, Department of Genetics at the University of the Free State

June 2014

Supervisor: **S.-R. Schneider**

Co-Supervisor: **Prof. J.J. Spies**

DECLARATION

I certify that the dissertation hereby submitted by me (Clement Malan) for the M.Sc. degree at the University of the Free State is my independent effort and had not previously been submitted for a degree at another university or faculty. I also waive copyright of the dissertation in favour of the University of the Free State.

Clement Malan

Date

ACKNOWLEDGEMENTS

The writing of this dissertation and its associated laboratory work has been the greatest academic challenge I have faced. Without the help of my study leaders and the support from my friends and family this study would not have been completed.

- I would like to thank Mrs. S.-R. Schneider for her guidance, patience and providing me an amicable atmosphere for doing research. I'm very grateful to you and your husband, Mr. J. Schneider for taking me to the Eastern and Northern Cape for sampling purposes and for everything that entailed our trips.
- I would like to thank Professor J.J. Spies for providing me the opportunity to study in the human genetics field and assisting me in the literature writing process. I'm fortunate to have you involved in my study and grateful for your general guidance over the project.
- I would like to offer my thanks to everyone that assisted me at the Genetics department of the University of the Free State. Mrs. H van der Westhuizen for assisting me with sequencing. Professor J.P. Grobler and Mrs. P. Spies for assisting me with genetic software.
- I would like to express my gratitude to all the individuals that participated in this research project by providing saliva samples for DNA extraction.
- I would like to thank the National Research Foundation (NRF) and the German Academic Exchange Service (DAAD) for their financial contributions.
- I would like to thank all the friends I've made at the Genetics department for their friendship and support during the course of the study.
- I would like to thank my parents (Ernest Malan and Christine Malan) and my sister (Sintiché Naude) for their financial support and encouragement.
- Finally, I would like to honour my Heavenly Father for giving me the ability and perseverance to complete this study.

CONTENTS

LIST OF ABBREVIATIONS AND ACRONYMS

2n	Number of alleles
5-HTT	*Serotonin Transporter*
5-HTTLPR	Serotonin transporter-linked polymorphic region
AD	Alcohol dependence
ADH1B	*Alcohol Dehydrogenase*
ADHD	Attention deficit hyperactivity disorder
ADR	Adverse drug reactions
AIDS	Acquired immunodeficiency syndrome
APS	Ammonium persulfate
ALDH2	*Aldehyde Dehydrogenase*
ALFRED	ALlele FREquency Database
COMT	*Catechol-O-Methyl Transferase*
CNS	Central nervous system
CYP4502E1	*Cytochrome P-4502E1*
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DRD2	*Dopamine Receptor D2*
DRD3	*Dopamine Receptor D3*
DRD4	*Dopamine Receptor D4*
FAS	Foetal alcohol syndrome
FOXP2	*Forkhead Box P2*

F_{ST}	Genetic distances
GRB2	Growth Factor Receptor-Bound Protein 2
GC	Guanine and cytosine
HWE	Hardy-Weinberg equilibrium
IQ	Intelligence quotients
Kb	Kilobase
kDa	Kilo Daltons
Km	Kilometres
L	Long
MAO	Monoamine Oxidase
MAO-A	*Monoamine Oxidase A*
mg	Milligram
MHPG	3-methoxy-4-hydroxyphenylglycol
ml	Millilitre
mM	Millimolar
mRNA	Messenger ribonucleic acid
NCK	Non-catalytic region of the Tyrosine Kinase Adaptor Protein
NET	*Norepinephrine Transporter*
ng	Nanogram
OCD	Obsessive compulsive disorder
PADPRP	*Poly (ADP-ribose) Polymerase*
PAGE	Polyacrylamide gels

PCA	Principal component analysis
PCR	Polymerase chain reactions
PFC	Prefrontal cortex
PRL	Prolactin inhibitor
R^2	Coefficient of determination
RCF	Relative centrifugal force
rpm	Revolutions per minute
S	Short
SADF	South African Defence Force
SASHG	Southern African Society of Human Genetics
SDF-1	*Stromal Cell-Derived Factor-1*
SH3	Src homology 3
SIDS	Sudden infant death syndrome
SLC6A4	*Solute Carrier Family 6 Member 4*
SNP	Single nucleotide polymorphism
SSCP	Single strand conformation polymorphism
STR	Short tandem repeat
TAE	Tris/acetic acid/EDTA
TBE	Tris/Borate/EDTA
TDT	Transmission disequilibrium test
TEMED	Tetramethylethylenediamine
uVNTR	Upstream VNTR

VNTR	Variable number tandem repeat
VL	Very long
XL	Extra-long
µl	Microlitres
°C	Degree Celsius

LIST OF FIGURES

travelled and the allele frequencies of the 5-HTTLPR and *DRD4* VNTR. Allele frequency was chosen as the dependent factor and distance migrated the independent factor. A) A positive correlation (unadjusted R^2=0.569) was identified for the S allele of the 5-HTTLPR and migration distance. B) No correlation (unadjusted R^2=0.002) was identified for the 2-repeat allele of *DRD4* VNTR and migration distance. C), D) and E) Positive correlations (unadjusted R^2=0.534, 0.517 and 0.534) were identified for the 7-repeat, >5-repeat and combined 2- and >5-repeat allele frequencies.

LIST OF TABLES

CHAPTER ONE

MOTIVATION FOR STUDY

Introduction

Genetic analyses of population groups world-wide have benefited behaviour, population and medical fields. However, genetic analyses and information on South African populations are limited, in particular information on neurotransmitter genes. Limited information exist pertaining to behavioural and neurotransmitter gene associations for South African populations. Neuropsychiatric diseases and behavioural disorders are some of the greatest causes of disability-adjusted life years in South Africa. Lack of research on South African populations has contributed to adverse drug reactions (ADR) and drug inefficacy. The current literature motivates the research of South African populations in the neurotransmitter genetic field. In the motivation, information affirming the lack of research is conversed. The benefits of research and the adverse effects resulting from a lack of research are elaborated on. Possible benefits of research on certain South African ethnic groups with regards to neurotransmitter genes are stated. The garnered information prompts an impetus of research analyses in the genetic field pertaining to neurotransmission. Information on referencing style and nomenclature used in the current study concludes the motivation.

Motivation

Limited information is available on the genetic variation and basis of disease in African populations due to populations not being genetically analysed. The inference of underlying genetic components of diseases could improve diagnostic and treatment strategies for African populations. The pharmacology field can benefit from studies on genetic diversity due to the genetic involvement in drug metabolism (Reed & Tishkoff, 2006; Campbell & Tishkoff, 2008; Wright *et al*., 2011).

A lack of research with regards to psychiatric genetics exists for African and South

African populations (Wright *et al*., 2011). Compounding the problem, revised burden of disease estimates rank neuropsychiatric disorders as some of the most prominent causes of disability-adjusted life years in South Africa (Norman *et al*., 2006). Alcoholism, obsessive compulsive disorder (OCD), panic disorder, schizophrenia, bipolar and unipolar disorder rank among the 20 top causes of disability-adjusted life years in South Africa (Norman *et al*., 2006). These mentioned disorders are associated with the *Dopamine Receptor D4* (*DRD4*) variable number tandem repeat (VNTR), *Monoamine Oxidase A* (*MAO-A*) upstream VNTR (uVNTR) and the serotonin transporter-linked polymorphic region (5-HTTLPR) of the *Serotonin Transporter* (*5-HTT*) gene.

Foetal alcohol syndrome (FAS) brought on by alcoholism is a disorder problematic in South Africa. The Western Cape community of South Africa has a FAS prevalence of 39.2 cases per 1 000 individuals for the six and seven year old age groups (May *et al*., 2000). In Wellington in the Western Cape the prevalence of FAS was determined 68-89.2 cases per 1 000 individuals in a primary school cohort (May *et al*., 2007). In comparison, among African Americans the prevalence was determined 2.3 cases per 1 000 births (Abel, 1995).

Foetal alcohol syndrome is indirectly linked to the 5-HTTLPR due to its association with alcohol dependence (AD) (Feinn *et al*., 2005; Wang *et al*., 2011). The 5-HTTLPR is situated approximately 1.4 kilobase (kb) upstream from the *5-HTT* transcription site and consists of various lengths of repetitive guanine and cytosine (GC) rich sequences (Ramamoorthy *et al*., 1993). A number of behavioural disorders including autism, depression and schizophrenia are linked to neurotransmitter genes and polymorphisms such as the *5-HTT* and 5-HTTLPR (Lung *et al*., 2002; Cohen *et al*., 2003; Feinn *et al*., 2005; Lasky-Su *et al*., 2005; Wang *et al*., 2011) indicating the necessity of further analysing such genes.

The pharmacological treatment of disorders in South Africa is inadequate and limited with regards to ADRs and drug inefficacy. The ADR rate and especially the fatal ADR rate frequencies are much higher in South Africa than in the United States of America (0.140 versus (vs) 0.067 and 0.015 vs 0.003) (Lazarou *et al*., 1998; Wright *et al*., 2011). The analysis of genes involved in the treatment process and the redesigning of more efficient medicine could reduce ADRs.

Inadequate information is available on allelic frequencies with regards to neuropsychiatric disorders such as schizophrenia, OCD and depression (Wright *et al*., 2011). Research on neurotransmitter genes such as the *DRD4*, *MAO-A* and *5-HTT* could provide valuable information required to make informed decisions with regards to drug treatment among South African populations (Gerretsen & Pollock, 2008; Wimbiscus *et al*., 2010). Genetic variation between and among ethnic groups has significant effects and present opportunities to develop better strategies against disease (Gu *et al*., 1999; Grimaldi *et al*., 2010).

A study on lung cancer patients from different ethnic groups making use of the *Poly (ADP-ribose) Polymerase* (*PADPRP*) was performed (Gu *et al*., 1999). Lung cancer patients totalling 288 from the American Caucasian, American Mexican and African American population groups were analysed. Results revealed significant allelic variation among the different ethnic groups. The susceptible genotypes (Aa or aa) were associated with cancer in the Mexican American population but not in the African American or American Caucasian populations. The susceptible genotypes increased the risk for large cell carcinoma (10.79-fold) and adenocarcinoma (3.21-fold) in American Mexicans.

In another study, the allele frequency of a G-to-A mutation in the *Stromal Cell-Derived Factor-1* (*SDF-1*) gene was determined among different Brazilian ethnic groups (Grimaldi *et al*., 2010). The particular mutation associated with acquired immunodeficiency syndrome (AIDS) progression was determined varying among populations. The analysed ethnic groups were from Joinville (German ancestry), Tiriyó tribe (Asian ancestry), Waiampi Amerindian tribe (Asian ancestry) and Salvador (Portuguese and African mixture). The *SDF1-3'* A frequencies for the populations were 0.051 for the Waiampi tribe, 0.165 for the Salvador population, 0.210 for the Joinville population and 0.237 for the Tiriyó tribe revealing significant genetic variation among the populations. The studies by Gu *et al*. (1999) and Grimaldi *et al*. (2010) highlight how genetic variation among ethnic groups can play a significant role in the diagnosis and treatment of diseases.

Similar to the studies based on the American populations, analysis of South African ethnic groups with the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR would provide considerable information with regards to genetic variation, differentiation

and indirectly to behavioural traits and disorders. Populations analysed in the current study are the Khwe, Xhosa, Sotho and Afrikaner. The Khwe population is genetically diverse and possibly associated with the first modern humans (Chen *et al*., 2000; Schlebusch & Soodyall, 2012; Schlebusch *et al*., 2012). Although the Khwe speak a Khoe-San language (de Almeida, 1965) they are significantly genetically related to Bantu populations (Chen *et al*., 2000; Schlebusch & Soodyall, 2012). Analyses of Khwe individuals will provide genetic insight into a population with an admixed genetic history. A Khwe population currently lives in Platfontein near Kimberly in the Northern Cape of South Africa and was sampled in the current study.

The Xhosa analysed in the current study are representatives of the Bantu. The Bantu speaking group constitute the majority of the South African population. The Xhosa ethnic group migrated to southern Africa approximately 2 000 years ago (Peires, 1981). At present a large population of Xhosa individuals live in the Eastern Cape Province of South Africa especially in the areas previously known as the Transkei and Ciskei, which were Xhosa homelands prior to democracy in South Africa (Davenport, 1978; Slater, 2002).

The Sotho ethnic group, like the Xhosa is a Bantu speaking population that migrated to South Africa (Davenport, 1978). Bantu migration into southern Africa took place in a series; therefore distinctions greater than linguistic differences are present among Bantu populations. Among seven South African Bantu speaking populations (Zulu, Xhosa, Southern Sotho, Tsonga/Shangaan, Tswana, Pedi and Venda) linguistic distances were determined correlating with genetic distances (Davenport, 1978; Lane *et al*., 2002).

The interrelationship of the South African Bantu speaking groups based on serogenetic deoxyribonucleic acid (DNA), autosomal DNA, autosomal short tandem repeat (STR) and Y-chromosome haplotype data were determined (Lane *et al*., 2002). Low genetic differentiation between populations and high within group differences suggests that the seven Bantu ethnic groups (Zulu, Xhosa, Southern Sotho, Tsonga/Shangaan, Tswana, Pedi and Venda) originated from a common ancestral population. The mentioned Bantu groups were not isolated for long, even though their languages diverged within the past 2 000 years (Lane *et*

al., 2002). Limited genetic information is available on Bantu populations even though they represent the majority of the South African population. The analysis of the Xhosa and Sotho populations would be a good initial genetic reference point for South African Bantu populations.

The Afrikaner population originated mostly from Dutch, French and German Caucasian settlers in southern Africa. However, according to genetic analyses African and Asian contributions to the Afrikaner genome also occurred (Greeff, 2007). The Afrikaner, according to law, was not allowed to marry individuals from a different race from 1685 till the abolishment of that law and its various amendments, however, prior to 1685 admixture occurred. Occasional admixture occurred regardless of the law (Davenport, 1978; Greeff, 2007).

Initial admixture and founder effects made the Afrikaner a popular candidate population for investigating genetic diseases (Davenport, 1978; Botha & Beighton, 1983; Tipping *et al*., 2001; Greeff, 2007). The Afrikaner population is assumed a homogenous population due to the number of generations (average 12 generations) that have passed since the arrival of the first European settlers in South Africa and due to limited intermarriage with other races (Davenport, 1978; Greeff, 2007). However, the observation of admixture levels in Afrikaners will be of interest due the fact that the law (1685) was established years after the arrival of the first settlers.

The aim of the current study is to determine the allelic frequency of the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR in the four South African ethnic groups (Khwe, Xhosa, Sotho and Afrikaner). The dissertation consists of four chapters which include the motivation, a literature review and two research chapters. Chapter two, being the literature review, presents background information on the *DRD4*, *MAO-A* and *5-HTT* (5-HTTLPR) and focuses on the association of the genes and polymorphs with disorders and traits. Information on the four ethnic groups (Khwe, Xhosa, Sotho and Afrikaner) with regards to their history is also included in the literature review. The aim of chapter three, the first research article, was to determine the genetic variation and differentiation among the four South African ethnic groups. Chapter three includes the genetic analyses of the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR regions of 349 individuals from the

mentioned ethnic groups living in South Africa. The existing levels of genetic diversity and differentiation were discussed and placed into context with world-wide populations. The aim of chapter four, the second research article, was to determine the possible associations between the *DRD4* VNTR, 5-HTTLPR and the migratory behavioural trait. Objectives included correlating the allelic frequencies with distances populations migrated and establishing whether a selective trend exists. A general conclusion of the dissertation follows the fourth chapter. A summary in both English and Afrikaans concludes the dissertation.

In the current study words comprising abbreviations and acronyms are written out fully in the first instance in text, there after abbreviated. Words forming abbreviations and acronyms are not written out fully for each chapter in the dissertation. The Harvard referencing style is used in the current study.

In the current study the term Coloured is used when referring to the Cape mixed ancestry population of South Africa following advisement from the Southern African Society of Human Genetics (SASHG). The mixed ancestry population referred to has genetic contributions from Khoe-San, European, Asian and Bantu populations (Nurse *et al*., 1985; Shell, 1994; Mountain, 2003; Quintana-Murci *et al*., 2010). The term Bantu used in scientific literature (Lane *et al*., 2002; Tishkoff & Williams, 2002; Reed & Tishkoff, 2006; Campbell & Tishkoff, 2008; Tishkoff *et al*., 2009) will be used in the current study although it is sometimes perceived as a derogatory term in South Africa. It is understood that words such as Bantu and Native are not directly insulting but may have a derogatory connotation and may be resented (van den Berghe, 1965; Booth, 1998).

The use of the term Bantu in the current study referring to black African individuals or the language that black African individuals from southern, central and some parts of western Africa speak is simply for identification and nomenclature purposes. The term Bantu simply meaning “people” was first used by Wilhelm Heinrich Immanuel Bleek (1827-1875) (Bleek, 1862). This name was fitting due to the many languages in the Bantu language group. Similarly in the current study Bantu will refer only to black African populations that are Bantu speaking in order to differentiate them from hunter gather populations such as the Khoe-San and other African language populations that are also indigenously African but not

Bantu. Khoe-San individuals make use of various Khoe-San languages and are to an extent physically distinguishable from black African Bantu populations (Schlebusch & Soodyall, 2012; Schlebusch *et al*., 2012).

CHAPTER TWO

LITERATURE REVIEW

Abstract

Genetic ethnicity studies are useful in providing information on population variation, differentiation and structure. Analyses performed with neurotransmitter genes provide valuable information on behavioural disorders and traits in addition to providing information on population variation and differentiation. Information on South African populations with regards to the allelic frequency of neurotransmitter genes are however usually inadequately based on one or two population groups analysed with only one gene. In the current chapter the available literature with regards to the neurotransmitter genes *DRD4*, *MAO-A* and *5-HTT* were reviewed. Information pertaining to the *DRD4*, *MAO-A* and *5-HTT* polymorphisms and allelic frequencies were compiled. In addition, information on the Khwe, Xhosa, Sotho and Afrikaner ethnic groups of South Africa were garnered with regards to their history and population structure. The genes and population groups were elaborated on separately and in some cases in conjunction with each other. Behavioural associations with the mentioned neurotransmitter genes were revised.

Introduction

The African continent has significant ethnic group variations and bears the potential to better understand genetic disorders. African populations, however, are largely underrepresented with regards to genetic studies (Cavalli-Sforza & Feldman, 2003; Campbell & Tishkoff, 2008; Tishkoff *et al*., 2009). There is also a lack of information on African populations in overlapping fields of genetics, such as behavioural genetics and pharmacology (Wright *et al*., 2011). South Africa, the leader in African biomedical research, contributes significant information though still limited in comparison to high research output countries (Patel & Kim, 2007).

The connections between genes, psychiatric disorders and behavioural traits have long been known, but still neglected in a South African context. A literature search

based on genes and polymorphisms that have an effect on behaviour such as the *DRD4* VNTR, *MAO-A*-uVNTR and the 5-HTTLPR yield sparse results when considering South African populations. The mentioned polymorphisms are ideal candidates for studies on behaviour due to their neurotransmission association. Neurotransmitter genes are associated with a range of disorders and traits including AD, autism, depression, novelty seeking, migratory behaviour and schizophrenia (Chen *et al*., 1999; Lung *et al*., 2002; Cohen *et al*., 2003; Feinn *et al*., 2005; Lasky-Su *et al*., 2005; Matthews & Butler, 2011; Wang *et al*., 2011).

Genetic analysis of neurotransmitter genes of a population is beneficial in providing information on allelic variation and differentiation in addition to the mentioned behavioural associations (Chang *et al*., 1996; Mansoor *et al*., 2008). World-wide populations and ethnic groups, within the same country, have been analysed in relationship to the *DRD4* VNTR and 5-HTTLPR providing significant genetic information (Chang *et al*., 1996; Mansoor *et al*., 2008; Bisso-Machado *et al*., 2013). South Africa is a prime candidate for a genetic study based on neurotransmitter genes, not only due to it being genetically neglected with regards to research but also due to its relative population size and ethnic diversity (>51 million South African individuals)[1]. Migration into southern Africa has resulted in South Africa having one of the most diverse population structures with contributions from the Nguni (Zulu, Xhosa and Ndebele populations), Sotho-Tswana (Tswana, Southern Sotho and Northern Sotho populations) and Afrikaans speaking (Caucasian and Coloured populations) communities constituting the majority of the population (Davenport, 1978).

The literature review comprises of the following segments: genetic ethnicity studies, neurotransmitter studies in South African ethnic groups, Dopamine Receptor D4, Monoamine Oxidase A, Serotonin Transporter and a summary. The genetic ethnicity study segment provides background information on genetic population studies pertaining to population analyses and structure with some emphasis on neurotransmitter genes. The Khwe, Xhosa, Sotho and Afrikaner are discussed with regards to their history and population structure. The segment: neurotransmitter studies in South African ethnic groups provide a background on

[1] Statistics South Africa Census 2011, www.Statssa.gov.za

the neurotransmitters dopamine, serotonin and the enzyme Monoamine Oxidase (MAO). Information relating to neurotransmitter studies performed on South African and world-wide populations are conversed. The Dopamine Receptor D4, Monoamine Oxidase A and Serotonin Transporter segments provide information on the mentioned neurotransmitters genetic variations and associations. The summary concludes the literature review.

Genetic ethnicity studies

The genetic analyses of racial and ethnic groups are useful for generating and exploring hypotheses regarding genetic risk factors and important medical outcomes (Burchard *et al*., 2003). Along with social, environmental and lifestyle factors genetic variations have an effect on disease among different populations (Taylor & Wright, 2005). Genetic variation and differentiation among populations are observed as populations from the same continent tend to group together (Burchard *et al*., 2003). Human population differentiation, a consequence of admixture and population stratification is influenced by mating patterns. Mating patterns in turn are influenced by cultural, geographical, social and/or religious barriers (Tang *et al*., 2005).

Studies that investigate genetic variation among populations typically begin with the sampling of “predefined” populations. The definition of a population is usually based on the culture or geography of an ethnic group or population, not necessarily on the underlying genetic relationship (Foster & Sharp, 2002). In order to infer more on a particular population’s structure in relation to other populations, genetic variant data (allelic frequencies) from populations from different geographical regions are obtained and analysed (Novembre & Stephens, 2008).

Analysed genetic data can be interpreted with a principal component analysis (PCA). When performing a PCA, an allele-frequency map, representing allele frequencies across dimensional space, is plotted. This high dimensional data when summarised with a PCA is distilled into a number of synthetic maps allowing

the visual analysis and interpretation of data (Novembre & Stephens, 2008; Wang *et al*., 2010).

A PCA performed on world-wide populations and the allele frequency of the *DRD4* VNTR, a gene implicated in neurotransmission (Chang *et al*., 1996) provided notable results (Figure 2.1). Analysis with the *DRD4* VNTR revealed genetic differentiation among geographic regions. Native South American populations (Karitiana, Ticuna and Surui) were peripherally located in the PCA resembling their peripheral geographic location. The peripheral location of Native American populations in the PCA is possibly due to bottlenecks along human migration routes from Africa to America via Europe and Asia. Populations from Europe, Africa and the Middle East grouped together in the PCA indicative of the modern human expansion pattern out of Africa (Chang *et al*., 1996; Tishkoff & Williams, 2002; Friedlaender *et al*., 2008; Tishkoff *et al*., 2009).

Ethnic groups from the same country have been genetically characterised with the *DRD4* VNTR, elucidating significant geographic and linguistic information (Ghosh & Seshardi, 2005; Mansoor *et al*., 2008). Clarification with regards to Pakistani ethnic groups was obtained with the use of the *DRD4* VNTR. Pakistan being a gateway to Southeast Asia was repeatedly invaded during its long history resulting in a complex scenario with regards to its ethnic group make-up. Invaders belonged to various population groups including Aryans, Turks and Arabs (Mansoor *et al*., 2008).

The genetic variation and differentiation of South African ethnic group can be determined with the use of neurotransmitter genes such as the *DRD4* VNTR (Mansoor *et al*., 2008). Genetic variants selected during the past several thousand years resulted in certain variants being unique to particular geographic regions (Tishkoff *et al*., 2007; Grossman *et al*., 2010; Seldin *et al*., 2011). Although genetic differences between populations represent only a small fraction of genetic variation, many simple and even complex diseases and disorders have prevalence's linked to genetic ancestry (Smith & O'Brien, 2005; Florez *et al*., 2009).

Variation and differentiation among South African ethnic groups have not been significantly investigated therefore meaningful opportunities exist (Wright *et al*., 2011). The analyses of South African ethnic groups with genes such as the *DRD4* VNTR can improve medical treatment and diagnosis due to its mentioned neurotransmission and behavioural association (Seedat *et al*., 2008). The analyses of ethnic groups should however be performed in a manner that considers population structure and history. The validity of comparisons depends on the history and population structure of the subjects (Ewens & Spielman, 1995).

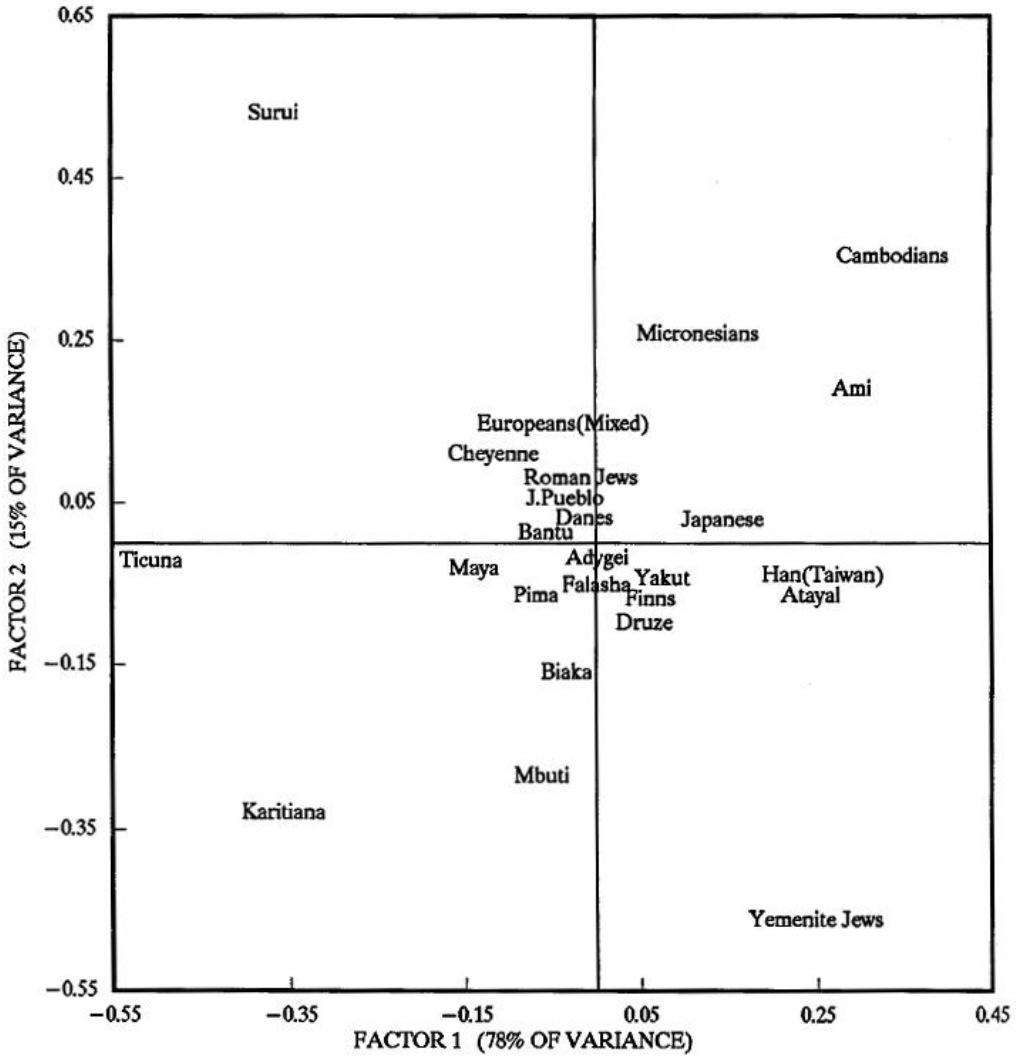

Figure 2.1 Principal component analyses of world-wide populations by Chang *et al*. (1996). Factor one accounted for 78% of the variance whereas factor two accounted for 15% of the variance.

False associations can occur due to geographically related populations being genetically related (Chang *et al*., 1996; Lane *et al*., 2002; Tishkoff *et al*., 2009; Coop *et al*., 2010). A particular polymorphism, (*AGT* M235T) associated with cooler climates, correlated with distances from the equator, thus demonstrating a geographic location representation instead of a trait association (Figure 2.2). Populations from the same continent grouped together instead of correlating with cooler climates for the particular polymorphism (Rosenberg *et al*., 2002; Coop *et al*., 2010).

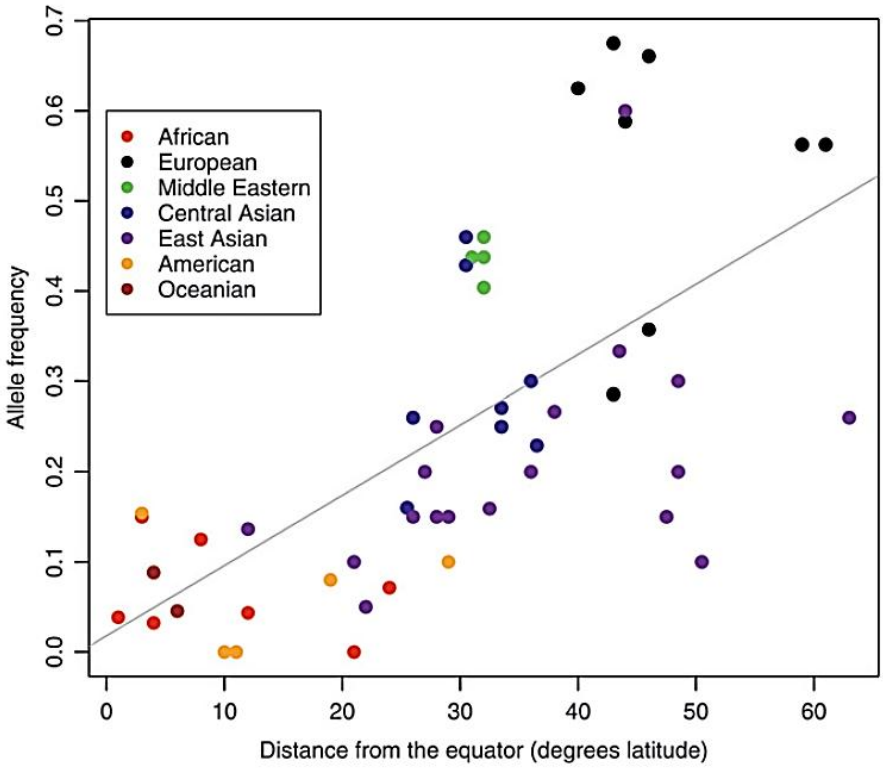

Figure 2.2 Correlations of the distances from the equator and the allele frequency of *AGT* M235T (associated with cooler climates) for different populations. Geographic regions (colour coded) are represented by 52 human populations. Graph obtained from Coop *et al.* (2010) with data from Rosenberg *et al*. (2002) and Thompson *et al*. (2004).

Population subdivision disregarding ethnicity can also result in false associations (Pritchard & Rosenberg, 1999). North American Pima individuals with high degrees of Caucasian ancestry were determined less susceptible to diabetes

(Lander & Schork, 1994). This led to the false conclusion that any allele that was present at a higher frequency in the Pima than among Caucasians was associated with disease. In order to avoid false association it is important that the structure of the subject populations or ethnic groups should be considered.

A genetic population structure exists among South African Bantu ethnic groups (Lane *et al*., 2002). The South African Bantu ethnic group's autosomal and Y-chromosomal DNA was investigated by Lane *et al*. (2002). Seven language groups (Pedi, Southern Sotho, Tsonga/Shangaan, Tswana, Venda, Xhosa and Zulu) were assessed. Analyses revealed that six of the seven language groups clustered together according to their linguistic groups. In addition, genetic distances were determined to an extent, correlating with the geographic locations of the populations (Lane *et al*., 2002). With the assistance of the already somewhat defined linguistic structure of South African populations, the genetic analyses of the populations allowed for the development of a more defined genetic based population structure.

When performing population studies, deviations in structure can be determined by observing the Hardy-Weinberg equilibrium (HWE) (Schaid & Jacobsen, 1999). Genotyping errors in particular can be minimised by adhering to the HWE (Hosking *et al*., 2004). A large number of studies do not report their HWE results and when investigated they proof to deviate (Salanti *et al*., 2005). Deviant studies lose their statistical significance when corrected for the HWE (Trikalinos *et al*., 2006). Studies that adhere to HWE are statistically informative, therefore when investigating neurotransmitter genes such as the *DRD4* VNTR (Ghosh & Seshadri, 2005; Mansoor *et al*., 2008) it is important that the HWE test is adhered too.

Studies on neurotransmitter genes and behavioural traits have been performed on large cohorts of different populations (Bellivier *et al*., 1998; Hoefgen *et al*., 2005; Denys *et al*., 2006). In order to avoid any false association and deviation from the HWE it is important as mentioned to review the history of the subject populations when such information is available. South African ethnic groups (Khoe-San (Khwe), Xhosa, Sotho and Afrikaner) will be discussed briefly with regards to their history as they are the primary subjects of the current study based on neurotransmitter genes.

Khoe-San (Khwe)

The Khwe population has the physical appearance of a Bantu population yet make use of a Khoe-San (Khoe and San) language (de Almeida, 1965). The language spoken by the Khwe belongs to the Kalahari-Khoe division of the Khoe-San (Schlebusch & Soodyall, 2012). The Khwe language has its origins in Angola and northern Namibia (Mesthrie, 1995). Khwe populations predominantly live in the Okavango swamp area and surrounding regions comprising northern Botswana, southern Angola and western Zimbabwe (Schlebusch, 2010).

According to genetic evidence the Khwe possibly originated during, or as a result of Bantu migration into southern Africa. Genetic analyses reveal a southern African Bantu genetic component in the Khwe population supporting this fact (Chen *et al*., 2000; Schlebusch *et al*., 2012; Schlebusch & Soodyall, 2012). It has been suggested that the Khwe also known as the Khoe-speaking San may be descendants of pastoralists that introduced the pastoralists culture to the region located in the present day northern Botswana (Schlebusch, 2010). The Khwe sampled in the current study have their origins in Angola and were employed by the South African Defence Force (SADF) before they were relocated to South Africa (Sharp & Douglas, 1996).

Khoe-San populations thinly populated southern Africa prior to the arrival of Bantu and European migrants. The Khoe were pastoralists owning cattle and sheep, while the San depended on hunting and food gathering. The San lived in small, loosely-knit patrilineal bands of only a few hundred widely dispersed individuals and may never have exceeded 20 000 individuals south of the Orange River. The Khoe expansion was southwards along the eastern edge of the Kalahari to the point where the Orange and Vaal River meet. At the arrival of the first Dutch settlers the Khoe were estimated at a 100 000 individuals (Raven-Hart, 1967; Elphick, 1972; Davenport, 1978; Schrire, 1980; Schlebusch & Soodyall, 2012).

According to suggestions the Khoe originated from the San. However, even if the Khoe did not originate from the San they still share linguistic and physical characteristics proving past interaction (Theal, 1910; Davenport, 1978; Tishkoff *et al*., 2007b). Interactions between Khoe-San and Bantu populations occurred as

mentioned, demonstrated by genetic associations, and the clicks used in the Xhosa and Zulu languages (Herbert, 1990; Tishkoff *et al*., 2009; Schlebusch & Soodyall, 2012). The association between South African Bantu (Xhosa and Zulu) and Khoe-San populations provide evidence for a similar association with the Khwe. Genetic evidence points to a greater Bantu than Khoe-San contribution for the Khwe (Chen *et al*., 2000).

Bantu (Xhosa and Sotho)

The South African Xhosa and Sotho populations form part of the southern branch of the eastern Bantu language group (Lane *et al*., 2002). According to archaeological evidence the Bantu arrived in southern Africa after the Khoe-San but prior to European settlers. Bantu migration (excluding slave trading) into southern Africa took place in a series of swirling currents and backwashes, rather than a constant flow of individuals in one direction (Schapera, 1930; Maggs, 1991). Migration of Bantu individuals according to genetic analyses took place within the last 4 000 years (Maggs, 1991; Tishkoff & Williams, 2002; Reed & Tishkoff, 2006). Possible migration routes into southern Africa indicate a north-eastern entry from a region in present day Cameroon (Campbell & Tishkoff, 2008; Tishkoff *et al*., 2009) (Figure 2.3).

Bantu migrants were funnelled into the Highveld region and coastal hinterland of South Africa by the shifting tsetse-fly belt of the Limpopo valley region and the Kalahari Desert. Four Bantu groups (Venda, Nguni, Tsonga and Sotho-Tswana) relocated below the Limpopo (Davenport, 1978). The Nguni was established in the present day South African provinces of KwaZulu Natal (Zulu population) and Eastern Cape (Xhosa population). The Sotho-Tswana language group was established in the interior of South Africa (Davenport, 1978).

Europeans (Afrikaner)

The Afrikaner population is an amalgamation of mostly Dutch, German and French

settlers in southern African (Davenport, 1978). Although the Afrikaners' European ancestors were not the first to arrive in southern Africa, they were the first to establish a base (Dutch East India Company) (Davenport, 1978; Cox *et al.*, 2001). The Dutch East India Company at first did not intend to establish a fully-fledged colony but later did after economic considerations. European settlement increased in the Cape during the latter stages of the 17th century, drawing mostly individuals of Dutch, German and French origin (Davenport, 1978).

From the Cape Colony the Afrikaner population migrated inland. The "trekboers" (frontier farmers) penetrated the interior of the country along three lines: into the Little Karoo and eventually down the Lown Kloof, up the Hex River pass into the Great Karoo or through the south-western districts to Mossel Bay. The Afrikaner advance destroyed many of the Khoe-San populations, drove some into the interior or absorbed them (Davenport, 1978; Keim, 1995).

The transfer of the Cape Colony from the Dutch East India Company to British control took place in periods from 1795 onwards (Cox *et al*., 2001). In 1820, parties from England, Ireland, Scotland and Wales were brought in. A large scale Afrikaner exodus out of the Cape Colony occurred in the 1830's, due to the British occupation. Afrikaner individuals from Albany, Cradock, Graaff-Reinet, Somerset East and Uitenhage constituted the migrant majority (Davenport, 1978). These migrants, known as the "Voortekkers", moved into two main directions, northwards towards Transvaal (today the Gauteng, Limpopo, Mpumalanga and North-West provinces of South Africa) or over the Drankensberg mountains into Natal (KwaZulu-Natal today). Approximately 15 000 "Voortrekkers" relocated northwards as far as the Limpopo River (Davenport, 1978; Keim, 1995).

Human migration into southern Africa and the interactions associated with it affected present day South Africa. Migrations resulted in South Africa having one of the largest and most diverse population structures in Africa consisting of various ethnic groups[2]. A population such as the Xhosa which migrated to the Eastern Cape of South Africa constitutes the majority of the present day Eastern Cape population. The Zulu population is the most prominent in KwaZulu Natal and the Sotho is the most prominent in the Free State (Davenport, 1978).

[2] Statistics South Africa Census 2011, www.Statssa.gov.za

The subjective characterisation of populations based on their past histories with regards to geographic and linguistic features is valuable when performing ethnic group studies. This initial characterisation lays a platform to perform more natural assignments, such as those based on genetic diversity (Pritchard *et al*., 2000). The genetic analyses of populations improve and provide more information with regards to past historical accounts (Cavalli-Sforza, 1997; Tishkoff *et al*., 2009) and in certain cases (*DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR) also provide information that can benefit the behavioural and psychiatric field (Chang *et al*., 1996; Sabol *et al*., 1998; Ghosh & Seshadri, 2005; Hu *et al*., 2006; Mansoor *et al*., 2008).

Figure 2.3 Bantu migration into southern Africa [adapted from Reed and Tishkoff (2006) and Campbell and Tishkoff (2008)].

Neurotransmitters studies in South African ethnic groups

Genes affecting neurotransmission and subsequently human behaviour (traits and disorders) have been investigated in a number of studies on different populations (Muramatsu *et al*., 1996; Chen *et al*., 1999; Samochowiec *et al*., 1999; Schmidt *et al*., 2000; Feinn *et al*., 2005; Herman *et al*., 2005; Lahti *et al*., 2005; Contini *et al*., 2006; Philibert *et al*., 2008; Matthews & Butler, 2011; Wang *et al*., 2011; Creswell *et al*., 2012). Lack of genetic research among African populations however has contributed to ADRs and drug inefficacy (Wright *et al.*, 2011). The analysis of neurotransmitter genes among South African ethnic groups can benefit the medical and pharmacological field (Tishkoff & Williams, 2002; Tishkoff *et al*., 2009; Wright *et al*., 2011).

Behavioural disorders are one of the top causes of disability-adjusted life years in South Africa (Seedat *et al*., 2008). Lack of genetic information with regards to disease diagnosis and treatment has hampered the behavioural and psychiatric disorder fields (Wright *et al*., 2011). Genetic variation among population groups contributes to different reactions with regards to disease treatment (Chang *et al*., 1996; Ghosh & Seshadri, 2005; Mansoor *et al*., 2008). Identifying the genetic variation of implicated genes among populations is one of the first steps towards developing better treatment and drug strategies.

Dopamine and serotonin are neurotransmitters that specifically have an effect on behaviour. Monoamine Oxidase, an enzyme involved in neurotransmission, is also associated with behaviour (Wise & Rompre, 1989; Chang *et al*., 1996; Lesch *et al*., 1996; Sabol *et al*., 1998). Dopamine belongs to a class of neurotransmitters known as catecholamines, which have a catechol ring and an amine side chain. Catecholamines communicate between neurons and act within the anatomically confined spaces of synapses. In neural tissue, dopamine is primarily synthesized in the central nervous system (CNS), limited production also occurs in the adrenal medulla. The function of dopamine being a prolactin inhibitor (PRL) was not determined until the 1970's (Ben-Jonathan & Hnasko, 2001).

Monoamine Oxidase A oxidises dietary amines and neurotransmitters, with a focus on serotonin, norepinephrine and dopamine as substrates (Grimsby *et al.,*

1991). The drug Iproniazid which is an inhibitor of the enzyme MAO was the first drug used for the treatment of depression. Treatment with MAO inhibitors increases the concentrations of serotonin and noradrenaline (norepinephrine) (Mongeau *et al*., 1997).

Serotonin mediated neurotransmission contributes to psychological functions such as food intake, motor activity and sleep (Mongeau *et al*., 1997; Lucki, 1998). Most serotonin containing neurons are localised along the midline of the brain stem and send long axons to innervate a wide distribution of receiving areas throughout the nervous system. The 5-HTT regulates the magnitude of serotonergic response which is important for the fine tuning of the brains serotonergic neurotransmission (Lesch *et al*., 1996). Expression of the 5-HTT is abundant in the limbic area of the brain which is involved in the emotional aspects of behaviour (Mongeau *et al*., 1997).

The human dopaminergic system is associated with different personality disorders, neuropsychiatric disorders and diseases (Risch & Merikangas, 1996; Swanson *et al.*, 1998; Jardemark *et al.*, 2002; Hoenicka *et al*., 2007). Association studies between diseases, behaviours and genes are popular due to their efficiency (Pritchard & Rosenberg, 1999; Tang *et al*., 2005). However, the detection of complex diseases such as bipolar disorder and schizophrenia are more complicated than just simple association due to the possible interaction of more than one gene (Risch & Merikangas, 1996).

In the current study the *DRD4* VNTR (Seaman *et al*., 1999; Okuyama *et al*., 2000; Mitsuyasu *et al*., 2007), the *MAO-A*-uVNTR (Sabol *et al*., 1998; Huang *et al*., 2004; Guo *et al*., 2008) and the 5-HTTLPR (Kunugi *et al*., 1997; Michaelovsky *et al*., 1999; Nakamura *et al*., 2000) will be investigated, due to the effect that they have on the neurotransmitters dopamine and serotonin and the enzyme MAO. Studies have already focused on the function of these particular genes associated with behavioural traits and disorders. However, limited studies have made use of South African populations as the target populations.

When investigating the Afrikaner it will be beneficial to review studies on Caucasian populations from the Netherlands, France and Germany, primarily due

to the fact that Afrikaner ancestors originated from these countries (Elphick, 1972; Davenport, 1978; Botha & Beighton, 1983). The 7-repeat allele of the *DRD4* VNTR for example was associated with maternal insensitivity and externalizing (aggressive, oppositional) behaviour in pre-schoolers from the Netherlands (Bakermans-Kranenburg *et al*., 2006). Externalizing behaviour was determined increasing six-fold in children that were exposed to insensitive care and that were born with the 7-repeat allele compared to children without these combined risks (Bakermans-Kranenburg *et al*., 2006). In another study using Dutch individuals, the relationship between *DRD4* VNTR and startle reactivity, gambling performance and sensation seeking were investigated. No association between the *DRD4* and the mentioned behaviours were determined in that particular study (Nederhof *et al*., 2011).

A German population was investigated to determine the association between the *MAO-A*-uVNTR polymorphism and depression (Schulze *et al*., 2000). It was determined that the high activity *MAO-A* alleles predisposed Germans to major depressive disorders (Schulze *et al*., 2000). In another study comprising German individuals no associations were determined between antisocial behaviour, alcoholism and the *MAO-A*-uVNTR (Koller *et al*., 2003). The genotype distribution was determined similar for the controls and alcoholics in that particular study. Among the alcoholics, 57 (0.337) individuals had the 3-repeat allele, 109 (0.645) had 4-repeat allele and three individuals (0.018) had 5-repeat allele. Twenty-eight individuals (0.389) from the control group had the 3-repeat allele, 44 (0.611) had the 4-repeat allele and none had the 5-repeat allele (Koller *et al*., 2003).

A family based association study was performed on the 5-HTTLPR with regards to anxiety, neuroticism and depression, making use of Dutch individuals (Middeldorp *et al*., 2007). In total 1 245 participants along with 559 of their parents were analysed. No additive effect was determined for the short (S) allele of the 5-HTTLPR in relation to depression. The allelic frequency of the long (L) allele was 0.570 whereas the frequency of the S allele was 0.430. The results demonstrated no direct association between the 5-HTTLPR and the mentioned behavioural traits (Middeldorp *et al*., 2007). The 5-HTTLPR allele frequencies were analysed in three South African populations (Esau *et al*., 2008). Samples were obtained from

68 Bantu, 96 Caucasian and 178 Coloured individuals. Among the Bantu population the L allele had a frequency of 0.840 and the S allele had a frequency of 0.160. The Caucasian population had an L allele frequency of 0.610 and S allele frequency of 0.390. The Coloured population had an L allele frequency of 0.860 and an S allele frequency of 0.140 (Esau *et al*., 2008).

The *DRD4* VNTR allele distribution was analysed world-wide with the use of 1 327 individuals from 36 different populations (Chang *et al*., 1996). The South African Bantu population reported in the study had a frequency of 0.610 for the 4-repeat allele. The 2-, 5-, 7- and 8-repeat allele frequencies for the same population were respectively 0.050, 0.040, 0.190 and 0.110. A small Khoe-San population was also analysed, only two alleles were identified, the 4- and the 6-repeat. The 4-repeat allele had a frequency of 0.910 compared to the 6-repeat having a frequency of 0.090 (Chang *et al*., 1996). It should be noted that Khoe-San sample population consisted of only 22 individuals, thus underscoring the genetic diversity of the population (Chang *et al*., 1996; Chen *et al*., 2000; Tishkoff *et al*., 2009; Schlebusch *et al*., 2012).

The *DRD4*, *MAO-A* and *5-HTT* (5-HTTLPR) will be individually discussed at greater length in the following segments. Descriptions of polymorphs associated with neurotransmitters will be reviewed along with behavioural association studies. Special mention will be made on the association of the genes with behaviour.

Dopamine Receptor D4

The DRD4 is a G protein-coupled receptor belonging to the dopamine D2-like receptor family (Oldenhof *et al*., 1998). The D2-like DRD4 encodes a 387 amino acid protein having seven transmembrane domains, several Src homology 3 (SH3) binding domains and a potential N-linked glycosylation site (Van Tol *et al.*, 1991). The SH3 domains are modular binding domains for protein-protein interactions which are essential for full functional activity, recognizing proline-rich sequence motifs (Ren *et al*., 1993). The D4 receptor can react with a large variety of SH3 domains of different origin, with the strongest interactions being with the Growth Factor Receptor-Bound Protein 2 (GRB2) and the non-catalytic region of

the Tyrosine Kinase Adaptor Protein (NCK) (Oldenhof *et al*., 1998).

The *DRD4* gene is located on chromosome 11p15.5. Various polymorphisms are present within the *DRD4* coding region including the mentioned VNTR sequence in exon III (Van Tol *et al*., 1992; Catalano *et al*., 1993; Nöthen *et al*., 1994; Seeman *et al*., 1994; Seaman *et al*., 1999; Okuyama *et al*., 2000; Mitsuyasu *et al*., 2007) (Figure 2.4). The alleles of the exon III polymorphism vary from 2- to 11-repeat units of 48 base pair (bp) (Mansoor *et al*., 2008). The 4-repeat allele is the most prevalent in human populations followed by the 7- and 2-repeat alleles. All the other alleles are sporadically distributed in various populations (Chang *et al*., 1996).

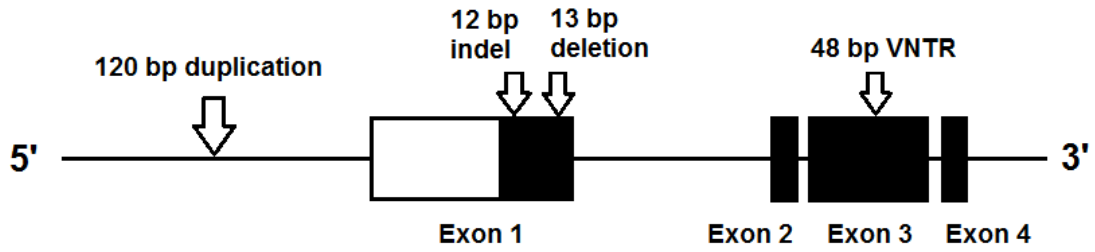

Figure 2.4 A schematic representation of polymorphisms in the *DRD4* gene [adapted from Okuyama *et al*. (2000) and Mitsuyasu *et al*. (2007)]. Indicated are the 120 bp duplication, 12 bp insertion/deletion, 13 bp deletion, 48 bp VNTR and correlated exons.

Variability exists among the different alleles of the *DRD4* VNTR. The 7-repeat allele is associated with a partial loss of *DRD4* mediated prefrontal inhibition in comparison to the 4-repeat allele (Asghari *et al*., 1995; Wang *et al*., 2004). The 7-repeat allele originated after the 4-repeat allele and prior to the upper Palaeolithic era (40 000-50 000 years ago) according to intra-allelic comparisons. African individuals have the most polymorphisms within the repeat sequences of the VNTR according to the comparisons (Wang *et al*., 2004). With the use of the *DRD4* VNTR and a multidimensional scale it was determined that peripheral geographically located populations are more divergent from centrally located populations (Figure 2.5). Native Pacific, North and South American populations were determined the most divergent. African, European and Asian populations grouped closer together in the multidimensional scale (Mansoor *et al*., 2008).

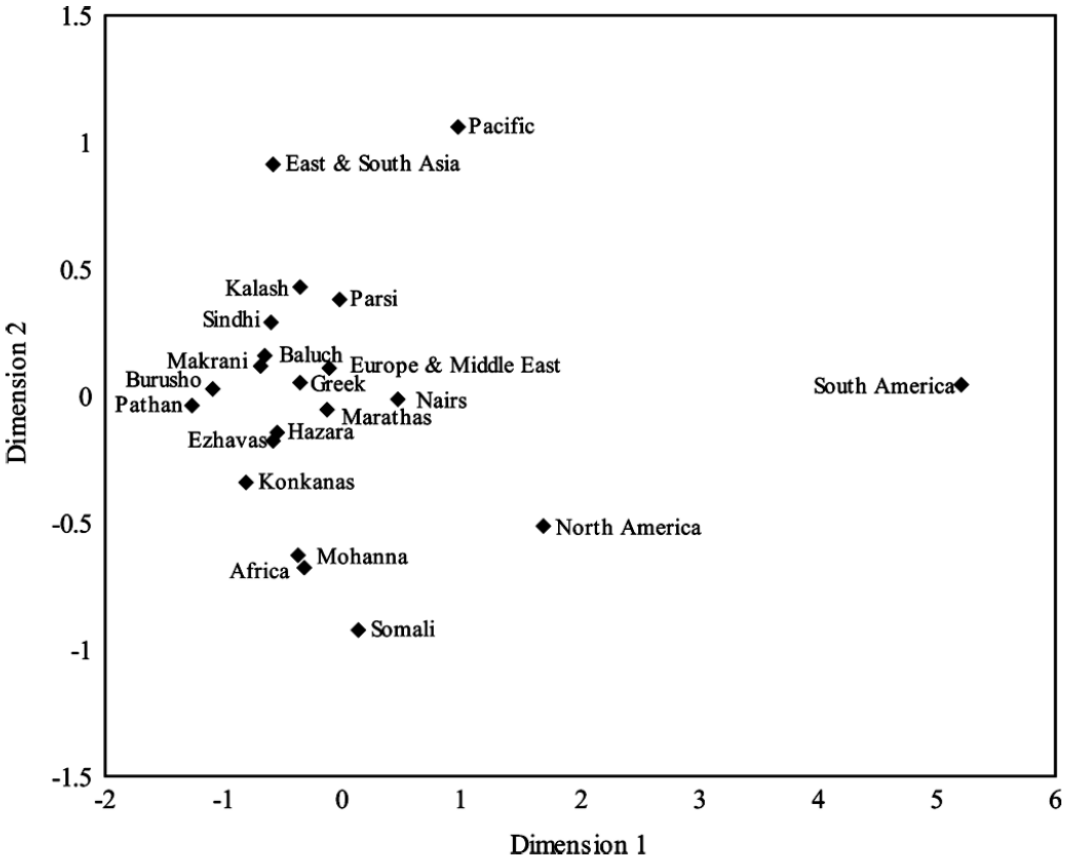

Figure 2.5 Multidimensional scaling of world-wide populations with the use of the *DRD4* VNTR. Pacific, North and South American populations were determined the most divergent. Majority of the analysed European, African and Asian populations grouped together (Mansoor *et al*., 2008).

In addition to the *DRD4* VNTR in exon III, a deletion consisting of 13 bp was found in exon I of the *DRD4* coding sequence (Figure 2.4). This particular deletion results in an altered reading frame from amino acid 79. A stop codon is generated 20 amino acids downstream of the deletion as a consequence of the altered reading frame (Nöthen *et al*., 1994). Somatic ailments such as obesity, acusticus neurinoma and disturbances of the autonomic nervous system were some symptoms of an individual homozygous for this mutation. The absence of a functional DRD4 protein may be responsible for these symptoms (Nöthen *et al*., 1994).

A polymorphic 12 bp repeat unit was found in exon I in addition to the mentioned 13 bp deletion (Catalano *et al*., 1993) (Figure 2.4). A sequence of four amino acids in the extracellular N-terminal part of the receptor characterises this particular mutation. The 12 bp unit occurs as a single 12 bp sequence or a more common two-fold repeat. Patients with delusional disorder have the rare single repeat at higher frequencies (Catalano *et al*., 1993).

At amino acid 194 of the DRD4 a val194-to-gly variant was observed (Seeman *et al*., 1994). The amino acid substitution, the result of a T to G transversion, is located one amino acid away from a serine amino acid. This particular serine amino acid is important for the binding of dopamine to the dopamine D2 receptor (Seeman *et al*., 1994). The gly194 variant is two orders of magnitude less sensitive to Clozapine, dopamine and Olanzapine than the wild type receptor. No pubic hair, no axillary hair, sickle cell disease and obesity were some symptoms of an individual homozygous for this variant (Liu *et al*., 1996).

A 120 bp polymorphic tandem duplication was found 1.2 kb upstream of the *DRD4* initiation codon (Seaman *et al*., 1999) (Figure 2.4). This particular polymorphism was observed in 371 children with attention deficit hyperactivity disorder (ADHD) (McCracken *et al*., 2000). A long (240 bp) and a short (120 bp) allele was determined for this particular polymorphism. Preferential transmission of the long allele for probable or definite ADHD was determined by a transmission disequilibrium test (TDT) (McCracken *et al*., 2000).

Nucleotide variants in the promoter region of *DRD4* ranging from the -1302 to -123 bp were observed in Japanese individuals (Okuyama *et al*., 2000) (Figure 2.6). Seven nucleotide variants: -521C to T, -602(G) 8-9, -603T insertion/deletion, -616C to G, -809G to A, -1123C to T and -1217G insertion/deletion were revealed by direct sequencing and single strand conformation polymorphism (SSCP). All the variants were polymorphic among the Japanese subjects except the -1123C to T which was only observed in one individual (Okuyama *et al*., 2000).

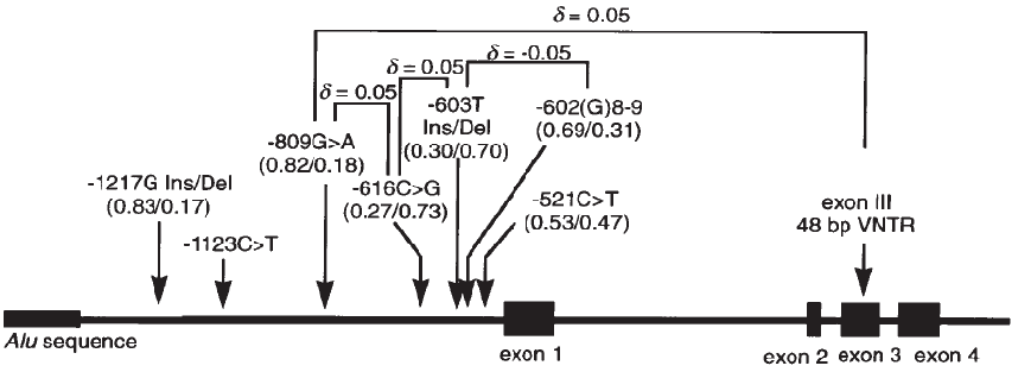

Figure 2.6 Location of variants identified in the promoter region of the *DRD4* (Okuyama *et al*., 2000).

The *DRD4* gene was extensively analysed for published or novel variants in 216 schizophrenic patients (Mitsuyasu *et al*., 2007) (Figure 2.7). Several insertion/deletion polymorphisms and a number of single nucleotide polymorphisms (SNP) were identified, including four novel SNPs and a novel mononucleotide repeat. The polymorphisms were the 48 bp VNTR (+2689 to +2880), 21 bp deletion (+106 to +126), 12 bp repeat (+64 to +87), +31G to C, -11C to T, -128G to T, -234C to A, -291C to T, -364A to G, -376C to T, -521T to C, -597(G) 2-5, -598G to T, -600G to C, -603T insertion/deletion, -615A to G, -616G to C, -713C to T, -768G to A, -809G to A, -906T to C, -930C to G or T, -1102G to A, -1106T to C, -1123C to T, -1217G insertion/deletion and the 120 bp tandem repeat (-1480 to -1240) (Mitsuyasu *et al*., 2007).

The extensively studied 48 bp *DRD4* VNTR is associated with a number of behavioural disorders such as schizophrenia, OCD, alcoholism and traits such as novelty seeking and risk taking (Muramatsu *et al*., 1996; Ebstein *et al*., 1997; Tomitaka *et al*., 1999; Kluger *et al*., 2002; Lung *et al*., 2002; Lahti *et al*., 2005; Camarena *et al*., 2007; Dreber *et al*., 2009; Kuhnen & Chiao, 2009; Gonçalves *et al*., 2012). The long alleles (≥6-repeat alleles) of the *DRD4* VNTR are especially associated with schizophrenia among Caucasians (Lung *et al*., 2002). An increased density and number of D4 receptors are present in the brains of schizophrenic individuals. Delayed schizophrenia onset was determined among females with the 7-repeat allele of the *DRD4* VNTR (Gonçalves *et al*., 2012). The

DRD4 protein is a major target for the drug Clozapine, used in the treatment of schizophrenia (Jardemark *et al*., 2002). Clozapine has the brand name Cloment in South Africa.

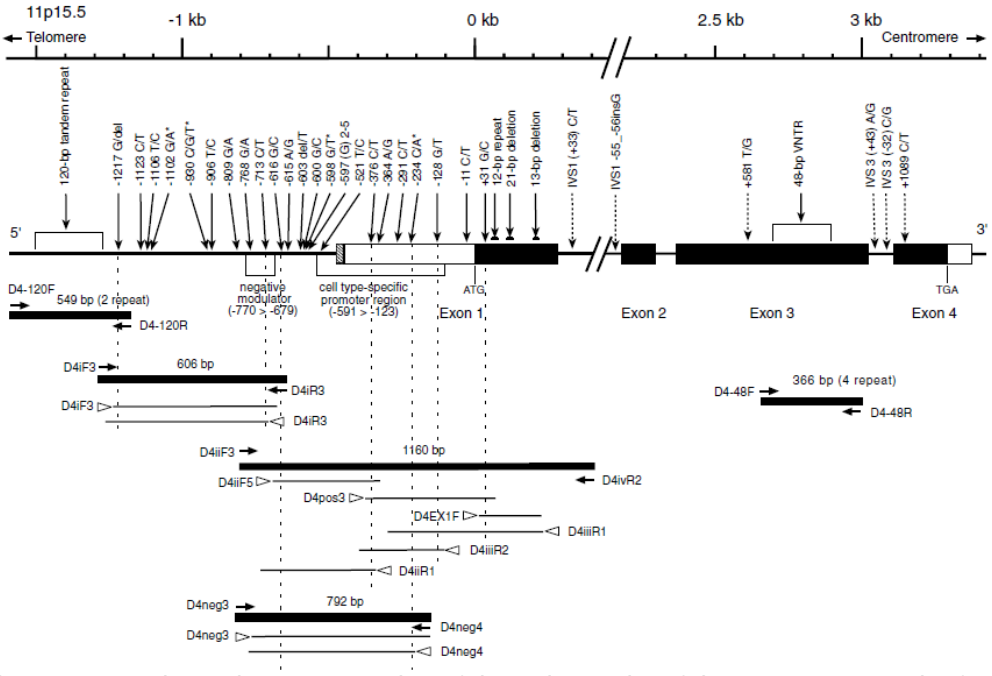

Figure 2.7 A schematic representation of the polymorphs of the *DRD4* gene. The four exons are indicated by black boxes. The white boxes are indicative of untranslated regions. Polymorphisms are indicated by closed arrows. Sequenced regions and the orientation of primers are indicated by thin open arrows. The five fragments genotyped are indicated by bold lines (Mitsuyasu *et al*., 2007).

The long alleles (≥5-repeat) of the *DRD4* VNTR are associated with OCD (Camarena *et al*., 2007). The 7-repeat allele was determined a factor in the phenotypic variance of tics among individuals with OCD (Cruz *et al*., 1997). The 2-repeat allele has a protective effect against OCD symptoms (Millet *et al*., 2003). Obsessive compulsive symptoms have emerged among individuals treated with Clozapine (Kim & Kim, 2006).

Alcoholism in Japanese individuals is linked to the 5-repeat allele of the *DRD4* VNTR. Alcoholics with the 5-repeat allele are also prone to other drug abuse (Muramatsu *et al*., 1996). The relationship between *DRD4*, alcohol consumption and social bonding was investigated (Creswell *et al*., 2012). Individuals with the 7-

repeat allele reported higher social bonding levels during alcohol consumption compared to 7-absent allele carriers. Novelty seeking was determined to be partially mediating alcohol abuse among young adults (Ray *et al*., 2008).

A gene-environment interaction was observed for alcoholism, novelty seeking and the *DRD4* VNTR (Lahti *et al*., 2005). High novelty seeking levels were reported for individuals with short alleles (2- or 5-repeat) that had fathers with frequent alcohol consumption or drunkenness episodes during their childhood. Novelty seeking levels were not associated with the *DRD4* alleles when less frequent alcohol consumption or drunkenness episodes occurred during childhood (Lahti *et al*., 2005).

Olanzapine reduces alcohol craving in individuals that carry either the short (no ≥ 7-repeat) or long (≥7-repeat) *DRD4* VNTR alleles (Hutchison *et al*., 2003). Naltrexone reduces the percentage of drinking days and the percentage of heavy drinking days in individuals with the *DRD4* long alleles (≥7-repeat). Individuals treated with Naltrexone had 8% fewer drinking days (Tidey *et al*., 2008). Olanzapine is sold under the trade names Lanzek, Zypadhera and Zyprexa. The trade name Zyprexa is used for Olanzapine in South Africa. Naltrexone trade names are Revia, Depade and Vivitrol, when marketed in its generic from, naltrexone hydrochloride.

Risk loving (novelty seeking characteristic) among men was associated with the 7-repeat allele of the *DRD4* VNTR (Dreber *et al*., 2009). Individuals that carry the 7-repeat allele took up to 25% more risk (Kuhnen & Chiao, 2009). Japanese women with the long *DRD4* VNTR alleles also had high novelty seeking levels (Tomitaka *et al*., 1999). The long *DRD4* alleles were associated with high novelty seeking scores in 13 out of 20 studies (Kluger *et al*., 2002). However, no difference was found in risk taking between individuals that had the 7-repeat allele absent or present in a study comprising 237 tournament contract bridge players (Dreber *et al*., 2011). No association between the *DRD4* polymorphs and suicidal behaviour (risk taking behaviour) has been determined (Zalsman *et al*., 2004). Migratory behaviour (novelty seeking characteristic) has been linked to the *DRD4* VNTR (Chen *et al*., 1999; Matthews & Butler, 2011). The 2- and 7-repeat *DRD4* alleles in particular were associated with migration in an extreme discordant phenotype

strategy. Significantly higher novelty seeking scores were observed for individuals with the 2- or/and 7-repeat alleles (Reist *et al*., 2007).

Low neuronal reactivity and increased exploratory behaviour are associated with certain *DRD4* polymorphisms indicating the biological influence of the *DRD4* on novelty seeking behaviour (Matthews & Butler, 2011). The *DRD4* VNTR is associated with sexual behaviours, particularly promiscuity and infidelity. Individuals with at least one 7-repeat *DRD4* allele report a greater categorical rate of promiscuous sexual behaviour (Garcia *et al*., 2010). Differences in sexual behaviour are likely partially mediated by individual genetic variation in genes coding for motivation and reward in the brain (Garcia *et al*., 2010).

Monoamine Oxidase A

Monoamine Oxidase was first identified in 1928 by Mary Hare (later Mary Bernheim). The oxidation and deamination of tyramine was observed by an unknown enzyme then named Tyramine Oxidase (Hare, 1928) and later renamed MAO by E.A. Zeller (Slotkin, 1999). Monoamine Oxidase A, an isoenzyme of MAO, oxidises dietary amines and neurotransmitters, with a focus on serotonin, norepinephrine and dopamine as substrates (EC 1.4.3.4; Grimsby *et al.,* 1991). The MAO-A protein spans 527 amino acids with a molecular mass of 59.7 kilo Daltons (kDa) (Hsu *et al*., 1989) and has a subcellular location within the outer mitochondrial membrane (Schnaitman *et al*., 1967). The *MAO-A* gene is located on the X-chromosome (Kochersperger *et al*., 1986) and is 1 937 nucleotides long (Hsu *et al*., 1988), consisting of 15 exons (Grimsby *et al*., 1991) (Figure 2.8).

Figure 2.8 A schematic representation of *MAO-A* depicting the location of exons. The unfilled bars represent untranslated regions, while the filled bars represent the coding regions. Below the bars exon numbers are indicated. Sequenced areas are indicated by horizontal arrows (Grimsby *et al*., 1991).

A number of polymorphisms are associated with the *MAO-A* gene (Brunner *et al*., 1993; Sabol *et al*., 1998; Gilad *et al*., 2002). The mentioned *MAO-A*-uVNTR is 1.2 kb upstream of the *MAO-A* coding sequence and consists of variants having 2, 3, 3.5, 4, 5 or 6 copies of a 30 bp sequence (ACCGGCACCGGCACCAGTACCCGCACCAGT) (Sabol *et al*., 1998; Huang *et al*., 2004; Guo *et al*., 2008) (Figure 2.9). The 3- and 4-repeat alleles are the most common among human populations (Sabol *et al*., 1998).

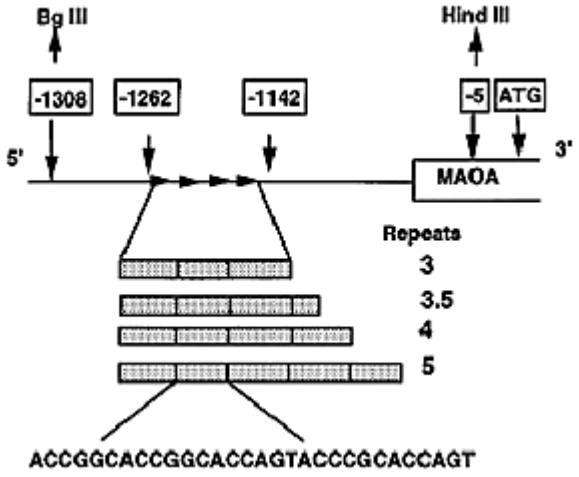

Figure 2.9 A schematic representation of *MAO-A* and the *MAO-A*-uVNTR repeat sequence (Sabol *et al*., 1998). The 30 bp repeat sequence is indicated.

The 3-repeat allele of the *MAO-A*-uVNTR had a frequency of 0.591 among African Americans, whereas the 4-repeat allele had a frequency of 0.364. The 3-repeat allele (0.610) was more prevalent than the 4-repeat allele (0.378) in Asian/Pacific Islanders whereas the 4-repeat allele (0.648) was twice as common as the 3-repeat allele (0.331) in White/Non-Hispanic populations. The 3.5- or 4-repeat alleles transcribe 2 to 10 times more efficiently than those with 3- or 5-repeat alleles (Sabol *et al*., 1998). The uVNTR alleles are usually divided into either high activity or low activity genotypes (Fergusson *et al*., 2011; Lung *et al*., 2011).

Five regions of the *MAO-A* gene were analysed in 56 males from 7 different ethnic groups (Gilad *et al*., 2002) (Figure 2.10). Results revealed 41 segregating sites, 39 were single nucleotide substitutions of which 11 were only identified once in a 90 kb region. Forty-six distinct haplotypes were identified among the 56 males.

Linkage disequilibrium was determined greater than expected under panmixia which is consistent with differentiation between ethnic groups (Gilad *et al*., 2002).

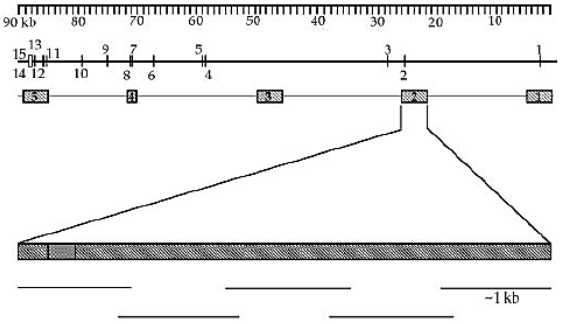

Figure 2.10 Schematic representation of the genomic structure of *MAO-A* and the sequencing strategy used by Gilad *et al*. (2002). The five sequenced products are indicated (overlapping segments) (Gilad *et al*., 2002).

An additional *MAO-A* polymorphism, the gln296-to-ter substitution resulting in Brunner syndrome was determined (Brunner *et al*., 1993). The defect for this condition is present in the p11-p12 region of the X-chromosome and was identified in a Dutch family. Males with the particular substitution all had aggressive and sometimes violent behaviour in response to fear, anger or/and frustration (Brunner *et al*., 1993).

Studies focussing on male individuals are easily interpreted due to the location of *MAO-A* on the X-chromosome (Cohen *et al*., 2003; Koen *et al*., 2004; de Milander *et al*., 2009; McDermott *et al*., 2009; Ferguson *et al*., 2011). The possibility of either X-chromosome inactivation or non-inactivation resulted in conflicting results regarding females (Carrel *et al*., 1999; Nordquist & Oreland, 2006; Pinsonneault *et al*., 2006). Suggestions that *MAO-A* belongs to a set of genes that are not X-chromosomally inactivated (Carrel & Willard, 2005) have not been confirmed (Nordquist & Oreland, 2006). The inactivating ratio of a certain *MAO-A* polymorphism (30 bp uVNTR) does not correlate with the X-chromosome inactivation ratio determined for the X-linked *Androgen Receptor* indicating the possibility of dosage compensation (Pinsonneault *et al*., 2006).

Regulation of *MAO-A* by methylation is evident. The extent of methylation in females is highly variable and correlates with allelic expression imbalance indicative of methylated regulated gene expression (Pinsonneault *et al*., 2006). Methylation brought on by life events, medication and/or unknown mediators have been demonstrated (Stabellini *et al*., 2009).

Methylation ratios for the *MAO-A*-uVNTR are consistently higher among females than males when testing for nicotine and alcohol dependence (Philibert *et al*., 2008). However, neither female nor male patients have a significant change in expression of *MAO-A* messenger ribonucleic acid (mRNA) when they are classified into groups based on their genotypes (e.g. male 3-repeat allele and female 3-, 3-repeat alleles or male 4-repeat allele and female 4-, 4-repeat alleles) (Sun *et al*., 2012). Heterozygotes though exhibit patterns of expression greater than corresponding homozygotes (Rockman & Wray, 2002).

Monoamine Oxidase A polymorphs, in particular the extensively investigated *MAO-A*-uVNTR, are associated and linked to mental and behavioural disorders. The associations between the uVNTR, disorders and traits vary with strength. Disorders such as schizophrenia, autistic behaviour, bipolar disorder, anxiety disorders, antisocial behaviour, aggressive behaviour and alcoholism and traits such as neuroticism and endurance performance have all been implicated (Lim *et al*., 1994; Kawada *et al*., 1995; Deckert *et al*., 1999; Samochowiec *et al*., 1999; Schmidt *et al*., 2000; Caspi *et al*., 2002; Yirmiya *et al*., 2002; Cohen *et al*., 2003; Eley *et al*., 2003; Huang *et al*., 2004; Jones *et al*., 2004; Koen *et al*., 2004; Samochowiec *et al*., 2004; Herman *et al*., 2005; Contini *et al*., 2006; Ducci *et al*., 2008; McDermott *et al*., 2009; Qiu *et al*., 2009; Cohen *et al*., 2011; Fergusson *et al*., 2011; Pardini *et al*., 2011).

Monoamine Oxidase A possibly contributes to an increased risk for paranoid schizophrenia but the precise mechanism is not understood (Sun *et al*., 2012). A study based on Chinese schizophrenic individuals and *MAO-A* revealed that males with the -941T polymorph had an increased risk for schizophrenia. However, in the same study no significant association was found between the *MAO-A*-uVNTR and schizophrenia (Qiu *et al*., 2009). No association between the *MAO-A*-uVNTR and schizophrenia was observed in two meta-analyses studies

(Norton *et al*., 2002; Li & He, 2008). It was concluded that *MAO-A*-uVNTR does not play any major role in the manifestation of violence in schizophrenia.

The relationship between the *MAO-A*-uVNTR and autism was investigated using 41 males with an average age of 12.6 years (Cohen *et al*., 2003). Children with the low activity *MAO-A* 3-repeat allele had more severe autistic behaviour and lower intelligence quotients (IQ). During follow up readings the individuals with the low activity allele showed reduced IQ levels (Cohen *et al*., 2003). Among the affected boys the 4-repeat allele (0.560) was the most common followed by the 3- and 3.5-repeat alleles (0.360 and 0.090).

Male boys with the low activity *MAO-A*-uVNTR 3-repeat allele were determined having more severe sensory behaviours, aggression and arousal regulation problems than boys with the high activity alleles (Cohen *et al*., 2011). Boys with the 4-repeat allele (high activity allele) with mothers homozygous for the 4-repeat allele displayed the mentioned behaviours at an increased frequency in comparison to individuals born to heterozygous mothers (Cohen *et al*., 2011). Maternal genotypes at the *MAO-A* locus may possibly modify the IQ of autistic children through the intrauterine environment (Jones *et al*., 2004). An association between IQ levels and *MAO-A*-uVNTR genotypes were found in autistic subjects with affected siblings (Yirmiya *et al*., 2002).

A weak association was determined between bipolar patients and a dinucleotide repeat polymorphism associated with *MAO-A* (Lim *et al*., 1994; Kawada *et al*., 1995). In another study, no association was determined between *MAO-A* and bipolar disorder (Turecki *et al*., 1999). Furthermore, no association was determined between bipolar disorder and the *MAO-A*-uVNTR (Müller *et al*., 2007).

Panic disorder has been associated with the *MAO-A*-uVNTR. A study on Italian and German women found that the longer alleles (3a-, 4- or/and 5-repeat) were more prominent in panic disorder patients than controls (Deckert *et al*., 1999). The 3a allele contains three copies of the 30 bp sequence (ACCGGCACCGGCACCAGTACCCGCACCAGT) with an additional 18 bp partial repeat sequence (CCAGTACCGGCACCGGCA). In another study, female patients suffering from panic attacks and anxiety disorders were also associated

with the longer *MAO-A*-uVNTR alleles (>3-repeat) (Samochowiec *et al*., 2004). In a study comprising 620 individuals, no relationship was observed between *MAO-A*-uVNTR alleles and panic disorder (Hamilton *et al*., 2000). A possible reason why no association was found by Hamilton *et al*. (2000) compared to Deckert *et al*. (1999) may be due to Hamilton *et al*. (2000) making use of a family based study while Deckert *et al*. (1999) made use of a population based study.

A large sample of male children were analysed in order to determine why certain maltreated children develop antisocial behaviour (Caspi *et al*., 2002). Results indicated that the *MAO-A*-uVNTR moderates the effect of maltreatment, with a significant gene-environment interaction. The effect of childhood maltreatment on antisocial behaviour was significantly weaker among males with the high activity *MAO-A* alleles (3.5- or/and 4-repeat) (Caspi *et al*., 2002). Individuals abused during childhood having low activity *MAO-A* alleles (2.5- or/and 3-repeat) were significantly more likely to be hostile (Fergusson *et al*., 2011). An association between childhood adversity and the low activity *MAO-A* alleles (2-, 3- or/and 5-repeat) resulting in antisocial behaviour was also established (Weder *et al*., 2009). However, in a separate study, no association was determined between *MAO-A*-uVNTR alleles, maltreatment and conduct problems (Haberstick *et al*., 2005).

The influence of the *MAO-A* gene promoter and rearing experience on aggressive behaviour was studied in 45 male rhesus monkeys (Newman *et al*., 2005). The study indicated that behavioural expression of the *MAO-A* gene is sensitive to social experiences early in development. Mother-raised monkeys with the high activity *MAO-A* alleles (5- or/and 6-repeat) had higher aggression scores (Newman *et al*., 2005). Human aggression manifests with greater frequency and intensity when provocation is experimentally manipulated upwards among individuals with the low activity *MAO-A*-uVNTR alleles (2.5- or/and 3-repeat) (McDermott *et al*., 2009).

A significant association between the low activity *MAO-A*-uVNTR alleles (2-, 3- or/and 5-repeat) and aggression was determined among individuals with brain injuries (Pardini *et al*., 2011). Control individuals carrying the low activity alleles revealed higher aggression levels than individuals with the high activity alleles (3.5- or/and 4-repeat) (Pardini *et al*., 2011). Individuals with prefrontal cortex

(PFC) lesions had no aggressive differences with or without the low activity *MAO-A* alleles. However, higher aggression levels were observed among the non-PFC lesion individuals with the high activity alleles (3.5- or/and 4-repeat). It was concluded that PFC integrity is required for the modulation of aggressive behaviours genetically (Pardini *et al*., 2011).

Antisocial behaviour was associated with the 3-repeat *MAO-A*-uVNTR allele in male alcoholics of German descent (Samochowiec *et al*., 1999; Schmidt *et al*., 2000). A similar association was observed among Brazilian males (European descent) (Contini *et al*., 2006). Earlier onset of alcoholism and AD with the low activity allele (3-repeat) was also observed (Contini *et al*., 2006). No association exists between *MAO-A* and alcoholism, with or without antisocial behaviour, among Chinese males (Lu *et al*., 2003). In another study testosterone analogues interacted with the *MAO-A* promoter *in vitro* and *in vivo*, influencing gene transcription as well as the cerebrospinal fluid levels of the MAO metabolite 3-methoxy-4-hydroxyphenylglycol (MHPG) (Sjöberg *et al*., 2008). Cerebrospinal testosterone interacts with the *MAO-A*-uVNTR to predict antisocial behaviour (Sjöberg *et al*., 2008).

A gene-environment interaction occurs between the *MAO-A*-uVNTR and alcoholism (Ducci *et al*., 2008). Sexually abused homozygous 3-repeat allele females had higher levels of antisocial personality disorder and alcoholism than women with other analysed genotypes. Heterozygous females displayed an intermediate risk pattern; no association exists in non-abused females (Ducci *et al*., 2008). A gene-gene interaction was established between the *MAO-A*-uVNTR and the 5-HTTLPR in relation to drinking behaviour among university female subjects (Herman *et al*., 2005). The highest risk for drinking occurred when a female was homozygous for the S allele of 5-HTTLPR and had one of the high activity *MAO-A* alleles (3.5- or/and 4-repeat). High activity *MAO-A* alleles along with the L allele of 5-HTTLPR were determined among females with the lowest risk (Herman *et al*., 2005).

Among Han Chinese men a gene-gene interaction occurred for the *MAO-A*-uVNTR and the *Dopamine Receptor D2* (*DRD2*). Individuals with the 3-repeat *MAO-A* allele and the A1/A1 genotype of *DRD2* were 3.48 times more likely to

suffer from AD, anxiety and depression in comparison to individuals that have the *DRD2* A2/A2 genotype (Huang *et al*., 2007). It was concluded that *MAO-A*-uVNTR may modify the association between *DRD2* and AD, anxiety and depression (Huang *et al*., 2007).

The 3-repeat allele of the *MAO-A*-uVNTR has been associated with higher reward dependence (Samochowiec *et al*., 2004). A gene-gene interaction between the *MAO-A*-uVNTR and the *Norepinephrine Transporter* (*NET*) exist and effects novelty seeking scores (Lee *et al*., 2008). The 4-repeat allele of the *MAO-A*-uVNTR has been associated with neuroticism in males (Eley *et al*., 2003). However, in additional studies (Swedish and Japanese populations) no association was observed with neuroticism (Garpenstrand *et al*., 2002; Hakamata *et al*., 2005; Urata *et al*., 2007). A slight gene-gene interaction occurs for the *MAO-A*-uVNTR, *Catechol-O-Methyl Transferase* (*COMT*) and the *Dopamine Receptor D3* (*DRD3*) in association with agreeableness (Urata *et al*., 2007).

Serotonin Transporter

Transportation of serotonin is performed by the 5-HTT, a highly selective chloride and sodium dependent transporter. Expression of the 5-HTT occurs in the brain, gastrointestinal cells and blood. The 5-HTT is encoded by a single gene, the *Solute Carrier Family 6 Member 4* (*SLC6A4*) and maps to chromosome 17q11.1-17q12 (Ramamoorthy *et al*., 1993; Lesch *et al*., 1996; Murphy *et al*., 2004) (Figure 2.11).

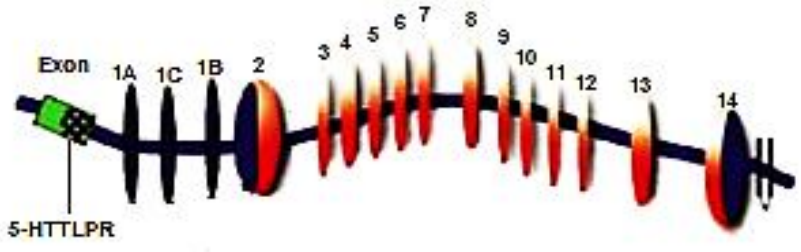

Figure 2.11 A schematic representation of the *5-HTT* gene. Coding exons (orange), intronic and non-coding areas (blue) are indicated along with the 5-HTTLPR (Murphy *et al*., 2004).

A number of polymorphisms exist in the coding region of the *5-HTT* gene, including the mentioned 5-HTTLPR (Ramamoorthy *et al*., 1993; Heils *et al*., 1995; Ogilvie *et al*., 1996; Nakamura *et al*., 2000; Weese-Mayer *et al*., 2003; Hu *et al*., 2006). Approximately 1.4 kb upstream from the *5-HTT* gene transcription site the 5-HTTLPR consists of various lengths of repetitive GC rich sequences (Ramamoorthy *et al*., 1993). The 5-HTTLPR is characterised by a 44 bp deletion or insertion commonly resulting in either an S or L allele (Figure 2.12). The S allele consists of 14 complex repeat sequences of 20-23 bp whereas the L allele consists of 16 (Lesch *et al*., 1996).

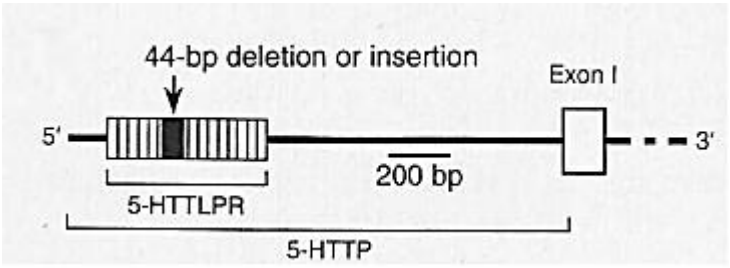

Figure 2.12 A schematic representation of the 5-HTTLPR. The 44-bp is indicated within the 5-HTTLPR (Lesch *et al*., 1996).

Five rare 5-HTTLPR alleles (15-, 18-, 19-, 20- and 22-repeat elements of 20-23 bp) were identified (Kunugi *et al*., 1997; Michaelovsky *et al*., 1999; Nakamura *et al*., 2000). Four 14-repeat and six 16-repeat allele variants with different sequence variations were also identified and named 14-A, 14-B, 14-C, 14-D, 16-A, 16-B, 16-C, 16-D, 16-E and 16-F (Nakumara *et al*., 2000). Low allelic frequencies were observed for the 14-B, 14-C, 14-D, 16-E and 16-F in Caucasian and Japanese populations. The 15-, 19-, 20- and 22-repeat alleles were only present in a Japanese population (Nakumara *et al*., 2000).

Functional variation exists among the 5-HTTLPR alleles. The S allele reduces the transcriptional efficiency of the promoter of the *5-HTT* gene and 5-HTT expression. Reduced serotonin uptake also occurred in lymphoblasts due to the S allele. Higher rates of serotonin uptake and higher levels of 5-HTT mRNA are observed in lymphoblasts of L homozygous individuals (Heils *et al*., 1996; Lesch *et al*., 1996; Greenberg *et al*., 1999).

A substitution (A→G) within the L allele of the 5-HTTLPR was determined (Heils *et al*., 1995; Nakamura *et al*., 2000; Hu *et al*., 2006) (Figure 2.13). The A→G substitution at nucleotide 1 273 of the *5-HTT* coding sequence results in Val425 being substituted for Ile425 (Ozaki *et al*., 2003). Both the L_A and L_G alleles consist of 16 complex repeat sequences. However, the L_G allele similarly to the S allele has reduced expression levels in comparison to the wild type L_A allele. Among Caucasian individuals with OCD, the L_AL_A genotype was the most common (Hu *et al*., 2006).

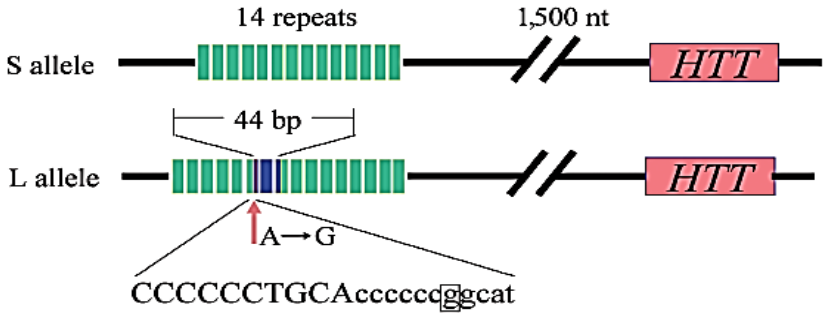

Figure 2.13 A schematic representation of the A to G substitution in the 5-HTTLPR (Hu *et al*., 2006).

A VNTR containing 9, 10 or 12 copies of a 16-17 bp repeat sequence was observed in intron II of the *5-HTT* (Figure 2.14). This particular polymorphism regulated gene expression (Ogilvie *et al*., 1996; Weese-Mayer *et al*., 2003). The allele with 9 copies of the repeat sequence is associated with unipolar disorder. The 9-repeat allele confers susceptibility to unipolar depression in more than 10% of individuals that are affected (Olgivie *et al*., 1996). The *5-HTT* intron II polymorphism is also associated with sudden infant death syndrome (SIDS) (Weese-Mayer *et al*., 2003). In a cohort of 90 cases, a significant association existed between SIDS and the 12-repeat allele among African Americans (Weese-Mayer *et al*., 2003).

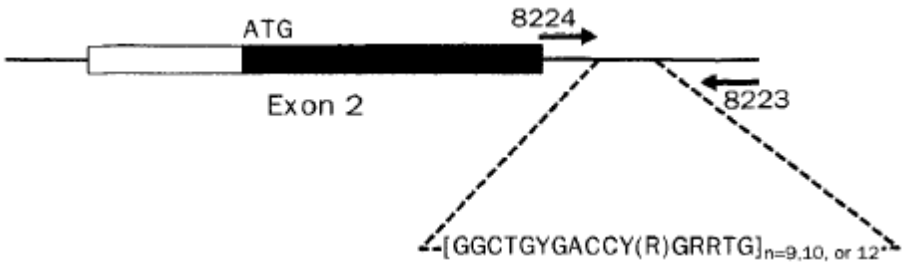

Figure 2.14 A schematic representation of the polymorphic VNTR in intron II of the *5-HTT* (Ogilvie *et al*., 1996).

Autism, AD, anxiety disorder, bipolar disorder, unipolar disorder and depression are associated with the 5-HTTLPR in varying degrees (Hariri *et al*., 2002; Feinn *et al*., 2005; Lasky-Su *et al*., 2005; Arieff *et al*., 2010; Clarke *et al*., 2010; Wang *et al*., 2011). The S allele of 5-HTTLPR, in particular, has been associated with autistic individuals from South Africa (Arieff *et al*., 2010). Among autistic participants the homozygous S genotype was determined at a frequency of 0.330 whereas the particular genotype was not present among the control subjects. The odds of developing autism with the S genotype compared to the L homozygous genotype increased 2.74-fold in a Caucasian population and 10.15-fold in a Coloured population (Arieff *et al*., 2010). Research has determined that autistic individuals with the S allele have increased frontal lobe grey matter volume (Wassink *et al*., 2007).

No association between the 5-HTTLPR and autism was observed in a separate study (Yonan *et al*., 2006). High magnesium concentrations can affect the amplification of the L allele in polymerase chain reactions (PCR) which possibly results in unwanted bias (Yonan *et al*., 2006). A meta-analysis of 16 studies showed no significant association between the 5-HTTLPR polymorphism and autism (Huang & Santangelo, 2008).

Alcohol dependence was associated with the 5-HTTLPR (Wang *et al*., 2011). Polymorphisms in the *Alcohol Dehydrogenase* (*ADH1B*), *Aldehyde Dehydrogenase* (*ALDH2*), *Cytochrome P-4502E1* (*CYP4502E1*), *COMT* and *5-HTT* were associated with AD in a Yunnan Han population from China (Wang *et al*., 2011). An association between AD and the 5-HTTLPR was found in a meta-

analysis comprising 17 published articles, including 3 489 alcoholics and 2 325 controls (Feinn *et al*., 2005). The S allele was significantly associated with AD. Individuals with early onset AD have more severe AD subtypes or a co-morbid psychiatric condition (Feinn *et al*., 2005). The homozygous L genotype reduced AD (Wang *et al*., 2011).

Ondansetron (originally marketed as Zofran) reduced alcohol consumption among AD individuals homozygous for the L allele (Kenna *et al*., 2009). Late onset or low vulnerability alcoholics reported fewer drinking and heavy drinking days after treatment with Sertraline. Individuals with early onset or high vulnerability alcoholics had fewer drinking and heavy drinking days when administered a placebo (Kranzler *et al*., 2011). Sertraline did not reduce alcohol use among individuals with either S/S or S/L genotypes of the 5-HTTLPR (Kenna *et al*., 2009). The homozygous L genotype was identified as the genotype with the greatest potential for a 5-HTT blockade (Kranzler *et al*., 2011). Dantron is the brand name used for Ondansetron in South Africa. A derivative of Ondansetron, odansetron hypochloride dehydrate is used in South Africa. Sertraline have the trade names Zoloft and Lustral. Adco-Zerta is Sertraline's trade name in South Africa.

Anxiety-related personality traits are associated with the 5-HTTLPR polymorphism (Hariri *et al*., 2002). The 5-HTTLPR accounts for about 3-4% of total variation and 7-9% for inherited variation for anxiety-related personality traits (Lesch *et al*., 1996). The association between 5-HTTLPR and anxiety-related personality traits was examined in a meta-analysis of 23 studies (Sen *et al*., 2004). The study suggested an association between increased anxiety-related personality scores and the S allele.

The relationship between two *5-HTT* polymorphisms, unipolar and bipolar disorder was investigated in a meta-analysis study (Lasky-Su *et al*., 2005). The polymorphisms analysed were the 5-HTTLPR and the 17 bp VNTR in intron II of the *5-HTT*. A significant correlation exists between the 5-HTTLPR and bipolar disorder. However, no significant association was found for the 17 bp VNTR (Lasky-Su *et al*., 2005). A small correlation between the 5-HTTLPR and unipolar depression exist (Clarke *et al*., 2010). The association remained significant as long as the samples (European and East Asian individuals) were analysed

separately; evidence for significant between-study heterogeneity. The relationship between bipolar disorder or suicidal behaviour and the 5-HTTLPR was investigated (Neves *et al*., 2008). The S allele was associated with more frequent violent suicide attempts. In a separate study, no association was found between the 5-HTTLPR and major depression or bipolar disorder in Brazilian patients (Oliveira *et al*., 2000).

The relationship between the 5-HTTLPR and hostile/aggressive traits was examined in depressed patients (Gonda *et al*., 2011). Women with the S allele had a greater risk for depression and aggression/hostility. Association was more profound in the presence of depression. The homozygous S genotype increased the risk for developing depression (Cervilla et al., 2006; Gonda et al., 2011).

Internalising and externalising behaviours are associated with the 5-HTTLPR. Caucasian children with one or two copies of the 5-HTTLPR S allele and a long (6- to 8-repeats) *DRD4* VNTR allele exhibited more internalising and externalising behaviour (Schmidt *et al*., 2007). The 5-HTTLPR L allele serves as a protective factor against internalising and externalising behaviour when joined with long (6- to 8-repeats) *DRD4* VNTR alleles (Schmidt *et al*., 2007). Children with the 5-HTTLPR S allele tend to show more attachment disorganisation when raised in institutions compared to children homozygous for the 5-HTTLPR L allele. A definite protective effect exists for the 5-HTTLPR L allele (Bakermans-Kranenburg *et al*., 2011).

The interaction between the S allele of the 5-HTTLPR and depression was investigated in a Pomeranian population (Grabe et al., 2005). A significant interaction exists between chronic disease, genotype and unemployment among female subjects. No independent association of the genotype with physical and mental health was observed (Grabe *et al*., 2005).

The 5-HTTLPR moderated the relationship between depression and stress (meta-analysis consisting of 54 studies and 40 749 subjects) (Karg *et al*., 2011). The S allele was associated with an increased risk for the development of depression under stress. An association exists between increased stress sensitivity due to childhood maltreatment in the presence of the S allele (Karg *et al*., 2011).

The thamlus and subregions of the brain were investigated in order to determine size differences among individuals having major depressive disorder in relation to the 5-HTTLPR. The 5-HTTLPR and major depressive disorders were independent factors contributing to the enlargement of the thalamus. The homozygous S genotype was linked to an enlarged pulvinar area (Young *et al*., 2008). The relationship between depression, antidepressants and 5-HTTLPR were investigated. Individuals with the L allele had an improved response to Escitalopram compared to individuals with the S allele. Males homozygous for the L allele in particular responded better to Escitalopram than males homozygous for the S allele (Huezo-Diaz *et al*., 2009). Escitalopram has the common brand names Lexapro and Cipralex, in South Africa it is sold under the name Lexamil. The active ingredient is escitalopram oxalate.

Summary

The literature review revealed that the *DRD4*, *MAO-A* and *5-HTT* (5-HTTLPR) have a substantial effect on human behaviour. Behavioural association studies based on African populations were limited. Psychiatric genetic research in relation to South Africa populations was identified as a particular area lacking information. The population genetics field was identified as a potential beneficiary of a study on South African ethnic groups. Research on the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR will provide information in areas that desperately require impetus.

South Africa has a great amount of ethnic diversity. It is therefore vital that South African populations are genetically analysed and more importantly to analyse the three mentioned neurotransmitter genes. The Khwe is related to some of the oldest representatives of the human race and it would be interesting to determine its genetic variation and to identify possible novel alleles. Admixture would be present for the Khwe in particular due to its relation with Bantu and other Khoe-San populations. Studies performed on the South African Bantu populations indicate low genetic difference estimates within group differences which suggest that Bantu populations arose from a common recent ancestral population. However, the amount of research information regarding the three mentioned

neurotransmitter genes in the South African Bantu populations is limited.

The Afrikaner is an amalgamated European population and has genetic contributions from African and Asian populations. Genetic studies performed on Caucasians with the use of the mentioned neurotransmitter genes have revealed association with some behavioural traits and therefore holds promise for the Afrikaner. Due to its small founder population, the Afrikaner population's genetic diversity is expected to be lower than that of European populations. A study that could bring more clarity to the amount of genetic diversity among South African ethnic groups would be informative for the rest of the scientific community especially those in behavioural, medical and population genetics fields.

CHAPTER THREE

ALLELIC FREQUENCIES OF THREE NEUROTRANSMITTER GENES IN FOUR SOUTH AFRICAN ETHNIC GROUPS

Abstract

South African ethnic groups are ideal candidates for genetic studies due to their large diversity and unique composition. The genetic characterisation of South African ethnic groups with neurotransmitter genes such as the *DRD4*, *MAO-A* and *5-HTT* has not been performed. This chapter reports on a genetic analysis of samples from the Khwe, Xhosa, Sotho and Afrikaner ethnic groups. Participants provided saliva samples from which DNA was isolated. The *DRD4* VNTR, *MOA-A*-uVNTR and 5-HTTLPR were successfully studied in 349 samples. The Xhosa and Sotho populations correlated the strongest with each other. The Khwe also associated with the Xhosa and Sotho but with a lower correlation. The Khwe, Xhosa and Sotho correlated with previously analysed African populations revised in literature. The Afrikaner population correlated with European populations.

Introduction

Genes, such as the *DRD4*, *MAO-A* and *5-HTT,* all effect neurotransmission and consequently human behaviour. Population analyses of genes are beneficial in providing information on population structures (Chang *et al*., 1996; Lesch *et al*., 1996; Sabol *et al*., 1998; Chen *et al*., 1999; Mansoor *et al*., 2008). The genetic characterisation of populations is essential to determine population structure and to provide information retrospectively on a population's current and past circumstances (Pritchard *et al*., 2000; Reed & Tishkoff, 2006; Campbell & Tishkoff, 2008; Wright *et al*., 2011).

The neurotransmitter, dopamine, influences mental and physiological functions including cognition, emotion, endocrine regulation, food intake, locomotor activity and positive reinforcement (Missale *et al*., 1998). One of the most important genes in the dopaminergic system is the *DRD4* gene which has a 48 bp VNTR sequence in its exon III. The alleles found at this particular VNTR range from 2 to

11 complex repeats of 48 bp (Van Tol *et al*., 1991; 1992; Mansoor *et al*., 2008). The 4-repeat allele of the VNTR is the most common in human populations followed by the 7- and 2-repeat alleles. The 4-repeat allele is prevalent in African (0.779) and European/Middle Eastern (0.709) populations (Chang *et al*., 1996). Native North (0.293) and South American (0.663) populations have substantial 7-repeat allele frequencies (Chang *et al*., 1996).

Gene expression variation exists among the alleles of the *DRD4* VNTR. The 7-repeat allele is associated with a loss of *DRD4* mediated prefrontal inhibition in comparison to the 4-repeat allele (Asghari *et al*., 1995; Wang *et al*., 2004). Due to functional variation, the *DRD4* VNTR is associated with behavioural disorders such as alcoholism, OCD and schizophrenia (Muramatsu *et al*., 1996; Cruz *et al*., 1997; Jardemark *et al*., 2002; Lung *et al*., 2002; Millet *et al*., 2003; Lahti *et al*., 2005; Camarena *et al*., 2007; Ray *et al*., 2008; Creswell *et al*., 2012; Gonçalves *et al*., 2012). The *DRD4* VNTR is also associated with novelty seeking behaviours including risk taking and migratory behaviour (Ebstein *et al*., 1997; Chen *et al*., 1999; Tomitaka *et al*., 1999; Kluger *et al*., 2002; Reist *et al*., 2007; Dreber *et al*., 2009; Kuhnen & Chiao, 2009; Matthews & Butler, 2011).

Monoamine Oxidase A, similarly to dopamine, is involved in neural regulation (Knoll & Magyar, 1972; Bach *et al*., 1988; Sabol *et al*., 1998; Shih & Thompson, 1999; Jacob *et al*., 2005). Neurotransmitter regulation with the MAO-A enzyme allows the termination of dopamine, norepinephrine and serotonin (Shih & Thompson, 1999; Buckholtz & Meyer-Lindenberg, 2008). Monoamine Oxidase A activity is influenced by a polymorphism upstream of the *MAO-A* coding sequence. The particular polymorphism consists of a 30 bp VNTR sequence present in 2, 3, 3.5, 4, 5 or 6 repeat allele copies (Sabol *et al*., 1998; Huang *et al*., 2004; Guo *et al*., 2008). Among Caucasians the 4-repeat allele frequency (0.648) of the uVNTR is twice as common as the 3-repeat allele frequency (0.331). African Americans have a higher 3-repeat allele frequency (0.591) than 4-repeat allele frequency (0.364) (Sabol *et al*., 1998).

The 3.5- and 4-repeat alleles of the *MAO-A*-uVNTR transcribe 2 to 10 times more efficiently in comparison to the other alleles (Sabol *et al*., 1998) resulting in a number of behavioural association studies being performed (Cohen *et al*., 2011;

Ferguson *et al*., 2011; Pardini *et al*., 2011). The *MAO-A*-uVNTR is linked to disorders such as aggression, alcoholism, antisocial behaviour, anxiety and autism (Deckert *et al*., 1999; Samochowiec *et al*., 1999; Schmidt *et al*., 2000; Caspi *et al*., 2002; Yirmiya *et al*., 2002; Cohen *et al*., 2003; Jones *et al*., 2004; Maron *et al*., 2004; Samochowiec *et al*., 2004; Herman *et al*., 2005; Contini *et al*., 2006; Huang *et al*., 2007; Ducci *et al*., 2008; Philibert *et al*., 2008; Sjörberg *et al*., 2008; McDermott *et al*., 2009; Cohen *et al*., 2011; Ferguson *et al*., 2011; Pardini *et al*., 2011). The *MAO-A*-uVNTR is also associated with traits such as neuroticism and novelty seeking (Eley *et al*., 2003; Samochowiec *et al*., 2004; Lee *et al*., 2008)

Neurotransmission mediated by serotonin and regulated by the 5-HTT also contributes to physiological functions and emotional states (Lesch *et al*., 1996; Missale *et al*., 1998). A polymorphism upstream of the *5-HTT* coding sequence, the 5-HTTLPR affects the efficiency of the 5-HTT during neurotransmission. The particular polymorphism is due to a 44 bp deletion or insertion characterised by either an S or L allele. The L allele consists of 16 complex repeat sequences of 20-23 bp whereas the S allele consists of 14 (Heils *et al*., 1995; 1996; Lesch *et al*., 1996). Less common variants [very long (VL) and extra-long (XL) alleles] also exist (Kunugi *et al*., 1997; Michaelovsky *et al*., 1999; Nakamura *et al*., 2000; Hu *et al*., 2006; Ehli *et al*., 2012).

The S allele of the 5-HTTLPR results in lower 5-HTT expression in membranes and 5-HTT mRNA expression in comparison to the L allele (Heils *et al*., 1995; 1996; Lesch *et al*., 1996). In addition, the S allele also contributes to less serotonin reuptake according to single photon emission computed tomography of the brain (Heinz *et al*., 2000). A SNP within the L allele (A→G) results in an allele (L_G) with expression levels closer to that of the S allele (Ozaki *et al*., 2003; Ehli *et al*., 2012). As a result of functional variation, the 5-HTTLPR is associated with behavioural disorders such as AD, anxiety, autism, bipolar disorder and depression (Lesch *et al*., 1996; Bellivier *et al*., 1998; Hariri *et al*., 2002; Caspi *et al*., 2003; Sen *et al*., 2004; Feinn *et al*., 2005; Grabe *et al*., 2005; Lasky-Su *et al*., 2005; Brune *et al*., 2006; Cervilla *et al*., 2006; Wassink *et al*., 2007; Neves *et al*., 2008; Young *et al*., 2008; Arieff *et al*., 2010; Clarke *et al*., 2010; Gonda *et al*., 2011; Karg *et al*., 2011; Wang *et al*., 2011).

The 5-HTTLPR L allele is generally more prevalent among Caucasian (0.572-0.580) and African (~0.838) populations (Gutiérrez *et al*., 1998; Szekely *et al*., 2004; Esau *et al*., 2008). Asian populations (0.698-0.785) have very high S allele frequencies (Kunugi *et al*., 1997; Esau *et al*., 2008). Due to its global variation the 5-HTTLPR provides information on etiological and evolutionary theories of humans. Within the pharmacology field possible population specific medication could be developed with regards to the 5-HTTLPR in the future (Gelernter *et al*., 1999; Pickar & Rubinow, 2001; Noskova *et al*., 2008; Matthews & Butler, 2011; Lazary *et al*., 2011). Genetic research on the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR can possibly reduce ADR levels among South African populations.

The high levels of neuropsychiatric disorders in South Africa and the inadequate amount of studies on behavioural and population fields need addressing (Tishkoff & Williams, 2002; Norman *et al*., 2006; Reed & Tishkoff, 2006; Campbell & Tishkoff, 2008; Wright *et al*., 2011). Studies on the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR indicate relevance in the fields of psychology, physiology and pharmacology and can thereby contribute significant information (Mongeau *et al*., 1997; Jardemark *et al*., 2002; Aguirre-Samudio & Nicolini, 2005; Pardini *et al*., 2011; Serretti *et al*., 2011; Kenna *et al*., 2012; Colucci *et al*., 2013). Since allelic variations of these genes have important implications the acquirement of such relevant data is important.

The aim of the chapter is to determine the genetic diversity of the Khwe, Xhosa, Sotho and Afrikaner ethnic groups of South Africa by determining the allelic frequencies of the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR for each ethnic group.

Materials and methods

Study populations

Information leaflets and consent forms in lay terms were distributed to potential participants [Appendix A1 (English) and A2 (Afrikaans)]. Basic participant information (e.g. sex, age) along with written consent was obtained. The study

was performed anonymously. Participants were informed that they can withdraw from the study at any time without consequences.

A cohort of 349 individuals from the Khwe, Xhosa, Sotho and Afrikaner ethnic groups were included in the study. Participants belonged to the third generation (all grandparents from the same ethnic group) of a particular ethnic group or greater. Saliva samples were collected in 15 or 50 millilitre (ml) sterile tubes from participants. Khwe individuals were sampled in the Northern Cape Province. Xhosa individuals were sampled in the Eastern Cape and Free State Provinces. Sotho individuals were sampled in the Free State as were the Afrikaner individuals. Ethical clearance for the study was obtained from the Medical Ethics Research Committee of the University of the Free State (ECUFS 152/2011).

Extraction protocols

Deoxyribonucleic acid was isolated from samples either by using an adapted protocol described in Quinque *et al*. (2006) or using the AquaGenomic™ (MultiTarget Pharmaceuticals, Salt Lake City, Utah) extraction method and protocol. Modifications to Quinque *et al*. (2006) involved the addition of 20 microlitres (µl) [0.1 milligram (mg)/µl] of Proteinase K (Roche Germany) instead of 30 µl (0.02 mg/µl). Centrifugation speed for the 15 and 50 ml tubes were adapted to 3 300 relative centrifugal force (RCF) for 15 minutes instead of the prescribed speed and duration mentioned in Quinque *et al*. (2006). Subsequent centrifugation steps took place at 18 440 RCF for 5 minutes. A pellet wash step was included. Samples were placed on an orbital shaker [99 revolutions per minute (rpm) for 30 minutes] and washed with 70% ethanol.

Modifications to the AquaGenomic™ protocol were made with regards to the speed of the centrifugation steps. The first 2 centrifugation steps were adapted to 3 300 RCF for 15 minutes. The remaining centrifugation steps were performed at 18 440 RCF for 5 minutes. An incubation step [30 minutes at 75 degree Celsius (°C)] was included and followed the addition of the AquaGenomic™ solution to the samples. A pellet wash step was also included. The samples were placed on an orbital shaker (99 rpm for 30 minutes) and washed with 70% ethanol. The rest of

the steps were performed in accordance to the AquaGenomic™ protocol. Samples were resuspended in T.$_1$E (10 millimolar (mM) Tris-HCl pH 7.6, 0.1mM EDTA) elution buffer in both extraction methods. Deoxyribonucleic acid quantification of samples was performed by using the Thermo Scientific NanoDrop™ Lite spectrophotometer (Thermo Scientific). The average of 2 readings were taken and used to dilute DNA samples to ~100 nanogram (ng) per microlitre (ng/µl).

Genotyping

Polymerase chain reaction was performed on the samples by making use of the G-Storm GS1 thermalcycler (Gene Technologies Ltd.). Primers used for amplification were all prepared by Whitehead Scientific. The primer sequences (5'-GCGACTACGTGGTCTACTCG-3' and 5'-AGGACCCTCATGGCCTTG-3') from Dreber *et al*. (2009) were used to amplify the *DRD4* VNTR; 5'-ACAGCCTGACCGTGGAGAAG-3' and 5'-GAACGGACGCTCCATTCGGA-3' for *MAO-A*-uVNTR (Sabol *et al*., 1998); 5'-GGCGTTGCCGCTCTGAATTGC-3' and 5'-GAGGGACTGAGCTGGACAACCCAC-3' for 5-HTTLPR (Heils *et al*., 1996).

For the amplification of *DRD4* VNTR, 10 µl KAPA HiFi HotStart DNA Polymerase (KAPA Biosystems, South Africa), 0.43 mM forward and reverse primer (final concentration), 5.75 µl nuclease free H_2O (Thermo Scientific, Fermentas™), 1.25 µl dimethyl sulfoxide (DMSO) and 17.39 ng/µl of DNA (final concentration) were required in a total volume of 23 µl. Reactions for the *MAO-A*-uVNTR samples were performed in a total volume of 20 µl consisting of 10 µl DreamTaq PCR Master Mix 2x (Thermo Scientific, Fermentas™), 0.25 mM forward and reverse primer (final concentration), 7 µl nuclease free H_2O (Thermo Scientific, Fermentas™) and 10 ng/µl DNA (final concentration). The reagents required to amplify 5-HTTLPR were 10 µl KAPA HiFi HotStart DNA Polymerase (KAPA Biosystems, South Africa), 0.22 mM forward and reverse primer (final concentration), 6.75 µl nuclease free H_2O (Thermo Scientific, Fermentas™), 1.25 µl DMSO and 17.39 ng/µl of DNA (final concentration) having a total volume of 23 µl.

The *DRD4* VNTR thermal cycling conditions [modified from Dreber *et al*. (2009) and Bachmann *et al*. (2003)] made use of an initial activation step of 98°C for 5 minutes followed by 3 cycles of 98°C (30 seconds), 68°C (30 seconds) and 72°C (40 seconds). Subsequent cycles made use of different annealing temperatures [3 cycles of 67°C (30 seconds), 3 cycles of 66°C (30 seconds), 3 cycles of 65°C (30 seconds), 3 cycles of 64°C (30 seconds) and 15 cycles of 66°C (30 seconds)]. A final extension step at 72°C for 5 minutes concluded the reaction. Thermal cycling conditions for *MAO-A*-uVNTR [modified from de Milander *et al*. (2009)] included an initial activation step of 95°C (3 minutes) followed by 30 cycles of 95°C (1 minute), 66°C (30 seconds) and 72°C (90 seconds). A final extension step of 72°C for 10 minutes concluded the reaction. The ramp temperature for *MAO-A*-uVNTR was set to 3°C per second for all temperature steps. The cycling conditions for 5-HTTLPR [modified from Heils *et al*. (1996)] included an initial activation at 95°C (7 minutes) followed by 3 cycles of 98°C (30 seconds), 63°C (30 seconds) and 72°C (1 minute). The first 3 cycles were followed by: 25 cycles of 98°C (30 seconds), 61°C (30 seconds) and 72°C (1 minute). A final extension step at 72°C for 10 minutes concluded the reaction. The 5-HTTLPR and *DRD4* VNTR ramp temperatures for the denaturation and annealing steps were set to 2.5°C and 1.5°C per second (Bachmann *et al*., 2003).

Gel electrophoresis and sequencing

The *DRD4* VNTR and 5-HTTLPR PCR products were separated on 2% agarose gels, which were prepared from Seakem® LE Agarose (Lonza) and Tris/acetic acid/EDTA (TAE). *Monoamine Oxidase A*-uVNTR PCR products were separated on either 8% polyacrylamide gels (PAGE) or 2% agarose gels. The 8% PAGE gels were required to discriminate between alleles having very similar bp sizes. Polyacrylamide gels were prepared by using 5.9 ml distilled H_2O, 2 ml 5 x Tris/Borate/EDTA (TBE), 2 ml of 40% acrylamide, 100 µl of 10% ammonium persulfate (APS) and 10 µl of tetramethylethylenediamine (TEMED). A 50 bp ladder (Thermo Scientific) was run along with the samples in order to differentiate between variant alleles. Bands were visualised with GelRed™ and bromophenol blue.

Polymerase chain reaction product clean-up was performed with either the Biospin PCR Purification Kit (BioFlux), NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel), GeneJET™ Gel Extraction and DNA Cleanup Micro Kit (Thermo Scientific) or the DNA Clean & Concentrator™ (Zymo Research). Sequence reactions were performed by making use of the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems®). For the *DRD4* VNTR and 5-HTTLPR sequencing reactions, 0.8 mM either forward or reverse primer (final concentration), 6.55 µl nuclease free H_2O (Thermo Scientific, Fermentas™), 0.75 µl Ready Reaction Premix (Applied Biosystems®), 1 µl 5 x Sequencing Buffer (Applied Biosystems®), 0.4 µl DMSO and 1 µl of PCR product were required. Sequence reactions for *MAO-A*-uVNTR samples were performed in total volumes of 10 µl with each reaction containing either 3.2 mM forward or reverse primer (final concentration), 4.3 µl nuclease free H_2O (Thermo Scientific, Fermentas™), 0.5 µl Ready Reaction Premix (Applied Biosystems®), 1 µl 5 x Sequencing Buffer (Applied Biosystems®) and 1 µl of PCR product. Variation in primer and PCR product concentration were necessary for some samples. The BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems®) prescribed protocol conditions was used for amplification.

Post sequencing PCR clean-up was performed in a 96-well plate by making use of an EDTA/ethanol precipitation protocol provided by the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems®). Centrifugation speeds for this particular protocol were modified. The first and third centrifugation steps speeds were adapted to 10 545 RCF with durations of 60 and 30 minutes. The second and fourth centrifugation steps were adjusted to 981 RCF for several seconds. Following clean-up, samples were analysed with the ABI 3130 (Applied Biosystems®) DNA sequencer.

The number of repeats for the *DRD4* VNTR gene were considered to be the number of times the 48 bp complex repeat units were present in the sequenced area of each of the analysed homozygous samples (Ding *et al*., 2002; Grady *et al*., 2003). Similarly the 20-23 bp complex repeat sequences of the 5-HTTLPR (Ehli *et al*., 2012) were determined. The number of times the 30 bp repeat sequence of the *MAO-A*-uVNTR (Gallardo-Pujol *et al*., 2013) occurred in a sequence was also counted for each homozygous or hemizygous sample.

Selected alleles of heterozygous samples were sequenced when no homozygous or hemizygous (in the case of *MAO-A*) representative alleles were available for the 3 analysed genes. Heterozygous sequencing required the cutting out of a PCR product from either an agarose or low temperature melting agarose gel. The extraction of PCR products from excised bands were performed by making use of either the NucleoSpin® Gel PCR Clean-up Kit (Macherey-Nagel) or GeneJET™ Gel Extraction and DNA Cleanup Micro Kit (Thermo Scientific). The excised PCR fragments were sequenced in similar fashion as the homozygotes. All sequencing results were viewed and edited by making use of purchased Geneious 6.1.6 software (Biomatters Ltd.). Allele frequencies, heterozygosities and the HWE were determined by Arlequin 3.5.1.2 (Laurent Excoffier). For the HWE the number of steps in the Markov chain was set at a 1 000 000 and the number of dememorisation steps was set at a 100 000.

Results and discussion

From the 349 initial DNA samples 200 *DRD4* VNTR, 255 *MAO-A*-uVNTR and 231 5-HTTLPR samples were successfully analysed (Table 3.1). A total of 686 PCR samples were successfully analysed. The Xhosa population contributed the most individuals whereas the Khwe contributed the least. The sampled populations were in HWE for the majority of the genes, using Arlequin 3.5.1.2 (Appendix B). The HWE test was not performed on the male *MAO-A* samples due to the presence of the gene on the X-chromosome; male and female *MAO-A* samples were grouped separately.

Table 3.1 The number of successfully analysed *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR samples for each ethnic group.

Ethnic group:	Number of samples successfully analysed for each gene:				Total number:
	DRD4	*MAO-A* male	*MAO-A* female	5-HTTLPR	
Khwe	45	16	28	46	135
Xhosa	58	28	43	63	192
Sotho	49	21	44	69	183
Afrikaner	48	34	41	53	176

Dopamine Receptor D4

The 2-, 3-, 4-, 5-, 6-, 7-, 8- and 10-repeat *DRD4* alleles were identified on 2% agarose gels (Figure 3.1) following sequencinig. A variant (4-repeat) identified on a 2% agarose gel (Figure 3.1) was compared to a reference *DRD4* 4-repeat sequence (HM191418) from GenBank (Figure 3.2) confirming that the fragments are indeed *DRD4* VNTR alleles. The order of the variant 48 bp repeat sequences differs for each of the *DRD4* alleles therefore in order to save space all variant alleles observed (2-, 4-, 5-, 6-, 7- and 8-repeat alleles) are provided in Appendix C. Two 2-repeats, six 4-repeats, one 5-repeat, one 6-repeat, six 7-repeats and one 8-repeat allele were successfully sequenced.

The sequencing of possible 3- and 10-repeat alleles was troublesome and no results were obtained in the current project, in spite of the use of various troubleshooting methods. Possible reason for sequence failure could be attributed to the high GC content of *DRD4* VNTR repeat sequences (Lichter *et al*., 1993; Choi *et al*., 1999). Another possible reason could be due to the alleles being present in only heterozygous genotypes (e.g. 3, 4; 4, 10) requiring a gel-cut method, resulting in decreased DNA concentration. The exposure of excised fragments to ultraviolet light during gel-cut out could also have resulted in sequence failure (Gründemann & Schömig, 1996). The combination of the mentioned factors possibly contributed to the overall sequence failure. However, with the use of a 50 bp ladder it was evident that the 3- and 10-repeat alleles were present in the populations due to the relation of the possible 3- and 10-repeat fragments with the other sequenced alleles. Identified 3- and 10-repeat fragments were included in the current study as 3- and 10-repeat alleles.

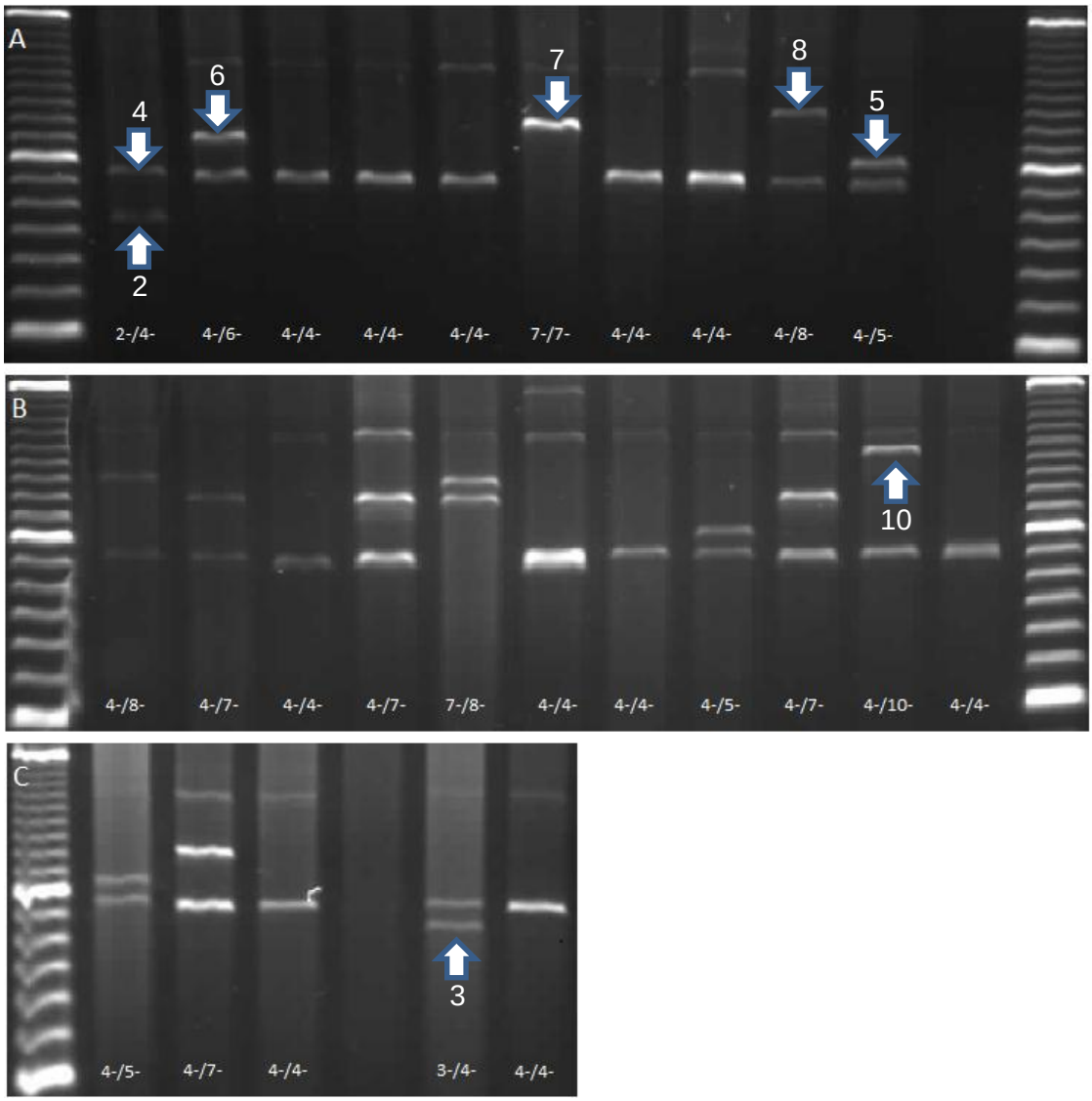

Figure 3.1 The visualisation of *DRD4* VNTR PCR products on 2% agarose gels, representing the 2-, 4-, 5-, 6-, 7-, 8- and possible 3- and 10-repeat alleles along with a 50 bp ladder. A) 2-, 4-, 5-, 6-, 7- and 8-repeats B) 10-repeat C) 3-repeat as indicated by arrows.

Among the analysed ethnic groups the 4-repeat allele of the *DRD4* VNTR was the most common with a frequency ranging from 0.586 to 0.711 (Table 3.2). The analysed population frequencies were similar to populations from the same geographic regions. The Khwe, Xhosa and Sotho correlated with African populations such as the Biaka, Mbuti (central Africa) and Falasha (eastern Africa) whereas the Afrikaner correlated with European (German and French) populations (Chang *et al*., 1996; Sander *et al*., 1997; Chen *et al*., 1999; Millet *et al*., 2003; Mansoor *et al*., 2008). Afrikaner frequencies for the *DRD4* 2-, 4- and 7-repeat

alleles were determined 0.104, 0.708 and 0.125 whereas German allele frequencies were 0.071, 0.701 and 0.152 (Sander *et al*., 1997). French allele frequencies for the 2-, 4- and 7-repeat alleles were 0.103, 0.722 and 0.008 (Millet *et al*., 2003). The similarity between the Afrikaner and European (German and French) populations were expected as the Afrikaner is of European ancestry (Davenport, 1978; Botha & Beighton, 1983).

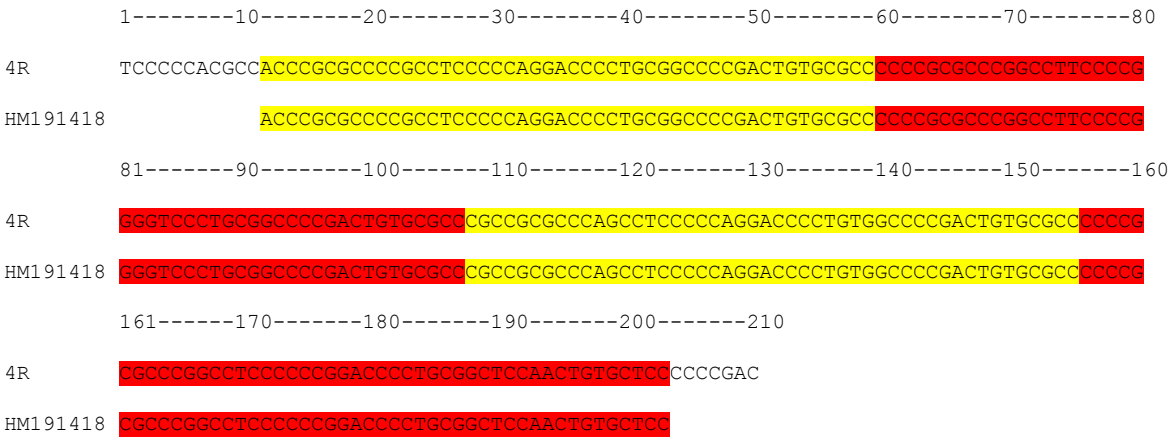

Figure 3.2 Nucleotide sequence alignment of the *DRD4* 4-repeat allele identified on the 2% agarose gel and a 4-repeat allele (Accession number HM191418) from GenBank, aligned with Geneious 6.1.6. The 48 bp repeat regions are consecutively highlighted yellow and red.

Following the 4-repeat allele, the 7- and the 2-repeat alleles were the most common *DRD4* VNTR alleles (Table 3.2). The Sotho population had the highest frequency (0.267) for the 7-repeat allele. This was not expected as the frequency of the 7-repeat allele among African populations (Biaka 0.140, Mbuti 0.160, San Bushmen 0, Bantu 0.190 and Falasha 0.110) analysed by Chang *et al*. (1996) ranged between 0-0.190. The Bantu population analysed by Chang *et al*. (1996) was from South Africa but was not subdivided into ethnic groups. The South African Bantu population analysed by Chang *et al*. (1996) had a similar 7-repeat allele frequency as the Xhosa and Khwe in the current study.

The Sotho migrated from a region in present day Cameroon into southern Africa via a south-eastern route (Davenport, 1978; Tishkoff *et al*., 2009). The association of the 7-repeat with migratory distance (Chang *et al*., 1996; Chen *et al*., 1999; Matthews & Butler, 2011) could be reason for the high frequency determined

among the Sotho population. However, this relationship is complex as the Xhosa population with a less prevalent 7-repeat allele frequency and greater 8-repeat allele frequency migrated a similar distance into southern Africa. This leads to the possibility that some sort of compensation occurred between the 7- and 8-repeat alleles (Chen *et al*., 1999). In a number of studies the 5- to 11-repeat alleles of *DRD4* are grouped together as long alleles and correlate with novelty seeking, schizophrenia and OCD (Chen *et al*., 1999; Lung *et al*., 2002; Camarena *et al*., 2007). Similarly the 5- to 11-repeat alleles may potentially correlate with migratory behaviour explaining why either the 7- or 8-repeat alleles were present at substantial frequencies for the Sotho and Xhosa.

Table 3.2. Allelic frequency of the *DRD4* VNTR in the analysed South African ethnic groups.

Ethnic group:	*DRD4* VNTR allele frequency:								Number of alleles (2n):
	2	3	4	5	6	7	8	10	
Khwe	0.056	0.000	0.711	0.011	0.033	0.156	0.000	0.033	90
Xhosa	0.031	0.000	0.673	0.031	0.020	0.174	0.061	0.010	98
Sotho	0.034	0.009	0.586	0.052	0.017	0.267	0.026	0.009	116
Afrikaner	0.104	0.010	0.708	0.000	0.031	0.125	0.021	0.000	96

The *DRD4* VNTR expected and observed heterozygosity were 0.470 and 0.578 for the Khwe, 0.515 and 0.510 for the Xhosa, 0.585 and 0.397 for the Sotho and 0.475 and 0.417 for the Afrikaner. The overall observed heterozygosity of *DRD4* is lower in comparison to STRs analysed in Balamurugan *et al*. (2010) and Liu *et al*. (2011). The average African *DRD4* observed heterozygosity (0.380) determined by Chang *et al*. (1996) is much lower than that of the Khwe (0.578). The Khwe population is the only population having a higher observed heterozygosity than expected heterozygosity for the *DRD4* gene in the current study which could be due to its distinct history.

A study by Chen *et al*. (2000) revealed that although the Khwe speak a Khoe-San language (de Almeida, 1965) they are more genetically related to Bantu populations than to other Khoe-San populations. It is proposed that the Khwe may have originated from the Bantu migration into southern Africa and may have

adopted the Khoe-San language upon contact with the Khoe-San populations (Chen *et al*., 2000). Possible admixture between the Khoe-San and Bantu groups may be the reason why the Khwe in the current study has such a high level of observed heterozygosity in comparison to its expected heterozygosity.

The HWE *P*-values were 0.775 for the Khwe, 0.501 for the Xhosa, 0.007 for the Sotho and 0.322 for the Afrikaner. The Sotho population was the only population not adhering to the HWE at the particular locus. Non-adherence is possibly due to factors including inbreeding, genotyping error, small sample population size and population stratification. In the absence of migration, mutation, natural selection, small population size and assortative mating, HWE will be adhered (Caballero & Hill, 1992; Wigginton *et al*., 2005; Chen *et al*., 2010). Population stratification, selection or/and genotyping errors are the likely reasons for the Sotho HWE non-adherence. Selection due to migration distance possibly had an effect, however this was not observed among the other populations. The Xhosa who migrated a similar distance as the Sotho into southern Africa adhered to the HWE. Genotyping error is another possible reason for non-adherence but not expected, identical procedures were utilised for each population. Duplicate analysis of samples can give accurate estimates of genotyping error (Leal, 2005) but this was not feasible in the current study.

The Sotho population is composed of three sublanguage populations (Southern Sotho, Tswana and Northern Sotho) indicative of possible population stratification. Population stratification, however, is expected to increase heterozygosity levels not decrease levels as observed for the Sotho (Leal, 2005). The sampling of the sublanguage groups together in the current study might be the actual reason for non-adherence not population stratification, as the subpopulations were treated as one but in actual fact are representative of three. The grouping together of the language groups possibly resulted in the under representation of each sublanguage population thereby affecting the HWE. The grouping together may have resulted in a “forced” population stratification.

The possibility that participants indicated that they have three generations of Sotho ancestry but in actual fact have mixed ancestry could also contribute to HWE non-adherence. The Sotho-Tswana language groups are similar and

admixture between these population groups are not uncommon. It has been determined that the Southern Sotho and Tswana populations are closely genetically related (Lane *et al*., 2002). The relation between the groups is possibly due to sharing recent ancestors or due to recent admixture or both. Determining whether participants provided misleading information is however impossible.

Monoamine Oxidase A

Monoamine Oxidase A-uVNTR PCR products were identified on 2% agarose gels (Figure 3.3) following sequencing. Sequences from analysed gels and a reference sequence (KC609431) from GenBank were compared (Figure 3.4). From the sequencing results 2.5-, 3.5-, 4.5- and 5.5-repeats were successfully identified instead of the 2-, 3-, 4- and 5-repeats alleles which were identified by previous authors (Sabol *et al*., 1998; Huang *et al*., 2004; Guo *et al.*, 2008). More than 10% (43 out of 255) of the *MAO-A* DNA samples were successfully sequenced. Three 2.5-repeat alleles, 17 3.5-repeat alleles, 22 4.5-repeat alleles and one 5.5-repeat allele were successfully sequenced.

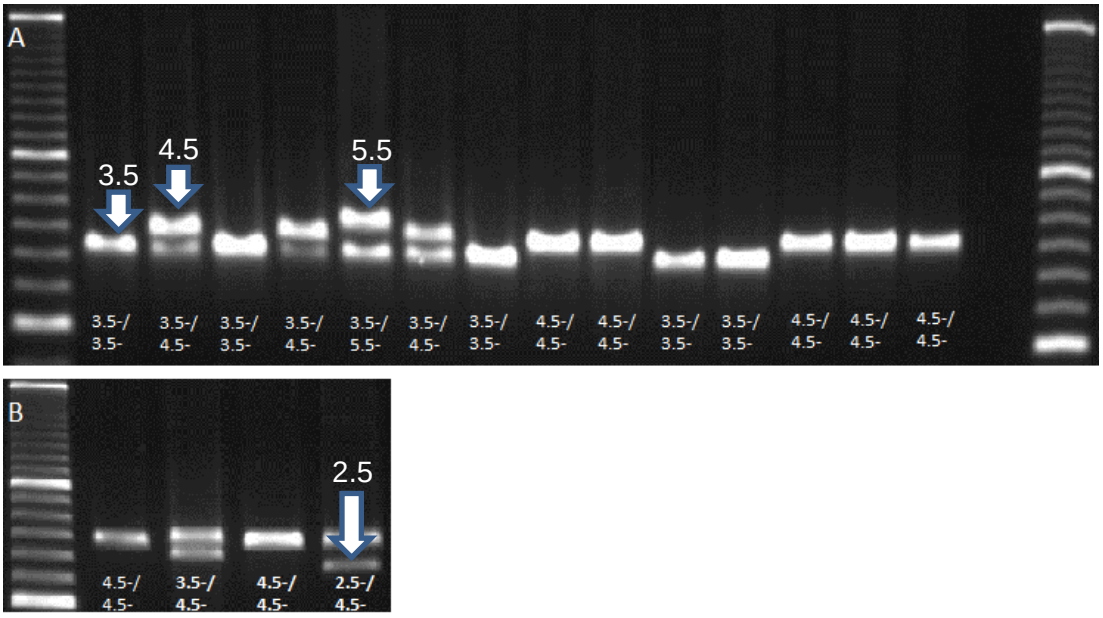

Figure 3.3 The visualisation of *MAO-A*-uVNTR PCR products on 2% agarose gels, representing the 2.5-, 3.5-, 4.5- and 5.5-repeat allele bands including a 50 bp ladder. A) 3.5-, 4.5- and 5.5-repeat B) 2.5-repeat as indicated by arrows.

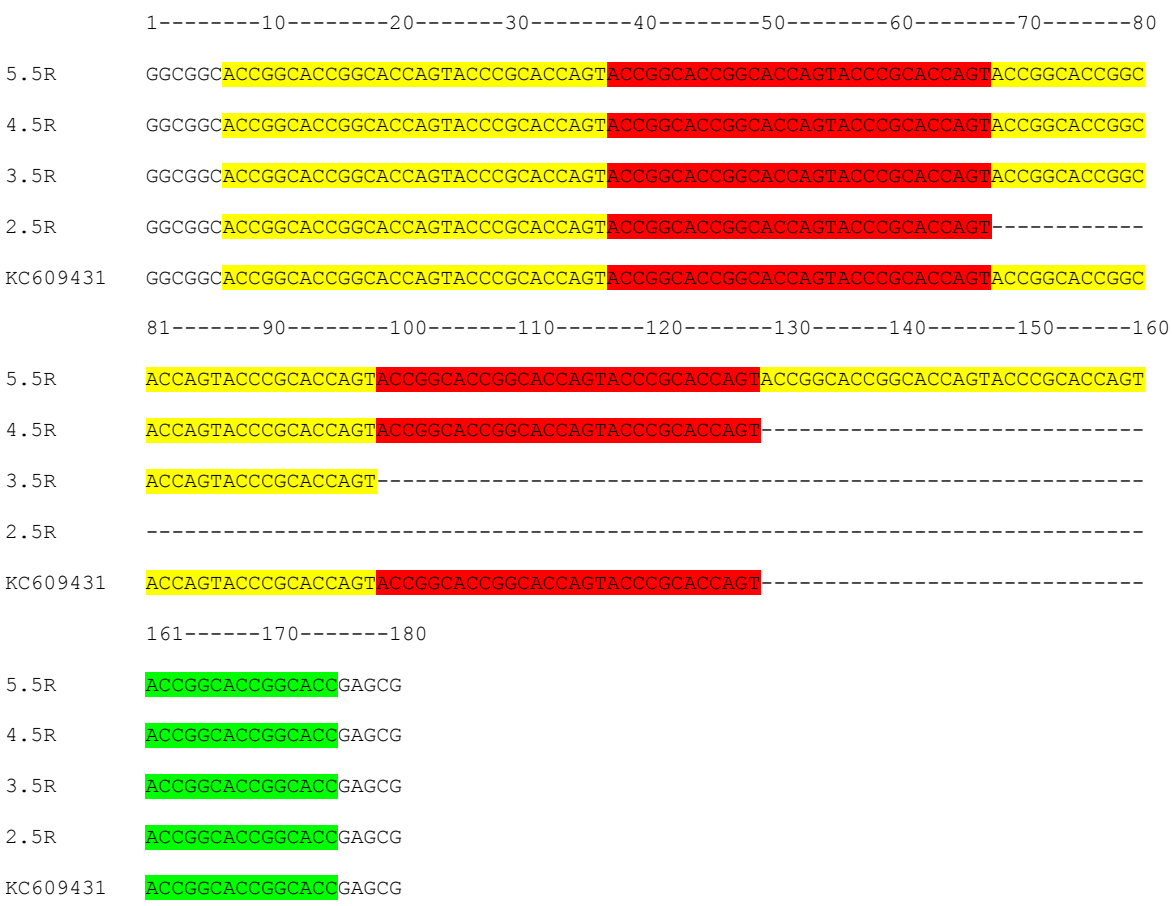

Figure 3.4 Nucleotide sequence alignment of *MAO-A*-uVNTR sequences identified (2.5-, 3.5-, 4.5- and 5.5-repeat alleles) on 2% agarose gels and a reference sequence from GenBank (Accession number KC609431), aligned with Geneious 6.1.6. The 30 bp repeat sequences are highlighted consecutively yellow and red, half repeat disregarded by original authors indicated by green.

The difference between results from this study and the previous mentioned authors were due to a discrepancy regarding half of the 30 bp repeat sequence. The original authors did not include half of the repeat when allele calling was performed. Authors such as Jorm *et al*. (2000), Das *et al*. (2006) and Laubscher (2012) have also noted this discrepancy. The 2-, 3-, 3.5-, 4-, 5- and 6-repeat alleles reported by Sabol *et al*. (1998), Huang *et al*. (2004), and Guo *et al*. (2008) therefore actually refer to 2.5-, 3.5-, 4-, 4.5-, 5.5- and 6.5-repeat alleles.

Male and female results were analysed separately due to the location of *MAO-A*-uVNTR on the X-chromosome. Female alleles contributed to 75.91% of the total sample (Table 3.3). The 4.5-repeat allele was the most common having a

frequency ranging from 0.464 to 0.794 for the males and 0.489 to 0.695 for the females in the analysed populations. The Khwe, Xhosa and Sotho populations were determined having four *MAO-A*-uVNTR variants (2.5-, 3.5-, 4.5- and 5.5-repeat alleles) whereas the Afrikaner (3.5-, 4.5- and 5.5-repeat alleles) had only three. The male (0.794) and female (0.695) Afrikaner populations had higher frequencies for the 4.5-repeat in comparison to North American Caucasians (0.648) (Sabol *et al*., 1998).

Table 3.3 Allelic variation of the *MAO-A*-uVNTR in South African ethnic groups for both male and female populations.

Ethnic group:	Sex:	*MAO-A*-uVNTR allele frequency:				2n:
		2.5	3.5	4.5	5.5	
Khwe	M	0.063	0.312	0.625	0.000	16
	F	0.018	0.429	0.535	0.018	56
	Total	0.028	0.403	0.555	0.014	72
Xhosa	M	0.000	0.393	0.464	0.143	28
	F	0.047	0.360	0.523	0.070	86
	Total	0.035	0.368	0.509	0.088	114
Sotho	M	0.143	0.286	0.571	0.000	21
	F	0.091	0.420	0.489	0.000	88
	Total	0.101	0.394	0.505	0.000	109
Afrikaner	M	0.000	0.177	0.794	0.029	34
	F	0.000	0.256	0.695	0.049	82
	Total	0.000	0.233	0.724	0.043	116

The 3.5-repeat allele among the Khwe (0.403), Xhosa (0.368) and Sotho (0.394) ethnic groups was not as dominant as among the African American population (0.591) studied by Sabol *et al*. (1998). Sabol *et al*. (1998) analysed a small number of individuals and due to African Americans having their own distinct history correlations were possibly affected (Tishkoff *et al*., 2009). African populations are not thoroughly analysed for the *MOA-A*-uVNTR therefore no significant correlation could be made.

The expected and observed heterozygosities determined for the female *MAO-A*-uVNTR samples were 0.538 and 0.607 for the Khwe, 0.596 and 0.605 for the Xhosa, 0.583 and 0.727 for the Sotho and 0.454 and 0.512 for the Afrikaner. The

HWE *P*-value for the Khwe was 0.675, Xhosa 0.159, Sotho 0.042 and Afrikaner 0.052. Only the Sotho population did not adhere to the HWE possibly due to the same reasons mentioned earlier ("forced" population stratification, sample population size or/and selection). Genotyping errors are not suspected, *DRD4* VNTR and *MAO-A*-uVNTR non-adherence was only determined for the Sotho population. Population stratification and sample population size are suspected the main reasons for the observed heterozygosity levels of the Sotho.

Serotonin transporter-linked polymorphic region

The sequencing of the 5-HTTLPR allowed the identification of the S and L alleles, however, two additional variants (VL and XL) of greater size (550 bp and 600 bp) were not successfully sequenced. Sequence failure was possibly due to the same reasons mentioned for the 3- and 10-repeat alleles of *DRD4* (high GC content and the utilization of a gel cut-out method). The VL and XL alleles were included in the results, as the fragments observed in this study correlate with alleles sequenced in previous studies (Kunugi *et al*., 1997; Gelernter *et al*., 1999; Narita *et al*., 2001; Hu *et al*., 2006; Ehli *et al*., 2012). Variant 5-HTTLPR PCR products were visualised on agarose gels (Figure 3.5). An example of an S allele (Figure 3.5) was sequenced to confirm its identification with a reference sequence (EU035982) obtained from GenBank (Figure 3.6). Two S and three L alleles were successfully sequenced.

The SNP (A→G) determined by Ozaki *et al*. (2003) was not identified among the analysed populations, however, it should be noted that not all L alleles were sequenced in the current project due to limited project time frame and budget (Appendix C). The L allele was the most prominent among the analysed populations ranging from 0.585 to 0.880 (Table 3.4). The Khwe population had the most diverse alleles (S, L, VL and XL) possibly due to its admixed history and/or relation to ancestral modern human populations (Khoe-San) (Chen *et al*., 2000). The Afrikaner population had the highest frequency for the S allele similar to European populations. The Afrikaner S allele frequency (0.396) was similar to German (0.399), Dutch (0.420) and French (0.414) control populations (Bellivier *et al*., 1998; Hoefgen *et al*., 2005; Denys *et al*., 2006). The Khwe population (0.065)

had the lowest frequency for the S allele with the Sotho (0.222) and Xhosa (0.210) populations being intermediate between the Afrikaner and the Khwe. The allelic frequencies obtained for the Sotho, Xhosa and Khwe were similar to that found for other African populations (Mbuti, South African Bantu) (Gelernter *et al*., 1999; Esau *et al*., 2008).

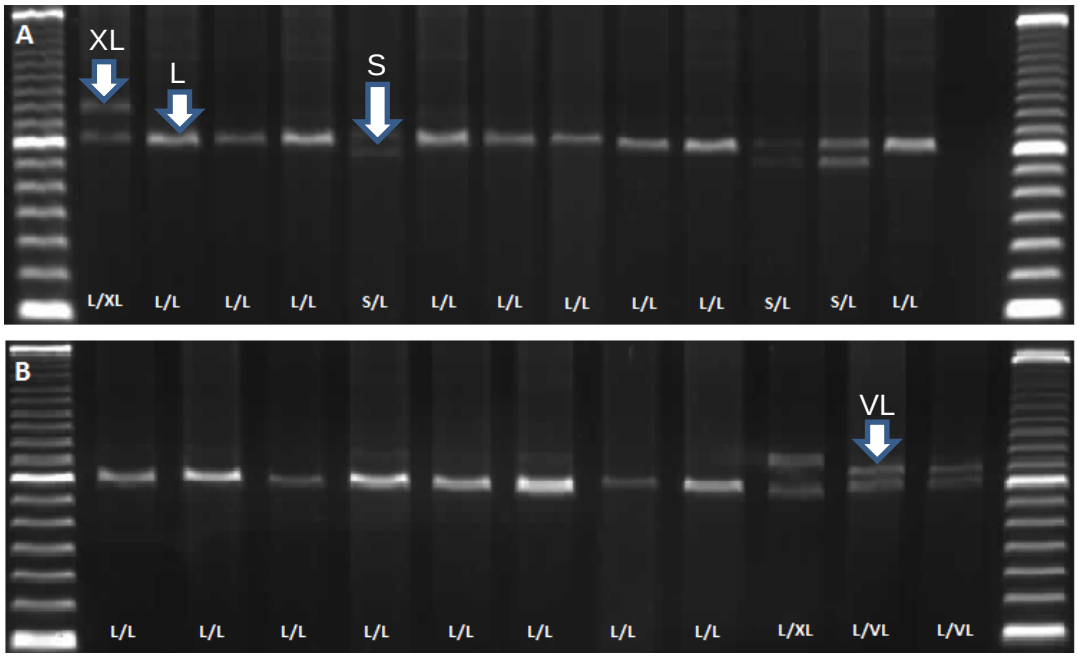

Figure 3.5 The visualisation of 5-HTTLPR PCR products on 2% agarose gels, representing the S, L, VL and XL allele bands including a 50 bp ladder. A) S, L and XL B) VL alleles indicated by arrows.

Expected and observed heterozygosities determined were 0.221 and 0.239 for the Khwe, 0.334 and 0.275 for the Xhosa, 0.373 and 0.286 for the Sotho and 0.505 and 0.415 for the Afrikaner. The HWE *P*-value for the Khwe was 1, Xhosa 0.154, Sotho 0.113 and Afrikaner 0.008. The Afrikaner population was the only population not adhering to HWE at this particular locus even though it had the highest level of observed heterozygosity. Suspected reason for non-adherence is possibly due to the Afrikaner being an amalgamated population and having a small founder population (Botha & Beighton, 1983; Stine & Smith, 1990; Hall *et al*., 2002). The Afrikaner had almost identical frequencies to other European populations; therefore the results based on the Afrikaner should not be ignored. The Khwe population was the only population with a higher observed

heterozygosity than expected heterozygosity for the 5-HTTLPR similarly as for the *DRD4* VNTR and *MAO-A*-uVNTR due to already mentioned reason (Chen *et al*., 2000).

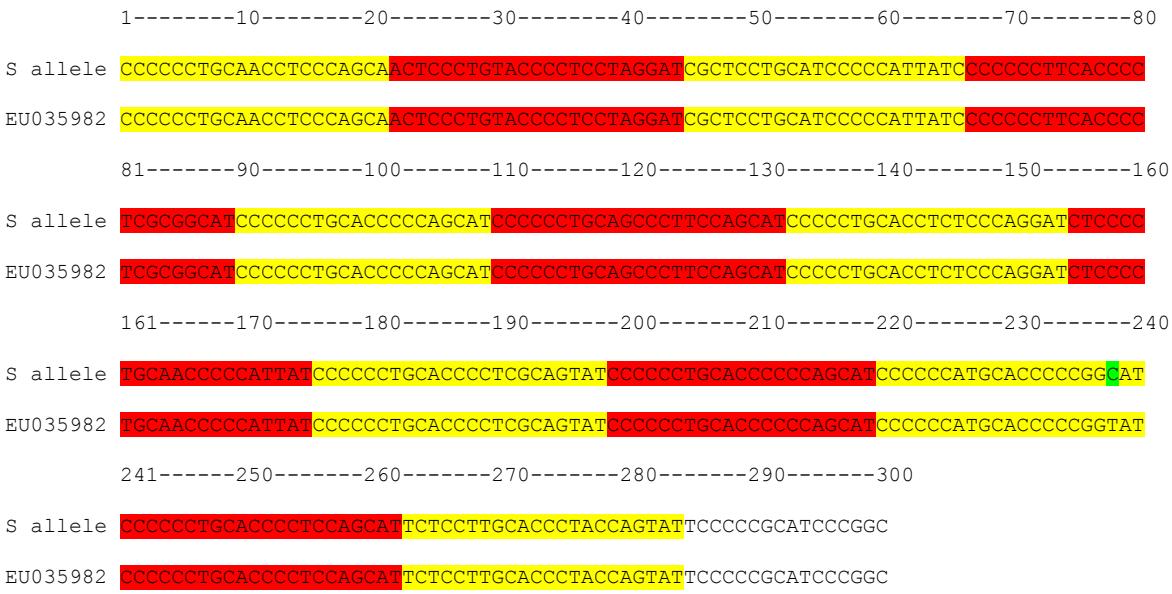

Figure 3.6 Nucleotide sequence alignment of the 5-HTTLPR S allele identified on the 2% agarose gel and a reference S allele sequence from GenBank (Accession number EU035982), aligned with Geneious 6.1.6. Reading motif starts from the second repeat of the 5-HTTLPR, the 20-23 bp complex repeat units are consecutively highlighted yellow and red. A single nucleotide variant was identified in the comparison (highlighted green).

Table 3.4 Allelic variation of the 5-HTTLPR in South African ethnic groups.

Ethnic group:	5-HTTLPR allele frequency:				2n:
	S	L	VL	XL	
Khwe	0.065	0.880	0.033	0.022	92
Xhosa	0.210	0.790	0.000	0.000	138
Sotho	0.222	0.762	0.000	0.016	126
Afrikaner	0.396	0.585	0.019	0.000	106

Conclusion

The *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR indicated that the Sotho and Xhosa populations are closely related due to the similarity among their allele frequencies. This correlation was expected as Sotho and Xhosa populations are

both of Bantu origin migrating along similar paths into southern Africa (Davenport, 1978; Tishkoff *et al*., 2009). A study by Lane *et al*. (2002) based on autosomal and Y-chromosome DNA determined that even though linguistically the Sotho and Xhosa have separated within the last 2 000 years they stem from a common ancestor population and were not isolated for long.

The Khwe population was represented by the smallest number of samples in this study. The Khwe, however, had the most diverse alleles at the 5-HTTLPR and was the only population with higher observed than expected heterozygosity frequencies for every analysed gene. The past history of the Khwe population from Platfontein in the Northern Cape with reference to admixture and association with the first modern human populations possibly contributed to the Khwe having substantial observed heterozygosity frequencies. A large scale study of the Khwe population would be of interest especially if DNA samples were to be collected at the various locations where the Khwe are presently living in southern Africa. The Khwe has been associated with ancestral human populations and have a history of migration (Chen *et al*., 2000).

The Afrikaner had the lowest genetic diversity among the analysed populations (5- and 10-repeat alleles of *DRD4* VNTR, 2.5-repeat allele of *MAO-A*-uVNTR and the XL allele of 5-HTTLPR were absent). The slight reduction in genetic diversity determined among the Afrikaner population was possibly due to past bottlenecks in human populations with regards to human migration from Africa to Europe (Jorde *et al*., 1998; 2000; Tishkoff & Williams, 2002; Lohmueller *et al*., 2008). Founder effects are also a possible reason for reduced genetic diversity among the Afrikaner (Botha & Beighton, 1983; Stine & Smith, 1990; Hall *et al*., 2002).

The Afrikaner population in the current study had similar allele frequencies to other European populations. The Afrikaner being an amalgamated European population was particularly related to its founder populations (Dutch, German and French) with regards to the *DRD4* VNTR and 5-HTTLPR. According to a study by Greeff (2007), Asian and African populations have also contributed to the Afrikaner genome. The exact measures of contributions to the Afrikaner genome has not been fully determined by extensive genetic analysis, therefore a study on this particular population with the analysis of more loci could be beneficial to

human population studies. The allelic frequencies for the three genes in this study have important implications for different behavioural disorders and might help in designing proper diagnostic strategies for different South African ethnic groups. The behavioural and psychiatric fields in particular may benefit from the genetic information provided with regards to genes that play a role in neuropsychiatric disorders.

For the first time information is available on four South African ethnic groups with regards to three neurotransmitter genes and the use of the same sampled population groups. From a population genetic view this study includes a population that is associated with ancestral human populations (Khwe); two of the largest population groups in South Africa (Xhosa and Sotho) have been studied and a population that has been amalgamated from various European populations (Afrikaner) was included. Possible pitfalls are the small population sizes utilized and the small number of loci analysed (Hong & Park, 2012). However, the majority of the genes analysed adhered to HWE. Genetic differentiation based on the allele frequencies of the analysed genes were possible and will be investigated to a greater extent in the following chapter. The following chapter will investigate the relationship of the analysed genes with migratory behaviour.

CHAPTER FOUR

PUTATIVE CORRELATION BETWEEN FREQUENCIES OF NEUROTRANSMITTER ALLELES AND HUMAN MIGRATION

Abstract

The *DRD4* VNTR is associated with the migration distance of modern humans. Comparisons of the *DRD4* VNTR make limited use of African populations. This is primarily due to the lack of neurotransmitter studies on African populations. In this chapter the *DRD4* VNTR, as well as an additional polymorphism (5-HTTLPR) were investigated among different South African populations in an attempt to compare allele frequencies with migration patterns. South African and revised world-wide populations' allele frequencies for the *DRD4* VNTR and 5-HTTLPR were compared. Linear correlations were calculated in order to determine whether the allele frequencies of the populations correlated with distance migrated. Positive correlations were identified for the 5-HTTLPR (S allele) and the *DRD4* VNTR (7-repeat, >5-repeat and 2- and >5-repeat alleles combined), indirectly providing information on the migratory behaviour trait and selection. Both the *DRD4* VNTR and 5-HTTLPR were associated with novelty seeking which is correlated with migratory behaviour. A clear relationship was observed between allele frequencies and migration distance. Significant genetic correlations between geographically related populations were identified.

Introduction

Archaeological and genetic studies support the fact that modern humans originated in Africa within the last ~200 000 years (Ingman *et al*., 2000; Tishkoff & Williams, 2002; Cavalli-Sforza & Feldman, 2003; Relethford, 2008; Tattersall, 2009). Studies on human migration point to an "out of Africa" pattern for modern humans (Watson *et al*., 1997; Quintana-Murci *et al*., 1999; Semino *et al*., 2002; Tishkoff & Williams, 2002; Cavalli-Sforza & Feldman, 2003). During migration from Africa certain individuals were conceivably better adapted to certain encountered environments, carrying beneficiary genetic compositions (Bamshad & Wooding, 2003; Lesch & Gutknecht, 2005; Sabeti *et al*., 2006).

Human phenotypes, such as the resistance to malaria and lactose tolerance, have adaptive value in certain environments and were possibly selected for or against, depending on environment (Hamblin & Di Rienzo, 2000; Tishkoff *et al*., 2007). A potential selective target is the *Forkhead Box P2* (*FOXP2*) gene (Enard *et al*., 2002). Variations within the *FOXP2* are linked with difficulty in expressive and receptive language. The dominant world-wide *FOXP2* variant was possibly fixed in modern humans prior to expansion from Africa as linguistic communication became more important (Enard *et al*., 2002). Selection on skeletal and neurological development predated modern human expansion (Schlebusch & Soodyall, 2012). Migration and expansion within and out of Africa were possibly influenced by the selection for certain traits.

Initial tests for selection with genetic markers were based on identifying loci with substantial allelic frequency variation between populations (Lewontin & Krakauer, 1973; Cavalli-Sforza, 1997). The goal of this chapter differs as we correlate allelic frequency variation between populations with a behavioural trait (migratory behaviour). Human population overlap and non-independence are some of the problems encountered during this study. Geographic related populations are usually genetically closely related, thus resulting in false positives (Chang *et al*., 1996; Lane *et al*., 2002; Tishkoff *et al*., 2009; Schlebusch & Soodyall, 2012). Genetic drift, brought on by population bottlenecks, contributes to the genetic association among geographic populations. Population bottlenecks occurred during human migration “out of Africa” (Tishkoff & Williams, 2002; DeGiorgio *et al*., 2009; Cai *et al*., 2011; Henn *et al*., 2011).

Certain behavioural traits are beneficial in certain circumstances. Individuals with high novelty seeking levels benefit when novel ecological and social circumstances are encountered (Chen *et al*., 1999). Individuals with high novelty seeking levels are characterised as being impulsive, excitable, explorative and extravagant, while individuals with low novelty seeking levels are more conservative being slow-tempered, reflective and rigid (Ebstein *et al*., 1997). The *DRD4* VNTR and 5-HTTLPR are associated with novelty seeking which in turn is associated with migratory behaviour (Ebstein *et al*., 1997; Chen *et al*., 1999; Tomitaka *et al*., 1999; Strobel *et al*., 2003; Reist *et al*., 2007; Kogan *et al*., 2010; Matthews & Butler, 2011).

The 5-HTTLPR possibly provides information on etiological and evolutionary theories of humans due to its global variation (Gelernter *et al*., 1999; Matthews & Butler, 2011). The 5-HTTLPR, as mentioned, is characterised by a 44 bp deletion or insertion resulting in either an S or L allele (14 or 16 repeat elements) (Lesch *et al*., 1996). A SNP determined within the L allele (A→G) results in an allele (L_G) with expression levels similar to that of the S allele (Ozaki *et al*., 2003; Ehli *et al*., 2012). The varying expression levels of the 5-HTTLPR alleles have a potential biological effect on novelty seeking behaviour (Strobel *et al*., 2003).

Caucasian populations have slightly higher frequencies for the L allele (0.572-0.580). Among African populations the L allele (~0.838) is even more prevalent. Asian individuals have very high S allele frequencies (0.698-0.785) (Kunugi *et al*., 1997; Gutiérrez *et al*., 1998; Szekely *et al*., 2004; Kim *et al*., 2007; Li *et al*., 2007; Esau *et al*., 2008). The wildtype L allele (L_A) has an average frequency of 0.503 in Caucasian populations (United States and Finland) with the less efficient L_G allele having a frequency of 0.114 (Hu *et al*., 2006). Among Native North American Indians (Plains Indians and Southwest Indians) the L_A allele occurs at a frequency of 0.350 and the L_G at a frequency of 0.010. African American individuals have the L_A allele at a frequency of 0.510 and L_G allele at 0.240 (Hu *et al*., 2006).

Studies generally do not differentiate between the two L alleles; generally only the term L allele is used. Reason for this is the majority of behavioural studies focus on the S allele frequency. The S allele is associated with various disorders and traits including novelty seeking as mentioned (Lesch *et al*., 1996; Strobel *et al*., 2003; Sen *et al*., 2004; Feinn *et al*., 2005; Cervilla *et al*., 2006; Young *et al*., 2008; Kogan *et al*., 2010; Gonda *et al*., 2011; Karg *et al*., 2011; Wang *et al*., 2011).

The *DRD4* gene is informative in a number of studies focussing on migration, due to its allelic variation among populations (Chen *et al*., 1999; Mansoor *et al*., 2008; Matthews & Butler, 2011). Alleles at the VNTR as mentioned range from 2 to 11 copies of variant 48 bp sequences (Van Tol *et al*., 1991; 1992; Mansoor *et al*., 2008). Across all human populations the 4-repeat allele of the *DRD4* VNTR is the most prevalent followed by the 7- and 2-repeat alleles (Chang *et al*., 1996). Among African, Caucasian and Asian populations the 4-repeat allele (>0.700) is most prevalent in comparison to Native North and South American populations that have substantial 7-repeat allele frequencies (0.483) (Chang *et al*., 1996; Chen

et al., 1999, Mansoor *et al*., 2008). The 7-repeat allele possibly arose 40 000 to 50 000 years ago, which correlates with modern human expansion from Africa and the higher frequency of the 7-repeat allele outside of Africa (Wang *et al*., 2004). Various behavioural studies have been performed on *DRD4* VNTR (Muramatsu *et al*., 1996; Cruz *et al*., 1997; Jardemark *et al*., 2002; Lung *et al*., 2002; Millet *et al*., 2003; Lahti *et al*., 2005; Camarena *et al*., 2007; Ray *et al*., 2008; Dreber *et al*., 2009; Kuhnen & Chiao, 2009; Creswell *et al*., 2012; Gonçalves *et al*., 2012).

Population studies on the *DRD4* VNTR and 5-HTTLPR have indicated similar allelic frequencies among geographically related populations, which is indicative of migratory behaviour (Chang *et al*., 1996; Kunugi *et al*., 1997; Gutiérrez *et al*., 1998; Sabol *et al*., 1998; Chen *et al*., 1999; Willeit *et al*., 2003; Szekely *et al*., 2004; Wang *et al*., 2004; Kim *et al*., 2007; Li *et al*., 2007; Esau *et al*., 2008; Mansoor *et al*., 2008; Nonnis Marzano *et al*., 2008; Noskova *et al*., 2008). The allelic frequency variation of the *DRD4* VNTR cannot be explained for by neutral processes such as drift and admixture (Chen *et al*., 1999; Matthews & Butler, 2011). With the use of a least squares linear model based on neutral loci from Tishkoff *et al*. (2009), Matthews and Butler (2011) determined that the frequency of the 2- and 7-repeat alleles of *DRD4* VNTR are more associated with migratory distance than expected by chance. Due to South African populations not being substantially genetically analysed, the loci used in Tishkoff *et al*. (2009) could not be included in a comparative analyses between possible selected and neutral loci. However, South African populations can be compared with world-wide populations.

The aim of the chapter is to identify the relationship between migration distance, migratory behaviour and the allelic frequencies of the 5-HTTLPR and *DRD4* VNTR. Objectives include determining the genetic distance among South African ethnic groups with the mentioned genes. Defining the relationship between the migration distance and allelic frequency of the current study populations and world-wide populations. Identifying whether positive correlations exist between the allelic frequencies and distances migrated with linear regressions. Determining if selection took place for novelty seeking alleles in relation to migration and migratory behaviour.

Migration analysis

A migration origin point (17.5°S, 12.5°E) for modern humans exists in south-western Africa (Tishkoff *et al*., 2009). Migration patterns out of Africa all indicate expansion either through northern or/and eastern Africa (Cavalli-Sforza *et al*., 1997; Tishkoff & Williams, 2002; Reed & Tishkoff, 2006; Sirugo *et al*., 2008; Tishkoff *et al*., 2009). The current study populations fall into two possible historic migration routes: a) migration of individuals (Xhosa and Sotho) from south-western Africa into a region in present day Cameroon then back into southern Africa in a south easterly direction as indicated in Figure 2.3 b) migration of individuals (Afrikaner) from south-western Africa to a region in present day Cameroon then into northern Africa and exiting through Egypt into Euroasia.

The Khwe has been located for the largest part of their history in south-western Africa and due to their association with the first modern humans it is not expected that the Khwe have migrated any significant distance during their history. The migration distance from the migration origin point (17.5°S, 12.5°E) to where the Khwe is located in the current study is therefore only a short distance. However, it should be noted that other populations that have admixed with the Khwe could have possibly migrated great distances. In order to avoid complications, groups that may have admixed with the Khwe, migration distances were not included.

Migration out of Africa followed the following basic line: Africa → Europe and/or Asia, Asia → North America → South America. This simplified migration pattern allowed for minimum migration distances of populations from the origin of human migration to the present locations. Migration distances from the migration origin point and distances between populations were measured in straight lines (as the crow flies).

Determining the origin points of South African populations were not feasible (origin of population during migration or origin of population at specific region) as African populations have a complex history of admixture and stratification (Schlebusch *et al*., 2012). Instead average geographical points were determined for each population using locations where the respective populations have been located for long periods of time and previously sampled with the ALlele FREquency Database (ALFRED) (Table 4.1).

Khwe individuals have been living in northern Namibia, southern Angola and northern Botswana (Chumbo & Mmaba, 2002) and recently a number of Khwe have been relocated to Platfontein in South Africa (den Hertog, 2013). The Okavango in Ngamiland was chosen as the Khwe representative geographic location due to the Khwe presiding there for a significant part of their recorded history (Chumbo & Mmaba, 2002). The location used for the Xhosa was close to Mthatha in the Eastern Cape while the location used for the Sotho was near Bloemfontein in the Free State. A random geographical location within one of Afrikaners ancestral countries was chosen as its representative point (Amsterdam) (Davenport, 1978). The recent migrations back to Africa for the Afrikaner ancestors were ignored. Origin (location) points for additional world-wide populations from previous studies were correlated with their sampling locations or by making use of ALFRED.

Table 4.1 Populations sampled and sampling location. The geographic points determined from ALFRED are provided.

Population:	Geographic location, ALFRED:
Khwe	18°S, 23°E
Xhosa	32°S, 27.5°E
Sotho	30°S, 26.5°E
Afrikaner	52.5°N, 4.9°E*

*Afrikaner location not obtained by ALFRED instead a representative location point in Amsterdam (Netherlands) was used.

Khwe individuals were sampled at Platfontein in the Northern Cape, Xhosa, Sotho and Afrikaner individuals were sampled in Bloemfontein during this study. Some Xhosa individuals were also sampled in the Eastern Cape. Sampling location and representative geographic location may differ for the same population; the primary reason for this is to avoid taking recent migrations into account. Recent migrations between Africa, Europe and the Americas did not have an effect on the allelic frequency of the 5-HTTLPR and *DRD4* VNTR. Migrations within the last 1 000 years were generally driven by slavery (African American), religious persecution (Afrikaner) and famines (Irish) not by novelty seeking behaviour

(Woodham-Smith, 1962; Berlin, 1998; Parra *et al*., 1998; Tipping *et al*., 2001). Recent migrations that occurred within the last 1 000 were ignored during this study.

Materials and methods

Ethical clearance for the study was obtained from the Medical Ethics Research Committee of the University of the Free State (ECUFS 152/2011). After obtaining informed consent from voluntary donors, saliva samples were collected and analysed for the 5-HTTLPR and *DRD4* VNTR, see Chapter 3. Genetic distances (F_{ST}) between the South African populations were determined with POPTREE2 (Latter, 1972; Takezaki *et al*., 2010).

Additional world-wide populations (Adygei, Apalaì, Biaka, Cambodians, Cheyenne, Danes, Dutch, Falasha, Finns, French, Germans, Han Chinese Taiwan, Japanese, Kachari, Malay, Maya, Mbuti, Muskoke, Nasioi Melanesians, Quechuan, Rondonian Surui, Siberian Eskimos, Ticuna and Yakut) from previous studies (Chang *et al*., 1996; Kunugi *et al*., 1997; Sander *et al*., 1997; Bellivier *et al*., 1998; Chen *et al*., 1999; Gelernter *et al*., 1999; Millet *et al*., 2003; Hoefgen *et al*., 2005; Denys *et al*., 2006; Bisso-Machado *et al*., 2013) were compared with the four South African populations. Phylogenetic trees were drawn for the populations in relation to their analysed allelic frequencies with POPTREE2 and demonstrated with GenGis 2.1.1. Linear regression comparisons between allelic frequencies and migration distances were performed with XLstat.

Results

A total of 231 5-HTTLPR and 200 *DRD4* VNTR samples were successfully analysed for 349 South African individuals (see Chapter 3). Low F_{ST} values were determined among the analysed South African populations (Table 4.2), similarly to those observed for other human populations (Pereira *et al*., 2012; Uchigasaki *et al*., 2013). The lowest genetic distance was between the Xhosa and Sotho primarily due to originating from the same ancestral Bantu population (Lane *et al*., 2002; Schlebusch *et al*., 2012). The Xhosa population was the closest relative of the Khwe population. The Khwe population is not as related to the Xhosa as the

Sotho possibly due to its relation with the first modern human populations (Khoe-San) (Chen *et al*., 2000; Schlebusch *et al*., 2012). The Afrikaner population was the most genetically distant population due to its mostly European ancestry (Davenport, 1978; Botha & Beighton, 1983; Greeff, 2007).

Table 4.2 Genetic distances for the Khwe, Xhosa, Sotho and Afrikaner ethnic groups based on the allele frequencies of the 5-HTTLPR and *DRD4* VNTR.

Population:	Xhosa	Sotho	Afrikaner
Khwe	0.016	0.031	0.094
Xhosa		0.009	0.042
Sotho			0.053

South African individuals were compared to populations from around the world [Figures 4.1 (5-HTTLPR) and 4.2 (*DRD4* VNTR)]. A particular continental similarity was identified (Chang *et al*., 1996; Chen *et al*., 1999). Populations from the same geographical area were genetically related in most cases. African populations (Afrikaner excluded) clustered together in the phylogenetic tree based on the predominance of the 5-HTTLPR L allele (Figure 4.1). The Afrikaner population grouped with European populations as expected. The L allele is slightly more prevalent than the S allele among European populations (Afrikaner included) as determined in revised literature (Bellivier *et al*., 1998; Hoefgen *et al*., 2005; Denys *et al*., 2006). The L allele was not split into the L_A and L_G alleles as information differentiating between the two L alleles were not available.

The Native South American populations grouped together with the Maya population with regards to the 5-HTTLPR (Figure 4.1). The Japanese were related to the Native American populations, due to substantial S allele frequencies. Siberian Eskimos, usually considered to be a link between Asian and North American populations (Malyarchuk *et al*., 2011; Achilli *et al*., 2013; Regueiro *et al*., 2013) surprisingly correlated with European populations.

With regards to the *DRD4* VNTR, African populations (Afrikaner excluded) grouped together due to prevalent 4-repeat allele frequencies (Figure 4.2). The

Xhosa, Sotho and Khwe, in comparison to other African populations, had slightly higher 7-repeat allele frequencies which are possibly due to admixture of southern African Bantu with ancestral Khoe-San groups (Schlebusch *et al*., 2012) or possibly due to selection. The Afrikaner grouped together with European populations due to substantial 2-repeat allele frequencies in addition to prevalent 4-repeat allele frequencies.

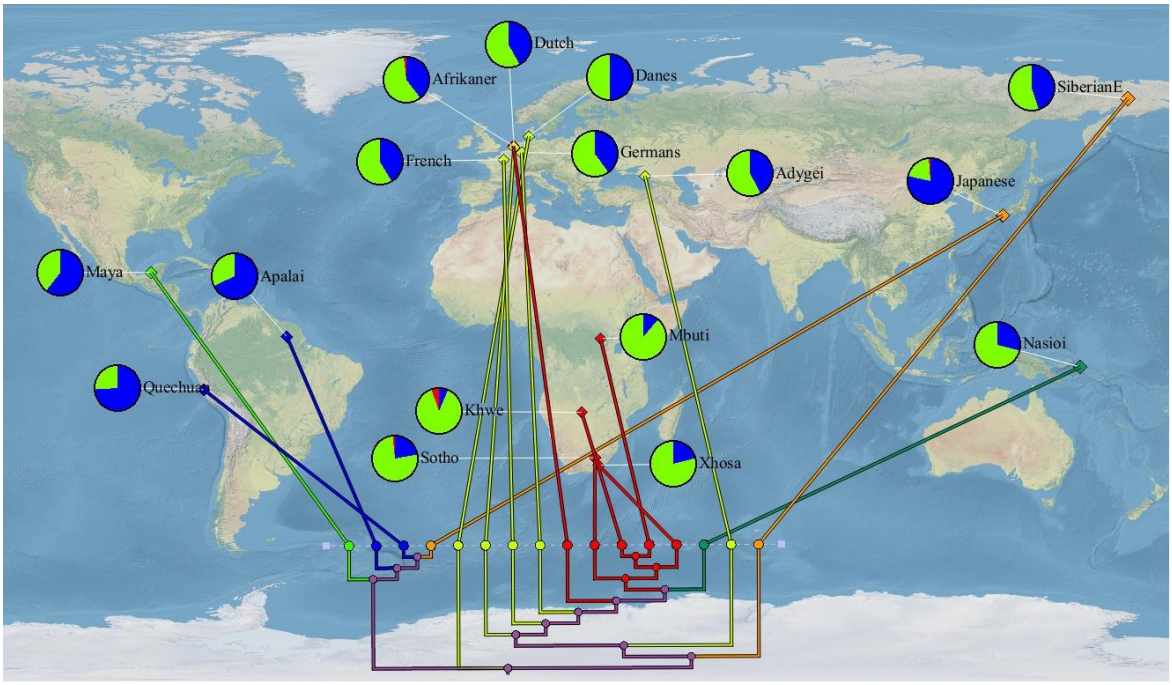

Figure 4.1 Allele frequencies of the analysed and world-wide populations for the 5-HTTLPR demonstrated by GenGis 2.1.1 in pie charts. Green is indicative of the L allele and blue is indicative of the S allele. Other alleles (VL and XL) are indicated by red. A somewhat continental resemblance was identified with the use of the phylogenetic tree imported from POPTREE2.

*SiberianE (Siberian Eskimos), *Nasioi (Nasioi Melanesians)

Asian populations had significant 2-repeat allele *DRD4* VNTR frequencies similar to European populations (Figure 4.2). However, Nasioi Melanesians in comparison to other Asian populations had reduced 2-repeat allele frequencies and greater 3-repeat allele frequencies. Nasioi Melanesians have a distinct genetic composition due to human migration between islands (Cox *et al*., 2010; Luca *et al*., 2011; Naka *et al*., 2012). Similarities among North and South American populations were based on very high 7-repeat allele frequencies. A number of migrations took place from Asia to North America which could account for the *DRD4* allele frequency variation among Native American populations

(Achilli *et al*., 2013).

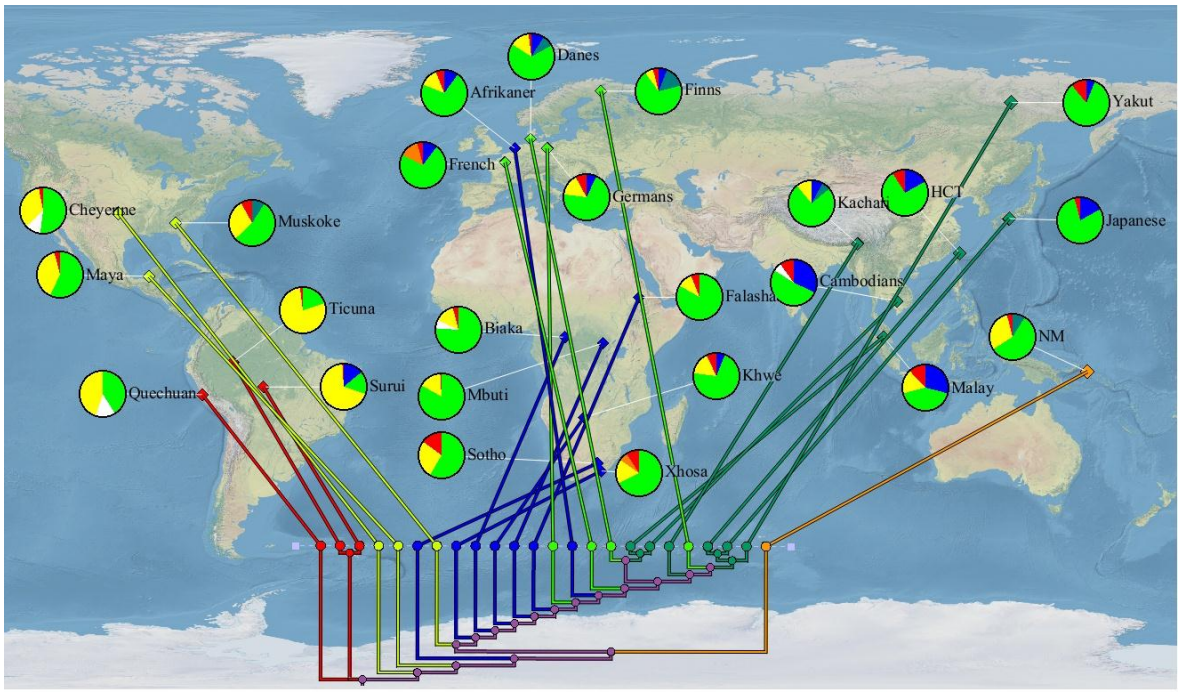

Figure 4.2 Allele frequencies of the *DRD4* VNTR for the analysed and world-wide populations, demonstrated by GenGis 2.1.1 in pie charts. The 4-repeat (green), 7-repeat (yellow), 2-repeat (blue), 6-repeat (white), 3-repeat (dark green), 8-repeat (orange) and other repeats 5- and 10-combined (red) alleles are all indicated. As in Figure 4.1 a somewhat continental resemblance was identified with the use of the phylogenetic tree imported from POPTREE2.

*HCT (Han Chinese Taiwan), *NM (Nasioi Melanesians)

Correlations based on migration distance and the allele frequencies of the 5-HTTLPR (S allele) and *DRD4* VNTR (2-repeat, 7-repeat, >5-repeat and combined 2- and >5-repeat alleles) were determined (Tables 4.3 and 4.4). African populations travelled the shortest distance from the modern human migration point (17.5°S, 12.5°E) followed by European, Asian, North and South American populations. Increase in migration distance for the populations correlated with substantial increase in the 5-HTTLPR S allele frequency (Table 4.3). The Khwe had the lowest S allele frequency (0.065) while the Japanese (0.785) had the highest. A continental resemblance was identified, with the average S allele frequency of African populations (Afrikaner excluded) being 0.166, Europe (Afrikaner included) 0.405, Asia 0.732, North America 0.605 and South America 0.712. The substantial S allele frequency for Asia was due to the Japanese population having an extremely high S allele frequency.

Table 4.3 Migration distance in kilometres (Km) along with the 5-HTTLPR S allele frequency for each population. The number of alleles (2n) analysed for each population is indicated.

Population:	Geographic location:	Distance from 17.5°S, 12.5°E[3]:	5-HTTLPR 2n:	S:
Khwe	18°S, 23°E*	1113	92	0.065
Mbuti	1.5°N, 28°E*	2711	72	0.111
Xhosa	32°S, 27.5°E*	6436	138	0.210
Sotho	30°S, 26.5°E*	6217	126	0.222
French	48.84°N, 2.36°E'	8516	198	0.414
Germans	50.73°N, 7.1°E'	8528	1654	0.399
Dutch	52.1°N, 5.11°E'	8727	134	0.420
Afrikaner	52.5°N, 4.9°E'	8772	106	0.396
Danes	55°N, 9°E[2]	8904	66	0.500
Adygei	44.5°N, 39.75°E*	7630	90	0.422
Japanese	34.1°N, 134.5°E*	14131	424	0.785
SE	65°N, 167.5°E*	18214	20	0.450
NM	6°S, 155°E*	19084	38	0.289
Maya	19°N, 91°W[2]	26838	76	0.605
Apalaí	2°N, 55°W*	31188	66	0.680
Quechuan	12°S, 77°W[2]	30611	74	0.740

'Coordinates obtained by making use of the iTouchMap.com (itouchmap.com/latlong.html).

*Coordinates obtained from ALFRED, the average of the given points were used.

[2]Coordinates obtained from Chen *et al*. (1999).

[3]Distances between points obtained by making use of Google Earth Antarctica (http://www.ig.utexas.edu/outreach/googleearth/latlong.html).

SE (Sibrian Eskimos), NM (Nasioi Melanesians)

The 2-repeat *DRD4* VNTR allele was scarcely present in the analysed African and Native American populations (Table 4.4). European and certain Asian populations had more prevalent frequencies for the 2-repeat allele. The 7-repeat allele had substantial frequencies in European populations and high frequencies in Native

American populations. When combining the 2-repeat and >5-repeat alleles a correlation was identified with migration distance. The correlation is particularly evident among Native American populations that travelled the greatest distances.

The coefficient of determination (R^2) for the 5-HTTLPR (S allele) and *DRD4* VNTR (7-repeat, >5-repeat and combined 2- and >5-repeat alleles) indicated positive correlations (Figure 4.3). The 2-repeat allele of *DRD4* on its own indicated no positive correlation. It is suggested that the 2-repeat allele functions in conjunction with the 7-repeat allele as previously determined (Matthews & Butler 2011). The R^2 value for the 7-repeat allele is lower but similar to that identified in Chen *et al*. (1999). Reason for difference is possibly due to Chen *et al*. (1999) differentiating between micro and macro migrations, which allowed for an increase in R^2 value.

The R^2 value for the >5-repeat allele is slightly lower than that for the 7-repeat allele, an indication that the majority of the long *DRD4* alleles (6-, 7-, 8- and 10-repeat alleles) are possibly associated with migration distance (Figure 4.3). Four populations are located on the horizontal line in regression C when testing for the 7-repeat allele alone while none touches when testing for the >5-repeat allele in regression D. A slight increase in the R^2 value is indicated when the 2-repeat allele is included along with the >5-repeat allele in regression E. This is an indication that the 2-repeat allele, in conjunction with the long alleles, is associated with migration distance (Chen *et al*., 1999; Matthews & Butler 2011). The R^2 values obtained for the S allele of the 5-HTTLPR is similar to that obtained for the 7-repeat allele of the *DRD4* VNTR an indication that the 5-HTTLPR is involved in migratory behaviour as the *DRD4* 7-repeat has previously been associated with migration distance and behaviour (Chen *et al*., 1999; Matthews & Butler, 2011).

Table 4.4 Migration distances for each of the populations along with the *DRD4* allele frequencies (2-repeat, 7-repeat, >5-repeat and combined 2- and >5-repeat alleles). The number of alleles (2n) analysed for each population is indicated.

Population:	Geographic location:	Distance from 17.5°S, 12.5°E[3]:	*DRD4* 2n:	2:	7:	>5:	2- and >5:
Khwe	18°S, 23°E*	1113	90	0.056	0.156	0.222	0.278
Biaka	3°N, 18°E[2]	1720	124	0	0.140	0.210	0.210
Mbuti	1.5°N, 28°E*	2711	74	0	0.160	0.170	0.170
Falasha	13.5°N, 37.5°E'	4406	128	0.030	0.110	0.140	0.170
Xhosa	32°S, 27.5°E*	6436	98	0.031	0.174	0.264	0.295
Sotho	30°S, 26.5°E*	6217	116	0.034	0.267	0.319	0.353
French	48.86°N, 2.35°E'	8518	126	0.103	0.008	0.151	0.254
Germans	52.52°N, 13.41°E'	8543	394	0.071	0.152	0.165	0.236
Afrikaner	52.5°N, 4.9°E'	8772	96	0.104	0.125	0.177	0.281
Danes	55°N, 9°E[2]	8904	64	0.090	0.140	0.160	0.250
Finns	67.5°N, 27.5°E*	10050	66	0.060	0.060	0.080	0.140
Kachari	27.5°N, 94.75°E'	10515	36	0.080	0.110	0.110	0.190
Cambodians	12.5°N, 105°E*	12495	50	0.320	0	0.100	0.420
HCT	25.1°N, 121.5°E'	13190	84	0.170	0	0.050	0.220
Malay	3.1°N, 101.5°E'	13609	24	0.290	0.170	0.250	0.540
Japanese	34.1°N, 134.5°E*	14357	102	0.170	0.010	0.020	0.190
Yakut	64.5°N, 135°E*	15509	78	0.060	0.040	0.040	0.100
NM	6°S, 155°E*	19635	46	0	0.300	0.300	0.300
Cheyenne	35.5°N, 99°W*	23454	96	0.010	0.340	0.450	0.460
Muskoke	33°N, 84°W[2]	24162	24	0.040	0.290	0.290	0.330
Maya	19°N, 91°W[2]	25450	100	0.010	0.390	0.420	0.430
Ticuna	3.5°S, 69°W*	28922	128	0.020	0.780	0.780	0.800
Quechuan	12°S, 77°W[2]	29223	44	0	0.450	0.590	0.590
Surui	10°S, 61°W*	30056	90	0.140	0.690	0.690	0.830

'Coordinates obtained by making use of the iTouchMap.com (itouchmap.com/latlong.html).

*Coordinates obtained from ALFRED, the average of the given points were used.

[2]Coordinates obtained from Chen *et al*. (1999).

[3]Distances between points obtained by making use of Google Earth Antarctica (http://www.ig.utexas.edu/outreach/googleearth/latlong.html).

SE (Sibrian Eskimos), NM (Nasioi Melanesians)

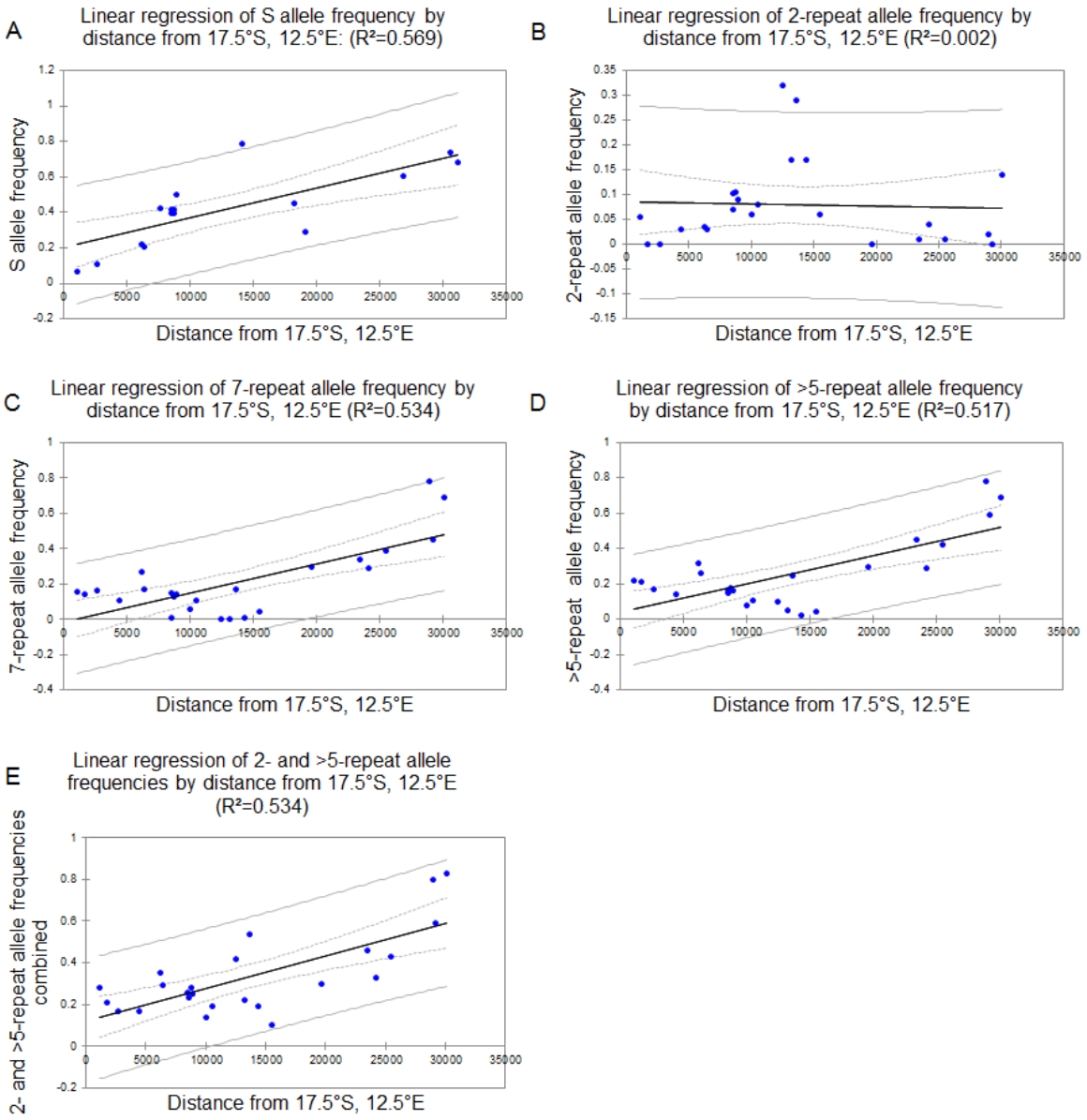

Figure 4.3 Linear regressions between the distance travelled and the allele frequencies of the 5-HTTLPR and *DRD4* VNTR. Allele frequency was chosen as the dependent factor and distance migrated the independent factor. A) A positive correlation (unadjusted R^2=0.569) was identified for the S allele of the 5-HTTLPR and migration distance. B) No correlation (unadjusted R^2=0.002) was identified for the 2-repeat allele of *DRD4* VNTR and migration distance. C), D) and E) Positive correlations (unadjusted R^2=0.534, 0.517 and 0.534) were identified for the 7-repeat, >5-repeat and combined 2- and >5-repeat allele frequencies.

Conclusion

The study revealed a significant genetic correlation between geographic related populations. South African populations excluding the Afrikaner were determined correlating with other African populations. A clear relationship was observed between allele frequencies and migration distance. The reason for the correlation between the alleles and distance travelled is of interest. Is the correlation a result of bottlenecks associated with past migrations or is it due to migrations driven on by novelty seeking alleles or other forces? Bottlenecks occurring during human migrations possibly contributed to the continental resemblance determined for the 5-HTTLPR and *DRD4* VNTR due to genetic drift and admixture. However, it is speculated that these founder effects would have resulted in abrupt changes in the allelic frequencies of populations (Chen *et al*., 1999) unlike the somewhat gradual increase identified in the majority of the analysed and reviewed populations. The study by Matthews and Butler (2011) revealed that the association between *DRD4* alleles and migration is of such nature that effects of admixture and drift could not account in total for the significant association.

The implication of the 5-HTTLPR and *DRD4* VNTR alleles in a number of behavioural studies and the functional variation identified holds key as to why the alleles possibly were selected for. The association of the mentioned genes with traits could have played a role in cultural selection. Cultural selection like natural selection selects the most appropriate trait for a particular environment (Wang *et al*., 2004). Populations with high novelty seeking levels benefitted in novel environments which possibly resulted in selection of this behavioural type and indirectly for genomic constituents associated with it.

Recent migrations did not have an impact on allelic frequency, as is evident in the Afrikaner population (African population with European ancestry). The Afrikaner ancestors recently migrated to southern Africa from Europe and have similar frequencies for the analysed genes as observed in European populations (generally low novelty seeking allele frequencies). Therefore novelty seeking alleles were probably selected for during historic migrations over long periods of time.

The prevalent frequency of VL and XL 5-HTTLPR alleles in the Khwe and its diminished presence in other world-wide populations is an indication that the Khwe is possibly related to ancestral human populations. The diminished presence of the VL and XL alleles are also possibly due to bottlenecks and the deselecting of longer alleles as expansion took place from Africa. The prevalent frequency of longer 5-HTTLPR alleles in the Khwe is possibly due to the population having not encountered conditions in which an allele (S allele) associated with high novelty seeking has been favoured.

The lack of information with regards to the frequency of the L allele as L_A or L_G alleles placed a limitation on this study. However, it is evident that with migration from Central Africa the S allele became more prominent, removing the factor of whether the L allele observed is actually a L_A or L_G allele. Migration from Africa resulted in a loss of genetic diversity which is possibly due to bottlenecks or selection.

Numerous *DRD4* VNTR alleles are associated with novelty seeking including the 2-, 6-, 7-, 8- and 10-repeat alleles. Correlations between migratory distances were determined the highest when all the *DRD4* novelty seeking alleles were included in the regression analysis. As determined by Chen *et al*. (1999) and Matthews and Butler (2011), *DRD4* novelty seeking alleles are associated with migration distance and the same was determent for the S allele of the 5-HTTLPR. Selection for the 7-repeat allele of *DRD4* was identified and therefore answers the question from the previous chapter. However, selection acts on a number of *DRD4* variants (Wang *et al*., 2004) and therefore does not explain why the Sotho did not adhere to HWE with regards to selection. Therefore the only possible reason as to why the Sotho deviated is due to the sampling procedure used.

CHAPTER FIVE

GENERAL CONCLUSION

The aim of this study was to determine the allelic diversity of some neurotransmitter genes in specific South African ethnic groups (Khwe, Xhosa, Sotho and Afrikaner) in order to determine their population structures. Answers to the following questions were pursued:

1. How much information is available on South African ethnic groups with regards to neurotransmitter genes (*DRD4*, *MAO-A* and *5-HTT*)?
2. What are the allelic frequencies of the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR for the Khwe, Xhosa, Sotho and Afrikaner ethnic groups?
3. What is the role of neurotransmitter genes with regards to selection, migration distance and migratory behaviour?

Information gained on South African populations was based on a limited amount of populations and neurotransmitter genes. Limited information was identified pertaining to South African populations with regards to neuropsychiatric disorders. The analyses of the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR in the Khwe, Xhosa, Sotho and Afrikaner ethnic groups therefore provided valuable allelic frequency information. For the first time a number of populations from different South African ethnic groups were analysed with three neurotransmitter genes.

Genetic variation between the populations was very low. The Xhosa and Sotho populations had similar allelic frequencies and shared the lowest genetic distance of the analysed populations. A greater genetic distance was determined between the Khwe and the analysed Bantu populations. The Afrikaner population correlated with European populations and was genetically the most distant.

The role of selection on neurotransmitter genes with regards to migration distance and novelty seeking was investigated and a positive correlation was determined. The 5-HTTLPR (S allele) and *DRD4* VNTR (7-repeat, >5-repeat and combined 2- and >5-repeat alleles) revealed association with selection. A particular correlation existed between the allele frequencies and migration distances of the current study and world-wide populations. The addition of South African populations

provided a perspective that was previously missing. Previous correlations on selection made no use of specific South African ethnic groups.

The analyses of the neurotransmitter genes in South African populations provided useful information to the behavioural and psychiatric fields. Comparing South African populations with world-wide populations allowed for further insight in evolutionary theories. With the inclusion of South African populations it appears that selection occurred for certain alleles of the *DRD4* VNTR as determined in Chen *et al*. (1999) and Matthews and Butler (2011).

The 5-HTTLPR was associated with migration distance for the first time in this study. The VL and XL alleles of the 5-HTTLPR had its highest frequency among the Khwe, no studied population had a higher percentage of VL and XL alleles. The genetic diversity of the Khwe and the presence of scarce alleles provide evidence that the Khwe is possibly related to ancestral human populations such as the Khoe-San. The increase of the 5-HTTLPR and *DRD4* VNTR novelty seeking alleles outside Africa with their highest frequency present in South America, links with the “out of Africa” theory. The frequency of the novelty seeking alleles increased gradually with distance from the migration origin point determined in south-western Africa by Tishkoff *et al*. (2009). The probably of bottlenecks or admixture contributing to the increase in frequency were not ruled out. However, it was concluded that such factors alone could not result in the substantial increased frequency.

Analyses of the current populations provided a substantial genetic background on South African populations. However, analyses with larger and more population groups should be performed in order to determine the whole picture. With regards to the Khwe it would be of interest and value to genetically analyse all the Khwe groups in southern Africa to determine the consequences of admixture in relation to recent and long term migrations. It would also be interesting to determine the genetic contributions made by the Khoe-San and Bantu populations to the Khwe population. Further analyses of the Xhosa and Sotho populations and other Bantu populations are also required as these groups consist of various language subgroups. The Sotho-Tswana language group for example consist of individuals belonging to the Southern Sotho, Tswana and Pedi groups. The analyses of the

Afrikaner with a large set of genes could bring forth more information on a population that was recently admixed from mostly European populations. The analyses of the Afrikaner could provide information on admixture dynamics. The Afrikaner is a popular population for genetic studies, due to its homogeneity and small founder population.

This study proved that neurotransmitter gene's allele frequencies differ between different South African ethnic groups. Linguistic related groups showed a higher degree of genetic relation in comparison to non-linguistic related groups. Neurotransmitter genes were determined significant factors in past human migrations. Populations that migrated the greatest distances had the highest levels of novelty seeking alleles.

CHAPTER SIX

SUMMARY ENGLISH

Inadequate information on the allelic frequency of neurotransmitter genes exist among South African ethnic groups. The analyses of neurotransmitter genes with populations world-wide have benefited behaviour, population and medical fields. Population genetic fields have benefited due to increased information on population differentiation, structure and variation.

Neuropsychiatric and behavioural disorders are prevalent among South African populations. In addition, research on neurotransmitter genes is limited for South African populations. Information on South African populations is inadequately based on one or two population groups and a limited number of neurotransmitter genes. This lack of information prompted the analyses of the Khwe, Xhosa, Sotho and Afrikaner ethnic groups with the *DRD4*, *MAO-A* and *5-HTT* (5-HTTLPR) in the current study.

Volunteers from the mentioned South African populations provided saliva samples. In total, 349 individuals were successfully analysed for the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR. Significant genetic diversity was determined among the different populations. The Xhosa and Sotho are the closest related populations in the study followed by the Xhosa and Khwe. The Afrikaner population, an African population with mostly European ancestral heritage is the most genetic divergent population. From the analysis of the South African populations a hypothesis was constructed that the *DRD4* and 5-HTTLPR allele frequencies may be associated with historic migration distance due to novelty seeking.

It was theorised that populations that have migrated the greatest distance from Africa would have the highest levels of novelty seeking alleles. With the use of the South African populations and world-wide populations from revised literature positive correlations were identified for the S allele of the 5-HTTLPR. Positive correlations were also determined for the 7-repeat, >5-repeat and the combined 2- and >5-repeat alleles of the *DRD4* VNTR.

The 4-repeat allele of the *DRD4* VNTR was determined the greatest among the analysed Khwe, Xhosa, Sotho and Afrikaner ethnic groups. The Khwe population had the highest frequency (0.711) for the 4-repeat allele among the analysed populations. The Khwe, Xhosa and Sotho populations *DRD4* frequencies were similar to other African populations. The Afrikaner population has similar frequencies to its ancestral populations for the *DRD4* VNTR.

A remarkable finding of the study is the high frequencies of the long 5-HTTLPR alleles (L, VL and XL) in the Khwe population. The Khwe 5-HTTLPR allele frequencies are similar to that of the pygmy Mbuti population from central Africa. The Khwe is the most genetically diverse population analysed in this study and is possibly associated with the first modern humans. In contrast to the Khwe, the Afrikaner population has reduced L allele frequencies and less genetic diversity.

The 4.5- and 3.5-repeat *MAO-A*-uVNTR alleles were the most common among the analysed populations. A significant finding in this study was that the Khwe, Xhosa and Sotho 3.5-repeat allele frequencies were significantly reduced compared to African Americans. The analysis of the study populations with the *MAO-A*-uVNTR provided information to an area that is lacking. Few correlations were previously possible due to the *MAO-A*-VNTR not being analysed in African populations.

This study proved that allele frequencies for neurotransmitter genes differ between South African ethnic groups. A high degree of genetic relation was determined among linguistically related groups (Xhosa and Sotho) compared to non-related groups (Afrikaner). Neurotransmitter genes were determined significant factors in past human migrations. Novelty seeking alleles were associated with populations that migrated great distances.

Key words: Afrikaner, Allele frequency, *Dopamine Receptor D4*, Ethnic groups, Khwe, *Monoamine Oxidase A*, Serotonin transporter-linked polymorphic region, Sotho, South Africa and Xhosa.

CHAPTER SEVEN

SUMMARY AFRIKAANS

Daar bestaan tans onvoldoende inligting in verband met neuro-oordrag gene onder Suid-Afrikaanse etniese groepe. Die ontleding van neuro-oordrag gene onder bevolkings wêreldwyd het gedrag, bevolking en mediese studievelde in die betrokke lande gebaat. Genetiese bevolkingstudies het voordeel getrek deur die verhoogde inligting oor die onderskeid, struktuur en variasie in die bevolkings.

Neuropsigiatriese afwykings en gedragsafwykings is algemeen onder Suid-Afrikaanse bevolkings. Daarbenewens is navorsing oor neuro-oordrag gene beperk vir Suid-Afrikaanse bevolkings. Neuro-oordrag inligting van Suid-Afrikaanse bevolkings is onvolledig en meestal gebaseer op ‘n beperkte aantal neuro-oordrag gene en fokus slegs op een of twee bevolkingsgroepe. Derhalwe is hierdie studie van die *DRD4*, *MAO-A* en *5-HTT* (5-HTTLPR) gebiede op die Khwe, Xhosa, Sotho en Afrikaner etniese groepe uitgevoer.

Vrywilligers van die betrokke Suid-Afrikaanse bevolkings groepe het speeksel monsters voorsien. ’n Totaal van 349 monsters is suksesvol ontleed vir die *DRD4*, *MAO-A* en 5-HTTLPR. Die genetiese diversiteit tussen die verskillende bevolkings is beduidend. Die Xhosa en Sotho is die mees verwante bevolkings in die studie, gevolg deur die Xhosa en Khwe. Die Afrikaner-bevolking, 'n Afrika-bevolking met ’n Europese oorsprong, is die mees geneties divergerende bevolking. Die data het gelei tot die formulering van 'n hipotese dat die *DRD4* ‘VNTR’ en 5-HTTLPR alleelfrekwensies geassosieer kan word met historiese migrasie afstand, moontlik as gevolg van ’n verhoogde drang na die soek van nuwe ervarings (‘novelty seeking’).

Dit word beweer dat bevolkings wat oor die grootste afstand uit Afrika migreer het, die hoogste vlak van die drang na nuwe ervarings sou hê. Deur die Suid-Afrikaanse bevolking met wêreld-wye bevolkings te kombineer, is positiewe korrelasies tussen die afstande en die alleelfrekwensie geïdentifiseer vir die S alleel van die 5-HTTLPR. Positiewe korrelasies is ook geïdentifiseer vir die 7-, >5- en die gekombineerde 2- en >5-herhaling allele van die *DRD4* ‘VNTR’.

Dit was bepaal dat die 4-herhaling alleel van die *DRD4* ‘VNTR’ die hoogste frekwensie onder die Khwe, Xhosa, Sotho en Afrikaner etniese groepe het. Die Khwe bevolking het die hoogste frekwensie (0.711) vir die 4-herhaling alleel onder die studie bevolkings. Die *DRD4* frekwensie van die Khwe, Xhosa en Sotho bevolkings is soortgelyk aan ander Afrika-bevolkings. Die Afrikaner-bevolking het soortgelyke frekwensies as sy voorvaderlike bevolkings vir die *DRD4* ‘VNTR’.

'n Merkwaardige bevinding van die studie is die hoë frekwensies van die lang 5-HTTLPR alleel (‘L’, ‘XL’ en ‘VL’) in die Khwe bevolking. Khwe frekwensies is soortgelyk aan dié van die dwerg Mbuti bevolking van Sentraal-Afrika. Die Khwe is die mees genetiese diverse bevolking ontleed in hierdie studie en hou moontlik verband met die eerste moderne mense. In teenstelling met die Khwe, het die Afrikaner-bevolking ‘n verminderde ‘L’ alleelfrekwensie en minder genetiese diversiteit.

Die 4.5- en 3.5-herhaling *MAO-A*-‘uVNTR’ allele was die mees algemeen onder die studie bevolkings. 'n Beduidende bevinding in die studie was dat die Khwe, Xhosa en Sotho 3.5-herhaling alleelfrekwensies aansienlik minder is in vergelyking met Afro-Amerikaners. Die ontleding van die studie bevolkings met die *MAO-A*-‘uVNTR’ het inligting verskaf om 'n gebied wat te kort het. Net ‘n paar korrelasies was voorheen moontlik as gevolg van die *MAO-A*-‘uVNTR’ wat nie in Afrikaanse bevolkings ontleed was nie.

Hierdie studie bewys dat alleelfrekwensies vir neuro-oordrag gene verskil tussen Suid-Afrikaanse etniese groepe. 'n Hoë graad van genetiese verhouding is bepaal onder taalkundig verwante groepe (Xhosa en Sotho) in vergelyking met nie-verwante groepe (Afrikaner). Neuro-oordrag gene was belangrike faktore bepaal onder menslike migrasies van die verlede. n’ Verhoogde drang na nuwe ervaring allele hou verband met bevolkings wat lang afstande gemigreer het.

Sleutelwoorde: Afrikaner, Alleelfrekwensie, *Dopamien Reseptor D4*, Etniese groepe, Khwe, *Monoamien Oksidase A*, Serotonin karweier-gekoppel polimorfiese streek, Sotho, Suid-Afrika en Xhosa.

CHAPTER EIGHT

REFERENCES

Abel, E.L. 1995. An update on incidence of FAS: FAS is not an equal opportunity birth defect. *Neurotoxicology and Teratology*, 17(4): 437-443.

Achilli, A., Perego, U.A., Lancioini, H., Olivieri, A., Gandini, F., Kashani, B.H., Battaglia, V., Grugni, V., Angerhofer, N., Rogers, M.P., Herrera, R.J., Woodward, S.R., Labuda, D., Smith, D.G., Cybulski, J.S., Semino, O., Malhi, R.S. and Torroni, A. 2013. Reconciling migration models to the Americas with the variation of North American native mitogenomes. *Proceedings of the National Academy of Sciences of the United States of America*, 110(35): 14308-14313.

Aguirre-Samudio, A.J. and Nicolini, H. 2005. *DRD4* polymorphism and the association with mental disorders. *Revista de Investigación Clínica*, 57(1): 65-75.

Arieff, Z., Kaur, M., Gameeldien, H., van der Merwe, L. and Bajic, V.B. 2010. 5-HTTLPR polymorphism: Analysis in South African autistic individuals. *Human Biology*, 82(3): 291-300.

Asghari, V., Sanyal, S., Buchwaldt, S., Paterson, A., Jovanovic, V. and Van Tol, H.H.M. 1995. Modulation of intracellular cyclic AMP levels by different human Dopamine D4 Receptor variants. *Journal of Neurochemistry*, 65(3): 1157-1165.

Bach, A.W., Lan, N.C., Johnson, D.L., Abell, C.W., Bembenek, M.E., Kwan, S.W., Seeburg, P.H. and Shih, J.C. 1988. cDNA cloning of human liver Monoamine Oxidase A and B: Molecular basis of differences in enzymatic properties. *Proceedings of the National Academy of Sciences of the United States of America*, 85(13): 4934-4938.

Bachmann, H.S., Siffert, W. and Frey, U.H. 2003. Successful amplification of extremely GC-rich promoter region using a novel ‘slowdown PCR’ technique. *Pharmacogenetics*, 13(12): 759-766.

Bakermans-Kranenburg, M.J., Dobrova-Krol, N. and van IJzendoorn, M. 2011. Impact of institutional care on attachment disorganisation and insecurity of

Ukrainian pre-schoolers: Protective effect of the long variant of the *Serotonin Transporter* gene (*5HTT*). *International Journal of Behavioral Development*: 1-8.

Bakermans-Kranenburg, M.J. and van IJzendoorn, M.H. 2006. Gene-environment interaction of the *Dopamine D4 Receptor* (*DRD4*) and observed maternal insensitivity predicting externalizing behavior in preschoolers. *Developmental Psychobiology*, 48(5): 406-409.

Balamurugan, K., Kanthimathi, S., Vijaya, M., Suhasini, G., Duncan, G., Tracey, M. and Budowle, B. 2010. Genetic variation of 15 autosomal microsatellite loci in a Tamil population from Tamil Nadu, Southern India. *Legal Medicine*, 12(6): 320-323.

Bamshad, M and Wooding, S.P. 2003. Signatures of natural selection in the human genome. *Nature Review*, 4(2): 99-111.

Bellivier, F., Henry, C., Szöke, A., Schürhoff, F., Nosten-Bertrand, M., Feingold, J., Launay, J.M., Leboyer, M. and Laplanche, J.L. 1998. *Serotonin Transporter* gene polymorphisms in patients with unipolar or bipolar depression. *Neuroscience Letters*, 255(3): 143-146.

Ben-Jonathan, N. and Hnasko, R. 2001. Dopamine as a prolactin (PRL) inhibitor. *Endocrine Reviews*, 22(6): 724-763.

Berlin, I. 1998. *Many thousands gone: The first two centuries of slavery in North America*. Cambridge, Massachusetts: Harvard University Press.

Bisso-Machado, R., Ramallo, V., Tarazona-Santos, E., Salzano, F.M., Bortolini, M.C. and Hünemeier, T. 2013. 5-HTTLPR genetic diversity and mode of subsistence in Native Americans. *American Journal of Physical Anthropology*, 151(3): 492-494.

Bleek, W.H.I.B. 1862. *A comparative grammar of South African languages*. Volume 2. London: Trübner & Co.

Booth, D. 1998. *The race game: Sport and politics in South Africa.* London: Frank Cass.

Botha, M.C. and Beighton, P. 1983. Inherited disorders in the Afrikaner population of southern Africa. Part I. Historical and demographic background,

cardiovascular, neurological, metabolic and intestinal conditions. *South African Medical Journal*, 64(16): 609-612.

Brune, C.W., Kim, S.J., Salt, J., Leventhal, B.L., Lord, C. and Cook Jr. E.H. 2006. 5-HTTLPR genotype-specific phenotype in children and adolescents with autism. *The American Journal of Psychiatry*, 163(12): 2148-2156.

Brunner, H.G., Nelen, M., Breakefield, X.O., Ropers, H.H. and van Oost, B.A. 1993. Abnormal behavior associated with a point mutation in the structural gene for Monoamine Oxidase A. *Science*, 262(5133): 578- 580.

Buckholtz, J.W. and Meyer-Lindenberg, A. 2008. *MAOA* and the neurogenetic architecture of human aggression. *Trends in Neurosciences*, 31(3): 120-129.

Burchard, E.G., Ziv, E., Coyle, N., Gomez, S.L., Tang, H., Karter, A.J., Mountain, J.L., Pérez-Stable, E.J., Sheppard, D. and Risch, N. 2003. The importance of race and ethnic background in biomedical research and clinical practice. *The New England Journal of Medicine*, 348(12): 1170-1175.

Caballero, A. and Hill, W.G. 1992. Effects of partial inbreeding on fixation rates and variation of mutant genes. *Genetics*, 131(2): 493-507.

Cai, X., Qin, Z., Wen, B., Xu, S., Wang, Y., Lu, Y., Wei, L., Wang, C., Li, S., Huang, X., Jin, L., Li, H. and the Genographic Consortium. 2011. Human migration through bottlenecks from Southeast Asia into East Asia during last glacial maximum revealed by Y chromosome. *PLoS ONE*, 6(8): 1-10.

Camarena, B., Loyzaga, C., Aguilar, A., Weissbecker, K. and Nicolini, H. 2007. Association study between the *Dopamine Receptor D_4* gene and obsessive-compulsive disorder. *European Nueropsychopharmacology*, 17(6-7): 406-409.

Campbell, M.C. and Tishkoff, S.A. 2008. African genetic diversity: Implications for human demographic history, modern human origins, and complex disease mapping. *Annual Review of Genomics and Human Genetics*, 9: 403-433

Carrel, L., Cottle, A.A., Goglin, K.C. and Willard, H.F. 1999. A first-generation X-inactivation profile of the human X chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, 96(25): 14440-14444.

Carrel, L. and Willard, H.F. 2005. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*, 434(7031): 400-

404.

Caspi, A., McClay, J., Moffitt, T.E., Mill, J., Martin, J., Craig, I.W., Taylor, A. and Poulton, R. 2002. Role of genotype in the cycle of violence in maltreated children. *Science*, 297(5582): 851-854.

Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A. and Poulton, R. 2003. Influence of life stress on depression: Moderation by a polymorphism in the *5-HTT* gene. *Science*, 301(5631), 386-389.

Catalano, M., Nobile, M., Novelli, E., Nöthen, M.M. and Smeraldi, E. 1993. Distribution of a novel mutation in the first exon of the human *Dopamine D4 Receptor* gene in psychotic patients. *Biological Psychiatry*, 34(7): 459-464.

Cavalli-Sforza, L.L. 1997. Genes, peoples, and languages. *Proceedings of the National Academy of Sciences of the United States of America*, 94(15): 7719-7724.

Cavalli-Sforza, L.L. and Feldman, M.W. 2003. The application of molecular genetic approaches to the study of human evolution. *Nature Genetics*, 33: 266-275.

Cervilla, J.A., Rivera, M., Molina, E., Torres-González, F., Bellón, J.A., Moreno, B., de Dios Luna, J., Lorente, J.A., de Diego-Otero, Y., King, M., Nazareth, I., Gutiérrez, B. and Predict Study Core Group. 2006. The 5-HTTLPR s/s genotype at the *Serotonin Transporter* gene (*SLC6A4*) increases the risk for depression in a large cohort of primary care attendees: The predict-gene study. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 141B(8): 912-917.

Chang, F. M., Kidd, J.R., Livak, K.J., Pakstis, A.J. and Kidd, K.K. 1996. The world-wide distribution of allele frequencies at the human *Dopamine D4 Receptor* locus. *Human Genetics*, 98(1): 91-101.

Chen, C., Burton, M., Greenberger, E. and Dmitrieva, J. 1999. Population migration and the variation of *Dopamine D4 Receptor* (*DRD4*) allele frequencies around the globe. *Evolution and Human Behavior*, 20 (5): 309-324.

Chen, L., Bruford, M.W., Xu, S., Zhou, K. and Yang, G. 2010. Microsatellite variation and significant population genetic structure of endangered finless porpoises (*Neophocaena phocaenoides*) in Chinese coastal waters and the Yangtze River. *Marine Biology*, 157(7): 1453-1462.

Chen, Y.S., Olckers, A., Schurr, T.G., Kogelnik, A.M., Huoponen, K. and Wallace, D.C. 2000. mtDNA variation in the South African Kung and Khwe-and their genetic relationships to other African populations. *American Journal of Human Genetics*, 66(4): 1362-1383.

Choi, J.S., Kim, J.S., Joe, C.O., Kim, S., Ha, K.S. and Park, Y.M. 1999. Improved cycle sequencing of GC-rich DNA template. *Experimental and Molecular Medicine*, 31(1): 20-24.

Chumbo, S. and Mmaba, K. 2002. *The Khwe of the Okavango panhandle: The past life*. Shakawe: The Teemacane Trust.

Clarke, H., Flint, J., Attwood, A.S. and Munafò, M.R. 2010. Association of the 5-HTTLPR genotype and unipolar depression: a meta-analysis. *Psychological Medicine*, 40(11): 1767-1778.

Cohen, I.L., Liu, X., Lewis, M.E.S., Chudley, A., Forster-Gibson, C., Gonzalez, M., Jenkins, E.C., Brown, W.T. and Holden, J.J.A. 2011. Autism severity is associated with child and maternal *MAOA* genotypes. *Clinical Genetics*, 79(4): 355-362.

Cohen, I.L., Liu, X., Schutz, C., White, B.N., Jenkins, E.C., Brown, W.T. and Holden, J.J.A. 2003. Association of autism severity with a *Monoamine Oxidase A* functional polymorphism. *Clinical Genetics*, 64(3): 190-197.

Colucci, R., Gambaccini, D., Ghisu, N., Rossi, G., Costa, F., Tuccori, M., De Bortoli, N., Fornai, M., Antonioli, L., Ricchiuti, A., Mumolo, M.G., Marchi, S., Blandizzi, C. and Bellini, M. 2013. Influence of the serotonin transporter 5HTTLPR polymorphism on symptom severity in irritable bowel syndrome. *PLoS ONE*, 8(2): 1-6.

Contini, V., Marques, F.Z.C., Garcia, C.E.D., Hutz, M.H. and Bau, C.H.D. 2006. *MAO*-uVNTR polymorphism in a Brazilian sample: Further support for the association with impulsive behaviors and alcohol dependence. *American*

Journal of Medical Genetics Part B (Neuropsychiatric Genetics), 141B(3): 305-308.

Coop, G., Witonsky, D., Di Rienzo, A. and Pritchard, J.K. 2010. Using environmental correlations to identify loci underlying local adaptation. *Genetics*, 185(4): 1411-1423.

Cox, G., Sealy, J., Schrire, C. and Morris, A. 2001. Stable carbon and nitrogen isotopic analyses of the underclass at the colonial Cape of Good Hope in the eighteenth and nineteenth centuries. *World Archaeology*, 33(1): 73-97.

Cox, M.P., Karafet, T.M., Lansing, J.S., Sudoyo, H. and Hammer, M.F. 2010. Autosomal and X-linked single nucleotide polymorphisms reveal a steep Asian-Melanesian ancestry cline in eastern Indonesia and a sex bias in admixture rates. *Proceedings of the Royal Society Biology*, 277(1687): 1589-1596.

Creswell, K.G., Sayette, M.A., Manuck, S.B., Ferrell, R.E., Hill, S.Y. and Dimoff, J.D. 2012. *DRD4* polymorphism moderates the effect of alcohol consumption on social bonding. *PLoS ONE*, 7(2): 1-9.

Cruz, C., Camarena, B., King, N., Páez, F., Sidenberg, D., de la Fuente, J.R. and Nicolini, H. 1997. Increased prevalence of the seven-repeat variant of the *Dopamine D4 Receptor* gene in patients with obsessive-compulsive disorder with tics. *Neuroscience Letters*, 231(1): 1-4.

Das, M., Bhowmik, A.D., Sinha, S., Chattopadhyay, A., Chaudhuri, K., Singh, M. and Mukhopadhyay, K. 2006. *MAOA* promoter polymorphism and attention deficit hyperactivity disorder (ADHD) in Indian children. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 141B(6): 637-642.

Davenport, T.R.H. 1978. *South Africa: A modern history*. 2nd edition. Hong Kong: The Macmillan Press Ltd.

De Almeida, A. 1965. *Bushmen and other non-Bantu peoples of Angola.* Johanesburg: Witwatersrand University Press.

Deckert, J., Catalano, M., Syagailo, Y.V., Bosi, M., Okladnova, O., Di Bella, D., Nöthen, M.M., Maffei, P., Franke, P., Fritze, J., Maier, W., Propping, P., Beckmann, H., Bellodi, L. and Lesch, K.P. 1999. Excess of high activity

Monoamine Oxidase A gene promoter alleles in female patients with panic disorder. *Human Molecular Genetics*, 8(4): 621-624.

DeGiorgio, M., Jakobsson, M. and Rosenberg, N.A. 2009. Explaining worldwide patterns of human genetic variation using a coalescent-based serial founder model of migration outward from Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 106(38): 16057-16062.

De Milander, L., Stein, D.J. and Collins, M. 2009. The *Interleukin-6*, *Serotonin Transporter*, and *Monoamine Oxidase A* genes and endurance performance during the South African ironman triathlon. *Applied Physiology, Nutrition, and Metabolism*, 34(5): 858-865.

Den Hertog, T.N. 2013. Diversity behind constructed unity: The resettlement process of the !Xun and Khwe communities in South Africa. *Journal of Contemporary African Studies*, 31(3): 345-360.

Denys, D., Van Nieuwerburgh, F., Deforce, D. and Westenberg, H.G.M. 2006. Association between serotonergic candidate genes and specific phenotypes of obsessive compulsive disorder. *Journal of Affective Disorders*. 91(1): 39-44.

Ding, Y.C., Chi, H.C., Grady, D.L., Morishima, A., Kidd, J.R., Kidd, K.K., Flodman, P., Spence, M.A., Schuck, S., Swanson, J.M., Zhang, Y.P. and Moyziz, R.K. 2002. Evidence of positive selection acting at the human *Dopamine Receptor D4* gene locus. *Proceedings of the National Academy of Sciences of the United States of America*. 99(1): 309-314.

Dreber, A., Apicella, C.L., Eisenberg, D.T.A., Garcia, J.R., Zamore, R.S., Lum, J.K. and Campbell, B. 2009. The 7R polymorphism in the *Dopamine D_4* gene (*DRD4*) is associated with financial risk taking in men. *Evolution and Human Bahavior*, 30(2): 85-92.

Dreber, A., Rand, D.G., Wernerfelt, N., Garcia, J.R., Lum, J.K. and Zeckhauser, R. 2011. The *Dopamine Receptor D4* gene (*DRD4*) and self-reported risk taking in the economic domain. *HKS Faculty Research Working Paper Series*: 1-16.

Ducci, F., Enoch, M.A., Hodgkinson, C., Xu, K., Catena, M., Robin, R.W. and Goldman, D. 2008. Interaction between a functional *MAOA* locus and childhood sexual abuse predicts alcoholism and antisocial personality disorder

in adult women. *Molecular Psychiatry*, 13(3): 334-347.

Ebstein, R.P., Nemanov, L., Klotz, I., Gritsenko, I. and Belmaker, R.H. 1997. Additional evidence for an association between the *Dopamine D4 Receptor* (*D4DR*) exon lll repeat polymorphism and the human personality trait of novelty seeking. *Molecular Psychiatry*, 2(6): 472-477.

Ehli, E.A., Hu, Y., Lengyel-Nelson, T., Hudziak, J.J. and Davies, G.E. 2012. Identification and functional characterization of three novel alleles for the serotonin transporter-linked polymorphic region. *Molecular Psychiatry*, 17(2): 185-192.

Eley, T.C., Tahir, E., Angleitner, A., Harriss, K., McClay, J., Plomin, R., Riemann, R., Spinath, F. and Craig, I. 2003. Association analysis of *MAOA* and *COMT* with neuroticism assessed by peers. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 120B(1): 90-96.

Elphick, R.H. 1972. *The Cape Khoi and the first phase of South African race relations*. Ph.d. thesis, Yale University, United Sates of America.

Enard, W., Przeworski, M., Fisher, S.E., Lai, C.S.L., Wiebe, V., Kitano, T., Monaco, A.P. and Pääbo, S. 2002. Molecular evolution of *FOXP2*, a gene involved in speech and language. *Nature*, 418(6900): 869-872.

Esau, L., Kaur, M., Adonis, L. and Arieff, Z. 2008. The 5-HTTLPR polymorphism in South African healthy populations: A global comparison. *Journal of Neural Transmission*, 155(5): 755-760.

Ewens, W.J. and Spielman, R.S. 1995. The transmission/disequilibrium test: History, subdivision, and admixture. *American Journal of Human Genetics*, 57(2): 455-464.

Feinn, R., Nellissery, M. and Kranzler, H.R. 2005. Meta-analysis of the association of a functional serotonin transporter promoter polymorphism with alcohol dependence. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 133B(1): 79-84.

Fergusson, D.M., Boden, J.M., Horwood, L.J., Miller, A.L. and Kennedy, M.A. 2011. *MAOA*, abuse exposure and antisocial behaviour: 30-year longitudinal study. *The British Journal of Psychiatry*, 198(6): 457-463.

Florez, J.C., Price, A.L., Campbell, D., Riba, L., Parra, M.V., Yu, F., Duque, C., Saxena, R., Gallego, N., Tello-Ruiz, M., Franco, L., Rodríguez-Torres, M., Villegas, A., Bedoya, G., Aguilar-Salinas, C.A., Tusíe-Luna, M.T., Ruiz-Linares, A and Reich, D. 2009. Strong association of socioeconomic status with genetic ancestry in Latinos: Implications for admixture studies of type 2 diabetes. *Diabetologia*, 52(8): 1528-1536.

Foster, M.W. and Sharp, R.R. 2002. Race, ethnicity, and genomics: Social classifications as proxies of biological heterogeneity. *Genome Research*, 12(6): 844-850.

Friedlaender, J.S., Friedlaender, F.R., Reed, F.A., Kidd, K.K., Kidd, J.R., Chambers, G.K., Lea, R.A., Loo, J.H., Koki, G., Hodgson, J.A., Merriwether, D.A. and Weber, J.L. 2008. The genetic structure of Pacific Islanders. *PLoS Genetics*, 4(1): 0173-0190.

Gallardo-Pujol, D., Andrés-Pueyo, A. and Maydeu-Olivares, A. 2013. *MAOA* genotype, social exclusion and aggression: An experimental test of a gene-environment interaction. *Genes, Brain and Behavior*, 12(1): 140-145.

Garcia, J.R., MacKillop, J., Aller, E.L., Merriwether, A.M., Wilson, D.S. and Lum, J.K. 2010. Associations between *Dopamine D4 Receptor* gene variation with both infidelity and sexual promiscuity. *PLoS ONE*, 5(11): 1-6.

Garpenstrand**,** H., Norton, N., Damberg, M., Rylander, G., Forslund, K., Mattila-Evenden, M., Gustavsson, J.P., Ekblom, J., Oreland, L., Bergman, H., Owen, M.J. and Jönsson, E.G. 2002. A regulatory *Monoamine Oxidase A* promoter polymorphism and personality traits. *Neuropsychobiology*, 46(4): 190-193.

Gelernter, J., Cubells, J.F., Kidd, J.R., Pakstis, A.J. and Kidd, K.K. 1999. Population studies of polymorphisms of the *Serotonin Transporter* protein gene. *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, 88(1): 61-66.

Gelernter, J., Kranzler, H., Coccaro, E., Siever, L., New, A. and Mulgrew, C.L. 1997. *D4 Dopamine-Receptor* (*DRD4*) alleles and novelty seeking in substance-dependent, personality-disorder, and control subjects. *American Journal of Human Genetics*, 61(5): 1144-1152.

Gerretsen, P and Pollock, B.G. 2008. Pharmacogenetics and the Serotonin Transporter in late-life depression. *Expert Opinion on Drug Metabolism & Toxicology*, 4(12): 1465-1478.

Ghosh, A. and Seshadri, M. 2005. Indian ethnic populations characterized by *Dopamine (D4) Receptor* VNTR polymorphism. *Annals of Human Biology*, 32(5): 574-584.

Gilad, Y., Rosenberg, S., Przeworski, M., Lancet, D. and Skorecki, K. 2002. Evidence for positive selection and population structure at the human *MAO-A* gene. *Proceedings of the National Academy of Sciences of the United States of* America, 99(2): 862-867.

Gonçalves, V.F., Tiwari, A.K., de Luca, V., Kong, S.L., Zai, C., Tampakeras, M., Mackenzie, B., Sun, L. and Kennedy, J.L. 2012. *DRD4* VNTR polymorphism and age at onset of severe mental illnesses. *Neuroscience Letters*, 519(1): 9-13.

Gonda, X., Fountoulakis, K.N., Csukly, G., Bagdy, G., Pap, D., Molnár, E., Laszik, A., Lazary, J., Sarosi, A., Faludi, G., Sasvari-Szekely, M., Szekely, A. and Rihmer, Z. 2011. Interaction of 5-HTTLPR genotype and unipolar major depression in the emergence of aggressive/hostile traits. *Journal of Affective Disorders*, 132(3): 432-437.

Grabe, H.J., Lange, M., Wolff, B., Völzke, H., Lucht, M., Freyberger, H.J., John, U. and Cascorbi, I. 2005. Mental and physical distress is modulated by a polymorphism in the *5-HT Transporter* gene interacting with social stressors and chronic disease burden. *Molecular Psychiatry*, 10(2): 220-224.

Grady, C.L., McIntosh, A.R., Beig, S., Keightley, M.L., Burian, H. and Black, S.E. 2003. Evidence from functional neuroimaging of a compensatory prefrontal network in Alzheimer's disease. *The Journal of Neuroscience*, 23(3): 986-993.

Greeff, J.M. 2007. Deconstructing Jaco: Genetic heritage of an Afrikaner. *Annals of Human Genetics*, 71(Pt 5): 674-688.

Greenberg, B.D., Tolliver, T.J., Huang, S.J., Li, Q., Bengel, D. and Murphy, D.L. 1999. Genetic variation in the serotonin transporter promoter region affects serotonin uptake in human blood platelets. *American Journal of Medical*

Genetics (Neuropsychiatric Genetics), 88(1): 83-87.

Grimaldi, R., Acosta, A.X., Machado, T.M.B., Bomfim, T.F. and Galvão-Castro, B. 2010. Distribution of *SDF1-3'*A polymorphisms in three different ethnic groups from Brazil. *The Brazilian Journal of Infectious Disease*, 14(2): 197-200.

Grimsby, J., Chen, K., Wang, L.J., Lan, N.C. and Shih, J.C. 1991. Human *Monoamine Oxidase A* and *B* genes exhibit identical exon-intron organization. *Proceedings of the National Academy of Sciences of the United States of* America, 88(9): 3637-3641.

Grossman, S.R., Shylakhter, I., Karlsson, E.K., Byrne, E.H., Morales, S., Frieden, G., Hostetter, E., Angelino, E., Garber, M., Zuk, O., Lander, E.S., Schaffner, S.F. and Sabeti, P.C. 2010. A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science*, 327(5967): 883-886.

Gründemann, D and Schömig, E. 1996. Protection of DNA during preparative agarose gel electrophoresis against damage induced by ultraviolet light. *BioTechniques*, 21(5): 898-903.

Gu, J., Spitz, M.R., Yang, F. and Wu, X. 1999. Ethnic differences in *poly* (*ADP-ribose*) *Polymerase* pseudogene genotype distribution and association with lung cancer risk. *Carcinogenesis*, 20(8): 1465-1469.

Guo, G., Ou, X.M., Roettger, M. and Shih, J.C. 2008. The VNTR 2 repeat in *MAOA* and delinquent behaviour in adolescence and young adulthood: Associations and *MAOA* promoter activity. *European Journal of Human Genetics*, 16(5): 626-634.

Gutiérrez**,** B., Pintor, L., Gastó, C., Rosa, A., Bertranpetit, J., Vieta, E. and Fañanás, L. 1998. Variability in the *Serotonin Transporter* gene and increased risk for major depression with melancholia. *Human Genetics*, 103(3): 319-322.

Haberstick, B.C., Lessem, J.M., Hopfer, C.J., Smolen, A., Ehringer, M.A., Timberlake, D. and Hewitt, J.K. 2005. *Monoamine Oxidase A* (*MAOA*) and antisocial behaviors in the presence of childhood and adolescent maltreatment. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 135B(1): 59-64.

Hakamata, Y., Takahashi, N., Ishihara, R., Saito, S., Ozaki, N., Honjo, S., Ono, Y. and Inada, T. 2005. No association between *Monoamine Oxidase A* promoter polymorphism and personality traits in Japanese females. *Neuroscience Letters*, 389(3): 121-123.

Hall, D., Wijsman, E.M., Roos, J.L., Gogos, J.A. and Karayiorgou, M. 2002. Extended intermarker linkage disequilibrium in the Afrikaner. *Genome Research*, 12(6): 956-961.

Hamblin, M.T. and Di Rienzo, A. 2000. Detection of the signatures of natural selection in humans: Evidence from the Duffy blood group locus. *American Journal of Human Genetics*: 66(5), 1669-1679.

Hamilton, S.P., Slager, S.L., Heiman, G.A., Haghighi, F., Klein, D.F., Hodge, S.E., Weissman, M.M., Fyer, A.J. and Knowles, J.A. 2000. No genetic linkage or association between a functional promoter polymorphism in the *Monoamine Oxidase-A* gene and panic disorder. *Molecular Psychiatry*, 5(5): 465-466.

Hare, M.L.C. 1928. Tyramine Oxidase: A new enzyme system in liver. *Biochemical Journal*, 22 (4): 968-979.

Hariri, A.R., Mattay, V.S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D., Egan, M.F. and Weinberger, D.R. 2002. *Serotonin Transporter* genetic variation and the response of the human amygdala. *Science*, 297(5580): 400-403.

Heils, A., Teufel, A., Petri, S., Seemann, M., Bengel, D., Balling, U., Riederer, P. and Lesch, K.P. 1995. Functional promoter and polyadenylation site mapping of the human *Serotonin* (*5-HT*) *Transporter* gene. *Journal of Neural Transmission*, 102(3): 247-254.

Heils, A., Teufel, A., Petri, S., Stöber, G., Riederer, P., Bengel, D. and Lesch, K.P. 1996. Allelic variation of human *Serotonin Transporter* gene expression. *Journal of Neurochemistry*, 66(6): 2621-2624.

Heinz, A., Jones, D.W., Mazzanti, C., Goldman, D., Ragan, P., Hommer, D., Linnoila. M. and Weinberger, D.R. 2000. A relationship between *Serotonin Transporter* genotype and *in vivo* protein expression and alcohol neurotoxicity. *Biological Psychiatry*, 47(7): 643-649.

Henn, B.M., Gignoux, C.R., Jobin, M., Granka, J.M., Macpherson, J.M., Kidd, J.M., Rodríguez-Botigué, L., Ramachandran, S., Hon, L., Brisbin, A., Lin, A.A., Underhill, P.A., Comas, D., Kidd, K.K., Norman, P.J., Parham, P., Bustamante, C.D., Mountain, J.L. and Feldman, M.W. 2011. Feature article: Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proceedings of the National Academy of Sciences of the United States of America*, 108: 5154-5162.

Herbert, R.K. 1990. The sociohistory of clicks in Southern Bantu. *Anthropological Linguistics*, 32(3-4): 295-315.

Herman, A.I., Kaiss, K.M., Ma, R., Philbeck, J.W., Hasan, A., Dasti, H. and DePetrillo, P.B. 2005. Serotonin transporter promoter polymorphism and *Monoamine Oxidase Type A* VNTR allelic variants together influence alcohol binge drinking risk in young women. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 133B(1): 74-78.

Hoefgen, B., Schulze, T.G., Ohlraun, S., von Widdern, O., Höfels, S., Gross, M., Heidmann, V., Kovalenko, S., Eckermann, A., Kölsch, H., Metten, M., Zobel, A., Becker, T., Nöthen, M.M., Propping, P., Heun, R., Maier, W. and Rietschel, M. 2005. The power of sample size and homogenous sampling: Association between the 5-HTTLPR *Serotonin Transporter* polymorphism and major depressive disorder. *Biological Psychiatry*, 57(3): 247-251.

Hoenicka, J., Aragüés, M., Ponce, G., Rodríguez-Jiménez, R., Jiménez-Arriero, M.A. and Palomo, T. 2007. From dopaminergic genes to psychiatric disorders. *Neurotoxicity Research*, 11(1): 61-71.

Hong, E.P. and Park, J.W. 2012. Sample size and statistical power calculation in genetic association studies. *Genomics & Informatics*, 10(2): 117-122.

Hosking, L., Lumsden, S., Lewis, K., Yeo, A., McCarthy, L., Bansal, A., Riley, J., Purvis, I. and Xu, C.F. 2004. Detection of genotyping errors by Hardy-Weinberg equilibrium testing. *European Journal of Human Genetics*, 12(5): 395-399.

Hsu, Y.P.P., Powel, J.F., Sims, K.B. and Breakfield, X.O. 1989. Molecular genetics of the Monoamine Oxidases. *Journal of Neurochemistry*, 53(1): 12-18.

Hsu, Y.P.P., Weyler, W., Chen, S., Sims, K.B., Rinehart, W.B., Utterback, M.C., Powell, J.F. and Breakefield, X.O. 1988. Structural features of human Monoamine Oxidase A elucidated from cDNA and peptide sequences. *Journal of Neurochemistry*, 51(4): 1321-1324.

Hu, X.Z., Lipsky, R.H., Zhu, G., Akhtar, L.A., Taubman, J., Greenberg, B.D., Xu, K., Arnold, P.D., Richter, M.A., Kennedy, J.L., Murphy, D.L. and Goldman, D. 2006. Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder. *The American Journal of Human Genetics*, 78(5): 815-826.

Huang, C.H. and Santangelo, S.L. 2008. Autism and *Serotonin Transporter* gene polymorphisms: A systematic review and meta-analysis. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 147B(6): 903-913.

Huang, S.Y., Lin, W.W., Wan, F.J., Chang, A.J., Ko, H.C., Wang, T.J., Wu, P.L. and Lu, R.B. 2007. *Monoamine Oxidase-A* polymorphisms might modify the association between the *Dopamine D_2 Receptor* gene and alcohol dependence. *Journal of Psychiatry & Neuroscience*, 32(3): 185-192.

Huang, Y.Y., Cate, S.P., Battistuzzi, C., Oquendo, M.A., Brent, D. and Mann, J.J. 2004. An association between a functional polymorphism in the *Monoamine Oxidase A* gene promoter, impulsive traits and early abuse experiences. *Neuropsychopharmacology*, 29(8): 1498-1505.

Huezo-Diaz, P., Uher, R., Smith, R., Rietschel, M., Henigsberg, N., Marusic, A., Mors, O., Maier, W., Hauser, J., Souery, D., Placentino, A., Zobel, A., Larsen, E.R., Czerski, P.M., Gupta, B., Hoda, F., Perroud, N., Farmer, A., Craig, I., Aitchison, K.J. and McGuffin, P. 2009. Moderation of antidepressant response by the *Serotonin Transporter* gene. *The British Journal of Psychiatry*, 195(1): 30-38.

Hutchison, K.E., Wooden, A., Swift, R.M., Smolen, A., McGeary, J., Adler, L. and Paris, L. 2003. Olanzapine reduces craving for alcohol: A *DRD4* VNTR polymorphism by pharmacotherapy interaction. *Neuropsychopharmacology*, 28(10): 1882-1888.

Ingman, M., Kaessmann, H., Pääbo, S. and Gyllensten, U. 2000. Mitochondrial genome variation and origin of modern humans. *Nature*, 408: 708-713.

Jacob, C.P., Müller, J., Schmidt, M., Hohenberger, K., Gutknecht, L., Reif, A., Schmidtke, A., Mössner, R. and Lesch, K.P. 2005. Cluster B personality disorders are associated with allelic variation of Monaomine Oxidase A activity. *Neuropsychopharmacology*, 30: 1711-1718.

Jardemark, K., Wadenberg, M.L., Grillner, P. and Svensson, T.H. 2002. Dopamine D3 and D4 receptor antagonists in the treatment of schizophrenia. *Current Opinion in Investigational Drugs*, 3(1): 101-105.

Jones, M.B., Palmour, R.M., Zwaigenbaum, L. and Szatmari, P. 2004. Modifier effects in autism at the *MAO-A* and *DBH* loci. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 126B(1): 58-65.

Jorde, L.B., Bamshad, M. and Rogers, A.R. 1998. Using mitochondrial and nuclear DNA markers to reconstruct human evolution. *BioEssays*, 20(2): 126-136.

Jorde, L.B., Watkins, W.S., Bamshad, M.J., Dixon, M.E., Ricker, C.E., Seielstad, M.T. and Batzer, M.A. 2000. The distribution of human genetic diversity: A comparison of mitochondrial, autosomal, and Y-chromosome data. *American Journal of Human Genetics*, 66(3): 979-988.

Jorm, A.F., Henderson, A.S., Jacomb, P.A., Christensen, H., Korten, A.E., Rodger, B., Tan, X. and Easteal, S. 2000. Association of a functional polymorphism of the *Monoamine Oxidase A* gene promoter with personality and psychiatric symptoms. *Psychiatric Genetics*, 10(2): 87-90.

Karg, K., Burmeister, M., Shedden, K. and Sen, S. 2011. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: Evidence of genetic moderation. *Archives of General Psychiatry*, 68(5): 444-454.

Kawada, Y., Hattori, M., Dai, X.Y. and Nanko, S. 1995. Possible association between *Monoamine Oxidase A* gene and bipolar affective disorder. *The American Journal of Human Genetics*, 56(1): 335-336.

Keim, C.A. 1995. Africa and Europe before 1900. In: Martin, P.M. and O'Meara, P. (eds.) *Africa*. 3rd edition. Bloomington, Indiana: Indiana University Press.

Kenna, G.A., Swift, R.M., Hillemacher, T. and Leggio, L. 2012. The relationship

of appetitive, reproductive and posterior pituitary hormones to alcoholism and craving in humans. *Neuropsychology Review*, 22(3): 211-28.

Kenna, G.A., Zywiak, W.H., McGeary, J.E., Leggio, L., McGeary, C., Wang, S., Grenga, A. and Swift, R.M. 2009. A within-group design of nontreatment seeking 5-HTTLPR genotyped alcohol-dependent subjects receiving Ondansetron and Sertraline. *Alcoholism: Clinical and Experimental Research*, 33(2): 315-323.

Kim, J.M., Stewart, R., Kim, S.W., Yang, S.J., Shin, I.S., Kim, Y.H. and Yoon, J.S. 2007. Interactions between life stressors and susceptibility genes (5-HTTLPR and *BDNF*) on depression in Korean Elders. *Biological Psychiatry*, 62(5): 423-428.

Kim, S.J. and Kim, C.H. 2006. The genetic studies of obsessive-compulsive disorder and its future directions. *Yonsei Medical Journal*, 47(4): 443-454.

Kluger, A.N., Siegfried, Z. and Ebstein, R.P. 2002. A meta-analysis of the association between *DRD4* polymorphism and novelty seeking. *Molecular Psychiatry*, 7(7): 712-717.

Knoll, J. and Magyar, K. 1972. Some puzzling pharmacological effects of Monoamine Oxidase inhibitors. *Advances in Biochemical Psychopharmacology*, 5: 393-408.

Kochersperger, L.M., Parker, E.L., Siciliano, M., Darlington, G.J. and Denney, R.M. 1986. Assignment of genes for human Monoamine Oxidase A and B to the X chromosome. *Journal of Neuroscience Research*, 16(4): 601-616.

Koen, L., Kinnear, C.J., Corfield, V.A., Emsley, R.A., Jordaan, E., Keyter, N., Moolman-Smook, J.C., Stein, D.J. and Niehaus, D.J.H. 2004. Violence in male patients with schizophrenia: Risk markers in a South African population. *Australian and New Zealand Journal of Psychiatry*, 38(4): 254-259.

Kogan, S.M., Beach, S.R.H., Philibert, R.A., Brody, G.H., Chen, Y.F. and Lei, M.K. 2010. 5-HTTLPR status moderates the effect of early adolescent substance use on risky sexual behavior. *Health Psychology*, 29(5): 471-476.

Koller, G., Bondy, B., Preuss, U.W., Bottlender, M. and Soyka, M. 2003. No association between a polymorphism in the promoter region of the *MAOA*

gene with antisocial personality traits in alcoholics. *Alcohol and Alcoholism*, 38(1): 31-34.

Kranzler, H.R., Armeli, S., Tennen, H., Covault, J., Feinn, R., Arias, A.J., Pettinati, H. and Oncken, C. 2011. A double-blind, randomized trial of Sertraline for alcohol dependence: Moderation by age of onset and 5-HTTLPR. *Journal of Clinical Psychopharmacology*, 31(1): 22-30

Kuhnen, C.M. and Chiao, J.Y. 2009. Genetic determinants of financial risk taking. *PLoS ONE*, 4(2): 1-4.

Kunugi, H., Hattori, M., Kato, T., Tatsumi, M., Sakai, T., Sasaki, T., Hirose, T. and Nanko, S. 1997. *Serotonin Transporter* gene polymorphisms: Ethnic difference and possible association with bipolar affective disorder. *Molecular Psychiatry*, 2(6): 457-462.

Lahti, J., Räikkönen, K., Ekelund, J., Peltonen, L., Raitakari, O.T. and Keltikangas-Järvinen, L. 2005. Novelty seeking: Interaction between parental alcohol use and *Dopamine D4 Receptor* gene exon III polymorphism over 17 years. *Psychiatric Genetics*, 15(2): 133-139.

Lander, E.S. and Schork, N.J. 1994. Genetic dissection of complex traits. *Science*, 265(5181): 2037-2048.

Lane, A.B., Soodyall, H., Arndt, S., Ratshikhopha, M.E., Jonker, E., Freeman, C., Young, L., Morar, B. and Toffie, L. 2002. Genetic substructure in South African Bantu-speakers: Evidence from autosomal DNA and Y-chromosome studies. *American Journal of Physical Anthropology*, 119: 175-185.

Lasky-Su, J.A., Faraone, S.V., Glatt, S.J. and Tsuang, M.T. 2005. Meta-analysis of the association between two polymorphisms in the *Serotonin Transporter* gene and affective disorders. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 133B(1): 110-115.

Latter, B.D.H. 1972. Selection in finite populations with multiple alleles. III. Genetic divergence with centripetal selection and mutation. *Genetics*, 70(3): 475-490.

Laubscher, N. 2012. *The role of emotional intelligence and a functional polymorphism in the* MAO-A *gene on aggression in humans*. Unpublished

M.Sc. thesis, Department of Genetics, UFS, South Africa.

Lazarou, J., Pomeranz, B.H. and Corey, P.N. 1998. Incidence of adverse drug reactions in hospitalized patients: A meta-analysis of prospective studies. *The Journal of the American Medical Association*, 279(15): 1200-1205.

Lazary, J., Juhasz, G., Hunyady, L. and Bagdy, G. 2011. Personalized medicine can pave the way for the safe use of CB_1 receptor antagonist. *Trends in Pharmacological Sciences*, 32(5): 270-280.

Leal, S.M. 2005. Detection of genotyping errors and pseudo-SNPs via deviations from Hardy-Weinberg equilibrium. *Genetic Epidemiology*, 29(3): 204-214.

Lee, B.C., Yang, J.W., Lee, S.H., Kim, S.H., Joe, S.H., Jung, I.K., Choi, I.G. and Ham, B.J. 2008. An interaction between the *Norepinephrine Transporter* and *Monoamine Oxidase A* polymorphisms, and novelty-seeking personality traits in Korean females. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 32: 238-242.

Lesch, K.P., Bengel, D., Heils, A., Sabol, S.Z., Greenberg, B.D., Petri, S., Benjamin, J., Müller, C.R., Hamer, D.H. and Murphy, D.L. 1996. Association of anxiety-related traits with a polymorphism in the *Serotonin Transporter* gene regulatory region. *Science*, 274(5292): 1527-1531.

Lesch, K.P. and Gutknecht, L. 2005. Pharmacogenetics of the Serotonin Transporter. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 29(6): 1062-1073.

Lewontin, R.C. and Krakauer, J. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 74(1): 175-195.

Li, D. and He, L. 2008. Meta-study on association between the *Monoamine Oxidase A* gene (*MAOA*) and schizophrenia. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 147B(2): 174-178.

Li, J., Wang, Y., Zhou, R., Zhang, H., Yang, L., Wang, B. and Faraone, S.V. 2007. Association between polymorphisms in *Serotonin Transporter* gene and attention deficit hyperactivity disorder in Chinese Han subjects. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 144B(1): 14-

19.

Lichter, J.B., Barr, C.L., Kennedy, J.L., Van Tol, H.H.M., Kidd, K.K. and Livak, K.J. 1993. A hypervariable segment in the human *Dopamine Receptor D_4* (*DRD4*) gene. *Human Molecular Genetics*, 2(6): 767-773.

Lim, L.C., Powell, J.F., Murray, R. and Gill, M. 1994. *Monoamine Oxidase A* gene and bipolar affective disorder. *The American Journal of Human Genetics*, 54(6): 1122-1124.

Liu, I.S.C., Seeman, P., Sanyal, S., Ulpian, C., Rodgers-Johnson, P.E.B., Serjeant, G.R. and Van Tol, H.H.M. 1996. Dopamine D4 Receptor variant in Africans, $D4_{Valine194Glycine}$, is insensitive to dopamine and clozapine: Report of a homozygous individual. *American Journal of Medical Genetics*, 61: 277-282.

Liu, J., Li, S.Y., Yin, J.Y., Zhang, W., Gao, B., Guo, L. and Qi, R. 2011. Allele frequencies of 6 autosomal STR loci in the Xibo nationality with phylogenetic structure among Chinese populations. *Gene*, 487(1): 84-87.

Lohmueller, K.E., Indap, A.R., Schmidt, S., Boyko, A.R., Hernandez, R.D., Hubisz, M.J., Sninsky, J.J., White, T.J., Sunyaev, S.R., Nielsen, R., Clark, A.G. and Bustamante, C.D. 2008. Proportionally more deleterious genetic variation in European than in African populations. *Nature*, 451(7181): 994-997.

Lu, R.B., Lin, W.W., Lee, J.F., Ko, H.C. and Shih, J.C. 2003. Neither antisocial personality disorder nor antisocial alcoholism is associated with the *MAO-A* gene in Han Chinese males. *Alcoholism: Clinical and Experimental Research*, 27(6): 889-893.

Luca, F., Hudson, R.R., Witonsky, D.B. and Di Rienzo, A. 2011. A reduced representation approach to population genetic analyses and applications to human evolution. *Genome Research*, 21(7): 1087-1098.

Lucki, I. 1998. The spectrum of behaviors influenced by serotonin. *Biological Psychiatry*, 44(3): 151-162.

Lung, F.W., Tzeng, D.S., Huang, M.F. and Lee, M.B. 2011. Association of the *MAOA* promoter uVNTR polymorphism with suicide attempts in patients with major depressive disorder. *BioMed Central Medical Genetics*, 12(74): 1-11.

Lung, F.W., Tzeng, D.S and Shu, B.C. 2002. Ethnic heterogeneity in allele

variation in the *DRD4* gene in schizophrenia. *Schizophrenia Research*, 57(2-3): 239-245.

Maggs, T. 1991. The early history of the black people in southern Africa. In Cameron, T. and Spies, S.B. (eds.) *A New Illustrated History of South Africa*. 2nd edition. Johannesburg: Southern Book Publishers.

Malyarchuk, B., Derenko, M., Denisova, G., Maksimov, A., Wozniak, M., Grzybowski, T., Dambueva, I. and Zakharov, I. 2011. Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup *Q1a*. *Journal of Human Genetics*, 56(8): 583-588.

Mansoor, A., Mazhar, K. and Qamar, R. 2008. VNTR polymorphism of the *DRD4* locus in different Pakistani ethnic groups. *Genetic Testing*, 12(2): 299-304.

Maron, E., Tasa, G., Tõru, I., Lang, A., Vasar, V. and Shlik, J. 2004. Association between serotonin-related genetic polymorphisms and CCK-4-induced panic attacks with or without 5-hydroxytryptophan pretreatment in healthy volunteers. *The World Journal of Biological Psychiatry*, 5(3): 149-154.

Matthews, L.J. and Butler, P.M. 2011. Novelty-seeking *DRD4* polymorphisms are associated with human migration distance out-of-Africa after controlling for neutral population gene structure. *American Journal of Physical Anthropology*, 145(3): 382-389.

May, P.A., Brooke, L., Gossage, J.P., Croxford, J., Adnams, C., Jones, K.L., Robinson, L. and Viljoen, D. 2000. Epidemiology of fetal alcohol syndrome in a South African community in the Western Cape Province. *American Journal of Public Health*, 90(12): 1905-1912.

May, P.A., Gossage, J.P., Marais, A.S., Adnams, C.M., Hoyme, H.E., Jones, K.L., Robinson, L.K., Khaole, N.C.O., Snell, C., Kalberg, W.O. and Hendricks, L. 2007. The epidemiology of fetal alcohol syndrome and partial FAS in a South African community. *Drug and Alcohol Dependence*. 88(2-3): 259-271.

McCracken, J.T., Smalley, S.L., McGough, J.J., Crawford, L., Del'Homme, M., Cantor, R.M., Liu, A. and Nelson, S.F. 2000. Evidence for linkage of a tandem duplication polymorphism upstream of the *Dopamine D4 Receptor* gene (*DRD4*) with attention deficit hyperactivity disorder (ADHD). *Molecular*

Psychiatry, 5(5): 531-536.

McDermott, R., Tingley, D., Cowden, J., Frazzetto, G. and Johnson, D.D.P. 2009. *Monoamine Oxidase A* gene (*MAOA*) predicts behavioural aggression following provocation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(7): 2118-2123.

Mehta, U., Durrheim, D.N., Blockman, M., Kredo, T., Gounden, R. and Barnes, K.I. 2007. Adverse drug reactions in adult medical inpatients in a South African hospital serving a community with a high HIV/AIDS prevalence: Prospective observational study. *British Journal of Clinical Pharmacology*, 65(3): 396-406.

Mesthrie, R. 1995. *Languages and social history: Studies in South African sociolinguistics.* Claremont: David Philip Publishers Ltd.

Michaelovsky, E., Frisch, A., Rockah, R., Peleg, L., Magal, N., Shohat, M. and Weizman. R. 1999. A novel allele in the promoter region of the human *Serotonin Transporter* gene. *Molecular Psychiatry*, 4(1): 97-99.

Middeldorp, C.M., de Geus, E.J.C., Beem, A.L., Lakenberg, N., Hottenga, J.J., Slagboom, P.E. and Boomsma, D.I. 2007. Family based association analyses between the *Serotonin Transporter* gene polymorphism (5-HTTLPR) and neuroticism, anxiety and depression. *Behavioral Genetics*, 37(2): 294-301.

Millet, B., Chabane, N., Delorme, R., Leboyer, M., Leroy, S., Poirier, M.F., Bourdel, M.C., Mouren-Simeoni, M.C., Rouillon, F., Loo, H. and Krebs, M.O. 2003. Association between the *Dopamine Receptor D4* (*DRD4*) gene and obsessive-compulsive disorder. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 116B(1): 55-59.

Missale, C., Nash, S.R., Robinson, S.W., Jaber, M. and Caron, M.G. 1998. Dopamine receptors: From structure to function. *Physiological Reviews*, 78(1): 189-225.

Mitsuyasu, H., Kawasaki, H., Ninomiya, H., Kinukawa, N., Yamanaka, T., Tahira, T., Stanton Jr., V.P., Springett, G.M., Hayashi, K., Tashiro, N. and Kanba, S. 2007. Genetic structure of the *Dopamine Receptor D4* gene (*DRD4*) and lack of association with schizophrenia in Japanese patients. *Journal of Psychiatric Research*, 41(9): 763-775.

Mongeau, R., Blier, P. and de Montigny, C. 1997. The serotonergic and noradrenergic systems of the hippocampus: Their interactions and the effects of antidepressant treatments. *Brain Research Reviews*, 23(3): 145-195.

Mountain, A. 2003. *The first people of the Cape: A look at their history and the impact of colonization on the Cape's indigenous people*. Claremont: David Phillip Publishers Ltd.

Müller, D.J., Serretti, A., Sicard, T., Tharmalingam, S., King, N., Artioli, P., Mandelli, L., Lorenzi, C. and Kennedy, J.L. 2007. Further evidence of *MAO-A* gene variants associated with bipolar disorder. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 144B(1): 37-40.

Muramatsu, T., Higuchi, S., Murayama, M., Matsushita, S. and Hayashida, M. 1996. Association between alcoholism and the *Dopamine D4 Receptor* gene. *Journal of Medical Genetics*, 33(2): 113-115.

Murphy, D.L., Lerner, A., Rudnick, G. and Lesch, K.P. 2004. Serotonin transporter: Gene, genetic disorders, and pharmacogenetics. *Molecular Interventions*, 4(2): 109-123.

Naka, I., Ohashi, J., Kimura, R., Furusawa, T., Yamauchi, T., Nakazawa, M., Natsuhara, K., Ataka, Y., Nishida, N., Ishida, T., Inaoka, T., Matsumura, Y. and Ohtsuka, R. 2012. *DRD4* VNTR polymorphism in Oceanic populations. *Anthropological Science*, 120(2): 151-155.

Nakamura, M., Ueno, S., Sano, A. and Tanabe, H. 2000. The human *Serotonin Tansporter* gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants. *Molecular Psychiatry*, 5(1): 32-38.

Narita, N., Narita, M., Takashima, S., Nakayama, M., Nagai, T. and Okado, N. 2001. *Serotonin Transporter* gene variation is a risk factor for sudden infant death syndrome in the Japanese population. *Pediatrics*, 107(4): 690-692.

Nederhof, E., Creemers, H.E., Huizink, A.C., Ormel, J. and Oldehinkel, A.J. 2011. L-*DRD4* genotype not associated with sensation seeking, gambling performance and startle reactivity in adolescents: The Trials study. *Neuropsychologia*, 49(5): 1359-1362.

Neves, F.S., Silveira, G., Romano-Silva, M.A., Malloy-Diniz, L., Ferreira, A.A., De

Marco, L. and Correa, H. 2008. Is the 5-HTTLPR polymorphism associated with bipolar disorder or with suicidal behavior of bipolar disorder patients? *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics*), 147B(1): 114-116.

Newman, T.K., Syagailo, Y.V., Barr, C.S., Wendland, J.R., Champoux, M., Graessle, M., Suomi, S.J., Higley, J.D. and Lesch, K.P. 2005. *Monoamine Oxidase A* gene promoter variation and rearing experience influences aggressive behavior in Rhesus monkeys. *Biological Psychiatry*, 57(2): 167-172.

Nonnis Marzano, F., Maldini, M., Filonzi, L., Lavezzi, A.M., Parmigiani, S., Magnani, C., Bevilacqua, G. and Matturri, L. 2008. Genes regulating the serotonin metabolic pathway in the brain stem and their role in the etiopathogenesis of the sudden infant death syndrome. *Genomics*, 91(6): 485-491.

Nordquist, N. and Oreland, L. 2006. Monoallelic expression of *MAOA* in skin fibroblasts. *Biochemical and Biophysical Research Communications*, 348(2): 763-767.

Norman, R., Bradshaw, D., Schneider, M., Pieterse, D. and Groenewald, P. 2006. *Revised burden of disease estimates for the comparative risk factor assessment, South Africa 2000. Methodological Note*. Cape Town: South African Medical Research Council, 2006.

Norton, N., Kirov, G., Zammit, S., Jones, G., Jones, S., Owen, R., Krawczak, M., Williams, N.M., O'Donovan, M.C. and Owen, M.J. 2002. Schizophrenia and functional polymorphisms in the *MAOA* and *COMT* genes: No evidence for association or epistasis. *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, 114(5): 491-496.

Noskova, T., Pivac, N., Nedic, G., Kazantseva, A., Gaysina, D., Faskhutdinova, G., Gareeva, A., Khalilova, Z., Khusnutdinova, E., Kovacic, D.K., Kovacic, Z., Jokic, M. and Seler, D.M. 2008. Ethnic differences in the serotonin transporter polymorphism (5-HTTLPR) in several European populations. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 32(7): 1735-1739.

Nöthen, M.M., Cichon, S., Hemmer, S., Hebebrand, J., Remschmidt, H., Lehmkuhl, G., Poustka, F., Schmidt, M., Catalano, M., Fimmers, R., Körner, J., Rietschel, M. and Propping, P. 1994. Human *Dopamine D4 Receptor* gene: Frequent occurrence of a null allele and observation of homozygosity. *Human Molecular Genetics*, 3(12): 2207-2212.

Novembre, J. and Stephens, M. 2008. Interpreting principal component analyses of spatial population genetic variation. *Nature Genetics*, 40(5): 646-649.

Nurse, G. T., Weiner, J. S. and Jenkins, T. 1985. *The peoples of southern Africa and their affinities*. Oxford: Oxford University Press.

Ogilvie, A.D., Battersby, S., Bubb, V.J., Fink, G., Harmar, A.J., Goodwin, G.M. and Smith, C.A.D. 1996. Polymorphism in *Serotonin Transporter* gene associated with susceptibility to major depression. *The Lancet*, 347(9003): 731-733.

Okuyama, Y., Ishiguro, H., Nankai, M., Shibuya, H., Watanabe, A. and Arinami, T. 2000. Identification of a polymorphism in the promoter region of *DRD4* associated with the human novelty seeking personality trait. *Molecular Psychiatry*, 5(1): 64-69.

Oldenhof, J., Vickery, R., Anafi, M., Oak, J., Ray, A., Schoots, O., Pawson, T., von Zastrow, M. and Van Tol, H.H.M. 1998. SH3 binding domains in the Dopamine D4 Receptor. *Biochemistry*, 37(45): 15726-15736.

Oliveira, J.R.M., Carvalho, D.R., Pontual, D., Gallindo, R.M., Sougey, E.B., Gentil, V., Lafer, B., Maia, L.G.S., Morias Jr., M.A., Matioli, S., Vallada, H., Moreno, R.A., Nishimura, A., Otto, P.A., Passos-Bueno, M.R. and Zatz, M. 2000. Analysis of the serotonin transporter polymorphism (5-HTTLPR) in Brazilian patients affected by dysthymia, major depression and bipolar disorder. *Molecular Psychiatry*, 5(4): 348-349.

Ozaki, N., Goldman, D., Kaye, W.H., Plotnicov, K., Greenberg, B.D., Lappalainen, J., Rudnick, G. and Murphy, D.L. 2003. Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype. *Molecular Psychiatry*, 8(11): 933-939.

Pardini, M., Krueger, F., Hodgkinson, C., Raymont, V., Ferrier, C., Goldman, D., Strenziok, M., Guida, S. and Grafman, J. 2011. Prefrontal cortex lesions and

MAO-A modulate aggression in penetrating traumatic brain injury. *Neurology*, 76(12): 1038-1045.

Parra, E.J., Marcini, A., Akey, J., Martinson, J., Batzer, M.A., Cooper, R., Forrester, T., Allison, D.B., Deka, R., Ferrell, R.E. and Shriver, M.D. 1998. Estimating African American admixture proportions by use of population-specific alleles. *The American Journal of Human Genetics*, 63(6): 1839-1851.

Patel, V. and Kim, Y.R. 2007. Contribution of low- and middle-income countries to research published in leading general psychiatric journals, 2002-2004. *British Journal of Psychiatry*, 190: 77-78.

Peires, J.B. 1981. *The house of Phalo: A history of the Xhosa people in the days of their independence*. Berkeley: University of California Press.

Pereira, V., Gusmão, L., Valente, C., Pereira, R., Carneiro, J., Gomes, I., Morling, N., Amorim, A. and João Prata, M. 2012. Refining the genetic portrait of Portuguese Roma through X-chromosomal markers. *American Journal of Physical Anthropology*, 148(3): 389-394.

Philibert, R.A., Gunter, T.D., Beach, S.R., Brody, G. and Madan, A. 2008. *MAOA* methylation is associated with nicotine and alcohol dependence in women. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 147B(5): 565-570.

Pickar, D. and Rubinow, K. 2001. Pharmacogenomics of psychiatric disorders. *Trends in Pharmacological Sciences,* 22(2): 75-83.

Pinsonneault, J.K., Papp, A.C. and Sadée, W. 2006. Allelic mRNA expression of X-linked *Monoamine Oxidase A* (*MAOA*) in human brain: dissection of epigenetic and genetic factors. *Human Molecular Genetics*, 15(17): 2636-2649.

Pritchard, J.K. and Rosenberg, N.A. 1999. Use of unlinked genetic markers to detect population stratification in association studies. *The American Journal of Human Genetics*, 65(1): 220-228.

Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945-959.

Qiu, H.T., Meng, H.Q., Song, C., Xiu, M.H., Chen, D.C., Zhu, F.Y., Wu, G.Y.,

Kosten, T.A., Kosten, T.R. and Zhang, X.Y. 2009. Association between *Monoamine Oxidase* (*MAO*)-*A* gene variants and schizophrenia in a Chinese population. *Brain Research*, 1287: 67-73.

Quinque, D., Kittler, R., Kayser, M., Stoneking, M. and Nasidze, I. 2006. Evaluation of saliva as a source of human DNA for population and association studies. *Analytical Biochemistry*, 353(2): 272-277.

Quintana-Murci, L., Harmant, C., Quach, H., Balanovsky, O., Zaporozhchenko, V., Bormans, C., van Helden, P.D., Hoal, E.G. and Behar, D.M. 2010. Strong maternal Khoisan contribution to the South African Coloured population: A case of gender-biased admixture. *The American Journal of Human Genetics*, 86(4): 611-620.

Quintana-Murci, L., Semino, O., Bandelt, H.J., Passarino, G., McElreavey, K. and Santachiara-Benerecetti, A.S. 1999. Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. *Nature Genetics*, 23(4): 437-441.

Ramamoorthy, S., Bauman, A.L., Moore, K.R., Han, H., Yang-Feng, T., Chang, A.S., Ganapathy, V. and Blakely, R.D. 1993. Antidepressant- and cocaine-sensitive human serotonin transporter: Molecular cloning, expression, and chromosomal localization. *Proceedings of the National Academy of Sciences of the United States of America*, 90(6): 2542-2546.

Raven-Hart, R. 1967. *Before van Riebeeck: Callers at South Africa from 1488 to 1652*. Capetown: C. Struik Ltd.

Ray, L.A., Bryan, A., MacKillop, J., McGeary, J., Hesterberg, K. and Hutchison, K.E. 2008. The *Dopamine D_4 Receptor* (*DRD4*) gene exon III polymorphism, problematic alcohol use and novelty seeking: Direct and mediated genetic effects. *Addiction Biology*, 14(2): 238-244.

Reed, F.A. and Tishkoff, S.A. 2006. African human diversity, origins and migrations. *Current Opinion in Genetics and Development*, 16(6): 597-605.

Regueiro, M., Alvarez, J., Rowold, D. and Herrera, R.J. 2013. On the origins, rapid expansion and genetic diversity of Native Americans from hunting-gatherers to agriculturalists. *American Journal of Physical Anthropology*,

150(3): 333-348.

Reist, C., Ozdemir, V., Wang, E., Hashemzadeh, M., Mee, S. and Moyzis, R. 2007. Novelty Seeking and the *Dopamine D4 Receptor* gene (*DRD4*) revisited in Asians: Haplotype characterization and relevance of the 2-repeat allele. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 144B(4): 453-457.

Relethford, J.H. 2008. Genetic evidence and the modern human origins debate. *Heredity*, 100(6): 555-563.

Ren, R., Mayer, B.J., Cicchetti, P. and Baltimore, D. 1993. Identification of a ten-amino acid proline-rich SH3 binding site. *Science*, 259(5098): 1157-1161.

Risch, N. and Merikangas, K. 1996. The future of genetic studies of complex human diseases. *Science*, 273(5281): 1516-1517.

Rockman, M.V and Wray, G.A. 2002. Abundant raw material for *CIS*-regulatory evolution in humans. *Molecular Biology and Evolution*, 19(11): 1991-2004.

Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A. and Feldman, M.W. 2002. Genetic structure of human populations. *Science*, 298(5602): 2381-2385.

Sabeti, P.C., Schaffner, S.F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., Palma, A., Mikkelsen, T.S., Altshuler, D. and Lander, E.S. 2006. Positive natural selection in the human lineage. *Science*, 312(5780): 1614-1620.

Sabol, S.Z., Hu, S. and Hamer, D. 1998. A functional polymorphism in the *Monoamine Oxidase A* gene promoter. *Human Genetics*, 103(3): 273-279.

Salanti, G., Sanderson, S. and Higgins, J.P.T. 2005. Obstacles and opportunities in meta-analysis of genetic association studies. *Genetics in Medicine*, 7(1): 13-20.

Samochowiec, J., Hajduk, A., Samochowiec, A., Horodnicki, J., Stępień, G., Grzywacz, A. and Kucharska-Mazur, J. 2004. Association studies of *MAO-A*, *COMT*, and *5-HTT* genes polymorphisms in patients with anxiety disorders of the phobic spectrum. *Psychiatric Research*, 128(1): 21-26.

Samochowiec, J., Lesch, K.P., Rottmann, M., Smolka, M., Syagailo, Y.V.,

Okladnova, O., Rommelspacher, H., Winterer, G., Schmidt, L.G. and Sander, T. 1999. Association of a regulatory polymorphism in the promoter region of the *Monoamine Oxidase A* gene with antisocial alcoholism. *Psychiatry Research*, 86(1): 67-72.

Sander, T., Harms, H., Dufeu, P., Kuhn, S., Rommelspacher, H. and Schmidt, L.G. 1997. *Dopamine D4 Receptor* exon III alleles and variation of novelty seeking in alcoholics. *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, 74(5): 483-487.

Schaid, D.J. and Jacobsen, S.J. 1999. Biased tests of association: Comparisons of allele frequencies when departing from Hardy-Weinberg proportions. *American Journal of Epidemiology*, 149(8): 706-711.

Schapera, I. 1930. *The Khoisan peoples of South Africa: Bushmen and Hottentots*. London: G. Routledge and Sons Ltd.

Schlebusch, C.M. 2010. *Genetic variation in Khoisan-speaking populations from southern Africa*. Ph.d. thesis, University of the Witwatersrand, South Africa.

Schlebusch, C.M., Skoglund, P., Sjödin, P., Gattepaille, L.M., Hernandez, D., Jay, F., Li, S., De Jongh, M., Singleton, A., Blum, M.G.B., Soodyall, H. and Jakobsson, M. 2012. Genomic variations in seven Khoe-San groups reveals adaptation and complex African history. *Science*, 338(374): 374-379.

Schlebusch, C.M. and Soodyall, H. 2012. Extensive population structure in San, Khoe and mixed ancestry populations from southern Africa revealed by 44 short 5-SNP haplotypes. *Human Biology*, 84(6): 695-724.

Schmidt, L.A., Fox, N.A. and Hamer, D.H. 2007. Evidence for a gene-gene interaction in predicting children's behaviour problems: Association of serotonin transporter short and dopamine receptor D4 long genotypes with internalising and externalising behaviours in typically developing 7-year-olds. *Development and Psychopathology*, 19: 1105-1116.

Schmidt, L.G., Sander, T., Kuhn, S., Smolka, M., Rommelspacher, H., Samochowiec, J. and Lesch, K.P. 2000. Different allele distribution of a regulatory *MAOA* gene promoter polymorphism in antisocial and anxious-depressive alcoholics. *Journal of Neural Transmission*, 107(6): 681-689.

Schnaitman, C., Erwin, V.G. and Greenawalt, J.W. 1967. The submitochondrial localization of Monoamine Oxidase. An enzymatic marker for the outer membrane of rat liver mitochondria. *The Journal of Cell Biology*, 32(3): 719-735.

Schrire, C. 1980. An inquiry into the evolutionary status and apparent identity of San hunter-gatherers. *Human Ecology*, 8(1): 9-32

Schulze, T.G., Müller, D.J., Krauss, H., Scherk, H., Ohlraun, S., Syagailo, Y.V., Windemuth, C., Neidt, H., Grässle, M., Papassotiropoulus, A., Heun, R., Nöthen, M.M., Maier, W., Lesch, K.P. and Rietschel, M. 2000. Association between a functional polymorphism in the *Monoamine Oxidase A* gene promoter and major depressive disorder. *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, 96(6): 801-803.

Seaman, M.I., Fischer, J.B., Chang, F.M. and Kidd, K.K. 1999. Tandem duplication polymorphism upstream of the *Dopamine D4 Receptor* gene (*DRD4*). *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, 88: 705-709.

Seedat, S., Stein, D.J., Herman, A., Kessler, R., Sonnega, J., Heeringa, S., Williams, S. and Williams, D. 2008. Twelve-month treatment of psychiatric disorders in the South African stress and health survey (World Mental Health Survey Initiative). *Social Psychiatry Psychiatric Epidemiology*, 43(11): 889-897.

Seeman, P., Ulpian, C., Chouinard, G., Van Tol, H.H.M., Dwosh, H., Lieberman, J.A., Siminovitch, K., Liu, I.S.C., Waye, J., Voruganti, P., Hudson, C., Serjeant, G.R., Masibay, A.S. and Seeman, M.V. 1994. Dopamine D4 Receptor variant, D4Glycine194, in Africans, but not in Caucasians: No association with schizophrenia. *American Journal of Medical Genetics*, 54(4): 384-390.

Seldin, M.F., Pasaniuc, B. and Price, A.L. 2011. New approaches to disease mapping in admixed populations. *Nature Reviews Genetics*, 12(8): 523-528.

Semino, O., Santachiara-Benerecetti, A.S., Falaschi, F., Cavalli-Sforza, L.L. and Underhill, P.A. 2002. Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. *American Journal of Human Genetics*, 70(1): 265-268.

Sen, S., Burmeister, M. and Ghosh, D. 2004. Meta-analysis of the association between a serotonin transporter promoter polymorphism (5-HTTLPR) and anxiety-related personality traits. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 127B(1): 85-89.

Serretti, A., Olgiati, P., Bajo, E., Bigelli, M. and De Ronchi, D. 2011. A model to incorporate genetic testing (5-HTTLPR) in pharmacological treatment of major depressive disorders. *The World Journal of Biological Psychiatry*, 12(7): 501-515.

Sharp, J. and Douglas, S. 1996. *Prisoners of their reputation? The veterans of the 'Bushmen' battalions in South Africa*. *Miscast. Negotiating the presence of the Bushmen.* Skotnes, P. Cape Town: University of Cape Town Press.

Shell, R.C.H. 1994. *Children of bondage: A social history of the slave society at the Cape of Good Hope, 1652-1838*. Hanover and London: Wesleyan University Press.

Shih, J.C. and Thompson, R.F. 1999. Monoamine Oxidase in neuropsychiatry and behavior. *The American Journal of Human Genetics*, 65(3): 593-598.

Sirugo, G., Hennig, B.J., Adeyemo, A.A., Matimba, A., Newport, M.J., Ibrahim, M.E., Ryckman, K.K., Tacconelli, A., Mariani-Costantini, R., Novelli, G., Soodyall, H., Rotimi, C.N., Ramesar, R.S., Tishkoff, S.A. and Williams, S.M. 2008. Genetic studies of African populations: An overview on disease susceptibility and response to vaccines and therapeutics. *Human Genetics*, 123(6): 557-598.

Sjöberg, R.L., Ducci, F., Barr, C.S., Newman, T.K., Dell'Osso, L., Virkkunen, M and Goldman, D. 2008. A non-additive interaction of a functional *MAO-A* VNTR and testosterone predicts antisocial behavior. *Neuropsychopharmacology*, 33(2): 425-430.

Slater, R. 2002. Differentiation and diversification: Changing livelihoods in Qwaqwa, South Africa, 1970-2000. *Journal of Southern African Studies*, 28(3): 599-614.

Slotkin, T.A. 1999. Mary Bernheim and the discovery of Monoamine Oxidase. *Brain Research Bulletin*, 50(5-6): 373.

Smith, M.W. and O'Brien, S.J. 2005. Mapping by admixture linkage disequilibrium: Advances, limitations and guidelines. *Nature Reviews Genetics*, 6(8): 623-632.

Stabellini, R., de Mello, J.C.M., Hernandes, L.M. and Pereira, LV. 2009. *MAOA* and *GYG2* are submitted to X chromosome inactivation in human fibroblasts. *Epigenetics*, 4(6): 388-393.

Stine, O.C. and Smith, K.D. 1990. The estimation of selection coefficients in Afrikaners: Huntington disease, Porphyria Variegata, and Lipoid Proteinosis. *The American Journal of Human Genetics*, 46(3): 452-458.

Strobel, A., Lesch, K.P., Jatzke, S., Paetzold, F. and Brocke, B. 2003. Further evidence for a modulation of novelty seeking by *DRD4* exon III, 5-HTTLPR, and *COMT* val/met variants. *Molecular Psychiatry*, 8(4): 371-372.

Sun, Y., Zhang, J., Yuan, Y., Yu, X., Shen, Y. and Xu, Q. 2012. Study of a possible role of the *Monoamine Oxidase A* (*MAOA*) gene in paranoid schizophrenia among a Chinese population. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 159B(1): 104-111.

Swanson, J.M., Sunohara, G.A., Kennedy, J.L., Regino, R., Fineberg, E., Wigal, T., Lerner, M., Williams, L., LaHoste, G.J. and Wigal, S. 1998. Association of the *Dopamine Receptor D4* (*DRD4*) gene with a refined phenotype of attention deficit hyperactivity disorder (ADHD): A family-based approach. *Molecular Psychiatry*, 3(1): 38-41.

Szekely, A., Ronai, Z., Nemoda, Z., Kolmann, G., Gervai, J. and Sasvari-Szekely, M. 2004. Human personality dimensions of persistence and harm avoidance associated with *DRD4* and 5-HTTLPR polymorphisms. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 126B: 106-110.

Takezaki, N., Nei, M. and Tamura, K. 2010. POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with windows interface. *Molecular Biology and Evolution*, 27(4): 747-752.

Tang, H., Quertermous, T., Rodriguez, B., Kardia, S.L.R., Zhu, X., Brown, A., Pankow, J.S., Province, M.A., Hunt, S.C., Boerwinkle, E., Schork, N.J. and

Risch, N.J. 2005. Genetic structure, self-identified race/ethnicity, and confounding in case-control association studies. *American Journal of Human Genetics* 76(2): 268-275.

Tattersall, A. 2009. Human origins: Out of Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 106(38): 16018-16021.

Taylor, A.L. and Wright, J.T. 2005. Should ethnicity serve as the basis for clinical trial design? Importance of race/ethnicity in clinical trials: Lessons from the African-American heart failure trial (A-HeFT), the African-American study of kidney disease and hypertension (AASK), and the antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT). *Circulation*, 112(23): 3654-3666.

Theal, G.M. 1910. *The yellow and dark-skinned people of Africa south of the Zambesi*. London: S.Sonnenschein & Co. Ltd.

Thompson, E.E., Kuttab-Boulos, H., Witonsky, D., Yang, L., Roe, B.A. and Di Rienzo, A. 2004. *CYP3A* variation and the evolution of salt-sensitivity variants. *American Journal of Human Genetics*, 75(6): 1059-1069.

Tidey, J.W., Monti, P.M., Rohsenow, D.J., Gwaltney, C.J., Miranda Jr, R., McGeary, J.E., MacKillop, J., Swift, R.M., Abrams, D.B., Shiffman, S. and Paty, J.A. 2008. Moderators of Naltrexone's effects on drinking, urge and alcohol effects in non-treatment-seeking heavy drinkers in the natural environment. *Alcoholism, Clinical and Experimental Research*, 32(1): 58-66.

Tipping, A.J., Pearson, T., Morgan, N.V., Gibson, R.A., Kuyt, L.P., Havenga, C., Gluckman, E., Joenje, H., de Ravel, T., Jansen, S. and Matthew, C.G. 2001. Molecular and genealogical evidence for a founder effect in Fanconi anemia families of the Afrikaner population of South Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10): 5734-5739.

Tishkoff, S.A., Gonder, M.K., Henn, B. M., Mortensen, H., Knight, A., Gignoux, C., Fernandopulle, N., Lema, G., Nyambo, T.B., Ramakrishnan, U., Reed, F.A. and Mountain, J. L. 2007b. History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. *Molecular Biology and Evolution*, 24(10): 2180-2195.

Tishkoff, S.A., Reed, F.A., Friedlaender, F.R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J.B., Awomoyi, A.A., Bodo, J.M., Doumbo, O., Ibrahim, M., Juma, A.T., Kotze, M.J., Lema, G., Moore, J.H., Mortensen, H., Nyambo, T.B., Omar, S.A., Powell, K., Pretorius, G.S., Smith, M.W., Thera, M.A., Wambebe, C., Weber, J.L. and Williams, S.M. 2009. The genetic structure and history of Africans and African Americans. *Science*, 324(5930): 1035-1044.

Tishkoff, S.A., Reed, F.A., Ranciaro, A., Voight, B.F., Babbitt, C.C., Silverman, J.S., Powell, K., Mortensen, H.M., Hirbo, J.B., Osman, M., Ibrahim, M., Omar, S.A., Lema, G., Nyambo, T.B., Ghori, J., Bumpstead, S., Pritchard, J.K., Wray, G.A. and Deloukas, P. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genetics*, 39(1): 31-40.

Tishkoff, S.A. and Williams, S.C. 2002. Genetic analysis of African populations: Human evolution and complex disease. *Nature Reviews Genetics*, 3: 611-621.

Tomitaka, M., Tomitaka, S., Otuka, Y., Kim, K., Matuki, H., Sakamoto, K. and Tanaka, A. 1999. Association between novelty seeking and *Dopamine Receptor D4* (*DRD4*) exon III polymorphism in Japanese subjects. *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, 88(5): 469-471.

Trikalinos, T.A., Salanti, G., Khoury, M.J. and Ioannidis, J.P.A. 2006. Impact of violations and deviations in Hardy-Weinberg equilibrium on postulated gene-disease associations. *American Journal of Epidemiology*, 163(4): 300-309.

Turecki, G., Grof, P., Cavazzoni, P., Duffy, A., Grof, E., Ahrens, B., Berghöfer, A., Müller-Oerlinghausen, B., Dvoráková, M., Libigerová, E., Vojtechovsky, M., Zvolsky, P., Joober, R., Nilsson, A., Prochazka, H., Licht, R.W., Rasmussen, N.A., Schou, M., Vestergaard, P., Holzinger, A., Schumann, C., Thau, K., Rouleau, G.A. and Alda, M. 1999. *MAOA*: Association and linkage studies with lithium responsive bipolar disorder. *Psychiatric Genetics*, 9(1): 13-16.

Uchigasaki, S., Tie, J. and Takahashi, D. 2013. Genetic analysis of twelve X-chromosome STRs in Japanese and Chinese populations. *Molecular Biology Reports*, 40(4): 3193-3196.

Urata, T., Takahashi, N., Hakamata, Y., Iijima, Y., Kuwahara, N., Ozaki, N., Ono, Y., Amano, M. and Inada, T. 2007. Gene-gene interaction analysis of

personality traits in a Japanese population using an electrochemical DNA array chip analysis. *Neuroscience Letters*, 414(3): 209-212.

Van den Berghe, P.L. 1965. *South Africa: A study in conflict*. Middletown, Connecticut: Wesleyan University Press.

Van Tol, H.H.M., Bunzow, J.R., Guan, H.C., Sunahara, R.K., Seeman, P., Niznik, H.B. and Civelli, O. 1991. Cloning of the gene for a human Dopamine D4 Receptor with high affinity for the antipsychotic clozapine. *Nature*, 350(6319): 610-614.

Van Tol, H.H.M., Wu, C.M., Guan, H.C., Ohara, K., Bunzow, J.R., Civelli, O., Kennedy, J., Seeman, P., Niznik, H.B. and Jovanovic, V. 1992. Multiple *Dopamine D4 Receptor* variants in the human population. *Nature*, 358(6382): 149-152.

Wang, C., Szpiech, Z.A., Degnan, J.H., Jakobsson, M., Pemberton, T.J., Hardy, J.A., Singleton, A.B. and Rosenberg, N.A. 2010. Comparing spatial maps of human population-genetic variation using procrustes analysis. *Statistical Applications in Genetics and Molecular Biology*, 9(1): 1-20.

Wang, E., Ding, Y.C., Flodman, P., Kidd, J.R., Kidd, K.K., Grady, D.L., Ryder, O.A., Spence, M.A., Swanson, J.M. and Moyzis, R.K. 2004. The genetic architecture of selection at the human *Dopamine Receptor D4* (*DRD4*) gene locus. *American Journal of Human Genetics*, 74(5): 931-944.

Wang, X., Zhong, S., Gao, L., Dou, S., He, G., Ran, Y., Bao, J. and Jing, Q. 2011. An association study between polymorphism of *Alcohol Dehydrogenase* (*ADH1B*), *Aldehyde Dehydrogenase* (*ALDH2*), *Cytochrome* (*CYP4502E1*), *Catechol-O-Methyltransferase* (*COMT*) and *5-Hydroxytryptamine Transporter* (*5-HTT*) genes in Yunnan Han population with alcohol dependence. *African Journal of Biotechnology*, 10(57): 12164-12170.

Wassink, T.H., Hazlett, H.C., Epping, E.A., Amdt, S., Dager, S.R., Schellenberg, G.D., Dawson, G. and Piven, J. 2007. Cerebral cortical gray matter overgrowth and functional variation of the *Serotonin Transporter* gene in autism. *Archives of General Psychiatry*, 64(6): 709-717.

Watson, E., Forster, P., Richards, M. and Bandelt, H.J. 1997. Mitochondrial

footprints of human expansions in Africa. *The American Journal of Human Genetics*, 61(3): 691-704.

Weder, N., Yang, B.Z., Douglas-Palumberi, H., Massey, J., Krystal, J.H., Gelernter, J. and Kaufman, J. 2009. *MAOA* genotype, maltreatment, and aggressive behavior: The changing impact of genotype at varying levels of trauma. *Biological Psychiatry*, 65(5): 417-424.

Weese-Mayer, D.E., Zhou, L., Berry-Kravis, E.M., Maher, B.S., Silvestri, J.M. and Marazita, M.L. 2003. Association of the *Serotonin Transporter* gene with sudden infant death syndrome: A haplotype analysis. *American Journal of Medical Genetics*, 122A(3): 238-245.

Wigginton, J.E., Cutler, D.J. and Abecasis, G.R. 2005. A note on exact tests of Hardy-Weinberg equilibrium. *American Journal of Human Genetics*, 76(5): 887-893.

Willeit, M., Praschak-Rieder, N., Neumeister, A., Zill, P., Leisch, F., Stastny, J., Hilger, E., Thierry, N., Konstantinidis, A., Winkler, D., Fuchs, K., Sieghart, W., Aschauer, H., Ackenheil, M., Bondy, B. and Kasper, S. 2003. A polymorphism (5-HTTLPR) in the *Serotonin Transporter Promoter* gene is associated with DSM-IV depression subtypes in seasonal affective disorder. *Molecular Psychiatry*, 8: 942-946.

Wimbiscus, M., Kostenko, O. and Malone, D. 2010. MAO inhibitors: Risks, benefits, and lore. *Cleveland Clinical Journal of Medicine*, 77(12): 859-882.

Wise, R.A. and Rompre, P.P. 1989. Brain dopamine and reward. *Annual Review of Psychology,* 40: 191-225.

Woodham-Smith, C. 1962. *The great hunger: Ireland 1845-1849*. New York and Evanston: Harper and Row.

Wright, G.E.B., Niehaus, D.J.H., Koen, L., Drögemöller, B.I and Warnich, L. 2011. Psychiatric genetics in South Africa: Cutting a rough diamond. *African Journal of Psychiatry*, 14(5): 355-366.

Yirmiya, N., Pilowsky, T., Tidhar, S., Nemanov, L., Altmark, L. and Ebstein, R.P. 2002. Family-based and population study of a functional promoter-region *Monoamine Oxidase A* polymorphism in autism: Possible association with IQ.

American Journal of Medical Genetics (Neuropsychiatric Genetics), 114(3): 284-287

Yonan, A.L., Palmer, A.A. and Gilliam, T.C. 2006. Hardy-Weinberg disequilibrium identified genotyping error of the *Serotonin Transporter* (*SLC6A4*) promoter polymorphism. *Psychiatric Genetics*, 16(1): 31-34.

Young, K.A., Bonkale, W.L., Holcomb, L.A., Hicks, P.B. and German, D.C. 2008. Major depression, 5HTTLPR genotype, suicide and antidepressant influences on thalamic volume. *The British Journal of Psychiatry*, 192(4): 285-289.

Zalsman, G., Frisch, A., Lewis, R., Michaelovsky, E., Hermesch, H., Sher, L., Nahshoni, E., Wolovik, L., Tyano, S., Apter, A., Weizman, R. and Weizman, A. 2004. *DRD4* receptor gene exon III polymorphism in inpatient suicidal adolescents. *Journal of Neural Transmission*, 111(12): 1593-1603.

UNIVERSITY OF THE
FREE STATE
UNIVERSITEIT VAN DIE
VRYSTAAT
YUNIVESITHI YA
FREISTATA

UFS·UV
NATURAL AND
AGRICULTURAL SCIENCES
NATUUR- EN
LANDBOUWETENSKAPPE

CHAPTER NINE

APPENDICES

Appendix A1: Cover letter, informed consent and general participant information.

Departement Genetika (116) / Department of Genetics (116)

Fakulteit Natuur- en Landbouwetenskappe

Faculty of Natural and Agricultural Sciences

Posbus / P.O. Box 339

Bloemfontein 9300

South Africa

E-pos / E-mail: Genetics@ufs.ac.za

☎ +27-(0)51-401-2595

🖷 +27-(0)86-518-7317

Information for participant

You have been invited to partake in Clement Malan's Magister Scientiae (M.Sc.) project. The main focus of the project is the analyses of the allele frequency of three neurotransmitter genes in different ethnic populations. The reason for studying genetic variation in different populations is to help determine why some populations are more susceptible to certain diseases than others and therefore help with the treatment of these diseases.

A participant would be required to donate a saliva sample, from which DNA would be extracted. The information required for this particular project would be obtained from the DNA from each participant. Participants are required to belong to the San, Afrikaner, Xhosa or Sotho ethnic groups and must be older than 18 years.

No risks are involved in the collection of saliva as it only requires the participant to rinse his/her mouth with mouthwash and then to deposit the mouthwash into a tube. The saliva samples obtained from each ethnic group will be labeled accordingly for the project. Participant names will not be included in the study therefore all participants will remain anonymous. No results will be given to participants; however, the research data would be accessible to the public after publication. Participation in the project is free of charge.

Similarly, you will also not receive any money or any other compensation if you participate in this study. You are allowed to withdraw from the study at any time without penalization.

The duration of the project would be at least two years. The DNA obtained will be used with your consent for future studies.

Please feel free to contact Prof. Spies or Mrs. Schneider with any questions you may have regarding this study. (051 4012261 or 051 4013730).

Thank you for your cooperation.

Informed Consent

Please read or listen to the information presented below, sign if you agree.

I declare that I understand and agree with the following information:

- What the research project entails.
- I willingly participate in the study without compensation.
- I agree to give a saliva sample.
- The saliva sample will only be used to extract DNA.
- The information obtained from the analysis of the DNA will be included in the research project.
- The DNA and information can be used in future studies.
- All the information collected will be treated confidentially.
- The results will not be given to participants.
- I am not pressurized to participate in the study.
- The participant may withdraw from the study at any time without penalization.
- I have read or listened to and understood the information supplied about the project and know that giving a sample is my choice.

I, the participant, voluntary agree(s) to participate in the above mentioned research project.

Signed at.....on.....20.....

....

Signature or Witness

right thumb print of participant

General participant information

Sample I.D.: .....

Mark the correct box

Sex: Male: ☐ Female: ☐

Ethnic group: Afrikaner: ☐

San: ☐

Southern Sotho: ☐

Xhosa: ☐

Age (years): 18 - 27: ☐

28 - 37: ☐

38 - 47: ☐

48 - 57: ☐

58 - 67: ☐

68 - 77: ☐

77 +: ☐

Home language:

Afrikaans: ☐

English: ☐

IsiNdebele: ☐

IsiXhosa: ☐

IsiZulu: ☐

Sesotho sa Leboa: ☐

Sesotho: ☐

Setswana: ☐

siSwati: ☐

Tshivenda: ☐

Xitsonga: ☐

Other: .....

Indicate province of origin:

Eastern Cape: ☐ Mpumalanga: ☐

Free State: ☐ North West: ☐

Gauteng: ☐ Northern Cape: ☐

KwaZulu - Natal: ☐ Western Cape: ☐

Limpopo: ☐

Indicate where you grew up:

Farm: ☐

Rural area: ☐

Town: ☐

City: ☐

UNIVERSITY OF THE
FREE STATE
UNIVERSITEIT VAN DIE
VRYSTAAT
YUNIVESITHI YA
FREISTATA

UFS·UV
NATURAL AND
AGRICULTURAL SCIENCES
NATUUR- EN
LANDBOUWETENSKAPPE

Appendix A2: Afrikaans translation of Appendix A1.

Departement Genetika (116) / Department of Genetics (116)

Fakulteit Natuur- en Landbouwetenskappe

Faculty of Natural and Agricultural Sciences

Posbus / P.O. Box 339

Bloemfontein 9300

South Africa

E-pos / E-mail: Genetics@ufs.ac.za

☎ +27-(0)51-401-2595

🖷 +27-(0)86-518-7317

Deelnemer inligting

U word uitgenooi om deel te neem in Clement Malan se Magister Scientiae (M.Sc.) projek wat sal plaas vind by die Departement van Genetika, Universiteit van die Vrystaat. Die projek se hoofdoel is om die allele frekwensies van drie neurotransmitter gene in verskillende etniese bevolkings te bepaal. Die rede hoekom die genetiese variasie in verskillende bevolkings bestudeer word is om te help analiseer hoekom verskillende bevolkings meer vatbaar is an sekere siektes as ander dus help met die behandeling van siektes.

Deelnemers word vereis om n' speeksel monster te skenk, waarvan DNS onttrek kan word. Die inligting benodig vir die projek word verkry deur die DNS van elke deelnemer. Deelnemers word vereis om aan of die San, Afrikaner, Xhosa of Sotho etniese groep te behoort en ok om oor die ouderdom van 18 te wees.

Geen risiko's is betrokke in die versameling van speeksel. Deelnemers word vereis om hul monde net met n' mondspoelmiddel te spoel en dan om die spoel middel in n' buis te deponeer. Die speeksel monsters wat ontvang word van elke etniese groep sal gevolglik gemerk word. Die name van die deelnemers sal nie in die studie ingesluit word nie daarom sal al die deelnemers anoniem bly. Geen resultate sal aan deelnemers verskaf word nie, navorsing inligting sal egter beskikbaar wees na publikasie. Deelname in die projek is verniet. Eweneens u sal geen geld of enige ander vergoeding kry vir deelname in die projek. U word toegelaat om te onttrek van die projek enige tyd sonder verantwoording stelling.

Die tydsduur van die projek sal ten minste twee jaar wees. Die DNS verkry sal met u toestemming in toekomstige studies gebruik word.

Voel vry om Prof. Spies of Mev. Schneider to kontak as daar enige vra oor die studie is. (051 4012261 of 051 4013730).

Dankie vir u samewerking.

Ingeligte toestemming

Lees of luister na die inligting gegee hier onder, gee u handtekening as n' saamstem

Ek verklaar dat ek verstaan en saam stem met die volgende inligting:

- Wat die navorsing projek behels.
- Ek vrywillig deelneem in die studie sonder betaling.
- Ek saamstem om n' speeksel monster te verskaf.
- Die speeksel monster net vir die onttrekking van DNS gebruik sal word.
- Die inligting verkry van analise van die DNS sal ingesluit word in die navorsing projek.
- Die DNS en inligting kan in toekomstige studies gebruik word.
- Alle inligting versamel sal vertroulik hanteer word.
- Resultate sal nie vir deelnemers gegee word nie.
- Ek word nie gedruk om deel te neem in die studie
- Die deelnemer mag enige tyd onttrek van die studie sonder verantwoording stelling.
- Ek het gelees of geluister en ek verstaan die gegewe inligting oor die projek en ek weet dat om n' monster te gee is my eie keuse.

Ek, die deelnemer, neem vrywillige deel in die bogenoemde navorsing projek.

Geteken by .....op.....20.....

....

Handtekening of Getuienis

regte duimafdruk van deelnemer

Algemene deelnemer inligting

Monster I.D.: .....

Verskaf asseblief die volgende inligting deur die regte leë blokkie te voltooi.

Geboortedatum: …… S.A. burger: Ja:☐ Nee: ☐

Geslag: Manlik:☐ Vroulik:☐

Ras: Asiatiese: ☐ Blanke:☐ Kleuring:☐

Ander: …… Swart:☐

Ouderdom (jaar):

18 - 22:☐ 57 – 61:☐

23 - 27:☐ 62 – 66:☐

28 - 32:☐ 67 – 71:☐

33 - 37:☐ 72 – 76:☐

38 - 42:☐ 77 +:☐

47 - 51:☐

52 - 56:☐

Huistaal:

Afrikaans:☐ Sesotho:☐

English:☐ Setswana:☐

IsiNdebele:☐ SiSwati:☐

IsiXhosa:☐ Tshivenda:☐

IsiZulu:☐ Xitsonga:☐

Sesotho sa Leboa:☐ Ander: .....

Dui die provinsie aan van waar u afkomstig is:

Oos – Kaap: ☐ Mpumalanga: ☐

Vrystaat: ☐ Noord Wes: ☐

Gauteng: ☐ Noord – Kaap: ☐

KwaZulu - Natal: ☐ Wes Kaap: ☐

Limpopo: ☐

Dui aan waar u groot geword het:

Plaas: ☐

Landelike gebied (platteland): ☐

Dorp: ☐

Stad: ☐

Appendix B: Hardy-Weinberg equilibrium data obtained for the *DRD4* VNTR, *MAO-A*-uVNTR females and 5-HTTLPR.

DRD4 VNTR results

```
================================
== Hardy-Weinberg equilibrium: (Khwe)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000

-------------------------------------------------------------------------
Locus  #Genot       Obs.Het.     Exp.Het.   P-value     s.d.  Steps done
-------------------------------------------------------------------------
    1      45       0.57778      0.46991    0.77498  0.00037     1001000
-------------------------------------------------------------------------


================================
== Hardy-Weinberg equilibrium: (Xhosa)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000

-------------------------------------------------------------------------
Locus  #Genot       Obs.Het.     Exp.Het.   P-value     s.d.  Steps done
-------------------------------------------------------------------------
    1      49       0.51020      0.51546    0.50077  0.00051     1001000


================================
== Hardy-Weinberg equilibrium: (Sotho)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000

-------------------------------------------------------------------------
Locus  #Genot       Obs.Het.     Exp.Het.   P-value     s.d.  Steps done
-------------------------------------------------------------------------
    1      58       0.39655      0.58501    0.00673  0.00008     1001000
```

```
================================
== Hardy-Weinberg equilibrium: (Afrikaner)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000
```

Locus	#Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	48	0.41667	0.47522	0.32181	0.00045	1001000

MAO-A-uVNTR female results

```
================================
== Hardy-Weinberg equilibrium: (Khwe)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000
```

Locus	#Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	28	0.60714	0.53831	0.67480	0.00046	1001000

```
================================
== Hardy-Weinberg equilibrium: (Xhosa)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000
```

Locus	#Genot	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	43	0.60465	0.59617	0.15880	0.00037	1001000

```
================================
== Hardy-Weinberg equilibrium: (Sotho)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000

-------------------------------------------------------------------------
Locus  #Genot       Obs.Het.     Exp.Het.   P-value     s.d.  Steps done
-------------------------------------------------------------------------
    1      44       0.72727      0.58281    0.04241  0.00020     1001000
-------------------------------------------------------------------------


================================
== Hardy-Weinberg equilibrium: (Afrikaner)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000

-------------------------------------------------------------------------
Locus  #Genot       Obs.Het.     Exp.Het.   P-value     s.d.  Steps done
-------------------------------------------------------------------------
    1      41       0.51220      0.45438    0.05193  0.00023     1001000
-------------------------------------------------------------------------
```

5-HTTLPR results

```
================================
== Hardy-Weinberg equilibrium: (Khwe)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000

-------------------------------------------------------------------------
Locus  #Genot       Obs.Het.     Exp.Het.   P-value     s.d.  Steps done
-------------------------------------------------------------------------
    1      46       0.23913      0.22145    1.00000  0.00000     1001000
-------------------------------------------------------------------------
```

```
================================
== Hardy-Weinberg equilibrium: (Xhosa)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000

-------------------------------------------------------------------------
Locus  #Genot       Obs.Het.     Exp.Het.   P-value     s.d.  Steps done
-------------------------------------------------------------------------
    1      69       0.27536      0.33439    0.15382  0.00036     1001000


================================
== Hardy-Weinberg equilibrium: (Sotho)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000

-------------------------------------------------------------------------
Locus  #Genot       Obs.Het.     Exp.Het.   P-value     s.d.  Steps done
-------------------------------------------------------------------------
    1      63       0.28571      0.37283    0.11283  0.00029     1001000


================================
== Hardy-Weinberg equilibrium: (Afrikaner)
================================

           Guo, S. and Thompson, E. 1992.
           Levene H. (1949).

Exact test using a Markov chain (for all Loci):
Forecasted chain length     :1000000
Dememorization steps        :100000

-------------------------------------------------------------------------
Locus  #Genot       Obs.Het.     Exp.Het.   P-value     s.d.  Steps done
-------------------------------------------------------------------------
    1      53       0.41509      0.50530    0.00840  0.00010     1001000
-------------------------------------------------------------------------
```

Appendix C: Sequence data obtained for the *DRD4* VNTR, *MAO-A*-uVNTR and 5-HTTLPR.

The complex 48 bp *DRD4* VNTR sequences are consecutively highlighted yellow and green.

2-repeat allele

```
1--------10--------20--------30--------40--------50--------60--------70--------80--------90
TCCCCCACGCCACCCGCGCCCCGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCCCCCCGGGACCCCTGC
91-------100-----------114
GGCTCCAACTGTGCTCCCCCCGAC
```

4-repeat allele (3 variants were identified, variations determined in third 48 bp segments)

```
1--------10--------20--------30--------40--------50--------60--------70--------80--------90
TCCCCCACGCCACCCGCGCCCCGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTTCCCCGGGGTCCCTGC
91-------100-------110-------120-------130-------140-------150-------160-------170------179
GGCCCCGACTGTGCGCCCGCCGCGCCCAGCCTCCCCCAGGACCCCTGTGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCCCCCCGGA
180-------190-------200-------210
CCCCTGCGGCTCCAACTGTGCTCCCCCCGAC

1--------10--------20--------30--------40--------50--------60--------70--------80--------90
TCCCCCACGCCACCCGCGCCCCGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTTCCCCGGGGTCCCTGC
91-------100-------110-------120-------130-------140-------150-------160-------170------179
GGCCCCGACTGTGCGCCCGCCGCGCCCGGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCCCCCCGGA
180-------190-------200-------210
CCCCTGCGGCTCCAACTGTGCTCCCCCCGAC

1--------10--------20--------30--------40--------50--------60--------70--------80--------90
TCCCCCACGCCACCCGCGCCCCGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTTCCCCGGGGTCCCTGC
91-------100-------110-------120-------130-------140-------150-------160-------170------179
GGCCCCGACTGTGCGCCCGCCGCGCCCGGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCCCCCCGGA
180-------190-------200-------210
CCCCTGCGGCTCCAACTGTGCTCCCCCCGAC
```

5-repeat allele

```
1--------10--------20--------30--------40--------50--------60--------70--------80--------90
TCCCCCACGCCACCCGCGCCCCGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTTCCCCGGGGTCCCTGC

91-------100-------110-------120-------130-------140-------150-------160-------170------179
GGCCCCGACTGTGCGCCCGCCGCGCCCGGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCTCCCGCGCCCGGCCTTCCCCGGGG

180-------190-------200-------210-------220-------230-------240-------250-----258
TCCCTGTGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCCCCCCGGACCCCTGCGGCTCCAACTGTGCTCCCCCCGAC
```

6-repeat allele

```
1--------10--------20--------30--------40--------50--------60--------70--------80--------90
TCCCCCACGCCACCCGCGCCCCGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTTCCCCGGGGTCCCTGC
91-------100-------110-------120-------130-------140-------150-------160-------170------179
GGCCCCGACTGTGCGCCCGCCGCGCCCAGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTTCCCCGGGG
180-------190-------200-------210-------220-------230-------240-------250-------260-----268
TCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCC
269--------280-------290-------300---306
CCCCGGACCCCTGCGGCTCCAACTGTGCTCCCCCCGAC
```

7-repeat allele

```
1--------10--------20--------30--------40--------50--------60--------70--------80--------90
TCCCCCACGCCACCCGCGCCCCGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTTCCCCGGGGTCCCTGC
91-------100-------110-------120-------130-------140-------150-------160-------170------179
GGCCCCGACTGTGCGCCCGCCGCGCCCGGCCTCCCCCCGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCCCCCAGGA
180-------190-------200-------210-------220-------230-------240-------250-------260-----268
CCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTTCCCCGGGGTCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCC
269--------280-------290-------300-------310-------320-------330-------340-----------354
CCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCCCCCCGGACCCCTGCGGCTCCAACTGTGCTCCCCCCGAC
```

8-repeat allele

```
1--------10--------20--------30--------40--------50--------60--------70--------80--------90
TCCCCCACGCCACCCGCGCCCCGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTTCCCCGGGGTCCCTGC
```

```
91-------100-------110-------120-------130-------140-------150-------160-------170------179
GGCCCCGACTGTGCGCCCGCCGCGCCCGGCCTCCCCCCGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCCCCCAGGA
180-------190-------200-------210-------220-------230-------240-------250-------260-----268
CCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCCGGCCTCCCCCCGGACCCCTGCGGCTCCAACTGTGCGCCCCCCGCGCCCGGCCTTC
269--------280-------290-------300-------310-------320-------330-------340-------350----357
CCCGGGGTCCCTGCGGCCCCGACTGTGCGCCCGCCGCGCCCGGCCTCCCCCAGGACCCCTGCGGCCCCGACTGTGCGCCCCCCGCGCCC
358---------370-------380-------390----------403
GGCCTCCCCCCGGACCCCTGCGGCTCCAACTGTGCTCCCCCCCGAC
```

The 30 bp *MAO-A*-uVNTR sequences are consecutively highlighted yellow and red, half repeat highlighted green.

2.5-repeat allele

```
1--------10--------20-------30--------40--------50--------60--------70-------80------88
GGCGGCACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCGAGCG
```

3.5-repeat allele

```
1--------10--------20-------30--------40--------50--------60--------70-------80--------90
GGCGGCACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCAGTACCC
91-------100-------110-----118
GCACCAGTACCGGCACCGGCACCGAGCG
```

4.5-repeat allele

```
1--------10--------20-------30--------40--------50--------60--------70-------80--------90
GGCGGCACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCAGTACCC
91-------100-------110-------120-------130-------140-----148
GCACCAGTACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCGAGCG
```

5.5-repeat allele

```
1--------10--------20-------30--------40--------50--------60--------70-------80--------90
GGCGGCACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCAGTACCC
91-------100-------110-------120-------130-------140-------150-------160-------170-----178
GCACCAGTACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCAGTACCCGCACCAGTACCGGCACCGGCACCGAGCG
```

The 5-HTTLPR 20-23 bp complex VNTR sequences are consecutively highlighted yellow and red. Reading motif starts from the second repeat.

S allele

```
1--------10--------20--------30--------40--------50--------60--------70--------80--------90
CCCCCCTGCAACCTCCCAGCAACTCCCTGTACCCCTCCTAGGATCGCTCCTGCATCCCCCATTATCCCCCCCTTCACCCCTCGCGGCATC
91-------100-------110-------120-------130-------140-------150-------160-------170------179
CCCCCTGCACCCCCAGCATCCCCCCTGCAGCCCTTCCAGCATCCCCCCTGCACCTCTCCCAGGATCTCCCCTGCAACCCCCATTATCCCC
180-------190-------200-------210-------220-------230-------240-------250-------260-----268
CCTGCACCCCTCGCAGTATCCCCCCTGCACCCCCCAGCATCCCCCATGCACCCCCGGCATCCCCCCTGCACCCCTCCAGCATTCTCCT
269--------280-----------294
TGCACCCTACCAGTATTCCCCCGCAT
```

L allele

```
1--------10--------20--------30--------40--------50--------60--------70--------80--------90
CCCCCCTGCAACCTCCCAGCAACTCCCTGTACCCCTCCTAGGATCGCTCCTGCATCCCCCATTATCCCCCCCTTCACCCCTCGCGGCATC
91-------100-------110-------120-------130-------140-------150-------160-------170------179
CCCCCTGCACCCCCAGCATCCCCCCTGCAGCCCCCCCAGCATCTCCCCTGCACCCCCAGCATCCCCCCTGCAGCCCTTCCAGCATCCCC
180-------190-------200-------210-------220-------230-------240-------250-------260-----268
CTGCACCTCTCCCAGGATCTCCCCTGCAACCCCCATTATCCCCCCTGCACCCCTCGCAGTATCCCCCCTGCACCCCCCAGCATCCCCC
269--------280-------290-------300-------310-------320-------330----338
ATGCACCCCCGGCATCCCCCCTGCACCCCTCCAGCATTCTCCTTGCACCCTACCAGTATTCCCCCGCAT
```