An Investigation into the Molecular Aetiology of Parkinson's Disease in South African Patients

By

Brigitte Glanzmann

Thesis presented in partial fulfilment of the requirements for the degree Master of Science in Medical Science (Human Genetics) in the Faculty of Medicine and Health Sciences at the University of Stellenbosch

Supervisor: Professor Soraya Bardien
Faculty of Medicine and Health Sciences
Department of Biomedical Sciences

Co-supervisor: Professor Jonathan Carr
Faculty of Medicine and Health Sciences
Division of Neurology

March 2013

DECLARATION

By submitting this thesis/dissertation, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: ……..

Date: ……...

Abstract

Parkinson's disease (PD) is a severely debilitating neurodegenerative disorder that results in motor circuit dysregulation and ultimately, causes impairment of movement. This condition is due to the selective degradation of the dopaminergic neurons in the substantia nigra pars compacta in the midbrain, which subsequently results in the pathological symptoms namely bradykinesia, resting tremor, postural instability and rigidity. It was initially hypothesized that individuals who develop PD were exposed to an environmental trigger(s) that caused the onset of the disease, but more recently, a significant genetic component, coupled to environmental factors have been implicated in disease pathogenesis. Currently, there are eight genes (*Parkin, PINK1, LRRK2, SNCA, DJ-1, ATP13A2, EIF4G1* and *VPS35*) that have been directly implicated in PD.

Worldwide, the prevalence of neurodegenerative disorders is increasing as populations are living longer. In Europe, Canada and USA, it has been projected that the prevalence of PD may increase by a factor of two between 2010 and 2050; approximately a 92% increase. In Tanzania (the only study done in sub-Saharan Africa) an even larger increase of 184% between 2005 and 2025 is predicted, due to the fact that the speed of populations ageing in developing countries, will exceed that of developed countries. Research into the causes and risk factors underlying neurodegenerative disorders such as PD is therefore urgently needed for policy makers and governments in developing nations to take appropriate action to deal with this impending health care problem.

The aim of the present study was to investigate the molecular aetiology of a group of South African PD patients. A total of 262 patients from various ethnic backgrounds were recruited for the study, and 35% had a positive family history of PD with the average age at onset (AAO) being 54.3 years of age (SD = 12.5 years). Mutation screening of the known PD genes (*Parkin, PINK1, LRRK2, SNCA* and *DJ-1*) was performed using high resolution melt and Sanger sequencing. Genotyping was done using fluorescently-labelled PCR primers followed by electrophoresis on an ABI 3130xl genetic analyser (for CTG repeats in *JPH3*) and with a KASP™ Genotyping Assay (for a 16bp indel in *DJ-1*). In order to identify a novel PD-causing gene, whole exome sequencing (WES) was conducted on three Afrikaner probands with an Illumina Genome Hiseq 2000™ and the sequences were aligned using the

NCBI Human Reference Genome 37.2. The BORG (Bio-Ontological Relationship Graph) semantic database, which models the relationship of human and model organism genes to functions, pathways and phenotypes, was used to filter and prioritise genetic variants shared between the three PD exomes.

It was determined that the known PD genes do not play a significant role in disease pathogenesis in the South African patients as only 15/262 (5.7%) of the patients harboured mutations: seven in *Parkin*, one in *PINK1*, six in *LRRK2* and one in *SNCA*. Only one of the patients harboured a 16bp indel variant at the transcription start site of *DJ-1*. None of the Black PD patients had pathogenic repeat expansions in *JPH3* thereby excluding Huntington disease-like 2 as a cause of the disease phenotype.

Genealogical analysis revealed that six of the apparently unrelated Afrikaner PD probands were related to a founder couple that immigrated to South Africa in the 1600s which suggests that there is a possible founder effect for the disease. Bioinformatics analysis of WES data on three of the probands identified 21 variants in 12 genes that were present in all three PD exomes and fulfilled various criteria. Sanger sequencing was used for verification of five variants and of these, two (in *CDC27* and *NEDD4*) were found to be artefacts. The remaining three (in *HECDT1*, *TBCC* and *RNF40*) were excluded based on the lack of co-segregation with disease and the high frequency of the allele in controls. Further work is necessary to verify the presence of the remaining sixteen variants and to characterise each of them for their possible pathogenicity.

The discovery of novel PD-causing genes is important as this may shed light on the pathways or processes that are involved. A current hypothesis implicates the lysosome-dependent pathway as a unifying biochemical pathway that can account for the phenotypic spectrum within PD. Notably, although Mendelian forms are thought to account for only about 10-15% of cases, the study of Mendelian inherited variants is likely to provide insight into the pathophysiology of the more common sporadic form of this condition. Dissecting the key molecular mechanisms underlying PD will provide critical information for improved treatment strategies and drug interventions that will ultimately prevent or halt neuronal cell loss in susceptible individuals.

Opsomming

Parkinson se siekte (PS) is 'n erge neurodegeneratiewe bewegings-siekte, wat motorstroombaan disregulasie veroorsaak. Dit lei uiteindelik tot beperkte bewegings vermoëns. Hierdie toestand word veroorsaak weens die selektiewe agteruitgang van die dopaminergeniese neurone in die substantia nigra pars compacta in die midbrein, wat later lei tot die patologiese simptome naamlik: bradykinesia, rustende spiersametrekkings, posturale onstabiliteit en rigiditeit. Daar is aanvanklik vermoed dat individue wat PS ontwikkel, aan 'n omgewingsfaktor(e) blootgestel is wat die aanvang van die siekte veroorsaak het, terwyl meer onlangs is daar 'n aansienlike genetiese komponent tesame met omgewingsfaktore geïdentifiseer, wat betrokke is by die patogenese van die siekte. Tans is daar agt gene (*Parkin*, *PINK1*, *LRRK2*, *SNCA*, *DJ-1*, *ATP13A2*, *EIF4G1* en *VPS35*) wat direk by PS geïmpliseer is.

Wêreldwyd is daar ‘n toenemende voorkoms van neurodegeneratiewe siektes aangesien bevolkings langer leef. In Europa, Kanada en die VSA, is daar geprojekteer dat die voorkoms van PS tussen 2010 en 2050 met 'n faktor van twee verhoog kan word. Dit is ongeveer 'n 92%-verhoging. In Tanzanië (die enigste studie wat tot dusver in sub-Sahara Afrika gedoen is) word daar selfs ‘n groter toename, van 184% tussen 2005 en 2025 voorspel. Dit is te danke aan die feit dat die bevolkings- veroudering in ontwikkelende lande die van ontwikkelde lande sal oorskry. Ondersoeke na die oorsake en risiko-faktore onderliggend aan neurodegeneratiewe siektes, byvoorbeeld PS, word dus dringend benodig deur beleidmakers en regerings in ontwikkelende lande, sodat hulle die nodige stappe kan neem om hierdie dreigende gesondheidsorg-probleem op te los.

Die doel van die huidige studie was om ondersoek in te stel na die molekulêre etiologie van 'n groep Suid-Afrikaanse PS pasiënte. 'n Totaal van 262 pasiënte van verskillende etniese agtergronde, is gewerf vir die studie. Hiervan het 35% 'n positiewe familiegeskiedenis van PS en die gemiddelde aanvangs ouderdom (AAO) was 54,3 jaar (SD = 12,5 jaar). Mutasie-analise van die bekende PS gene is uitgevoer met behulp van hoë resolusie smelt en Sanger volgordebepaling. Genotipering is gedoen met behulp van fluoresserend geëtiketteerde PKR inleiers met elektroforese, op 'n ABI 3130xl genetiese analiseerder (CTG herhalings in *JPH3*), en met 'n KASP ™ Genotipering toets (vir 'n 16bp indel in *DJ-1*). Ten einde, om 'n nuwe PS-veroorsakende geen te identifiseer was heel eksoom volgordebepaling (WES) uitgevoer op drie Afrikaner PS positiewe pasiënte met 'n Illumina Genome Hiseq 2000™ en die volgorders is

gerangskik met behulp van die NCBI Menslike Verwysings Genoom 37.2. Die BORG (Bio-Ontologiese Verhoudings Grafiek) semantiese databasis, wat gebaseer is op die verhouding van die mens en model organisme gene funksies, paaie en fenotipes, en is gebruik om genetiese variante, wat gedeel word tussen die drie PS exome te filtreer en te prioritiseer.

Daar is vasgestel dat die bekende PS gene nie 'n belangrike rol in die patogenese van die siekte in die Suid-Afrikaanse pasiënte speel nie. Dit is aangesien slegs 15/262 (5.7%) van die pasiënte bekende mutasies dra: sewe in *Parkin*, een in *PINK1*, ses in *LRRK2* en een in *SNCA*. Slegs een van die pasiënte het 'n 16bp delesie variant in die transkripsie promotor area van *DJ-1* gedra. Geen van die Swart PS pasiënte het patogeniese herhalings in *JPH3* vertoon nie. Gevolglik is Huntington siekte-agtige 2 uitgesluit as 'n oorsaak van die siekte fenotipe.

Genealogiese analise het getoon dat ses van die skynbaar onverwante Afrikaner PS pasiënte verwant is aan 'n stigter paartjie wat in die 1600's na Suid-Afrika geïmigreer het, wat daarop dui dat daar 'n moontlike stigter effek vir die siekte is. Bioinformatiese analise van WES data vir drie van die pasiënte, het 21 variante in 12 gene geïdentifiseer, wat in al drie PS exome teenwoordig was en verskeie kriteria vervul het. Sanger volgordebepaling is gebruik vir die bevestiging van vyf variante en van hierdie, is twee (in *CDC27* en *NEDD4*) bevind om artefakte te wees. Die oorblywende drie (in *HECDT1*, *TBCC* en *RNF40*) is uitgesluit gebaseer op die gebrek aan gesamentlike-segregasie met die siekte en die hoë frekwensie van die allele in die kontrole groep. Verdere werk is nodig om die teenwoordigheid van die oorblywende variante te verifieer en om elkeen van hulle te karakteriseer vir hulle moontlike patogenisiteit.

Die ontdekking van die nuwe PS-veroorsakende gene is belangrik aangesien dit lig kan werp op die stelsels of prosesse wat betrokke is. 'n Huidige hipotese impliseer die lisosoom-afhanklike pad as 'n verenigende biochemiese padweg, wat verantwoordelik is vir die fenotipiese spektrum binne PS. Alhoewel Mendeliese vorms vermoedelik verantwoordelik is vir slegs omgeveer 10-15% van die gevalle, is die studie van Mendelse gene geneig om insig te verkry in die patofisiologie van die meer algemene sporadiese vorm van hierdie toestand. Ontleding van die kern molekulêre meganismes onderliggend aan PS sal kritiese inligting vir beter strategieë vir behandeling en geneesmiddel-intervensies voorsien, wat gevolglik neuronale sel verlies in vatbare individue sal voorkom of beëindig.

Table of Contents

Acknowledgements

There have been so many individuals who have actively contributed to this thesis, directly and indirectly, past and present, that the acknowledgements could easily become the most intricate part of my thesis.

I would like to thank Professor Soraya Bardien, my supervisor, for her unfailing support and encouragement throughout the course of my thesis. You have needed to do and say very little to inspire me and the effort that you made with the many drafts that I sent you was far more than I could ever have expected from you. The respect that I have for you cannot be measured and I look forward to the future projects we aim to tackle together.

I also owe thanks to Dr. Craig Kinnear and Professor Paul van Helden; Craig for making the stormier days a little bit easier and Professor van Helden for understanding when I needed a step away from the Department, even when it was expected of me to do otherwise.

To my colleagues in the MAGIC lab; there have been days when I have been less than palatable and many of you have provided much needed advice and motivation when it was necessary.

This study would not have been possible without the support of the Harry Crossley Foundation, Medical Research Council, National Research Foundation; being afforded an opportunity to pursue a career in research comes with great responsibility and I will strive to do what I can to live up to this responsibility.

Two people have shared the past two years with me personally; my sister Claudia for keeping my feet firmly on the ground and my head facing in the right direction and my mother, Catherine for your patience and support. I have not forgotten the sacrifices that you have made for me to be here, nor shall I.

List of Abbreviations

AAO	Age at onset
AD	Autosomal dominant
ANK	Ankyrin repeat domain
AP/MS	Affinity Purification or Mass Spectrometry
AR	Autosomal recessive
ARM	Armadillo domain
ATP13A2	ATPase type 13 A2
BBB	Blood brain barrier
BE	Biliary epithelial
BORG	Bio-Ontological Relationship Graph
CAF	Central Analytical Facility
CaMK	Calcium/calmodulin dependant kinases
CDC27	Cell division cycle protein 27
CNS	Central nervous system
CNV	Copy number variations
COR	Carboxy terminal of ROC
Ct	Cycle threshold
CVCD	Common variants underlie common diseases
DALY	Disability adjusted life year
DBS	Deep brain stimulation
ddNTP	Di-deoxyribonucleotide triphosphate
DHODH	dihydroorotate dehydrogenase
DJ-1	Daisuke-Junko-1
dNTP	Deoxyribonucleotide triphosphate
dsDNA	Double stranded DNA
DUB	Deubiquitinating enzyme
EIF4G	Eukaryotic translation initiation factor 4 gamma
ExoI	Exonuclease I

FRET	Fluorescent resonance energy transfer
GEOPD	Genetic Epidemiology of Parkinson's disease
GO	Gene Ontology
GTP	Guanosine triphosphate
GWAS	Genome wide association studies
HD	Huntington's disease
HDL2	Huntington's disease-like 2
HECTD1	HECT domain containing 1
HEK23	Human embryonic kidney
HP	Human Phenotype
HRM	High resolution melt
IBR	In-between RING
IDT	Integrated DNA Technologies
ITGAX	Integrin alpha X
JPH3	Junctophilin-3
LB	Lewy body
LRR	Leucine rich repeat domain
LRRK2	Leucine rich repeat kinase 2
MAO	Monamine oxidase
MAP	Microtubule associated protein
MAPKKK	Mitogen–activated protein kinase kinase kinase
MAPT	Microtubule-associated protein tau
MLPA	Multiplex ligation-dependent probe amplification
MNS	Mental, neurological and substance abuse
MP	Mammalian Phenotype
MPTP	1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine
MRI	Magnetic resonance imaging
MTS	Mitochondrial targeting domain
NAC	Non-amyloid-B component
NEDD4	Neural precursor cell expressed developmentally down-regulated protein 4
NGS	Next generation sequencing

NHLS	National Health Laboratory Services
PARK2	Parkin
PCR	Polymerase chain reaction
PD	Parkinson's disease
PINK1	PTEN-induced kinase 1
PW	Pathway
RFLP	Restriction fragment length polymorphism
RING	Really interesting new gene
RNF40	Ring finger protein 40
ROC	Ras of complex proteins
ROS	Reactive oxygen species
RVCD	Rare variants underlie common disorders
SANBI	South African National Bioinformatics Institute
SAP	Shrimp alkaline phosphatase
SCA	Spinocerebellar ataxia
scaRNA	small Cajal body-specific RNA
SD	Standard deviation
SDS	Sequence detection system
SNCA	α-synuclein
snoRNA	small nucleolar RNA
SNP	Single nucleotide polymorphism
SNpc	Substantia nigra pars compacta
SNV	Single nucleotide variant
SPECT	Single photon emission computerized tomography
SSA	sub-Saharan Africa
ssDNA	Single stranded DNA
Ta	Annealing temperature
TBCC	Tubulin folding cofactor C
Tm	Melting temperature
TM	Transmembrane region
TNF	Tumour necrosis factor

TRAP	TNF receptor associated protein
Ub	Ubiquitin
UBL	Ubiquitin-like domain
UPD	Unique Parkin domain
UPS	Ubiquitin proteasome system
VPS35	Vacuolar protein sorting-associated protein 35
WES	Whole exome sequencing
WHO	World Health Organization
WT	Wild type
αSYN	α-synuclein protein

List of Figures

Page

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 8

Chapter 9

List of Tables

Page

Chapter 8

Chapter 9

Outline of the Thesis

In order to improve the flow of the different sub-sections of the study, we have deviated from convention by dividing the thesis into nine chapters and the layout is as follows:

- **Chapter One** provides a comprehensive background on what is currently known about Parkinson's disease (PD) worldwide, previous findings on the South African patients and the overall aims and objectives of the present study.

- **Chapters Two** to **Eight** are the experimental chapters, each including results.
    - **Chapters Two, Three, Four, Five** and **Six** outline the mutation screening of the known PD genes
    - **Chapter Seven** describes a study into PD in Black African individuals and its possible clinical overlap with Huntington's disease-like 2.
    - **Chapter Eight** encompasses whole exome sequencing, the first of its kind to be performed on South African PD patients, with the underlying hypothesis that a founder effect may exist for PD in the South African Afrikaner population.

- **Chapter Nine** provides the general conclusions that can be drawn from the entire study, stresses the need for more studies to be carried out on the sub-Saharan African populations and also describes future work which may be carried out to further the knowledge on South African PD patients.

Chapter 1: General Introduction

Page

1.1 Introduction

The 4th November 2006 marked the 100th anniversary of the first documentation of proteinaceous depositions in the brains of patients (Heemels 2006). The investigation into this broad and complex field began when Alois Alzheimer identified proteinaceous amyloid plaques in the brains of his severely demented patients on post-mortem (Alzheimer 1892). More than 100 years later, other neurodegenerative disorders such as Parkinson's disease, Huntington's disease and amyotrophic lateral sclerosis have been differentiated and it has been demonstrated that patients suffering from these disorders also have proteinaceous inclusions and depositions. To date, definitive answers have yet to be provided to Alzheimer's original question of whether the depositions are a causal factor in neurodegeneration (Rubinsztein 2006). There are currently two hypotheses that exist regarding the development of these disorders: 1) Neurodegeneration causes the proteinaceous plaque deposits, but protein aggregation has no causal role in the development of the disease; 2) Proteinaceous plaques form aggregations and deposits which subsequently result in neurodegeneration (Lansbury and Lashuel 2006).

According to a report published by the World Health Organization (WHO) in 2008, it has been estimated that collectively mental, neurological and substance abuse (MNS) disorders contribute to as much as 13% of the global disease burden (Mathers et al. 2008) (Table 1.1), which is greater than the burden of both cardiovascular disorders and cancer. Although scientific research has provided meaningful insight into MNS disorders, the inability to provide a cure or to provide possible preventative strategies for these devastating illnesses, is an indication of the lack of understanding of the brain and its pathways - at both a molecular and biochemical level (Ferri et al. 2005; Mathers et al. 2008). There are few available treatment strategies for MNS disorders, but even these remain unavailable to the thousands of patients who are in need of them, particularly in the developing countries. It has been estimated that in some countries as many as 25% of patients do not have access to anti-epileptic drugs and alarmingly, as many as 83% of patients are unable to receive anti-Parkinsonian medication; these figures are estimates for low income or impoverished countries such as those on the African continent, the East Asian regions and areas within the South Americas (Mathers et al. 2008).

Currently, the lack of understanding regarding MNS disorders has led to the exploration of the impact of war, natural disasters and other additional environmental factors on disease development (Saxena et al. 2008). The investigation into the role of genes and the environment and their interactions will also provide valuable insight into the aetiology of MNS disorders. However, the translation of scientific research into clinical practice remains a stumbling block for both treatment and cure of these disorders (Palop, Chin, and Mucke 2006).

Parkinson's disease (PD) (OMIM # 168600), the focus of the present study (highlighted in Table 1.1), is among the top 14 MNS disorders worldwide and is the second most common neurodegenerative disorder, preceded only by Alzheimer's disease (Beddington et al. 2008). It is characterised by a range of motor symptoms inevitably resulting in an impairment of the affected individual's motor skills (Gasser 2001). This disorder was first identified in 1817 by the English physician Dr. James Parkinson in his essay entitled *"An essay on the shaking palsy"* (Parkinson 1817). This appears to be the first known documentation of PD. Approximately 60 years after Dr. Parkinson's description, the French neurologist Dr. Jean Martin Charcot formally recognised Parkinson's disease and named it as such.

Table 1.1 Global burden of mental, neurological and substance abuse (MNS) disorders.

	Worldwide		High - income countries		Low and middle - income countries	
Rank	**Cause**	**DALYs[1] (millions)**	**Cause**	**DALYs (millions)**	**Cause**	**DALYs (millions)**
1	Unipolar depressive disorders	65.5	Unipolar depressive disorders	10.0	Unipolar depressive disorders	55.5
2	Alcohol abuse disorders	23.7	Alzheimer's and dementias	4.4	Alcohol abuse disorders	19.5
3	Schizophrenia	16.8	Alcohol abuse disorders	4.2	Schizophrenia	15.2
4	Bipolar affective disorder	14.4	Drug abuse disorders	1.9	Bipolar affective disorder	12.9
5	Alzheimer's and dementias	11.2	Schizophrenia	1.6	Epilepsy	7.3
6	Drug abuse disorders	8.4	Bipolar affective disorder	1.5	Alzheimer's and dementias	6.8
7	Epilepsy	7.9	Migraine	1.4	Drug abuse disorders	6.5
8	Migraine	7.8	Panic disorder	0.8	Migraine	6.3

9	Panic disorder	7.0	Insomnia	0.8	Panic disorder	6.2
10	Obsessive compulsive disorder	5.1	**Parkinson's disease**	**0.7**	Obsessive compulsive disorder	4.5
11	Insomnia	3.6	Obsessive compulsive disorder	0.6	Post traumatic stress disorder	3.0
12	Post traumatic stress disorder	3.5	Epilepsy	0.5	Insomnia	2.9
13	**Parkinson's disease**	**1.7**	Post traumatic stress disorder	0.5	Multiple sclerosis	1.2
14	Multiple sclerosis	1.5	Multiple sclerosis	0.3	**Parkinson's disease**	**1.0**

Data taken from the World Health Organization *The Global Burden of Disease, 2004.* [1]Disability Adjusted Life Year (DALYs) - a unit for measuring the amount of health lost due to disease or injury.

This debilitating disease is currently without a cure, and the pathological **degeneration** of the dopaminergic neurons of the substantia nigra pars compacta (SNpc) in the midbrain is the main cause for the movement impairment. Decreases in the levels of the neurotransmitter dopamine at the nerve terminals leads to a dysregulation of the motor circuits (Cookson and Bandmann 2010) (figure 1.1). Motor circuit dysregulation not only affects movement in several ways, but also has a major impact on the psychiatric and cognitive states of the patient (Cookson and Bandmann 2010).

The pathology of PD is relatively well understood but the aetiology of the disease remains largely unknown. Historically, the predominant view regarding the aetiology was that it was solely due to environmental factors (Dawson and Dawson 2003). In recent times, it is hypothesized that a combination of genetic and environmental factors may lead to disease development (Moore et al. 2005). For this reason, along with the significant clinical diversity between PD patients, it is likely that several molecular pathways involving different genes and effectors will influence the survival of the dopaminergic neurons (Moore et al. 2005). These pathways include the autophagy-lysosomal and ubiquitin proteasomal systems - both of which are involved in the degradation of excess, unwanted and misfolded proteins within the cell (Pan et al. 2008), cell signalling, mitochondrial electron transport, glucose utilization and glucose sensing (Betarbet et al. 2006). Mitochondrial dysfunction, impaired drug and toxin handling as well as protein misfolding/aggregation are alternate pathways which could lead to an increase in neurodegeneration and subsequent PD .

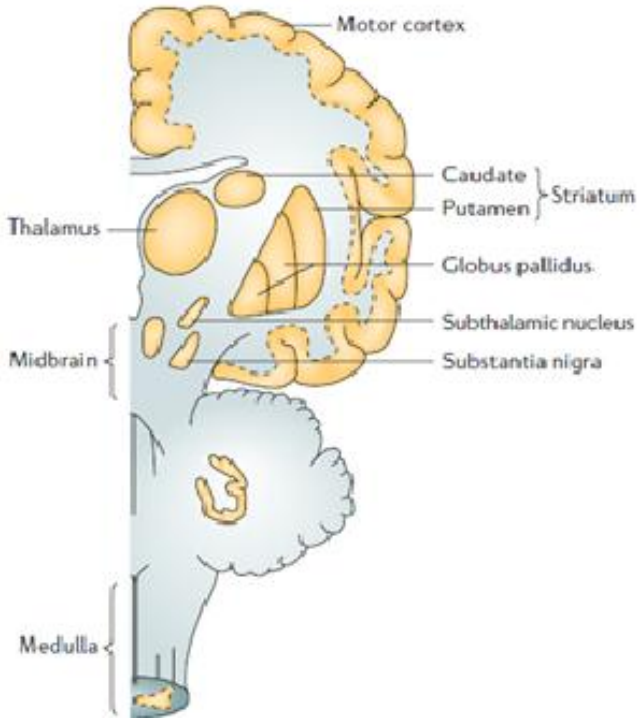


Figure 1.1 A representation of the brain regions affected by PD. The yellow regions are an indication of the affected regions – voluntary movements are initiated in the motor cortex and output regulated to the brain stem which is comprised of the midbrain, medulla and spinal cord. The transfer of the output signal is managed by so-called sub-cortical targets which include the thalamus, putamen and subthalamic nucleus (Taken from Farrer 2006).

1.2 Diagnosis, Symptoms and Treatment of PD

Diagnostic criteria used for the clinical diagnosis of PD have been established by the UK Parkinson's Disease Society Brain Bank (Gibb and Lees 1988). The clinical diagnosis is predominantly based on the motor symptoms presented by the patient (Fearnley and Lees 1991) and these include:

- bradykinesia
- resting tremor
- rigidity
- postural instability

The diagnosis of the disorder is dependent on the presence of three of the four above mentioned symptoms, but bradykinesia (the inability of a patient to start and continue movements, as well as the inability to adjust the positioning of the body) is an essential symptom (Brooks 2010). These motor symptoms are thought to arise in a patient when

approximately 80% of the striatal dopamine and 50% of the nigral neurons are lost (Fearnley and Lees 1991). Non–motor symptoms of PD include the development of dementia as well as other psychiatric disturbances (Siderowf 2001). Additionally, it has been reported that patients may have visual hallucinations, as well as aggression and paranoia, particularly towards family members (Naimark et al. 1996). A number of additional, early symptoms have been recorded in numerous patients who subsequently went on to develop PD and these are now considered when examining and diagnosing possible PD patients. These symptoms precede the onset of the disease by up to ten years and include olfactory disturbances (such as a loss of smell) incontinence, constipation, depression and sleep disturbances (Adler 2005; Parkinson's Disease Foundation). Sensory symptoms such as muscle soreness and numbness in the extremities (Samii, Nutt, and Ransom 2004), as well as variations in sleeping patterns, dyskinesias and excessive sweating have been documented but have also been shown to be exacerbated when treatment therapies are employed (Samii, Nutt, and Ransom 2004).

Treatment strategies for PD patients are relatively limited, with the administration of levodopa (a precursor of dopamine) being the most common and to date, most effective treatment strategy (Gibb and Lees 1988). Additional treatment approaches include monoamine oxidase (MAO) inhibitors, dopamine antagonists and amantadine (Jankovic 2006). Surgery is also an effective treatment for certain patients and may include the implantation of deep brain simulation (DBS) devices that are implanted into the thalamus, sub thalamic nucleus or palladium (University of Maryland Medical Centre). Additionally, advancements in stem cell therapy, the transplantation of foetal neurons and gene silencing techniques provide possible alternative treatment strategies for PD (Snyder and Olanow 2005; McCormack et al. 2010).

The distinction of PD from other Parkinsonian disorders is difficult due to overlapping symptoms and for this reason various diagnostic techniques using neuro-imaging have been developed. However, these have not been widely implemented and often a patient's rapid response to levodopa treatment, is typically used for confirming a PD diagnosis (Brooks et al. 2010).

1.3 Neuropathological Identification of PD

PD can only be diagnosed with certainty following an autopsy. Anatomically, the substantia nigra (figure 1.2) is examined as dopaminergic neuronal loss in this region is considered to be a pathological characteristic of PD (Zimprich et al. 2004). In an unaffected individual, the substantia nigra has a black appearance (figure 1.3), caused by the elevated levels of melanin in the dopaminergic neurons in this region (Dawson and Dawson 2003; Jankovic 2008). In a PD patient, the substantia nigra loses the black appearance due to the loss of the dopaminergic neurons (figure 1.3) (Adler et al. 2005; Jankovic 2008).

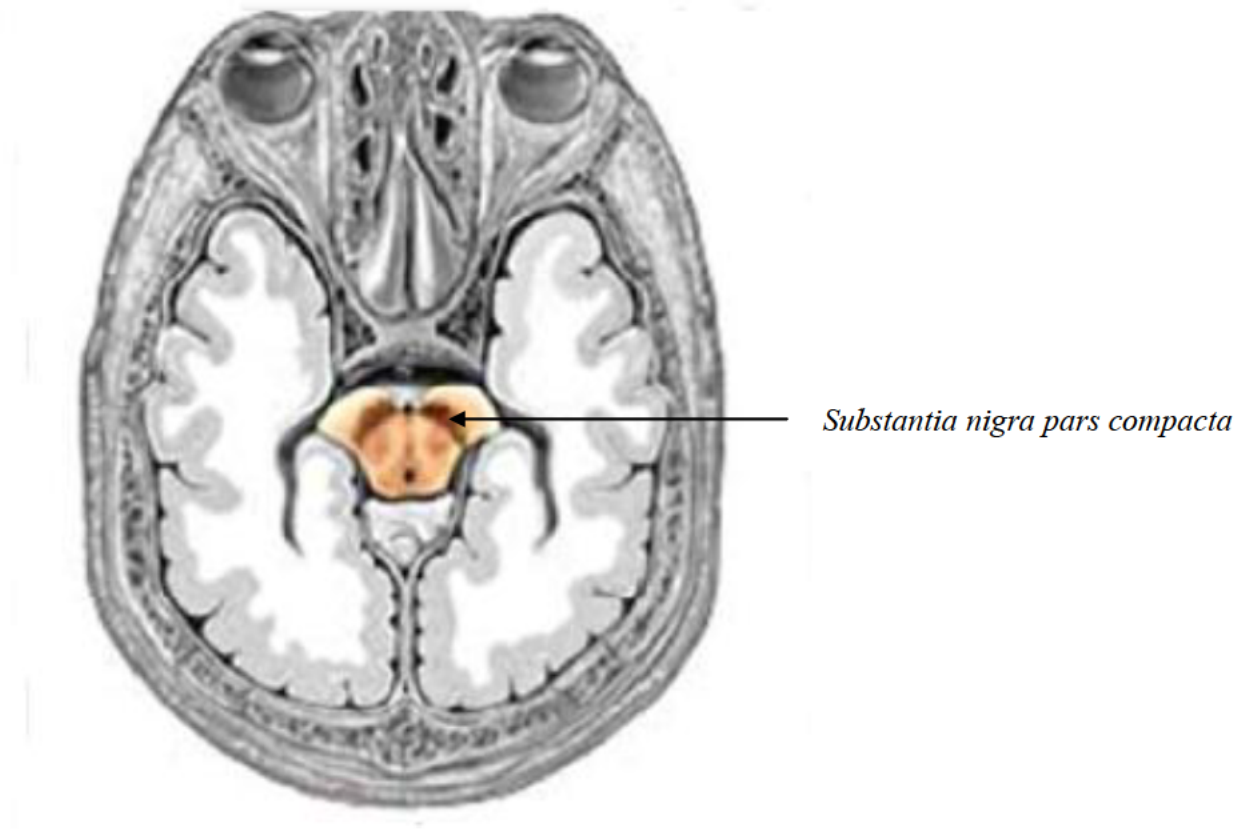


Figure 1.2 A representation of the location of the substantia nigra in the brain. A transverse/horizontal section through the midbrain shows the anatomical locality of the substantia nigra (Taken from http://health.allrefer.com/health/parkinsons-disease-substantia-nigra-and-parkinsons-disease html).

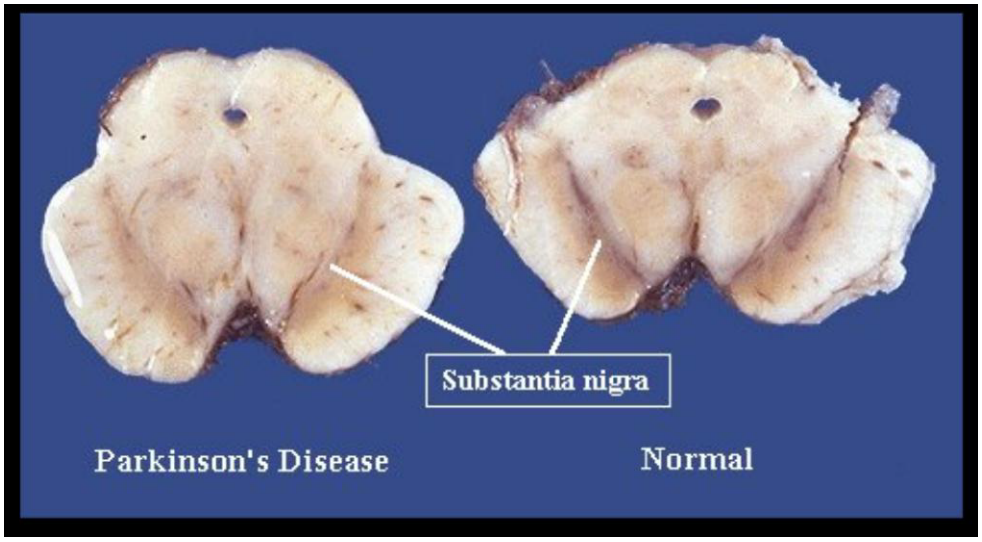


Figure 1.3 A representation of the substantia nigra of a PD patient and an unaffected individual. The substantia nigra is almost completely indistinguishable in an affected individual. This is due to the dopaminergic neuronal loss (Taken from http://gwc maricopa.edu/class/bio201/parkn/jcadis51 htm).

The degradation of the substantia nigra is not the only pathological characteristic that can be identified following autopsy in PD patients. The presence of Lewy bodies (LBs) (figure 1.4) and Lewy body neurites can be identified in the remaining neurons found in the substantia nigra (Jankovic 2008). Both can be visualized through immunohistochemical staining (Love 2005). LBs are intra-cytoplasmic inclusions that are highly proteinaceous and have a dense eosinophilic core, while Lewy body neurites are nerve cell processes that contain aggregates of α-synuclein and other proteins, all of which are highly ubiquinated (Gasser 2001). The major fibrillar component of LBs and Lewy body neurites is the protein, α-synuclein that is predominantly expressed in the neocortex, substantia nigra, cerebellum and thalamus (Dawson and Dawson 2003; J Jankovic 2008). The current hypothesis regarding these pathological hallmarks is that amino acid changes, whole gene duplications or triplications in α-synuclein, may lead to an increased possibility that aggregations of this protein will form, ultimately resulting in neuronal dysfunction and death (Goedert and Spillantini 1998; Karpinar et al. 2009).

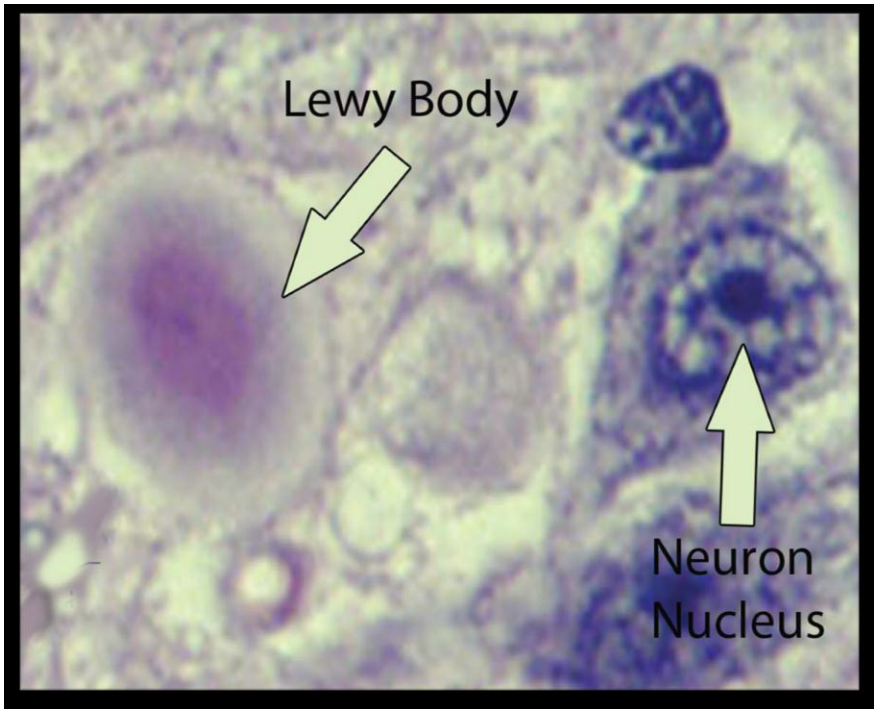


Figure 1.4 An immunohistochemical stain showing a Lewy body (LB) in an affected individual. LBs are intra-cytoplasmic inclusions which are identified in PD patients through autopsy (Taken from http://www.virtualmedstudent.com/links/neurological/parkinsons html).

1.4 Incidence and Prevalence of PD

The incidence of a disease is defined by the total number of new cases that present in a given population over a specified time. Prevalence is defined as the portion of individuals within a population that are affected by the disease (Sellbach et al. 2006). PD occurs worldwide, but the prevalence and incidence of this disorder differs significantly, particularly across geographic locations as well as across various ethnic groups – generally lower rates have been reported in Africa (Okubadejo et al. 2006). It has been hypothesized that this may be due to genetic and environmental variations as well as cultural and ethnic differences between patients (Okubadejo et al. 2006). The crude incidence rate of PD globally is estimated to be around 17 per 100 000 per year – in Africa, this figure is estimated to be at around 4.5 per 100 000 (Melcon et al. 1997), significantly lower than the global rates. Africa has a crude prevalence rate of between 7 and 43 per 100 000, while globally, this figure has been estimated to be between 7 and 657 per 100 000 (Melcon et al. 1997; Okubadejo et al. 2006). PD is not, however restricted to older individuals. Patients who present with clinical manifestations of the symptoms before the age of 50 years, are considered to have early onset PD (Periquet et al. 2003). Late onset PD occurs in patients with an age at onset (AAO) of 50

years and older. Juvenile PD has been recorded, where patients have an AAO of 20 years and younger. It is estimated that PD affects 1% of individuals over the age of 65 years; this figure increases significantly as the age of the population increases, with as many as 4% of individuals over the age of 85 being affected by the disorder (Eeden et al. 2003). The incidence rate of PD has also been recorded as being higher in men (almost double) than in women (Luecking et al. 2000). A proposed hypothesis for these gender differences is the fact that women produce oestrogen which may have a neuroprotective role (Gillies and McArthur 2010). The proposed mechanisms involving oestrogen include an increase in blood flow that promotes the removal of neurotoxins from the cranial regions, antioxidant functions and the inhibition of the MAO enzyme (Dluzen and McDermott 2000).

1.5 Environmental Risk Factors for PD

Neurodegeneration involves a loss of neuronal structure as well as function. Prior to advancements in molecular genetics, it was initially hypothesized that PD was caused by exposures to various environmental factors (Elbaz and Tranchant 2007). Studies have concluded that it is highly unlikely that neurodegenerative disorders such as PD are caused by exposures to a single environmental agent (Cannon and Greenamyre 2011). Instead, epidemiological studies have provided evidence that there are a number of compounds and risk factors that may play a role in the development of PD (Cannon and Greenamyre 2011). Exposure to pesticides, herbicides and fungicides such as rotenone and paraquat are some of the most well studied examples. Paraquat is commonly used in developing countries as a broad-spectrum herbicide and does not appear to enter the brain across the blood-brain barrier but is transferred through a neutral amino acid channel (McCormack 2003). Paraquat has been identified as a role player in redox cycling, thereby exerting a negative effect on dopaminergic neurons (Cannon and Greenamyre 2011) (figure 1.5). Rotenone is another toxin that has been widely studied and is used as an insecticide. It is a lipophilic compound that can cross the blood-brain barrier (figure 1.5); it is a selective inhibitor of mitochondrial complex I and has been implicated in the development of PD (Cannon and Greenamyre 2011; Tanner et al. 2011). Another molecule capable of crossing the blood brain barrier is the synthetic bi-product of heroin, 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) (Dauer and Przedborski 2003). This molecule is highly neurotoxic and when transported into the dompaminergic neurons of the substantia nigra, irreparable damage to these neurons results

in Parkinsonian features. MPTP is also used to generate primate models of PD (Dauer and Przedborski 2003).

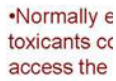


Figure 1.5 Toxin entry into the brain and the interactions with various cell types. Toxins may enter the brain and central nervous system (CNS) through transporters, if it shares structural similarity with another molecule. Alternatively, the toxins may be transported across the blood-brain barrier. Ex, example; BBB, blood brain barrier (Taken from Cannon and Greenamyre 2011).

Other environmental compounds that have been implicated in PD pathogenesis include organophoshates, organochlorines and carbamates but there is very limited epidemiological data available to support this (Cannon and Greenamyre 2011). Exposure to metals has also been investigated in various neurodegenerative disorders; increases in iron levels have been observed in the substantia nigra of PD patients but it is unclear whether accumulations of the compound is as a result of the disease process, or whether the increased iron levels contribute to the pathogenesis of the disorder (Oakley et al. 2007). Abnormally high exposures to manganese result in striatal dopamine depletion and subsequently Parkinsonism, but the phenotypic features differ distinctly from typical PD (Santamaria and Sulsky 2010). Parkinsonism is defined as the clinical manifestation of some PD symptoms but the predominant features are atypical or additional neurological features not commonly found in typical PD cases (Klein, Schneider, and Lang 2009).

Manganese exposure also plays a significant role in mitochondrial dysfunction as it accumulates in the mitochondria and impairs oxidative metabolism (Cannon and Greenamyre 2011). Additional environmental factors that may contribute to PD pathogenesis include rural living, well water drinking and head trauma (Sanyal et al. 2010). It has been suggested that patients who have suffered severe head trauma are two and a half times more likely to develop PD (Goldman et al. 2006; Elbaz and Tranchant 2007; Tanner et al. 2011). It should be noted however, that environmental exposures do not only predispose to disease development as certain factors such as smoking and caffeine intake have been shown to protect against the development of PD. Smoking is one of the most widely studied lifestyle factors and its role in PD has been extensively studied (Wirdefeldt et al. 2011). Numerous mechanisms have been proposed for the way in which smoking may play a neuroprotective role - the compounds contained in tobacco (anabasine, cotinine, hydroquinone, nicotine and nomicotine) may play a cumulative role in neuroprotection, but nicotine is proposed to be the most likely compound as it is a dopaminergic stimulant and relieves PD symptoms (Alves et al. 2008; Wirdefeldt et al. 2011). Interestingly, the combination of nicotine and hydroquinone have been shown to inhibit the formation of α-synuclein fibrils in mouse models, thereby providing strong evidence of the neuroprotective role of smoking in PD (Hong, Fink, and Uversky 2009; Wirdefeldt et al. 2011). Caffeine is the most widely used psychoactive stimulant because of its presence in substances such as coffee and energy drinks as well as other commercially available products (Douna et al. 2012). It has been shown to abate the depletions in striatal dopamine in mouse models that had been treated with MPTP to induce Parkinsonian features. Further exploration also revealed that continual intake of caffeine reduced MPTP toxicity and resulted in overall locomotor tolerance, further implicating its plausible neuroprotective effects (Xu et al. 2010). It is hypothesized that the neuroprotective effects of caffeine are due to the fact that it is an adenosine receptor antagonist - this is defined as a drug capable of attaching to a specific receptor and subsequently preventing the binding of any other molecule to that specific receptor site (Julien 2004; Wirdefeldt et al. 2011). It is thought that the adenosine receptor anagonism is the most likely way in which caffeine promotes neuroprotection and that this mechanism of action may be further mediated by the attenuation of neuronal glutamate release as the adenosine receptor antagonists facilitate excitotoxic glutamate release in the CNS (Xu et al. 2010; Douna et al. 2012).

1.6 Genetic Aetiology of PD

Over the past 15 years, advancements in molecular genetics have provided significant insight into the understanding of PD (Luecking et al. 2000). The theory that PD is largely sporadic and without a genetic aetiology has been discarded and genetic studies have strengthened the hypothesis that the disease has a significant genetic component (Polymeropoulos et al. 1996; Luecking et al. 2000). Mutations within six genes namely *SNCA*, *LRRK2*, *Parkin*, *DJ-1*, *PINK1* and *ATP13A2*, have been conclusively identified in cases of autosomal dominant and autosomal recessive PD (Table 1.2). More recently, two new genes namely *VPS35* and *EIF4G1* have been included in the list of PD causing genes but to date, only one pathogenic mutation has been identified in each of these genes (Vilariño-Güell et al. 2011; Chartier-Harlin et al. 2011).

The aetiology of PD still remains poorly understood, but the continual investigation into the identification of PD genes and development of new animal models have provided significant insight into the pathobiology of the disease (Moore et al. 2005). Advances in molecular research have provided evidence that mitochondrial dysfunction *(PINK1*, *DJ-1* and *Parkin)*, dysfunctions in the ubiquitin-proteasome system and the autophagy-lysosomal pathway *(Parkin*, *ATP13A2* and *SNCA)* may result in PD pathogenesis. Alterations in the kinase signalling pathways *(PINK1* and *LRRK2)* are additional pathways that have been implicated in the disease (Moore et al. 2005; Brooks 2010).

The known PD genes screened in this study will be discussed in more detail in Chapters Two to Five.

Table 1.2 List of the genes involved in familial PD.

Gene	Locus	Form of PD	Inheritance	AAO (years)	Mutations	Protein	Protein Function
SNCA	PARK1/4	Early onset	AD	20 – 85	A30P; E46K; A53T; gene duplications and tripliactions	Alpha synuclein	Plays a role in synaptic vesicle recycling, compartmentalization of neurotransmitters
Parkin	PARK2	Juvenile and early onset	AR	16 – 72	Various point mutations; exonic rearrangements	Parkin	Cell signalling, protein clearance and degradation
PINK1	PARK6	Early onset	AR	20 – 40	Various point mutations; rare, large deletions	PTEN putative induced kinase	Unknown; possible role in mitochondrial protection during oxidative stress
DJ – 1	PARK7	Early onset	AR	20 – 40	Point mutations; large deletions	Oncogene DJ-1	Unknown; possible role in cellular protection against oxidative stress
LRRK2	PARK8	Late onset	AD	32 – 79	7 point mutations	Leucine rich repeat kinase 2	Cellular and protein interactions and cell signalling
ATP13A2	PARK9	Juvenile and early onset	AR	11 – 16	Point mutations	P5 subfamily of ATPases	Unknown; cellular cation homeostasis and maintenance of neuronal integrity and possible lysosomal functioning

AD= autosomal dominant, AR = autosomal recessive; AAO= Age at onset

As described previously, PD can only be diagnosed with certainty at autopsy. For this reason, it is important to exclude other neurodegenerative disorders which present with similar overlapping clinical feartures or Parkinsonism. Some disorders share overlapping clinical features with PD particularly during the early stages of the disease and some of these diseases include Huntington's Disease-like 2 (HDL2), spinocerebellar ataxia (SCA) and fragile X Tremor Ataxia Syndrome (Table 1.3) (Klein, Schneider, and Lang 2009). HDL2 is a rare neurodegenerative disorder that has been reported in very few patients and is a result of triplet repeat expansions in the *JPH3* gene (Margolis 2001; Margolis et al. 2004). This neurodegenerative disorder is characterized by selective neuronal degeneration, dementia and movement abnormalities (Bardien et al. 2007). This disorder is discussed in more detail in Chapter 7. There are numerous subtypes of SCAs, all of which are characterized specifically by the lack of coordination of walking in the patients – this inability to coordinate walking is also coupled to a poor control of hand and eye movements, as well as speech impairments (Miyai et al. 2011). Due to advancements in genetics, genetic testing can now be used successfully to differentiate between most of the disorders that are listed in Table 1.3.

Table 1.3 Some examples of disorders that can present clinically with a Parkinsonian phenotype.

Condition	Gene Involved	Clinical Phenotype
PARK1	*SNCA*	Parkinsonism associated with the G209A mutation
PARK9/ Kufor Rakeb Syndrome	*ATP13A2*	Kufor Rakeb disease—levodopa-responsive parkinsonism with pyramidal degeneration, supranuclear gaze palsy and dementia
PARK14	*PLA2G6*	Adult-onset dystonia-parkinsonism
DYT3	*TAF1*	X-linked dystonia-parkinsonism / Lubag disease; response to deep brain stimulation surgery
DYT5 (DRD)	*Tyrosine hydroxylase*	Tyrosine hydroxylase deficiency in infancy
DYT12	*ATP1A3*	Rapid-onset dystonia-parkinsonism in an Irish kindred
DYT12	*ATP1A3*	The first Asian case of rapid-onset dystonia parkinsonism
SCA2	*Ataxin 2*	‘‘Apparently sporadic Parkinson’s disease’’
SCA2	*Ataxin 2*	Dopa-responsive parkinsonism
SCA3	*Ataxin 3*	Parkinsonism in Machado-Joseph disease
SCA6	*CACNA1A*	Parkinsonism and reduced nigrostriatal function
NBYA1/ PKAN	*PANK2*	Akinesia in two siblings compound, heterozygote with two missense mutations in the PANK2 gene
FXTAS/ Fragile X	*FMR1*	Fragile X Tremor Ataxia Syndrome
HDL2	*JPH3*	Clinical heterogeneity in a South African

		family
Aceruloplasminemia	*CP*	Parkinsonism
GM1 type 3 gangliosidosis	*GLB1*	Dystonia-parkinsonism
PKU	-	Phenylketonuria presenting as L-Dopa responsive parkinson syndrome

Taken from Klein, Schneider, and Lang 2009.

1.7 PD Molecular Research in South Africa

Extensive investigation into the genetic basis of PD in the European, Japanese and North American populations has been carried out, but little is known about the genetic cause of PD in Africa, particularly in the sub-Saharan African regions (Okubadejo 2008). The drive to identify other genes that may be involved in the development of PD or may act as susceptibility or protective factors, remains important (Farrer 2006; Gasser 2010). It is hypothesized that due to the demographic transition which Africa as a whole is experiencing, the population is likely to become significantly older and therefore the incidence of neurodegenerative disorders such as PD in African countries, will increase (Okubadejo et al. 2006).

Epidemiological data available for PD in Africa to date suggests that it is uncommon, but very few studies have focussed specifically on sub-Saharan African patients and subsequently there is very limited knowledge about the biochemical pathways that are affected in these patients. It has been suggested that PD in black patients may present with different clinical features, making disease identification rather than diagnosis a possible problem (Okubadejo et al. 2006; Haylett et al. 2012).

The PD genetics research group at the Division of Molecular Biology and Human Genetics at Stellenbosch University in Cape Town, South Africa is currently the only group studying the genetic aetiology of PD in South African patients. Results to date (Table 1.4) from the studies conducted by the group have concluded that mutations in the known PD genes have low frequencies in affected patients (Bardien et al. 2009; Keyser et al. 2010; Haylett et al. 2012).

Table 1.4 Mutations previously identified in South African PD patients.

PD genes implicated in autosomal recessive PD					
Gene	**Patient No.**	**Ethnicity**	**AAO**	**Mutation**	**Family History**
Parkin	37.12	Black	45	Heterozygous duplication of exon 2 + heterozygous deletion of exon 9	Yes
	53.44	Mixed Ancestry	27	Homozygous deletion of exons 3+4	No
	56.45	Caucasian (Afrikaner)	27	Homozygous deletion of exon 4	Yes
	77.60	Caucasian	25	Heterozygous P113fsX163 + heterozygous deletion of exon 3	Yes
	78.74	Black	56	Heterozygous G430D + heterozygous deletion of exon 3	No
	78.76	Caucasian	27	Heterozygous G430D + heterozygous deletion of exon 4	No
	81.03	Caucasian	48	Heterozygous duplication of exon 2-6 + heterozygous duplication of exon 5	Yes
PINK1	68.10	Indian	37	Homozygous Y258X	Yes

PD genes implicated in autosomal dominant PD					
LRRK2	60.47	Mixed Ancestry	47	Heterozygous G2019S	Yes
	68.06	Caucasian	42	Heterozygous G2019S	Yes
	81.64	Caucasian	70	Heterozygous G2019S	Yes
	82.57	Caucasian	63	Heterozygous G2019S	No
	84.25	Caucasian	58	Homozygous G2019S	Yes
	85.57	Mixed Ancestry	62	Heterozygous R1441C	Yes
SNCA	42.35	Caucasian	46	Whole gene triplication	Yes
Heterozygous variants identified with unknown pathogenic significance					
Parkin	42.06	Caucasian (Afrikaner)	37	R402C	No
	51.70	Caucasian (Afrikaner)	42	E310D	No
	55.54	Caucasian (Afrikaner)	56	Duplication of exon 2	No
	65.79	Caucasian	55	H200Q	Yes
	68.22	Mixed Ancestry	50	Duplication of exons 2+3	No
	81.27	Mixed Ancestry	49	Heterozygous deletion of exon 4	Yes
	84.47	Mixed Ancestry	61	Deletion of exons 3+4	No
DJ-1	50.31	Mixed Ancestry	38	g.6_+10 del (16bp deletion at transcription start site)	No
LRRK2	68.07	Caucasian (Afrikaner)	50	Q2089R	Yes

1.8 Mutation Detection Approaches used in the Present Study

There are a wide variety of different techniques that have been developed for mutation screening, each with their own advantages and limitations. The methods employed in the present study include high resolution melt, Sanger sequencing and WES in order to identify pathogenic mutations. The principles underlying these methods will be described briefly in this section.

1.8.1 High Resolution Melt

High resolution melt (HRM) is an analytical method in which DNA fragments are distinguished from each other through their melting behaviour. It is an expansion of existing DNA dissociation methods, that allow for the characterization of DNA fragements according to the way in which they dissociate ('melt'). Double stranded DNA (dsDNA) (pre-melt phase) is converted to single stranded (ss) DNA (post-melt phase) as it is subjected to increases in temperatures (figure 1.6). This is analysed or monitored by adding a fluorescent dye (e.g. EvaGreen, Syto 9 and Sybr Green) to the PCR reaction mixture that is allowed to intercalate within the dsDNA of the PCR products. As the strands separate, the dye is released, causing a decrease in fluorescence as the temperature increases. HRM instrumentation collects and analyses fluorescent signals in real time, thereby characterizing the different DNA fragments.

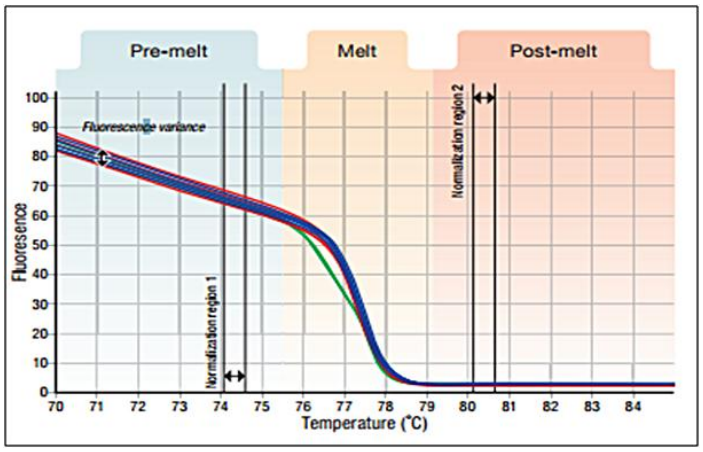


Figure 1.6 Illustration of the principle underlying high resolution melt. With an increase in temperature double stranded (ds) DNA melts and becomes single stranded (ss) DNA. As the melt progresses, an intercalating dye is released - the fluorescence produced is used to create a thermal denaturation profile which is unique for each DNA sequence. As the temperature increases, more of the dsDNA is converted to ssDNA. Fluorescence is plotted against temperature (Taken from Introduction to HRM Analysis http://www.kapabiosystems.com/public/pdfs/kapa-hrm-fast-pcr-kits/Introduction_to_High_Resolution_Melt_Analysis_Guide.pdf).

A thermal denaturation profile can be constructed in which fluorescence is plotted against the temperature; this profile is specific for the PCR product as well as the length of the sequence, base and GC content (Ye et al. 2010). Alterations in the nucleotide sequence will affect the way in which the fragment melts. This will allow fragments with nucleotide sequence alterations to be identified when they are compared to the wild type sample; these can then be sequenced in order to characterize the sequence variant (Ye et al. 2010). HRM data can be analysed as either Normalized Graphs (figure 1.7) or as Difference Graphs (figure 1.8).

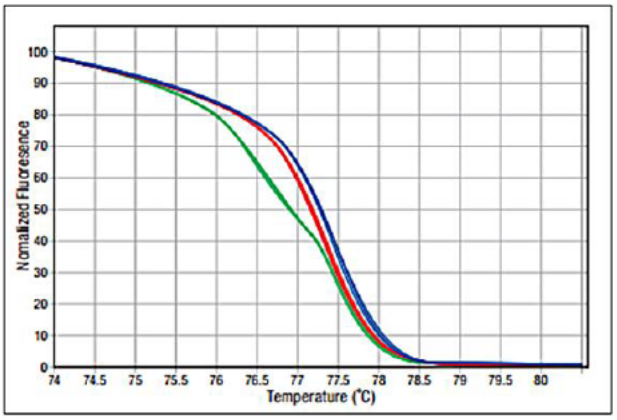


Figure 1.7 Example of a HRM Normalized graph. Temperature range in a specific region is selected to allow for identification of variation between different wild type and mutant samples respectively. Blue and red lines are homozygous wild type and homozygous mutant respecively; green line is a heterozygous mutant sample (Taken from Introduction to HRM Analysis http://www.kapabiosystems.com/public/pdfs/kapa-hrm-fast-pcr-kits/Introduction_to_High_Resolution_Melt_Analysis_Guide.pdf).

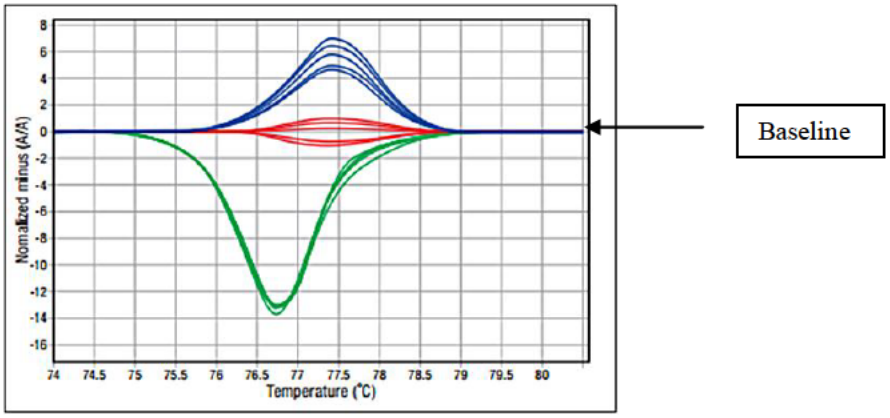


Figure 1.8 Example of a HRM Difference graph. This graph is used when a particular genotype is identified and used as a reference/baseline for the other DNA samples. The position of each sample relative to the reference is plotted against the temperature thereby showing differences between various samples. The homozygous wild-type is used as a reference (red) and the other samples are compared to it. Blue is the homozygous mutant and green is the heterozygous mutant (Taken from Introduction to HRM Analysis http://www.kapabiosystems.com/public/pdfs/kapa-hrm-fast-pcr-kits/Introduction_to_High_Resolution_Melt_Analysis_Guide.pdf).

HRM is a relatively simple and cost effective means to screen patients for known mutations. It is also a means of identifying novel variants and is sufficiently sensitive to identify single base pair changes. An additional advantage is that it is a closed-assay system, and no post PCR processing is therefore necessary.

1.8.2 Sanger Sequencing

Samples with altered denaturation profiles identified though the use of HRM were sequenced using direct sequencing to characterize the sequence variant. This was performed using the Sanger Sequencing method, otherwise known as the 'Chain terminator sequencing method'. In this technique, the DNA is denatured into single strands, and primers that are specific to the sequence are added and annealing takes place. The DNA fragments are truncated with di-deoxynucleotides (ddNTPs) that are added to the reaction mix in addition to the 'normal' dNTPs. The addition of the ddNTPs allows for the generation of DNA fragments that have been terminated at locations where these nucleotides are inserted as they prevent the addition of further nucleotides (Cawley 2005; Metzenberg 2008). The resulting fragments, varying by 1bp each, are then separated in glass capillaries which are filled with a viscous polymer and analysed using specialized software (Cawley 2005; Metzenberg 2008).

1.8.3 Whole Exome Sequencing

The human genome is composed of an estimated 3 billion nucleotides, but only approximately 1.22% of these are thought to form part of the protein coding regions, namely the exons (Ng et al. 2010). Whole exome sequencig (WES) is a relatively new molecular technique which allows for the identification of novel genes and disease causing mutations by sequencing only the exons; it is estimated that there is a total of approximately 180 000 exons in the human genome (Teer and Mullikin 2010). Essentially, the exome therefore represents a highly enriched subset of the genome in which to search for pathogenic mutations with large effect sizes.

There are several basic steps which are needed before the actual sequencing of the exons can be carried out. Genomic DNA is used as the input DNA and is used to construct a shotgun library; the DNA is randomly sheared and the fragments that are generated by this shearing process are flanked by specific adaptors. Library enrichment follows through a process known as aqueous-phase hybridization capture; this allows for fragments of sheared DNA,

which are complimentary to the exons to hybridize to biotinylated DNA in the presence of so-called 'blocking oligonucleotides' which are complimentary to the adaptors. Hybridized fragments are then isolated through the use of biotin-streptividin-based pull-down and this allows the target DNA to be captured, amplified and sequenced through the use of massively parallel sequencing approaches. This is then followed by the mapping, alignment and identification of sequence variants (figure 1.9) (Bamshad et al. 2011).

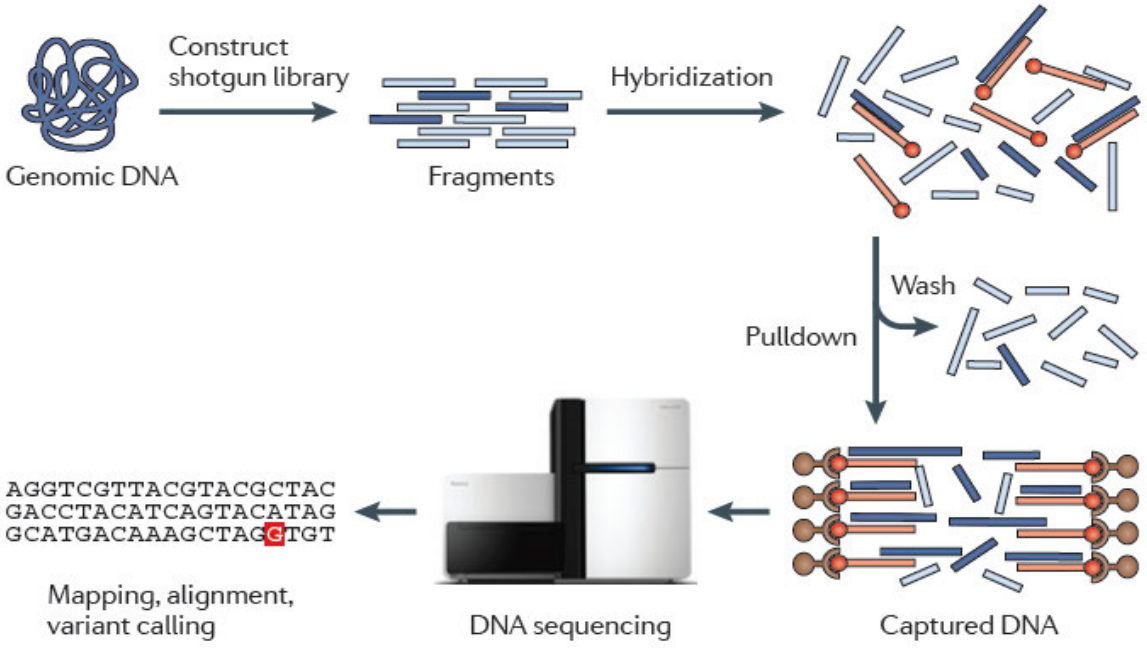


Figure 1.9 Basic workflow of whole exome sequencing (WES) (Taken from Bamshad et al. 2011)**.**

One of the biggest challenges currently faced regarding the use of WES as a method to identify novel disease-causing mutations, is the definition of the exome. Although significant advancements have been made, a considerable amount of uncertainty regarding which regions of the human genome are truly protein-coding regions still exists (Bamshad et al 2011). The initial approach when selecting the target regions for WES was a very conservative one, with only a 'high confidence' subset of genes being identified as targets and thus studied further - most of the exons have been annotated in the consensus CDS (CCDS) (http://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi) (Bamshad et al. 2011). As the capacity of sequencing has increased, so has the target range, with a number of proteins with a postulated function now being included in the target regions.

Three main vendors (Agilent, Illumina and NimbleGen) provide commercial kits that target exons obtained from the CCDS, as well as the additional hypothetical proteins. The

differences in these commercial kits reinforce the drawback of not being able to fully define the exome; Agilent provides a SureSelect Human All Exon 50 Mb Kit™, which targets exons annotated by the GENCODE (http://www.sanger.ac.uk/gencode/) project, CCDS and includes 10 base pairs of flanking sequence for each region. Non-coding RNAs are also included and these are obtained from miRBase and Rfam databases respectively - this kit also captures a large portion of the putative exome, with up to 50Mb being captured per single sequencing reaction (http://www.genomics.agilent.com/CollectionSubpage.aspx?PageType=Product&SubPageType=ProductData&PageID=2318).

Illumina provides the TruSeq™ Exome Enrichment Kit that allows for a greater region of the exome to be sequenced as 64Mb of the human genome can successfully be targeted. This kit covers exons defined in the CCDS, RefSeq (coding exons and exons plus short intronic sequences) (http://www.ncbi.nlm.nih.gov/RefSeq/), Encode/Gencode coding exons as well as predicted microRNA targets (http://www.microrna.org/microrna/getDown-loads.do). This kit therefore provides a broad coverage of all of the exons in the databases, but provides insight into non-coding DNA in the exon flanking regions such as the promoter regions as well as 3'UTRs (http://www.illumina.com/documents/products/datasheets/datasheet_truseq_exome_enrichment_kit.pdf).

The Roche NimbleGen SeqCap EZ Exome v3™ is another commercially available kit that can be used for exome capture. This kit also covers 64Mb of the human genome and makes use of six different databases to collect the information needed for the coding portions of the genome. It makes use of Gencode (http://www.gencodegenes.org/) and miRBASE (http://www.mirbase.org/) as two additional databases for exome identification whereas all the other kits only make use of dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/), Vega (http://vega.sanger.ac.uk/index.html), Ensembl (www.ensembl.org), RefSeq (http://www.ncbi.nlm.nih.gov/RefSeq/) and CCDS. This kit also targets microRNAs. For exons which are smaller than 100bp in length, this kit lengthens the reads to more than 100bp, thus covering a greater region of the genome (http://www.nimblegen.com/products/seqcap/ez/v3/index.html).

The development of commercial kits has not, however, come without a number of restrictions. The biggest and most significant issue encountered is the fact that unannotated genes or exons are missed. Additionally, the efficiency of the capture probes varies significantly and for this reason, some regions (especially those that are GC rich) may not be targeted with the same efficiency as others and certain regions may not be sequenced at all. Another cause for concern is that not all target fragments are amplified with identical efficiency thus making alignment and variant calling problematic.

However, despite the stumbling blocks that are encountered when using WES, this strategy provides a powerful means to identify possibly novel disease-causing mutations (Bilgüvar et al. 2010; Bamshad et al. 2011). The feasibility and plausibility of this approach as a new and successful method to identify novel candidate disease-causing mutations was demonstrated when a novel gene was identified and successfully implicated in Miller syndrome, a rare Mendelian disorder (Ng et al. 2010). This study was the first documentation of the use of WES as a means to identify novel genes in rare disorders; Ng and colleagues performed WES on four patients (two of the affected patients were siblings and the other two affected patients were unrelated). The exons, as well as all of the splice donor and acceptor sites were sequenced; the splice donor and acceptor sites were included as many variations in these regions have been linked to Mendelian disorders (Kobelka 2010). Synonymous variants were assumed to be non-pathogenic and in this study, only non-synonymous coding variants and variants identified in splice donor and acceptor sites were prioritized (Ng et al. 2010). These variants were then compared to various databases and a list of novel candidate genes subsequently identified. The novel variants identified in both siblings were then compared to those variants identified in the other two affected patients and based on the finding of overlapping novel genes among all four affected patients, a novel candidate disease-causing gene, *DHODH* was identified (Kobelka 2010) (figure 1.10).

The identification of the novel disease-causing gene identified by Ng and colleagues provided the first significant evidence that WES is a powerful tool for studying inherited disorders. In doing so, a framework for the identification of rare mutations can be developed and importantly, this approach can be used for small families not amenable to traditional approaches such as linkage analysis (Ng et al. 2010; Kobelka 2010).

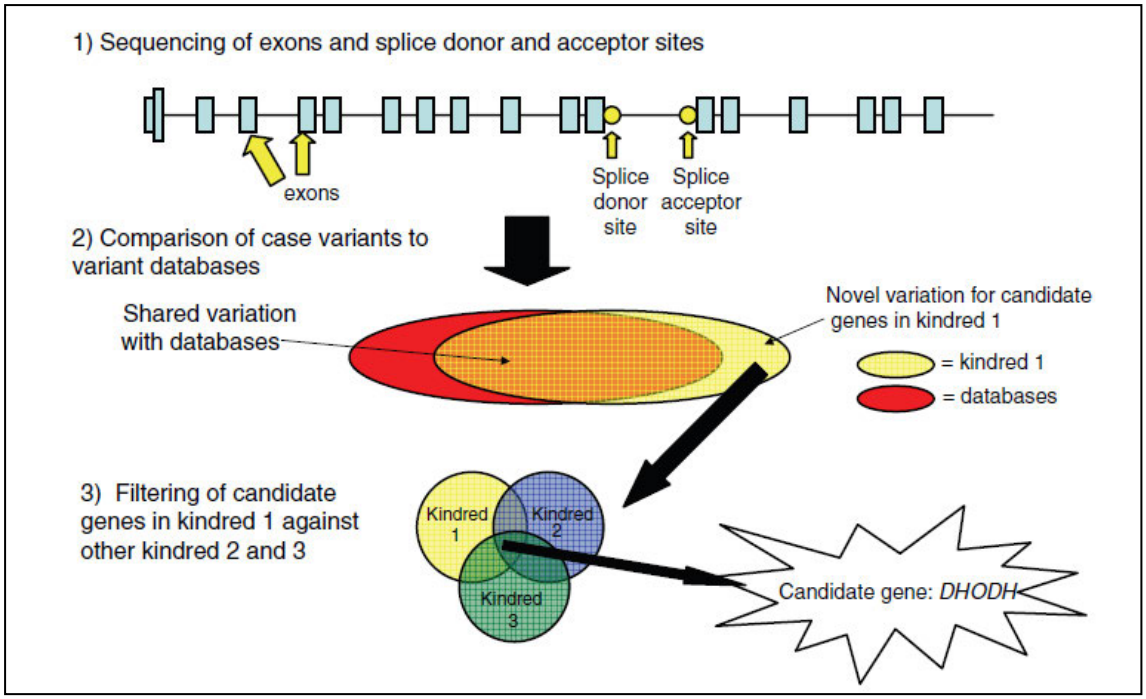


Figure 1.10 Basic filtering approach used to narrow down WES results to identify disease-causing genes. Kindred 1 is the affected sibling pair and kindreds 2 and 3 are unrelated, affected patients. Based on the overlapping novel genes in all four patients, the novel candidate gene was identified as *DHODH*.

1.9 The Present Study

There is a need to focus specifically on South African PD patients as they differ significantly in ancestral origins when compared to the rest of the world; and it is therefore likely that these patients may harbour novel mutations (Okubadejo et al. 2008; Keyser et al. 2010; Haylett et al. 2012). The identification of disease-causing mutations in patients is important in families as this will facilitate better clinical management of their families and high risk individuals can be identified.

For the present study a total of 262 South African PD patients, from diverse ethnic groups, were recruited for genetic analysis (Table 1.5). For the puropses of our study, the English-speaking Caucasians and Afrikaner Caucasians were analysed seperately as the Afrikaner population is unique to South Africa and due to the high number of reported PD cases in this this specific ethnic group, significant insight into PD may be gained through the independant study of these patients.

The various ethnic groups can be defined as follows:

- The English-speaking Caucasian population is comprised of individuals of European descent.
- The Afrikaner Caucasian population is unique to South African and is comprised of Afrikaans speaking individuals. These individuals are mainly of Dutch decent, but may also have German and French ancestry (Greeff 2007).
- The Mixed Ancestry population is defined as an admixture or a combination of various ethnic groups. These various combinations include immigrants from Western Europe, India, Malaysia and Madagascar as well as combination of ethnic groups which are indigenous to South Africa, such as San and Khoi – Khoi (Patterson et al. 2010).
- The Black African population is comprised of individuals whose ancestry can be directly traced to the African continent. This ethnic group is comprised of individuals who speak traditional African languages such as Zulu, Xhosa, Ndebele, Tsonga, Venda, Swazi, Northern Sotho, Tswana and Sesotho.
- The Indian population is comprised of individuals who migrated from colonial India to the African continent in the early 19th century.

Table 1.5 Ethnic breakdown of the 262 South African PD patients participating in the study

Ethnic group	n (% of 262)	Pos family history n (%) of each group	Average AAO (range) in years
English-speaking Caucasian	101 (38.6)	30 (29.7)	57.8 (25-80)
Afrikaner Caucasian	76 (29.0)	22 (29.3)	51.8 (17-76)
Mixed Ancestry	64 (24.4)	10 (15.6)	51.5 (20-80)
Black African	17 (6.5)	3 (17.6)	54 (30-74)
Indian	4 (1.5)	1 (25.0)	50 (35-68)

Pos, positive; AAO, age at onset of the disorder

1.9.1 Aims and Objectives

Our hypothesis was that South African PD patients harbour novel disease-causing mutations due to their unique ancestry. The aim of the present study was therefore to identify and characterise the disease-causing mutations in 262 South African patients with PD.

Five of the known PD genes, namely *Parkin* (all exons), *PINK1* (all exons), *LRRK2* (selected exons), *SNCA* (selected exons) and *DJ-1* (only one variant) were screened for mutations. *ATP13A2* was not included because patients with mutations in this gene have a very unusual phenotype, which is juvenile onset usually under 20 years and atypical PD symptoms such as behavioural problems, dementia, facial tremor and pyramidal tract dysfunction (Vilariño-Güell et al. 2008) and none of our patients presented with these specific symptoms. *EIF4G1* and *VPS35* were not included as to date, only one pathogenic mutation has been identified in each of these genes (Chartier-Harlin et al. 2011b; Zimprich et al. 2011). Mutations in *DJ-1* are very rare worldwide; accounting for <1% of all PD cases, so the exons of this gene were not screened. Instead, we screened for the presence of a functional variant, a 16bp indel (g.-6_+10del) in the 5'UTR of *DJ-1* that is thought to influence transcription (Keyser et al. 2009). Furthermore, we investigated triplet repeat expansions at the HDL2 locus in the *JPH3* gene, in only the Black PD patients. Of the 262 study participants, varying numbers of patients had previously been screened for mutations by other members of our research group. Therefore, in the present study subsets of the 262 patients were screened for different loci in an attempt to ensure that all the patients had been screened for these genes.

The objectives of the study were as follows:

1. To screen four of the known PD genes (*PARK2*, *SNCA*, *LRRK2*, and *PINK1*) for disease-causing mutations using HRM and Sanger sequencing.
2. To screen for the presence of the 16bp indel polymorphism in the *DJ-1* gene.
3. To determine whether Black patients diagnosed with PD harbour pathogenic repeat expansions at the HDL2 locus.
4. To perform WES on an extended Afrikaner kindred to identify a possible novel disease-causing mutation.

Chapter 2: Mutation Screening of the *Parkin* gene

Page

2.1 Introduction

Parkin (PARK2) (OMIM #600116) encodes a 465 amino acid protein (Luecking et al. 2000) that belongs to a family of E3 ubiquitin ligase proteins, which are also known as the in-between ring (IBR) family (Beasley, Hristova, and Shaw 2007). The five protein domains of *Parkin* are shown in figure 2.1. Parkin plays a pivotal role in the protein degradation part of the ubiquitin proteasomal system (UPS) by tagging proteins with ubiquitin (Beasley, Hristova, and Shaw 2007).

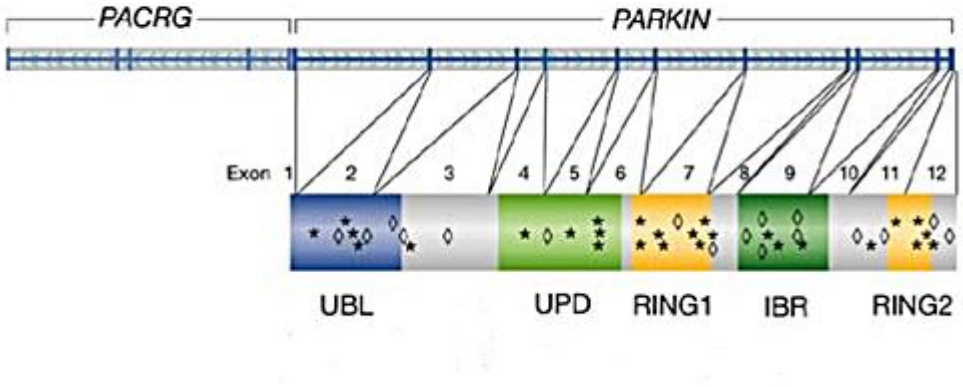


Figure 2.1 Protein domains of Parkin. The numbers indicate the exons. The five domains are N-terminal ubiquitin-like domain (UBL),a cysteine-rich unique parkin domain, two C-terminal RING (really interesting new gene) domains that are separated by an in-between-RING domain (IBR). The positions of the nonsense and frameshift mutations are indicated by the diamonds and the locations of the disease-associated mutations are indicated by the asterisks (Taken from Kahle and Haass 2004).

Parkin is located on chromosome 6, specifically at the 6q26 location and has 12 exons. The role of *Parkin* in PD was first described in 1998, when exonic deletions were identified through the use of linkage analysis in Japanese families who presented with autosomal recessive, early onset PD (Kitada et al. 1998). Mutations in *Parkin* have been implicated in approximately 20% all of early onset familial cases with autosomal recessive inheritance of PD (Wider and Wszolek 2007) as opposed to the 10% for other genes that have been documented as causal factors for early onset PD (Luecking et al. 2000; Wider and Wszolek 2007).

More than 150 mutations have been identified in the *Parkin* gene and these include point mutations and exonic rearrangements. Exonic rearrangements are thought to be due to the large size of the introns, resulting in errors in mRNA splicing ultimately leading to the

incorporation of insertions and deletions (Abbas et al. 1999; Gasser 2001; Gasser 2010). Documentations of *Parkin* mutations have shown that the mutations occur throughout the gene (figure 2.1) (Bonifati 2007; Gasser 2010). *Parkin* mutations are thought to possibly result in pathological changes which cause the degeneration of the substantia nigra as well as changes in the locus coeruleus. The changes in the locus coeruleus are considered to be milder and less significant than those in the substantia nigra (Gasser 2010). LBs were initially thought to be uncommon in patients with *Parkin* mutations (Shimura et al. 2000; Huynh et al. 2001). Recent studies using improved analytical methods have, however identified LB-positive PD patients with *Parkin* mutations, although they are considered to be uncommon (Farrer et al. 2001; Pramstaller et al. 2005). A noteworthy feature of patients who carry *Parkin* mutations is that the disease progression appears to be slower than in other cases of PD. Interestingly, when levodopa is administered as a management strategy, these patients show a greater responsiveness to the drug, at significantly lower doses without the additional treatment complications such as dyskinesias (Lohmann et al. 2003).

2.1.1 Expression Profile

The protein product of *Parkin* is principally a cytosolic protein and has been shown to co-localise with actin filaments providing evidence that it may be a cytoskeletal–associated protein (Huynh et al. 2001; Cookson 2003; Wang et al. 2005). Parkin is also expressed in the cell bodies of neurons in the midbrain, cerebellum, cerebral cortex as well as the basal ganglia (Shimura et al. 2000). It appears to be widely spread throughout many of the human tissues but significant expression can be seen in the brain (Kitada et al. 1998) (figure 2.2). Parkin has also been identified as a protein that is present in the synaptic vesicles, endoplasmic reticulum and the outer membrane of the mitochondria (Darios et al. 2003).

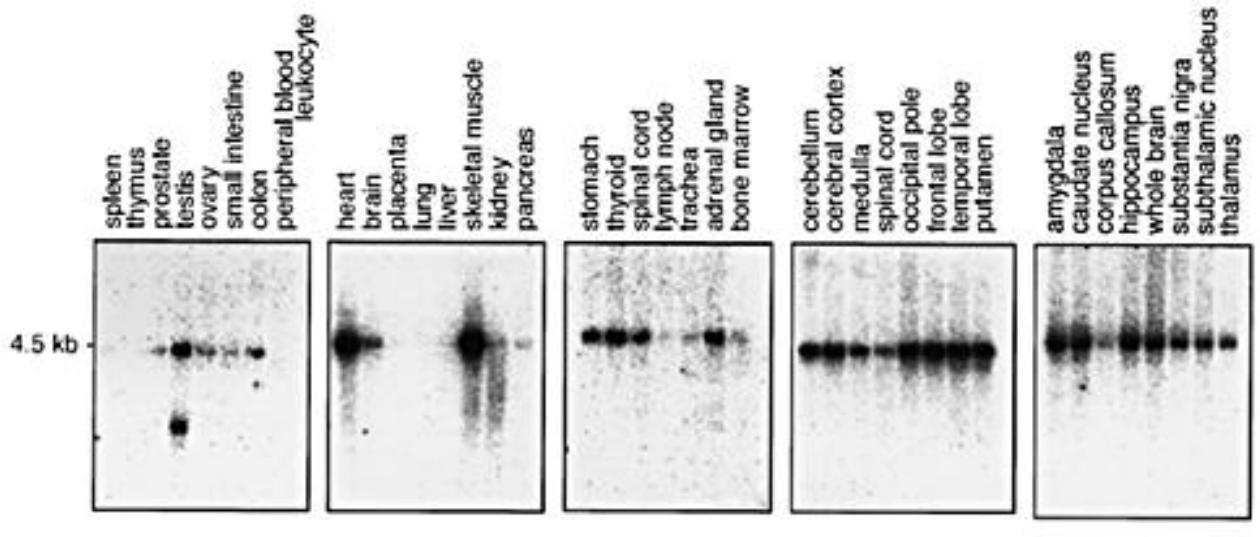


Figure 2.2 Northern blot analysis of Parkin expression. A Northern blot showing the expression of Parkin in the human body. It is expressed in nearly all anatomical tissues but significant expression can be seen in the cranial regions (Taken from Kitada et al. 1998).

2.1.2 Postulated Function

Parkin plays a pivotal role in the UPS, the predominant mechanism by which misfolded, damaged and short-lived proteins are removed from the cell. The UPS regulates protein turnover, therefore resulting in a dynamic proteome (Hershko and Ciechanover 1998). As mentioned before, Parkin belongs to the E3 ubiquitin ligase family due to the fact that it has an IBR domain. This domain is significant as it is the region that interacts with the ubiquitin–conjugating enzymes (E2) and catalyses the attachment of ubiquitin molecules to specific protein targets (Moore et al. 2005) (figure 2.3). This process allows for 'ubiquitin tagging' to take place in order to specify the destruction of specific proteins by the proteasome (Shimura et al. 2000). Ubiquitination results from the consecutive actions of the ubiquitin activating E1, E2 and E3 enzymes. Subsequent cycles of ubiquitination result in the formation of a poly–ubiquitin chain that can then be recognised by the 26S proteasome (Moore et al. 2005). E3 ubiquitin ligases provide substrate specificity to the ubiquitination process as each ligase binds to specific subsets of proteins. Defects in Parkin may thus interfere with the proteolytic pathway that could lead to the deleterious accumulation of particular proteins which in turn may contribute to the death of nigral neurons (Matsumine et al. 1997; Kitada et al. 1998). The tagging of proteins with ubiquitin may also occur for processes that are proteosome-independant: some of these roles include signal transduction and protein trafficking (Kahle and Haass 2004). Additionally, it was determined that Parkin is associated with

mitochondrial DNA in a neuroblastoma cell line as well as in cells that are undergoing proliferation (Rothfuss et al. 2009). The conclusions reached through various studies were that Parkin protects the mitochondrial DNA from oxidative damage and may act to stimulate mitochondrial repair (Rothfuss et al. 2009; da Costa et al. 2009). It has been reported that Parkin acts together with PINK1 in a pathway which promotes the maintenance of mitochondrial functioning and integrity (Rothfuss et al. 2009).

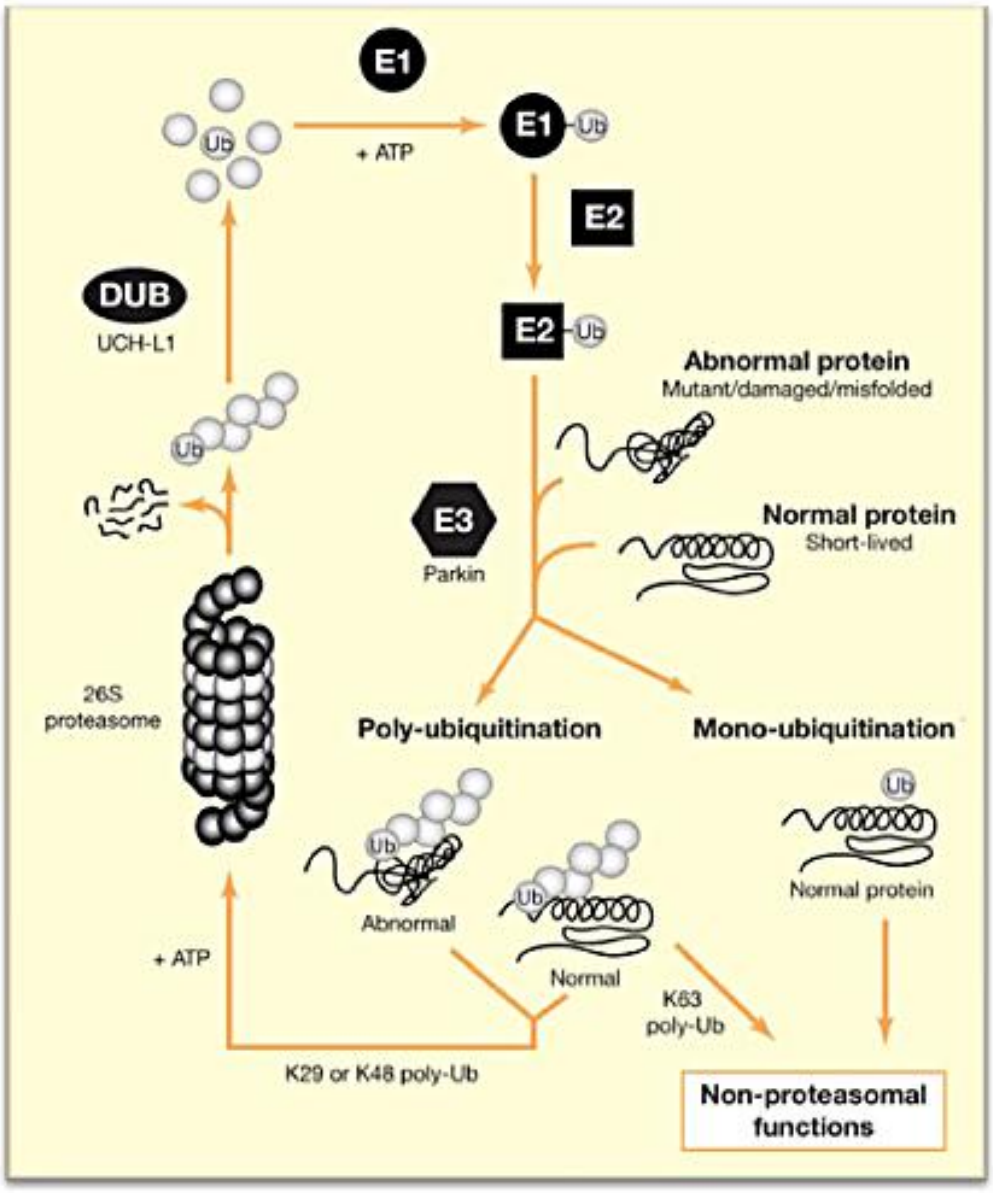


Figure 2.3 The ubiquitin-proteasome system (UPS) and ubiquitination. Parkin (an E3 protein) plays a specific role in this system as is mediates the transfer of ubiquitin monomers to E2. Mutations in *Parkin* are therefore thought to play a significant role in the UPS and may thus alter the functionality of the pathway due to protein mishandling. E1, ubiqitin activating enzyme; E2, ubiquitin conjugating enzyme; E3, Parkin; DUB, deubiquitinating enzyme; Ub, ubiquitin (Taken from Glickman and Ciechanover 2002).

2.1.3 Disease-causing Mutations and Susceptibility Alleles

2.1.3.1 Disease-causing Mutations

Mutations found in this gene include nonsense and missense mutations, as well as whole exon rearrangements, deletions and duplications (Hattori et al. 2004). Five of the most common mutations in *Parkin* are (1) deletions of exon 4; (2) various single base pair deletions in exon 2; (3) deletions of exons 3 and 4; (4) deletions of exon 3; (5) point mutations in exon 7 (Abbas et al. 1999; Kann 2002; Klein et al. 2003). These five variants account for approximately 35% of all *Parkin* mutations (Klein et al. 2003). Collectively, mutations in *Parkin* have been associated with 50% of early onset PD in some studies, and the mutation frequency decreases significantly with an increase in the age at onset (Kann 2002). In one study, 73 families who presented with PD before the age of 45 years were examined. It was determined that 36 of these patients had *Parkin* mutations – approximately 49% of the total patients (Wirdefeldt et al. 2011). Notably, 77% of the patients who were younger than 20 years of age when they were diagnosed with PD, had *Parkin* mutations. Further examination of the study participants revealed that the need for a molecular diagnosis for *Parkin*-associated PD is essential and that clinical manifestations alone are not sufficient for accurate and definitive diagnosis of the disease in patients (Luecking et al. 2000).

2.1.3.2 Susceptibility Alleles

A susceptibility allele increases an individual's risk of developing a particular disorder. *Parkin* is highly polymorphic and the current hypothesis is that polymorphisms in this gene may result in an increase in susceptibility to develop PD – whether early onset or late onset PD (Oliveira 2003). One study identified eighteen different polymorphisms in the patients that were studied, of which four were identified as novel and therefore plausible susceptibility alleles (Sun et al. 2006). However, larger scale studies are necessary in order to determine the role of heterozygous *Parkin* mutations and the determination of whether or not they act as susceptibility factors for PD.

2.1.4 The Present Study

The present study investigated the role of *Parkin* in South African PD patients and all 12 exons were screened using HRM and sequencing.

2.2 Materials and Methods

2.2.1 Study Participants

Ethics approval was obtained from the Committe for Human Research at Stellenbosch University, Cape Town (Protocol number: 2002/C059). A total of 262 PD patients had been recruited from the Movement Disorders Clinic at Tygerberg Hospital in Cape Town, as well as from the Parkinson's Association of South Africa. The patients were diagnosed according to the UK Brain Bank Diagnostic criterion which requires that the patients present with bradykinesia plus at least one of the following symptoms: resting tremor, rigidity and postural instability (Gibb and Lees 1988). All study participants met the criteria. The cohort included 161 (61.5%) male and 101 (38.5%) female patients. The average age at onset (AAO) of the patients was 54.3 years of age. The standard deviation (SD) is 12.46 years and the range of the AAO falls between 17 and 80. A total of 35% of these patients reported a positive family history while 65% could either not provide any information regarding possible family history, or had no known reported history of PD.

Written, informed consent was obtained from each of the patients and a blood sample was taken in order to obtain a DNA sample for the genetic analysis. A total of 132 control samples from each of the different ethnic groups namely Afrikaner Caucasian, English-speaking Caucasian, Mixed Ancestry and Black were additionally recruited from the Western Province Blood Transfusion Services and these patients were not examined for PD, but were used as a means to assess the frequency of specific sequence variants in each ethnic group.

Of the 262 patients, 229 had already been screened for mutations in *Parkin* in a previous study by our group (Haylett et al. 2012) and therefore in the present study the remaining 33 patients were investigated.

2.2.2 Genetic Analysis

Peripheral blood samples were collected from each patient and genomic DNA was extracted using the phenol-chloroform method (Appendix I).

Specific PCR primer sequences had been designed using Primer 3 software (Rozen 2000) and the sequences are provided in Table 2.1. A total volume of 25µl reaction mix was set up for each sample. The reaction mixture consisted of the following reagents: 20µM of each of the forward and reverse primers; 3.0µM $MgCl_2$ (Bioline, UK); 1x NH_4 buffer (Bioline, UK);

2µM SYTO9 fluorescent dye (Invitrogen, USA); 0.25 units BIOTAQ DNA polymerase (Bioline, UK) and 10ng template DNA. The reactions were set up using the *epMotion™ 5070* (Brinkmann Instruments, Canada) that allows for automated pipetting and preparation of PCR reactions.

All 12 exons of *Parkin* were PCR amplified and screened using HRM. The real-time PCR and HRM analysis was set up and carried out on a RotorGene 6000 instrument (Corbett Life Science, Australia) with the following cycling conditions: an initial step at 94°C for 5 min; 40 cycles with conditions of denaturation at 94°C for 15 s, varying annealing temperatures (Table 2.1) for 15 s and extension at 72°C for 20 s. Thereafter, two additional holding steps were included: 94°C for 1 min to allow for complete denaturation of the double stranded DNA and then at 50°C for 1 min to allow for renaturation of the DNA. HRM analysis was performed with melt temperatures ranging from 75°C to 95°C with the temperature increasing by 0.1°C increments at each step. A wild-type (WT) reference sample was included in every run and samples with altered HRM profiles were selected and Sanger sequenced in order to identify the sequence variant. Known variants in each exon were included as positive controls.

Samples exhibiting altered denaturation profiles were first cleaned by adding 0.5 units of Exonuclease I (Promega, USA) and SAP (Shrimp Alkaline Phosphatase) (Cleveland, Ohio, USA) to 8.0µl PCR product. This cocktail was then incubated at an initial temperature of 37°C for 15 min followed by 15 min at 80°C (to denature the enzymes) in a 2720 Thermal Cycler (Applied Biosystems, Foster City, USA). Exonuclease I was used to remove any excess primer or primer dimers and SAP served to remove excess dNTPs.

Sequencing was performed at the Central Analytical Facility (CAF), Stellenbosch University using the Sanger Sequencing Method, that made use of the BigDye Terminator Sequence Ready Reaction kit version 3.1 (Applied Biosystems, Foster City, USA). The reactions were then electrophoresed and analysed on a 3130 x1 Genetic Analyser (Applied Biosystems, Foster City, USA). Sequencing electropherograms were analysed using BioEdit version 7.0.1 (Hall 1999).

Table 2.1 Primers designed for the exonic sequences of the *PARK2* gene.

Exon	Primer Sequence (5´-3´)	%GC	Tm (°C)	PCR conditions (Ta in °C)	Size of PCR fragment (bp)
1	For: gaa cta cga ctc cca gca g	58	57.3	55	300
	Rev: ccc gtc att gac agt tgg	56	55.7		
2	For: cac cat ttaa ggg ctt cga g	50	56.5	55	313
	Rev:tca ggc atg aat gtc aga ttg	43	56.3		
3	For: tct cgc att tca tgt ttg aca	39	56.5	55	364
	Rev: gca gac tgc act aaa caa aca	43	57.3		
4	For: gct ttt aaa gag ttt ctt gtc	33	51.3	55	299
	Rev: ttt ctt ttc aaa gag ggg tga	38	55.2		
5	For: gga aac atg tct taa gga gt	40	52.5	55	223
	Rev: ttc ctg gca aac agt gaa ga	45	57.3		
6	For: cca aag aga ttg ttt act gtg	38	52.8	55	276
	Rev: ggg gga gtg atg cta ttt tt	45	55.3		
7	For: cct cca gga tta cag aaa ttg	43	54.0	55	280
	Rev: gtt ctt ctg ttc ttc att agc	38	52.6		
8	For: ggc aac act ggc agt tga ta	50	58.5	55	232
	Rev: ggg gag ccc aaa ctg tct	61	58.8		
9	For: tcc cat gca ctg tag ctc ct	55	60.2	55	297
	Rev: cca gcc cat gtg caa aag c	58	60.4		
10	For: caa gcc aga gga atg aat at	40	52.7	53	272
	Rev: gga act ctc cat gac ctc aaa	48	57.3		
11	For: ccg acg tac agg gaa cat aaa	48	57.8	55	300
	Rev: gga acc ttc aga cag cat at	55	54.8		
12	For: tct agg cta gcg tgc tgg tt	55	61.0	55	296
	Rev: gcg ttg tgt gtg tgt ttg a	47	57.3		

Tm, melting temperature; Ta, annealing temperature

2.3 Results

Through the use of HRM and direct sequencing, a number of polymorphisms were identified in the patients (Table 2.2). The majority of these polymorphisms have already been reported. A novel heterozygous variant of unknown pathogenic significance, T387P was identified in one of the Mixed Ancestry patients (figure 2.4).

Table 2.2 Sequence variants found in the *PARK2* gene in 33 South African patients.

Patient	Ethnicity	Age at Onset	Family History	Variant	Exon	In dbSNP	Frequency from dbSNP (n=number of chromosomes)
89.02	MA	54	No	P37P	2	rs143477190	C, 0.993; T, 0.007; n= 2276
88.99	MA	30	No	A46T	2	rs75860381	C, 0.958; T, 0.042; n=118
				S167N	4	rs1801474	G, 0.919; A,0.082; n= 4550
				L261L	7	rs143902760	G, 1.00; A, 0.00; n=4550
				V380L	10	rs1801582	G, 0.682; C 0.318; n=4550
92.60	Afrikaner	45	Yes	R42C	2	rs149699346	G, 1.00; A, 0.00; n=4552
90.94	MA	40	No	P153R	4	rs55654276	C, 0.989; G, 0.011; n=4450
				M192L	5	rs9456735	T, 0.941; G, 0.059; n=286
				L261L	7	rs143902760	G, 1.00; A, 0.00; n=4550
90.86	MA	70	Yes	M192L	5	rs9456735	T, 0.941; G, 0.059; n=286
				IVS12+16G>A	3'UTR	No	-
90.88	MA	80	No	M192L	5	rs9456735	T, 0.941; G, 0.059; n=286
92.94	MA	57	No	L261L	7	rs143902760	G, 1.00; A, 0.00; n=4550
				V380L	10	rs1801582	G, 0.682; C 0.318; n=4550
				T387P	**10**	**Novel**	-
91.85	MA	65	No	R256C	7	rs150562946	G, 0.997; A, 0.003; n=1322
92.97	MA	77	No	L261L	7	rs143902760	G, 1.00; A, 0.00; n=4550
				P437L	12	rs149953814	G, 0.997; A, 0.003; n= 4282
88.74	Afrikaner	36	Yes	IVS8+48C>T	3'UTR	No	-
89.01	Caucasian	47	Yes	IVS8+48C>T	3'UTR	No	-
88.52	MA	55	Yes	V380L	10	rs1801582	G, 0.682; C 0.318; n=4550
88.53	MA	59	No	V380L	10	rs1801582	G, 0.682; C 0.318; n=4550

90.95	MA	27	Yes	D394N	11	rs1801334	G, 0.945; A, 0.055; n=4550

MA, Mixed Ancestry. All frequencies are determined in the Exome Sequencing Project cohort population available in dbSNP; -, no frequency available. All variants identified in a heterozygous state.

The altered HRM Difference graph for the T387P variant is shown in figure 2.4 and the sequencing results are shown in figure 2.5. Variant Effect Predictor Tool from Ensembl (http://www.ensembl.org/tools.html) predicted the variant to be "benign". Using the SMART Protein Analyser (http://smart.embl-heidelberg.de/), it was determined that the variant falls between the IBR domain and the RING2 domain. It therefore falls outside of the domains in an unknown region. Sequence alignments show that the T387P residue is not evolutionarily conserved and does not reside in a conserved region (figure 2.6).

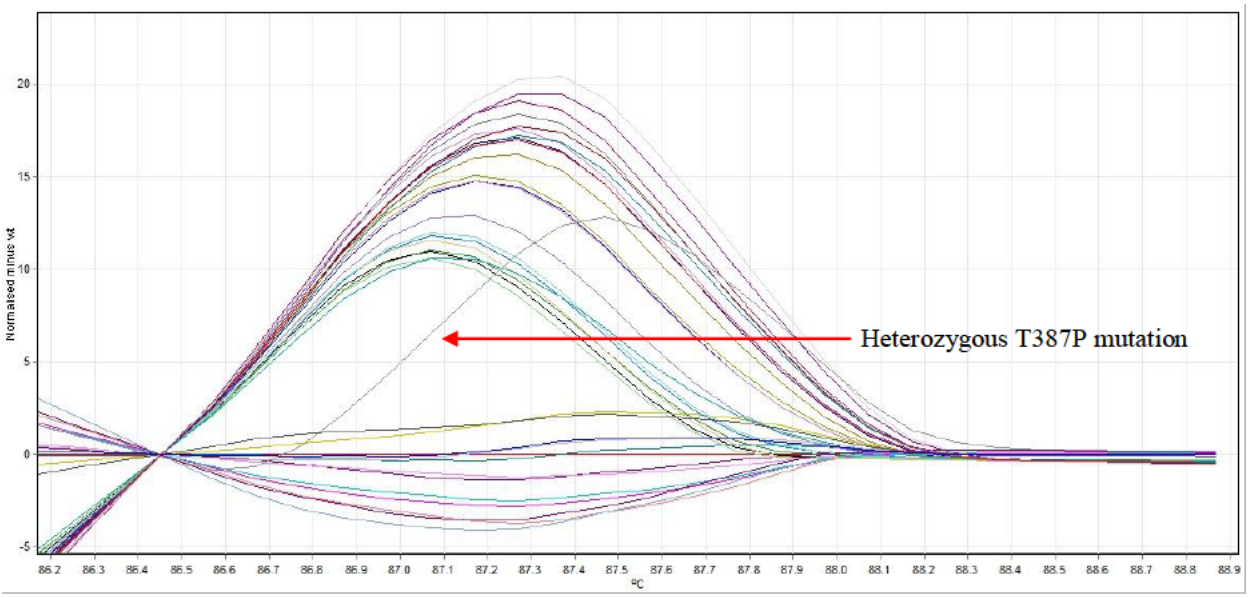
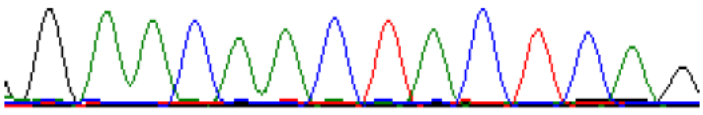


Figure 2.4 HRM difference graph indicating the *Parkin* T387P variant in one of the patients.

Wild Type

G A A C A A C T A C T C A G



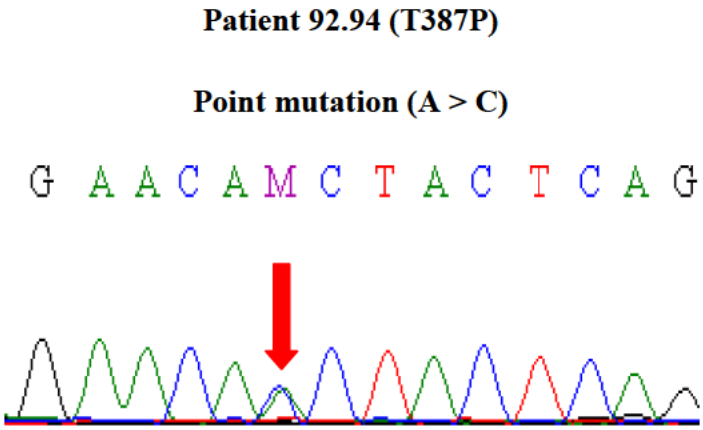


Figure 2.5 Sequence chromatogram indicating the position of the T387P variant in *Parkin*.

```
Parkinhuman     LLPEPDQRKV T CEGGNGLGCGFAFCRECKEAYHEGECSAVFEASGTTTQAYRVDER 396
Parkinchimp     LLPEPDQRKV T CEGGNGLGCGFAFCRECKEAYHEGECSALFEASGTTTQAYRVDER 396
Parkinmouse     LLPEQGQRKV T CEGGNGLGCGFVFCRDCKEAYHEGDCDSLLEPSGATSQAYRVDKR 394
Parkinrat       LLPEQGQKKV T CEGGNGLGCGVTFCRDCRMRHDLRLCDWLFENPSPTWQAYRVDQR 419
Parkinchicken   LLPEPEVRKI V CEPGNGIGCGFVFCRECKEEYHEGECSSFLSTQGAVAQGYVVDEN 393
                ****  :  * : ** ***:***..***:*:  :.   *. .::  ... *.* **:.
```

Figure 2.6 Sequence alignments of *Parkin* using ClustalW. Sequence alignments of *Parkin* comparing human (ENST00000366898), chimpanzee (ENSPTRT00000034690), mouse (ENSMUSG00000073465), rat (ENSRNOT00000049023) and chicken (ENSGALG00000011562). The box indicates the position of the T387P variant. The conserved regions have been indicated with an asterisk.

A total of 132 Mixed Ancestry controls were screened for the T387 variant in order to determine whether or not this is a common variant in this specific ethnic group. None of the controls were identified as carriers of this variant.

2.4 Discussion

The aim of the present study was to determine the role of *Parkin* in the South African PD patients. Using HRM and sequencing, 33 patients were screened and a number of polymorphisms (nine non-synonymous, two synonymous, one intronic and one in the 3'UTR) as well as one novel non-synonymous variant (T387P) were found. Notably, none of the patients screened had any of the known missense mutations in this gene.

Our research group had previously described that *Parkin* did not play a significant role in 229 South African PD patients and homozygous or compound heterozygous mutations were only identified in seven patients (Table 1.4; (Haylett et al. 2012). Therefore, taken together with

the results of the present study in which 33 additional patients were screened, the number of patients with *Parkin* mutations in our total cohort is only 2.7% (7/262). However, this frequency takes into account patients who also have late-onset PD (i.e. AAO >50 years). If only patients with a young AAO (i.e. AAO ≤ 50 years) are included in the study, then the frequency of *Parkin* mutations is 6.5% (6/93). Some studies have shown that *Parkin* mutations can be found in up to 20% of non-familial, early-onset PD cases and in up to 50% of familial early-onset PD cases (West et al. 2003). Frequencies of *Parkin* mutations can depend on the population group studied; in a recent study carried out on a Mexican population of PD patients, it was determined that 17.5% of the patients were carriers of simple heterozygous mutations and 25.4% were carriers of compound heterozygous mutations (Camacho et al. 2012). However, low frequencies of *Parkin* mutations have been reported in Serbian and Polish populations, where the frequencies of these mutations range between 1.3 and 3.8% in affected PD individuals (Djarmati et al. 2004; Koziorowski et al. 2010). The low frequency in the South African patients could therefore also be due to the numerous unique and diverse ethnic groups present in the country.

Notably, some of the non-synonymous variants such as S167N and M192L have previously been described as pathogenic in some sudies (Satoh and Kuroda 1999; Hedrich et al. 2002). However, given the high frequencies of these mutations in control individuals (7.9% for S167N and 6.7% for M192L (Haylett et al. 2012) it is unlikely that they are disease-causing and this suggests that further studies are needed before variants in this gene can be labelled as 'disease-causing'.

The present study also identified a novel variant, a heterozygous T387P in exon 10. This variant was identified in a Mixed Ancestry patient with an AAO of 57 years of age and no known family history. This patient was also identified as a carrier of two known polymorphisms, L161L and V380L (Table 2.2). The pathogenicity of T387P is not known as it was not identified in any of the 132 ethnic-matched control patients that were screened, but it was predicted to be benign through the use of the Variant Effect Predictor Tool (http://www.ensembl.org/tools.html). The T387P variant affects an amino acid which is evolutionarily conserved across human, chimp, mouse and rat but not in chicken. It is found between the IBR, is cysteine rich and separates the two RING domains (Shimura et al. 2000) - the RING finger motifs are essential for binding to UbcH7, which then conversely binds to the E2 conjugating enzyme thus forming an active component of the UPS (Shimura et al. 2000). This suggests that although this variant had been predicted to be benign, it could be

speculated that the location of the variant, as well as the fact that it is absent in the controls, may implicate it in PD susceptibility. Additional variants found in this region are A366T and T388L (Lohmann et al. 2003). Both of these variants have been predicted as non-pathogenic (http://www.ensembl.org/tools.html).

The large number of variants (fourteen in total) that were found, along with the fact that the positive controls for each exon were detected using HRM, is an indication that this mutation screening method is effective. However, it is still possible that some mutations could have been missed by not sequencing all of the exons in each of the patients, but this cost is currently prohibitively expensive in our particular setting.

Future studies should include screening for exonic deletions or insertions in *Parkin* in these 33 patients using a technique such as Multiplex Ligation-dependent Probe Amplification (MLPA) (Bardien et al. 2009; Keyser et al. 2010). In addition, functional studies such as yeast-two hybrid and AP/MS (Affinity Purification and Mass Spectometry) are warranted on the novel T387P variant to determine whether it possibly influences protein function.

In conclusion, heterozygous or homozygous point mutations in *Parkin* are not a major cause of PD in South African patients and therefore mutation screening of other PD genes are warranted. Identification of the disease-causing genes and mutations in these families is necessary as this will elucidate the biological pathways underlying PD in South African patients. In addition, this information would be important to confirm diagnosis in patients that present with atypical PD symptoms and also to identify at-risk family members.

Chapter 3: Mutation Screening of the *PINK1* gene

Page

3.1 Introduction

PTEN-putative induced kinase 1 (*PINK1*) (OMIM# 608309) has been implicated in autosomal recessive PD and was the first gene that successfully linked mitochondria to the disease (Valente et al. 2004). It is found on chromosome 1p36.12, is composed of 8 exons and encodes a protein of 581 amino acids in length, which is localized in the mitochondria. *PINK1* is composed of two main domains: a highly conserved serine/threonine kinase domain and a N-terminal mitochondrial targeting domain (MTS) and a transmembrane region is found between these two domains (Valente et al. 2001) (figure 3.1). The serine/threonine protein kinase found in this domain belongs to the calcium/calmodulin-dependant class of kinases (CaMKs) that are responsible for calcium transduction within cells (Abou-Sleiman et al. 2006; Ibanez 2006). The CaMKs are involved in a wide variety of cellular responses that are induced by factors such as neurotransmitters and other signalling mechanisms which involve hormones (Means 2001). Additionally, the CaMKs are significantly affected by calcium levels within the cells and so-called 'calcium-mobilizing stimuli' are employed in order to maintain intracellular calcium levels. Elevations in the levels of calcium thus act as a second messenger system (Means 2001). Although PINK1 has been identified as a protein kinase, the actual kinase activity of the protein has not yet been demonstrated (Moore et al. 2005).

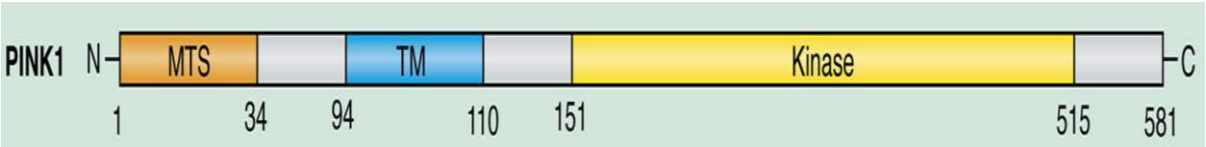


Figure 3.1 Protein domains of PINK1. The mitochondrial targeting domain (MTS), transmembrane region (TM) and the serine/threonine kinase domain are indicated (Taken from http://jcs.biologists.org/content/suppl/2012/03/25/125.4.795.DC1/JCS093849Panel2.jpg).

PINK1 was initially identified through linkage mapping in an Italian family that presented with autosomal recessive and early onset PD (Valente et al. 2001). The locus, PARK6 was identified in the family which had four definitely affected PD individuals with an AAO that ranged between 32 and 48 years of age (Valente et al. 2001). The PARK6 locus was confirmed in a further eight families across four different countries in Europe, all of whom had early onset PD (Brancati 2002).

Linkage to the PARK6 locus was also described in eight Asian families that were unrelated and had early onset PD - the AAO of these patients was reported between 18 and 56 years (Hatano et al. 2004). Noteworthy observations of this particular study was the documentation of levodopa responsiveness and slow disease progression (Hatano et al. 2004). This finding was further supported when 21 patients who had been positively identified as having *PINK1*-associated PD, were re-examined (Albanese et al. 2005). It was determined that 20 of the patients had slow disease progression, but that 84% of the patients had developed additional involuntary and impaired movements, which were concluded to be as a result of the levodopa treatment and not as a result of the disease (Albanese et al. 2005).

PINK1 provided the first evidence that a kinase signalling pathway may be actively involved in dopaminergic cell death (Moore et al. 2005). The serine/threonine kinase of PINK1 has been identified as a protective mechanism against oxidative stress and subsequent apoptosis (Deas, Plun-Favreau, and Wood 2009). PINK1 is critical for mitochondrial integrity as well as for the maintenance of mitochondrial homeostasis due to the presence of the aforementioned serine/threonine kinase (Valente 2004; Hoepken et al. 2007). Over expressed PINK1 has, however been localized to the mitochondria and has been further implicated in PD pathogenesis (Valente 2004; Abou-Sleiman et al. 2006).

Evidence for the involvement of mitochondria in PD first surfaced when heroin users presented with severe Parkinsonian features after the unintentional injection of a synthetic bi-product, MPTP (Langston 1983). MPP^+ (1-methyl-4-phenyl-pyridinium ion) is the active metabolite of MPTP and is able to cross the blood-brain barrier. This molecule is of particluar interest as it can be transported into the dopaminergic neurons (Dauer and Przedborski 2003). Interestingly, MPP^+ is an active inhibitor of mitochondrial complex I (Nicklas 1987) and the inhibition of this specific mitochondrial complex, is directly related to an increase in free radical generation. Free radical generation results in an increase in oxidative stress through changes in the electron transport chain (Schapira 1997; Schapira 2010). This discovery is of relevance to PD as continual studies have shown that PD patients have significantly lower activity in complex I but that this lack of activity is not due to levodopa treatment administered to the patients (Mann et al. 1994; Haas et al. 1995; Cooper et al. 1995). There are two current hypotheses involving mitochondrial complex I and neurodegenerative disorders: 1) the complex I inhibitors may be independently responsible for development of neurodegenerative disorders and 2) complex I inhibitors may contribute

towards the development of neurodegenerative disorders in conjunction with genetic and environmental factors (Schapira 2010).

Oxidative stress results in damage to intra- and extra-cellular cell structures, nucleic acids and proteins within the cells due to an excess of reactive oxygen species (ROS) (Storz and Imlayt 1999). It has been documented that increases in the ROS of the cells may be advantageous to the immune system and may play an important role in cell signalling (Zhou, Ma, and Sun 2008). The maintenance of ROS levels in the cells must remain carefully balanced - if the levels of ROS increase to such an extent that the cells cannot neutralize and then eliminate them from the targeted cells, they can damage intra-cellular structures, DNA, lipids and proteins (Zhou, Ma, and Sun 2008).

Under expression or accumulation of the protein is not the only mechanism by which PINK1 is thought to cause disease; PD-associated mutations that are found in *PINK1* may result in the inability of PINK1 to protect the mitochondria against oxidative stress and apoptosis (Deng et al. 2008; Haque et al. 2008) (figure 3.2). This is due to the fact that mutations in *PINK1* may prevent the phosphorylation of TRAP1 (mitochondrial chaperone tumour necrosis factor receptor-associated protein-1) and the protection against oxidative stress then becomes reduced (Pridgeon et al. 2007).

Figure 3.2 PINK1 and its protective role against oxidative stress. Should a mitochondrion be PINK1 deficient, it will show altered complex I activity. PINK1 is also necessary to phosphorylate TNF receptor associated protein 1 (TRAP1) in order to protect the cell against oxidative stress (Taken from Deas, Plun-Favreau, and Wood 2009).

Mitochondrial dysfunction is therefore thought to be the most likely way in which PINK1 plays a role in PD pathogenesis (Yao and Wood 2009). Interestingly, PINK1 and Parkin act in unison with each other in order to regulate mitochondrial function - and this pathway has been examined in PD pathogenesis (Yu et al. 2011). Should there be mutations in either *PINK1* or *Parkin*, a loss of function is exhibited (Cookson and Bandmann 2010). *Drosphila melanogaster* knockouts have been used to study the interactions of PINK1 and Parkin. Knockouts of *PINK1* result in mitochondrial abnormalities and apoptosis, both of which are predominant in the flight muscles of the fly as well as in the spermatid cells (Greene et al. 2003). *Parkin* knockouts have shown very similar phenotypic characteristics and additionally, in the absence of PINK1, Parkin is able to salvage the loss of PINK1 (Cookson and Bandmann 2010) thus providing evidence that the two proteins interact with each other in the same pathway (Clark et al. 2006; Park 2006). The converse cannot be seen and it has been concluded that PINK1 functions upstream of Parkin (Cookson and Bandmann 2010). The PINK1-Parkin pathway has been examined extensively and it has been determined that although Parkin is normally a cytosolic protein, if membrane potential decreases sufficiently, it can be recruited to the membrane of the mitochondria (Narendra et al. 2008) and then be involved in the promotion of mitochondrial autophagy (mitophagy) (Narendra et al. 2008; Geisler et al. 2010; Vives-Bauza et al. 2010).

3.1.1 Expression Profile

Numerous functional studies have been carried out on PINK1 and it has been determined that it is localized to the mitochondria, specifically in the mitochondrial matrix as well as in the intermembrane spaces (Gandhi, Chen, and Wilson-Delfosse 2009). PINK1 is expressed in numerous organs but high levels of this protein are found in the heart, skeletal muscle and testes (Unoki and Nakamura 2001) (figure 3.3).

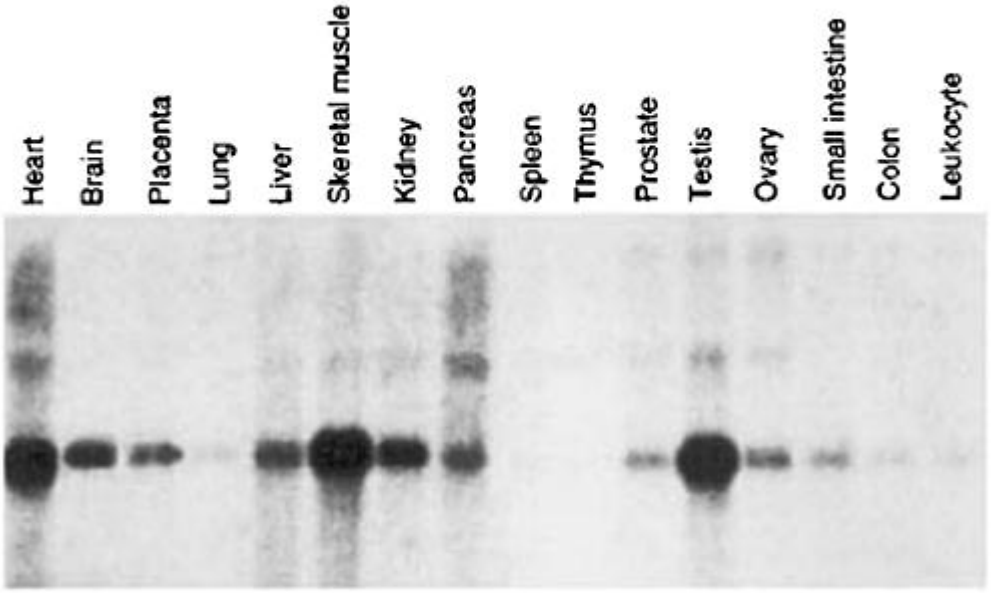


Figure 3.3 Northern Blot analysis of PINK1 expression. Northern blot analysis of PINK1 shows that the protein product of the gene is expressed in numerous tissues throughout the body. (Taken from Unoki and Nakamura 2001).

3.1.2 Postulated Function

Limited knowledge on the exact function of PINK1 is available. It has been postulated that it may play a role in neuroprotection, or more specifically, the prevention of mitochondrial dysfunction and apoptosis that is caused by the inhibition of proteins (Rohe 2004). The localization of PINK1 to the mitochondria has been shown through the use of a number of knockout models. As mentioned above, it has been noted that the loss of the specific phenotypes in *PINK1* knockout flies can all be directly associated with mitochondrial dysfunction and additionally have significant problems in coping with oxidative stress (Clark et al. 2006; Park 2006).

3.1.3 Disease-causing Mutations and Susceptibility Alleles

3.1.3.1 Disease-causing Mutations

PINK1 mutations have been identified in patients who have early onset PD - whether these cases are sporadic or familial. The frequency of *PINK1* mutations has been estimated to be between 0% and 15% across various ethnic groups (Ferraris et al. 2009). Additionally, *PINK1* mutations have been reported as the second most common cause of autosomal recessive, early onset PD and is preceded only by *Parkin*. Two *PINK1* mutations, G309D

and W437X were initially identified in three consanguineous families, and both mutations are found in the kinase domain (Valente et al. 2004). Additional mutations that have been identified are truncating mutations, frame shift mutations as well as several point mutations (Tan 2010; Bonifati 2012). It has been documented that the majority of the mutations which are found in *PINK1* are found in the kinase domain and decrease the enzymatic activity of the protein (Leutenegger 2006; Tan and Skipper 2007) (figure 3.4).

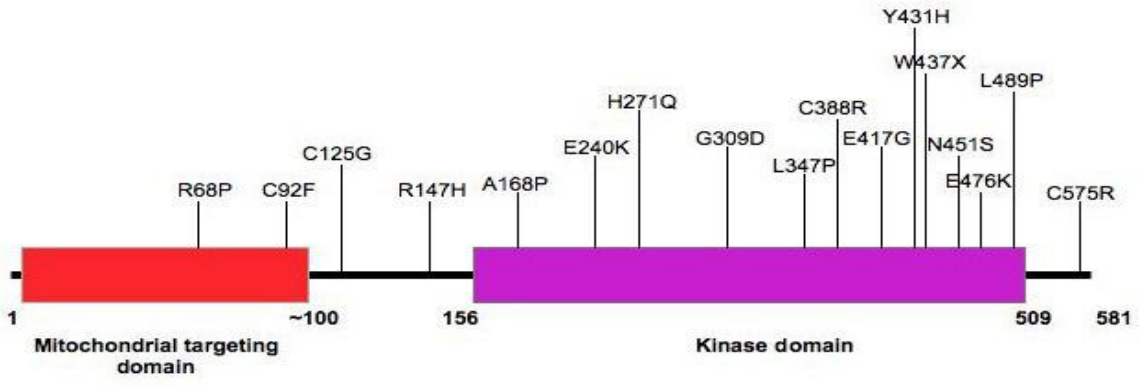


Figure 3.4 Positions of *PINK1* mutations in humans. A diagrammatic representation of some of the mutations that have been identified in the *PINK1* gene and that have been associated with PD. Note that the majority of the mutations are found in the kinase domain (Taken from http://www.ppu.mrc.ac.uk/research/?pid=11&sub1=research).

3.1.3.2 Susceptibility Alleles

Mutations in *PINK1* are relevant to understanding disease because of their frequency in PD patients; the frequency of these mutations is highly dependent on the ethnic group studied and may be as high as 15% in some groups of patients (Klein 2007). The current theories involving *PINK1* is that homozygous mutations cause autosomal recessive, early onset PD, while heterozygous mutations are considered to be susceptibility factors for the disease (Eggers et al. 2010). A number of susceptibility factors have been identified and these include L63L, A340T and N521T - the A340T variant has been associated with late-onset PD in the Chinese populations (Groen et al. 2004; Eggers et al. 2010). A340T has been extensively studied and it has been concluded that although the single nucleotide polymorphism (SNP) lies in a kinase domain, the corresponding amino acid residue (Threonine) is not evolutionarily conserved across species (Wang et al. 2006). Additionally, the N521T variant is found outside any coding region, but the Asparagine at position 521 is highly conserved across species (Rohe 2004). Susceptibility factors affect progression, onset

and severity of PD and are thought to have an additive effect - in *PINK1* it has been hypothesized that heterozygous mutations are more likely to increase the risk for the development of PD, while the homozygous or compound heterozygous mutations are disease-causing (Eggers et al. 2010).

3.1.4 The Present Study

The present study investigated the role of *PINK1* in South African PD patients. All eight exons were screened using HRM and sequencing.

3.2 Materials and Methods

3.2.1 Study Participants

A total of 155 patients had already been screened for possible mutations in exons 1 to 8 (R.Keyser, PhD thesis 2011; Keyser et al. 2010). In the present study, the remaining 107 of the 262 patients were screened for mutations in all exons.

3.2.2 Genetic Analysis

The HRM and sequencing methodology has been described in Chapter 2; section 2.2.2; page 34-35. Specific primer sequences had been designed for the PCR using Primer 3 software (Rozen 2000) and the primer sequence for each exon is provided in Table 3.1. In some cases, 5% of various additives (DMSO, Formamide and Betaine) were added for optimization of the PCR reaction.

Table 3.1 Primers designed for *PINK1* exons 1 to 8.

Region	Primer Sequence (5´-3´)	%GC	Tm (°C)	PCR conditions (Ta in °C)	Size of PCR fragment (bp)
Exon 1 (part A*)	For: agg cgc cat tac cag cat ag Rev: aag aag cgg aga cgg tta gg	55 55	60.3 59.5	58 (DMSO + Betaine)	389
Exon 1 (part B*)	For: agg ctg ggc cgc agg ac Rev: cct ccg ctc ggc tta gga c	76 68	64.7 62.1	55 (Formamide)	283
Exon 2 (part A*)	For: cct ttt ctt ggg cct tcc ta Rev: aat gta ggc atg gtg gct tc	50 50	57.1 58.2	55	204
Exon 2 (part B*)	For: agg gca gtc cat tgg taa gg Rev: ggg cat ttt gag aac atc tcc	55 47	59.4 57.3	55	250
Exon 3	For: agg cag ggc tta caa gga ac Rev: tgc tct caa aga agt ccc agt	55 47	60.0 59.0	55	220
Exon 4	For: agg tgt tgt atc tga tgc tg Rev: tcc cct tgg gag atg tat ca	45 50	55.2 57.1	55	287
Exon 5	For: cgt cga tgt gtg gta gcc Rev: tct agt gcc cct gga gag c	61 63	57.9 60.4	55	250
Exon 6	For: gaa gga ggg gag gag aaa tg Rev: tgc att cag tgg aca tgt gg	51 53	57.0 58.5	55	233
Exon 7	For: atg ggc ggg cag cgt gat gtc t Rev: ctg gaa cga gaa cag agg ttt c	63 50	69.1 58.9	55	339
Exon 8 (part A*)	For: gga cca gag aag gga aga cc Rev: ctt ctc tgt gag cct gtt gg	60 55	59.1 58.2	55 (DMSO)	248
Exon 8 (part B*)	For: aga tgg ttg gct ggc tcc t Rev: acc ctc acc att cac aga cc	58 55	60.9 59.1	55 (DMSO + Betaine)	248

Due to the large size of exons 1, 2 and 8, two sets of overlapping PCR fragments were produced for each of these exons. DMSO, dimethyl sulfoxide.

3.3 Results

After HRM analysis, a number of samples with altered HRM profiles were selected for sequencing. A total of five known as well as one novel variant was identified and the pathogenicity was assessed. The variants that were identified are shown in Table 3.2. The novel variant identified, A494V was analysed using the Variant Effect Predictor Tool from Ensembl (http://www.ensembl.org/tools.html) and was predicted to be benign (Table 3.3).

Sequence alignments showed that the A494 residue is evolutionarily conserved from man through to zebrafish, but that it does not occur in a conserved region of PINK1 (figure 3.5) although it lies in a kinase domain

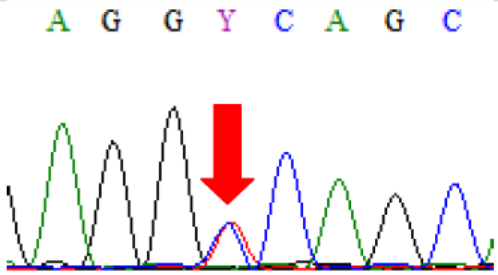
Table 3.2 Sequence variants which were identified in *PINK1* in South African PD patients.

Patient	Ethnicity	Age at Onset	Family History	Variant	Exon	In dbSNP	Frequency in dbSNP (n=number of chromosomes)
83.42	Caucasian	68	No	R312R	4	rs56200357	G, 0.998; A, 0.002; n=4550
81.67	Caucasian	63	No	R312R	4	rs56200357	G, 0.998; A, 0.002; n=4550
85.57	MA	62	No	S284Y	4	rs113092523	C, 0.999; A, 0.001; n=4552
91.95	MA	55	No	A340T	5	rs3738136	G, 0.910; A, 0.090; n=1324
81.68	MA	54	No	A340T	5	rs3738136	G, 0.910; A, 0.090; n=1324
84.66	MA	57	No	A340T (homo)	5	rs3738136	G, 0.910; A, 0.090; n=1324
84.47	MA	61	No	A494V	7	Novel	-
90.92	MA	75	Yes	R501Q	8	rs61744200	G, 0.988; A, 0.012; n=4546
90.98	Afrikaner	68	No	N521T	8	rs1043424	A, 0.688; C, 0.312; n=4544
98.77	Indian	55	Yes	3'UTR +37A>T	8	rs686658	A, 0.288; T, 0.712; n=120

				3'UTR+40G>A		rs115768147	A, 0.050; G, 0.950; n=120

MA, mixed ancestry; UTR, untranslated region; All frequencies are determined in the WESProject cohort population.

Table 3.3 Description of the non-synonymous novel A494V variant identified.

Patient Number	Mutation	Sequence Chromatogram	Zygosity	Variant Effect Predictor Tool
84.47	A494V		Heterozygote	Benign

```
PINK1Human        RALLQREASKRPS A RVAANVLHLSLWGEHILA-LKNLKLDKMVG 537
PINK1Chimp        RALLQREASKRPS A RVAANVLHLSLWGEHILA-LKNLKLDKMVG 537
PINK1Mouse        RSLLQREASKRPS A RLAANVLHLSLWGEHLLA-0LKNLKLDKMI 536
PINK1Chicken      RMLLQRDPNKRLS A RVAANVLHLSLWGESILA-SEALKPDQMTA 453
PINK1Zebrafish    KLLLRENPHKRPS A RVAANILHISLWGRRVLVGLDEVRMAEMMA 520
                  : **:.:. ** * * *:***:**:****. :*.  . ::  :* .
```

Figure 3.5 Sequence alignments of PINK1 using ClustalW. Sequence alignments of PINK1 for the human (ENS00000158828), chimpanzee (ENSPTRT000000052), mouse (ENSMUS0000030536), chicken (ENSGALT00000004909) and zebrafish (ENSDART00000006436). The box indicates the position of the A494V variant.

3.4 Discussion

PINK1 is the second most common cause of PD, preceded only by *Parkin*, and is responsible for the development of autosomal recessive, early onset PD (Deng et al. 2008). In the present study, we screened all eight exons of *PINK1* in the South African PD patients. A number of known variants were identified in the patients including four non-synonymous, one synonymous, two known variants in the 3'UTR and one novel variant (A494V).

In a previous study on *PINK1*, our group reported that of the 155 PD patients screened, mutations were identified in only one of the patients (a homozygous nonsense mutation, Y258X) (Chapter 1, Table 1.4; R. Keyser, PhD thesis, 2011; Keyser et al. 2010). Taken together with the findings of the present study, this reveals a frequency of only 0.4% of mutations in *PINK1* (1/262). The frequency of *PINK1* mutations in other studies ranges between 1 and 15%, but this is dependant on the population group which is studied (Bonifati et al. 2005). In a group of Italian patients, the frequency of *PINK1* mutations was to be found 7.7% (Bonifati et al. 2005), and in a Filipino PD population, 3% of the individuals who had been diagnosed with PD had mutations in *PINK1* (Rogaeva et al. 2004).

One patient of Mixed Ancestry and with no family history was found to have a novel heterozygous A494V variant in exon 7. The pathogenicity of this variant was assessed through the use of the Variant Effect Predictor Tool from Ensembl (http://www.ensembl.org/tools.html) and it was predicted to be "benign". The A494V variant is found in the kinase domain and occurs at an amino acid position that is evolutionarily conserved (figure 3.5). Interestingly, the majority of mutations that have been identified in *PINK1* have been found in the kinase domain and the enzymatic activity is reduced in these regions (Sim et al. 2006). Another observation has been that mutations which are located outside of the kinase domain may play a role in protein stability, therefore affecting kinase activity. No controls were screened to determine the frequency of the novel A494V variant because the patient (84.47) in which the variant was identified had already been identified as a carrier of an exonic deletion in *Parkin* (Table 1.4, Chapter 1).

A limitation of the present study is the fact that we did not screen for exonic rearrangements in this gene and therefore MLPA will be used in future studies of this gene. In conclusion, mutations in *PINK1* are thought to be a minor cause of PD in South African patients and other genes should be screened to identify causative genetic factors in these patients.

Chapter 4: Mutation Screening of the *LRRK2* gene

Page

4.1 Introduction

Leucine–rich repeat kinase 2 (*LRRK2*) (OMIM #609007) is implicated in the development of autosomal dominant PD. *LRRK2* was identified as a PD-causing gene in 2004 by two independant groups (Zimprich et al. 2004; Paisán-Ruíz et al. 2004). It encodes a multi–domain protein which is composed of 2 527 amino acids, and has 51 exons in total (Paisán Ruíz et al. 2005). It is located on chromosome 12p12 and the protein for which *LRRK2* encodes forms part of the ROCO protein family (Zimprich et al. 2004). LRRK2 is composed of a number of functional domains, including the armadillo domain (ARM); ankyrin repeat domain (ANK); leucine – rich repeat domain (LRR); Ras of complex proteins (ROC), carboxy terminal of ROC (COR) and mitogen–activated protein kinase kinase kinase (MAPKKK) (figure 4.1). The ROC domain binds to GTP and is essential for the functioning of MAPKKK and is noteworthy as it does not possess GTPase activity (figure 4.1). The final domain of LRRK2 is a WD40 domain which is rich in aspartate and tryptophan repeats (Zimprich et al. 2004).

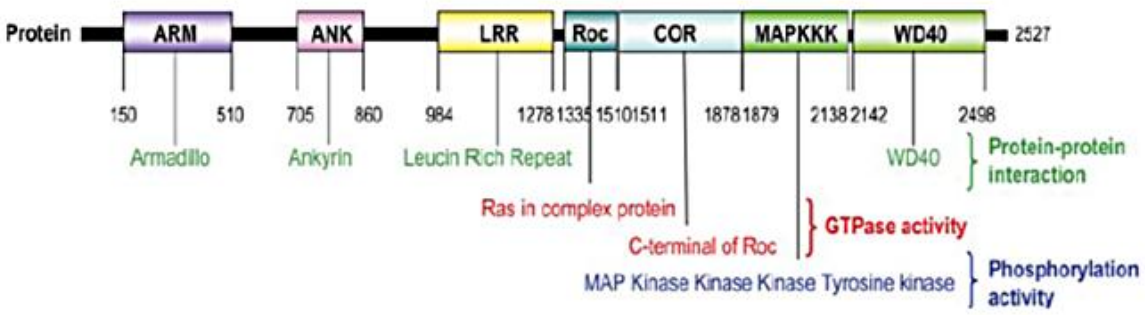


Figure 4.1 Protein domains of LRRK2. The LRRK2 protein contains 7 domains which are ARM, ANK, LRR, Roc, COR, MAPKKK and WD40. The numbers indicate the amino acids positions (Taken from Lesage and Brice 2009).

Mutations in the coding regions of *LRRK2* have been implicated and linked to the development of autosomal dominant, late onset PD (Farrer 2006). Pathologically, the patients are heterogeneous with documentations of both Lewy Body and tau pathologies. Additionally, motor neuron disease and neuronal loss without intracellular inclusions have also been reported (Zimprich et al. 2004). Tau pathologies are significant as these may also affect the outcome of tau proteins, which are highly soluble microtubule-associated proteins (MAPs) (Zimprich et al. 2004) that are significant role players in physiology as they act as stabilizers of microtubules through different isoforms and phosphorylation (MacLeod et al.

2006). Pathologies of the tau proteins may act as biochemical markers of neurofibrillary degeneration and this can be correlated directly with clinical manifestations of neurodegenerative disorders such as PD (Zimprich et al. 2004).

4.1.1 Expression Profile

LRRK2 is expressed in multiple tissues such as the central nervous system, peripheral leukocytes, the kidneys, lungs and liver as well as in the heart (figure 4.2) (Paisán Ruíz et al. 2005), but the pathological isoforms appear to be restricted to the brain regions, suggesting that a brain–specific context is crucial for the onset of *LRRK2* linked PD (Lin et al. 2009). An additional postulant is that LRRK2 that is produced in the brain may have specific biochemical characteristics that are directly associated with neuronal functions (Lin et al. 2009).

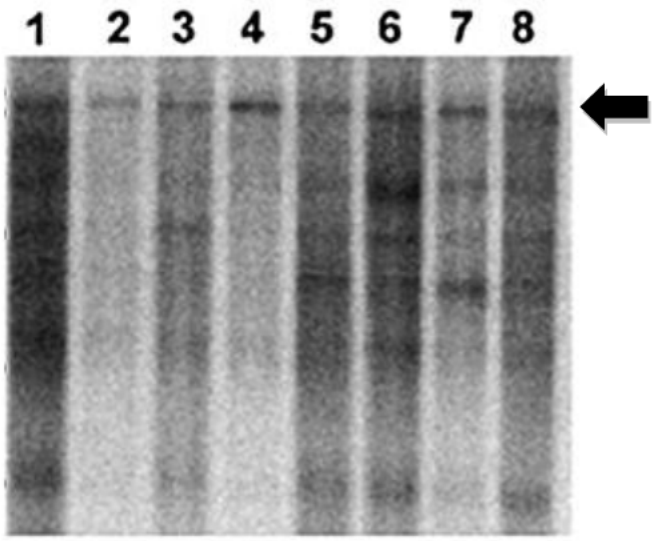


Figure 4.2 Northern Blot analysis of LRRK2 expression. Northern blot analysis shows that LRRK2 is expressed in many tissues. Various tissues are indicated by lane numbers: (1) heart; (2) brain; (3) placenta; (4) lung; (5) liver; (6) skeletal muscle; (7) kidney; (8) pancreas. The black arrow is in an indicator of LRRK2's position (Taken from Paisán Ruíz et al. 2005).

4.1.2 Postulated Function

Due to the presence of a number of functional domains, it has been suggested that LRRK2 may be involved in a number of functions, more specifically it may play a pivotal role in signal transduction pathways (Thomas 2007). However, the exact function is still unknown. LRRK2 has been identified as a tyrosine kinase–like protein; the ROC domain is thought to bind to GTP in order to regulate the kinase activity of MAPKKK (Deng et al. 2005) but it

does not possess GTPase activity (Ito, Shirai, and Hattori 2007). One study showed that LRRK2 presents with autophosphorylation activity as well as a mixed- lineage kinase activity (West et al. 2005). The presence of a number of protein interaction domains, suggests that LRRK2 may form the scaffolding complex which is necessary for the assembly of a multi–protein signal complex (Guo, Wang, and Chen 2006). Additional studies have made use of *Drosophila* showing that LRRK2 has the ability to interact with microRNA pathways as a means to regulate protein synthesis (Gehrke et al. 2010). LRRK2 has also been implicated in the endosomal–autophagic pathway and is a target gene for IFNγ which may be involved in additional pathogenic pathways (Gardet et al. 2010).

4.1.3 Disease-causing Mutations and Susceptibility Alleles

4.1.3.1 Disease-causing Mutations

To date, mutations in the *LRRK2* gene are significant as they are documented as the most common cause of dominantly inherited PD (Gasser 2010). *LRRK2* mutations have been implicated in approximately 10% of all cases of PD in Caucasians (Orr Urtreger 2007). These statistics do not exclude sporadic cases of PD, of which *LRRK2* mutations are estimated to appear in 3.6% of cases (Khan et al. 2005; Mata et al. 2005). It has been documented that certain variants identified in *LRRK2* may exert an increase in kinase activity (West et al. 2005). More than 100 sequence variants have been identified for this gene and a number of these variants have been directly and unequivocally implicated as disease causing variants by co-segregation in large families or via functional studies (Biskup et al. 2006). These variants include: N1437H located in exon 29; R1441C, R1441G, R1441H in exon 31; Y1699C in exon 35; I2020T and the most common mutation G2019S are both found in exon 41. These mutations are located in the functional domains of the protein (Hedrich et al. 2006).

G2019S has been identified as the single most common mutation in *LRRK2* and is frequent in both sporadic (1-2%) and familial (2-5%) cases of PD (Lesage and Brice 2009). Three different haplotypes have been found in G2019S carriers. Haplotype 1, a common 193kb genomic region (Zabetian et al. 2006), is shared by 95% of documented G2019S carriers from European, North and South African populations as well as Ashkenazi Jews (Lesage and Brice 2009). The mutation that these populations appear to share possibly developed in the Ashkenazi Jews. The hypothesis is that the mutation arose in this population much earlier than in the Europeans or in the North African Arabs (Lesage and Brice 2009). Haplotype 2

has been documented in five families of European ancestry, and haploype 3 is comprised of mainly Japanese carriers, but has also been recorded in a Turkish family (Lesage and Brice 2009). Interestingly, it appears as though the prevalence of G2019S may be affected by ethnicity – G2019S accounts for 30-40% of familial and sporadic cases of PD in patients from North Africa (Ozelius et al. 2006; Lesage et al. 2010), 10-30% of PD in Ashkenazi Jews (Orr Urtreger 2007; Ishihara et al. 2007), but is very rare in sub-Saharan Africa (Okubadejo et al. 2008), Asia and certain European countries such as Germany and Poland (Papapetropoulos et al. 2007; Lesage and Brice 2009). The penetrance of G2019S is also age specific – 28% penetrance at an age of 59 years, 51% at 69 years of age compared to 74% at an age of 79 years (Healy et al. 2008).

4.1.3.2 Susceptibility Alleles

LRRK2 is a highly polymorphic gene (Ross et al. 2008) and it was hypothesized that these polymorphisms may be risk factors in the development of PD. This hypothesis has since been proven, and it has been concluded that polymorphisms in *LRRK2* are susceptibility alleles for PD in certain patients and some polymorphisms may have a protective role in particular groups of individuals (Ross et al. 2011). In a study carried out on 15,540 individuals that were separated into varying ethnic groups, 121 *LRRK2* exonic variants were examined (Ross et al. 2011). It was determined that the M1646T variant identified in Caucasian PD patients and controls, was more likely to act as a susceptibility allele in all Caucasian patients from varying geographic locations except those individuals from South Africa. The same study found that three SNPs, namely N551K, R1398H and K1423K, although found in stronge linkage disequilibrium have a significant protective effect in Caucasians and Asian individuals, thus forming a so-called three-SNP haplotype (Ross et al. 2011).

A number of studies have been carried out on Asian populations in order to identify possible risk factors for PD (Di Fonzo et al. 2005). G2385R was found at a significantly higher frequency in PD patients who were of Chinese and Taiwanese ethnicity as opposed to the control populations (Di Fonzo et al. 2005). This observation was supported by independant studies, and it was subsequently suggested that the G2385R variant could be included as a common risk factor in the Asian populations (Tan et al. 2008). An additional SNP that has also been identified as a risk factor for late onset but dopamine responsive PD in the Asian populations is the R1628P variant. This SNP is found in a highly conserved region of the

gene (Ross et al. 2008) and in one study in Chinese populations was identified in 8.4% of PD positive patients vs. 3.4% of the controls (Tan et al. 2008). Subsequent *in vitro* analyses showed that G2385R and R1628P have increased autophosphorylation and kinase activity when compared to the wild type LRRK2 and could therefore increase the risk of PD via a similar mechanism to that of the pathogenic mutations (Tan 2010).

Genome wide association studies (GWAS) have been used to identify susceptibility alleles for PD. In GWAS, unrelated patients and controls who present with a particular phenotype as well as unaffected controls are genotyped for thousands of SNPs that are spread throughout the genome (Cookson 2010). Two noteworthy GWAS performed specifically on PD patients were on Caucasian patients and controls, and an additional study was performed on the Asian population (Cookson 2010). A few thousand cases were examined and it was determined that modest associations with PD could be detected (Simón-Sánchez et al. 2009). G2385R was found to be a common variant in the Asian population and was associated with PD. Also, for individuals who are from European decent and who present with a positive family history of PD, variation around *LRRK2* was consistently found to alter the risk for the development of PD (Simón-Sánchez et al. 2009).

4.1.4 The Present Study

In the present study, selected exons of *LRRK2* were screened for pathogenic mutations in the South African patients using HRM and sequencing. Exons 31 and 41 were screened as they contain the majority of the known mutations; R1441C/G/H (in exon 31), and G2019S and I2020T (both in exon 41). In addition, exon 42 was screened, as a putative pathogenic variant in this exon, Q2089R, had previously been found by our group in one Afrikaner PD patient (A. Marsberg, BScHons thesis, 2009).

4.2 Materials and Methods

4.2.1 Study Participants

For exon 31, a total of 194 patients had previously been screened for mutations (R. Keyser, PhD thesis, 2011) and therefore in the present study the remaining 68 patients were screened.

For exon 41, a total of 239 patients had previously been screened (Lesage et al. 2010; unpublished data) and therefore in the present study, the remaining 23 patients were screened.

The entire group of 262 patients were screened for the putative pathogenic variant, Q2089R, in exon 42. Additionally, 132 Afrikaner controls were screened to determine the frequency of this variant in the ethnically matched background population.

4.2.1 Genetic Analysis

The HRM and sequencing methodology has been described in Chapter 2; section 2.2.2; page 34-35. Specific primer sequences had been designed for the PCR using Primer 3 software (Rozen 2000) and the primer sequence for each exon is provided in Table 4.1.

Table 4.1 Primers designed for exons 31,41 and 42 of the *LRRK2* gene.

Region	Primer Sequence (5´-3´)	%GC	Tm (^{0}C)	PCR conditions (Ta in ^{0}C)	Size of PCR fragment (bp)
Exon 31	For: agc agg ccc agt ttg aaa g	53	58.3	55	173
	Rev: gaa ccc tcg ctt att cag ga	50	57.2		
Exon 41	For: gca cag aat ttt tga tgc ttg	38	54.9	55	331
	Rev: gag gtc agt ggt tat cca tcc	52	57.5		
Exon 42	For: gcc tcc ttg gat gta gta tga gc	52	60.6	55	289
	Rev: tga agc tgc tga tat taa gaa aa	30	54.4		

4.3 Results

Screening for mutations in exon 31

The R1441C mutation could easily be identified and distinguished from the wild type samples using HRM analysis (figure 4.3).

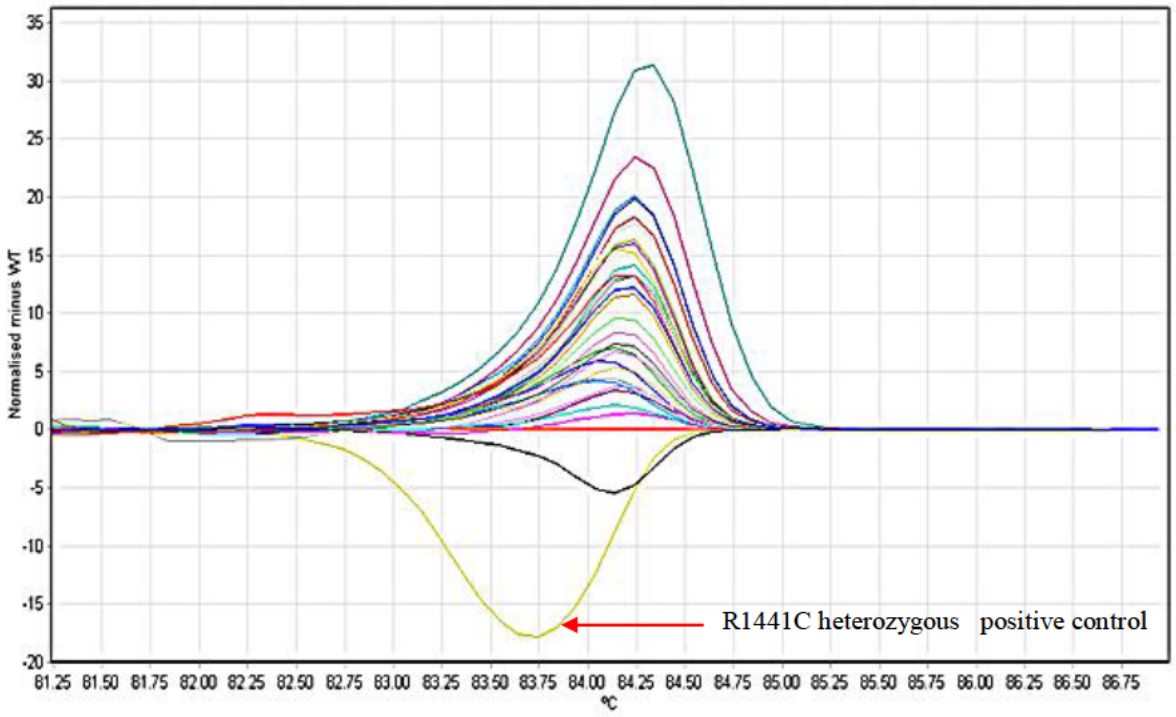


Figure 4.3 HRM difference graph indicating that the heterozygous *LRRK2* R1441C mutation can be identified due to the altered HRM profile.

Of the 68 patients that were screened, none were shown to harbour the R1441C, R1441G or R1441H mutations or any other mutations in exon 31. A total of five samples were sequenced and it was confirmed that there were no mutations or polymorphisms present. R1441C had previously been identified in one of the PD patients. Therefore, in the present study, the family members of this particular patient were sequenced in order to identify additional mutation carriers and it was determined that both siblings of the R1441C positive patient also carry the same mutation as well his niece (figure 4.4). Using the SMART Protein Analyser (http://smart.embl-heidelberg.de/), it was determined that the variant falls in the ROC domain of the protein and occurs in an evolutionarily conserved region (figure 4.5).

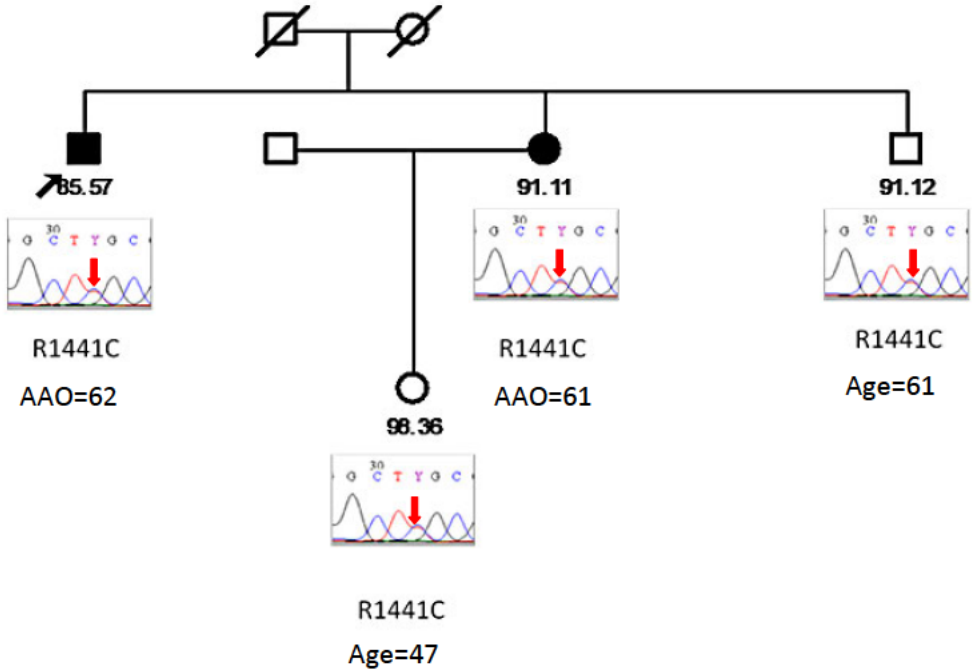


Figure 4.4 Sequencing results from the family carrying the R1441C mutation. The mutation is present in both of the siblings of the proband 85.57 as well as in the daughter of one of the patients.

```
LRRK2Human      KA R ASSSPVILVGTHLDVSDEKQRKACMSKITKELLNKRGFPAIRDYHFVNATEES 1498
LRRK2Chimp      KA R ASSSPVILIGTHLDVSDEKQRKACISKITKELLNKRGFPAIRDYHFVNATEES 1498
LRRK2Rat        KA R ASSSPVILVGTHLDVTDEKQRKACISKITKELLNKRGFPAIRDYHFVNATEES 1498
LRRK2Cow        KA R ASSSPVILVGTHLDVTDEKQRKACISKITKELLNKRGFPAIRDYHFVNATEES 1498
LRRK2Mouse      KA R ASSSPVILVGTHLDVSDEKQRKACISKITKELLNKRGFPTIRDYHFVNATEES 1498
                ** * ********:******:********:***************:************
```

Figure 4.5 Sequence alignments of LRRK2 (R1441C) using ClustalW. Sequence alignments of LRRK2 R1441C in human (ENSG00000188906), chimpanzee (ENSPTR0000004828), mouse (ENSMUS36273), rat (ENSRNOG0000004048) and cow (ENSBTAG00000016260). The box indicates the position of the R1441C variant. The conserved regions have been indicated with an asterisk.

Screening for mutations in exon 41

Of the 23 patients screened for G2019S, none were found to harbour this specific mutation, or any other mutation in exon 41 as the altered HRM profiles for the G2019S positive control sample could easily be identified (figure 4.6). A total of seven samples were selected and sequenced and this confirmed that no mutations or polymorphisms were present. Using the SMART Protein Analyser (http://smart.embl-heidelberg.de/), it was determined that the variant falls the MAPKKK domain. The protein alignment using ClustalW showed that G2019S falls in a well conserved region (figure 4.7).

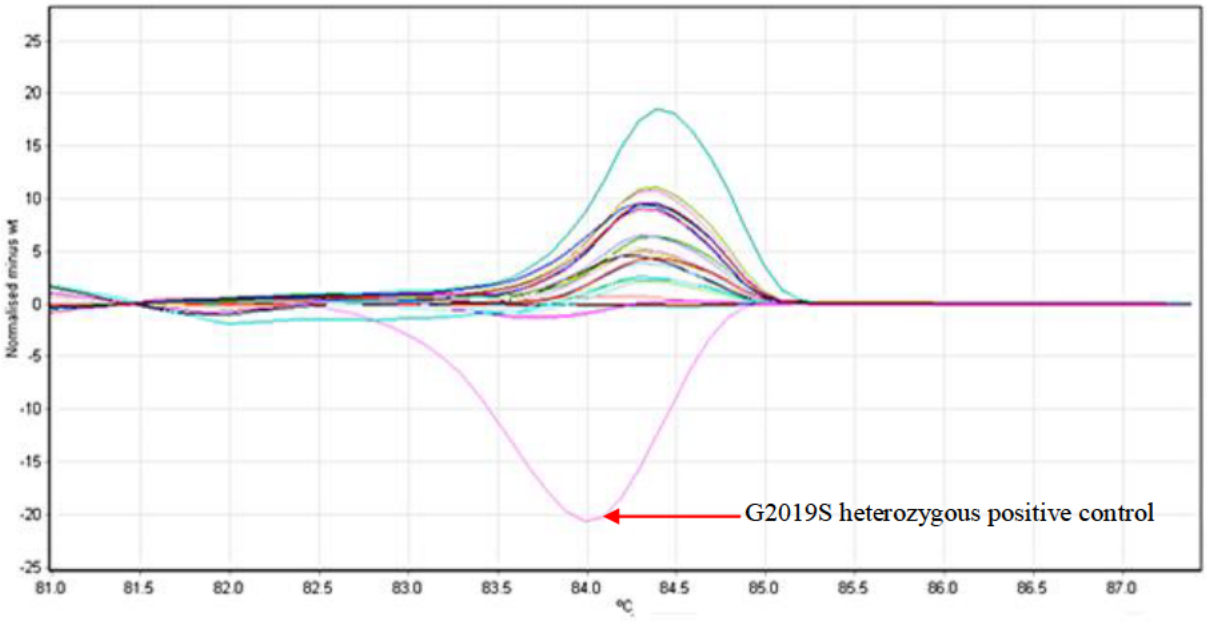


Figure 4.6 HRM difference graph indicating that the *LRRK2* G2019S mutation can be identified due to the altered HRM profile.

```
LRRK2Human      FTLYPNAAIIAKIADY G IAQYCCRMGIKTSEGTPGF 2038
LRRK2Chimp      FTLYPNAAIIAKIADY G IAQYCCRMGIKTSEGTPGF 2038
LRRK2Rat        FTLYPNAAIIAKIADY G IAQYCCRMGIKTSEGTPGF 2038
LRRK2Cow        FTLYPNAAIIAKIADY G IAQYCCRMGIKTSEGTPGF 2038
LRRK2Mouse      FTLYPNAAIIAKIADY G IAQYCCRMGIKTSEGTPGF 2038
                ****************   ********************
```

Figure 4.7 Sequence alignments of LRRK2 G2019S using ClustalW. Sequence alignments of LRRK2 G2019S of the human (ENSG00000188906), chimpanzee (ENSPTR0000004828), mouse (ENSMUS36273), rat (ENSRNOG0000004048) and cow (ENSBTAG00000016260). The box is an indication of the position of the G2019S variant. The conserved regions (where there are no amino acid changes) have been indicated with an asterisk.

Screening for the Q2089R sequence variant in exon 42

The Q2089R variant in exon 42 had previously been identified in one Afrikaner PD patient in a heterozygous state. In the present study, the entire group of 262 patients were screened using HRM and no additional patients harbouring the Q2089R variant were identified (figure 4.8).

Family members of the individual who was originally identified as a carrier of the Q2089R variant were sequenced in order to identify additional Q2089R–carrying individuals.

The variant was identified in both of the children of the proband (figure 4.9). However, no symptoms of PD in the children are apparent. At this stage, as the two children are younger than the identified AAO of their father (which is 50 years), it is not possible to rule out the fact that Q2089R may be pathogenic. As both parents of the proband are deceased, it is not possible to determine whether Q2089R co-segregates with the disorder. Interestingly, the Q2089R variant was absent in individual 70.25, the unaffected sister of the proband.

A total of 132 additional Afrikaner controls were screened using HRM in order to determine whether or not this is a common variant in this specific ethnic group. One of the controls exhibited an altered HRM profile and the sample was sequenced which confirmed the presence of the variant (figure 4.10). This control had been recruited at the age of 72 years and to our knowledge, did not exhibit any signs of PD.

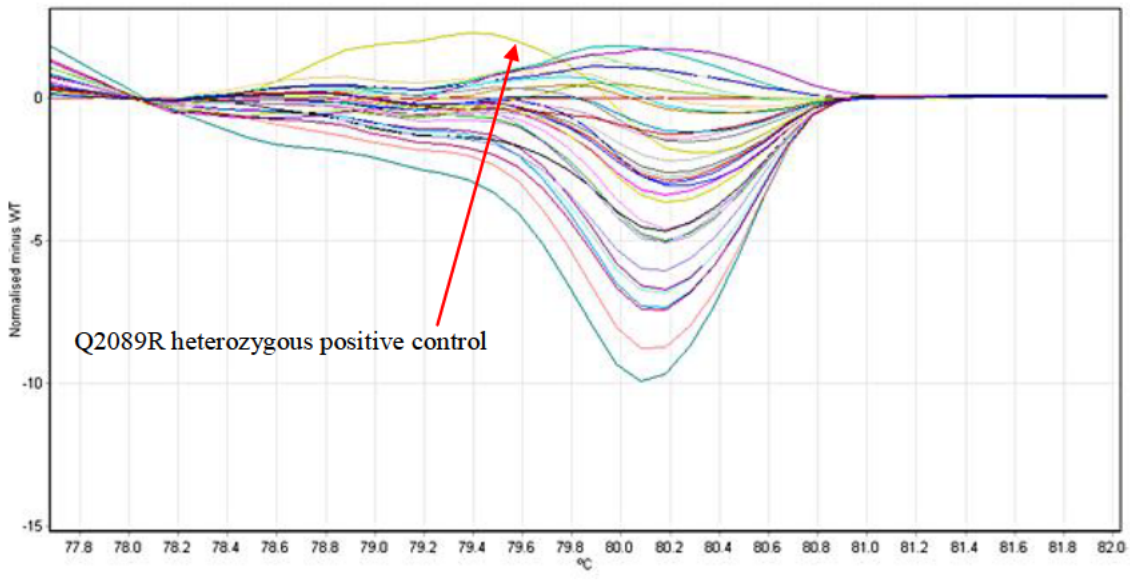


Figure 4.8 HRM difference graph indicating that the *LRRK2* Q2089R variant can be identified due to the altered HRM profile.

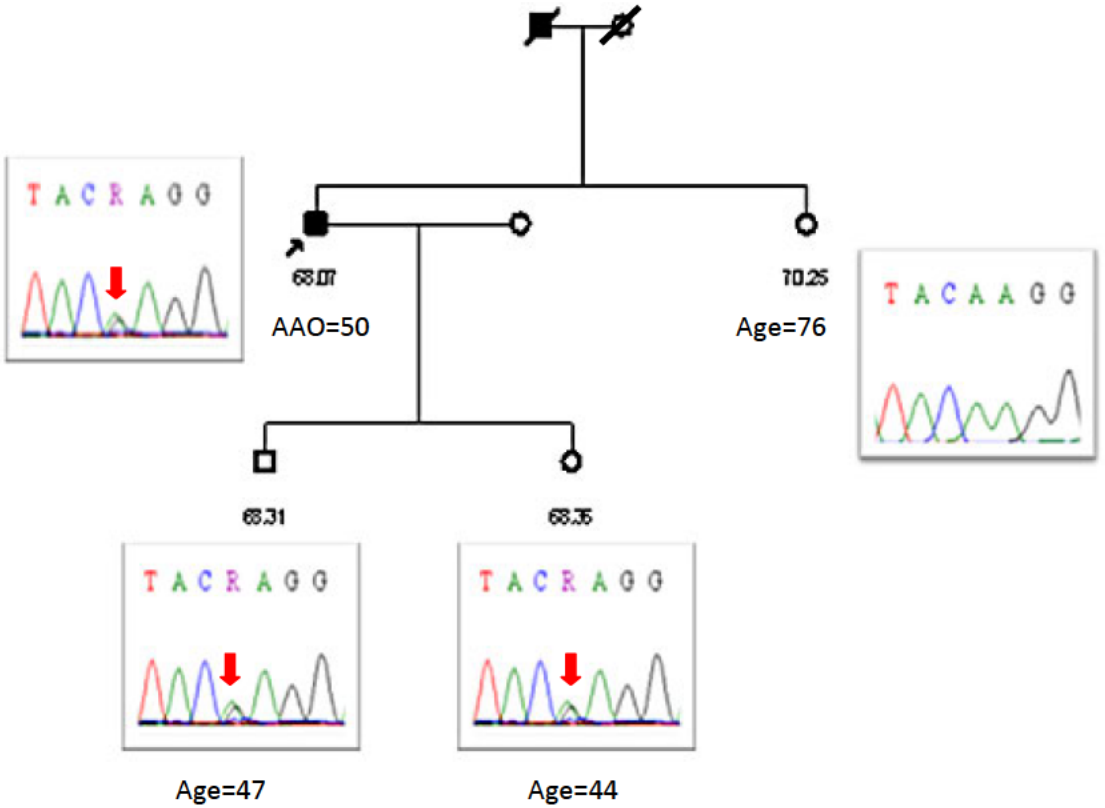


Figure 4.9 Sequencing results from the patient with the Q2089R variant his additional family members. The variant is present in both children of the proband, but not in his unaffected sibling.

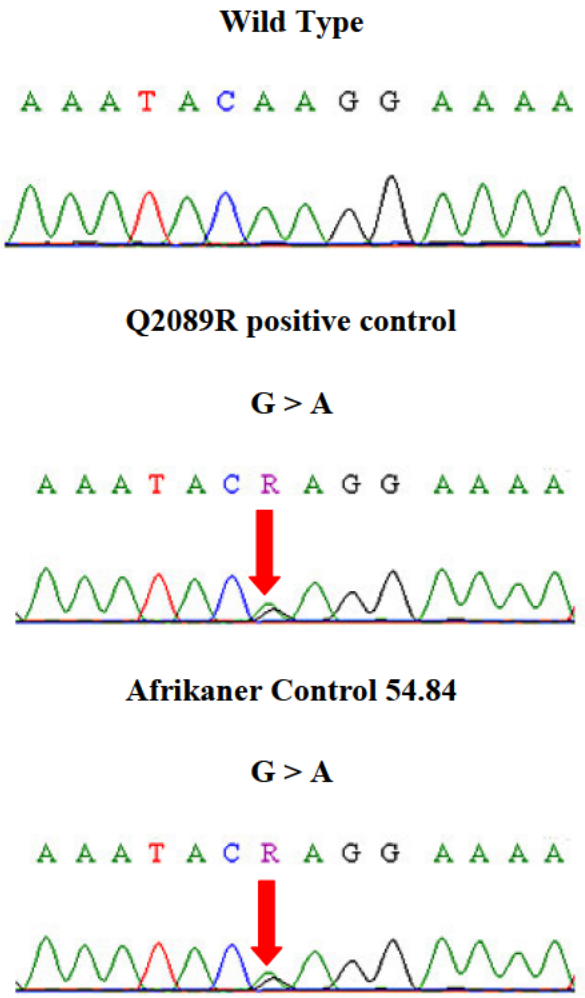


Figure 4.10 Sequence chromatogram indicating the position of the Q2089R variant in the wild type, patient and control.. The point mutation also present in the Afrikaner control.

```
LRRK2Human      YDILTTGGRIVEGLKFPNEFDELEI Q GKLPDPVKE 2098
LRRK2Chimp      YDILTTGGRIVEGLKFPNEFDELEI Q GKLPDPVKE 2098
LRRK2Rat        YDILTTGGRIAEGLKFPNEFDELAI Q GKLPDPVKE 2098
LRRK2Cow        YDILTTGGRIAEGLKFPNEFDELAI Q GKLPDPVKE 2098
LRRK2Mouse      HDIWTTGSRIMEGLRFPNEFDELAI Q GKLPDPVKE 2098
                :** ***.** ***:******** * * *********
```

Figure 4.11 Sequence alignments of LRRK2 Q2089R using ClustalW. Sequence alignments of LRRK2 Q2089R in human (ENSG00000188906), chimpanzee (ENSPTR0000004828), mouse (ENSMUS36273), rat (ENSRNOG0000004048) and cow (ENSBTAG00000016260). The box is indicates the position of the Q2089R variant. The conserved regions (where there are no amino acid changes) have been indicated with an asterisk.

The frequency of the Q2089R variant in ethnically matched controls is 0.38% (1/264 chromosomes) and using the Variant Effect Predictor Tool from Ensembl (http://www.ensembl.org/tools.html), it was predicted that the variant is "possibly damaging". This variant is found in the MAPKKK domain, a region that is evolutionarily conserved across a number of species (figure 4.11).

4.4 Discussion

The aim of the present study was to determine whether any of the South African patients harbour known or novel mutations in three exons (exons 31, 41 and 42) of the *LRRK2* gene. All 51 exons of LRRK2 were not screened due to the costs, and to date, the majority of the mutations have been found in exons 31 and 41. Mutations in *LRRK2* have been identified as a significant cause of autosomal dominant, late onset PD (Farrer 2006; Guedes et al. 2010).

A total of 68 patients were screened for mutations in exon 31 and none were found to carry the R1441C mutation. In a previous study, R1441C was found in one patient (Table 1.4, Chapter 1; R. Keyser, PhD thesis, 2011) and therefore in total, the frequency of the R1441C mutation is 0.4% (1/262) in this group of patients. Studies carried out on various population groups worldwide have reported different frequencies for the R1441C mutation - the frequency was reported to be as high as 20% in familial PD cases in a Spanish cohort (Paisán-Ruíz et al. 2005) and as low as 0.2% of familial PD cases in Northern Nebraska (Zimprich et al. 2004). Furthermore, the family members of the carrier (AAO = 62) of the R1441C mutation were screened using sequencing and the variant was identified in an affected sibling and an unaffected sibling (61 years of age) of the patient as well as his unaffected niece (47 years of age) (figure 4.4). The sibling was not interested in neurological

assessment to determine whether he had any signs of PD. The niece was examined and is currently unaffected but may well develop PD at a later stage. R1441C is found in the ROC domain of the protein (Lesage and Brice 2009), which contains a GTPase, a hydrolysis enzyme found in the highly conserved G domain, which along with the kinase domain, is hypothesized to play a role in signal transduction and other cellular signalling pathways (Vancraenenbroeck et al. 2012). This hypothesis is supported by the fact that the proteins found in the ROC domain include adenylyl cyclase which is a significant role player in a range of cellular processes such as signal transduction, cytoskeletal organization and nuclear transport (Shin et al. 2008; Vancraenenbroeck et al. 2012).

A total of 23 patients were screened for mutations in exons 41 and none had the common G2019S mutation. Previously G2019S had been found in 5 patients (Table 1.4; R. Keyser, PhD thesis, 2011) and therefore the total frequency of G2019S in our patients is 1.9% (5/262). The worldwide frequencies of the G2019S mutation are variable and the frequency observed in the South African patients is significantly lower than in many populations: the highest documented is in North African Arabs, where the frequency has been reported as 39% and the lowest rates have been reported as 0.1% in countries such as Greece and India (Guedes et al. 2010).

All 262 patients were screened for a novel heterozygous variant, Q2089R in exon 42 which had been previously identified in one of our Afrikaner Caucasian patients through direct sequencing of this exon in 30 patients (A. Marsberg, BSc Hons thesis, 2009). This variant was predicted as "possibly damaging" through the use of the Variant Effect Predictor Tool (http://www.ensembl.org/tools.html) and therefore the frequency in our entire PD cohort was examined. Q2089R was not identified in any additional patients. The available family members of the affected patient (AAO = 50) were screened and it was determined that the Q2089R variant is found in both of the unaffected children of the affected individual (aged 47 and 44 years of age respectively), but was not found in the unaffected sibling, who is aged 76 years (figure 4.9). Additionally, 132 Afrikaner controls were screened in order to determine the frequency in the population and one individual was found to harbour the variant, providing a frequency of 1/264 chromosomes (0.38%). The Q2089R variant is of importance as it falls in the MAPKKK domain (similar to that of G2019S and I2020T) and mutations in this region are hypothesized to have a significant effect on the kinase activity of the protein (West 2005; Vancraenenbroeck et al. 2012). In addition, Q2089 is evolutionarily conserved and resides within a highly conserved region (figure 4.11).

Limitations of the present study include the fact that not all of the exons of *LRRK2* were screened and therefore future studies should involve mutation screening of all 51 exons to determine the full spectrum of mutations in this gene. In addition, kinase assays should be performed for the Q2089R variant to determine whether it has a possible effect on the kinase function of LRRK2.

In conclusion, in the present study the R1441C and G2019S mutations were both found at a relatively low frequency when compared to other studies. It is plausible that other mutations in this gene are the cause of PD in our patients and it would be interesting to investigate whether founder effects may exist for novel mutations, as has been found for G2019S, particularly in the Afrikaner patients.

Chapter 5: Mutation Screening of the *SNCA* gene

Page

5.1 Introduction

α-Synuclein (*SNCA*) (OMIM# 163890) was the first gene to be identified as a pathogenic PD gene (Polymeropoulos et al. 1996). It has been implicated in autosomal dominant, early onset PD and is found on chromosome 4q21. *SNCA* is composed of 6 exons that encode a 140 amino acid protein called α-synuclein (SNCA) (Jo et al. 2000). SNCA is comprised of an amphipathic N-terminal region, a non-amyloid-B component (NAC) domain and a C-terminal region, that is highly acidic (figure 5.1) (Fortin et al. 2004; Bisaglia et al. 2009).

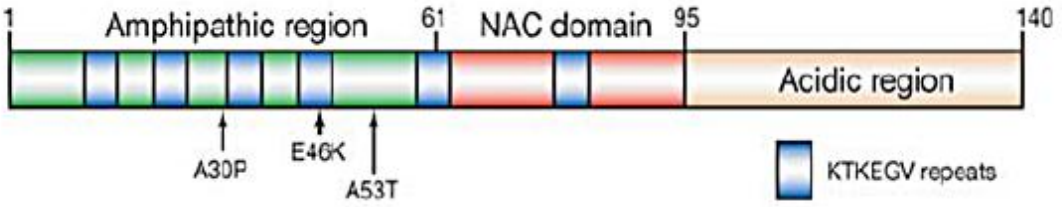


Figure 5.1 Protein domains of SNCA. The N-terminal amphipathic region, NAC domain and C-terminal acidic domain are indicated above. The positions of the three known pathogenic point mutations are indicated on the diagram (Taken from Moore et al. 2005).

SNCA is a member of the synuclein family and is one of three structurally related proteins. This family includes beta–synuclein (which has been implicated as an antagonist to α–synuclein) and gamma–synuclein (implicated in cancer as well as certain neurodegenerative disorders) (Surguchov and Jeon 2008). These proteins are small and soluble, are expressed in neural tissues and have also been identified in certain tumours. There are two noteworthy structural characteristics of these proteins: the presence of a repetitive but degenerative amino acid motif KTKEGV throughout the first 87 residues, as well as the acidic stretches that have been identified in the C – terminal region (Surguchov and Jeon 2008). SNCA is an intrinsically unfolded molecule that is able to assemble into Lewy body–like filaments, which separates it from other synucleins as it is the only protein capable of doing this (Dawson and Dawson 2003). Due to the small size of SNCA, it has a tendency to misfold and this misfolding leads to aggregation (Uversky and Eliezer 2009). The misfolding is thought to have a significant effect on the toxic effects of the protein leading to the increase in protein load, that will ultimately result in pathogenicity (Uversky and Eliezer 2009). SNCA has been identified as the major fibrillar component of Lewy bodies and Lewy body neurites, the pathological hallmarks of both sporadic and familial PD. Three point mutations, A30P, E46K and A53T in *SNCA* have been implicated in PD pathogenesis and pathogenic whole

gene duplications and triplications of the whole gene have also been reported in certain cases (Mueller et al. 2005; Uversky and Eliezer 2009).

5.1.1 Expression Profile

SNCA is expressed throughout the brain and is found in high concentrations in the presynaptic nerve terminals (George 2002). The major cranial regions that have been reported as having the highest levels of the protein are the cerebral neocortex, the hippocampus and the substantia nigra (George 2002).

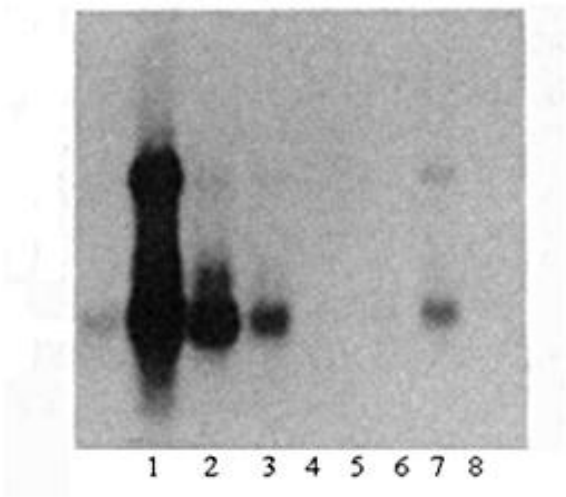


Figure 5.2 Northern Blot analysis of SNCA expression. A Northern Blot showing the expression of SNCA in the human body. SNCA is expressed predominantly in the heart and brain. Lane 1: heart, Lane 2: brain, Lane 3: placenta, Lane 4: lung, Lane 5: liver, Lane 6: skeletal muscle, Lane 7 : kidney, Lane 8: pancreas (Taken from Uéda et al. 1993).

5.1.2 Postulated Function

It has been postulated that SNCA may have a number of functions – however, a single core function is yet to be determined. SNCA is a significant component of Lewy bodies (Moore et al. 2005). It remains unclear whether the Lewy bodies themselves are toxic to the cells, or whether the so-called protofibrils (intermediates that have been fibrillized) are the major cause for toxicity (Miller et al. 2004). SNCA has been identified as a possible protein interaction partner of synphilin–1, that is present in numerous cranial regions (Engelender et al. 2000). Functional studies have been carried out in order to further investigate the interaction between SNCA and synphilin-1. It has been determined that the two proteins interact with each other in the neurons when *in vivo* and promote the formation of cytosolic inclusions (Engelender et al. 2000). When the proteins are co-transfected into HEK23 cells,

cytoplasmic eosinophilic inclusions can be observed (Engelender et al. 2000). This study has been supported by an additional study, where it was shown that synphilin-1 can be ubiquitinated by Parkin, subsequently resulting in ubiquitin positive, Lewy body–like structures (Chung et al. 2001).

SNCA has been identified as an active role player in oxidative stress (Moore et al. 2005). Additional studies show that over expression of the protein in P12, NT2 and SK-N-MC neuroblastoma cell lines may result in a decrease in proteasome activity, and therefore a subsequent decrease in the viability of the cell (Lee and Price 2001; Petrucelli et al. 2002; Fleming et al. 2004). SNCA has also been implicated in the induction of fibrillization of MAPT; animal models have provided evidence that the co-incubation of SNCA and tau will promote the fibrillization of both of these proteins jointly (Fleming et al. 2004). This is significant and suggests that there is an interaction between SNCA and tau that may account for the inclusions present in various neurodegenerative disorders.

5.1.3 Disease-causing Mutations and Susceptibility Alleles

5.1.3.1 Disease-causing Mutations

The first point mutation that was identified in *SNCA* was the A53T mutation and only two additional point mutations have been identified to date and they are A30P and E46K (Kruger 1998; Zarranz et al. 2004). Whole gene duplications and triplications have also been identified in patients, but it should be noted that all of the above-mentioned mutations are extremely rare. To date, no *SNCA* point mutations have been identified in patients with sporadic PD (Berg et al. 2005). It has been documented that patients with *SNCA* duplications present with late onset but levodopa responsive PD, that resembles typical PD. Patients who have *SNCA* triplications have much earlier onset PD, which is considered to be more severe, with dementia and a faster disease progression common in these individuals (Fuchs et al. 2007). Cumulatively, all results that have been documented to date show that SNCA is capable of triggering neurodegeneration when the overall expression levels are increased by even a modest level – thus the mutations may cause disease due to the fact that there is an increase in function (Gasser 2010).

5.3.1.2 Susceptibility Alleles

Two recent GWAS on PD have concluded that *SNCA* may play a role in both familial and sporadic PD, even though no *SNCA* point mutations have been identified in sporadic cases (Simón-Sánchez et al. 2009; Satake et al. 2009). This may be due to the fact that SNCA has been identified in Lewy bodies that are present in both forms of PD (Cookson and Bandmann 2010). It remains to be determined however, whether the increased risk for PD in sporadic cases from GWAS is mainly due to the SNPs identified using this method may affect dosage of SNCA, or whether there is an additional, underlying cause (Cookson and Bandmann 2010).

5.1.4 The Present Study

In the present study two of the six exons of *SNCA* were screened using HRM and sequencing. These exons were selected as the three missense mutations previously identified reside in these exons; A30P (in exon 2), and E46K and A53T (both in exon 3).

5.2 Materials and Methods

5.2.1 Study Participants

A total of 119 patients had already been screened for the A30P, E46K and A56T mutations (unpublished data). In the present study the remaining 143 of the 262 patients were screened for mutations in both exons. Due to the fact that the variations in *SNCA* are relatively rare, no positive controls were available for the HRM analyses.

5.2.2 Genetic Analysis

The HRM and sequencing methodology has been described in Chapter 2; section 2.2.2; pages 34-35. Specific primer sequences had been designed for the PCR using Primer 3 software (Rozen 2000) and the primer sequences for each exon are provided in Table 5.1.

Table 5.1 Primers designed for exons 2 and 3 of the *SNCA* gene.

Region	Primer Sequence (5´-3´)	%GC	Tm (^{0}C)	PCR conditions (Ta in ^{0}C)	Size of PCR fragment (bp)
Exon 2	For: ccc cga aag ttc tca ttc aa Rev: ccc atc act cat gaa gc	45 50	55.7 52.0	55; 3.0mM $MgCl_2$	235
Exon 3	For: ttt aag gct agc ttg aga ct Rev: cca cac taa tca cta gat ac	40 40	53.7 49.8	40; 2.5mM $MgCl_2$	146

5.3 Results

The screening of exons 2 and 3 in *SNCA* using HRM showed that no patients had any of the known mutations. None of the patients that were screened using HRM showed altered denaturation profiles (figure 5.3 and 5.5). However, for each exon, samples that appeared to have any shift in the melt curves, were selected and sequenced in order to verify that no sequence variants were present (figures 5.4, 5.6 and 5.7).

It was therefore concluded that none of the 143 patients that were screened for mutations in both exons 2 and 3 had any known or unknown pathogenic mutations – the patients were thus all wild type.

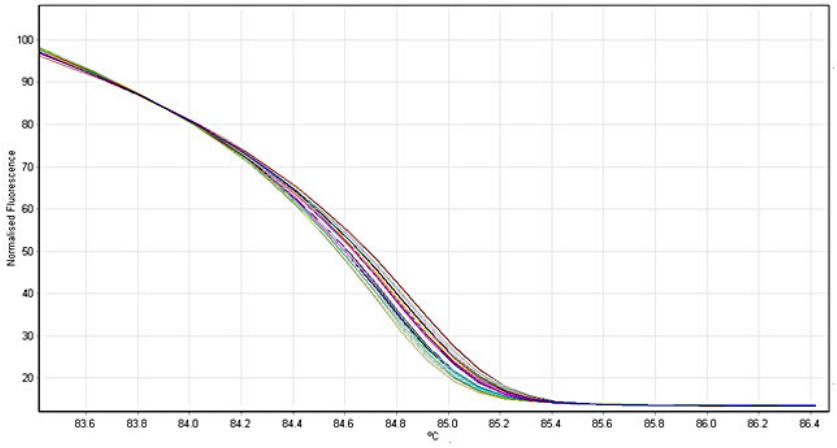


Figure 5.3 HRM normalized graph of the results obtained from the screening of exon 2 of *SNCA* for the A30P mutation.

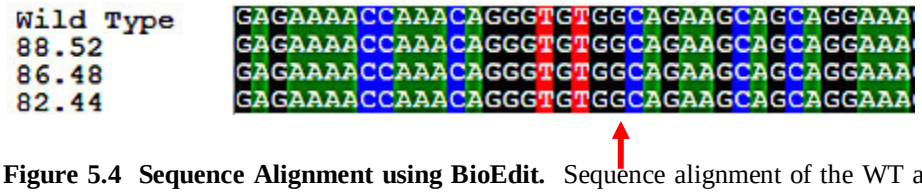


Figure 5.4 Sequence Alignment using BioEdit. Sequence alignment of the WT and patient samples for the A30P mutation in exon 2. None of the patients were found to have this variant and the position of the A30P variant is shown by the arrow.

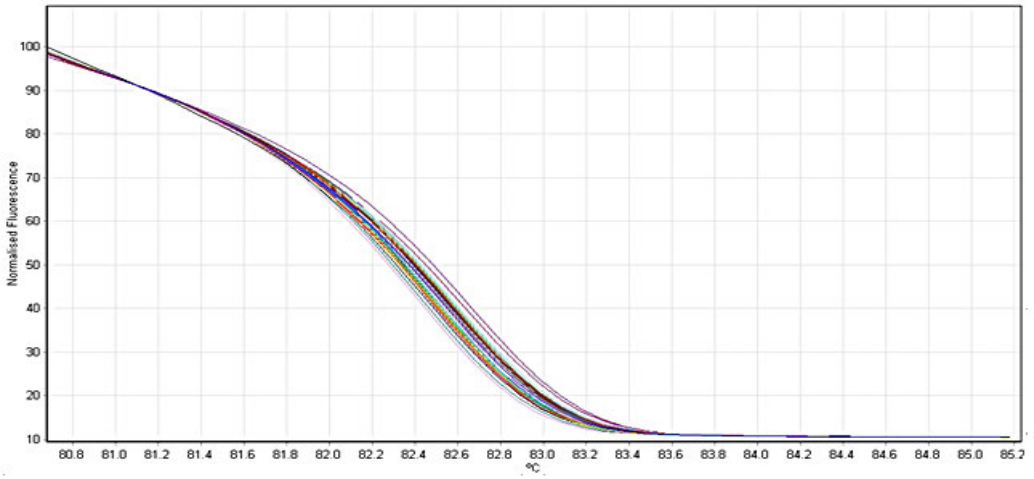


Figure 5.5 HRM normalized graph of the results obtained from the screening of exon 3of *SNCA* in order to determine whether or not the PD patients had any of the known mutations in this exon.

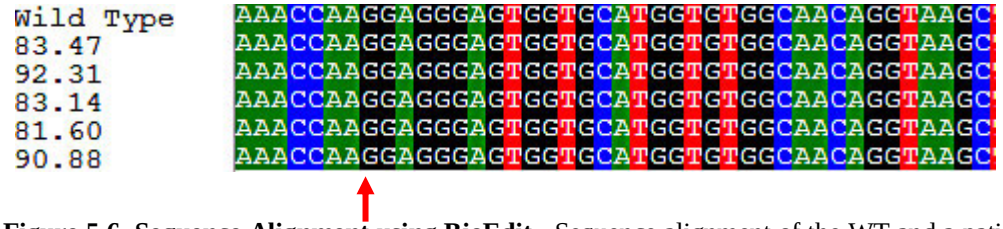


Figure 5.6 Sequence Alignment using BioEdit. Sequence alignment of the WT and a patient samples for the E46K mutation in exon 3. None of the patients were found to have this variant and the position of the E46K variant is shown by the arrow.

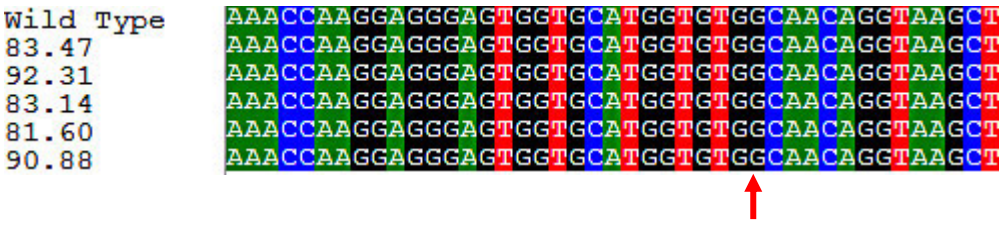


Figure 5.7 Sequence Alignment using BioEdit. Sequence alignment of the WT and a patient samples for the A53T mutation in exon 3. None of the patients were found to have this variant and the position of the A53T variant is shown by the arrow.

5.4 Discussion

Mutations in *SNCA* are rare and result in autosomal dominant, early-onset PD with the disease progressing rapidly in affected patients (Miller et al. 2004). Only three missense mutations (A30P in exon 2; E46K and A53T in exon 3) have been identified in this gene as well as whole gene duplications and triplications (Polymeropoulos et al. 1996; Zarranz et al. 2004). In the present study, we investigated whether these three mutations were present, but none of the 143 patients screened harboured any of these mutations or any other sequence variants.

Previously, our group had screened individuals for the point mutations in exons 2 and 3 as well as copy number variations using MLPA and one of the patients with a triplication of the entire gene was identified (Table 1.4; Keyser et al. 2010). Therefore, in total, we have identified mutations in *SNCA* in only 0.4% of our patients (1/262).

The missense mutations that have been identified in *SNCA* are all found in the N-terminal amphipathic region which contains six, imperfect eleven-residue repeats, denoted XKTKEGVXXXX (Sode et al. 2007). The KTKEGV is a highly conserved hexameric motif and has been shown to maintain the natural unfolded state of α-synuclein and this subsequently prevents fibril formation in the cells (Sode et al. 2007). Mutations in the amphipathic region, but particularly E46K as it is found in the KTKEGV motif, are thought to contribute significantly to disease pathogenesis because these defects prevent α-synuclein from maintaining the unfolded state and therefore results in protein accumulation and subsequent disease.

Limitations of the current study were that positive controls for the three mutations were not included in the HRM analysis, the fact that not all of the exons of *SNCA* were screened and that the investigation of duplications and triplications of the gene were not included in the study. Therefore, future work should address these particular shortcomings.

In conclusion, none of the 262 South African patients have point mutations in exons 2 or 3 of *SNCA*. However, the remaining exons should be sequenced to determine if novel mutations exist in these patients. Although mutations in *SNCA* are a rare cause of PD worldwide, due to its expression in presynaptic vesicles and it being an integral component of Lewy bodies, it remains an important gene for mutation screening.

Taken together, the results from Chapters Two to Five imply that mutations in the known PD genes *Parkin, LRRK2, SNCA* and *PINK1* that have been reported at relatively high frequencies in other global populations, are not a major cause of the disorder in the South African patients that were screened. Therefore, other approaches such a whole genome or whole exome sequencing is necessary to identify possible novel PD genes in our patients. The identification of a novel PD-causing gene(s) is important as this may shed light on the pathways leading to the disorder and also ultimately may lead to the development of improved therapeutic modalities.

Chapter 6: Investigation of a 16bp indel (g.6_+10 del) in the *DJ-1* gene

Page

6.1 Introduction

Mutations in *DJ-1* (Daisuke-Junko-1) (OMIM# 602533) have been identified in autosomal recessive, early onset PD (Bonifati et al. 2003). This gene is located on chromosome 1p36.23 and is composed of seven exons which encode a protein of 189 amino acids in length (van Duijn et al. 2001; Bonifati et al. 2003). It was initially identified as an oncogene and was implicated as the causal factor for male rat infertility before its role in PD was discovered (Nagakubo et al. 1997; Kim et al. 2005). The protein product of *DJ-1* is located in the cytoplasm and belongs to the DJ-1/Thi/PfpI protein superfamily due to the presence of the DJ-1/PfpI domain (figure 6.1). This family of proteins contains a well conserved catalytic cysteine residue which lies adjacent to a histidine residue and is part of the catalytic triad (which involves Cys-His-Asp/Glu) in the proteases (Wilson et al. 2004). The cysteine residue is significant because it is modified in the presence of reactive oxygen species (ROS) and becomes sulfinic acid, an oxoacid of sulphur with a distinct structure of RSO(OH) (Canet-Avilés et al. 2004). Interestingly, all of the proteins that belong to this family are oligomers that appear to be central to the maintenance of the biochemical activity and stability within the cell (Wilson et al. 2004). The oligomerization (the formation of an oligomer from monomers, dimers and trimers) of the DJ-1/Thi/PfpI protein family has been shown to be variable and it is hypothesized that this may be associated with the functional diversity that is observed within this family (Wilson et al. 2004).

Figure 6.1 Protein domains of DJ-1. DJ-1 is composed of a single domain, namely the DJ-1/PfpI domain. This domain possesses a critical cysteine residue, which is sensitive to ROS (Taken from Moore et al. 2005).

Mutations in *DJ-1* were initially identified in two families of European origin (Dutch and Italian) both of which presented with early-onset, autosomal recessive PD (Bonifati et al. 2002; Hague et al. 2003). In a subsequent study of 185 unrelated, early onset PD cases, with Ashkenzai Jewish and Afro-Carrabean ethnicities, only two of the individuals were found to have mutations in *DJ-1* (Abou-Sleiman et al. 2006). This study was supported when a mutation was identified in only one early onset PD patient of Hispanic and Jewish Ancestry from a total PD cohort of 953 patients screened for mutations in the gene (Alcalay 2010).

Mutations in *DJ-1* are rare and account for as less than 1% of early onset PD cases (Moore et al. 2005).

Due to the scarcity of PD patiets with *DJ*-1 mutations, there is limited knowledge regarding the clinical presentations of patients with these mutations and very little data regarding genotype-phenotype (Vásquez et al. 2004). Also the molecular mechanisms which triggers pathogenesis remains unclear (Wang et al. 2012). The abnormal accumulation of proteins such as SNCA in PD patients is thought to contribute significantly to neurodegeneration and disease progression (Taylor, Hardy, and Fischbeck 2002; Bonifati et al. 2003) and cytoskeletal proteins such as DJ-1 may clear misfolded proteins, thus preventing protein accumulation and lead to the prevention of neurodegeneration (Wang et al. 2012).

DJ-1 has also been identified as a critical factor in mitochondrial function and autophagy. Autophagy is a homeostatic process which is employed by the cell in order to break down the intracellular components - this strategy is essential for normal cell functioning (Levine, Mizushima, and Virgin 2011). DJ-1 functions together with the PINK1/Parkin pathway (Section 3.1 pg 45-46) in order to carry out this homeostatic process (Thomas and Beal 2011). It has been suggested that the disruption of the removal of damaged mitochondria from the cells through mitophagy may be a central part of the PD pathogenic process (Thomas and Beal 2011). This has been supported through studies of mammalian cells, that have proven that DJ-1 functions downstream of both PINK1 and Parkin. This is due to the fact that a loss of *PINK1* and *Parkin* will still result in an altered phenotype, irrespective of the fact that DJ-1 is still present in the cell (Dodson and Guo 2007; Exner et al. 2007). DJ-1 plays an active role in alleviating oxidative stress as it can decrease the amount of ROS within the cells through a self oxidation process (Taira et al. 2004). During increased periods of oxidative stress, DJ-1 located in the mitochondria may prevent mitochondrial damage.

6.1.1 Expression Profile

DJ-1 is a cytoplasmic protein but is capable of translocating to the mitochondria when elevated levels of ROS are expressed in the cells (Nagakubo et al. 1997). It is expressed throughout the body, with significant levels of protein expressed in the pancreas, skeletal muscle and heart (figure 6.2). Importantly, DJ-1 is found in neurons and astrocytes (Lee et al. 2003; Tao and Tong 2003).

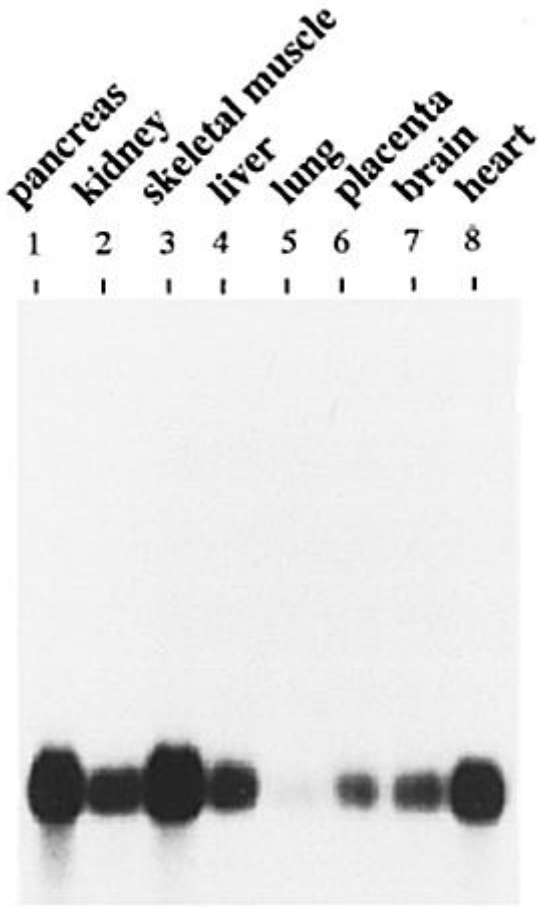


Figure 6.2 Northern Blot analysis of DJ-1 expression. Northern blot analysis of DJ-1 shows that it is found ubiquitously, with one exception being the lung (Taken from Nagakubo et al. 1997).

6.1.2 Postulated Function

To date, as mentioned, no definitive physiological functions have been defined for DJ-1 (Moore et al. 2005) but there are several that have been hypothesized. Prior to the identification of the role of *DJ-1* in PD, it was successfully linked to various biological processes such as oncogenesis, infertility, cellular responses to stressors, mRNA stability and gene transcription (Bonifati 2007). The catalytic triad involving Cys-His-Asp/Glu has been implicated as an antioxidant and therefore as a means to protect the cells against oxidative stress (Abou-Sleiman et al. 2003; Moore et al. 2005). The antioxidant properties are, however largely dependent on the cysteine residue, which is located at position 106 (Canet-Avilés et al. 2004). When oxidation takes place, the cysteine residue forms a disulphide bond, that has a reinforcing effect on the protein structure and stability (Canet-Avilés et al. 2004). It is also hypothesized that DJ-1 may prevent the aggregation of SNCA and may

additionally act as a protector of dopaminergic neurons that are exposed to ROS (Taira et al. 2004).

6.1.3 Disease-causing Mutations and Susceptibility Alleles

6.1.3.1 Disease-causing Mutations

Some of the pathogenic mutations in *DJ-1* include heterozygous and homozygous point mutations, truncations caused by nonsense mutations and exonic deletions, most of which result in a loss of function of the gene (Eeden et al. 2003). Some point mutations that have been identified in *DJ-1* result in structural changes of the protein such as a decrease in the antioxidant activity and unfolding of the protein which may lead to destabilization of the molecule (Takahashi-Niki et al. 2004; Tomiyama et al. 2009). The point mutation L166P is the most widely studied mutation and the structural and functional effects that occur as a result of this mutation, are well characterized (Tomiyama et al. 2009). It has been reported that L166P results in an altered secondary structure *in vitro* and this suggests a loss of function (Olzmann et al. 2004). Other mutations identified in *DJ-1* are M26I and D149A (Abou-Sleiman et al. 2003), A104T (Hague et al. 2003) and E46D (Hering et al. 2004). Limited information regarding the physiological and biochemical changes which occur as a result of these mutations, is currently available.

6.1.3.2 Susceptibility Alleles

The coding region of *DJ-1* is highly conserved across various species (Taira et al. 2004). Polymorphisms and susceptibility alleles have been observed and these are present in the non-coding regions of the gene (Bonifati 2007). One of the susceptibility alleles is an 18bp indel (g.168_185) which was initially identified in a Finnish PD population (Eerola et al. 2003). This indel is found in the 5'UTR and is hypothesized to affect gene expression because of its proximity to the promoter region and a transcription regulatory sequence (SP1) (figure 6.3) and has been identified in approximately 25% of PD cases (Eerola et al. 2003; Morris et al. 2003). Another variant thought to be a susceptibility factor is the R98Q variant, in exon 5 which has been found in 2% of patients vs. 0.7% in the control population (Djarmati et al. 2004). This variant was also found in a homozygous state in a control patient - suggesting that this variant is non-pathogenic (Bonifati 2007).

-1015 ggatccttct aagctcattc aagaattttg ggctttaact atttcctttg atttaacctg

-955 gtaccaggtg ccaactttag ataataggga tatctaatta cttctaaatt cctcagataa

-895 ggggcctgct tgatggtcac caggtgatct gtgctctcct taagagggaa taagacctag

-835 cgttggcaga gttctgtagg gtgactatag ttaacagtaa tctgttgtat attttaaaat

-775 gttattattg aagagagtaa ctggaatgtt cccagtataa agacaaatgt ttaaggtgat

-715 agagatctca tttaccctga tttaatcatt acacattata tgaaagtatc aaaataccac

-655 atgtacccag aaaacacata cgtctcttac atatcaataa atacaacttg agattatgat

-595 gtaaatacat ctgaccaact tggtacttat tagacttatg tgcgcagcac tgctctagtc

-535 ctgtgggtgc agcagcatca ggatcgttaa agaaaacaaa caatgctgag aaaaaaactc

-475 acacccctga gacatccggg tgtgaataaa tgcggcagag tcgcccgaga tcgggagacc

-415 aggcgtgggg gagaggtccg ggaggcctgg accagagtcc taacagacca gaggcgaaac

-355 gggaaggcgc gccagaaaag gaacaacgca aagggagcag gcgtgcacgg agcgcgaact

-295 aaggaacccc tctgacaacc ccagtccctc ggcagttcca gagaccggct cctcacggag

-235 ggtggcggta gagactgtta agccccgcgg gcgccggggc aggccggact gtgccattcg

-175 tgggggggtac catgtgggac cgagccgcct cacccagggc tgtccagcta gaaactcccc

-115 ggtgccaccc ccgcctcagt ccgaggtaga ctcggccgga cgtgacgcag cgtgaggcca

-55 aggcggcgtg agtctgcgca gtgtggggct gagggaggcc ggacggcgcg cgtgcgtgct

+5 ggcgtgcgtt cactttcagc ctggtgtggg gtgagtggta cccaacgggc cggggcgccg

+65 cgtccgcagg aagaggcgcg gggtgcaggt cagcgccagc gggggcgcgg cgcatgtgtg

+125 ggccgtggcg ctgggcggcg tgggggtgct ggacggtgtc cctgtgctgg acggtgtccc

+185 gctggctcag aaccggcgcg gggcctgggt cggggccgcc ctcgcttccg gcctcccagt

+245 cgggccctgt cgctggcgtt ggatttgact gaccgccagc gtggtggcaa cgctgaagcg

+305 tccagaatct tctgcctaac ctctcgccgg catggaactg gctagccgtt ttattaaact

+365 ctgttttgcg tggacggtaa accctccaga taatctgtaa ataggttaaa aaaaattcgg

+425 aacctcgttg agctgctgtc gttggcagtg agaactccgc gcagagagac agatgtagtt

+485 gggttgactt cagtgagggg atttccatct ttctcagtca ttaaaaaaag tgttcagaca

+545 tttaacactg ttgaccccca cacacaattt tttagtacag ttataactaa gaaaacaaaa

+605 atcccctcca aaaaattaca agttaattgc gaaagaccac atttaaattt ttgcccatga

+665 aattcagttt agtcgtttct ctgaaacagt gcttcaaaaa agactgtttc cccgcattgt

+725 gtgaaatgca ggagacccac gtacttgtat ttttaaaaaa cccatttgca acatactatt

+785 aaagttggat ttaagagaac atggtagaag aaaatctaag caatactaca ccttttagca

+845 ccctcattat gttttcatct cagagcaatt aaaactgcta tacaaatcaa cgttaagata

+905 actaaactgc tgcttttttc gtattcagtt gtctatgaaa accgtttccc taggaagtac

```
+965    ttactctgct tgaaaatgct cctaaacttt aaattttggg gtatctcagg gttgcaatga

+1025   aagttttttg aaatcttttt ttttttttt ttttaaggct tgtaaacata taacataaaa

+1085   atggcttcca aaagagctc
```

Figure 6.3 Nucleotide sequence of the promotor region of *DJ-1*. The position of the 18bp indel that has been identified in the 5'UTR of *DJ-1* is indicated in bold blue font and nucleotide base pair underlined and in bold, red font indicates the transcription start site. The 18bp indel is thought to affect gene expression because of its proximity to the start site. The SP1 site (a transcriptional regulation sequence) is indicated in bold, purple font and has a double underline. The position of the g.-6_+10del variant, the focus of the present study, is indicated by the bold, underlined font (GenBank Accession number: AB045294).

6.1.4 The Present Study

In a previous study, conducted by our group, a novel g.-6_+10del variant had been identified at the transcription start site in one of the South African Mixed Ancestry PD patients (figure 6.3) (Keyser et al. 2009). Genotyping of the family resulted in the identification of one homozygous and additional heterozygous individuals (figure 6.4). A functional study using a luciferase assay showed that the deletion reduces transcriptional levels by 40% (Keyser et al. 2009). The aim of the present study was to investigate the frequency of the 16bp deletion in the PD cohort and thereby possibly determine the role of this variant in the South African PD patients.

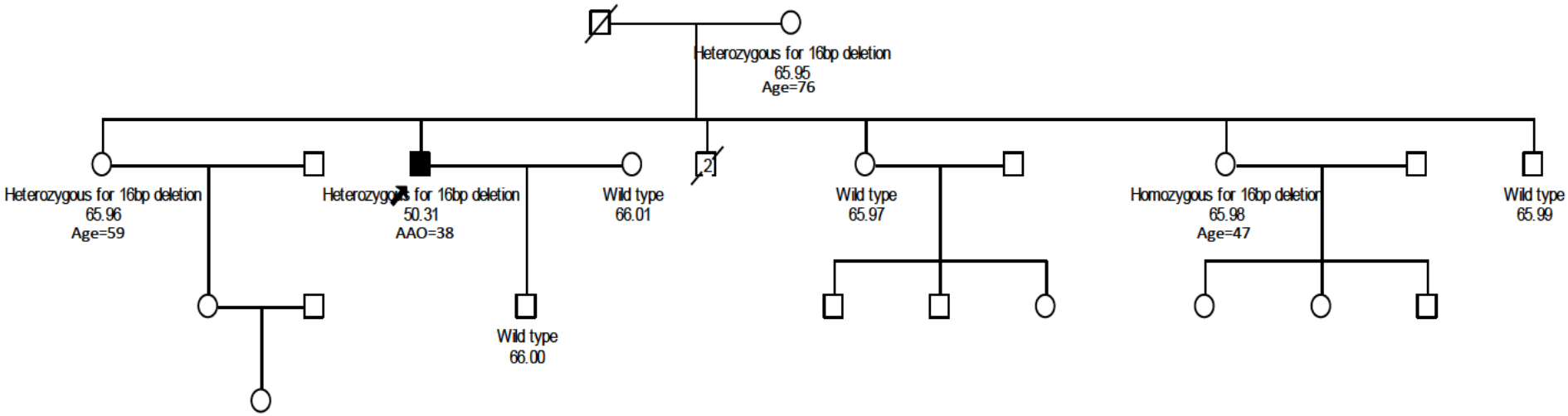


Figure 6.4 Pedigree of PD family who harbour the g.-6_+10del variant in *DJ-1*. Besides the proband, two unaffected family members are heterozygous for the deletion while another unaffected sibling is homozygous for the deletion. Four individuals were wild type as stated on the pedigree.

6.2 Materials and Methods

6.2.1 Study Participants

A total of 148 patients had previously been screened for the g.-6_+10del variant using a restriction fragment length polymorphism (RFLP) (Keyser et al. 2009). We aimed to re-genotype these patients and to include the remaining 114 patients in our cohort that had not yet been screened.

6.2.2 Genetic Analysis

A number of different techniques were employed in order to successfully genotype the g.-6_+10del variant in the South African patients. These methods included HRM, RFLPs, a Dual Labelled Probe Assay (Applied Biosystems, Foster City, USA) and finally, since none of these methods worked, a custom KASP™ Genotyping Assay was designed (LGC Genomics, Teddington, United Kingdom).

HRM Analysis

The HRM and sequencing methodology has been described in Chapter 2; section 2.2.2; pages 34-35. Specific primers had previously been designed for the PCR using Primer3 software (Rozen 2000) and the primer sequences are indicated in Table 6.1. Due to the high GC content of the fragment, the size of the PCR fragment as well as the non-specific binding of the primers to the target sequencing of DNA, a number of different primers needed to be designed for the HRM analysis of this variant (Table 6.1). Various additives (DMSO, Formamide and Betaine) were included at 5% concentrations to each of the PCR reactions.

Restriction fragment length polymorphism (RFLP) analysis

The PCR primer set labelled "A" generated a PCR product of 234bp in length (Table 6.1) was used for this analysis. Following PCR, a total of 5 units of *KpnI* (Promega, USA) was added to 8µl of the PCR product, and incubated overnight at 37°C. The digested products were then electrophoresed on 4% agarose gels as well as on a 12% polyacrylamide gels (Appendix V). Both gels were used to compare the resolution of the digested fragments. Silver staining was used to visualize the digested PCR products on the polyacrylamide gels and ethidium bromide was used to visualize the digested PCR product on the agarose gels. *KpnI* cuts at position 103 of the PCR fragment (figure 6.5), therefore it is expected that in a homozygous wild type individual, fragments of 131bp and 103bp would be visualized, whilst

in a patient with the heterozygous g.-6_+10del variant, three fragments would be generated 131bp, 103bp and 87bp - the 87bp is as a result of the g.-6_+10del variant that is missing from the sequence. In an individual homozygous for the deletion, two fragments of 131bp and 87bp would be visualized due to the fact that the g.-6_+10del variant is missing on both of the alleles.

Table 6.1 Primers designed for HRM analysis for the *DJ-1* gene.

Primer Name	Primer Sequence (5´-3´)	%GC	Tm (°C)	PCR conditions (Ta in °C)	Size of PCR fragment (bp)
Primer set A	For: caa ggc ggc gtg agt ctg	67	61.1	62 (Formamide and Betaine)	234
	Rev: gtc cag cac agg gac acc	67	60.0		
Primer set B	For: acc cag ggc tgt cca gct	67	63.0	60 (Betaine)	321
	Rev: ggt gtc cct gtg ctg gac	67	60.0		
Primer set C	For: aac gca cgc cag cac gca cg	70	69.6	65 (DMSO)	192
	Rev: ggt gtc cct gtg ctg gac	67	60.0		
Primer set D	For: gta cca ctc acc cca cac cag	62	62.3	62 (DMSO and Betaine)	195
	Rev: gtc cag cac agg gac acc	67	60.0		

ccaaggcggcgtgagtctgcgcagtgtggggctgagggaggccggacggcgc**gcgtgcgtgctggcg**

KpnI cut site

tgcgttcactttcagcctggtgtGGGGTGAGTGGTACCCAA/CGGGCCGGGGCGCCGCGTCCGCAGG

AAGAGGCGCGGGGTGCAGgtcagcgccagcgggggcgcggcgcatgtgtgggccgtggcgctgggcgc

gtgggggtgctggacggtgtccctgtgctggac

Figure 6.5 Nucleotide sequence of the 234bp PCR product of DJ-1 used for RFLP. The position of the g.-6_+10del variant is indicated by the blue, underlined section of sequence. *KpnI* cuts at position 103, as indicated on the sequence.

Dual Labelled Probe Assay

The samples were also genotyped on an ABI Prism 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, USA) using a custom designed DNA Dual Labelled Probe Assay (Integrated DNA Technologies, Coralville, Iowa). This assay works in exactly the same way as a conventional TaqMan® or 5' exonuclease assay. Probes that are composed of two fluorophores are used. One of these fluorophores is called a quencher (Q) and while the probe is intact, it causes a reduction of the fluorescence from the second fluorophore (F) which is referred to as the reporter. The fluorophore is located at the 5' end of the probe and the quencher at the 3' end. Once the probe has bound to the template DNA, after denaturation has taken place, the primers will anneal and the *Taq* polymerase will add the necessary nucleotides. Through its 5' exonuclease activity, *Taq* is able to cleave off the probe and the quencher and reporter will then be separated. The reporter is then able to emit its energy, which is quantified using digital software (figure 6.6) (www.bio.davidson.edu/courses/molbio/molstudents/spring2003/pierce/realtimepcr.htm). By using two probes with different reporter fluorophores, SNPs or indels can be genotyped using this approach.

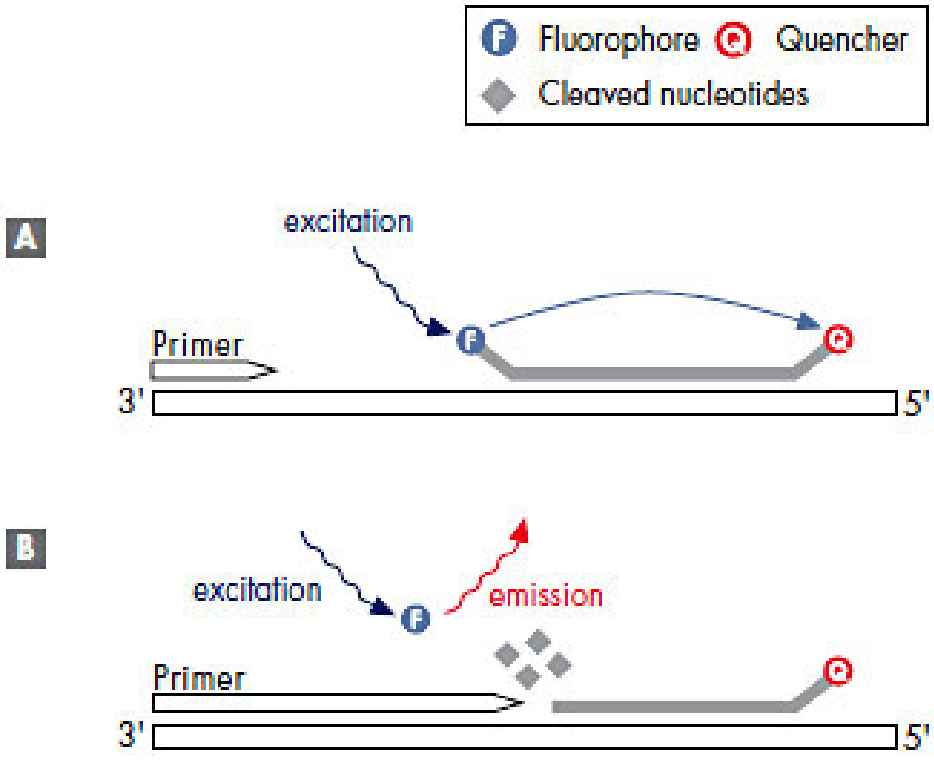


Figure 6.6 Principle behind conventional Taqman® Assay/Dual Labelled Probe Assay. Initially, both the probes and primers anneal to the target sequence and the proximity of the fluorphore (F) and quencher (Q) prevents fluorescence. When *Taq* polymerase reaches the probe, the fluorophore is cleaved from the quencher and the fluorescent signal can be measured. The signal obtained is directly proportional to the amount of PCR product produced (Taken from www.qiagen.com/hb/quantitectmultiplexpcr).

The PCR primer/probe sequences that were designed for detection of the 16bp deletion in *DJ-1* are indicated in Table 6.2. A total reaction volume of 25µl was set up for each of the samples using the real-time PCR kit, QuantiFast® Multiplex PCR kit (Qiagen, Hilden, Germany). The reaction mixture was comprised of 12.5µl QuantiFast® Multiplex PCR kit (Qiagen, Hilden, Germany), 20x primer/probe mix (20x primer probe mix is composed of 10µM forward primer, 10µM reverse primer and 4µM probe), 10ng template DNA and 7.0µl RNAse free water. The PCR was carried out on the ABI 7900HT (Applied Biosystems, Foster City, USA) and the cycling conditions were as follows: an initial enzyme activation step for 5 min at 95°C, followed by two-step cycling for 45 cycles with a denaturation step at 95°C for 30 s and an annealing/extension step of 60°C for 30 s. Allelic discrimination was performed on the ABI Prism 7900HT using the end-point analysis which was carried out using the Sequence Detection System (SDS) 2.4 software that has a 95% confidence level. This software allows for the fluorescence of the samples to be detected and calibrated and subsequently performs automatic allele calling through the generation of allelic discrimination plots.

Table 6.2 Primers and probes designed for the Dual Labelled Probe Assay.

	Sequence (5´-3´)	%GC	Tm (°C)	PCR conditions (Ta in °C)	Size of PCR fragment (bp)
PCR Forward Primer	ggc tgt cca gct aga aac tcc	57.1	60.4	60	187
PCR Reverse Primer	gtt ggg tac cac tca ccc	61.1	56.9		
Probe (wild type)	HEX-tga acg cgc gcc gtc	73.3	-	-	-
Probe (16bp del)	FAM-acg cca gca cgc acg c	75.0	-	-	-

The wild type probe was labelled with HEX which fluoresces green and the probe for the g.-6_+10del variant was labelled with FAM, which fluoresces blue on the ABI 7900HT.

KASP™ Genotyping Assay

Genotyping of the g.-6_+10del variant was performed by a commercial company, KBiosciences (LGC Genomics, Teddington, United Kingdom) using a custom designed

KASP™ Genotyping Assay. This assay is based on the fluorescent resonance energy transfer (FRET) system that is an effective means for SNP and indel detection (http://www.lgcgenomics.com/kasp-overview). This system is a probe based system that allows two fluorophores, a donor and acceptor that are tagged with a specific dye, to interact when bound to a target fragment within a specific proximity to each other. The proximity of the two fluorophores results in a FRET reaction. When the donor fluorophore is in close proximity to the acceptor fluorophore and is subsequently excited by a light source, the acceptor fluorophore is excited and a fluorescence is produced. This fluorescence is directly proportional to the amount of target DNA that is amplified (http://dnasoftware.com/FretAssays/tabid/139/Default.aspx) (figure 6.7).

A total reaction volume of 20µl was set up for each of the samples in a 384 well plate. This reaction volume was composed of 10ng template DNA, 2x KASP™ Master Mix and 0.055 µl KASP™ Primer Master Mix. The plate was then heat sealed and the PCR carried out on the Hydrocycler (LGC Genomics, Teddington, United Kingdom), which supports plate based PCR reactions. KASP™ makes use of a two step PCR rather than the conventional three step PCR. The initial activation step of 94°C for 15min is necessary for Hot-Start Taq activation, followed by 10 cycles of annealing which includes 20s at 94°C and 60s at a temperature of between 65-57°C (with the temperature being decreased by 0.8°C per cycle). The final extension step is performed at 94°C for 20s immediately followed by a lower temperature of 57°C for 60s - both of which take place over 26 cycles.

Following successful PCR, plate reading was performed in order to determine the fluorescence in each well. SNP calling was performed using KlusterKaller™ software that uses dual emission data imported from a fluorescent reader to generate a cluster graph for each assay (http://www.kbioscience.co.uk/software/klustercaller.html). Algorithms are built into the software in order to determine if the genotype calls are homozygous for one allele, heterozygous, homozygous for the second allele or inconclusive.

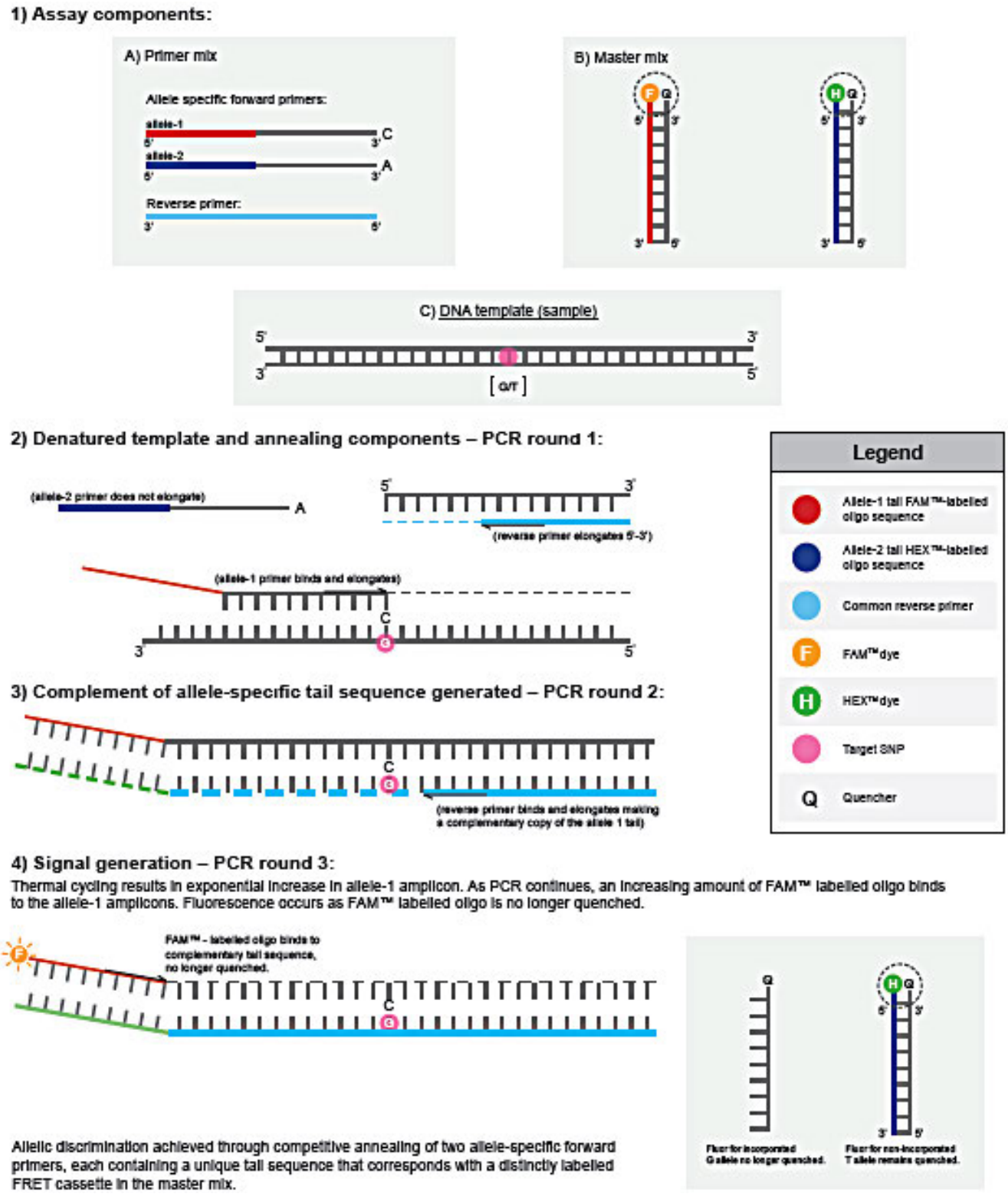


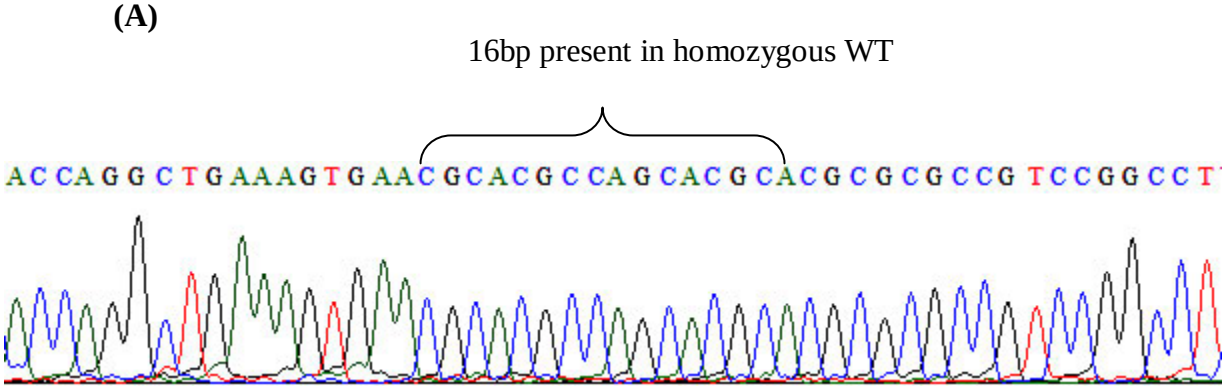
Figure 6.7 KASP™ Genotyping Chemistry. Fluorescence is emitted at a specific wavelength when the donor and acceptor (in this assay, HEX and FAM) are in close proximity to each other. Fluorescence will occur when the specific oligonucleotide is no longer quenched (Taken from http://www.lgcgenomics.com/how-does-kasp-work).

6.3 Results

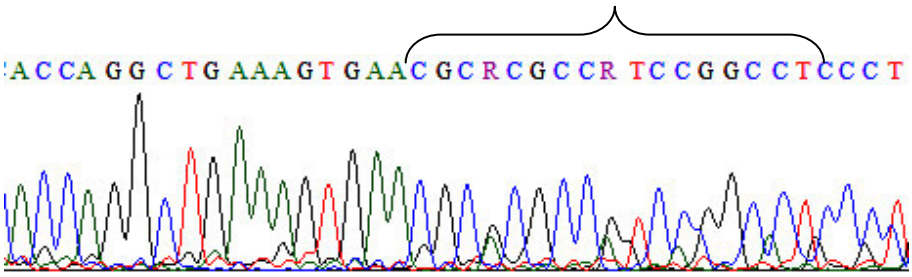
Sequencing of family member of patient 50.31

Initially, all family members of proband 50.31 were sequenced in order to verify the results that had been previously obtained. This was successfully achieved and the findings corroborated the previous results (figure 6.4). Sequencing results of a wild type, heterozygous and homozygous individual for the g.-6_+10del variant are shown in figure 6.8.

(A)



(B)



(C)

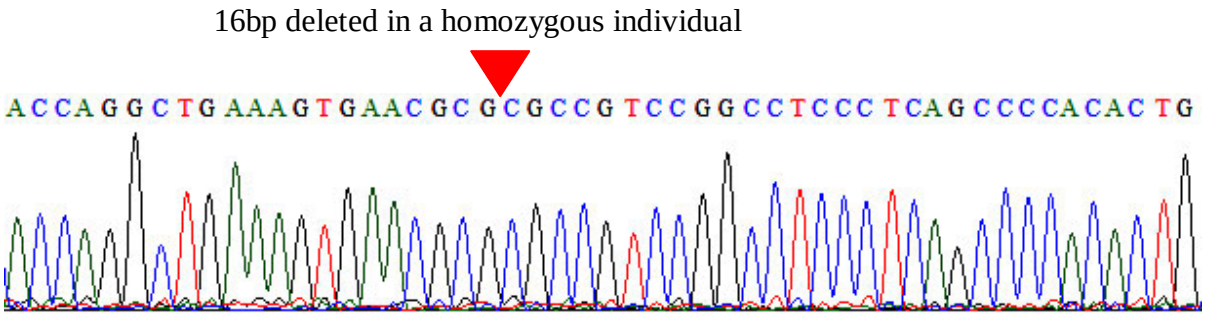


Figure 6.8 Sequence electropherograms of the family members of patient 50.31. (A) An individual homozygous for the wild-type allele; (B) An individual heterozygous for the g.-6_+10del variant; (C) An individual homozygous for the g.-6_+10del variant in *DJ-1*.

Figure 6.9 Sequence Alignment the g.-6_+10del variant using BioEdit. Sequence alignment of the homozygous WT and homozygote for the g.-6_+10del variant. The deletion can be clearly detected when the sequences are aligned.

Genotyping using HRM

Samples that were wild type, heterozygous for the deletion or homozygous for the deletion were used in the HRM experiments as positive controls using primer set A (Table 6.1). However, despite numerous attempts, the various genotypes could not be distinguished from each other using this method (figure 6.10 and 6.11). This was unexpected as a 16bp deletion should significantly alter melt profile of a DNA fragment. Designing alternative forward and reverse primers (Table 6.1; sets B, C and D) to change the size of the PCR product as well as to change the GC content of the fragment (as this is a GC rich region) also did not result in a detection of the g.-6_+10del variant. Identical results were obtained for each of the primer sets i.e. none of the genotypes could be distinguished from each other.

We therefore decided to genotype the g.-6_+10del variant using the RFLP method that had been used previously (Keyser et al. 2009).

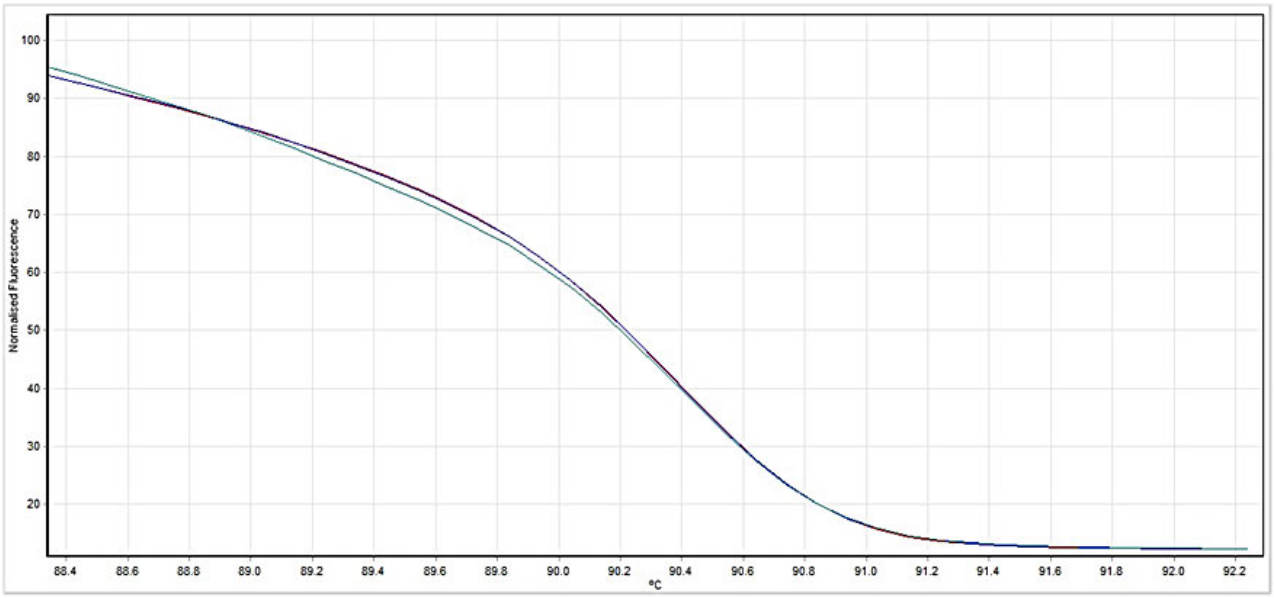


Figure 6.10 HRM normalized graph using the positive controls for *DJ-1*. A distinction between the WT, heterozygous and homozygous mutants could not be made.

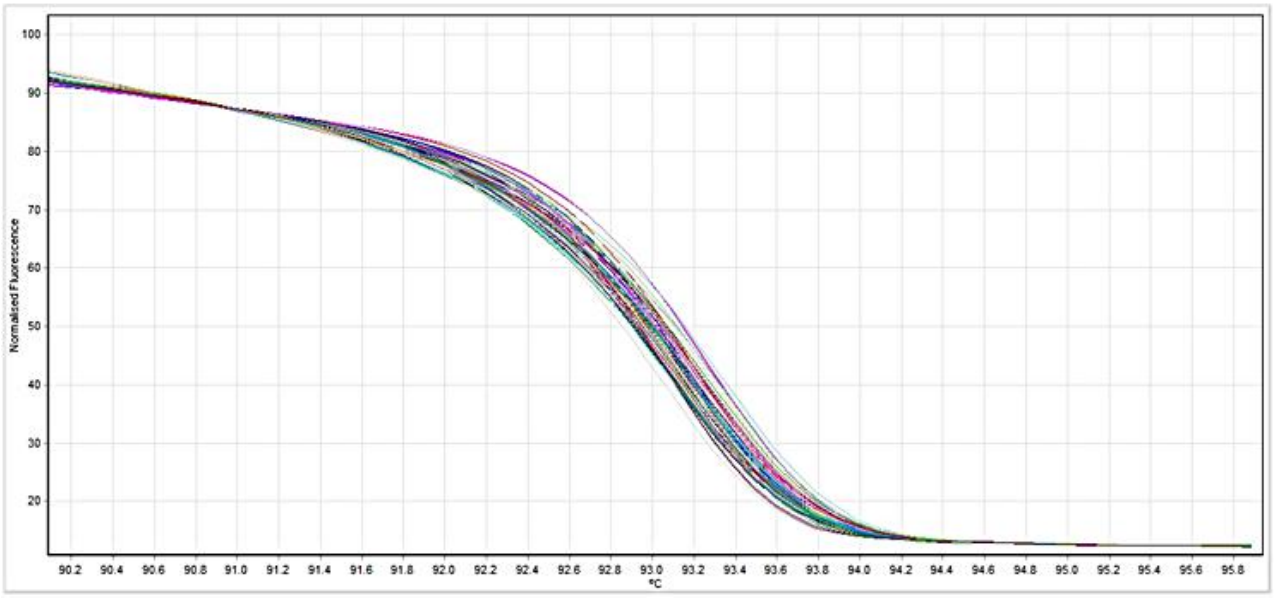


Figure 6.11 HRM normalized graph from the screening of the g.-6_+10del variant. A total of 68 patients were included in the run and no differences in the melt profiles were observed.

Genotyping using RFLP

An initial experiment, using only 8 samples, was attempted to determine whether this method could be used to successfully genotype the g.-6_+10del variant. The RFLP assay in this small sample number generated the expected results (figure 6.12). However, when the large

batches of samples were digested with *KpnI*, the assay became problematic and the positive controls could no longer be distinguished (figure 6.13). All samples appeared identical and it would seem as if the PCR did not work efficiently when cocktails were made for large batches of samples. Electrophoresing the digests on agarose or PAGE gels did not improve the resolution of the fragments.

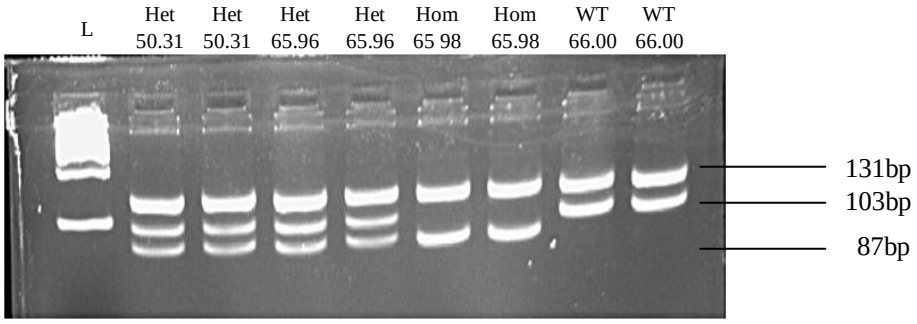


Figure 6.12 Digest of *DJ-1* using *KpnI* in the positive control samples on an agarose gel. The digest produced the expected fragment sizes. Het, heterozygous; Hom, homozygous; L, ladder (O'Range Ruler 100bp DNA ladder).

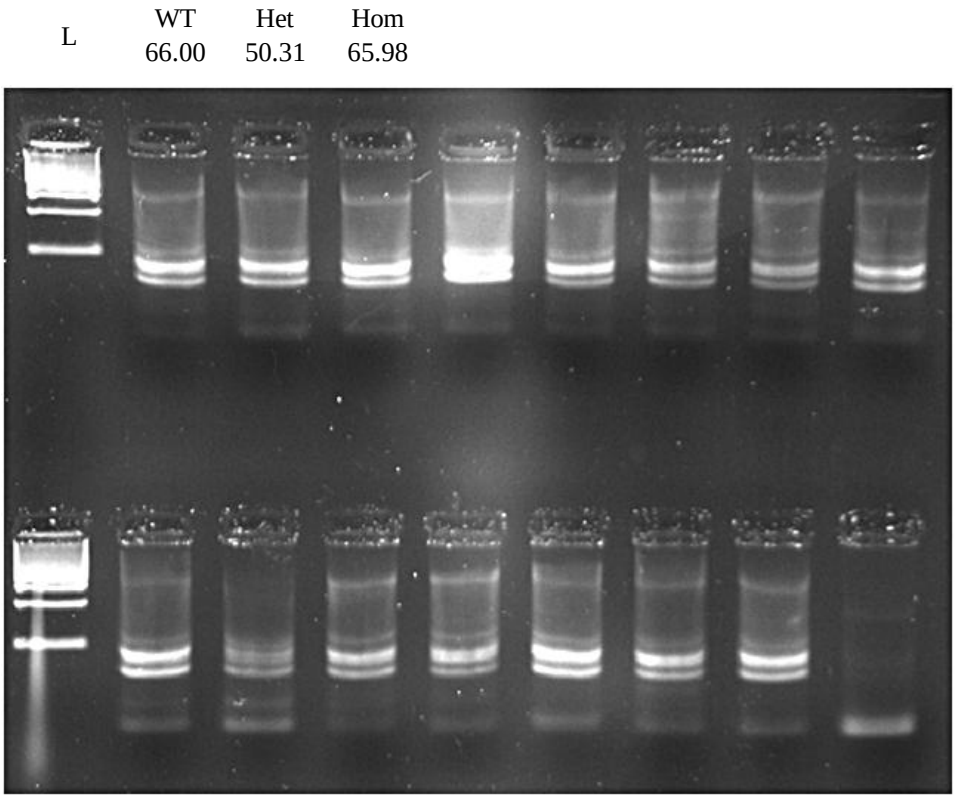


Figure 6.13 Digest of *DJ-1* fragment using *KpnI* in the patient samples. When the *KpnI* digest was performed in large batches of samples, the digest did not work. No distinction could be made between any of the samples and the controls.

Genotyping using a Dual Labelled Primer/Probe Assay

Next, since the aim was to produce a robust high through-put method for genotyping of the g.-6_+10del variant, a probe based approach was attempted. The PCR primers and probes were custom designed by Inetgrated DNA Technologies (IDT) (Coralville, Iowa, USA) and the assay was performed in our laboratory. Genotyping was carried out on the ABI 7900HT and a homozygous wild type control, a heterozygous as well as a homozygous deletion was included in each of the runs. Figure 6.14 shows the allele calling for the positive controls and patient samples. The assay was unsuccessful as the patients were all genotyped as homozygous for the g.-6_+10del variant and no distinction could be made between the homozygous wild type and the heterozygous control.

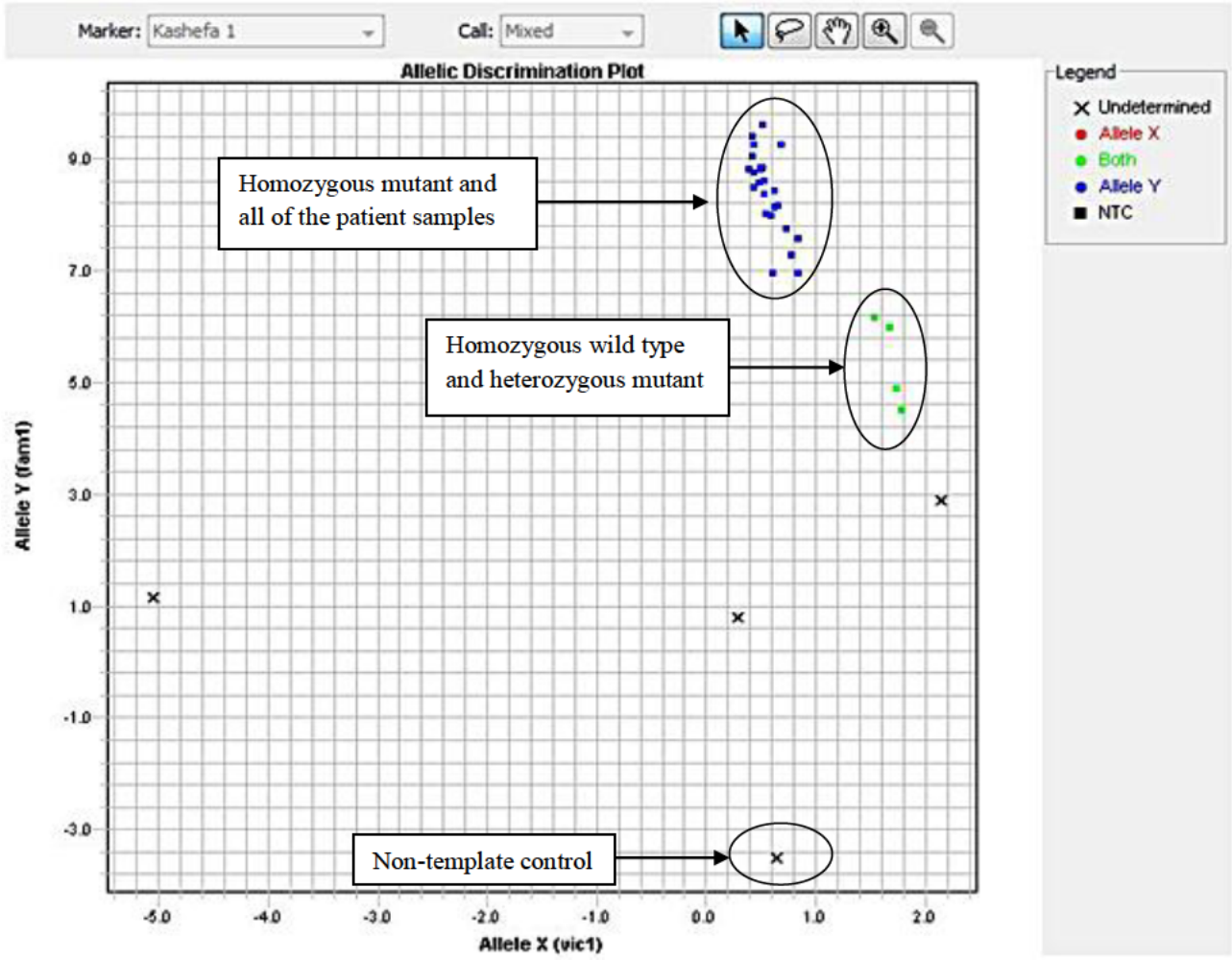


Figure 6.14 Allelic discrimination plot of a run performed on the ABI 7900HT. A homozygous positive control as well as heterozygous control were included in the run. No distinction could be made between the homozygous wild type and heterozygous control. The patient samples all grouped with the homozygous deletion sample.

Genotyping using the KASP™ Genotyping Assay

Finally, the KASP™ Genotyping Assay designed and genotyped by KBiosciences yielded promising genotyping results for the *DJ-1* g.-6_+10del variant (figure 6.15). All positive controls were successfully genotyped and none of the patient samples except for 50.31 harboured the deletion.

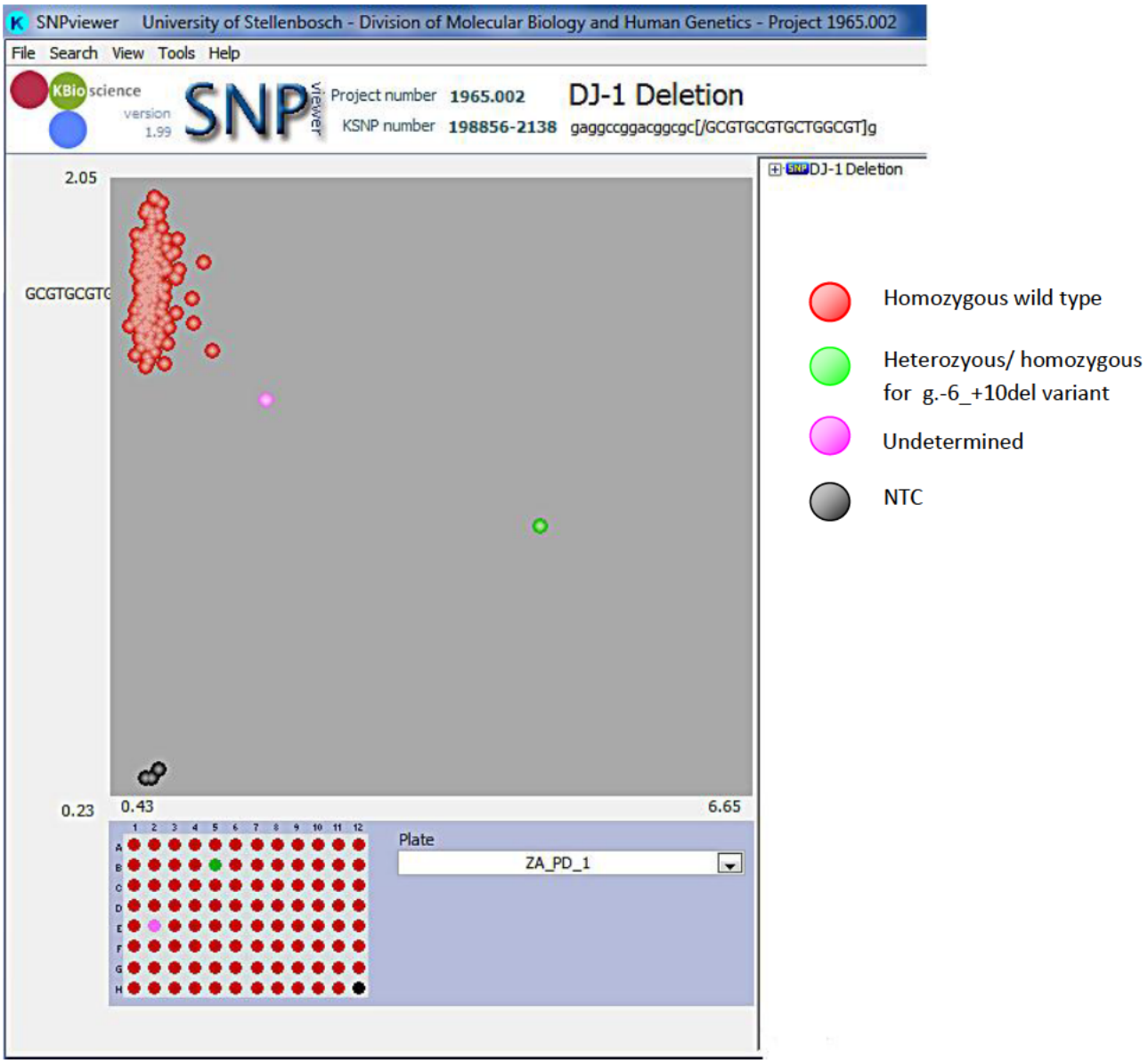


Figure 6.15 KASP™ Genotyping results viewed via SNPviewer. None of the 262 patients (except for 50.31) were identified as carriers of the g.-6_+10del variant. The majority of the patients are grouped as homozygous wild type with only one undetermined calling.

6.4 Discussion

DJ-1 has been implicated in autosomal recessive PD, however mutations in this gene are a rare cause of the disease, accounting for less than 1% of all cases (Tomiyama et al. 2009). For this reason, this gene is generally not routinely screened for mutations (Wirdefeldt et al. 2011). In the present study we did not screen the exons, but aimed to investigate the role of a functional variant in a group of South African PD patients. This variant had been previously identified in one of our affected PD patients and in unaffected family members. The significance of this discovery is the fact that this indel was found to span the transcription start site (Taira et al. 2004) and and was shown to reduce transcription of *DJ-1* (Keyser et al. 2009). This was done using a Dual-Luciferase Reporter assay; the 5'UTR region that spanned the g.-6_+10del variant was cloned into a pGL4.10-Basic luciferase vector and subsequently transfected into HEK293 and BE (2)- M17 neuroblastomal cells. It was determined that the g.-6_+10del variant caused a 47% decrease in luciferase activity in the HEK293 cells and a 60% reduction in the BE (2)-M17 cells when compared to the wild type. These findings emphasized the importance of g.-6_+10del variant in the regulation of transcription which in turn may influence the efficiency of translation and/or transcript stability (Keyser et al. 2009).

As the PCR of this region is problematic, it was decided that sequencing of the family members of 50.31 should be carried out to confirm the original results. We verified the presence of the g.-6_+10del variant in heterozygous state in the patient's unaffected mother (aged 76 years) and sister (aged 59 years) and in a homozygous state in another unaffected sister (aged 47 years) (figure 6.4). The ages of these family members are older than that of the AAO of the proband, which was documented as 38 years. This reduces the likelihood that the disorder in the proband is due to the presence of the deletion.

Numerous molecular analytical techniques were employed to identify the g.-6_+10del variant in our group of patients. It is surprising that the g.-6_+10del variant could not be detected using HRM as the variant is expected to impact significantly on the melt profile of the fragment. However, previous work in our laboratory suggests that HRM on the RotorGene 6000 can miss small deletions (unpublished data). We also hypothesized that the GC content of this fragment may play a significant role in the inability to detect the deletion. The GC content of the PCR fragment with and without the deletion is 79% and 73% respectively. It is therefore speculated that due to the high GC content and the presence of repetitive sequences,

the fragment produces secondary structures that have an impact on the melt profile, making HRM analysis ineffective under these circumstances. Furthermore, RFLP analysis failed to detect the g.-6_+10del variant in large batches of samples. We suspect that in large batches of PCR, the reaction works less efficiently and primer dimers are a significant contaminant. Therefore as the aim was to generate an efficient medium to high throughput assay for genotyping of the g.-6_+10del variant, we attempted to use a TaqMan® probe approach. The PCR fragment generated for this method was only 195bp but contained repeat sequences. The primers and probes were separate i.e. the primer was not labelled with the probe and due to the repeatative nature of the PCR fragment, it is likely that there was significant non-specific binding of the probes to the fragment, therefore resulting in the incorrect genotype calls (i.e. all patient samples were laballed as homozygous for the deletion).

Finally, a collaboration with KBiosciences was established in order to have a custom assay designed which resulted in the successful genotyping of the patients. As already mentioned, the KASP™ Genotyping Assay makes use of the FRET system to detect SNPs and indels, but is more efficient as it does not require a separation step. This method was effective because the primers are designed and validated *in silico* thereby allowing for 99% accuracy and maximum flexibility per reaction and the system is based on competative allele specific PCR, thus resulting in successful genotyping. The results obtained from the genotyping concluded that the g.-6_+10del variant was only found in one of the 262 South African PD patients.

An 18bp indel in the 5'UTR of *DJ-1* was indentified in a Finnish population with autosomal recessive, early onset PD (Eerola et al. 2003). The location of the g.-6_+10del variant lies 157bp upstream of the 18bp indel (g.168_185dup). A total of 147 patients with sporadic PD and 137 controls were genotyped for the 18bp indel and were analysed (Eerola et al. 2003). A fragment of 405bp was expected in homozygous wild type individuals and a fragment of 387bp was expected in patients with the deletion (Eerola et al. 2003). No significant differences were identified between the allele or genotype frequencies of the patients and controls and it was detemined that this 18bp indel did not act as a significant risk factor for PD (Eerola et al. 2003). To date, no case-control association study has been performed on the g.-6_+10del variant.

One of the major limitations of the present study is the fact that sequencing could not be performed on all of the patients to screen for pathogenic mutations because of financial constraints. It is proposed that for future work, the frequency of the 16bp indel should be

determined in ethnically matched controls. In addition, further functional studies should be performed on the variant to determine its effect on the antioxidant properties on *DJ-1*. Due to its position at the transcription start site, the 16bp indel warrants further investigation on transcriptional regulation of *DJ-1* and possible interaction with enhancers and repressor molecules.

Chapter 7: Screening for Pathogenic Repeat Expansions at the HDL2 locus

Page

7.1 Introduction

PD occurs across the globe, but far fewer cases have been reported in Black African individuals (Okubadejo et al. 2006). The crude incidence rate (i.e. the number of new cases of PD observed in a population of individuals over a specified length of time) worldwide has been estimated at around 17 per 100 000, whereas in Africa this number is estimated to be lower, at around 4.5 per 100 000 (Twelves, Perkins, and Counsell 2003). Similarly, the crude prevalence rate varies globally between 7 and 657 per 100 000 but in Africa this falls to between 7 and 43 per 100 000 (Osuntokun et al. 1987; Lanska et al. 1988; Haimanot 1990).

It could also be speculated that the PD phenotype in Black African individuals may be different to that of Caucasians due to environmental and cultural differences between the various populations (McInerney-Leo, Gwinn-Hardy, and Nussbaum 2004) and for this reason, many of these cases remain undiagnosed. It is also hypothesized that the disorder in Black Africans may be due to novel genetic causes or gene-environment interactions, that could subsequently affect disease development and progression (Okubadejo et al. 2006; Okubadejo et al. 2008).

Huntington's disease-like 2 (HDL2) is an autosomal dominant neurodegenerative disorder, which is found predominantly in individuals of Black African descent (Margolis 2005). Interestingly, HDL2 and PD have been shown to share many overlapping clinical features specifically in the early stages of both disorders (Bardien et al. 2007; Table 7.1). HDL2 was first described in 2001 and patients who have this disease have a trinucleotide repeat expansion (CTG/CAG) in the *junctophilin-3* (*JPH3)* gene that is found on chromosome 16q24.3 (Margolis 2001; Margolis 2005). The pathogenic repeats are found in exon 2A of *JPH3*, and are in a CTG orientation (Margolis 2003). The normal range for the CTG/CAG repeats is between 6 and 28, while affected patients have more than 40 of these repeats (Holmes et al. 2001).

HDL2 is a rare disorder, with less than 100 patients being reported worldwide (Margolis 2005). It has, to date, only been described in individuals of African ancestry (Bardien et al. 2007). In one study, all 28 American patients who were diagnosed with HDL2 identified themselves as African-American (Margolis 2005). Also, a Moroccan patient who was diagnosed with HDL2 was from a region in Morocco which has been predominantly occupied by African individuals (Margolis et al. 2004). In South Africa, a study was reported on 56 South African Black and Mixed Ancestry patients who had been referred to the

National Health Laboratory Services (NHLS)/University of Witwatersrand's genetic testing service for Huntington disease (HD) triplet repeat expansion and HDL2 (Margolis 2005). The results showed that only 38% of these patients had HD, whereas 27% had HDL2 (Margolis 2005). In another South African study, of 11 Black families (12 individuals in total), half of the individuals (6/12) were found to have mutations in *JPH3* and were subsequently diagnosed with HDL2 (Magazi et al. 2008). It has been speculated that HDL2 in Black Africans may be due to a founder effect.

Table 7.1 Comparison of the motor and non-motor symptoms of PD and HDL2.

	Parkinson's disease	Huntington's disease-like 2
Motor symptoms	Bradykinesia	Bradykinesia
	Resting tremor	Tremor
	Rigidity	Rigidity
	Postural instability	-
Non-motor symptoms	Autonomic dysfunction	-
	Cognitive and neurobehavioral problems	Cognitive and neurobehavioral problems
	Sensory and sleep disturbances	-
Causative genes	*LRRK2; PINK1; SNCA; DJ–1; Parkin; ATP13A2*	*JPH3*

The HDL2 phenotype is very similar to that of HD with progressive and severe dementia, and emotional and movement abnormalities being the predominant clinical symptoms of patients (Bardien et al. 2007). However Parkinsonian symptoms (bradykinesia, rigidity and tremor) may predominate in some families (Bardien et al. 2007).

HDL2 should therefore be considered in a wide spectrum of individuals who present clinically with abnormal movement and neuropsychiatric disturbances, so that the full spectrum of this disorder can be documented (Bardien et al. 2007; Magazi et al. 2008).

7.1.1 The Present Study

We hypothesized that the Black South African PD patients in our cohort that were excluded from the known PD genes may harbour pathogenic repeat expansions in the *JPH3* gene. Only the Black patients were screened as HDL2 expansions have to date been found exclusively in individuals of African ancestry.

7.2 Materials and Methods

7.2.1 Study Participants

Of the 262 patients, only the 17 Black patients were analyzed. These patients exhibited an average AAO of 54 years (SD = 11.72) and a range of 30-74 years of age. Approximately a quarter (24%) of these individuals had a positive family history of the disorder. The pedigrees of these 17 patients are shown in Appendix II. The DNA of one HDL2-positive individual with 42 repeats was included in the analysis as a positive control.

7.2.2 Genetic Analysis

The *JPH3* gene was screened for triplet repeat expansions and genotyping was carried out using fluorescently labelled (FAM) primers that had been designed to flank the CTG/CAG triplet repeat (Table 7.2). A total PCR reaction volume of 25µl reaction was set up for each sample and the reaction mixture consisted of the following reagents: 20µM each of the forward and reverse primers; 2.0mM $MgCl_2$ (Bioline, UK); 1x NH_4 buffer (Bioline, UK); 0.25 units BIOTAQ DNA polymerase (Bioline, UK) and 10ng template DNA. The PCR conditions were as follows: an initial denaturation step at 94°C for 5 min; 30 cycles with conditions of denaturation at 94°C for 30 s, 58°C annealing temperatures for 30 s and extension at 72°C for 45 s. The PCR products that were obtained were subsequently electrophoresed on an ABI 3130xl Genetic Analyser (Applied Biosystems, Foster City, USA) and analysed using GeneMapper Software version 3.7 (Applied Biosystems). Sanger sequencing of a selected number of patients was then used in order to confirm the counting of number of repeats. This sequencing was performed using the BigDye Terminator Sequence Ready Reaction kit version 3.1 (Applied Biosytems).

Table 7.2 Primers designed to detect CTG/CAG repeat expansions in the *JPH3* gene.

Gene	Primer Sequence (5´-3´)	%C	Tm (°C)	PCR conditions (Ta in °C)	Size of PCR fragment (bp)
JPH3	For: gga atc tcg tct ttc agt gg Rev: FAM-tga gga gtg gat atc gga gag	50.0 52.4	55.9 57.8	58	242 (14 CTG repeats)

7.3 Results

All 17 Black patients were successfully PCR-amplified and analysis on the ABI 3130xl Genetic Analyser facilitated distinction between alleles in the normal range and those in the pathogenic expanded range (figure 7.1 and figure. 7.2). A total of five individuals with homozygous genotypes were sequenced in order to verify that the repeats were counted accurately. Figure 7.3 shows the sequence of an individual who is homozygous for 14 CTG/CAG repeats. It should be noted that the repeats are uninterrupted.

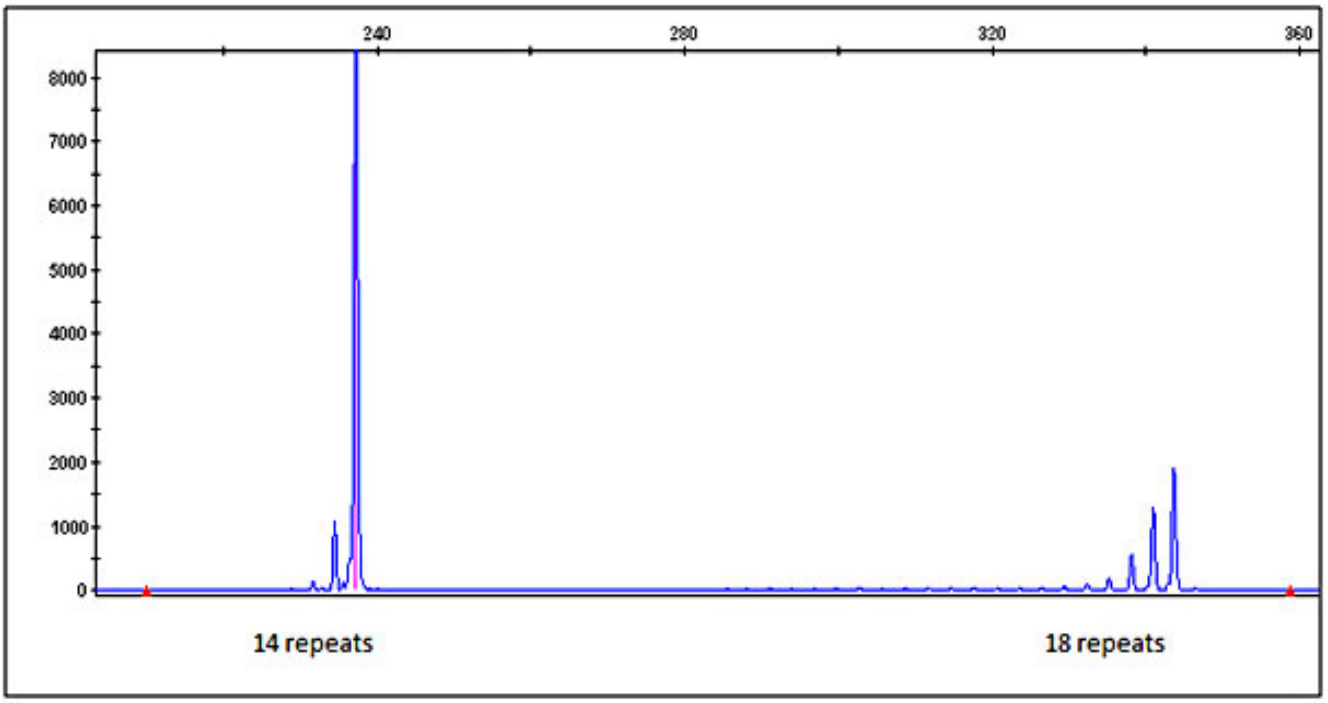


Figure 7.1 Chromatogram of the CTG/CAG repeats in *JPH3* in an unaffected individual. The two alleles comprising of 14 and 18 repeats are clearly visible. This individual is thus unaffected due to the fact that the number of repeats does not fall into the pathogenic range of more than 40. (More than one peak is visible for each allele- this is possibly due to slippage).

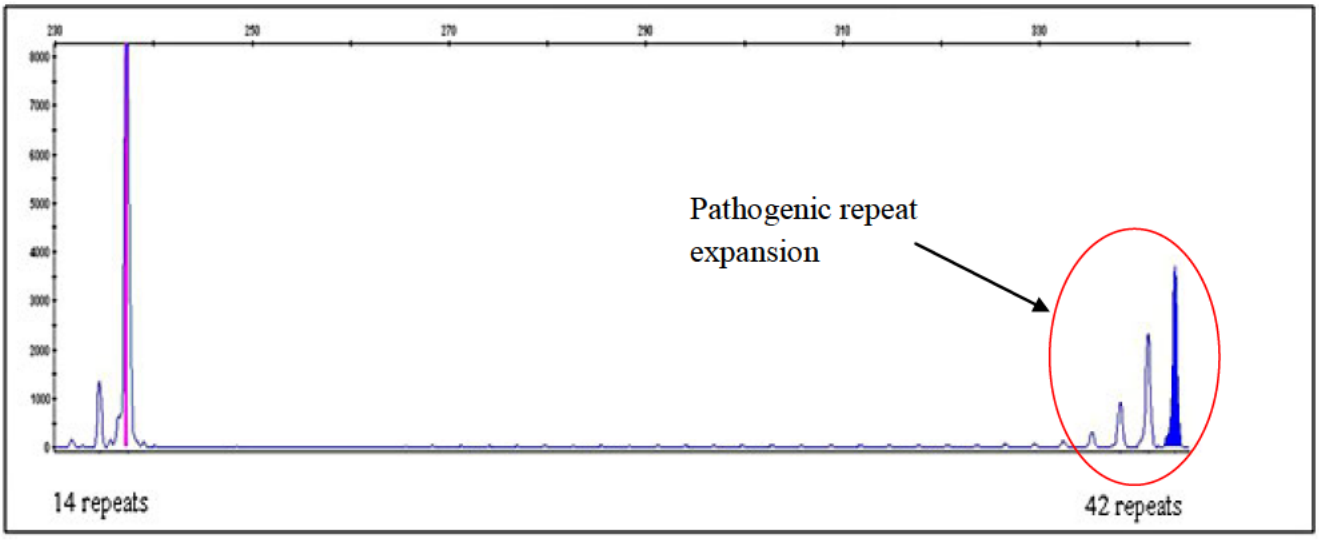


Figure 7.2 Chromatogram of the CTG/CAG repeats in *JPH3* in an HDL2-positive individual. The affected individual has two alleles, one allele of 14 repeats and another in the pathogenic range of 42 repeats.

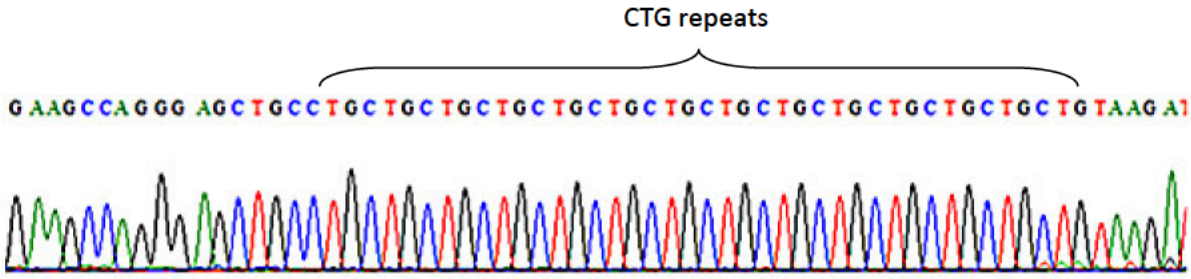


Figure 7.3 Sequence chromatogram of CTG/CAG repeats within the normal range of 14 repeats in the *JPH3* gene.

It was determined that none of the 17 Black PD patients had pathogenic repeat expansions in the *JPH3* gene. The number of repeats ranged between 8 and 18 in patients, with the mode being 14 repeats (figure. 7.4).

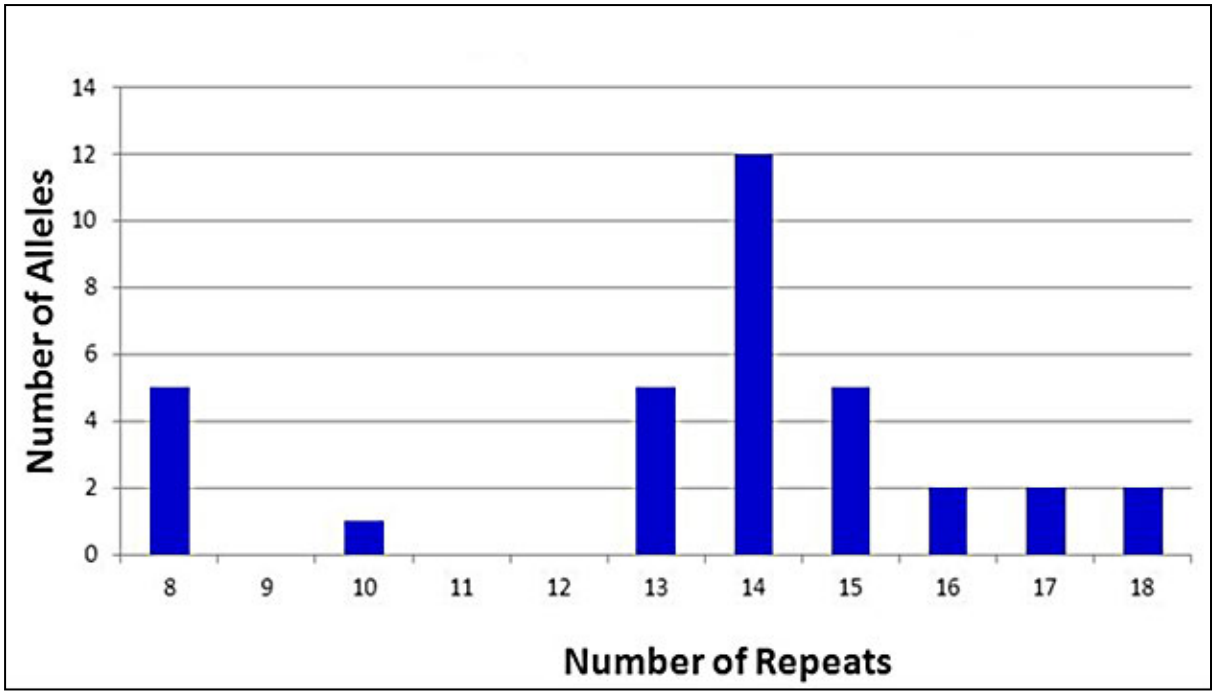


Figure 7.4 Graphic representation of the number of repeats in the *JPH3* gene in the 17 Black PD patients. It was determined that the mode was 14 repeats and the highest number of repeats in all of the patients was 18, which falls into the normal non-pathogenic range.

7.4 Discussion

In the present study we had hypothesised that Black African patients with a diagnosis of PD or Parkinsonism may harbour pathogenic repeat expansions in the gene for HDL2. HDL2 is due to repeat expansions (>40 repeats) in only one gene, *JPH3*, and therefore we screened this gene to determine the number of repeats in our patients. All 17 patients were excluded as having pathogenic expansions in this gene. Therefore, further work is necessary to determine the cause of the movement disorder in this group of patients.

Although none of the limited number of patients that we had screened had pathogenic repeat expansions in *JPH3*, it remains plausible that individuals with Parkinsonism may have pathogenic expansions in this gene as there are numerous overlapping clinical symptoms between these two neurodegenerative disorders. Due to the fact that HDL2 is found predominantly in individuals of African ancestry and appears to be relatively common in South African Black patients (Magazi et al. 2008), we propose that all Black individuals with a diagnosis of idiopathic PD or Parkinsonism be screened for repeat expansions in *JPH3*.

The results obtained from the present study cannot, however be considered representative of the Black population, particularly those who have been diagnosed with PD or Parkinsonism. The Black patients recruited for the study were predominantly from the Xhosa-speaking individuals. This ethnic group comes mainly from the Eastern Cape province where the infrastructure surrounding public health care in this province and no state-employed neurologists pose major problems regarding the identification of PD, leaving many patients undiagnosed or misdiagnosed (Prof. J. Carr, personal communication). Throughout the course of the present study, it was noted that there is an urgent need to expand the study, increase patient numbers from the Black African populations and from various provinces in South Africa for a better representation of the different ethnic groups.

It has been shown that PD patients may harbour pathogenic repeat expansions in the genes for SCA2 and SCA3 albeit that the repeats are smaller than those found in spinocerebellar ataxia patients (Gwinn-Hardy et al. 2001; Socal et al. 2009). These patients present clinically with features such as asymmetric onset, bradykinesia, rigidity and tremor - making the distinction between PD and spinocerebellar ataxia difficult. The importance and role of SCA mutations have been emphasized in numerous PD patients, particularly those of diverse ethnic backgrounds (Socal et al. 2009). Studies have reported that as many as 13% of individuals who had initially been diagnosed with PD, had pathogenic repeat expansions in SCA genes - suggesting that screening for other genes in diseases that present with Parkinsonian is important, particularly in specific ethnic groups where the frequency of the disease appears to be low (Gwinn-Hardy et al. 2001; Socal et al. 2009). In a previous study by a BSc(Hons) student from our group, all our Black patients had been excluded from the SCA2 and SCA3 loci (M.Morris, BSc Hons thesis 2010).

In conclusion, genetic testing can be used in conjunction with various clinical and imaging tests, to distinguish between disorders with overlapping clinical symptoms and phenotypes. As the underlying genetic basis of more disorders is delineated, this will aid in exclusion of disorders that share similar clinical features. Ultimately, a definitive diagnosis is critical as this would mean improved clinical management of the affected individual and identification of at-risk family members.

Chapter 8: Whole Exome Sequencing of Probands from an Extended Afrikaner Family

Page

8.1 Introduction

Next generation sequencing (NGS) is a powerful and effective strategy for discovering the genetic basis of human disorders that were previously intractable to conventional approaches such as positional cloning and linkage analysis. There are two main NGS strategies viz. whole genome sequencing and whole exome sequencing (WES) (Hedges et al. 2009). Whole genome sequencing is the characterization of the entire genome of an individual. This is a complex and costly process, as the data obtained is difficult to analyse because of the large volumes of information generated and interpretation of variants in the non protein-coding part of the genome is challenging (Ng et al. 2009; Robinson, Krawitz, and Mundlos 2011). Furthermore, the bioinformatics expertise and infrastructure needed to process whole genome sequencing data is beyond the scope of most laboratories (Bras and Singleton 2011).

On the other hand, WES is a targeted sequencing approach where only the approximately 180,000 exons (about 1.22% of the entire genome) are sequenced and screened for pathogenic mutations. WES requires only about 5% as much sequencing as whole genome sequencing and therefore is far less expensive. Besides its application to biomedical research, WES can also be applied to numerous other areas of research, including human evolution and biological research. Applications to human evolution allow for trace amounts of ancient, contaminated DNA to be sequenced, that could not be done through conventional sequencing methods such as shotgun Sanger sequencing (Burbano et al. 2010). Biological applications of WES have examined the effects of copy number variations (CNVs) in disease; low resolution methods have been used in conjunction with targeted exome capture to successfully identify breakpoints for a number of known CNVs and subsequently identify the most likely repair mechanisms employed by the biological system (Conrad et al. 2010). However, the application and success of WES has had the most impact on biomedical research with the discovery of new causal mutations for a number of Mendelian disorders such as Kabuki syndrome, Miller Syndrome, Fowler Syndrome and Freeman-Sheldon Syndrome (Ku, Naidoo, and Pawitan 2011).

Currently, a formidable challenge faced when analysing WES data is how to identify a pathogenic mutation amongst the background of polymorphisms and possible sequencing errors that are generated in each sequenced individual. This approach yields a high number of single nucleotide variants (SNVs); typically ~24,000 SNVs in African American samples compared to ~20,000 SNVs in Caucasian American samples (Bamshad et al. 2011) (Table

8.1). The difference in the numbers of SNVs identified is dependent on various factors, some of which include: locus heterogeneity, the specific population, family or ethnic group studied and the sequencing platform used (Bamshad et al. 2011). The most effective approach to analysing the data is to screen a relatively small number of affected individuals as well as a control group. The results obtained should then be analysed through a sequential filtering process (Biesecker 2010). The target group of patients should include (as far as possible) closely related affected individuals as well as unrelated affected individuals. A list of coding variants is obtained and these are compared to those found in the public databases such as dbSNP as well as those SNVs that are found in controls - any variants found in either the control group or the database are then eliminated as possible candidates. The final outcome of the filtering process is a shortlist of possible candidate disease-causing genes which can be analysed further (Biesecker 2010; Bamshad et al. 2011).

Table 8.1 Mean number of coding variants in two different populations.

Variant type	Mean number of variants (±SD) in African Americans (n=100)	Mean number of variants (±SD) in Caucasian Americans (n=100)
Novel Variants		
Missense	303 (±32)	192 (±21)
Nonsense	5 (±2)	5 (±2)
Synonymous	209 (±26)	109 (±16)
Splice	2 (±1)	2 (±1)
Total	*520 (±53)*	*307 (±33)*
Non-Novel Variants		
Missense	10 828 (± 342)	9 319 (±233)
Nonsense	98 (±8)	89 (±6)
Synonymous	12 567 (±416)	10 536 (±280)
Splice	36 (±4)	32 (±3)
Total	*23 529 (±751)*	*19 976 (±505)*
Total Variants		
Missense	11 131 (±364)	9 511 (±244)
Nonsense	103 (±8)	93 (±6)
Synonymous	12 776 (±434)	10 645 (±286)
Splice	38 (±5)	34 (±4)
Total	*24 049 (±791)*	*20 283 (±523)*

Taken from (Bamshad et al 2011)

Diseases such as PD are amenable to WES approaches with both rare Mendelian as well as common sporadic forms of the disorder being suitable for this type of analysis (Bras and Singleton 2011). For recessive forms of PD, as few as three individuals may provide significant insight into the disease when using WES; for clearly dominant disorders, as few as

four or five individuals may be sufficient to identify novel mutations (Wang et al. 2010; Glazov et al. 2011).

The success of this approach in identifying novel mutations in diseases such as PD has been shown by the identification of a novel gene *VPS35* (vacuolar sorting protein associated protein 35) in a Swiss-kindred with autosomal dominant late-onset PD (Vilariño-Güell et al. 2011). WES was performed on an affected pair of first degree cousins (Vilariño-Güell et al. 2011). The NimbleGen Sequence Arrays were used for exonic capture and sequencing performed on the Illumina Genome Analyzer. The number of variants identified in each patient was 34,754 and 29,952 respectively. Filtering was carried out using HapMap to filter the results further by eliminating additional polymorphisms and structural alterations such as CNVs were eliminated using the Database of Genomic Variants (version 6) and a total of 4,265 candidate variants remained. Upon further filtering, where variants found on the X and Y chromosomes as well as synonymous and non-coding variants that were already present in dbSNP (version 130) were excluded, a preliminary candidate list of disease-causing variants of 69 variants was identified (Vilariño-Güell et al. 2011). Notably, of these, 36 were found to be artefacts using Sanger sequencing, leaving 33 validated variants. Only two variants were identified as novel namely A1012V found in Integrin alpha X *(ITGAX)* and D620N, found in *VPS35*. Upon further screening of 4,326 PD patients and 3,309 controls, only four additional patients were identified as carriers of the novel variant in *VPS35* and none of the patients carried the *ITGAX* variant, but it was identified in one of the controls. None of the controls were found to carry the *VPS35* variant thus identifying it as a novel disease-causing mutation in PD (Vilariño-Güell et al. 2011). The use of first degree cousins and the specific filtering strategy employed was a proof of principle that WES could be used to successfully identify novel PD-causing genes.

The same variant in *VPS35* was identified in an Austrian family and in this case, two second degree cousins were selected for WES under the assumption that any shared rare variants identified in these patients would be plausible disease-causing mutations (Zimprich et al. 2011). Once the sequencing results were obtained and the sequences aligned, the SNVs were identified using dbSNP (version 131). Further filtering made use of SAMtools (version 0.1.7), which eliminated SNVs recorded in dbSNP as well as known indels (Zimprich et al. 2011). This approach resulted in only ten non-synonymous coding variants to be short-listed as candidates, possibly as more distantly-related individuals had been used (second degree cousins) as opposed to the first degree cousins which had been used for the first study

(Vilariño-Güell et al. 2011). The D620N change in the *VPS35* gene was observed in all eight patients available for genetic study but was not found in any of the 2,783 controls screened (Zimprich et al. 2011). This study provided further evidence that WES is an effective tool that can be used in the identification of novel disease genes even if the filtration processes to identify the mutations differs from study to study.

It is expected that in the near future, WES will provide significant insights into both complex and Mendelian disorders (Bras and Singleton 2011; Dixon-Salazar et al. 2012). For monogenic disorders, the best possible approach is to examine the family members of the affected individual thereby eliminating overlapping variants in unaffected individuals and pinpointing novel variants that co-segregate with the disorder (Bras and Singleton 2011). In the case of complex disorders, protein-coding genetic association studies can be carried out, thus identifying novel risk factors for the disease (Bras and Singleton 2011). To date, many WES strategies have been successful using a small number of closely related, affected individuals to find rare mutations in the same gene that is shared amongst these individuals. It is speculated that populations that exhibit genetic founder effects (many individuals sharing identical mutations by familial descent), such as the South African Afrikaner, would be especially tractable to this approach. The Afrikaner are unique to South Africa, and are descended from approximately 2,000 mainly male progenitors primarily of Dutch, German and French descent who settled in the Cape in the 17th and 18th centuries (Heese 1971). Research into the Afrikaner lineage indicates that the Dutch make up approximately 53% of the Afrikaner population, with 27% and 17% being attributed to German and French settlers, respectively. The remaining 3% is made up of so-called 'other' European population groups such as Swiss, Belgian and Polish settlers (Heese 1971). It is hypothesized that the forefathers of the Afrikaners who arrived on the continent in the 17th century contributed significantly to the Afrikaner lineage. This implies that settlers who had arrived with Jan van Riebeeck were more likely to contribute to the growth of the Afrikaner lineage than the individuals who had arrived on the continent at a later stage (Heese 1971). Due to language, religious and cultural differences, these settlers initially lived in very isolated communities. Many couples had more than four children and in subsequent generations, many consanguineous marriages followed (Heese 1971; Prof. G. Geldenhuys, personal communication).

Today, founder effects for a number of rare inherited disorders such as porphyria, Long QT syndrome and Huntington's disease are evident in the Afrikaner population (Karayiorgou et

al. 2004; Greeff 2007). Populations that exhibit founder effects are important in genetic research because the role of environmental factors as well as other genetic variations (modifiers) can be examined and could be used to explain phenotypic variations observed in affected patients (as the primary mutation often does not account for the phenotypic differences) (Brink et al. 2005). Therefore, the study of diseases which appear to have founder effects will provide a more comprehensive framework for the identification of phenotypic variations as well as provide insight into the role of additional genetic factors in the development of a disease.

8.1.1 The Present Study

Notably, it was observed that approximately 29% of the 262 South African patients in our cohort are Afrikaner. Given that founder effects for a number of different disorders have been found in the Afrikaner population, we hypothesized that a founder effect for PD may exist in this population. The aim of the present study was therefore to determine whether founder effects for PD exist in the Afrikaner population using genealogical analysis and whole exome sequencing.

8.2 Materials and Methods

8.2.1 Study Participants

Of the 262 study participants, 76 were self-identified as Afrikaner. The AAO of these patients was 51.8 years of age, SD ± 6.8 and a range of 17-76 years. A total of 68% of the patients were male and 32% were female. A total of 29% of these patients presented with a positive family history of PD.

8.2.2 Genetic Analysis

Extensive genealogical analysis was conducted by a genealogist, Prof. Gerhard Geldenhuys, on all recruited Afrikaner PD families with a positive family history of the disorder. Dutch Reformed church records, government archival records and books on Afrikaner families were reviewed in order to construct family trees for each of the families.

Genomic DNA of three of the Afrikaner probands were selected and subjected to WES in the laboratory of our collaborator, Prof. Owen Ross at the Department of Neuroscience at the

Mayo Clinic College of Medicine in Florida, USA. Exome capture was performed using the Agilent SureSelect Human All Exon Kit, a liquid-phase hybridization method that covers 1.22% of the human genome. This coverage includes all known genes, over 700 human miRNAs and over 300 non-coding RNAs, which include small nucleolar RNAs (snoRNAs) and small Cajal body-specific RNAs (scaRNAs) . WES was performed using an Illumina Genome Hiseq 2000™, by paired end reads. The input DNA was diluted and the DNA sheared (Agilent Technologies, Santa Clara, California, USA). Samples were purified using the QIAquick PCR Purification Kit (Agilent Technologies) and the quality of the DNA subsequently checked through the use of the Agilent 2100 Bioanalyser™ - DNA quality could be observed in the form of an electropherogram and samples with a distribution peak at a height of 150 ± 10% nucleotides were selected for further analysis. Further purification of the sheared DNA then took place and 'A' bases were then added to the 3' end of the fragments. The samples were then purified through the use of Qiagen MinElute PCR Purification Column (Qiagen, Hilden, Germany). The paired end adaptors were then ligated to the fragments and the samples further purified through the use of the AMPure DNA Purification Kit (Agilent Technologies). An adaptor ligated library was then generated, purified and the quality assessed and at this stage, a minimum of 500ng of library was needed for the hybridization amplification. The sequencing was then carried out following cluster amplification of the library (http://www.chem.agilent.com/Library/datasheets/Public/5990-6319en_lo.pdf).

This design covers approximately 37Mb of the genome and a minimum of thirty fold redundancy which provides a coverage of greater than 99% with a concordance of 99.9%. The sequences were then aligned using NCBI Human Reference Genome 37.2. The Illumina quality system was used to determine the presence of indels as well as point mutations. Common SNP variations were filtered using dbSNP (version 132) and structural variations such as CNVs were determined using the Database of Genomic Variants (http://projects.tcag.ca/variation/) as well data obtained from the 1000 Genomes Project (http://www.1000genomes.org/) using a paired-end method and clustering algorithms. Novel point mutations and indels were identified using CASAVA software provided by Illumina and visualized with GenomeStudio Software. The mrFAST (micro-read fast alignment search tool) was also used for the prediction of absolute CNVs for duplicated segments and genes.

The data obtained through this exome seqencing method was then subsequently analysed through the use of bioinformatics tools in collaboration with Dr. Junaid Gamieldien at the South African National Bioinformatics Institute (SANBI). A bioinformatics pipeline was designed in order to analyse the data. A semantic database called BORG (Bio-Ontological Relationship Graph) (figure 8.1) was used to filter and prioritize the variants that were shared between the three PD exomes. BORG, in combination with a list of so-called 'Parkinson's Disease related terms' (Appendix III) mined various online databases in order to identify human genes as well as orthologous genes in model organisms which are involved in PD-associated functions, pathway and phenotypes.

Primer sequences designed for screening the prioritized variants in the *CDC27, NEDD4, HECTD1, RNF40* and *TBCC* genes are shown in Appendix VI.

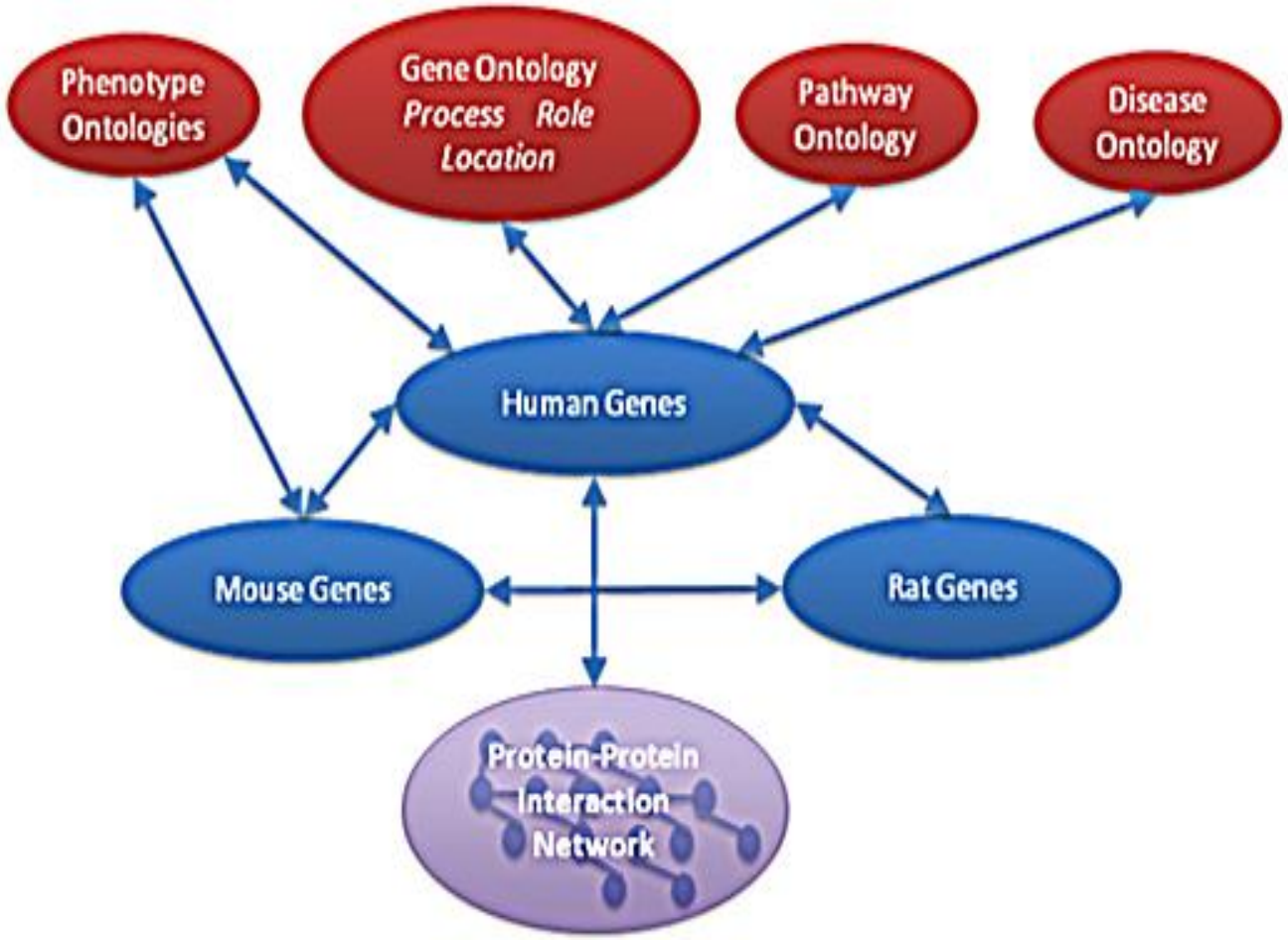


Figure 8.1 Diagram of relationships and networks on which BORG (Bio-Ontological Relationship Graph) is based.

8.3 Results

Genealogical data revealed that six of the apparently unrelated Afrikaner families could be successfully traced back to a founder couple that had immigrated to South Africa in the 1600s (Families ZA134, ZA111, ZA89, ZA106, ZA92 and ZA78; figure 8.2). The individual pedigrees of these six families are shown in Appendix IV. The founder couple to which all these patients can be linked were married in 1668 and it is documented that they arrived on the continent with the first settlers in 1652. This data therefore suggested that there is a possible founder effect for PD in these Afrikaner individuals, making them good candidates for WES as multiple affected individuals from the same 'extended family' can then be examined. Based on the genealogical data, three of the probands were selected (from families ZA111, ZA106 and ZA92; figure 8.2) and WES was performed on these individuals. Hereafter, these probands will be referred to as ZA111, ZA106 and ZA92. This analysis generated approximately 27,000 variants per patient (Table 8.2).

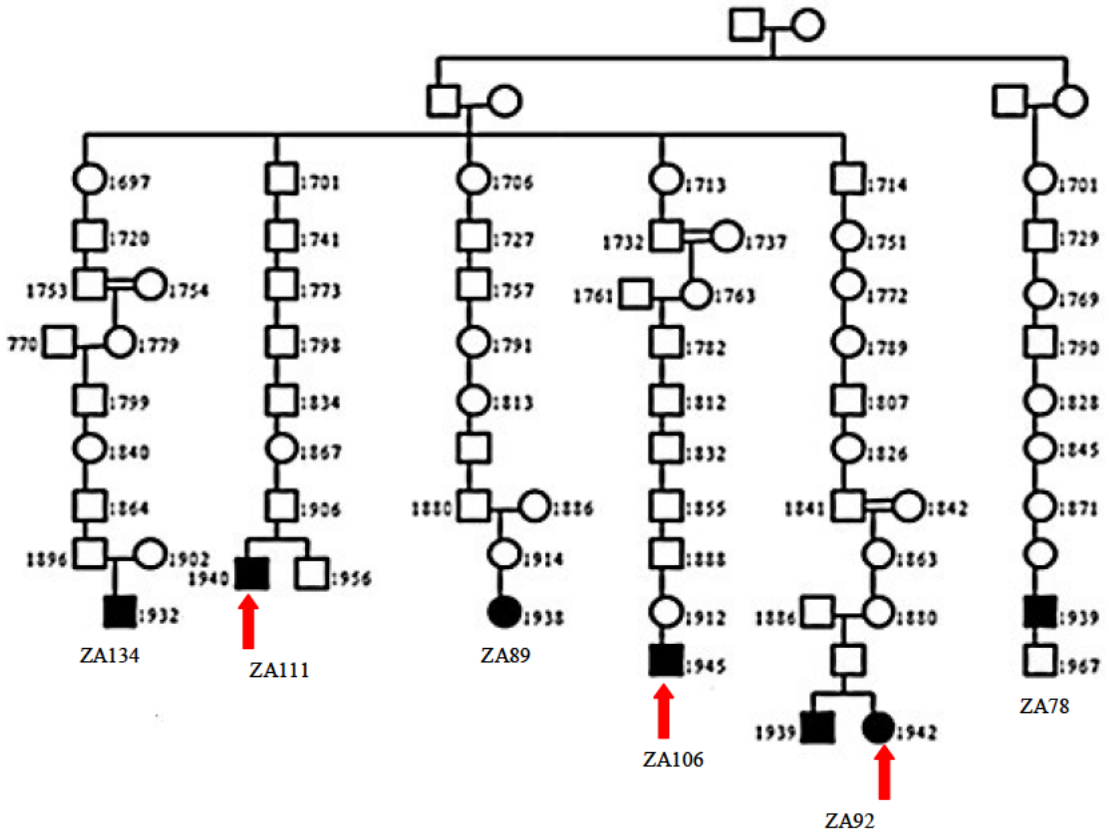


Figure 8.2 Pedigree of the six Afrikaner PD probands shown to be distantly related through genealogical studies. The probands used in this study are coloured in black and the three probands selected for exome sequencing are numbered ZA92, ZA106 and ZA111 and these are indicated by a red arrow. Males are denoted as squares and females are denoted as circles in the pedigree. Numerical values indicate date of birth and the double lines indicate consanguinity.

Table 8.2 List of variants identified in each of the three Afrikaner PD probands.

	ZA92	ZA106	ZA111
Total number of variants	26,724	27,405	25,815
Total number of SNVs	24,853	25,524	23,933
Total number of Indels	1,871	1,881	1,882
Total number of SNVs not in dbSNP	2,982	2,760	2,559
Total number of Indels not in dbSNP	878	816	850

SNVs , single nucleotide variants

Analysis of all the known PD genes revealed a number of known SNPs and one novel variant in the 5'UTR of *EIF4G1* (Table 8.3). No variants were found in *DJ-1* or *SNCA*. Notably, the findings exclude all the known PD genes from causing the disorder in these individuals. This means that it is likely that they harbour a mutation(s) in a novel PD-causing gene.

Table 8.3 Variants detected in the known PD genes in the three PD patients ZA92, ZA106 and ZA111.

	Variant	In dbSNP	Frequency (n=no. chromosomes)	Present in		
				ZA92	ZA106	ZA111
Parkin	V189L	rs1801582	(ESP) C,0.174; G,0.826; n=4550	Yes	Yes	Yes
	3'UTR +118G>A	rs6812138	No frequencies available	No	No	Yes
PINK1	Non-genic*	rs3131713	No frequencies available	Yes	Yes	Yes
	N521T	rs1042434	(ESP) C,0.677; G,0.333; n=4552	Yes	Yes	Yes
	3'UTR+37A>T	rs686658	No frequencies available	No	Yes	Yes
	3'UTR+181 C>G	rs513414	No frequencies available	Yes	Yes	No
	3'UTR+40A>G	rs6893	No frequencies	No	No	Yes

			available			
LRRK2	R50H	rs2256408	(ESP) A, 0.923; G, 0.077; n=4550	Yes	Yes	Yes
	N551K	rs7308720	(ESP) C, 0.898; G0.102; n=4510	No	Yes	Yes
	R1398H	rs7133914	(ESP) A,0.100; G, 0.900; n=4540	No	Yes	No
	S1647T	rs3459182	No frequencies available	No	Yes	No
	M2397T	rs3761863	(ESP) A,0.384; G,0.616; n=4554	Yes	Yes	Yes
	3'UTR+140 C>T	rs6673790	No frequencies available	Yes	Yes	Yes
SNCA	-	-	-	-	-	-
DJ-1	-	-	-	-	-	-
ATP13A2	P1172P	rs3170740	No frequencies available	Yes	Yes	No
	Non-genic*	rs7531163	N/A	No	No	Yes
VPS35	3'UTR+281 C>A	rs808078	No frequencies available	No	No	Yes
EIF4G1	T161A	rs1331914	(YRI) C,0.5; T, 0.5; n=2	No	No	Yes
	M432V	rs2178403	(ESP) C,0.841; T,0.159; n=4552	No	No	Yes
	5′UTR	Novel	No frequencies available	No	No	Yes

*Non-genic refers to areas on contigs that do not contain any genes; ESP, Exome Sequencing Project; YRI, Yoruba Ibidan, Nigeria;

Next, in this preliminary analysis of the data, the variants were analysed to determine what was shared across all three probands and not present in the public databases, namely dbSNP and 1000 Genomes. A dominant model of inheritance was assumed due to the parent-to-child transmission observed in family ZA111 (Appendix IV). Variants in the 5' UTR and 3' UTR as well as on the X and Y chromosomes were excluded. This produced a list of 1,135 variants and after further filtering with BORG, 175 of the variants were prioritized for further study (Table 8.4).

Table 8.4 Number of missense, nonsense and indels identified across all three patients.

	Variants shared across three probands	Shared variants that were filtered by BORG
Novel coding indels	25	3
Novel missense	1 108	171
Novel nonsense	2	1
Total	**1 135**	**175**

Of these shared variants, we further prioritized according to whether they met the following criteria:

1) were non-synonymous or nonsense variants, AND

2) the average allele frequency of the variant in the databases had to be ≤ 0.01, AND

3) the read depth had to be ≥ 30, OR

4) if the variants were predicted as being potentially deleterious according to SIFT and PolyPhen2.

Read depth refers to the number of times each base was sequenced in total. This is predetermined by the platform used and for the Illumina Human All Exon Kit, it is predicted that each base will be covered a minimum of 30x - therefore anything below this coverage may be an artefact rather than a true variant (Charier et al. 2012). These criteria reduced the list to only 21 variants (Table 8.5) of which five were chosen for Sanger sequencing verification based on their function being related to that of the known PD genes. Verification with Sanger sequencing is a necessary and important step as NGS technology is known to produce a significant number of artefacts mainly due to the short read lengths.

The results of the Sanger sequencing are shown in figure 8.3 and Table 8.6. In total, seven individuals were sequenced; all six probands in figure 8.2 as well as the affected sibling of ZA92. These results showed that two of the five variants (R625X in *CDC27* and P294L in *NEDD4*) were artefacts, as when the sequences were aligned to the Human Reference Genome 37.2, all the patients contained the wild-type alleles. This finding further emphasizes the fact that Sanger sequencing must be used for validation of WES data. The remaining three variants (L2029P in *HECDT1*, Q615R in *RNF40* and V65A in *TBCC*) were found to be real as they were confirmed in the patients. However, it was determined that

these variants are likely to be non-pathogenic. In the case of the L2029P and Q615R variants, all the patients were homozygous for the common (major) allele; in both instances the reference genome harboured the rare (minor) allele (figure 8.3C and D; Table 8.6). For V65A, individual ZA92 had the heterozygous variant but it was not present in the affected sibling of this individual or the other five probands (figure 8.3E; Table 8.6).

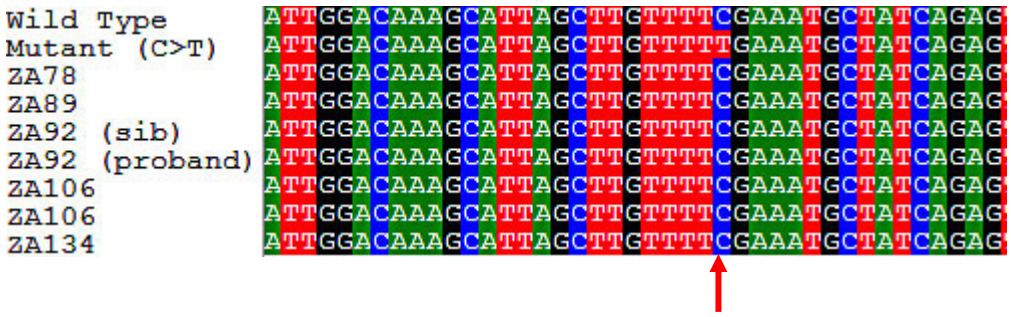
Table 8.5 List of variants prioritized for further analysis in the six Afrikaner probands.

Gene	SNV	rs#	Polyphen and SIFT prediction	Ave. Allelic Frequency (n = no of chromosomes)	Read Depth	Function
CDC27	R625X	rs77685276	Unknown	(YRI & CEU) A,0.5; G,0.5; n=4 *	157	Plays a significant role in G1 phase of the cell cycle - E3 ubiquitin ligase that controls progression through mitosis. Protein ubiquitination.
CDC27	W638R	rs74348171	Probably damaging	(YRI) A,0.5; G,0.5; n=2 *	84	
CDC27	H609Q	rs75661039	Probably damaging	(CSAgilent) T,0.007; A,0.993; n=129	171	
CDC27	H609R	rs76926116	Possibly damaging	(YRI) C,0.5; T,0.5; n=2 *	169	
CDC27	A273G	-	Possibly damaging	-	143	
CDC27	I235T	-	Possibly damaging	-	145	
DDX11	A848V	-	Probably damaging	-	54	Necessary for the E2 ubiquitination protein to the mitotic chromosomes. Also necessary for the maintenance of chromosome segregation . Cellular response to unfolded protein.
GRIN3A	R1111Q	rs76232475	Damaging (low confidence)	(CSAgilent) A,0.003; G,0.997; n=1323	98	Plays a role in the development of dendritic spines and function is dependent on Ca^{2+} and Mg^{2+} in the cell. Abnormal synaptic transmission.
NEDD4	P294L	-	Damaging	-	118	E3 ubiquitin-protein ligase that is able to accept ubiquitin from E2 ubiquitin-conjugating enzyme and then transfers the ubiquitin to a target molecule. Directly involved in ubiquiting binding.
PREX1	T1111A	-	Unknown	-	40	Functions as a Rac guanine nucleotide exchange factor which converts GDP to GTP.
RP1L1	E1328V	-	Unknown	-	100	Necessary for the differentiation of

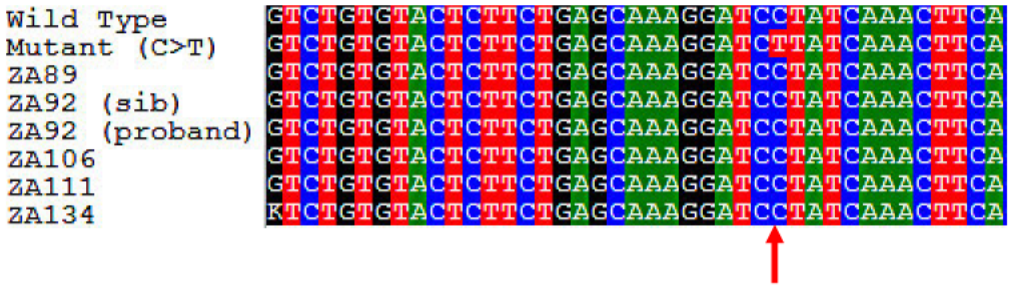
						photoreceptor cells into both rods and cones.
RP1L1	E1328K	-	Unknown	-	100	
RP1L1	T1327I	-	Unknown	-	96	
RP1L1	T1327A	-	Unknown	-	95	
RP1L1	A1319T	rs73201156	Unknown	-	97	
TRIP11	N701S	-	Unknown	-	241	Involved in the maintenance of the *cis* Golgi structure.
CLCN6	E198G	rs198400	Probably damaging	(ESP) A,0.01; G,0.99; n=4550	56	Functions as an antiporter in the chloride transport system.
GPR126	Q1127R	rs1262686	Probably damaging	(ESP) A,0.01; G,0.99; n=4550	346	Necessary for the promylenation of Schwann cells and the myelination of axons.
HECTD1	L2029P	rs1315794	Possibly damaging	(ESP) T,0.01; C,0.99; n=4550	161	E3 ubiquitin-protein ligase which accepts ubiquitin from an E2 ubiquitin-conjugating enzyme Similar functioning as NEDD4.
RNF40	Q615R	rs7195142	Probably damaging	(ESP) A,0.01; G,0.99; n=4550	64	E3 ubiquitination activity. Manages monoubiquitination of H2BK120ub1.
TBCC	V65A	rs2234026	Possibly damaging	(ESP) T,0.01; C,0.99; n=4550	78	Plays a role in the final step of the tubulin folding pathway.

SNVs highlighted were selected for verification in the six Afrikaner probands using Sanger sequencing.
* Allele frequency >0.01 but only 1 individual had been genotyped.
CSA Agilent, This population includes 662 participants of European descent from the ClinSeq project, all of whom have undergone WES using Agilent's 38Mb or 50Mb capture kit. YRI, Yoruba Ibidan, Nigeria. ESP, Exome Sequencing Project. '-'; no frequency data available.

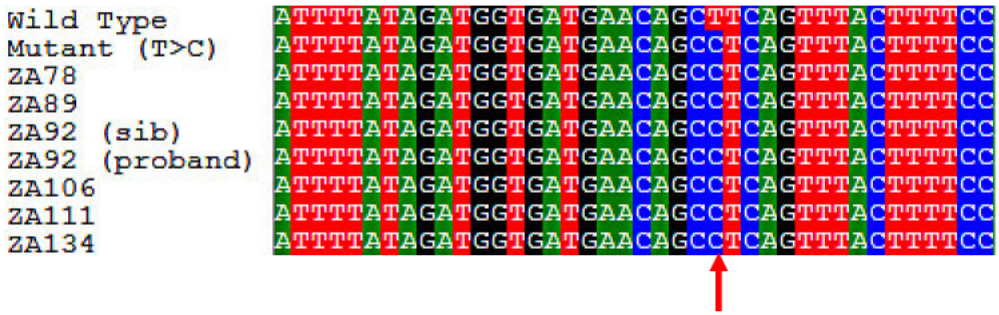
(A) Sequence Alignment of CDC27



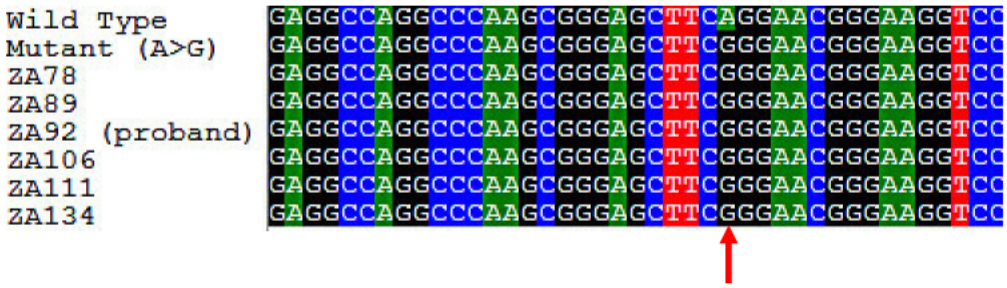
(B) Sequence Alignment of NEDD4



(C) Sequence Alignment of HECTD1



(D) Sequence Alignment of RNF40



(E) Sequence Alignment of TBCC

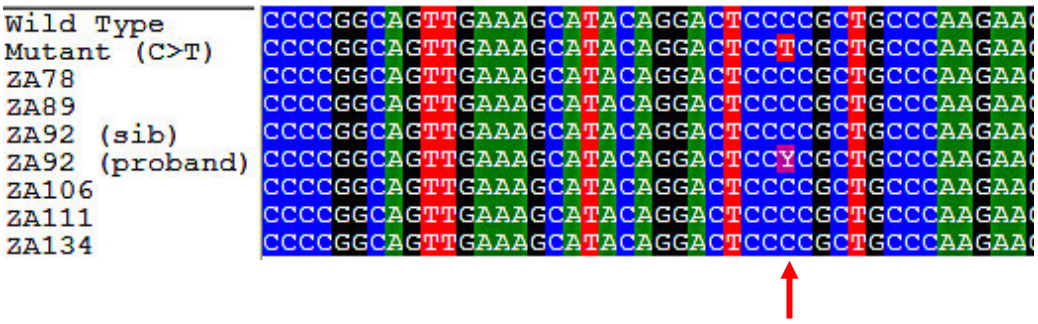


Figure 8.3 Sequence alignments of each of the five prioritized variants. The location of the variant is indicated by the red arrow. Three of the five variants were found to be real through the use of Sanger sequencing. Sib, sibling (the affected sibling of ZA92, figure 8.2). The wild type is the reference sample, the mutant is the sample in which a homozygous change would be present and the accession numbers are as follows: for CDC27 (ENSG00000004897), for NEDD4 (ENSG00000069869), for HECTD1 (ENSG00000092148), for RNF40 (ENSG00000103549) and TBCC (ENSG00000124659).

Table 8.6: Results obtained from Sanger sequencing validation.

Gene	SNV	Codon	ZA78	ZA89	ZA92 (Proband)	ZA92 (Sib)	ZA106	ZA111	ZA134
CDC27	R625X	CGA>TGA	×	×	×	×	×	×	×
NEDD4	P294L	CCT>CTT	DNW	×	×	×	×	×	×
HECTD1	L2029P	CTT>CCT	✓(hom)	✓(hom)	✓(hom)	✓(hom)	✓(hom)	✓(hom)	✓(hom)
RNF40	Q615R	CAG>CGG	✓(hom)	✓(hom)	✓(hom)	DNW	✓(hom)	✓(hom)	✓(hom)
TBCC	V65A	GCT>GTT	×	×	✓ (het)	×	×	×	×

✓, variant present; ×, variant absent; DNW, did not work; het, heterozygous; hom, homozygous; Sib, affected sibling of ZA92 (figure 8.2)

8.4 Discussion

The present study aimed to determine whether founder effects for PD exist in the Afrikaner population and this was analysed using genealogical analysis and WES. The basis for this hypothesis was the fact that approximately 30% of our PD cohort are Afrikaner and that founder effects have been implicated in this population for numerous diseases including Huntington's disease, porphyria, Long QT Syndrome and Gaucher's disease. Six apparently unrelated Afrikaner probands were found to be distantly related though genealogical analysis

and three of these patients were exome sequenced. Approximately 27,000 variants were identified in each of these individuals and through the use of bioinformatics, 21 variants in 12 genes were identified as candidate variants, that warrant further study to assess their possible pathogenicity.

This is the first WES project to be performed on South African PD patients. Our previous work had excluded the known PD genes in these three patients and this was confirmed by the WES results because no pathogenic mutations were found in any of the known PD genes (Table 8.3). From the shortened list comprised of 21 candidate disease-causing variants, five were screened using Sanger sequencing. These variants had been prioritized according to the function of the protein. Parkin (one of the predominant causes of early-onset PD) functions as an E3 ubiquitin protein ligase and four of the five namely CDC27, NEDD4, HECTD1 and RNF40 are E3 ubiquitin protein ligases. CDC27 (cell division cycle protein 27) is a functional component of the anaphase promoting complex/cyclosome (APC/C), regulated by the cell-cycle and controls progression through mitosis and the G1 phase of the cell cycle (Ahlskog et al. 2010). NEDD4 (neural precursor cell expressed developmentally down-regulated protein 4) transfers ubiquitin from an E2 ubiquitin-conjugating enzyme to specific targeted substrates (Kwak et al. 2012). HECTD1 (HECT domain containing 1) has functional characteristics identical to those of NEDD4 (Zohn, Anderson, and Niswander 2007). RNF40 (Ring finger protein 40) actively mediates the monoubiquitination of 'Lys-120' of histone H2B - this adds a tag for epigenetic transcriptional activation and acts as a precursor for histone H3 ('Lys-4' and 'Lys 79') methylation (Jääskeläinen et al. 2012). TBCC (tubulin folding cofactor C), although not an E3 ubiquitin protein ligase, plays a role in tubulin binding- this is significant as heat shock proteins Hsp60 and Hsp90 have both been directly implicated in PD pathogenesis (Ebrahimi-Fakhari, Wahlster, and McLean 2011). Despite the results obtained which confirms the presence of two of the variants in all the affected patients (L2029P in *HECDT1* and Q615R in *RNF40*), the data suggests that these variants are likely to be non-pathogenic due to their frequencies in controls in the online databases. In both cases, the more common allele (which appears at a frequency of 99% in the control population screened) was found in all of the affected PD patients thereby concluding that the SNV is likely to be non-pathogenic due to the frequency and notwithstanding the function. In the case of R625X in *CDC27* and P294L in *NEDD4*, it was concluded that the results obtained through WES were, in fact, sequencing artefacts which are due to short sequencing read lengths (typically < 100bp) particularly when SNV coverage

is low, general quality scores (such as read depth) are low or when there is strand bias i.e. when SNVs are covered only by sequences found on the same strand. The generation of artefacts, and thereby false positives, is a common feature of WES, with as many as 52% (36/69) of the prioritized SNVs being excluded as artefacts through Sanger sequencing validation (Vilariño-Güell et al. 2011). This is why Sanger sequencing must be used to verify the presence of all prioritized variants (Bras and Singleton 2011; Chahrour et al. 2012; Dixon-Salazar et al. 2012).

Limitations of employing WES as a method for novel mutation detection include the fact that a significant portion of the human genome is not examined (98.8%). Also, the definition of the so-called "exome" differs widely but this is dependent on the commercial kit that is used, CNVs are difficult to detect using this method, unannotated genes can be missed, all exons are not sequenced equally therefore variant calling may become a problem and, as mentioned, false positives and negatives are also often encountered.

Future work for this project encompasses verification of the remaining sixteen variants using Sanger sequencing and in variants that are real, checking for co-segregation and the frequency of the specific allele in the online databases. Once a putative pathogenic variants(s) is found, screening of the rest of the South African PD cohort and large numbers of ethnically matched controls should be performed in order to determine the frequency in the population as well as screening of approximately 10 000 PD patients that can potentially be obtained through collaboration with researchers affiliated to the international PD consortium, GEOPD (http://www.geopd.org), of which we are a member. Finally, functional studies are needed to determine the effect of the variant on protein function, but the nature of these experiments is dependent on the specific function of the protein involved as well as the domain in which the variant resides.

In conclusion, the use of NGS, more specifically the use of WES to identify novel disease-causing mutations is effective and highly applicable to South African Afrikaner PD patients. This specialized technique, along with appropriate bioinformatics pipelines for analysis provides a framework for the identification of disease genes that would not have been possible using traditional approaches. The present study provided important insights into PD in the Afrikaner population and the basis for identification of a possible founder mutation.

Chapter 9: General Conclusion

Page

9.1 Introduction

The impact of neurodegenerative disorders on the global disease burden is a significant one (Beddington et al. 2008; Collins et al. 2011). Global statistics indicate that brain disorders account for the highest number of disability adjusted life years (DALYs) (Collins et al. 2011). In addition, the high levels of continuous care that is needed by these patients as well as the increases in medical overheads and losses of output to the economy makes this one of the more costly groups of disorders (Beddington et al. 2008; Collins et al. 2011) . For this reason, it is quintessential that accurate projections for neurodegeneration are made in order for countries to budget appropriately for healthcare initiatives and interventions. In a study carried out on populations with movement disorders from Europe, Canada and USA, it has been projected that the prevalence of PD may increase by a factor of two between 2010 and 2050; approximately a 92% increase in the number of affected patients (Dorsey et al. 2007; Bach et al. 2011). In developing countries such as those on the African continent, the prediction of the number of affected patients is even more significant because of the expected faster increase in the number of aged individuals (Dotchin et al. 2012). In Tanzania, it is estimated that the prevalence of PD will increase by 184% by 2025 (Dotchin et al. 2012); this emphasizes the need for further research on PD pertaining to the genetic basis, as this may pave the way towards improved disease management, targeted treatment strategies and ultimately a cure.

The investigation into the molecular aetiology of PD in the South African population has provided important insights into this disease. It has been determined that five of the known PD-causing genes do not appear to play a significant role in disease development in these patients (Table 9.1) which prompted a study to identify a novel PD gene using WES.

Table 9.1 Results obtained from mutation screening of the known PD genes in the South African patients.

Gene	Exons screened	Number of patients screened*	Results- number patients with mutations*	Comment
Parkin	All 12 exons	262	7/262 (2.7%) Various exon rearrangements and point mutations	Remaining 33 patients need to be screened for exon dosage using MLPA.
PINK1	All 8 exons	262	1/262 (0.4%) 1 with Y258X	Remaining 33 patients should be screened using MLPA.
LRRK2	31, 41 and 42	262	6/262 (2.3%) 5 with G2019S; 1 with R1441C	All 51 exons should be screened.
SNCA	2 and 3	262	1/262 (0.4%) 1 with whole gene triplication	Remaining 33 patients need to be screened using MLPA. All five exons should be screened.
DJ-1	Promoter region	262	1/262 (0.4%) 1 with a 16bp del (pathogenicity unknown)	Frequency in controls from different ethnic groups to be determined. All eight exons should be screened.

***** Combination of findings from the present study and previous studies by our group.

The largest number of pathogenic mutations was found in *Parkin* (2.7%) which is rather low when compared to many other published studies (Djarmati et al. 2004; Koziorowski et al. 2010). *LRRK2* was the second largest contributor (2.3%) and all five of G2019S-positive patients are known to share a common founder haplotype (Bardien et al. 2011; R. Keyser PhD thesis 2010). For the *PINK1* and *SNCA* genes, mutations in only single patients were found. One patient harboured a 16bp deletion over the transcription start site of *DJ-1* but the pathogenicity of this variant is still undetermined. Differences in socio-economic statuses of affected patients and environmental exposures of the various South African ethnic groups as well as the lack of infrastructure regarding medical healthcare systems and novel disease genes may all impact on the low numbers of mutation-positive individuals that were identified.

WES analysis was performed on three PD probands from six Afrikaner families that were shown to be related to a common founder couple through genealogical analysis. This analysis revealed 1,135 novel variants that were shared by the PD patients, which is higher than that which is expected in three unrelated individuals (Cirulli et al. 2010). Various filtering strategies were employed that reduced the list of shared variants to 175, one of which was a nonsense variant (R625X in *CDC27*; Table 8.4). It is postulated that five of these genes, namely *CDC27, NEDD4, HECTD1, RNF40* and *TBCC* are the most likely candidates because each of them have an active role in the UPS. CDC27 plays a role in the cell cycle regulation and mediates the progression of mitosis to G1 phase (Pérez-Pérez et al. 2008); NEDD4 plays a crucial role in neural development (Kwak et al. 2012); HECTD1 plays an active role in the regulation of intracellular localization and secretion of Hsp90, a component directly implicated in PD pathogenesis (Sarkar and Zohn 2012); RNF40 is a histone H2B monoubiquitination enzyme, responsible for DNA repair (Jääskeläinen et al. 2012). Lastly, TBCC was selected based on the role that it has to play as a molecular chaperone (Hage-Sleiman et al. 2010) Proteins involved in the ubiquitination process were specifically prioritized as Parkin is a key enzyme in this protein degradation pathway. The ubiquitination of proteins is important to maintain cell homeostasis by tagging the unwanted and excess proteins within the cell (Glickman and Ciechanover 2002). However, through the use of Sanger sequencing, it was determined that all five of the SNVs in each of the genes could be excluded as possible PD-causing mutations. For this reason, further work is necessary to identify which of the variants that have been identified through WES are real as opposed to false positives and ultimately which one is the novel disease-causing mutation.

To date, most WES projects have analysed affected individuals that are closely related to one another in the quest to discover rare disease-causing mutations in the same gene in these individuals. Similarly, populations that demonstrate founder effects may be amenable to WES approaches. Founder populations are significant as they originate from a relatively limited number of individuals and have expanded over several generations in relative isolation (Karayiorgou et al. 2004). As a result thereof, it is reasonable to hypothesize that there are a limited number of independent susceptibility alleles or mutations which may segregate in these populations (Karayiorgou et al. 2004). The Afrikaner population was founded by immigrants who initially settled in the Cape in 1652 and they later spread inland, where they then became geographically isolated from other communities; it should be noted that geographic separation was not the only means in which these individuals isolated themselves from others: language, cultural and religious practices contributed significantly to the isolation of these communities (Karayiorgou et al. 2004). Consanguinity was common, particularly in the early generations and population growth can be contributed largely to reproduction and immigration (Hayden 1980; Goldman 1996; Warnich et al. 1996). The demographic history of this population is observed in the increased prevalence of certain rare Mendelian disorders such as porphyria, Long QT syndrome, Huntington's disease and Gaucher's disease (Greeff 2007; Karayiorgou et al. 2004). Additionally, low allelic diversity has been associated with these disease loci, specifically in the Afrikaner population and it is hypothesized that this is due to the genetic drift because of the population size (Karayiorgou et al. 2004). For these reasons, the South African population has been identified as a good candidate population for this pilot study on WES (Chapter 8).

9.2 Common variants underlie common diseases (CVCD) vs. rare variants underlie common diseases (RVCD) hypothesis for PD

One of the long-standing theories regarding PD pathogenesis is that there are common variants that underlie this common disorder (Goldstein 2009). This theory has been extensively studied in PD through the use of genome wide association studies (GWAS). These GWAS examine thousands of SNPs with minor allele frequencies of typically greater than 5% in unrelated affected individuals, in order to determine whether or not a variant can be associated with the disorder (Simón-Sánchez et al. 2009). The results obtained have shown that SNPs in genes such as *SNCA*, *MAPT* and *LRRK2* are major contributors to PD

susceptibility, thereby supporting previous hypotheses (Satake et al. 2009; Do et al. 2011). However, other loci have not been verified in independent GWAS studies and in general the susceptibility loci have been shown to have modest effect sizes with odd ratios of less than two (Bras and Singleton 2011). Also the exact genes and variants driving most of the associations are still unknown. Therefore, more recently emphasis has been placed on the multiple rare variants underlie common diseases (RVCD) hypothesis for PD (Tsuji 2010; Farrer, GEO-PD submission 2012).

Empirical data provides evidence for the RVCD hypothesis with variants in *LRRK2* (G2385R, R1628P and M1646T) (Tan et al. 2008; An et al. 2008, 2; Ross et al. 2011) found to have the highest population and allelic attributable risk. In many of these cases, evidence for a common ancestral founder was identified, through haplotype analysis using polymorphic microsatellite markers, and this has provided further support for pathogenicity (Tan et al. 2008; An et al. 2008, 2; Ross et al. 2011). More intensive efforts should therefore be focussed on PD families with multiple affected individuals to identify novel disease genes and mutations which can then be genotyped across large panels of patients and controls from diverse populations. The development of high-throughput NGS technologies has expedited this approach (Tsuji 2010; Farrer, GEO-PD submission 2012). The discovery of D620N, in *VPS35* is one example that supports the RVCD hypothesis. Initially, D620N was discovered in one Swiss and one Austrian kindred using WES (Vilariño-Güell et al. 2011; Zimprich et al. 2011). A subsequent study in which D620N was genotyped in 4,326 PD patients and 3,309 controls found that mutation was not found in the controls, but was identified in five of the familial and importantly, also in two of the sporadic PD positive patients (Vilariño-Güell et al. 2011). This provided the 'proof of principle' that rare mutations such as D620N is also implicated in the sporadic form of PD and justifies the study of Mendelian forms of PD. Novel NGS approaches such as WES can be an effective strategy to prove the RVCD theory in various complex disorders.

9.3 A current unifying model of PD pathogenesis

PD can be considered to be a multifactorial disorder and there are numerous factors that are hypothesized to play a role in disease development and progression. The current evidence from molecular and cellular biology has provided evidence that several pathways may be involved and continual research into the field suggests that possible targets can be identified

and used for the diagnosis of the disease before the onset of major motor symptoms (Tofaris 2012). Genetic studies have provided the necessary insight to demonstrate that PD cannot be considered as a single clinical entity, but should rather be considered as a heterogeneous group of diseases, with varying associated pathologies, clinical signs and symptoms (Corti, Lesage, and Brice 2011). By this it is meant that some patients may present with atypical PD symptoms such as dystonia or early onset dementia and that only between 5-10% of patients formally diagnosed with PD will have a 'textbook' clinical presentation with symptoms of bradykinesia, resting tremor, postural instability and rigidity and will carry mutations in the known PD genes (Farrer 2006; Corti, Lesage, and Brice 2011). Despite this clinicopathological heterogeneity, it is hypothesized that PD is due to defects in a single, unifying biochemical pathway. The investigation into the relationship between autosomal dominant and autosomal recessive PD genes may disentangle the pathways of convergence between different PD forms thereby tipping the scales towards pathology and ultimately towards a cure (Corti, Lesage, and Brice 2011; Alberio, Lopiano, and Fasano 2012; Tofaris 2012).

The development of disease models are proposed to provide a more systematic approach to PD pathogenesis. One such model is illustrated by focussing specifically on the autosomal recessive genes involved in PD. It has been demonstrated through numerous genetic and functional studies that the clinical similarities of PD caused by mutations in *Parkin*, *PINK1* and *DJ-1*, the unique characteristics of each gene as well as the identification of specific functions of each of these genes within common biological pathways provides the evidence needed to conclude that these genes may interact with each other and as a result, are cumulatively involved in PD pathogenesis (Corti, Lesage, and Brice 2011). PINK1 and Parkin interact with each other and in cases where there is a loss of function in either of these genes due to mutations, the maintenance of mitochondria is no longer possible; conversely mutations in DJ-1 have been shown to directly affect mitochondrial function and therefore is hypothesized to interact with the PINK1/Parkin pathway although it remains unclear as to how exactly this interaction may or may not occur (Canet-Avilés et al. 2004; Corti, Lesage, and Brice 2011). Other models suggest that all the recessive and dominant genes interact, albeit at different levels, giving rise to the phenotypic differences observed (figure 9.1).

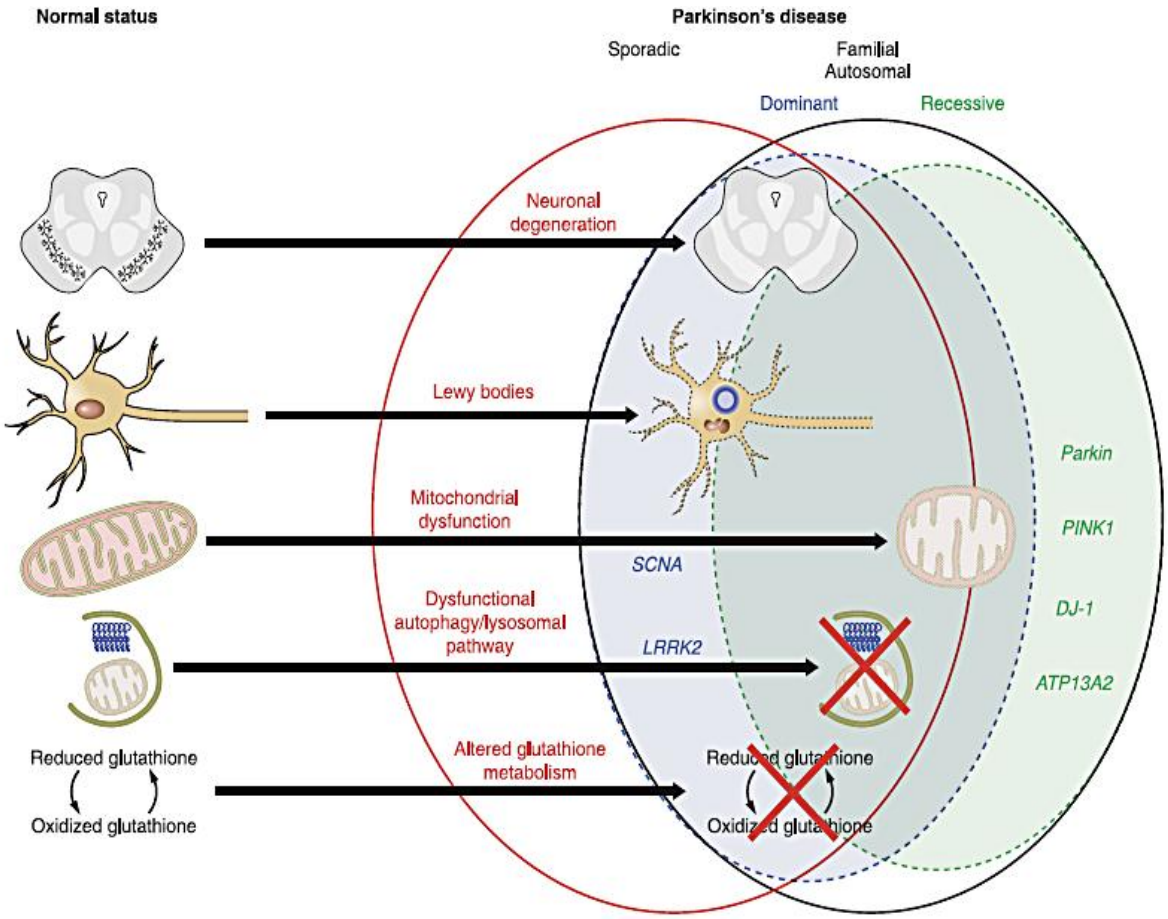


Figure 9.1 Proposed interactive model involving PD pathogenesis. It is hypothesized that autosomal dominant and recessive genes interact in such a way as to result in PD pathogenesis (Taken from Olga Corti, Lesage, and Brice 2011).

More recently, specific focus has been given to the dysfunctional lysosome-dependant pathway as a key unifying mechanism in PD pathogenesis (figure 9.2). Lysosomes are dynamic, membrane bound organelles that are actively involved in degradation within numerous pathways and they are specifically regulated by protein/organelle interactions (Tofaris 2012). They are actively involved in the degradation of damaged mitochondria through mitophagy and in PD, this pathway may be disrupted, thus leading to pathogenesis (Tofaris 2012). The accumulation of SNCA is hypothesized to act as a trigger for the pathogenic cascade involved in PD and therefore, directly affects the lysosome. The impact hereof, is a direct effect on PINK1, Parkin and DJ-1 whereby the mitochondrial integrity can no longer be maintained and in the cases of PINK1 and Parkin, mitophagy is affected. This then produces a knock-on effect where autophagy (the self degradation of the contents of a cell through the use of lysosomal machinery) and endosomal trafficking is affected in LRRK2, the endosomal recycling in VPS35 is affected and the ATP13A2 and GBA needed

for lysosomal functioning may lead to pathogenesis (Tofaris 2012). Mitochondrial membrane damage, lysosomes and synaptic vesicles possibly trigger functions even further downstream because of the increase in ROS, proteases and neurotransmitters.

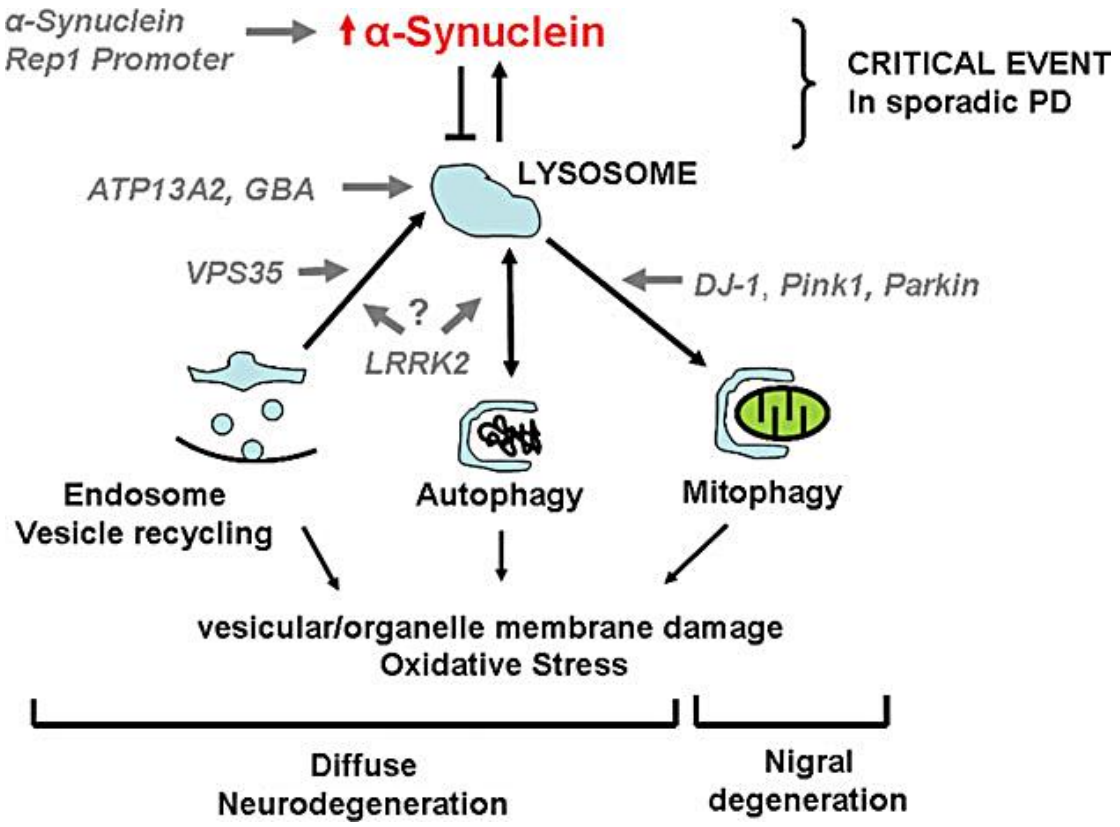


Figure 9.2 Lysosome-dependant pathway as a model for PD pathogenesis. The known PD genes are hypothesized to interact in the lysosomal pathway subsequently resulting in a change in the reaction of the cell to oxidative stress and the processing of excess proteins (Taken from Tofaris 2012).

9.4 Neurodegenerative disorders with overlapping clinical symptoms

A confounding factor in studies on PD is the number of neurodegenerative disorders with similar clinical symptoms; often complicating diagnosis of the disorder. PD is typically defined phenomenologically as the presence of bradykinesia, rigidity, postural instability and resting tremor (although, not all features are mandatory) together with an excellent and sustained response to levodopa as well as the development of motor complications after many years of levodopa treatment. The definitive diagnosis requires the demonstration of loss of dopaminergic neurons and the presence of Lewy-body synucleinopathy at autopsy which

obviously is not available in the clinical setting. This means that clinicians have to consider an increasing range of differential diagnosis when confronted with a patient with Parkinsonism and a positive family history. Parkinsonism can be defined as a variety of dissimilar underlying clinical features that can cause Parkinson's-like symptoms, such as tremor, stiffness, slowing of movement and balance problems, but the predominant features are atypical or additional neurological features that are not found in the majority of PD cases (Klein, Schneider, and Lang 2009). Disease classification and identification is largely dependent on the clinician's expertise and the clinical manifestations of the patient at the time of the consultation. Disorders such as HDL2 and SCAs can share a significant clinical overlap with PD and for this reason, algorithms have been designed as a means to clarify PD diagnosis, particularly in early onset or atypical cases, using clinical observations and results from MRI scans. Also, the ethnic background can be useful as there are a number of diseases which are specifically related to specific populations and genetic testing can be performed in such instances, so as to rule out disorders that present with signs of Parkinsonism or even mimic idiopathic PD (figure 9.3) (Klein, Schneider, and Lang 2009).

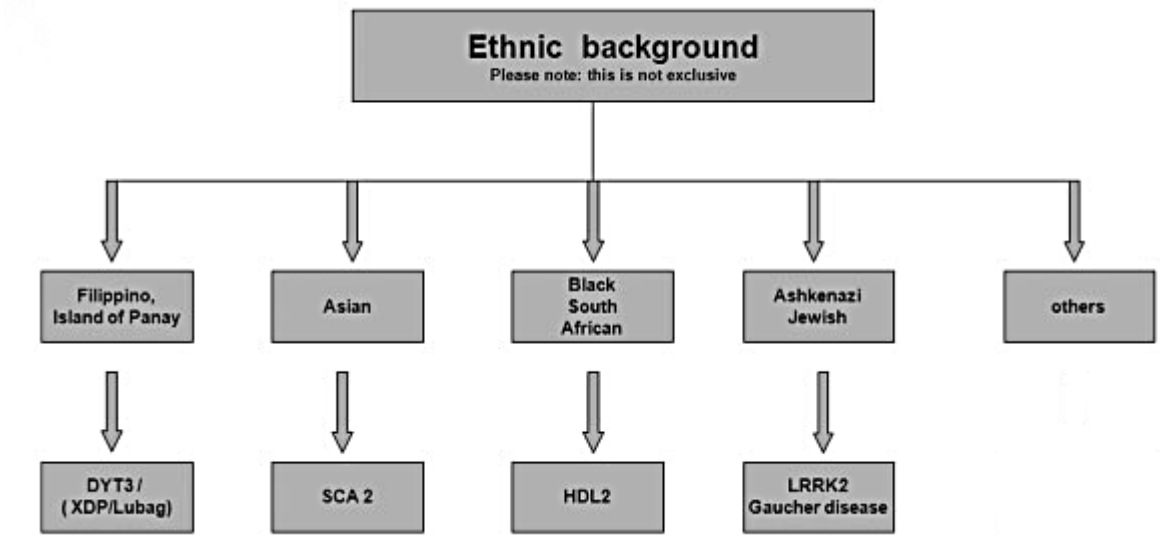


Figure 9.3 Suggested algorithm for clinical diagnosis based on ethnic background (Klein, Schneider, and Lang 2009).

Along with the suggested algorithm, a relatively new development into PD and its separation from other Parkinsonian disorders, is the discovery and identification of biomarkers that can then be used as a reliable and accurate way to identify at risk populations (Agarwal and

Stoessl 2012; Jankovic 2012). The biomarker projects aim to investigate blood, genetic, cerebrospinal fluid, imaging and neurophysiologic abnormalities in patients that may be affected (Jankovic 2012). Additionally, new developments into PD involve modifications of the single photon emission computerized tomography (SPECT) scan; these have been used to identify pre-synaptic dopamine levels in the nigrostriatal terminals, thereby providing a means to differentiate PD from similar Parkinsonian disorders.

9.5 Problems with PD research in sub-Saharan Africa (SSA)

The prevalence of PD in sub-Saharan African **(**SSA) countries is reported to be significantly lower than that found in the developed world. In Africa, the prevalence has been reported to be between 7 and 43 per 100 000 (Okubadejo et al. 2006) compared to figures of between 7 and 657 for the rest of the world. However, these figures for SSA are likely to be an underestimation due to a number of reasons including methodological problems with some of the studies (hospital-based studies are thought to underestimate PD as most patients are in the community and are not in a hospital or clinical environment) and the fact that many patients are either misdiagnosed or undiagnosed.

Although PD can only be identified with certainty following autopsy, the diagnosis thereof can be made according to a set of clinical criteria such as the UK PD Society Brain Bank criteria. Moreover, it is of paramount importance that patients be diagnosed by a movement disorder specialist and undergo specialized and specific neurological examinations. In progressed forms of the disease, diagnosis becomes more straightforward, while definitive PD diagnoses are more difficult in the early stages (Dotchin et al. 2012). In developed countries technological applications such as SPECT scans have been developed and successfully utilized to aid in diagnosis. This approach is used in cases where diagnosis is uncertain and gamma rays are used to measure the dopamine re-uptake which is significantly reduced in patients where neuronal degeneration has taken place, as is the case in PD patients (Isaias and Antonini 2010). Additionally, dopaminergic transporter imaging scans have been developed and these have a sensitivity approaching 100% in the diagnosis of neurodegenerative Parkinsonism. Cumulatively, the availability of modern technologies, trained specialists and functional healthcare infrastructures allows for the identification of PD and subsequently appropriate management of the disorder (Dotchin et al. 2012). This is not possible in SSA as these specialised imaging modalities are not widely available and few

patients go on to post-mortem examination so the only way in which diagnosis of the disease can take place is through clinical examination (Mshana, Dotchin, and Walker 2011). This poses another problem as due to cultural beliefs and practices of many tribes in SSA, many individuals who are affected by PD seek treatment from traditional healers, as the symptoms presented may be believed to be a curse or as a result of witchcraft (Mshana, Dotchin, and Walker 2011). Other patients do not seek medical treatment due to the misconception that their symptoms are part of the normal ageing process (Dotchin et al. 2012).

Several SSA countries do not have neurologists, movement disorder specialists or neurosurgeons. Training of other healthcare professionals to identify PD would be one way of overcoming the lack of specialists and would be an innovative way of raising awareness of the disorder. In one study carried out in Tanzania, a PD specialist nurse who had been trained in the UK, was used to educate patients and carers about the disease. Furthermore, a manual entitled 'Where there is no neurologist' developed by the World Federation of Neurology is available for free via the internet and can be used by paramedical professionals to give basic advice, prescribe treatment and identify conditions that need to be referred to a doctor (http://www.wfneurology.org/the-african-experience).

Another critical challenge faced in SSA is that for the minority of PD patients who are diagnosed, anti-Parkinsonian drug therapies are often unavailable or unaffordable (Dotchin et al. 2012). Typically, hospitals in SSA stock outdated PD drugs with a levodopa and carbidopa formulation that contains a lower proportion of dopa-decarboxylase inhibitor than used in developed countries. This often results in side effects including nausea and postural hypotension leading to noncompliance with the medication. In general, modern, efficacious treatments for PD such as nonergot dopamine agonists and monoamine oxidase B inhibitors are not available in SSA. PD does not only affect the motor systems of patients; non-motor symptoms may precede the onset of disease , continue throughout its course and can be more troubling to the patient than the motor symptoms (Dotchin, Msuya, and Walker 2010). Some of these non-motor symptoms include depression, cognitive decline, changes in sleeping patterns, constipation and pain and these may lead to a reduced quality of life even if the motor symptoms are adequately controlled (Parkinson's Disease Foundation 2012). Notably, the cognitive impairment and mood disturbances are challenging for these patients as well as their carers; and diagnosis of cognitive impairment is difficult in the African setting, as the tools are not suitable for use in populations with low educational levels (Dotchin et al. 2012).

To date, only a minor number of studies regarding the underlying genetic factors of PD in SSA have been reported. The number of published studies is as little as nine, with the majority of these stemming from our group in South Africa (Bardien et al. 2009; Keyser et al. 2009; Keyser 2010a; Keyser 2010b; Haylett et al. 2012). The focus of these studies was predominantly on the known PD causing genes and small number of mutations identified in these patients has provided strong evidence that the known PD causing genes do not play a major role in the disorder and therefore the hypothesis that novel genes may be involved is a plausible one. The use of alternative NGS technologies may be particularly feasible in these populations for the identification of these novel disease-causing mutations.

Recently, the international Movement Disorder Society launched an Interest Group for Movement Disorders in SSA. Their goal is to build a network of interested parties to provide education and training in movement disorders for improving patient care. Regular meetings and conferences to encourage networking between African clinicians and scientists with an interest in PD and other movement disorders would also be a major boost to more research activities in this field. In Mali, the Health Minister has announced an innovative programme allowing all PD patients to be provided with anti-Parkinsonian drugs free of charge (Dotchin et al. 2012). Taken together, these initiatives will help raise awareness of PD in SSA countries, may aid in the identification of the disease at an earlier stage, and may also reduce stigmas for patients and their family members.

9.6 Limitations of the present study

Limitations of the present study included the limited numbers of study participants particularly from the Black ethnic group, for which only 17 patients had been recruited. This may be due to recruitment bias as study participants were only recruited from The Movement Disorders Clinic at Tygerberg Hospital in Cape Town. In addition, there are no state-employed neurologists in the Eastern Cape Province, from which almost all of our Black patients originate, which means that many Black patients will not be diagnosed. Another limitation was the fact that there were not many families available for study that had more than two affected individuals. Co-segregation with disease within a family was therefore not possible and the pathogenicity of novel variants could not be determined. Furthermore, due to financial constraints not all study participants could be assessed using comprehensive

imaging analysis such as MRI and SPECT scans to determine the extent and progression of their condition.

Another major limitation of the study is that Sanger sequencing could not be performed on all of the patients. Ideally, if financial constraints had not been a factor, all exons of all eight of the known PD genes would have been screened for pathogenic mutations. However, this was not feasible and for this reason, alternative cost-effective methods had to be employed. HRM was chosen as the most cost-effective, medium throughput method to be used for mutation screening but it is possible that pathogenic mutations were missed using this approach. False positives and negatives are generated when primer dimers are formed and the normalized graphs are re-adjusted to incorporate the primer dimer formation, thus possibly resulting in aberrant results. Additionally, it is of paramount importance that the cycle threshold (Ct) values of the pre-amplification template be less than 30. Ct can be defined as the number of cycles that are required for fluorescent signals to cross the threshold (i.e. the background). The lower the Ct value, the greater the amount of target nucleic acid per reaction; if these Ct values are not below 30, then poor RT-PCR products are generated and poor melt curves are the result, thus making the curves difficult to analyse (Wittwer 2009; Ye et al. 2010). Also, HRM was not effective in the identification of a g.-6_+10del variant in *DJ-1* with a significant GC content. It has been reported that HRM is not effective in identifying homozygous substitutions and A>T substitutions do not cause significant shifts in the melt curves (less than a 0.2°C, difference), thus making these specific substitutions difficult to detect (Wittwer 2009; Ye et al. 2010). It should also be noted that samples with the same or similar melt curves may not have the same variant and all samples with altered melt profiles should thus be sequenced to characterize the sequence variant.

Of the 262 patients that were included in the study, 33 of them were not screened for exon dosage changes using MLPA assays. Exonic rearrangements, whole gene duplications and triplications are common in PD patients (Hedrich et al. 2002) but are not detected by techniques such as HRM or Sanger sequencing and all patients should thus be subjected to MLPA analysis to exclude copy number variations or rearrangements within the known PD genes.

The WES approach, although extremely useful in the identification of novel disease-causing mutations, comes with a number of requirements that need to be taken into account when

employing this technique (Cirulli et al. 2010; Kim 2010). The biggest concern is the fact that approximately 98.8% of the genome is not analysed as the assumption is that the mutation is in the coding regions. Added to this, variations with regards to the definition of the 'exome' also poses a significant challenge and different commercial kits vary in their specific targets. Moreover, WES does not cover all genes; some of the genes that are found in highly repetitive regions, genes that have a high degree of sequence similarity with other genomic regions including pseudogenes and as yet unidentified or unannotated genes are missed. CNVs are also not effectively detected through the use of this method. Sanger sequencing to verify the presence of variants detected is still an important step in this process as due to the short read lengths many fragments could be mapped to the wrong genomic region. Whole genome sequencing is therefore an attractive alternative to WES and generation of whole genome data from all the South African PD patients would be a significant resource. However, until this approach becomes more affordable and the appropriate bioinformatics infrastructure is in place to deal with the large amounts of data, WES remains the more viable option.

9.7 Future work

A relatively small number of novel variants were identified in the known PD genes, but these should be analysed further to conclusively determine their role in PD pathogenesis. Functional studies and analysis of the novel sequence variants identified in *Parkin* (T387P) and *LRRK2* (Q2089R) are proposed. The frequency of each of the variants should first be established in large number of ethnically matched controls and given the outcome, possible functional studies should be carried out. It is plausible to consider this because the Q2089R variant in *LRRK2* resides in the highly conserved kinase domain and it has been documented that mutations in this gene may decrease GTPase activity and increase kinase activity (Tan et al. 2008; Aasly et al. 2010). In the case of *Parkin,* the effect of the T387P variant on ubiquitination and protein aggregation could be carried out through the use of immunohistochemistry and ubiquitylation assays.

Improved and effective, medium-throughput methods for the identification of g.-6_+10del variant in *DJ-1* are also a necessary consideration. This variant was genotyped by a commercial company (KBiosciences, LGC Genomics, Teddington, United Kingdom), but an effective assay should be established locally. The frequency of the g.-6_+10del variant can

then be determined in controls from the different ethnic groups once an appropriate screening method has been developed. Moreover, given the proximity of the indel to the transcription start site, it is warranted that further investigation be carried out on the effect on transcriptional regulation using South western blotting or electrophoretic mobility shift assays (Siu, Lee, and Chow 2008). Patients and controls should also be screened for the frequency of the 18bp indel which had previously been identified in a Finnish population in the promoter region (Eerola et al. 2003).

Another facet of future work which may be applied, is a targeted resequencing approach (Chan et al. 2012). This is a relatively new concept, but one which has not yet been widely applied to many NGS projects. This approach investigates a relatively small number of genes across large populations and is complex in that a large amount of PCR is necessary to amplify all the exons targeted while at the same time managing to keep track of which PCR product belongs to which patient - this is done through DNA bar coding, which allows multiplexing of numerous samples and this, in turn, decreases the overall costs involved in the sequencing process (Chan et al. 2012). This means that even when as few as 10 genes with 10 exons each are examined in 100 individuals, approximately 10,000 PCR reactions and 10,000 sequencing reactions are the result. An NGS approach would theoretically amplify the genomic regions, combine all information and finally run a library prep protocol (i.e. fragment and attach linkers). The advantages of this approach are that any region of interest can be targeted and the results are relatively easy to analyse. Future work should involve targeted resequencing of all known PD genes in all of the South African PD patients, and it is anticipated that this work could be done locally on the Ion Torrent (Life Technologies, Carlsbad, California, USA), which is available at the Central Analytical Facility (CAF) at Stellenbosch University. In conjunction with targeted resequencing, MLPA should be performed for all patients. This is important as current software used in the analysis of targeted resequencing data is not effective at detecting CNVs. CNVs are thought to influence gene expression and can be directly associated with a range of phenotypes and diseases. Also as new genes are discovered, CNVs analysis of these genes should be included in the mutation screening strategies.

The WES study on South African PD patients in is the first of its kind and has yielded promising results. Future work will involve the verification through the use of Sanger sequencing of the remaining sixteen sequence variants that were identified as plausible

candidate disease-causing mutations. Additionally, the entire PD cohort should be screened for these variants and the pathogenicity assessed within the population. The frequency of the identified pathogenic mutations can then be assessed in approximately 10 000 PD patients potentially available from the GEOPD consortium. And finally, the effect of the identified variants on protein function should be assessed through the use of various assays such as protein kinase assays if the protein has a kinase function. The present study has highlighted that there is an urgent need for the development of a 'universal bioinformatics pipeline' that can be applied to all future WES projects so as to ensure consistency and optimal functionality of this approach.

9.8 Concluding remarks

Given the significant economic and social burden of neurodegenerative disorders such as PD, more research into the underlying causes and risk factors, is needed. The investigation into the molecular aetiology of PD in the South African patients has revealed that five of the known genes do not play a significant role in these patients and it is plausible to assume that the underlying mutations in many of these patients are in novel PD-causing genes. Only 15 patients were identified as carriers of mutations and of these, the majority (eight) of these are Caucasian.

Since the South African population encompasses numerous diverse ethnic groups, each with unique ancestral origins it is plausible that they may harbour novel PD genes that could potentially provide important insights into disease mechanisms (Okubadejo et al. 2008). Notably, the Black population has been understudied worldwide and therefore more studies are warranted on this particular group of patients. However, there are only 17 Black patients in our group of study participants and therefore more intensive efforts are needed for the identification and recruitment of additional patients. In the quest to find disease genes, NGS approaches are an important strategy as large families with multiple affected individuals are not required, and this is significant because these are rare in a late-onset disorder such as PD. The drive to identify more PD genes is of utmost important as this information is needed to help piece together all the pieces of the puzzle of the disease pathogenesis. Current hypotheses implicate lysosome-dependent pathways as playing an integral role (Tofaris 2012) and it would be interesting to see if new PD genes fit in with this hypothesis or implicate other pathways. Taken together, the pathways in which PD genes are involved and data on

pathological as well as clinical symptoms of patients may reveal exactly why the dopaminergic neurons in the substantia die in these patients – which, almost 200 years after Dr James Parkinson first described the disorder, remains a key unanswered question.

References

Aasly, J. O., C. Vilariño-Güell, J. C. Dachsel, P. J. Webber, A. B. West, K. Haugarvoll, K. K. Johansen, et al. 2010. "Novel Pathogenic LRRK2 p.Asn1437His Substitution in Familial Parkinson's Disease." *Movement Disorders* 25 (13) (October 15): 2156–2163. doi:10.1002/mds.23265.

Abbas, N., C. B. Lücking, S. Ricard, A. Dürr, V. Bonifati, G. De Michele, S. Bouley, et al. 1999. "A Wide Variety of Mutations in the Parkin Gene Are Responsible for Autosomal Recessive Parkinsonism in Europe." *Human Molecular Genetics* 8 (4) (April 1): 567–574. doi:10.1093/hmg/8.4.567.

Abou-Sleiman, P. M., D. G. Healy, N. Quinn, A. J. Lees, and N. W. Wood. 2003. "The Role of Pathogenic DJ-1 Mutations in Parkinson's Disease." *Annals of Neurology* 54 (3): 283–286. doi:10.1002/ana.10675.

Abou-Sleiman, P. M., M. M. K. Muqit, N. Q. McDonald, Y. Xiang Yang, S. Gandhi, D. G. Healy, K. Harvey, et al. 2006. "A Heterozygous Effect for PINK1 Mutations in Parkinson's Disease?" *Annals of Neurology* 60 (4) (September 12): 414–419. doi:10.1002/ana.20960.

Adler, C. 2005. "Nonmotor Complications in Parkinson's Disease." *Movement Disorders* 20 (S11) (May): S23–S29. doi:10.1002/mds.20460.

Agarwal, P. A., and A. J. Stoessl. 2012. "Biomarkers for Trials of Neuroprotection in Parkinson's Disease." *Movement Disorders*: n/a–n/a. doi:10.1002/mds.25065.

Ahlskog, J. K., J. K. Björk, A. N. Elsing, C. Aspelin, M. Kallio, P. Roos-Mattjus, and L. Sistonen. 2010. "Anaphase-Promoting Complex/Cyclosome Participates in the Acute Response to Protein-Damaging Stress." *Molecular and Cellular Biology* 30 (24) (December 15): 5608–5620. doi:10.1128/MCB.01506-09.

Albanese, A., E. M. Valente, L. M. Romito, E. Bellacchio, A. E. Elia, and B. Dallapiccola. 2005. "The PINK1 Phenotype Can Be Indistinguishable from Idiopathic Parkinson Disease." *Neurology* 64 (11) (June 14): 1958–1960. doi:10.1212/01.WNL.0000163999.72864.FD.

Alberio, T., L. Lopiano, and M. Fasano. 2012. "Cellular Models to Investigate Biochemical Pathways in Parkinson's Disease." *FEBS Journal* 279 (7): 1146–1155. doi:10.1111/j.1742-4658.2012.08516.x.

Alcalay, R. N. 2010. "Genetic Insights into Early-onset Parkinson's Disease." *European Neurological Review* 5 (1): 30–33.

Alves, G., E. B. Forsaa, K. F. Pedersen, M. D. Gjerstad, and J. P. Larsen. 2008. "Epidemiology of Parkinson's Disease." *Journal of Neurology* 255 (S5) (September 12): 18–32. doi:10.1007/s00415-008-5004-3.

Alzheimer, A. 1892. "Ueber Einen Fall Von Spinaler Progressiver Muskelatrophie Mit Hinzutretender Erkrankung Bulbärer Kerne Und Der Rinde." *Archiv Für Psychiatrie Und Nervenkrankheiten* 23 (2): 459–485. doi:10.1007/BF01843464.

An, X.-K., R. Peng, T. Li, J.-M. Burgunder, Y. Wu, W.-J. Chen, J.-H. Zhang, et al. 2008. "LRRK2 Gly2385Arg Variant Is a Risk Factor of Parkinson's Disease Among Han-Chinese from Mainland China." *European Journal of Neurology* 15 (3) (March): 301–305. doi:10.1111/j.1468-1331.2007.02052.x.

Arias Vásquez, A., K. Sleegers, M. C.J. Dekker, W. A. van Gool, J. C. van Swieten, Y. S. Aulchenko, B. A. Oostra, and C. M. van Duijn. 2004. "A Deletion in DJ-1 and the Risk of Dementia—a Population-based Survey." *Neuroscience Letters* 372 (3) (December 6): 196–199. doi:10.1016/j.neulet.2004.09.040.

Bach, J-P., U. Ziegler, G. Deuschl, R. Dodel, and G. Doblhammer-Reiter. 2011. "Projected Numbers of People with Movement Disorders in the Years 2030 and 2050." *Movement Disorders* 26 (12): 2286–2290. doi:10.1002/mds.23878.

Bamshad, M. J., S. B. Ng, A. W. Bigham, H. K. Tabor, M. J. Emond, D. A. Nickerson, and J. Shendure. 2011. "Exome Sequencing as a Tool for Mendelian Disease Gene Discovery." *Nature Reviews Genetics* 12 (11) (November 1): 745–755. doi:10.1038/nrg3031.

Bardien, S., F. Abrahams, H. Soodyall, L. van der Merwe, J. Greenberg, T. Brink, and J. Carr. 2007. "A South African Mixed Ancestry Family with Huntington Disease-like 2: Clinical and Genetic Features." *Movement Disorders* 22 (14): 2083–2089. doi:10.1002/mds.21672.

Bardien, S., R. Keyser, Y. Yako, D. Lombard, and J. Carr. 2009. "Molecular Analysis of the Parkin Gene in South African Patients Diagnosed with Parkinson's Disease." *Parkinsonism & Related Disorders* 15 (2) (February): 116–121. doi:10.1016/j.parkreldis.2008.04.005.

Bardien, S., S. Lesage, A. Brice, and J. Carr. 2011. "Genetic Characteristics of Leucine-rich Repeat Kinase 2 (LRRK2) Associated Parkinson's Disease." *Parkinsonism & Related Disorders* 17 (7) (August): 501–508. doi:10.1016/j.parkreldis.2010.11.008.

Beasley, S. A., V. A. Hristova, and G. S. Shaw. 2007. "Structure of the Parkin In-between-Ring Domain Provides Insights for E3-Ligase Dysfunction in Autosomal Recessive

Parkinson's Disease." *Proceedings of the National Academy of Sciences of the United States of America* 104 (9) (February 27): 3095–3100.

Beddington, J., C. L. Cooper, J. Field, U. Goswami, F. A. Huppert, R. Jenkins, H. S. Jones, T. B. L. Kirkwood, B. J. Sahakian, and S. M. Thomas. 2008. "The Mental Wealth of Nations." *Nature* 455 (7216) (October 22): 1057–1060. doi:10.1038/4551057a.

Berg, D., M. Niwar, S. Maass, A. Zimprich, J. C. Möller, U. Wuellner, T. Schmitz-Hübsch, et al. 2005. "Alpha-synuclein and Parkinson's Disease: Implications from the Screening of More Than 1,900 Patients." *Movement Disorders* 20 (9): 1191–1194. doi:10.1002/mds.20504.

Betarbet, R., R. M. Canet-Aviles, T. B. Sherer, P. G. Mastroberardino, C. McLendon, J-H. Kim, S. Lund, et al. 2006. "Intersecting Pathways to Neurodegeneration in Parkinson's Disease: Effects of the Pesticide Rotenone on DJ-1, A-synuclein, and the Ubiquitin–proteasome System." *Neurobiology of Disease* 22 (2) (May): 404–420. doi:10.1016/j.nbd.2005.12.003.

Biesecker, L. G. 2010. "Exome Sequencing Makes Medical Genomics a Reality." *Nature Genetics* 42 (1) (January): 13–14. doi:10.1038/ng0110-13.

Bilgüvar, K., A. K. Öztürk, A. Louvi, K. Y. Kwan, M. Choi, B. Tatlı, D. Yalnızoğlu, et al. 2010. "Whole-exome Sequencing Identifies Recessive WDR62 Mutations in Severe Brain Malformations." *Nature* 467 (7312) (September 9): 207–210. doi:10.1038/nature09327.

Bisaglia, M., I. Tessari, S. Mammi, and L. Bubacco. 2009. "Interaction Between Alpha-synuclein and Metal Ions, Still Looking for a Role in the Pathogenesis of Parkinson's Disease." *Neuromolecular Medicine* 11 (4): 239–251. doi:10.1007/s12017-009-8082-1.

Biskup, S., D. J. Moore, F. Celsi, S. Higashi, A. B. West, S. A. Andrabi, K. Kurkinen, et al. 2006. "Localization of LRRK2 to Membranous and Vesicular Structures in Mammalian Brain." *Annals of Neurology* 60 (5): 557–569. doi:10.1002/ana.21019.

Bonifati, V. 2012. "Autosomal Recessive Parkinsonism." *Parkinsonism & Related Disorders* 18 (SUPPL. 1).

Bonifati, V., P. Rizzu, F. Squitieri, E. Krieger, N. Vanacore, J.C. Swieten, A. Brice, et al. 2003. "DJ-1 (PARK7), a Novel Gene for Autosomal Recessive, Early Onset Parkinsonism." *Neurological Sciences* 24 (3): 159–160. doi:10.1007/s10072-003-0108-0.

Bonifati, V., C. F. Rohe, G. J. Breedveld, E. Fabrizio, M. De Mari, C. Tassorelli, A. Tavella, et al. 2005. "Early-onset Parkinsonism Associated with PINK1 Mutations: Frequency, Genotypes, and Phenotypes." *Neurology* 65 (1) (July 11): 87–95. doi:10.1212/01.wnl.0000167546.39375.82.

Bonifati, V. 2007. “Genetics of Parkinsonism.” *Parkinsonism & Related Disorders* 13, Supplement 3 (0): S233–S241. doi:10.1016/S1353-8020(08)70008-7.

Bonifati, V., G. J. Breedveld, F. Squitieri, N. Vanacore, P. Brustenghi, B. S. Harhangi, P. Montagna, et al. 2002. “Localization of Autosomal Recessive Early-onset Parkinsonism to Chromosome 1p36 (PARK7) in an Independent Dataset.” *Annals of Neurology* 51 (2): 253–256. doi:10.1002/ana.10106.

Brancati, F. 2002. “PARK6 Is a Common Cause of Familial Parkinsonism.” *Neurological Sciences* 23 (SUPPL. 2). doi:10.1007/s100720200097.

Bras, J. M., and A. B. Singleton. 2011. “Exome Sequencing in Parkinson’s Disease.” *Clinical Genetics* 80 (2) (June 16): 104–109. doi:10.1111/j.1399-0004.2011.01722.x.

Brink, P., L. Crotti, V. Corfield, A. Goosen, G. Durrheim, P. Hedley, M. Heradien, et al. 2005. “Phenotypic Variability and Unusual Clinical Severity of Congenital Long-QT Syndrome in a Founder Population.” *Circulation* 112 (17) (October 25): 2602–2610. doi:10.1161/CIRCULATIONAHA.105.572453.

Brooks, D. J. 2010. “Examining Braak’s Hypothesis by Imaging Parkinson’s Disease.” *Movement Disorders* 25 (S1) (January 1): S83–S88. doi:10.1002/mds.22720.

Burbano, H. A., E. Hodges, R. E. Green, A. W. Briggs, J. Krause, M. Meyer, J. M. Good, et al. 2010. “Targeted Investigation of the Neandertal Genome by Array-Based Sequence Capture.” *Science* 328 (5979) (May 7): 723–725. doi:10.1126/science.1188046.

Camacho, J. L. G., N. M. Jaramillo, P. Y. Gómez, M. R. Violante, C. B. Woehrlen, E. A. Vilatela, and M. López López. 2012. “High Frequency of Parkin Exon Rearrangements in Mexican-mestizo Patients with Early-onset Parkinson’s Disease.” *Movement Disorders* 27 (8): 1047–1051. doi:10.1002/mds.25030.

Canet-Avilés, R. M., M. A. Wilson, D. W. Miller, R. Ahmad, C. McLendon, S. Bandyopadhyay, M. J. Baptista, D. Ringe, G. A. Petsko, and M. R. Cookson. 2004. “The Parkinson’s Disease Protein DJ-1 Is Neuroprotective Due to Cysteine-sulfinic Acid-driven Mitochondrial Localization.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (24) (June 15): 9103–9108. doi:10.1073/pnas.0402959101.

Cannon, J. R., and J. T. Greenamyre. 2011. “The Role of Environmental Exposures in Neurodegeneration and Neurodegenerative Diseases.” *Toxicological Sciences* 124 (2) (September 13): 225–250. doi:10.1093/toxsci/kfr239.

Cawley S. 2005. “Sanger’s Method.” http://stat-www.berkeley.edu/users/terry/Classes/s260.1998/Week8b/week8b/node9.html.

Chahrour, M. H., T.W. Yu, E. T. Lim, B. Ataman, M. E. Coulter, R. S. Hill, C. R. Stevens, et al. 2012. "Whole-Exome Sequencing and Homozygosity Analysis Implicate Depolarization-Regulated Neuronal Genes in Autism." *PLoS Genet* 8 (4) (April 12): e1002635. doi:10.1371/journal.pgen.1002635.

Chan, M., S. M. Ji, Z. X. Yeo, L. Gan, E. Yap, Y. S. Yap, R. Ng, et al. 2012. "Development of a Next-Generation Sequencing Method for BRCA Mutation Screening: A Comparison Between a High-Throughput and a Benchtop Platform." *The Journal of Molecular Diagnostics* 14 (6) (November): 602–612. doi:10.1016/j.jmoldx.2012.06.003.

Chartier-Harlin, M-C., J. C. Dachsel, C. Vilariño-Güell, S. J. Lincoln, F. Leprêtre, M. M. Hulihan, J. Kachergus, et al. 2011. "Translation Initiator EIF4G1 Mutations in Familial Parkinson Disease." *American Journal of Human Genetics* 89 (3) (September 9): 398–406. doi:10.1016/j.ajhg.2011.08.009.

Chung, K. K., Y. Zhang, K. L. Lim, Y. Tanaka, H. Huang, J. Gao, C. A. Ross, V. L. Dawson, and T. M. Dawson. 2001. "Parkin Ubiquitinates the Alpha-synuclein-interacting Protein, Synphilin-1: Implications for Lewy-body Formation in Parkinson Disease." *Nature Medicine* 7 (10) (October): 1144–1150. doi:10.1038/nm1001-1144.

Cirulli, E., A. Singh, K. Shianna, D. Ge, J. Smith, J. Maia, E. Heinzen, J. Goedert, D. Goldstein. 2010. "Screening the Human Exome: a Comparison of Whole Genome and Whole Transcriptome Sequencing." *Genome Biology* 11 (5) (May 28): R57. doi:10.1186/gb-2010-11-5-r57.

Clark, I. E., M. W. Dodson, C. Jiang, J. H. Cao, J. R. Huh, J. H. Seol, S. J. Yoo, B. A. Hay, and M. Guo. 2006. "Drosophila Pink1 Is Required for Mitochondrial Function and Interacts Genetically with Parkin." *Nature* 441 (7097) (June 29): 1162–1166. doi:10.1038/nature04779.

Collins, P. Y., V. Patel, S. S. Joestl, D. March, T. R. Insel, and A. S. Daar. 2011. "Grand Challenges in Global Mental Health." *Nature* 475 (7354) (July 6): 27–30. doi:10.1038/475027a.

Conrad, D. F., C. Bird, B. Blackburne, S. Lindsay, L. Mamanova, C. Lee, D. J. Turner, and M. E. Hurles. 2010. "Mutation Spectrum Revealed by Breakpoint Sequencing of Human Germline CNVs." *Nature Genetics* 42 (5) (May): 385–391. doi:10.1038/ng.564.

Cookson, M. R., and O. Bandmann. 2010. "Parkinson's Disease: Insights from Pathways." *Human Molecular Genetics* 19 (R1) (April 26): R21–R27. doi:10.1093/hmg/ddq167.

Cookson, M. R. 2003. “Parkin’s substrates and the pathways leading to neuronal damage.” *Neuromolecular Medicine* 3 (1) (February): 1–13. doi:10.1385/NMM:3:1:1.

Cookson, M. R. 2010. “Chapter 3 - Unravelling the Role of Defective Genes.” In *Recent Advances in Parkinson’s Disease: Basic Research*, Volume 183:43–57. Elsevier. http://www.sciencedirect.com/science/article/pii/S0079612310830031.

Cooper, J. M., S. E. Daniel, C. D. Marsden, and A. H. Schapira. 1995. “L-dihydroxyphenylalanine and Complex I Deficiency in Parkinson’s Disease Brain.” *Movement Disorders: Official Journal of the Movement Disorder Society* 10 (3) (May): 295–297. doi:10.1002/mds.870100311.

Correia Guedes, L., J.J. Ferreira, M.M. Rosa, M. Coelho, V. Bonifati, and C. Sampaio. 2010. “Worldwide Frequency of G2019S LRRK2 Mutation in Parkinson’s Disease: A Systematic Review.” *Parkinsonism & Related Disorders* 16 (4) (May): 237–242. doi:10.1016/j.parkreldis.2009.11.004.

Corti, O., S. Lesage, and A. Brice. 2011. “What Genetics Tells Us About the Causes and Mechanisms of Parkinson’s Disease.” *Physiological Reviews* 91 (4) (October): 1161–1218. doi:10.1152/physrev.00022.2010.

Da Costa, C. A., C. Sunyach, E. Giaime, A. West, O. Corti, A. Brice, S. Safe, et al. 2009. “Transcriptional Repression of P53 by Parkin and Impairment by Mutations Associated with Autosomal Recessive Juvenile Parkinson’s Disease.” *Nature Cell Biology* 11 (11) (November): 1370–1375. doi:10.1038/ncb1981.

Darios, F., O. Corti, C. B. Lücking, C. Hampe, M-P. Muriel, N. Abbas, W-J. Gu, et al. 2003. “Parkin Prevents Mitochondrial Swelling and Cytochrome c Release in Mitochondria-dependent Cell Death.” *Human Molecular Genetics* 12 (5) (March 1): 517–526.

Dauer, W., and . Przedborski. 2003. “Parkinson’s Disease: Mechanisms and Models.” *Neuron* 39 (6) (September 11): 889–909. doi:10.1016/S0896-6273(03)00568-3.

Dawson, T. M., and V. L. Dawson. 2003. “Molecular Pathways of Neurodegeneration in Parkinson’s Disease.” *Science* 302 (5646) (October 31): 819–822. doi:10.1126/science.1087753.

Deas, E., H. Plun-Favreau, and N. W. Wood. 2009. “PINK1 Function in Health and Disease.” *EMBO Molecular Medicine* 1 (3) (June 12): 152–165. doi:10.1002/emmm.200900024.

Deng, H., M. W. Dodson, H. Huang, and M. Guo. 2008. "The Parkinson's Disease Genes Pink1 and Parkin Promote Mitochondrial Fission And/or Inhibit Fusion in Drosophila." *Proceedings of the National Academy of Sciences* 105 (38) (September 23): 14503–14508. doi:10.1073/pnas.0803998105.

Deng, H., W. Le, Y. Guo, C. B. Hunter, W. Xie, and J. Jankovic. 2005. "Genetic and Clinical Identification of Parkinson's Disease Patients withLRRK2 G2019S Mutation." *Annals of Neurology* 57 (6) (June): 933–934. doi:10.1002/ana.20510.

Dixon-Salazar, T. J., J. L. Silhavy, N. Udpa, J. Schroth, S. Bielas, A. E. Schaffer, J. Olvera, et al. 2012. "Exome Sequencing Can Improve Diagnosis and Alter Patient Management." *Science Translational Medicine* 4 (138) (June 13): 138ra78. doi:10.1126/scitranslmed.3003544.

Djarmati, A., K. Hedrich, M. Svetel, N. Schäfer, V. Juric, S. Vukosavic, R. Hering, et al. 2004. "Detection of Parkin (PARK2) and DJ1 (PARK7) Mutations in Early-onset Parkinson Disease: Parkin Mutation Frequency Depends on Ethnic Origin of Patients." *Human Mutation* 23 (5) (May): 525. doi:10.1002/humu.9240.

Dluzen, D.E., and J.L McDermott. 2000. "Gender Differences in Neurotoxicity of the Nigrostriatal Dopaminergic System: Implications for Parkinson's Disease." *The Journal of Gender Specific Medicine* (6) (September): 36-42.

Do, C. B., J.Y. Tung, E. Dorfman, A. K. Kiefer, E. M. Drabant, U. Francke, J. L. Mountain, et al. 2011. "Web-Based Genome-Wide Association Study Identifies Two Novel Loci and a Substantial Genetic Component for Parkinson's Disease." *PLoS Genet* 7 (6) (June 23): e1002141. doi:10.1371/journal.pgen.1002141.

Dodson, M. W, and M. Guo. 2007. "Pink1, Parkin, DJ-1 and Mitochondrial Dysfunction in Parkinson's Disease." *Current Opinion in Neurobiology* 17 (3) (June): 331–337. doi:10.1016/j.conb.2007.04.010.

Dorsey, E. R., R. Constantinescu, J. P. Thompson, K. M. Biglan, R. G. Holloway, K. Kieburtz, F. J.Marshall, et al. 2007. "Projected Number of with Parkinson Disease in the Most Populous Nations, 2005 Through 2030." *Neurology* 68 (5) (January 30): 384–386. doi:10.1212/01.wnl.0000247740.47667.03.

Dotchin, C., A. Jusabani, W. K. Gray, and R. Walker. 2012. "Projected Numbers of People with Movement Disorders in the Years 2030 and 2050: Implications for sub-Saharan Africa, Using Essential Tremor and Parkinson's Disease in Tanzania as an Example." *Movement Disorders: Official Journal of the Movement Disorder Society* 27 (9) (August): 1204–1205; author reply 1207. doi:10.1002/mds.25097.

Dotchin, C., O. Msuya, and R. Walker. 2010. “The Use of Smell Identification Tests in the Diagnosis of Parkinson’s Disease in Brazil.” *Movement Disorders* 25 (4): 510–511. doi:10.1002/mds.22906.

Douna, H., B. M. Bavelaar, H. Pellikaan, B. Olivier, and T. Pieters. 2012. “Neuroprotection in Parkinson’s Disease: A Systematic Review of the Preclinical Data.” *The Open Pharmacology Journal* 6 (1) (May 9): 12–26.

Van Duijn, C. M., M. C. Dekker, V. Bonifati, R. J. Galjaard, J. J. Houwing-Duistermaat, P. J. Snijders, L. Testers, et al. 2001. “Park7, a Novel Locus for Autosomal Recessive Early-onset Parkinsonism, on Chromosome 1p36.” *American Journal of Human Genetics* 69 (3) (September): 629–634. doi:10.1086/322996.

Ebrahimi-Fakhari, D., L. Wahlster, and P. J. McLean. 2011. “Molecular Chaperones in Parkinson’s Disease – Present and Future.” *Journal of Parkinson’s Disease* 1 (4) (November): 299–320.

Eeden, Van Den, S. K., C. M. Tanner, A. L. Bernstein, R. D. Fross, A. Leimpeter, D. A. Bloch, and L. M. Nelson. 2003. “Incidence of Parkinson’s Disease: Variation by Age, Gender, and Race/Ethnicity.” *American Journal of Epidemiology* 157 (11) (June 1): 1015–1022. doi:10.1093/aje/kwg068.

Eerola, J., D. Hernandez, J. Launes, O. Hellström, S. Hague, C. Gulick, J. Johnson, et al. 2003. “Assessment of a DJ-1 (PARK7) Polymorphism in Finnish PD.” *Neurology* 61 (7) (October 14): 1000–1002.

Eggers, C., A. Schmidt, J. Hagenah, N. Bruggemann, J. C. Klein, V. Tadic, L. Kertelge, et al. 2010. “Progression of Subtle Motor Signs in PINK1 Mutation Carriers with Mild Dopaminergic Deficit.” *Neurology* 74 (22) (May 31): 1798–1805. doi:10.1212/WNL.0b013e3181e0f79c.

Elbaz, A., and C. Tranchant. 2007. “Epidemiologic Studies of Environmental Exposures in Parkinson’s Disease.” *Journal of the Neurological Sciences* 262 (1–2) (November 15): 37–44. doi:10.1016/j.jns.2007.06.024.

Engelender, S., T. Wanner, J. J. Kleiderlein, K. Wakabayashi, S. Tsuji, H. Takahashi, R. Ashworth, R. L. Margolis, and C. A. Ross. 2000. “Organization of the Human Synphilin-1 Gene, a Candidate for Parkinson’s Disease.” *Mammalian Genome* 11 (9) (September 1): 763–766. doi:10.1007/s003350010123.

Exner, N., B. Treske, D. Paquet, K. Holmström, C. Schiesling, S. Gispert, I. Carballo-Carbajal, et al. 2007. “Loss-of-function of Human PINK1 Results in Mitochondrial Pathology and Can Be Rescued by Parkin.” *The Journal of Neuroscience: The*

Official Journal of the Society for Neuroscience 27 (45) (November 7): 12413–12418. doi:10.1523/JNEUROSCI.0719-07.2007.

Farrer, M., P. Chan, R. Chen, L. Tan, S. Lincoln, D. Hernandez, L. Forno, et al. 2001. "Lewy Bodies and Parkinsonism in Families with Parkin Mutations." *Annals of Neurology* 50 (3) (August 3): 293–300. doi:10.1002/ana.1132.

Farrer, M. J. 2006. "Genetics of Parkinson Disease: Paradigm Shifts and Future Prospects." *Nature Reviews Genetics* 7 (4) (April): 306–318. doi:Article.

Fearnley, J.M, and A.J Lees. 1991. "Ageing and Parkinson's Disease: Substantia Nigra Regional Selectivity." *Brain* 114 (5): 2283-2301. doi:10/1093/brain/114.5.2283.

Ferraris, A., T. Ialongo, G. C. Passali, M. T Pellecchia, L. Brusa, M. Laruffa, A. Guidubaldi, et al. 2009. "Olfactory Dysfunction in Parkinsonism Caused by PINK1 Mutations." *Movement Disorders: Official Journal of the Movement Disorder Society* 24 (16) (December 15): 2350–2357. doi:10.1002/mds.22816.

Ferri, C. P., M. Prince, C. Brayne, H. Brodaty, L. Fratiglioni, M. Ganguli, K. Hall, et al. 2005. "Global Prevalence of Dementia: a Delphi Consensus Study." *The Lancet* 366 (9503) (December 17): 2112–2117. doi:10.1016/S0140-6736(05)67889-0.

Fleming, S. M., J. Salcedo, P-O. Fernagut, E. Rockenstein, E. Masliah, M. S. Levine, and M-F. Chesselet. 2004. "Early and Progressive Sensorimotor Anomalies in Mice Overexpressing Wild-type Human Alpha-synuclein." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 24 (42) (October 20): 9434–9440. doi:10.1523/JNEUROSCI.3080-04.2004.

Di Fonzo, A., C.Tassorelli, M. De Mari, H. F. Chien, J. Ferreira, C. F. Rohé, G. Riboldazzi, et al. 2005. "Comprehensive Analysis of the LRRK2 Gene in Sixty Families with Parkinson's Disease." *European Journal of Human Genetics* 14 (3) (December 7): 322–331. doi:10.1038/sj.ejhg.5201539.

Fortin, D. L., M. D. Troyer, K. Nakamura, S. Kubo, M. D. Anthony, and R. H. Edwards. 2004. "Lipid Rafts Mediate the Synaptic Localization of Alpha-synuclein." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 24 (30) (July 28): 6715–6723. doi:10.1523/JNEUROSCI.1594-04.2004.

Fuchs, J., C. Nilsson, J. Kachergus, M. Munz, E-M. Larsson, B. Schüle, J. W. Langston, et al. 2007. "Phenotypic Variation in a Large Swedish Pedigree Due to SNCA Duplication and Triplication." *Neurology* 68 (12) (March 20): 916–922. doi:10.1212/01.wnl.0000254458.17630.c5.

Gandhi, P.N., S. G. Chen, and A. L. Wilson-Delfosse. 2009. "Leucine-rich Repeat Kinase 2 (LRRK2): A Key Player in the Pathogenesis of Parkinson's Disease." *Journal of Neuroscience Research* 87 (6) (May 1): 1283–1295. doi:10.1002/jnr.21949.

Gardet, A., Y. Benita, C. Li, B. E. Sands, I. Ballester, C. Stevens, J. R. Korzenik, et al. 2010. "LRRK2 Is Involved in the IFN-gamma Response and Host Response to Pathogens." *Journal of Immunology (Baltimore, Md.: 1950)* 185 (9) (November 1): 5577–5585. doi:10.4049/jimmunol.1000548.

Gasser, T. 2010. "Identifying PD-causing Genes and Genetic Susceptibility Factors: Current Approaches and Future Prospects." *Progress in Brain Research* 183 (C): 2–20. doi:10.1016/S0079-6123(10)83001-8.

Gasser, Thomas. 2001. "Genetics of Parkinson's Disease." *Journal of Neurology* 248 (April 5): 833–840.

Gehrke, S., Y. Imai, N. Sokol, and B. Lu. 2010. "Pathogenic LRRK2 Negatively Regulates microRNA-mediated Translational Repression." *Nature* 466 (7306) (July 29): 637–641. doi:10.1038/nature09191.

Geisler, S., K. M. Holmström, A. Treis, D. Skujat, S. S. Weber, F. C. Fiesel, P. J. Kahle, and W. Springer. 2010. "The PINK1/Parkin-mediated Mitophagy Is Compromised by PD-associated Mutations." *Autophagy* 6 (7) (October): 871–878.

George, J. M. 2002. "The Synucleins." *GenomeBiology.com* 3 (1).

Gibb, W.R.G, and A.J Lees. 1988. "A Comparison of Clinical and Pathological Features of Young- and Old-onset Parkinson's Disease." *Neurology* 38 (9) (September): 1402-1406.

Gillies, G. E., and S.McArthur. 2010. "Independent Influences of Sex Steroids of Systemic and Central Origin in a Rat Model of Parkinson's Disease: A Contribution to Sex-specific Neuroprotection by Estrogens." *Hormones and Behavior* 57 (1) (January): 23–34. doi:10.1016/j.yhbeh.2009.06.002.

Glazov, E. A., A. Zankl, M. Donskoi, T. J. Kenna, G. P. Thomas, G. R. Clark, E. L. Duncan, and M. A. Brown. 2011. "Whole-Exome Re-Sequencing in a Family Quartet Identifies POP1 Mutations As the Cause of a Novel Skeletal Dysplasia." *PLoS Genet* 7 (3) (March 24): e1002027. doi:10.1371/journal.pgen.1002027.

Glickman, M. H., and A. Ciechanover. 2002. "The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction." *Physiological Reviews*, April 1.

Goedert, M., and M.G. Spillantini. 1998. “Lewy Body Diseases and Multiple System Atrophy as A-synucleinopathies.” *Molecular Psychiatry*, November.

Goldman, A. 1996. “Founder Effect and the Prevalence of Myotonic Dystrophy in South Africans: Molecular Studies.” *American Journal of Human Genetics* 59 (2): 445–452.

Goldman, S.M, C.M. Tanner, D. Oakes, G. S. Bhudhikanok, A. Gupta, and J. W. Langston. 2006. “Head Injury and Parkinson’s Disease Risk in Twins.” *Annals of Neurology* 60 (1) (July): 65–72. doi:10.1002/ana.20882.

Goldstein, D. B. 2009. “Common Genetic Variation and Human Traits.” *New England Journal of Medicine* 360 (17): 1696–1698. doi:10.1056/NEJMp0806284.

Greeff, J. M. 2007. “Deconstructing Jaco: Genetic Heritage of an Afrikaner.” *Annals of Human Genetics* 71 (5): 674–688. doi:Article.

Greene, J. C., A. J. Whitworth, I. Kuo, L. A. Andrews, M.B. Feany, and L. J. Pallanck. 2003. “Mitochondrial Pathology and Apoptotic Muscle Degeneration in Drosophila Parkin Mutants.” *Proceedings of the National Academy of Sciences* 100 (7) (April 1): 4078–4083. doi:10.1073/pnas.0737556100.

Groen, J. L., T. Kawarai, A. Toulina, C. Rivoiro, S. Salehi-Rad, C. Sato, A. Morgan, et al. 2004. “Genetic Association Study of PINK1 Coding Polymorphisms in Parkinson’s Disease.” *Neuroscience Letters* 372 (3) (December 6): 226–229. doi:10.1016/j.neulet.2004.09.043.

Guo, L., W. Wang, and S. G. Chen. 2006. “Leucine-rich Repeat Kinase 2: Relevance to Parkinson’s Disease.” *The International Journal of Biochemistry & Cell Biology* 38 (9) (January): 1469–1475. doi:10.1016/j.biocel.2006.02.009.

Gwinn-Hardy, K., A. Singleton, P. O’Suilleabhain, M. Boss, D. Nicholl, A. Adam, J. Hussey, P. Critchley, J. Hardy, and M. Farrer. 2001. “Spinocerebellar Ataxia Type 3 Phenotypically Resembling Parkinson Disease in a Black Family.” *Archives of Neurology* 58 (2) (February): 296–299.

Haas, R. H., F. Nasirian, K. Nakano, D. Ward, M. Pay, R. Hill, and C. W. Shults. 1995. “Low Platelet Mitochondrial Complex I and Complex II/III Activity in Early Untreated Parkinson’s Disease.” *Annals of Neurology* 37 (6) (June): 714–722. doi:10.1002/ana.410370604.

Hage-Sleiman, R., S. Herveau, E-L. Matera, J-F. Laurier, and C. Dumontet. 2010. "Tubulin Binding Cofactor C (TBCC) Suppresses Tumor Growth and Enhances Chemosensitivity in Human Breast Cancer Cells." *BMC Cancer* 10: 135. doi:10.1186/1471-2407-10-135.

Hague, S. 2003. "Early-onset Parkinson's Disease Caused by a Compound Heterozygous DJ-1 Mutation." *Annals of Neurology* 54 (2): 271–274. doi:10.1002/ana.10663.

Hague, S., E. Rogaeva, D. Hernandez, C. Gulick, A. Singleton, M. Hanson, J. Johnson, et al. 2003. "Early-onset Parkinson's Disease Caused by a Compound Heterozygous DJ-1 Mutation." *Annals of Neurology* 54 (2) (August): 271–274. doi:10.1002/ana.10663.

Hall, T.A. 1999. "BioEdit: a User-friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT." *Nucleic Acid Symosium Series* 41 (May): 95-98.

Haque, M. E., K. J. Thomas, C. D'Souza, S. Callaghan, T. Kitada, R. S. Slack, P. Fraser, M. R. Cookson, A. Tandon, and D.S. Park. 2008. "Cytoplasmic Pink1 Activity Protects Neurons from Dopaminergic Neurotoxin MPTP." *Proceedings of the National Academy of Sciences* 105 (5) (February 5): 1716–1721. doi:10.1073/pnas.0705363105.

Hatano, Y., Y. Li, K. Sato, S. Asakawa, Y. Yamamura, H. Tomiyama, H. Yoshino, et al. 2004. "Novel PINK1 Mutations in Early-onset Parkinsonism." *Annals of Neurology* 56 (3) (September): 424–427. doi:10.1002/ana.20251.

Hattori, N., T. Kitada, H. Matsumine, S. Asakawa, Y. Yamamura, H. Yoshino, T. Kobayashi, et al. 2004. "Molecular Genetic Analysis of a Novel Parkin Gene in Japanese Families with Autosomal Recessive Juvenile Parkinsonism: Evidence for Variable Homozygous Deletions in the Parkin Gene in Affected Individuals." *Annals of Neurology* 44 (6) (October 8): 935–941. doi:10.1002/ana.410440612.

Hayden, M. R. 1980. "The Prevalence of Huntington's Chorea in South Africa." *The South African Medical Journal* 58 (5): 193–196.

Haylett, W. L., R.J. Keyser, M. C. du Plessis, C. van der Merwe, J. Blanckenberg, D. Lombard, J.Carr, and S. Bardien. 2012. "Mutations in the Parkin Gene Are a Minor Cause of Parkinson's Disease in the South African Population." *Parkinsonism & Related Disorders* 18 (1) (January): 89–92. doi:10.1016/j.parkreldis.2011.09.022.

Healy, D. G., M. Falchi, S. S. O'Sullivan, V. Bonifati, A. Durr, S. Bressman, A. Brice, et al. 2008. "Phenotype, Genotype, and Worldwide Genetic Penetrance of LRRK2-

associated Parkinson's Disease: a Case-control Study." *Lancet Neurology* 7 (7) (July): 583–590. doi:10.1016/S1474-4422(08)70117-0.

Hedges, D. J., D. Hedges, D. Burges, E. Powell, C. Almonte, J. Huang, S. Young, et al. 2009. "Exome Sequencing of a Multigenerational Human Pedigree." *PloS One* 4 (12): e8232. doi:10.1371/journal.pone.0008232.

Hedrich, K., K. Marder, J. Harris, M. Kann, T. Lynch, H. Meija-Santana, P. P. Pramstaller, et al. 2002. "Evaluation of 50 Probands with Early-onset Parkinson's Disease for Parkin Mutations." *Neurology* 58 (8) (April 23): 1239–1246.

Hedrich, K., S. Winkler, J. Hagenah, K. Kabakci, M. Kasten, E. Schwinger, J. Volkmann, et al. 2006. "Recurrent LRRK2 (Park8) Mutations in Early-onset Parkinson's Disease." *Movement Disorders: Official Journal of the Movement Disorder Society* 21 (9) (September): 1506–1510. doi:10.1002/mds.20990.

Heemels, M-T. 2006. "Neurodegeneration." *Nature* 443 (7113) (October 18): 767–767. doi:10.1038/443767a.

Heese, J.A. 1971. *Die Herkoms Van Die Afrikaner*. A.A Balkema.

Hering, R., K. M. Strauss, X. Tao, A. Bauer, D.Woitalla, E-M. Mietz, S. Petrovic, et al. 2004. "Novel Homozygous p.E64D Mutation in DJ1 in Early Onset Parkinson Disease (PARK7)." *Human Mutation* 24 (4) (October): 321–329. doi:10.1002/humu.20089.

Hershko, A., and A. Ciechanover. 1998. "THE UBIQUITIN SYSTEM." *Annual Review of Biochemistry* 67 (1) (July): 425.

Hoepken, H.-H. 2007. "Mitochondrial Dysfunction, Peroxidation Damage and Changes in Glutathione Metabolism in PARK6." *Neurobiology of Disease* 25 (2): 401–411. doi:10.1016/j.nbd.2006.10.007.

Hoepken, H-H., S. Gispert, B. Morales, O. Wingerter, D. D. Turco, A. Mülsch, R. L. Nussbaum, et al. 2007. "Mitochondrial Dysfunction, Peroxidation Damage and Changes in Glutathione Metabolism in PARK6." *Neurobiology of Disease* 25 (2) (February): 401–411. doi:10.1016/j.nbd.2006.10.007.

Holmes, S. E., E. O'Hearn, A. Rosenblatt, C. Callahan, H. S. Hwang, R. G. Ingersoll-Ashworth, A. Fleisher, et al. 2001. "A Repeat Expansion in the Gene Encoding Junctophilin-3 Is Associated with Huntington Disease-like 2." *Nature Genetics* 29 (4) (December): 377–378. doi:10.1038/ng760.

Hong, D-P., A. L. Fink, and V.N. Uversky. 2009. "Smoking and Parkinson's Disease: Does Nicotine Affect Alpha-synuclein Fibrillation?" *Biochimica Et Biophysica Acta* 1794 (2) (February): 282–290. doi:10.1016/j.bbapap.2008.09.026.

Huynh, D. P., D. R. Scoles, T. H. Ho, D. Bigio, R. Marc, and S.M. Pulst. 2001. "Parkin Is Associated with Actin Filaments in Neuronal and Nonneural Cells." *Annals of Neurology* 48 (5) (April 4): 737–744. doi:10.1002/1531-8249(200011)48:5<737::AID-ANA7>3.0.CO;2-D.

Ibanez, P. 2006. "Mutational Analysis of the PINK1 Gene in Early-onset Parkinsonism in Europe and North Africa." *Brain* 129 (3) (January 6): 686–694. doi:10.1093/brain/awl005.

Isaias, I. U., and A. Antonini. 2010. "Single-photon Emission Computed Tomography in Diagnosis and Differential Diagnosis of Parkinson's Disease." *Neuro-degenerative Diseases* 7 (5): 319–329. doi:10.1159/000314498.

Ishihara, L., R. A. Gibson, L. Warren, R. Amouri, K. Lyons, C. Wielinski, C. Hunter, et al. 2007. "Screening for Lrrk2 G2019S and Clinical Comparison of Tunisian and North American Caucasian Parkinson's Disease Families." *Movement Disorders: Official Journal of the Movement Disorder Society* 22 (1) (January): 55–61. doi:10.1002/mds.21180.

Ito, S., W. Shirai, and T. Hattori. 2007. "Evaluating Posterolateral Linearization of the Putaminal Margin with Magnetic Resonance Imaging to Diagnose the Parkinson Variant of Multiple System Atrophy." *Movement Disorders: Official Journal of the Movement Disorder Society* 22 (4) (March 15): 578–581. doi:10.1002/mds.21329.

Jääskeläinen, T., H. Makkonen, T. Visakorpi, J. Kim, R. G. Roeder, and J. J. Palvimo. 2012. "Histone H2B Ubiquitin Ligases RNF20 and RNF40 in Androgen Signaling and Prostate Cancer Cell Growth." *Molecular and Cellular Endocrinology* 350 (1) (March): 87–98. doi:10.1016/j.mce.2011.11.025.

Jankovic, J. 2008. "Parkinson's Disease: Clinical Features and Diagnosis." *Journal of Neurology, Neurosurgery & Psychiatry* 79 (4) (April 1): 368–376. doi:10.1136/jnnp.2007.131045.

Jankovic, J. 2006. "An Update on the Treatment of Parkinson's Disease." *Mount Sinai Journal of Medicine* 73 (4) (July): 682–689. doi:Article.

Jankovic, J. 2012. “Current Concepts in Parkinson’s Disease and Other Movement Disorders.” *Current Opinion in Neurology* 25 (4) (August): 429–432. doi:10.1097/WCO.0b013e3283550cdd.

Jo, E., J. McLaurin, C. M. Yip, P. St George-Hyslop, and P.E. Fraser. 2000. “α-Synuclein Membrane Interactions and Lipid Specificity.” *Journal of Biological Chemistry* 275 (44) (November 3): 34328–34334. doi:10.1074/jbc.M004345200.

Julien M.A. 2004. *A Primer of Drug Action*. 10th ed. Worth Publishers.

Kahle, P. J., and C. Haass. 2004. “How Does Parkin Ligate Ubiquitin to Parkinson’s Disease?” *EMBO Reports* 5 (7) (July): 681–685. doi:10.1038/sj.embor.7400188.

Kann, M. 2002. “The Parkin Gene Is Not Involved in Late-onset Parkinson’s Disease [1] (multiple Letters).” *Neurology* 58 (5): 835.

Karayiorgou, M., M. Torrington, G. R. Abecasis, H. Pretorius, B. Robertson, S. Kaliski, S.Lay, et al. 2004. “Phenotypic Characterization and Genealogical Tracing in an Afrikaner Schizophrenia Database.” *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 124B (1): 20–28. doi:10.1002/ajmg.b.20090.

Karpinar, D. P., M. B. G. Balija, S. Kugler, F. Opazo, N. Rezaei-Ghaleh, N. Wender, H-Y. Kim, et al. 2009. “Pre-fibrillar [alpha]-synuclein Variants with Impaired [beta]-structure Increase Neurotoxicity in Parkinson’s Disease Models.” *EMBO J* 28 (20) (October 21): 3256–3268. doi:10.1038/emboj.2009.257.

Keyser, R. J., S. Lesage, A. Brice, J. Carr, and S. Bardien. 2010. “Assessing the Prevalence of PINK1 Genetic Variants in South African Patients Diagnosed with Early- and Late-onset Parkinson’s Disease.” *Biochemical and Biophysical Research Communications* 398 (1) (July): 125–129. doi:10.1016/j.bbrc.2010.06.049.

Keyser, R., D.Lombard, R. Veikondis, J. Carr, and S. Bardien. 2010. “Analysis of Exon Dosage Using MLPA in South African Parkinson’s Disease Patients.” *Neurogenetics* 11 (3): 305–312. doi:10.1007/s10048-009-0229-6.

Keyser, R., L. van der Merwe, M. Venter, C. Kinnear, L. Warnich, J. Carr, and S. Bardien. 2009. “Identification of a Novel Functional Deletion Variant in the 5’-UTR of the DJ-1 Gene.” *BMC Medical Genetics* 10 (1) (October 13): 105. doi:10.1186/1471-2350-10-105.

Khan, N. L., S. Jain, J. M. Lynch, N. Pavese, P. Abou-Sleiman, J. L. Holton, D. G. Healy, et al. 2005. "Mutations in the Gene LRRK2 Encoding Dardarin (PARK8) Cause Familial Parkinson's Disease: Clinical, Pathological, Olfactory and Functional Imaging and Genetic Data." *Brain* 128 (12) (December 1): 2786–2796. doi:10.1093/brain/awh667.

Kim, D. W. 2010. "Whole Human Exome Capture for High-throughput Sequencing." *Genome* 53 (7): 568–574. doi:10.1139/G10-025.

Kim, S. W., H. S. Ko, V. L. Dawson, and T.M. Dawson. 2005. "Recent Advances in Our Understanding of Parkinson's Disease." *Drug Discovery Today: Disease Mechanisms* 2 (4): 427–433. doi:10.1016/j.ddmec.2005.11.015.

Kitada, T., S. Asakawa, N. Hattori, H. Matsumine, Y. Yamamura, S. Minoshima, M. Yokochi, Y. Mizuno, and N.Shimizu. 1998. "Mutations in the Parkin Gene Cause Autosomal Recessive Juvenile Parkinsonism." *Nature* 392 (6676) (April 9): 605–608. doi:10.1038/33416.

Klein, C. 2007. "Impact of Recent Genetic Findings in Parkinson's Disease." *Current Opinion in Neurology* 20 (4): 453–464. doi:10.1097/WCO.0b013e3281e6692b.

Klein, C., K. Hedrich, C. Wellenbrock, M. Kann, J. Harris, K. Marder, A. E. Lang, et al. 2003. "Frequency of Parkin Mutations in Late-onset Parkinson's Disease." *Annals of Neurology* 54 (3): 415–416. doi:10.1002/ana.10737.

Klein, C., S. A. Schneider, and A. E. Lang. 2009. "Hereditary Parkinsonism: Parkinson Disease look-alikes—An Algorithm for Clinicians to 'PARK' Genes and Beyond." *Movement Disorders* 24 (14): 2042–2058. doi:10.1002/mds.22675.

Kobelka, C. E. 2010. "Exome Sequencing: Expanding the Genetic Testing Toolbox." *Clinical Genetics* 78 (2): 132–134. doi:10.1111/j.1399-0004.2010.01452.x.

Koziorowski, D., D. Hoffman-Zacharska, J. Sławek, W. Szirkowiec, P. Janik, J. Bal, and A.Friedman. 2010. "Low Frequency of the PARK2 Gene Mutations in Polish Patients with the Early-onset Form of Parkinson Disease." *Parkinsonism & Related Disorders* 16 (2) (February): 136–138. doi:10.1016/j.parkreldis.2009.06.010.

Kruger, R. 1998. "Ala30Pro Mutation in the Gene Encoding A-synuclein in Parkinson's Disease." *Nature Genetics* 18 (2): 106–108.

Ku, C-S., N. Naidoo, and Y.Pawitan. 2011. "Revisiting Mendelian Disorders Through Exome Sequencing." *Human Genetics* 129 (4): 351–370. doi:10.1007/s00439-011-0964-2.

Kwak, Y-D., B. Wang, J. J. Li, R. Wang, Q. Deng, S. Diao, Y. Chen, et al. 2012. "Upregulation of the E3 Ligase NEDD4-1 by Oxidative Stress Degrades IGF-1 Receptor Protein in Neurodegeneration." *The Journal of Neuroscience* 32 (32) (August 8): 10971–10981. doi:10.1523/JNEUROSCI.1836-12.2012.

Langston, J. W. 1983. "Parkinson's Disease in a Chemist Working with 1-methyl-4-phenyl-1,2,5,6-tetrahydropyridine." *The New England Journal of Medicine* 309 (5): 310.

Lansbury, P. T., and H. A. Lashuel. 2006. "A Century-old Debate on Protein Aggregation and Neurodegeneration Enters the Clinic." *Nature* 443 (7113) (October 19): 774–779. doi:10.1038/nature05290.

Lanska, D. J., L. Lavine, M. J. Lanska, and B. S. Schoenberg. 1988. "Huntington's Disease Mortality in the United States." *Neurology* 38 (5) (May 1): 769–769. doi:10.1212/WNL.38.5.769.

Lee, M. K., and D. L. Price. 2001. "Advances in Genetic Models of Parkinson's Disease." *Clinical Neuroscience Research* 1 (6) (December): 456–466. doi:10.1016/S1566-2772(01)00024-X.

Lee, S-J., S. J. Kim, I-K. Kim, J. Ko, C-S. Jeong, G-H. Kim, C. Park, et al. 2003. "Crystal Structures of Human DJ-1 and Escherichia Coli Hsp31, Which Share an Evolutionarily Conserved Domain." *Journal of Biological Chemistry* 278 (45) (November 7): 44552–44559. doi:10.1074/jbc.M304517200.

Lesage, S., and A. Brice. 2009. "Parkinson's Disease: From Monogenic Forms to Genetic Susceptibility Factors." *Human Molecular Genetics* 18 (R1) (April 15): R48–R59. doi:10.1093/hmg/ddp012.

Lesage, S., E. Patin, C. Condroyer, A. L. Leutenegger, E. Lohmann, N. Giladi, A. Bar-Shira, et al. 2010. "Parkinson's Disease-related LRRK2 G2019S Mutation Results from Independent Mutational Events in Humans." *Human Molecular Genetics* 19 (10) (March 2): 1998–2004. doi:10.1093/hmg/ddq081.

Leutenegger, A.-L. 2006. "Juvenile-onset Parkinsonism as a Result of the First Mutation in the Adenosine Triphosphate Orientation Domain of PINK1." *Archives of Neurology* 63 (9): 1257–1261. doi:10.1001/archneur.63.9.1257.

Levine, B., N. Mizushima, and H. W. Virgin. 2011. “Autophagy in Immunity and Inflammation.” *Nature* 469 (7330) (January 20): 323–335. doi:10.1038/nature09782.

Lin, X., L. Parisiadou, X-L. Gu, L. Wang, H. Shim, L. Sun, C. Xie, et al. 2009. “Leucine-Rich Repeat Kinase 2 Regulates the Progression of Neuropathology Induced by Parkinson’s-Disease-Related Mutant A-synuclein.” *Neuron* 64 (6) (December 24): 807–827. doi:10.1016/j.neuron.2009.11.006.

Lohmann, E., M. Periquet, V. Bonifati, N. W. Wood, G. De Michele, A-M. Bonnet, V. Fraix, et al. 2003. “How Much Phenotypic Variation Can Be Attributed to Parkin Genotype?” *Annals of Neurology* 54 (2) (May 30): 176–185. doi:10.1002/ana.10613.

Love, S. 2005. “Neuropathological Investigation of Dementia: a Guide for Neurologists.” *Journal of Neurology, Neurosurgery & Psychiatry* 76 (suppl_5) (December 1): v8–v14. doi:10.1136/jnnp.2005.080754.

Luecking, C. B., A. Duerr, V. Bonifati, J. Vaughan, G. De Michele, T. Gasser, B. S. Harhangi, et al. 2000. “Association between early-onset Parkinson’s disease and mutations in the parkin gene.” *New England Journal of Medicine* 342 (21): 1560–1567.

MacLeod, D., J. Dowman, R. Hammond, T. Leete, K. Inoue, and A. Abeliovich. 2006. “The Familial Parkinsonism Gene LRRK2 Regulates Neurite Process Morphology.” *Neuron* 52 (4) (November): 587–593. doi:10.1016/j.neuron.2006.10.008.

Magazi, D. S., A. Krause, V. Bonev, M. Moagi, Z. Iqbal, M. Dludla, and C. H. van der Meyden. 2008. “Huntington’s Disease Genetic Heterogeneity In African Patients.” *South African Medical Journal* 98 (3) (October 1): 200. Ethnicity (black subjects). doi:10.7196/samj.289.

Mann, V. M., J. M. Cooper, S. E. Daniel, K. Srai, P. Jenner, C. D. Marsden, and A. H. V. Schapira. 1994. “Complex I, Iron, and Ferritin in Parkinson’s Disease Substantia Nigra.” *Annals of Neurology* 36 (6): 876–881. doi:10.1002/ana.410360612.

Margolis, R. L. 2001. “A Disorder Similar to Huntington’s Disease Is Associated with a Novel CAG Repeat Expansion.” *Annals of Neurology* 50 (3): 373–380. doi:10.1002/ana.1124.

Margolis, R. L. 2003. “Huntington’s Disease-like 2: A Clinical, Pathological, and Molecular Comparison to Huntington’s Disease.” *Clinical Neuroscience Research* 3 (3): 187–196. doi:10.1016/S1566-2772(03)00061-6.

Margolis, R. L. 2005. “Huntington’s Disease Like-2: Review and Update.” *Acta Neurologica Sinica* 14 (1): 1–8.

Margolis, R. L., S. E. Holmes, A. Rosenblatt, L. Gourley, E. O’Hearn, C. A. Ross, W. K. Seltzer, et al. 2004. “Huntington’s Disease-like 2 (HDL2) in North America and Japan.” *Annals of Neurology* 56 (5) (November): 670–674. doi:10.1002/ana.20248.

Mata, I. F., J. M. Kachergus, J. P. Taylor, S. Lincoln, J. Aasly, T. Lynch, M. M. Hulihan, et al. 2005. “Lrrk2 Pathogenic Substitutions in Parkinson’s Disease.” *Neurogenetics* 6 (4) (September 20): 171–177. doi:10.1007/s10048-005-0005-1.

Mathers, C., D. M. Fat, World Health Organization, and J. T. Boerma. 2008. *The Global Burden of Disease: 2004 Update*. World Health Organization.

Matsumine, H., M. Saito, S. Shimoda-Matsubayashi, H. Tanaka, A. Ishikawa, Y. Nakagawa-Hattori, M. Yokochi, et al. 1997. “Localization of a Gene for an Autosomal Recessive Form of Juvenile Parkinsonism to Chromosome 6q25.2-27.” *American Journal of Human Genetics* 60 (3) (March): 588.

McCormack, A. L. 2003. “Effects of L-dopa and Other Amino Acids Against Paraquat-induced Nigrostriatal Degeneration.” *Journal of Neurochemistry* 85 (1): 82–86.

McCormack, A. L., S. K. Mak, J. M. Henderson, D. Bumcrot, M. J. Farrer, and D. A. Di Monte. 2010. “α-Synuclein Suppression by Targeted Small Interfering RNA in the Primate Substantia Nigra.” *PLoS ONE* 5 (8) (August 11): e12122. doi:10.1371/journal.pone.0012122.

McInerney-Leo, A., K. Gwinn-Hardy, and R. L. Nussbaum. 2004. “Prevalence of Parkinson’s Disease in Populations of African Ancestry: a Review.” *Journal of the National Medical Association* 96 (7) (July): 974–979.

Means, A. R. 2001. “Defining Ca2 /Calmodulin-dependent Protein Kinase Cascades in Transcriptional Regulation.” *Journal of Biological Chemistry* 276 (5): 2975–2978. doi:10.1074/jbc.R000027200.

Melcon, M. O., D. W. Anderson, R. H. Vergara, and W. A. Rocca. 1997. “Prevalence of Parkinson’s Disease in Junín, Buenos Aires Province, Argentina.” *Movement Disorders* 12 (2) (March 1): 197–205. doi:10.1002/mds.870120210.

Metzenberg S. 2008. “Sanger Method.” http://www.csun.edu/%7Ehcbio027/biotechnology/lec3/sanger.html.

Miller, D. W., S. M. Hague, J. Clarimon, M. Baptista, K. Gwinn-Hardy, M. R. Cookson, and A. B. Singleton. 2004. "α-Synuclein in Blood and Brain from Familial Parkinson Disease with SNCA Locus Triplication." *Neurology* 62 (10) (May 25): 1835–1838. doi:10.1212/01.WNL.0000127517.33208.F4.

Miyai, I., M. Ito, N. Hattori, M. Mihara, M. Hatakenaka, H. Yagura, G. Sobue, M. Nishizawa, and for the Cerebellar Ataxia Rehabilitation Trialists Collaboration. 2011. "Cerebellar Ataxia Rehabilitation Trial in Degenerative Cerebellar Diseases." *Neurorehabilitation and Neural Repair* 26 (5) (December 2): 515–522. doi:10.1177/1545968311425918.

Moore, D. J., A. B. West, V. L. Dawson, and T. M. Dawson. 2005. "Molecular Pathophysiology of Parkinson's Disease." *Annual Review of Neuroscience* 28 (1) (July): 57–87. doi:Article.

Morris, C. M., K.K. O'Brien, A. M. Gibson, J. A. Hardy, and A. Singleton. 2003. "Polymorphism in the Human DJ-1 Gene Is Not Associated with Sporadic Dementia with Lewy Bodies or Parkinson's Disease." *Neuroscience Letters* 352 (2) (December 4): 151–153. doi:10.1016/j.neulet.2003.08.037

Mshana, G., C. Dotchin, and R. Walker. 2011. "'We Call It the Shaking Illness': Perceptions and Experiences of Parkinson's Disease in Rural Northern Tanzania." *BMC Public Health* 11 (1) (April 8): 219. doi:10.1186/1471-2458-11-219.

Mueller, J. C., J. Fuchs, A. Hofer, A. Zimprich, P. Lichtner, T. Illig, D. Berg, U. Wüllner, T. Meitinger, and T. Gasser. 2005. "Multiple Regions of A-synuclein Are Associated with Parkinson's Disease." *Annals of Neurology* 57 (4) (April): 535–541. doi:10.1002/ana.20438.

Nagakubo, D., T. Taira, H. Kitaura, M. Ikeda, K. Tamai, S. M.M. Iguchi-Ariga, and H.Ariga. 1997. "DJ-1, a Novel Oncogene Which Transforms Mouse NIH3T3 Cells in Cooperation Withras." *Biochemical and Biophysical Research Communications* 231 (2) (February 13): 509–513. doi:10.1006/bbrc.1997.6132.

Naimark, D., E. Jackson, E. Rockwell, and D.V. Jeste. 1996. "Psychotic Symptoms in Parkinson's Disease Patients with Dementia." *American Geriatric Society* 44 (3) (March): 296-299.

Narendra, D., A. Tanaka, D-F. Suen, and R. J. Youle. 2008. "Parkin Is Recruited Selectively to Impaired Mitochondria and Promotes Their Autophagy." *Journal of Cell Biology* 183 (5) (December 1): 795–803. doi:10.1083/jcb.200809125.

Ng, S. B., K. J. Buckingham, C. Lee, A. W. Bigham, H. K. Tabor, K. M. Dent, C. D. Huff, et al. 2010. "Exome Sequencing Identifies the Cause of a Mendelian Disorder." *Nature Genetics* 42 (1) (January): 30–35. doi:Article.

Ng, S. B., E. H. Turner, P. D. Robertson, S. D. Flygare, A. W. Bigham, C. Lee, T. Shaffer, et al. 2009. "Targeted Capture and Massively Parallel Sequencing of 12 Human Exomes." *Nature* 461 (7261) (August 16): 272–276. doi:10.1038/nature08250.

Nicklas, W. J. 1987. "MPTP, MPP and Mitochondrial Function." *Life Sciences* 40 (8): 721–729.

Oakley, A. E., J. F. Collingwood, J. Dobson, G. Love, H. R. Perrott, J. A. Edwardson, M. Elstner, and C. M. Morris. 2007. "Individual Dopaminergic Neurons Show Raised Iron Levels in Parkinson Disease." *Neurology* 68 (21) (May 22): 1820–1825. doi:10.1212/01.wnl.0000262033.01945.9a.

Okubadejo, N., A. Britton, C. Crews, R. Akinyemi, J. Hardy, A. Singleton, and J. Bras. 2008. "Analysis of Nigerians with Apparently Sporadic Parkinson Disease for Mutations in LRRK2, PRKN and ATXN3." Ed. Amanda Ewart Toland. *PLoS ONE* 3 (10) (October 17): e3421. doi:10.1371/journal.pone.0003421.

Okubadejo, N. U. 2008. "An Analysis of Genetic Studies of Parkinson's Disease in Africa." *Parkinsonism & Related Disorders* 14 (3) (April): 177–182. doi:10.1016/j.parkreldis.2007.08.006.

Okubadejo, N. U., J. H. Bower, W. A. Rocca, and D. M. Maraganore. 2006a. "Parkinson's Disease in Africa: A Systematic Review of Epidemiologic and Genetic Studies." *Movement Disorders* 21 (12): 2150–2156. doi:10.1002/mds.21153.

Olzmann, J. A., K. Brown, K. D. Wilkinson, H. D. Rees, Q. Huai, H. Ke, A. I. Levey, L. Li, and L-S Chin. 2004. "Familial Parkinson's Disease-associated L166P Mutation Disrupts DJ-1 Protein Folding and Function." *Journal of Biological Chemistry* 279 (9) (February 27): 8506–8515. doi:10.1074/jbc.M311017200.

Orr Urtreger, A. 2007. "The LRRK2 G2019S Mutation in Ashkenazi Jews with Parkinson Disease: Is There a Gender Effect?" *Neurology* 69 (16): 1595–1602. doi:10.1212/01.wnl.0000277637.33328.d8.

Osuntokun, B., A. Adeuja, B. Schoenberg, O. Bademosi, V. Nottidge, A. Olumide, O. Lge, F. Yaria, and C. Bolis. 1987. "Neurological Disorders in Nigerian Africans: a Community-based Study." *Acta Neurologica Scandinavica* 75 (1): 13–21. doi:10.1111/j.1600-0404.1987.tb07883.x.

Ozelius, L. J., G. Senthil, R. Saunders-Pullman, E. Ohmann, A. Deligtisch, M. Tagliati, A. L. Hunt, et al. 2006. "LRRK2 G2019S as a Cause of Parkinson's Disease in Ashkenazi Jews." *New England Journal of Medicine* 354 (4): 424–425. doi:10.1056/NEJMc055509.

Paisán-Ruíz, C., A. E. Lang, T. Kawarai, C. Sato, S. Salehi-Rad, G. K. Fisman, T. Al-Khairallah, P. St George-Hyslop, A. Singleton, and E. Rogaeva. 2005. "LRRK2 Gene in Parkinson Disease Mutation Analysis and Case Control Association Study." *Neurology* 65 (5) (September 13): 696–700. doi:10.1212/01.WNL.0000167552.79769.b3.

Palop, J. J., J. Chin, and L. Mucke. 2006. "A Network Dysfunction Perspective on Neurodegenerative Diseases." *Nature* 443 (7113) (October 19): 768–773. doi:10.1038/nature05289.

Pan, T., S. Kondo, W. Le, and J. Jankovic. 2008. "The Role of Autophagy-lysosome Pathway in Neurodegeneration Associated with Parkinson's Disease." *Brain* 131 (8) (January 10): 1969–1978. doi:10.1093/brain/awm318.

Papapetropoulos, S., A. A. Argyriou, G. Mitsi, and E. Chroni. 2007. "The G2019S LRRK2 Mutation Is Uncommon Amongst Greek Patients with Familial Parkinson's Disease." *European Journal of Neurology* 14 (11): e6–e6. doi:10.1111/j.1468-1331.2007.01933.x.

Park, J. 2006. "Mitochondrial Dysfunction in Drosophila PINK1 Mutants Is Complemented by Parkin." *Nature* 441 (7097): 1157–1161. doi:10.1038/nature04788.

Parkinson, J. 1817. "An Essay on the Shaking Palsy." *Journal of Neuropsychiatry* 14 (2): 223–236. doi:10.1176/appi.neuropsych.14.2.223.

Parkinson's Disease Foundation. 2012. "Nonmotor Symptoms - Parkinson's Disease Foundation (PDF)." Accessed June 24. http://www.pdf.org/symptoms_nonmotor_early.

Patterson, N., D. C. Petersen, V. Der Ross, E. Richard, H. Sudoyo, R. H. Glashoff, S. Marzuki, D. Reich, and V. M. Hayes. 2010. "Genetic Structure of a Unique Admixed Population: Implications for Medical Research." *Human Molecular Genetics* 19 (3) (February 1): 411–419. doi:10.1093/hmg/ddp505.

Pérez-Pérez, J. M., O. Serralbo, M. Vanstraelen, C. González, M-C. Criqui, P. Genschik, E. Kondorosi, and B. Scheres. 2008. "Specialization of CDC27 Function in the

Arabidopsis Thaliana Anaphase-promoting Complex (APC/C)." *The Plant Journal* 53 (1): 78–89. doi:10.1111/j.1365-313X.2007.03312.x.

Periquet, M., M. Latouche, E. Lohmann, N. Rawal, G. De Michele, S. Ricard, H. Teive, et al. 2003. "Parkin Mutations Are Frequent in Patients with Isolated Early-onset Prkinsonism." *Brain: a Journal of Neurology* 126 (Pt 6) (June): 1271–1278.

Petrucelli, L., C. O'Farrell, P. J. Lockhart, M. Baptista, K. Kehoe, L. Vink, P. Choi, et al. 2002. "Parkin Protects Against the Toxicity Associated with Mutant Alpha-synuclein: Proteasome Dysfunction Selectively Affects Catecholaminergic Neurons." *Neuron* 36 (6) (December 19): 1007–1019.

Polymeropoulos, M. H., J. J. Higgins, L. I. Golbe, W. G. Johnson, S. E. Ide, G. Di Irio, G. Sanges, et al. 1996. "Mapping of a Gene for Parkinson's Disease to Chromosome 4q21-q23." *Science*, November 15.

Pramstaller, P. P., M. G. Schlossmacher, T. S. Jacques, F. Scaravilli, C. Eskelson, I. Pepivani, K. Hedrich, et al. 2005. "Lewy Body Parkinson's Disease in a Large Pedigree with 77 Parkin Mutation Carriers." *Annals of Neurology* 58 (3) (August 29): 411–422. doi:10.1002/ana.20587.

Pridgeon, J. W., J. A. Olzmann, L-S. Chin, and L. Li. 2007. "PINK1 Protects Against Oxidative Stress by Phosphorylating Mitochondrial Chaperone TRAP1." *PLoS Biology* 5 (7) (July). doi:10.1371/journal.pbio.0050172. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1892574/.

Robinson, P.N., P. Krawitz, and S. Mundlos. 2011. "Strategies for Exome and Genome Sequence Data Analysis in Disease-gene Discovery Projects." *Clinical Genetics* 80 (2) (August): 127–132. doi:10.1111/j.1399-0004.2011.01713.x.

Rogaeva, E., J. Johnson, A. E. Lang, C. Gulick, K. Gwinn-Hardy, T. Kawarai, C. Sato, et al. 2004. "Analysis of the PINK1 Gene in a Large Cohort of Cases with Parkinson Disease." *Archives of Neurology* 61 (12) (December): 1898–1904. doi:10.1001/archneur.61.12.1898.

Rohe, C. F. 2004. "Homozygous PINK1 C-terminus Mutation Causing Early-onset Parkinsonism." *Annals of Neurology* 56 (3): 427–431. doi:10.1002/ana.20247.

Ross, O. A, A. I. Soto-Ortolaza, M. G. Heckman, J. O. Aasly, N. Abahuni, G. Annesi, J. A. Bacon, et al. 2011. "Association of LRRK2 Exonic Variants with Susceptibility to Parkinson's Disease: a Case-control Study." *Lancet Neurology* 10 (10) (October): 898–908. doi:10.1016/S1474-4422(11)70175-2.

Ross, O. A., Y-R. Wu, M-C. Lee, M. Funayama, M-L. Chen, A. I. Soto, I. F. Mata, et al. 2008. "Analysis of Lrrk2 R1628P as a Risk Factor for Parkinson's Disease." *Annals of Neurology* 64 (1) (July): 88–92. doi:10.1002/ana.21405.

Rothfuss, O., H. Fischer, T. Hasegawa, M. Maisel, P. Leitner, F. Miesel, M. Sharma, et al. 2009. "Parkin Protects Mitochondrial Genome Integrity and Supports Mitochondrial DNA Repair." *Human Molecular Genetics* 18 (20) (October 15): 3832–3850. doi:10.1093/hmg/ddp327.

Rozen, S. 2000. "Primer3 on the WWW for General Users and for Biologist Programmers." *Methods in Molecular Biology* 132: 365–386.

Rubinsztein, D. C. 2006. "The Roles of Intracellular Protein-degradation Pathways in Neurodegeneration." *Nature* 443 (7113) (October 19): 780–786. doi:10.1038/nature05291.

Samii, A., J. G. Nutt, and B. R. Ransom. 2004. "Parkinson's Disease." *The Lancet* 363 (9423) (May 29): 1783–1793. doi:10.1016/S0140-6736(04)16305-8.

Santamaria, A. B., and S. I. Sulsky. 2010. "Risk Assessment of an Essential Element: Manganese." *Journal of Toxicology and Environmental Health. Part A* 73 (2): 128–155. doi:10.1080/15287390903337118.

Sanyal, J., D. P. Chakraborty, B. Sarkar, T. K. Banerjee, S. C. Mukherjee, B. C. Ray, and V.R. Rao. 2010. "Environmental and Familial Risk Factors of Parkinsons Disease: Case-control Study." *The Canadian Journal of Neurological Sciences. Le Journal Canadien Des Sciences Neurologiques* 37 (5) (September): 637–642.

Sarkar, A.A., and I. E. Zohn. 2012. "Hectd1 Regulates Intracellular Localization and Secretion of Hsp90 to Control Cellular Behavior of the Cranial Mesenchyme." *The Journal of Cell Biology* 196 (6) (March 19): 789–800. doi:10.1083/jcb.201105101.

Satake, W., Y.Nakabayashi, I. Mizuta, Y. Hirota, C. Ito, M. Kubo, T. Kawaguchi, et al. 2009. "Genome-wide Association Study Identifies Common Variants at Four Loci as Genetic Risk Factors for Parkinson's Disease." *Nature Genetics* 41 (12) (November 15): 1303–1307. doi:10.1038/ng.485.

Satoh, J., and Y. Kuroda. 1999. "Association of Codon 167 Ser/Asn Heterozygosity in the Parkin Gene with Sporadic Parkinson's Disease." *Neuroreport* 10 (13) (September 9): 2735–2739.

Saxena, S., G.Thornicroft, M. Knapp, and H. Whiteford. 2008. "Resources for Mental Health: Scarcity, Inequity, and Inefficiency." *The Lancet* 370 (9590): 878–889. doi:10.1016/S0140-6736(07)61239-2.

Schapira, A. H. V. 1997. "Complex I Function in Familial and Sporadic Dystonia." *Annals of Neurology* 41 (4): 556–559. doi:10.1002/ana.410410421.

Schapira, A.H.V. 2010. "Complex I: Inhibitors, Inhibition and Neurodegeneration." *Experimental Neurology* 224 (2) (August): 331–335. doi:10.1016/j.expneurol.2010.03.028.

Sellbach, A. N., R. S. Boyle, P. A. Silburn, and G. D. Mellick. 2006. "Parkinson's Disease and Family History." *Parkinsonism & Related Disorders* 12 (7) (October): 399–409. doi:10.1016/j.parkreldis.2006.03.002.

Shimura, H., N.Hattori, S. Kubo, Y. Mizuno, S. Asakawa, S. Minoshima, N. Shimizu, et al. 2000. "Familial Parkinson Disease Gene Product, Parkin, Is a Ubiquitin-protein Ligase." *Nature Genetics* 25 (3) (July): 302.

Shin, N., H. Jeong, J. Kwon, H. Y. Heo, J.J. Kwon, H. J. Yun, C-H. Kim, et al. 2008. "LRRK2 Regulates Synaptic Vesicle Endocytosis." *Experimental Cell Research* 314 (10) (June 10): 2055–2065. doi:10.1016/j.yexcr.2008.02.015.

Siderowf, A. 2001. "Parkinson's Disease: Clinical Features, Epidemiology and Genetics." *Neurologic Clinics* 19 (3) (August): 565–578, vi.

Sim, C. H., D. S. S. Lio, S. S. Mok, C. L. Masters, A. F. Hill, J. G. Culvenor, and H-C Cheng. 2006. "C-terminal Truncation and Parkinson's Disease-associated Mutations Down-regulate the Protein Serine/threonine Kinase Activity of PTEN-induced Kinase-1." *Human Molecular Genetics* 15 (21) (November 1): 3251–3262. doi:10.1093/hmg/ddl398.

Simón-Sánchez, J., C. Schulte, J. M. Bras, M. Sharma, J. R. Gibbs, D. Berg, C. Paisan-Ruiz, et al. 2009. "Genome-wide Association Study Reveals Genetic Risk Underlying Parkinson's Disease." *Nature Genetics* 41 (12): 1308–1312. doi:10.1038/ng.487.

Siu, F. K. Y., L. T. O. Lee, and B. K. C. Chow. 2008. "Southwestern Blotting in Investigating Transcriptional Regulation." *Nature Protocols* 3 (1): 51–58. doi:10.1038/nprot.2007.492.

Snyder, B. J., and C. Warren Olanow. 2005. “Stem Cell Treatment for Parkinson’s Disease: An Update for 2005.” *Current Opinion in Neurology* 18 (4) (August): 376–385.

Socal, M. P., V. E. Emmel, C. R. M. Rieder, A. Hilbig, M. L. Saraiva-Pereira, and L. B. Jardim. 2009. “Intrafamilial Variability of Parkinson Phenotype in SCAs: Novel Cases Due to SCA2 and SCA3 Expansions.” *Parkinsonism & Related Disorders* 15 (5) (June): 374–378. doi:10.1016/j.parkreldis.2008.09.005.

Sode, K., S. Ochiai, N. Kobayashi, and E. Usuzaka. 2007. “Effect of Reparation of Repeat Sequences in the Human Alpha-synuclein on Fibrillation Ability.” *International Journal of Biological Sciences* 3 (1): 1–7.

Storz, G., and J. A. Imlayt. 1999. “Oxidative Stress.” *Current Opinion in Microbiology* 2 (2) (April): 188–194. doi:10.1016/S1369-5274(99)80033-2.

Sun, M., J. C. Latourelle, G. Wooten, M. F. Lew, C. Klein, H. A. Shill, L. I. Golbe, et al. 2006. “Influence of Heterozygosity for Parkin Mutation on Onset Age in Familial Parkinson Disease: The GenePD Study.” *Archives of Neurology* 63 (6) (June): 826–832. doi:10.1001/archneur.63.6.826.

Surguchov, A., and K. W. Jeon. 2008. “Chapter 6 Molecular and Cellular Biology of Synucleins.” In *International Review of Cell and Molecular Biology*, 270:225–317.

Taira, T., Y. Saito, T. Niki, S. M. M. Iguchi-Ariga, K. Takahashi, and H. Ariga. 2004. “DJ-1 Has a Role in Antioxidative Stress to Prevent Cell Death.” *EMBO Reports* 5 (2) (February): 213–218. doi:10.1038/sj.embor.7400074.

Takahashi-Niki, K., T. Niki, T. Taira, S.M. M. Iguchi-Ariga, and H. Ariga. 2004. “Reduced Anti-oxidative Stress Activities of DJ-1 Mutants Found in Parkinson’s Disease Patients.” *Biochemical and Biophysical Research Communications* 320 (2) (July 23): 389–397. doi:10.1016/j.bbrc.2004.05.187.

Tan, E-K., and L. M. Skipper. 2007. “Pathogenic mutations in Parkinson disease.” *Human Mutation* 28 (7) (July): 641–653. doi:10.1002/humu.20507.

Tan, E-K., M. Tang, L. C. Tan, Y-R. Wu, R-M. Wu, O. A. Ross, and Y. Zhao. 2008. “Lrrk2 R1628P in non-Chinese Asian Races.” *Annals of Neurology* 64 (4) (October): 472–473. doi:10.1002/ana.21467.

Tan, X. 2010. “Risk Factors for Parkinson’s Disease and Interaction Analysis in Chinese Patients.” *Neural Regeneration Research* 5 (6): 476–480. doi:10.3969/j.issn.1673-5374.2010.06.013.

Tanner, C. M., F. Kamel, G. W. Ross, J.A. Hoppin, S. M. Goldman, M. Korell, C. Marras, et al. 2011. “Rotenone, Paraquat, and Parkinson’s Disease.” *Environmental Health Perspectives* 119 (6) (January 26): 866–872. doi:10.1289/ehp.1002839.

Tao, X., and L. Tong. 2003. “Crystal Structure of Human DJ-1, a Protein Associated with Early Onset Parkinson’s Disease.” *The Journal of Biological Chemistry* 278 (33) (August 15): 31372–31379. doi:10.1074/jbc.M304221200.

Taylor, J. P., J. Hardy, and K. H. Fischbeck. 2002. “Toxic Proteins in Neurodegenerative Disease.” *Science (New York, N.Y.)* 296 (5575) (June 14): 1991–1995. doi:10.1126/science.1067122.

Teer, J. K., and J. C. Mullikin. 2010. “Exome Sequencing: The Sweet Spot Before Whole Genomes.” *Human Molecular Genetics* 19 (R2) (August 12): R145–R151. doi:10.1093/hmg/ddq333.

Tekle Haimanot, R. 1990. “Community-based Study of Neurological Disorders in Rural Central Epithopia.” *Neuroepidemiology* 9 (5): 263–277.

Thomas, B. 2007. “Parkinson’s Disease.” *Human Molecular Genetics* 16 Spec No. 2. doi:10.1093/hmg/ddm159.

Thomas, B., and M. F. Beal. 2011. “Molecular Insights into Parkinson’s Disease.” *F1000 Medicine Reports* 3: 7. doi:10.3410/M3-7.

Tofaris, G. K. 2012. “Lysosome-dependent Pathways as a Unifying Theme in Parkinson’s Disease.” *Movement Disorders* 27 (11): 1364–1369. doi:10.1002/mds.25136.

Tofaris, G. K. 2012. “Lysosome-dependent Pathways as a Unifying Theme in Parkinson’s Disease.” *Movement Disorders* 27 (11): 1364–1369. doi:10.1002/mds.25136.

Tomiyama, H., Y. Li, H. Yoshino, Y. Mizuno, S. Kubo, T. Toda, and N. Hattori. 2009. “Mutation Analysis for DJ-1 in Sporadic and Familial Parkinsonism: Screening Strategy in Parkinsonism.” *Neuroscience Letters* 455 (3) (May 22): 159–161. doi:10.1016/j.neulet.2009.03.033.

Tsuji, S. 2010. “Genetics of Neurodegenerative Diseases: Insights from High-throughput Resequencing.” *Human Molecular Genetics* 19 (R1) (April 22): R65–R70. doi:10.1093/hmg/ddq162.

Twelves, D., K. S. M. Perkins, and C. Counsell. 2003. “Systematic Review of Incidence Studies of Parkinson’s Disease.” *Movement Disorders: Official Journal of the Movement Disorder Society* 18 (1) (January): 19–31. doi:10.1002/mds.10305.

Uéda, K., H. Fukushima, E. Masliah, Y. Xia, A. Iwai, M. Yoshimoto, D. A. Otero, J. Kondo, Y. Ihara, and T. Saitoh. 1993. “Molecular Cloning of cDNA Encoding an Unrecognized Component of Amyloid in Alzheimer Disease.” *Proceedings of the National Academy of Sciences of the United States of America* 90 (23) (December 1): 11282–11286.

University of Maryland Medical Centre. 2012. “Surgery as a Treatment for Parkinson’s Disease.” *University of Maryland Medical Center*. Accessed June 24. http://www.umm.edu/parkinsons/surgery.htm?src=fblike.

Unoki, M., and Y. Nakamura. 2001. “Growth-suppressive Effects of BPOZ and EGR2, Two Genes Involved in the PTEN Signaling Pathway.” *Oncogene* 20 (33) (July 27): 4457–4465. doi:10.1038/sj.onc.1204608.

Uversky, V. N., and D. Eliezer. 2009. “Biophysics of Parkinson’s Disease: Structure and Aggregation of Alpha-synuclein.” *Current Protein & Peptide Science* 10 (5) (October): 483–499.

Valente, E. M., P. M. Abou-Sleiman, V. Caputo, M. M. K. Muqit, K. Harvey, S. Gispert, Z. Ali, et al. 2004. “Hereditary Early-onset Parkinson’s Disease Caused by Mutations in PINK1.” *Science (New York, N.Y.)* 304 (5674) (May 21): 1158–1160. doi:10.1126/science.1096284.

Valente, E. M., A. R. Bentivoglio, P. H. Dixon, A. Ferraris, T. Ialongo, M. Frontali, A. Albanese, and N. W. Wood. 2001. “Localization of a Novel Locus for Autosomal Recessive Early-Onset Parkinsonism, PARK6, on Human Chromosome 1p35-p36.” *American Journal of Human Genetics* 68 (4) (April): 895–900.

Vancraenenbroeck, R., E. Lobbestael, S. D. Weeks, S. V. Strelkov, V. Baekelandt, J. M. Taymans, and M. De Maeyer. 2012. “Expression, Purification and Preliminary Biochemical and Structural Characterization of the Leucine Rich Repeat Namesake Domain of Leucine Rich Repeat Kinase 2.” *Biochimica Et Biophysica Acta* 1824 (3) (March): 450–460. doi:10.1016/j.bbapap.2011.12.009.

Vilariño-Güell, C., A. I. Soto, S. J. Lincoln, S. B. Yahmed, M. Kefi, M. G. Heckman, M. M. Hulihan, et al. 2008. "ATP13A2 Variability in Parkinson Disease." *Human Mutation* 30 (3) (December 9): 406–410. doi:10.1002/humu.20877.

Vilariño-Güell, C., C. Wider, O. A. Ross, J. C. Dachsel, J. M. Kachergus, S. J. Lincoln, A. I. Soto-Ortolaza, et al. 2011. "VPS35 Mutations in Parkinson Disease." *American Journal of Human Genetics* 89 (1) (July 15): 162–167. doi:10.1016/j.ajhg.2011.06.001.

Vives-Bauza, C., R. L. A. de Vries, M. Tocilescu, and S. Przedborski. 2010. "PINK1/Parkin Direct Mitochondria to Autophagy." *Autophagy* 6 (2) (February): 315–316.

Wang, C., H. S. Ko, B. Thomas, F. Tsang, K. C. M. Chew, S-P. Tay, M. W. L. Ho, et al. 2005. "Stress-Induced Alterations in Parkin Solubility Promote Parkin Aggregation and Compromise Parkin's Protective Function." *Human Molecular Genetics* 14 (24) (December 15): 3885–3897. doi:10.1093/hmg/ddi413.

Wang, F., X. Feng, J. Ma, H. Zou, and P. Chan. 2006. "A Common A340T Variant in PINK1 Gene Associated with Late-onset Parkinson's Disease in Chinese." *Neuroscience Letters* 410 (2) (December 20): 121–125. doi:10.1016/j.neulet.2006.09.080.

Wang, J. L., X. Yang, K. Xia, Z. M. Hu, L. Weng, X. Jin, H. Jiang, et al. 2010. "TGM6 Identified as a Novel Causative Gene of Spinocerebellar Ataxias Using Exome Sequencing." *Brain: a Journal of Neurology* 133 (Pt 12) (December): 3510–3518. doi:10.1093/brain/awq323.

Wang, X., T. G. Petrie, Y. Liu, J. Liu, H. Fujioka, and X. Zhu. 2012. "Parkinson's Disease-associated DJ-1 Mutations Impair Mitochondrial Dynamics and Cause Mitochondrial Dysfunction." *Journal of Neurochemistry* 121 (5) (June): 830–839. doi:10.1111/j.1471-4159.2012.07734.x.

Warnich, L., M. J. Kotze, I. M. Groenewald, J. Z. Groenewald, M. G. van Brakel, C. J. van Heerden, J. N. P. de Villiers, et al. 1996. "Identification of Three Mutations and Associated Haplotypes in the Protoporphyrinogen Oxidase Gene in South African Families with Variegate Porphyria." *Human Molecular Genetics* 5 (7) (July 1): 981–984. doi:10.1093/hmg/5.7.981.

West, A. B., P. J. Lockhart, C. O'Farell, and M. J. Farrer. 2003. "Identification of a Novel Gene Linked to Parkin via a Bi-directional Promoter." *Journal of Molecular Biology* 326 (1) (February 7): 11–19. doi:10.1016/S0022-2836(02)01376-1.

West, A B., D. J. Moore, S. Biskup, A. Bugayenko, W. W. Smith, C. A. Ross, V. L. Dawson, and T. M. Dawson. 2005. "Parkinson's Disease-associated Mutations in Leucine-rich Repeat Kinase 2 Augment Kinase Activity." *Proceedings of the National Academy of Sciences of the United States of America* 102 (46) (November 15): 16842–16847. doi:10.1073/pnas.0507360102.

Wider, C., and Z. K. Wszolek. 2007. "Clinical Genetics of Parkinson's Disease and Related Disorders." *Parkinsonism & Related Disorders* 13, Supplement 3 (0): S229–S232. doi:10.1016/S1353-8020(08)70007-5.

Wilson, M. A., C. V. St. Amour, J. L. Collins, D. Ringe, and G. A. Petsko. 2004. "The 1.8-Å Resolution Crystal Structure of YDR533Cp from Saccharomyces Cerevisiae: A Member of the DJ-1/ThiJ/PfpI Superfamily." *Proceedings of the National Academy of Sciences of the United States of America* 101 (6) (February 10): 1531–1536. doi:10.1073/pnas.0308089100.

Wirdefeldt, K. K., H-O. Adami, P. P. Cole, D. D. Trichopoulos, and J.J. Mandel. 2011. "Epidemiology and etiology of Parkinson's disease: a review of the evidence." *European journal of epidemiology* 26 Suppl 1: S1–58.

Wittwer, C.T. 2009. "High-resolution DNA Melting Analysis: Advancements and Limitations." *Human Mutation* 30 (6): 857–859. doi:10.1002/humu.20951.

Xu, K., Y-H. Xu, J-F. Chen, and M. A .Schwarzschild. 2010. "Neuroprotection by Caffeine: Time Course and Role of Its Metabolites in the MPTP Model of Parkinson's Disease." *Neuroscience* 167 (2) (May 5): 475–481. doi:10.1016/j.neuroscience.2010.02.020.

Yao, Z., and N. W. Wood. 2009. "Cell Death Pathways in Parkinson's Disease: Role of Mitochondria." *Antioxidants & Redox Signaling* 11 (9) (September): 2135–2149. doi:10.1089/ARS.2009.2624.

Ye, M. H., Chen, J.L, Zhao, G.P, Zheng, M.Q, and Wen, J. 2010. "Sensitivity and Specifcity of High-resolution Melting Analysis in Screening Unknown Snps and Genotyping a Known Mutation." *Animal Science Papers and Reports* 28 (2): 161–170.

Yu, W., Y. Sun, S. Guo, and B. Lu. 2011. "The PINK1/Parkin Pathway Regulates Mitochondrial Dynamics and Function in Mammalian Hippocampal and Dopaminergic Neurons." *Human Molecular Genetics* 20 (16) (August 15): 3227–3240. doi:10.1093/hmg/ddr235.

Zabetian, C. P., H. Morino, H. Ujike, M. Yamamoto, M. Oda, H. Maruyama, Y. Izumi, et al. 2006. "Identification and Haplotype Analysis of LRRK2 G2019S in Japanese Patients

with Parkinson Disease." *Neurology* 67 (4) (August 22): 697–699. doi:10.1212/01.wnl.0000227732.37801.d4.

Zarranz, J.J, J. Alegre, J. C. Gómez-Esteban, E. Lezcano, R. Ros, I. Ampuero, L. Vidal, et al. 2004. "The New Mutation, E46K, of Alpha-synuclein Causes Parkinson and Lewy Body Dementia." *Annals of Neurology* 55 (2) (February): 164–173. doi:10.1002/ana.10795.

Zhou, L., X. Ma, and F.Sun. 2008. "The Effects of Protein Interactions, Gene Essentiality and Regulatory Regions on Expression Variation." *BMC Systems Biology* 2 (1) (June 26): 54. doi:10.1186/1752-0509-2-54.

Zimprich, A., A. Benet-Pagès, W. Struhal, E.Graf, S. H. Eck, M. N. Offman, D. Haubenberger, et al. 2011. "A Mutation in VPS35, Encoding a Subunit of the Retromer Complex, Causes Late-Onset Parkinson Disease." *The American Journal of Human Genetics* 89 (1) (July 15): 168–175. doi:10.1016/j.ajhg.2011.06.008.

Zimprich, A., S.Biskup, P. Leitner, P. Lichtner, M.Farrer, S. Lincoln, J. Kachergus, et al. 2004. "Mutations in LRRK2 Cause Autosomal-Dominant Parkinsonism with Pleomorphic Pathology." *Neuron* 44 (4) (November 18): 601–607. doi:10.1016/j.neuron.2004.11.005.

Zohn, I. E., K. V. Anderson, and L. Niswander. 2007. "The Hectd1 Ubiquitin Ligase Is Required for Development of the Head Mesenchyme and Neural Tube Closure." *Developmental Biology* 306 (1) (June): 208–221. doi:10.1016/j.ydbio.2007.03.018.

URL Reference List

- Virtual MedStudent.com: [http://www.virtualmedstudent.com/links/neurological/parkinsons html] Date accessed: 03/03/2011
- KapaBiosystems: [http://www.kapabiosystems.com/public/pdfs/kapa-hrm-fast-pcr-kits/Introduction_to_High_Resolution_Melt_Analysis_Guide.pdf] Date accessed: 19/10/2011
- *CCDS* Report for Consensus *CDS*: [http://www.ncbi nlm nih.gov/CCDS/CcdsBrowse.cgi] Date accessed: 10/02/2012
- Wellcome Trust Sanger Institute: [http://www.sanger.ac.uk/gencode/] Date accessed: 10/02/2012
- Agilent Technologies: [http://www.genomics.agilent.com/CollectionSubpage.aspx?PageType=Product&SubPageType=ProductData&PageID=2318] Date accessed: 10/02/2012
- NCBI Reference Sequence: [http://www ncbi.nlm nih.gov/RefSeq/] Date accessed: 10/02/2012
- MicroRNA.org: [http://www.microrna.org/microrna/getDown-loads.do] Date accessed: 10/02/2012
- Illumina.com :[http://www.illumina.com/documents/products/datasheets/datasheet_truseq_exome_enrichment_kit.pdf] Date accessed: 10/02/2012
- Gencodegenes.org: [http://www.gencodegenes.org/] Date accessed: 10/02/2012
- Mirbase.org: [http://www.mirbase.org/] Date accessed: 10/02/2012

- dbSNP: [http://www.ncbi.nlm.nih.gov/projects/SNP/)] Date accessed: 10/02/2012
- Vega- Wellcome Trust Sanger Institute: [http://vega.sanger.ac.uk/index html] Date accessed: 10/02/2012
- Ensebl.org:[www.ensembl.org] Date accessed:10/02/2012
- Roche NimbleGen: [http://www nimblegen.com/products/seqcap/ez/v3/index html] Date accessed: 10/02/2012
- Variant Effect Predictor Tool: [http://www.ensembl.org/tools.htm] Date accessed: 25/03/2011
- SMART Protein Analyser: [http://smart.embl-heidelberg.de/] Date accessed: 25/03/2011
- MRC Protein Phosphorylation Unit: [http://www.ppu mrc.ac.uk/research/?pid=11&sub1=research] Date accessed: 04/04/2011
- Real Time PCR: [www.bio.davidson.edu/courses/molbio/molstudents/spring2003/pierce/realtimepcr htm] Date accessed: 06/04/2012
- QuantiTect® Multiplex PCR Handbook [www.qiagen.com/hb/quantitectmultiplexpcr] Date accessed: 18/09/2012
- KASP™ Genotyping Overview: [http://www.lgcgenomics.com/kasp-overview] Date accessed: 18/09/2012
- DNA Software: [http://dnasoftware.com/FretAssays/tabid/139/Default.aspx] Date accessed: 18/09/2012
- KBiosciences KlusterKaller Software: [http://www.kbioscience.co.uk/software/klustercaller html] Date accessed: 18/09/2012
- LGC Genomics: How does KASP™ work? [http://www.lgcgenomics.com/how-does-kasp-work] Date accessed: 20/10/2012
- GEOPD: [http://www.geopd.org] Date accessed: 20/10/2012
- World Federation of Neurology: [http://www.wfneurology.org/the-african-experience] Date accessed: 20/10/2012

Appendix 1: DNA isolation from blood using the phenol/chloroform method

Extraction of nuclei from whole blood

Blood from two 5ml EDTA tubes per patient was transferred into a 50ml Falcon tube. The tube was then filled to 20 ml with ice-cold lysis buffer and inverted gently a few times. Subsequently, the sample was incubated on ice for 5-10 min. The sample was then centrifuged at 2500-3000 rpm at room temperature in a Beckman model TJ-6 centrifuge (Scotland, UK). The supernatant was discarded and the pellet was resuspended in 20ml, ice-cold lysis buffer which was then followed by another round of incubation and centrifugation. The supernatant was discarded and the pellet resuspended in DNA extraction buffer, after which the nuclei were either immediately used for DNA extraction, or stored at -70°C until DNA was required for genetic testing.

Extraction of DNA from nuclei

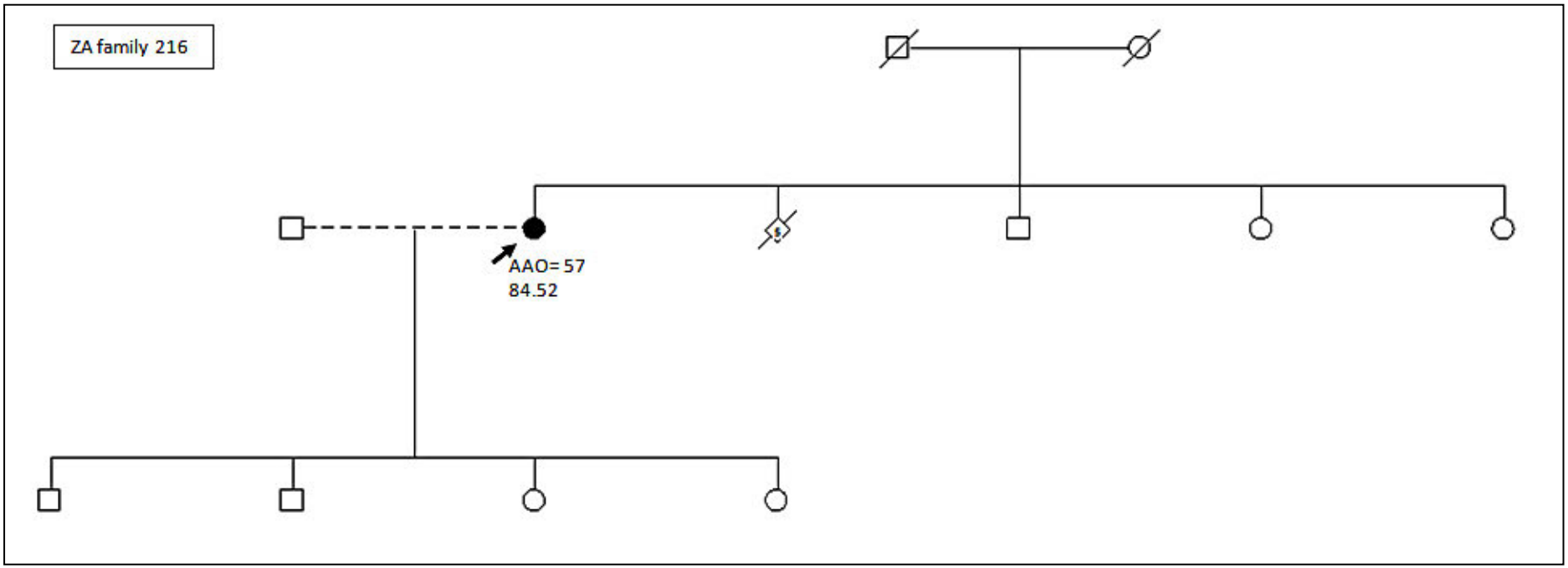
A total volume of 100µl of proteinase K (10µg/ml) was added to newly prepared or defrosted nuclei and the mixture was incubated overnight at 37°C. After this step, 2ml distilled water, 500µl 3M sodium-acetate and 25µl phenol/chloroform were added to the sample. The tubes were subsequently inverted and mixed gently for 10 min on a Voss rotator (Voss of Maldon, England) at 4°C. The mixture was then transferred to a glass Corex tube so that the aqueous phase could be clearly distinguished from the organic phase, followed by centrifugation in a Sorvall RC-5B refrigerated super-speed centrifuge (rotor SS 34, Dupont Instruments) at 8000 rpm at 4°C for 10 min.

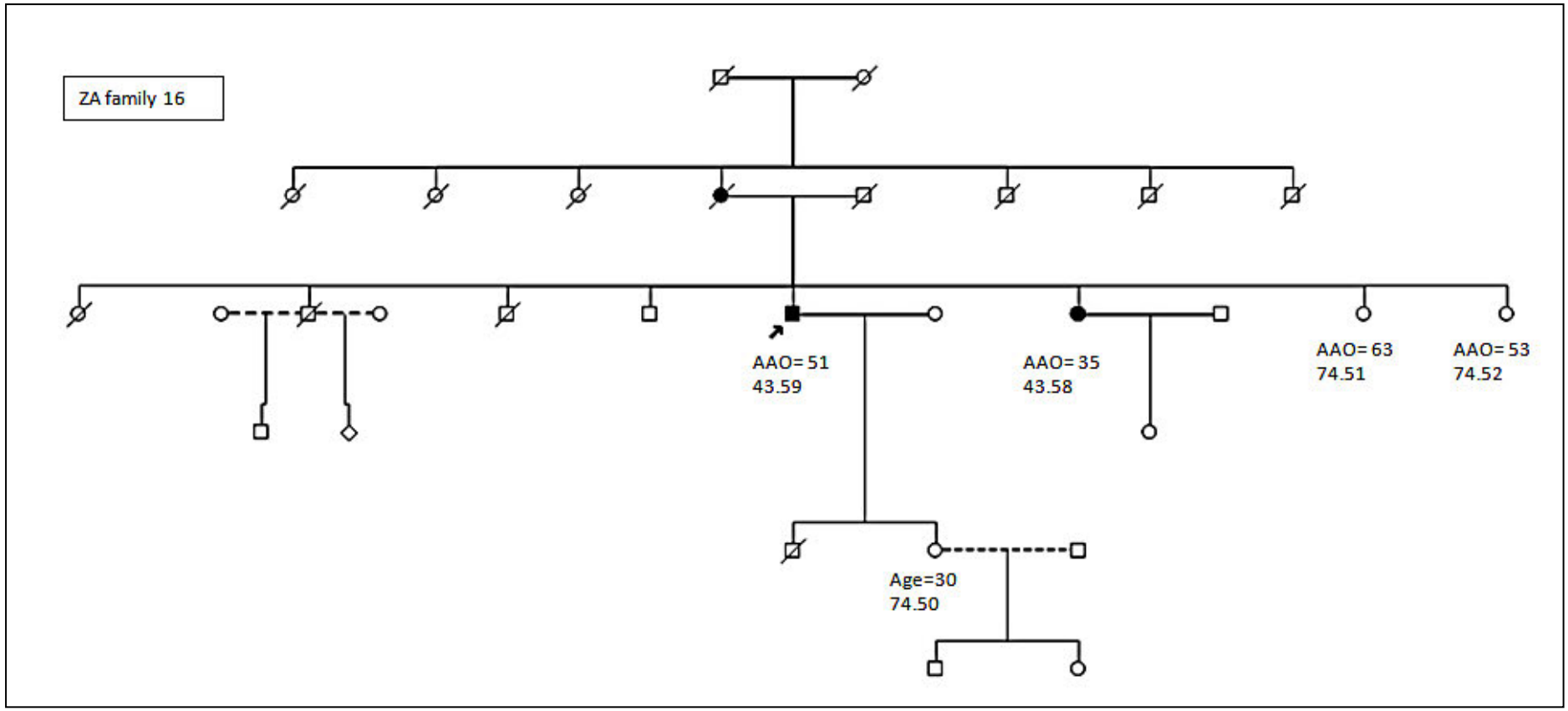
The upper aqueous phase contained the DNA and was transferred to a clean Corex tube using a sterile plastic Pasteur pipette, while taking care not to disturb the interface or the organic phase. Approximately 25ml chloroform/octanol was added to the aqueous phase after which the tube was closed with a polypropylene stopper and gently inverted for 10 min. This mixture was centrifuged at 4°C, followed by the removal of the upper aqueous phase as described earlier. The DNA was then ethanol precipitated by adding two volumes of ice-cold 96% ethanol and inverting gently until DNA strands appeared as a white precipitate.

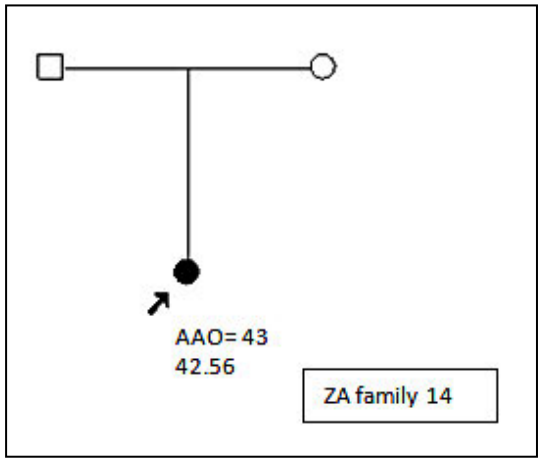
The DNA strands were removed using a yellow-tipped Gilson pipette and placed in a clean, 1.5ml Eppendorf microfuge tube. One millilitre 70% ethanol was then added to the DNA and the mixture centrifuged in a Beckman microfuge for 3 min at 13000 rpm. The ethanol was carefully decanted and the 70% ethanol wash step was repeated one more time in order to remove any excess salts. After careful removal of most of the ethanol, the DNA pellet was air-dried for 30-60 min at room temperature by inverting the Eppendorf microfuge tube on Carlton paper. Two hundred microlitres Tris-EDTA buffer was added and the DNA was resuspended, initially by stationary incubation at 37°C overnight and subsequently by gentle mixing in a Voss rotator at 4°C for a further 3 days. This was followed by stationary incubation at 4°C until the DNA had been fully resuspended.

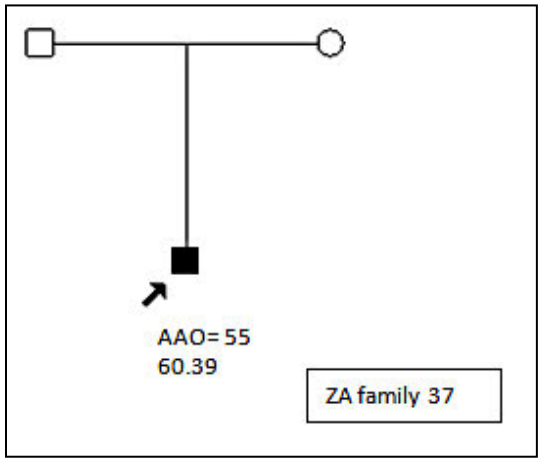
After 1-2 weeks, when the DNA had completely resuspended in the buffer, the optical density (OD) of the DNA was determined in a Milton Roy series 120i spectrophotometer (USA) at 260nm (OD260). The DNA concentration, in μg/μl, was determined by diluting 10μl of DNA in 500μl of TE and multiplying the measured OD260 by a factor of 2.5, while the purity of the DNA was monitored by the OD260//OD280 ratio, which should be approximately 1.8 for pure DNA.

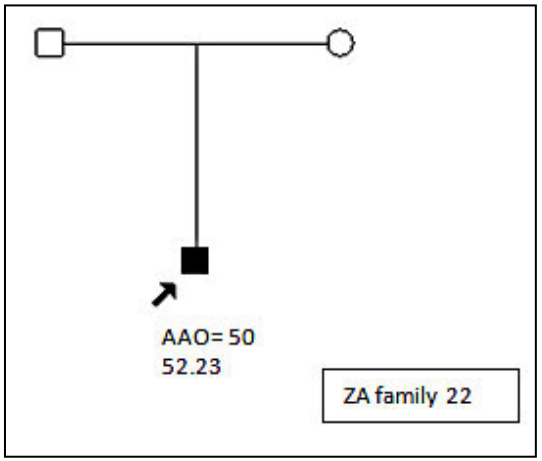
Appendix 11: Pedigrees of the Black African Patients

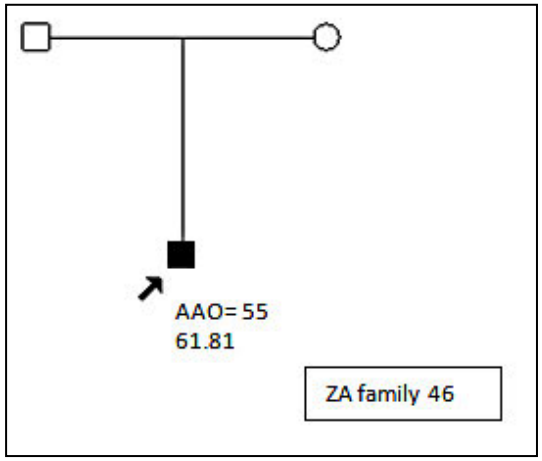


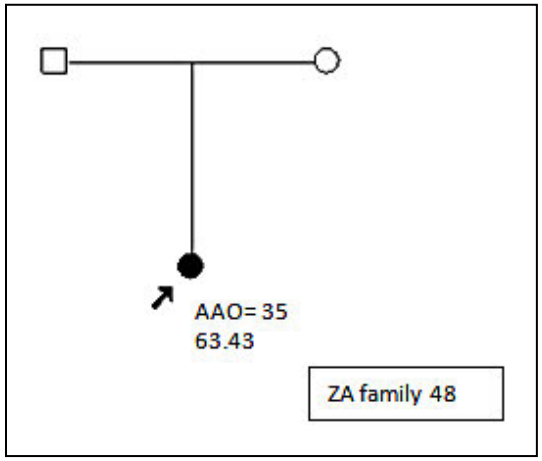

ZA family 16
AAO= 51
43.59
AAO= 35
43.58
AAO= 63
74.51
AAO= 53
74.52
Age=30
74.50

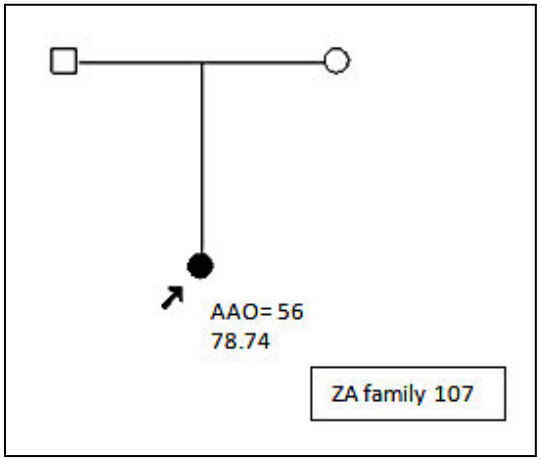

AAO= 43
42.56
ZA family 14

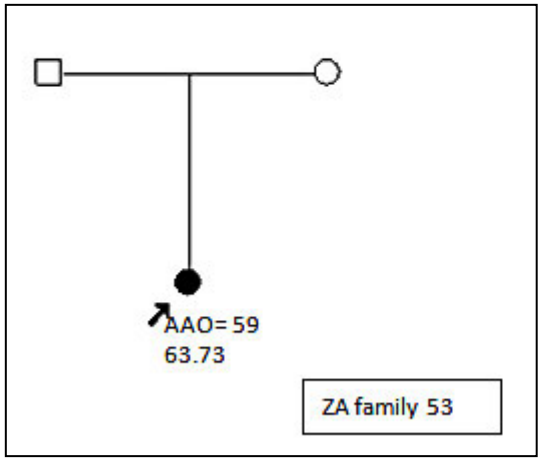

AAO= 55
60.39
ZA family 37

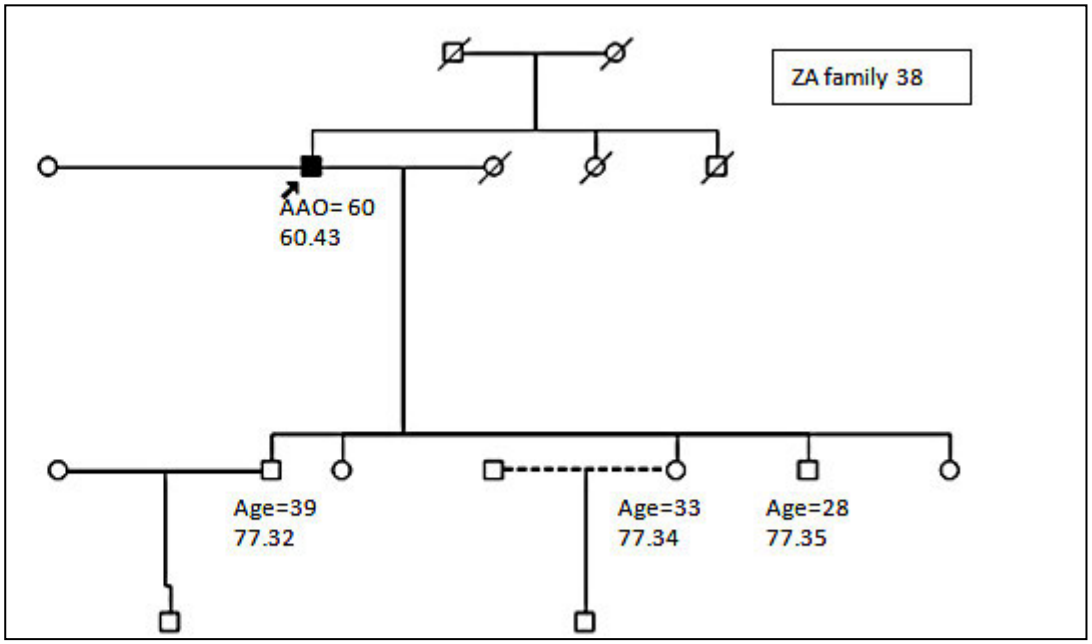

AAO= 50
52.23
ZA family 22

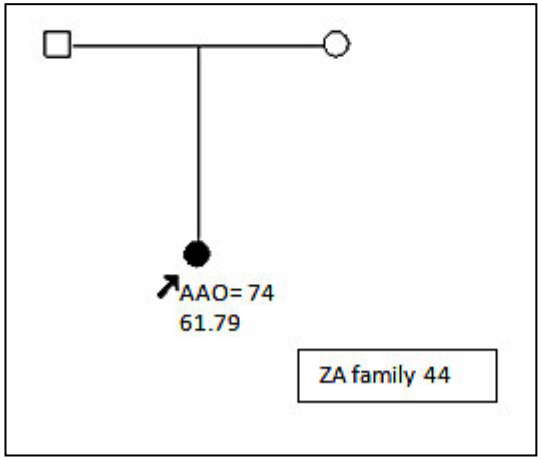

AAO= 55
61.81
ZA family 46

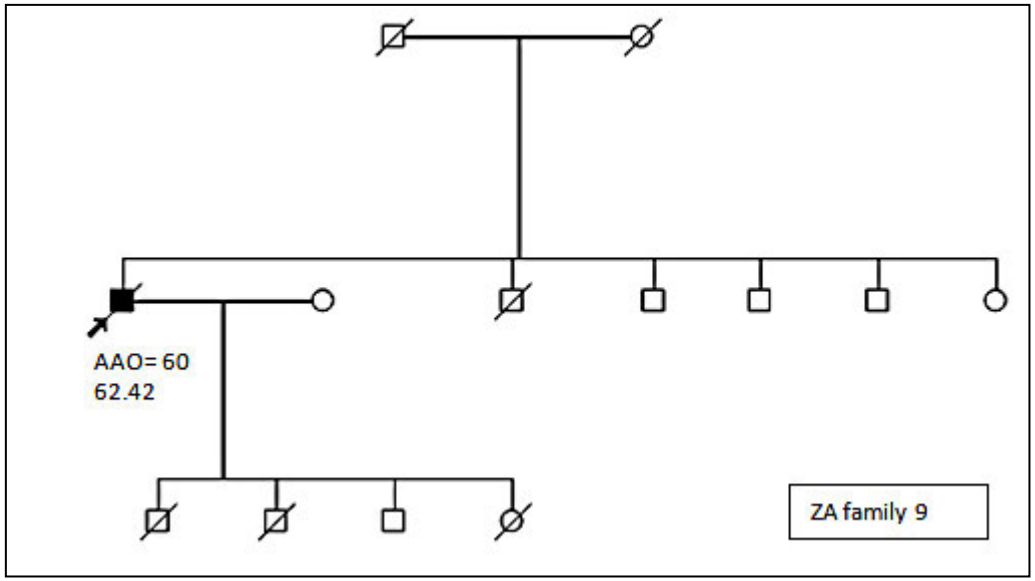

AAO= 35
63.43
ZA family 48

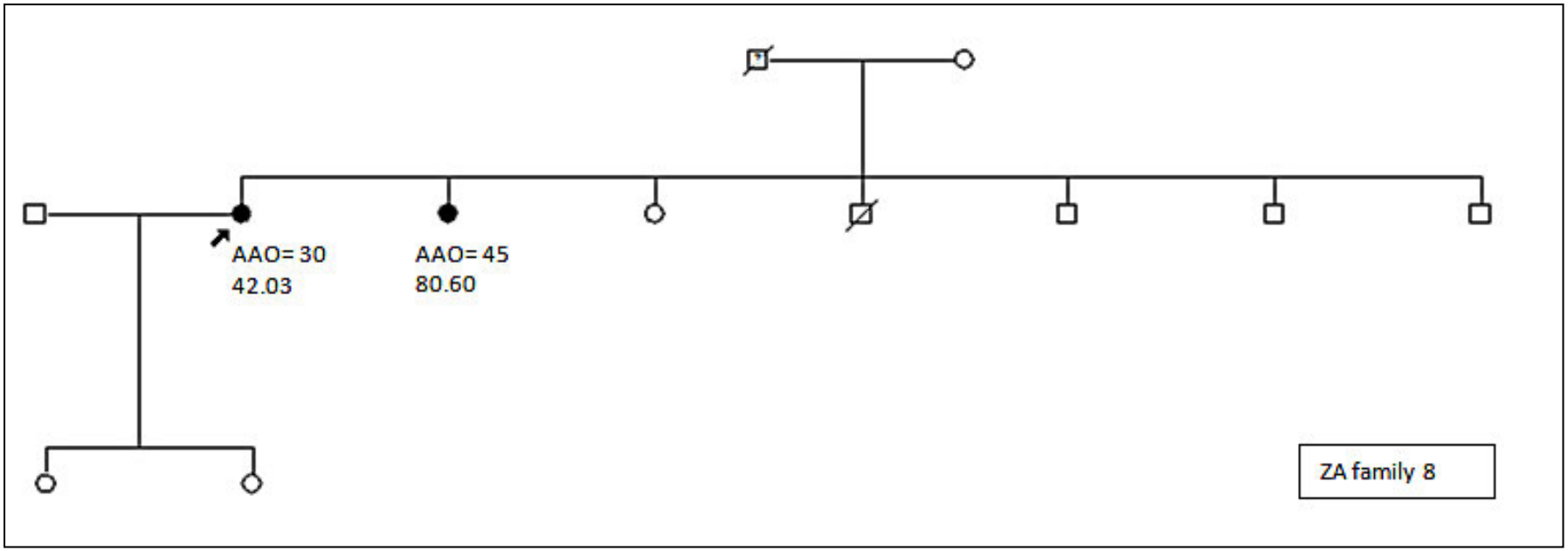

AAO= 56
78.74
ZA family 107

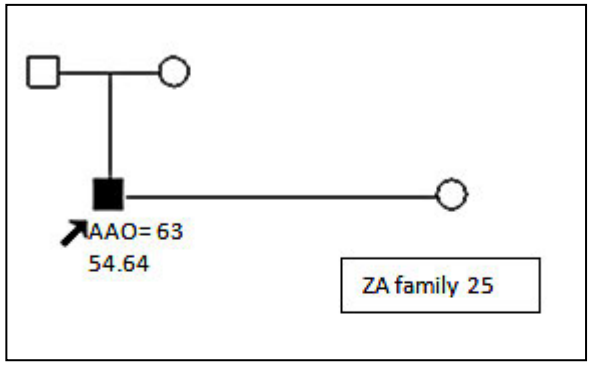

AAO= 59
63.73
ZA family 53

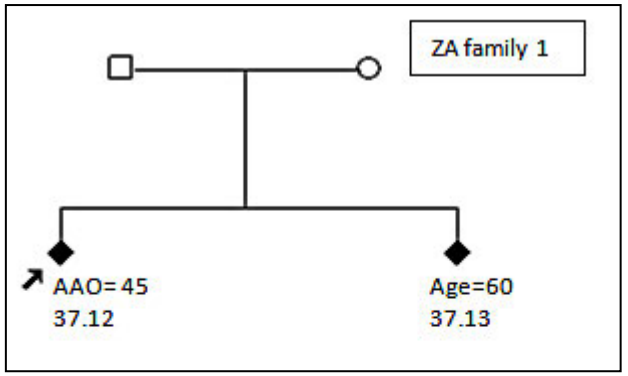

ZA family 38
AAO= 60
60.43
Age=39
77.32
Age=33
77.34
Age=28
77.35

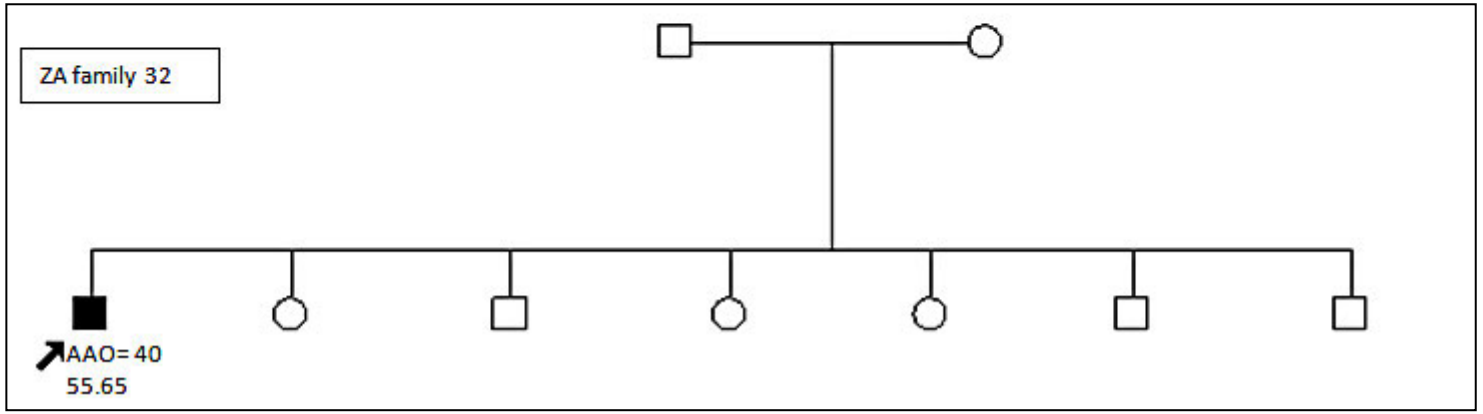

AAO= 74
61.79
ZA family 44

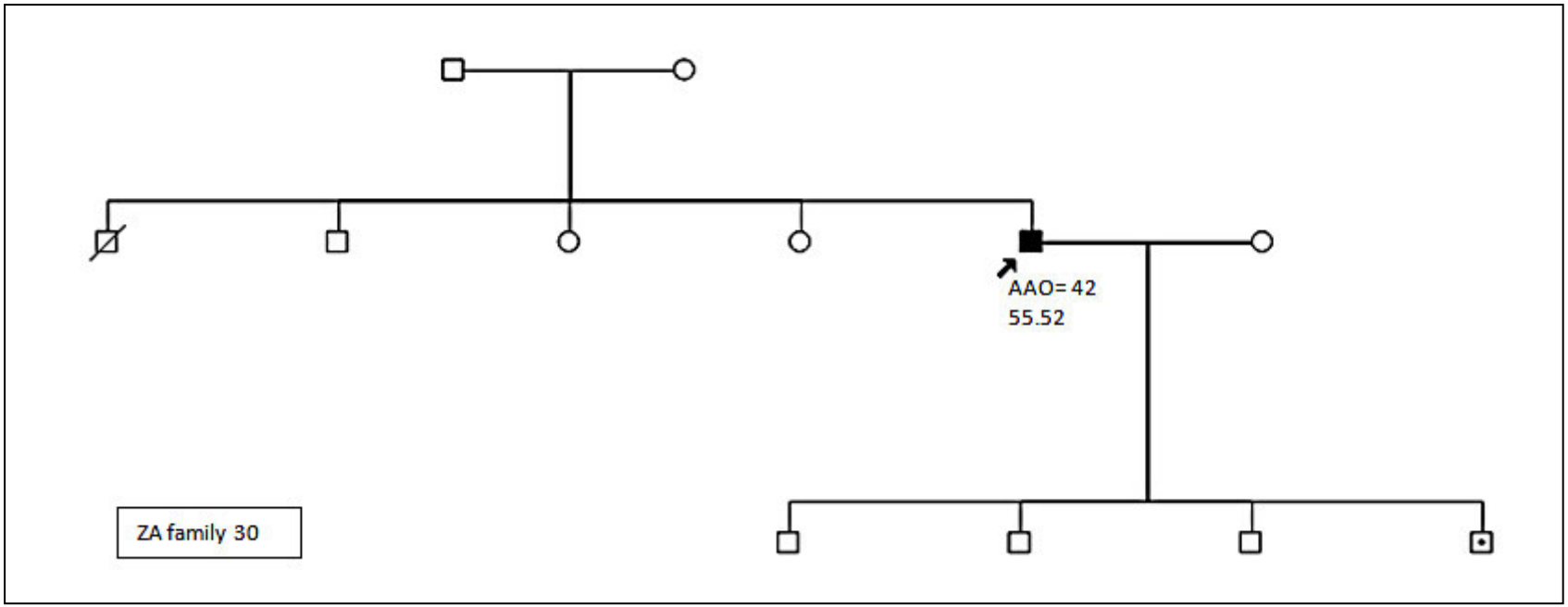

AAO= 60
62.42
ZA family 9


AAO= 30
42.03
AAO= 45
80.60
ZA family 8


AAO= 63
54.64
ZA family 25


ZA family 1
AAO= 45
37.12
Age=60
37.13


ZA family 32
AAO= 40
55.65


AAO= 42
55.52
ZA family 30

Appendix III: List of PD terms used for BORG Analysis

GO:0000151	ubiquitin ligase complex
GO:0000152	nuclear ubiquitin ligase complex
GO:0000153	cytoplasmic ubiquitin ligase complex
GO:0000209	protein polyubiquitination
GO:0000422	mitochondrion degradation
GO:0000502	proteasome complex
GO:0000835	ER ubiquitin ligase complex
GO:0000836	Hrd1p ubiquitin ligase complex
GO:0000837	Doa10p ubiquitin ligase complex
GO:0000838	Hrd1p ubiquitin ligase ERAD-M complex
GO:0000839	Hrd1p ubiquitin ligase ERAD-L complex
GO:0001588	dopamine receptor activity, coupled via Gs
GO:0001591	dopamine receptor activity, coupled via Gi/Go
GO:0001963	synaptic transmission, dopaminergic
GO:0001964	startle response
GO:0003832	beta-alanyl-dopamine hydrolase activity
GO:0003833	beta-alanyl-dopamine synthase activity
GO:0004221	ubiquitin thiolesterase activity
GO:0004500	dopamine beta-monooxygenase activity
GO:0004839	ubiquitin activating enzyme activity
GO:0004842	ubiquitin-protein ligase activity
GO:0004843	ubiquitin-specific protease activity
GO:0004952	dopamine receptor activity
GO:0005329	dopamine transmembrane transporter activity
GO:0005330	dopamine:sodium symporter activity
GO:0005838	proteasome regulatory particle
GO:0005839	proteasome core complex
GO:0006464	protein modification process

GO:0006511	ubiquitin-dependent protein catabolic process
GO:0006513	protein monoubiquitination
GO:0006585	dopamine biosynthetic process from tyrosine
GO:0006986	response to unfolded protein
GO:0007014	actin ubiquitination
GO:0007191	activation of adenylate cyclase activity by dopamine receptor signaling pathway
GO:0007195	inhibition of adenylate cyclase activity by dopamine receptor signaling pathway
GO:0007212	dopamine receptor signaling pathway
GO:0008537	proteasome activator complex
GO:0008540	proteasome regulatory particle, base subcomplex
GO:0008541	proteasome regulatory particle, lid subcomplex
GO:0010390	histone monoubiquitination
GO:0010499	proteasomal ubiquitin-independent protein catabolic process
GO:0010798	regulation of multivesicular body size involved in ubiquitin-dependent protein catabolism
GO:0010992	ubiquitin homeostasis
GO:0010993	regulation of ubiquitin homeostasis
GO:0010994	free ubiquitin chain polymerization
GO:0010995	free ubiquitin chain depolymerization
GO:0014046	dopamine secretion
GO:0014059	regulation of dopamine secretion
GO:0014069	postsynaptic density
GO:0015872	dopamine transport
GO:0016567	protein ubiquitination
GO:0016574	histone ubiquitination
GO:0016578	histone deubiquitination
GO:0016579	protein deubiquitination
GO:0019005	SCF ubiquitin ligase complex
GO:0019773	proteasome core complex, alpha-subunit complex
GO:0019774	proteasome core complex, beta-subunit complex
GO:0022623	proteasome-activating nucleotidase complex

GO:0022624	proteasome accessory complex
GO:0030579	ubiquitin-dependent SMAD protein catabolic process
GO:0031144	proteasome localization
GO:0031145	anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process
GO:0031146	SCF-dependent proteasomal ubiquitin-dependent protein catabolic process
GO:0031371	ubiquitin conjugating enzyme complex
GO:0031396	regulation of protein ubiquitination
GO:0031397	negative regulation of protein ubiquitination
GO:0031398	positive regulation of protein ubiquitination
GO:0031461	cullin-RING ubiquitin ligase complex
GO:0031462	Cul2-RING ubiquitin ligase complex
GO:0031463	Cul3-RING ubiquitin ligase complex
GO:0031464	Cul4A-RING ubiquitin ligase complex
GO:0031465	Cul4B-RING ubiquitin ligase complex
GO:0031466	Cul5-RING ubiquitin ligase complex
GO:0031467	Cul7-RING ubiquitin ligase complex
GO:0031593	polyubiquitin binding
GO:0031595	nuclear proteasome complex
GO:0031596	ER proteasome complex
GO:0031597	cytosolic proteasome complex
GO:0031598	nuclear proteasome regulatory particle
GO:0031599	ER proteasome regulatory particle
GO:0031600	cytosolic proteasome regulatory particle
GO:0031601	nuclear proteasome core complex
GO:0031602	ER proteasome core complex
GO:0031603	cytosolic proteasome core complex
GO:0031606	cytosolic proteasome core complex, alpha-subunit complex
GO:0031607	nuclear proteasome core complex, beta-subunit complex
GO:0031608	ER proteasome core complex, beta-subunit complex
GO:0031609	cytosolic proteasome core complex, beta-subunit complex

GO:0031610 nuclear proteasome regulatory particle, base subcomplex

GO:0031611 ER proteasome regulatory particle, base subcomplex

GO:0031612 cytosolic proteasome regulatory particle, base subcomplex

GO:0031613 nuclear proteasome regulatory particle, lid subcomplex

GO:0031614 ER proteasome regulatory particle, lid subcomplex

GO:0031615 cytosolic proteasome regulatory particle, lid subcomplex

GO:0031624 ubiquitin conjugating enzyme binding

GO:0031625 ubiquitin protein ligase binding

GO:0031748 D1 dopamine receptor binding

GO:0031749 D2 dopamine receptor binding

GO:0031750 D3 dopamine receptor binding

GO:0031751 D4 dopamine receptor binding

GO:0031752 D5 dopamine receptor binding

GO:0032225 regulation of synaptic transmission, dopaminergic

GO:0032226 positive regulation of synaptic transmission, dopaminergic

GO:0032227 negative regulation of synaptic transmission, dopaminergic

GO:0032434 regulation of proteasomal ubiquitin-dependent protein catabolic process

GO:0032435 negative regulation of proteasomal ubiquitin-dependent protein catabolic process

GO:0032436 positive regulation of proteasomal ubiquitin-dependent protein catabolic process

GO:0033134 ubiquitin activating enzyme binding

GO:0033182 regulation of histone ubiquitination

GO:0033183 negative regulation of histone ubiquitination

GO:0033184 positive regulation of histone ubiquitination

GO:0033522 histone H2A ubiquitination

GO:0033523 histone H2B ubiquitination

GO:0033602 negative regulation of dopamine secretion

GO:0033603 positive regulation of dopamine secretion

GO:0033768 SUMO-targeted ubiquitin ligase complex

GO:0034450 ubiquitin-ubiquitin ligase activity

GO:0034515 proteasome storage granule

GO:0035240 dopamine binding

GO:0035361	Cul8-RING ubiquitin ligase complex
GO:0035518	histone H2A monoubiquitination
GO:0035519	protein K29-linked ubiquitination
GO:0035520	monoubiquitinated protein deubiquitination
GO:0035521	monoubiquitinated histone deubiquitination
GO:0035522	monoubiquitinated histone H2A deubiquitination
GO:0035523	protein K29-linked deubiquitination
GO:0035616	histone H2B conserved C-terminal lysine deubiquitination
GO:0035800	ubiquitin-specific protease activator activity
GO:0035871	protein K11-linked deubiquitination
GO:0042053	regulation of dopamine metabolic process
GO:0042416	dopamine biosynthetic process
GO:0042417	dopamine metabolic process
GO:0042420	dopamine catabolic process
GO:0042551	neuron maturation
GO:0042787	protein ubiquitination involved in ubiquitin-dependent protein catabolic process
GO:0043130	ubiquitin binding
GO:0043161	proteasomal ubiquitin-dependent protein catabolic process
GO:0043162	ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway
GO:0043223	cytoplasmic SCF ubiquitin ligase complex
GO:0043224	nuclear SCF ubiquitin ligase complex
GO:0043248	proteasome assembly
GO:0043328	protein targeting to vacuole involved in ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway
GO:0043329	protein targeting to membrane involved in ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway
GO:0043494	CLRC ubiquitin ligase complex
GO:0044313	protein K6-linked deubiquitination
GO:0044314	protein K27-linked ubiquitination
GO:0044382	CLRC ubiquitin ligase complex localization to heterochromatin
GO:0045963	negative regulation of dopamine metabolic process

GO:0045964	positive regulation of dopamine metabolic process
GO:0046928	regulation of neurotransmitter secretion
GO:0050372	ubiquitin-calmodulin ligase activity
GO:0050780	dopamine receptor binding
GO:0051087	chaperone binding
GO:0051436	negative regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle
GO:0051437	positive regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle
GO:0051438	regulation of ubiquitin-protein ligase activity
GO:0051439	regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle
GO:0051440	regulation of ubiquitin-protein ligase activity involved in meiotic cell cycle
GO:0051441	positive regulation of ubiquitin-protein ligase activity involved in meiotic cell cycle
GO:0051442	negative regulation of ubiquitin-protein ligase activity involved in meiotic cell cycle
GO:0051443	positive regulation of ubiquitin-protein ligase activity
GO:0051444	negative regulation of ubiquitin-protein ligase activity
GO:0051583	dopamine uptake
GO:0051584	regulation of dopamine uptake
GO:0051585	negative regulation of dopamine uptake
GO:0051586	positive regulation of dopamine uptake
GO:0051587	inhibition of dopamine uptake
GO:0051865	protein autoubiquitination
GO:0055105	ubiquitin-protein ligase inhibitor activity
GO:0055106	ubiquitin-protein ligase regulator activity
GO:0060158	activation of phospholipase C activity by dopamine receptor signaling pathway
GO:0060159	regulation of dopamine receptor signaling pathway
GO:0060160	negative regulation of dopamine receptor signaling pathway
GO:0060161	positive regulation of dopamine receptor signaling pathway
GO:0060162	negative regulation of phospholipase C-activating dopamine receptor signaling pathway
GO:0070086	ubiquitin-dependent endocytosis
GO:0070530	K63-linked polyubiquitin binding
GO:0070534	protein K63-linked ubiquitination

GO:0070535 histone H2A K63-linked ubiquitination

GO:0070536 protein K63-linked deubiquitination

GO:0070537 histone H2A K63-linked deubiquitination

GO:0070628 proteasome binding

GO:0070682 proteasome regulatory particle assembly

GO:0070842 aggresome assembly

GO:0070844 polyubiquitinated protein transport

GO:0070845 polyubiquitinated misfolded protein transport

GO:0070936 protein K48-linked ubiquitination

GO:0070979 protein K11-linked ubiquitination

GO:0071108 protein K48-linked deubiquitination

GO:0071542 dopaminergic neuron differentiation

GO:0071596 ubiquitin-dependent protein catabolic process via the N-end rule pathway

GO:0071629 cytoplasm-associated proteasomal ubiquitin-dependent protein catabolic process

GO:0071630 nucleus-associated proteasomal ubiquitin-dependent protein catabolic process

GO:0071795 K11-linked polyubiquitin binding

GO:0071796 K6-linked polyubiquitin binding

GO:0071894 histone H2B conserved C-terminal lysine ubiquitination

GO:0071947 protein deubiquitination involved in ubiquitin-dependent protein catabolic process

GO:0075346 modification by symbiont of host protein by ubiquitination

GO:0080008 CUL4 RING ubiquitin ligase complex

GO:0080129 proteasome core complex assembly

GO:0085020 protein K6-linked ubiquitination

GO:0090085 regulation of protein deubiquitination

GO:0090086 negative regulation of protein deubiquitination

GO:0090363 regulation of proteasome core complex assembly

GO:0090364 regulation of proteasome assembly

GO:0097027 ubiquitin-protein ligase activator activity

GO:0097039 protein linear polyubiquitination

GO:1900044 regulation of protein K63-linked ubiquitination

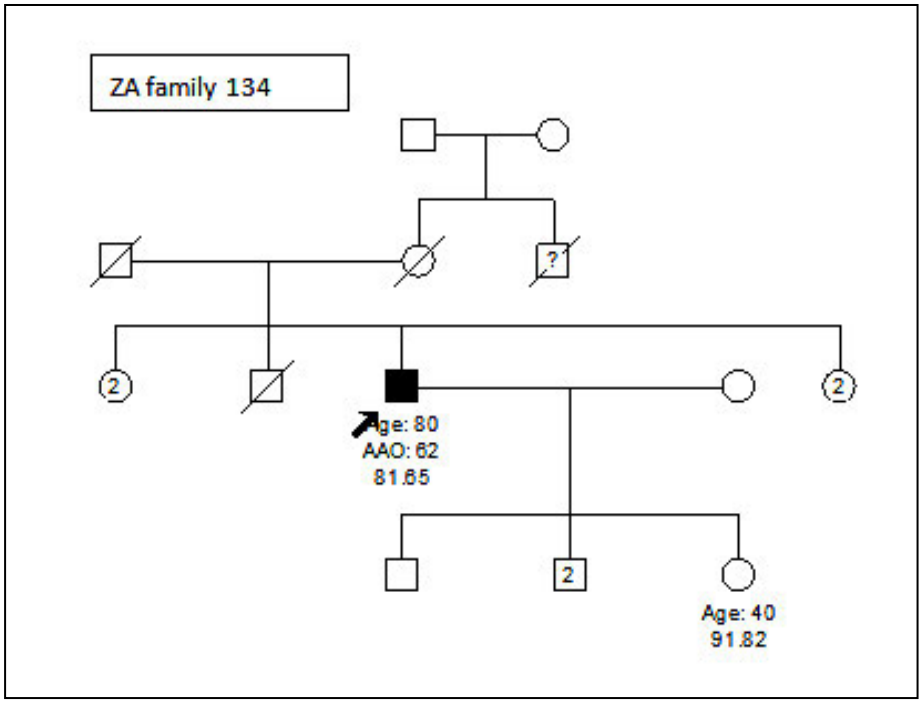
GO:1900045 negative regulation of protein K63-linked ubiquitination

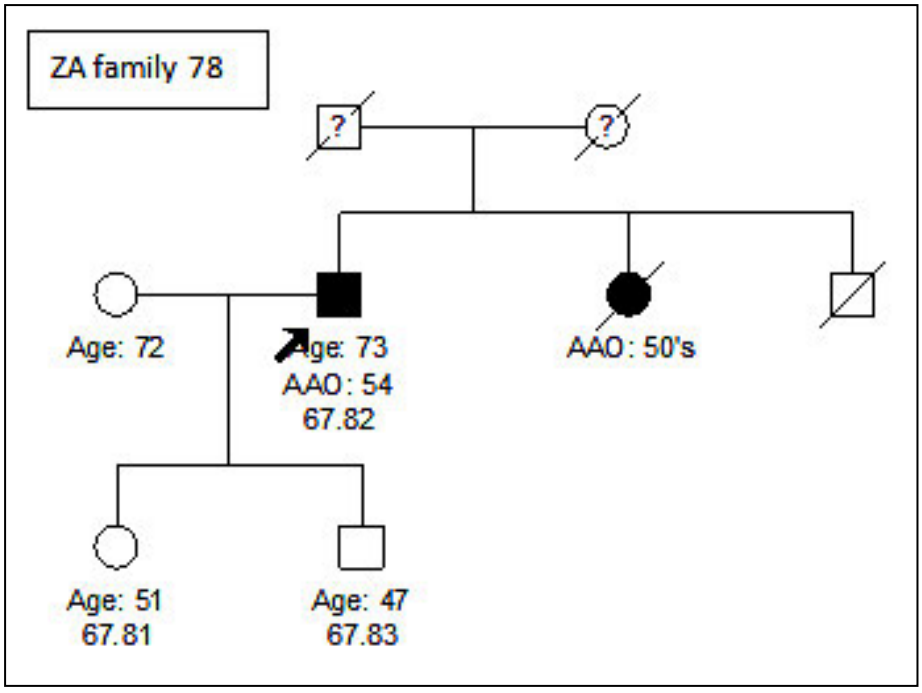
GO:2000058	regulation of protein ubiquitination involved in ubiquitin-dependent protein catabolic process
GO:2000059	negative regulation of protein ubiquitination involved in ubiquitin-dependent protein catabolic process
GO:2000060	positive regulation of protein ubiquitination involved in ubiquitin-dependent protein catabolic process
GO:2000152	regulation of ubiquitin-specific protease activity
GO:2000157	negative regulation of ubiquitin-specific protease activity
GO:2000158	positive regulation of ubiquitin-specific protease activity
GO:2000395	regulation of ubiquitin-dependent endocytosis
GO:2000396	negative regulation of ubiquitin-dependent endocytosis
GO:2000397	positive regulation of ubiquitin-dependent endocytosis
GO:2000777	positive regulation of proteasomal ubiquitin-dependent protein catabolic process involved in cellular response to hypoxia
GO:2001166	regulation of histone H2B ubiquitination
GO:2001167	negative regulation of histone H2B ubiquitination
GO:2001168	positive regulation of histone H2B ubiquitination
GO:2001173	regulation of histone H2B conserved C-terminal lysine ubiquitination
GO:2001174	negative regulation of histone H2B conserved C-terminal lysine ubiquitination
GO:2001175	positive regulation of histone H2B conserved C-terminal lysine ubiquitination
HP:0001276	Hypertonia
HP:0001288	Gait disturbance
HP:0001295	Involuntary rhythmic myoclonic movements ('tremor') of the distal extremities, usually fingers
HP:0001300	Parkinsonism
HP:0001309	Movements ('tremors') characterized by 8 to 10-Hz discharges
HP:0001332	Dystonia
HP:0001337	Tremor
HP:0001347	Hyperreflexia
HP:0002063	Rigidity
HP:0002067	Bradykinesia
HP:0002080	Intention tremor
HP:0002172	Postural instability

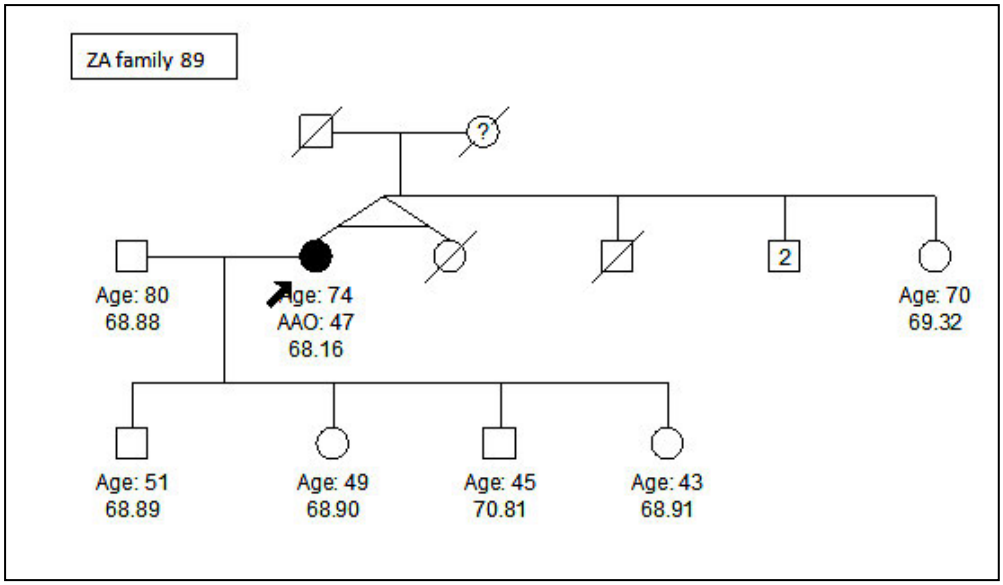
HP:0002174	Postural tremor
HP:0002275	Poor motor coordination
HP:0002322	Resting tremor
HP:0002333	Motor deterioration
HP:0002345	Action tremor
HP:0002346	Head tremor
HP:0002361	Psychomotor degeneration
HP:0002378	Hand tremor
HP:0002489	Psychomotor regression
HP:0006925	Postural tremor, slow, irregular
HP:0007197	Action and postural tremor
HP:0007297	Postural tremor of arms
HP:0007351	Upper limb postural tremor
HP:0100022	Abnormality of movement
MP:0000745	tremors
MP:0000811	hippocampal neuron degeneration
MP:0000836	abnormal substantia nigra morphology
MP:0000952	abnormal CNS glial cell morphology
MP:0001363	increased anxiety-related response
MP:0001388	abnormal stationary movement
MP:0001405	impaired coordination
MP:0001905	abnormal dopamine level
MP:0001906	increased dopamine level
MP:0002066	abnormal motor capabilities/coordination/movement
MP:0002272	abnormal nervous system electrophysiology
MP:0002882	abnormal neuron morphology
MP:0003203	increased neuron apoptosis
MP:0003224	neuron degeneration
MP:0003243	abnormal dopaminergic neuron morphology
MP:0003244	loss of dopaminergic neurons
MP:0003313	abnormal locomotor activation

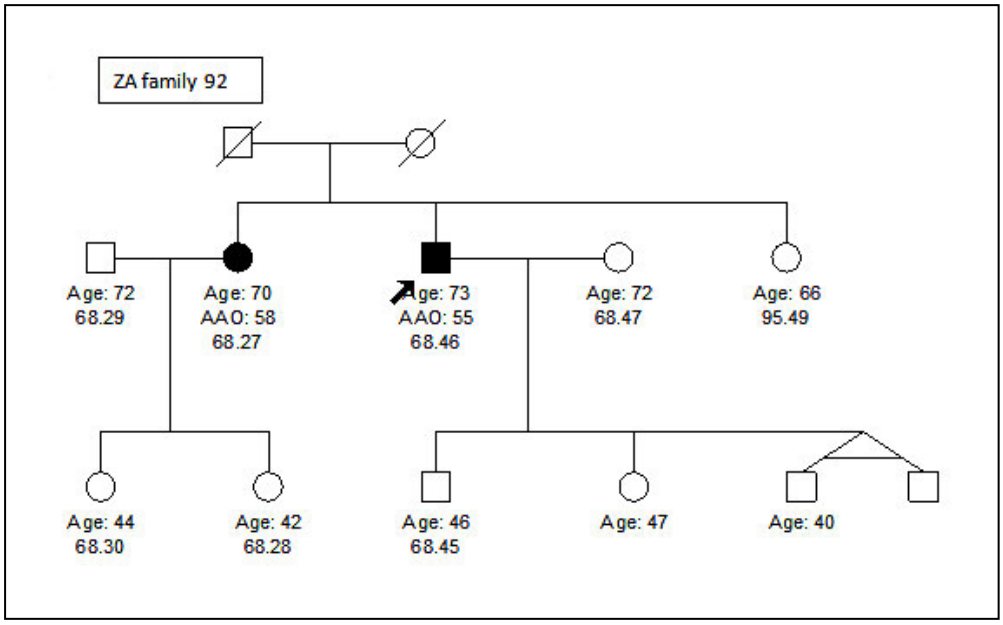
MP:0003491	abnormal voluntary movement
MP:0003492	abnormal involuntary movement
MP:0003635	abnormal synaptic transmission
MP:0003964	abnormal noradrenaline level
MP:0004250	tau protein deposits
MP:0005424	jerky movement
MP:0005643	decreased dopamine level
MP:0006035	abnormal mitochondrial morphology
MP:0010149	abnormal synaptic dopamine release
MP:0011448	decreased dopaminergic neuron number
MP:0011449	increased dopaminergic neuron number
MP:0011450	ectopic dopaminergic neuron
MP:0011451	increased susceptibility to dopaminergic neuron neurotoxicity
MP:0011452	decreased susceptibility to dopaminergic neuron neurotoxicity
PW:0000144	ubiquitin/proteasome degradation pathway
PW:0000182	lysosomes based pathway of protein degradation
PW:0000294	altered ubiquitin/proteasome degradation pathway
PW:0000325	protein degradation pathway
PW:0000326	altered protein degradation pathway
PW:0000394	dopamine signaling pathway
PW:0000395	altered dopamine signaling pathway
PW:0000409	dopamine metabolic pathway
PW:0000414	protein degradation pathway via the 'core' 20S proteasome pathway
PW:0000415	proteasome degradation pathway involving cullin-dependent ubiquitin ligases
PW:0000417	ubiquitin, ubiquitin-like/proteasome degradation pathway
PW:0000418	altered ubiquitin, ubiquitin-like degradation pathway
PW:0000433	protein modification pathway via conjugation with ubiquitin and ubiqutin-like molecules
PW:0000434	ubiquitination pathway
PW:0000802	dopamine biosynthetic pathway
PW:0000851	dopamine signaling pathway via D1 family of receptors

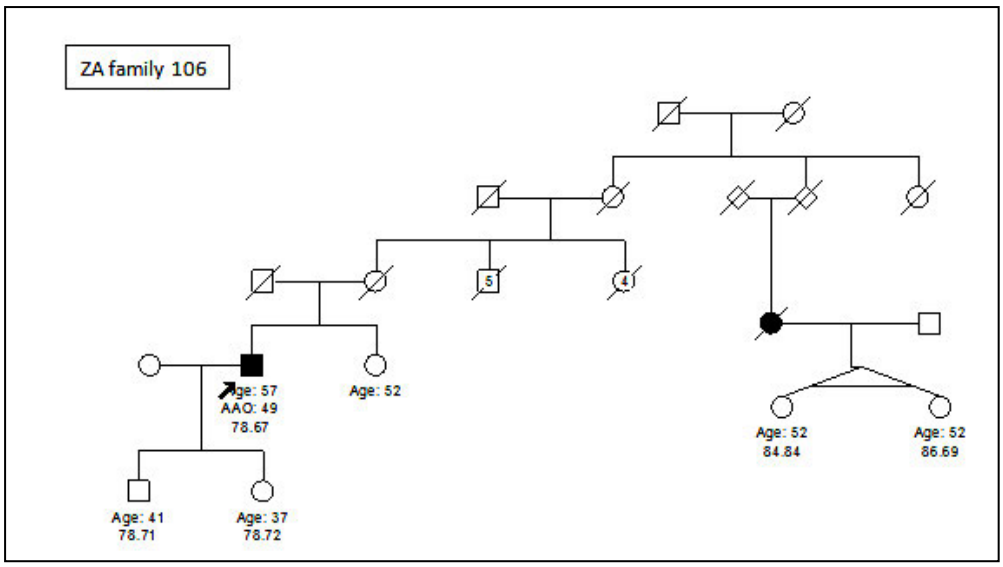
Appendix IV: The individual pedigrees of each of the Afrikaner individuals included in Chapter 8

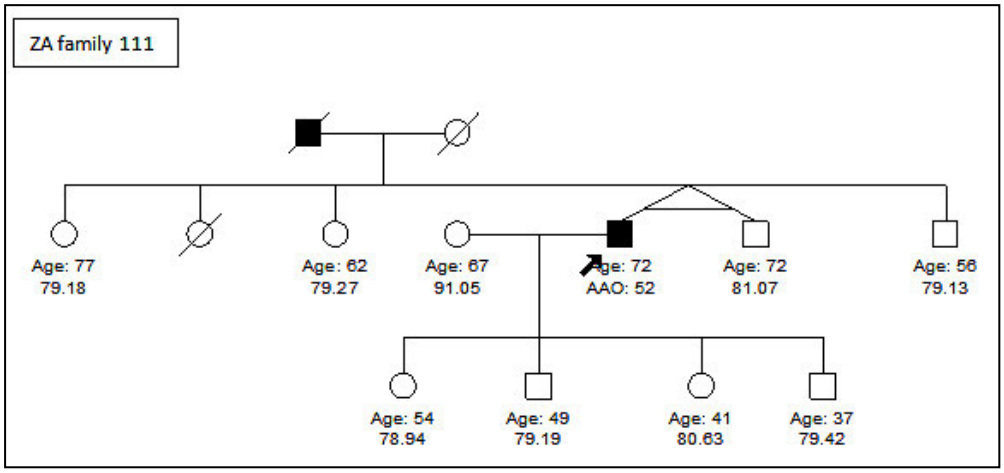












Appendix V: Reagents and Solutions

Cresol Loading Dye

2% (v/v) 10mg/ml cresol stock solution

0.9933M sucrose

10x TBE Electrophoresis Buffer (pH 8.3)

0.0890M Trizma Base

0.0890M Boric Acid

0.0020M EDTA

12% Polyacrylamide gel electrophoresis (PAGE) gels

H_2O	10.2 ml
1.5 M Tris-HCl (pH 8.8)	7.5 ml
20% (w/v) SDS	0.15 ml
Acrylamide/Bis-acrylamide (30%/0.8% w/v)	12.0 ml
10% (w/v) ammonium persulfate (APS)	0.15 ml
TEMED	0.02 ml

Stacking Gel Solution (4% Acrylamide):

H_2O	3.075 ml
0.5 M Tris-HCl, pH 6.8	1.25 ml
20% (w/v) SDS	0.025 ml
Acrylamide/Bis-acrylamide (30%/0.8% w/v)	0.67 ml
10% (w/v) ammonium persulfate (APS)	0.025 ml
TEMED	0.005 ml

Appendix VI: Primers Designed for Validation of WES Results

Gene	Primer Sequence (5´-3´)	%GC	Tm (^{0}C)	PCR conditions (Ta in ^{0}C)	Size of PCR fragment (bp)
CDC27	For: GAG TAT CTA CTG AAG CTC CTC CAT	45.8	58.9	55	384
	Rev: CCA GCA CCA TCA ATA CGA CTT TGT C	48.0	63.4		
NEDD4	For: TTC CTA GGA GTC CTT CAA CTA CCC GA	50.0	64.2	58	477
	Rev: TTT GGG AGG ATG GCA AAC ATG CAG	50.0	64.2		
HECTD1	For: ATC CCA GGA AGg tct gta aga agc	50.0	61.9	55	290
	Rev: TCA GAG GGA AGG GAA AGA TGG TGA	50.0	62.1		
RNF40	For: ACC ATC TGA CTT CAT CCC TCT TCC	50.0	67.7	58	491
	Rev: AAG GAG ACA GGA ACA GAG CCT CA	52.3	62.9		
TBCC	For: GGG CTG CAG AAA CTA ATC AAC GAC	50.0	62.4	55	478
	Rev: AGT CAC TGC AGT CCT CCA GGA AA	53.4	63.4		